A Session-based Approach to Autonomous Database Tuning

Krisztián Mózsi, Attila Kiss

Eötvös Loránd University, Faculty of Informatics
Pázmány Péter sétány 1/C, 1117 Budapest, Hungary
mozsik@inf.elte.hu, kiss@inf.elte.hu

Abstract: By using autonomous tuning tools to optimize database systems, a lot of time-consuming, manual work can be automated. However, self-tuning database systems are trying to optimize global metrics of efficiency, they may set back rare, but critical functions of applications that use the database. The priority of application functions cannot be expressed in existing solutions, therefore, another approach may be needed. In this paper, a session-based method is presented, where application functions are represented as sessions, by building and using language models based on previous observations. With this technique, a similarity measure can also be defined, to interpret minor differences between sessions caused by program logic, as similarity. If usage patterns appear on user level as well, it is reasonable to construct user groups along similar behavior, to utilize such patterns. As the most significant part of an autonomous solution is forecasting, a method is also presented to predict future workload characteristics, by identified user groups. Then, this approach has been evaluated in practice, mainly to determine the optimal corpus size and validate session recognition.

Keywords: self-tuning database; autonomous system; dynamic database tuning; response time optimization

1 Introduction and Related Work

Optimization of database systems is quite a hard and diversified task, thus several subtopics can be identified. The main goal of them is usually common, namely to improve a numeric value, which characterizes database performance well. This metric may be throughput, response time or a sort of resource utilization of executed queries.

It is desirable to be concerned with the exploitable opportunities on the level of database management systems, and particularly the application which uses the database. Database administrators often try to find the most proper configuration and index set manually, but for this purpose, hundreds of parameters, helper

structures and settings should be chosen appropriately, furthermore future use cases and their correlations, patterns should be known in advance. This may be a very hard and time-consuming activity. Therefore, automatic solutions may be used to make this task less painful.

Offline advisor tools have already been implemented in commercial database management systems [1, 2]. These tools require a preliminary workload history, which describes future usage patterns well. This approach can only be used in really special cases, as it only generates an initially optimal configuration (e.g. index set), thus changes in workload patterns will not be taken into consideration.

In practice, it has proven to be a strict constraint to work with. Consequently a new, online approach was needed to handle dynamically changing workload patterns. The idea is based on the offline approach, but recommendations are generated repeatedly for predefined time intervals. This method would be more accurate when changes are tracked by observing the actual workload, thus new advises are generated by need. Two ways can be identified to manifest the recommended configuration changes: a database administrator can check these changes, or they can be executed in an automatic way. As generation and deployment might be a time-consuming task, adapting the database configuration to the workload change might be late [3].

By forecasting the expected change points, automatic, self-tuning database systems can be achieved. For this, it is essential that the change points are following some identifiable pattern and are not random. Such solutions are already implemented as part of commercial database systems, such as SQL Server 17+ and Oracle 18c auto-tuning functions, but an open-source solution called Peloton is also available.

Automatic tools are based upon an assumption that databases are often used by a higher-level application. Therefore, usage patterns appear along application functions. Nevertheless, these applications might have non-functional requirements that defines some features response time-critical. For these kinds of applications, it is not always a good idea to choose global response time as a value to optimize, because an automatic tool with such strategy would set back the response time of a rare, but critical function. As an example, consider a critical, insert-heavy operation, and a noncritical read operation on a concrete column of a table. Let us suppose that the read operation is called much more often than the insert. In this case, the global optimization technique is going to decide to create an index to decrease the response time of reads, even though response time of the critical function is going to increase because of the maintenance cost of the created index.

Nevertheless, all of the already existing tools mentioned above are trying to optimize a global value. Accordingly, a new approach is needed to get rid of this limitation.

The main contributions of this paper are the following:

- a method to overcome the above mentioned limitation,
- a technique to characterize users and their operations, to take users into consideration, if some of them behave similarly,
- a prototype implementation of the proposed approach, and its evaluation.

2 Sessions and User Groups

If priority values were assigned to each function of the application, the limitation, caused by the global optimization technique, could be solved during the deployment of indices. At the same time, as initial knowledge comes only from query logs, it is not trivial to identify these functions on the level of individual queries. For this, a language model is going to be constructed.

A function can be modeled as a sequence of query-transactions, which is called a session, and a transaction is defined as a sequence of queries. It is clearly visible that a method is needed to transform sequences of queries to sessions, based on the log, furthermore to create groups of the identified sessions, as multiple instances of the same – or almost the same – session are expected to appear in the log.

Sessions are issued by users of the upper-level application. It is desired to take users into consideration, if some of them are behaving similarly, for example, several bank clerks in an administration system of a bank are invoking similar functions of the system, as they fulfill the same work. A user can be described by the type of the invoked sessions, and users who invoke similar sessions can be grouped by this fact.

3 Life-Cycle of Self-Tuner Database Systems

The high-level behavior can be described by the observation-forecasting-reaction cyclically repeating triplet. Autonomous databases actively watch query history to reveal past correlations, then give predictions for the future based on the recognized patterns, and finally perform the corrections at a suitable time.

Autonomous database systems are expected to be fully automatic, that is, they should operate reliably without human intervention. This method does not require any database administrator to continuously fine-tune the system according to the

changes in the patterns. Note that the primary guarantee of the automatism is an appropriate forecaster component, which predicts based on past observations.

These systems proceed on a query log, which contains various information about past queries, such as the time when the query was executed, query text, executed physical plan and response time. The sequence of queries found in the log is characterized to retrieve relevant properties, on which a grouping could be based. A dimensionality reduction is needed, because making predictions for each query one by one would be too expensive. Sought solutions have identified three types of properties altogether, and each of them selected one, which suited their goals the most. By selecting physical properties to characterize queries, it is assumed that the most important attribute of a query is its resource utilization. It might be true, but it is not rewarding to build prediction systems upon physical properties, because changes of the schema, table sizes or the amount of resource affect these values. Another option is to choose logical properties, namely accessed columns and syntax tree of the executed query. Unfortunately, pattern recognition is quite hard in this case [4]. Grouping by similar arrival rate history of templates can be seen in QueryBot 5000 [4], which founds on the Peloton framework [3]. Firstly, the arrival rate patterns of query templates are collected. Then, a collection is created for each template by sampling these arrival rates, thus templates are clustered based on cosine similarity of the collections. It can be observed that queries of the same transaction belong to the same cluster, but disadvantages are exposed during clustering. Unfortunately, infrequent, but long-running queries are regarded as noise, furthermore similar transactions may belong to different groups because of minor differences caused by program logic.

After picking a predictable property regarding the constructed groups, forecasting models are built. By using such trained models, information about future behavior is obtained, then optimization is done based on that. Self-tuner systems should prepare for changing behavior, that is, new queries may appear, already observed ones may be replaced to another group or even disappear, therefore prediction models possibly become invalid.

4 Session Identification and Clustering

Identifying sessions is an appropriate way to group queries as well. In opposition to QueryBot 5000 [4], program logic differences are handled well, moreover, important queries are not considered as noise.

4.1 Preprocessing Query Logs

To handle similar queries together, the text of the queries should be canonized. The first step is to apply semantics-preserving transformations, such as regularizing alias names and transforming logical expressions into normal form. Existing solutions can be found [5] for this purpose. After this step, the parameters of the queries should be masked.

Equivalence of canonized queries can be defined based on the equivalence of the query texts, the similarity of execution time and size of the result table. By aggregating query templates that were considered equivalent, the number of considered data decreases. Unique identifiers are assigned for each template to reference them in the next steps.

4.2 Identifying Sessions

The goal would be to retrieve sessions from the previously cleaned data to recognize application functions. This process has two phases. Firstly, a statistical language model is built, then the bounds of sessions can be defined using the language model [6].

4.2.1 Language Modeling

However, defining a timeout would be an obvious way to separate a query sequence into sessions, it is not sufficient, as a timeout does not always indicate the beginning of a new session. More accurate splitting could be given by recognizing queries that frequently occur consecutively. For this purpose, a language model is going to be used, which is a probability distribution over arbitrary word sequences. By using a language model, the probability of a word sequence can be estimated.

Let us formulate the analogy with the language modeling domain, then clarify the previous statement formally. The identifiers of query templates can be regarded as words, and sessions made from them as sentences of a language. A sequence of sessions denotes the whole text. Let $s = \langle q_1, \ldots, q_n \rangle$ be a session candidate sequence, consisting of template identifiers. The probability whether s is a valid sentence of the language can be defined using the chain rule.

$$P(s) = P(q_1) \ldots P(q_n | q_1 \ldots q_{n-1}) = \prod_{i=1}^{n} P(q_i | q_1 \ldots q_{i-1}) \quad (1)$$

A type of language models is called n-gram, which assumes that the probability of the n^{th} element of a sequence depends only on the previous n-1 elements. That is, P(s) can be estimated as follows.

$$\prod_{i=1}^{n} P(q_i | q_1 \ldots q_{i-1}) \approx \prod_{i=1}^{n} P(q_i | q_{i-n+1} \ldots q_{i-1}) \quad (2)$$

The method works as follows. To train the language model, a long corpus is needed, which is preferably separated into sentences. Then, the sentences are separated to n-tuples of words, and also statistics are calculated. After that, the probability of n-grams can be determined from their relative frequency, and the probability of a sequence can be estimated using the equation above.

As training data for the model, session sequences are needed that describe patterns between queries well. Initially, sentence boundaries are not known, therefore well-separated training data is not available. Yet, login/logout of users or a predefined timeout would help to find boundary points, thus more accurate result can be achieved. This technique is called semi-supervised learning.

4.2.2 Neural Language Models

As n-gram models calculate probability by the number of query co-occurrences, complex correlations unfortunately cannot be noticed. For the purpose, neural language models can be used, which became popular in the field of language modeling [8]. Often recurrent neural networks, especially LSTMs are chosen. The task is to estimate the probability of a word based on a given context, which can be considered as a multiclass classification problem. Defining the length of the context (l_c) is not trivial, as it depends on the observable patterns. The simplest case is when $l_c = 1$, that is, the probability of each word is calculated by the previous word.

To train the model, 3-dimension input tensors are created with rank (batch size, l_c, 1), furthermore the expected template identifier is defined for each input. Note that the expected identifier should be one-hot encoded. The architecture consists of an embedding layer, which maps positive identifiers to vectors, a hidden LSTM layer, and finally a dense output layer with softmax activation function. Thus, for a template identifier input sequence with length l_c, a vector of probabilities is given as output, which can be thought of as a probability distribution. The i^{th} element of the vector is the probability that the i^{th} query template follows the input context based on the learnt language. Thus, the probability of a given template can be selected from this output vector. Note that if l_c is greater than the minimal size of a session, then a pseudo-identifier is needed to expand the size of the input sequence according to the rank of the input tensor.

4.2.3 Determining Session Boundaries

At this point, a trained language model is given, which is able to estimate the probability of arbitrary template sequences based on the foreshown training data. The primary assumption is that the level of uncertainty is roughly constant in a session, and it is measurable by empirical entropy [6]. Let SC be a sequence of template identifiers. Then, the probability, which is noted by P(SC), can be

estimated by the trained language model. In this case, empirical entropy is defined as follows.

$$H(SC) = -\frac{1}{n} \cdot \log P(SC) \tag{3}$$

With this, moving along an observed template identifier sequence, an identifier is added to SC in each step. In addition, the probability of the sequence is calculated in each step. Thus, the entropy can also be defined, and when it changes more than a predefined threshold, the start of a new session is found. Note that another dataset is needed for this operation than for language model training.

With this method, a long, raw sequence of query template identifiers can be separated into coherent subsequences, thus sessions can be retrieved that represent higher-level tasks of users.

4.3 Session Clustering

During the previous step, some potentially equivalent sessions are identified. Therefore, it would be desired to filter these duplicated sessions, and create groups which consist of similar sessions. This is a two-phase operation.

The first step is to form session classes. Let s_1 and s_2 be identified sessions. They belong to the same class if either of them is a subsequence of the other, or their Jaccard similarity is sufficiently high (e.g. exceeds a threshold limit, which is near 1), furthermore long-running queries should come from the same templates. High Jaccard similarity means that they share almost the same query templates. Let the representant element of the class be the longest session, which usually has the largest execution time as well. This step is inspired to correct minor, a few queries long sliding errors occurred during the identification phase, and to reduce the complexity of the next step. As it can be observed, the number of session representants to maintain is significantly reduced after this step, and converges to the number of task-types issued by users. The algorithm is the following. Initially, each session belongs to a separate class. Then, by using pairwise comparison, the condition of joining is checked. If it is true, the classes should be joined together. For this approach, the actual location of each session should be tracked. As pairwise comparisons are used, the method finishes in $O(n^2)$ time, where n means the number of sessions.

As queries often arrive by a program logic, several similar sessions may be found among class representants. For example, let s_1 = <1, 1, 2, 3, 4> and s_2 = <1, 1, 1, 2, 1, 4>, sessions represented by template identifiers. They can be considered similar, because clues of program logic are present. The number of the first elements are probably originated from a loop, and the cause of the difference between the second last identifiers may be a conditional statement. Let us call this program logic based similarity. This value can be measured by the Needleman-Wunsch algorithm [6, 7], which is a common method in bioinformatics. The goal

is to find the most exact matching between two sequences based on a score to maximalize. The scoring system is defined as follows. Matches are rewarded by 3 points, mismatches and insertion/deletion actions worth less (e.g. 1) points, thus sessions with more mismatches can be considered less similar. It is easily discernible that mismatches and insertion/deletion actions mean loops and branches.

The final output score is then defined as the weighted sum of points. The number of matches (nMatches), loops (nLoops) and branches (nBranches) are also considered during the score calculation. To construct a similarity measure (noted by PLS), this value should be normalized. For this, the highest reachable score is used.

$$PLS(s_1, s_2) = \frac{3 * nMatches(s_1, s_2) + nLoops(s_1, s_2) + nBranches(s_1, s_2)}{3 * \max\{|s_1|, |s_2|\}} \quad (4)$$

The borders of the function values and other similarity properties are easily verifiable.

$$0 \leq PLS(s_1, s_2) \leq 1 \quad (5)$$

By using program logic similarity and Jaccard similarity (noted by JS), a distance function is defined to group session classes.

$$d_{SC}(sc_i, sc_j) = 1 - \varepsilon_1 \cdot JS(sc_i, sc_j) - \varepsilon_2 \cdot PLS(sc_i, sc_j) \quad (6)$$

ε weights should be determined to sum up to 1. Since weighted similarity measures are subtracted from 1, d_{SC} is trivially a distance function.

Finally, an appropriate clustering algorithm should be picked. Methods that need the number of clusters as a parameter are certainly not suitable for this problem, as it cannot even be estimated in general. Furthermore, the algorithm should not reckon a session class as noise, just because it does not have enough neighbours to form a new group.

Regarding these aspects, using hierarchical, agglomerative clustering is reasonable. Initially, each session class forms a cluster, then by calculating the distance between them, some of them are merged. The procedure stops when there is only one cluster left. An appropriate clustering can be obtained by stopping earlier, using stopping criteria.

To put it into practice, besides the distance function between session classes, a distance function between clusters and stopping criteria is needed. Two possible options are identified to define distance between clusters, by reducting to elementwise distance. One of them is the centroid method, but choosing a central element is not trivial, as there is only a relative distance measure is defined between session classes. Therefore, single linkage method is chosen, which means the distance of the closest elements. Let $C_1 = \{sc_{11},\ldots,sc_{1m}\}$ and $C_2=\{sc_{21},\ldots,sc_{2n}\}$

be clusters of session classes. Then, D_{min} is defined as follows, based on d_{SC} distance.

$$D_{\min}(C_1, C_2) = \min\left\{d_{SC}\left(sc_{1i}, sc_{2j}\right) \middle| \; i \in [1..m], \; j \in [1..n]\right\} \quad (7)$$

Note that as $0 \leq d_{SC}(a,b) \leq 1$ for all a and b session class, $0 \leq D_{min}(A,B) \leq 1$ is also true for all clusters of session classes.

Using this function, distance-based stopping criteria can be also defined. The algorithm should stop when no more clusters are found near each other. Nearness is bounded by an ε_3 threshold.

5 User Clustering

If patterns are observable on the user level, it is desired to create user groups based on their behavior. The behavior is characterized by the distribution of the executed session types. Therefore. an n-dimensional vector is defined for each recognized user, where n is the number of recognized session clusters. The i^{th} element means the number of observed sessions which belong to the i^{th}session cluster. By dividing these values by the total count, ratios are obtained. These ratios are useful to determine the probability that an incoming template belongs to the given session group. Thus, a discrete probability distribution is defined for each user. Based on the similarity of the corresponding distribution, users can be categorized. Several methods can be found to define similarity between distributions. One of them is Hellinger-distance, which is defined using Bhattacharyya-coefficient [9]. Let p_1 and $p_2$probability mass functions of different discrete distributions. Then, Hellinger-distance is defined as follows.

$$H(p_1, p_2) = \sqrt{1 - BC(p_1, p_2)}, \; where \; BC(p_1, p_2) = \sum_{i=1}^{n} \sqrt{p_1(i) \cdot \; p_2(i)} \quad (8)$$

Since conditions are similar to clustering of session classes, hierarchical, agglomerative clustering is appropriate by using distance metric H here as well. Nearness is bounded by an ε_4 threshold.

6 Predicting Workload Patterns

The most important part of a self-tuning database system is the forecasting component, which predicts future behavior by recent recognitions. The concrete property to forecast depends on the chosen optimizable value and the method used during optimization. In QueryBot 5000 [4], a prediction model is created for each query template cluster, which predicts the amount of arriving queries for a given time interval ahead, based on recently collected arrival rates.

Similarly, it might be necessary to gain information about future queries for automatical index maintenance, e.g. how many of them will be executed and by which users. To utilize usage pattern similarity of users, let us create prediction models by user clusters. Then, the estimation of three potential properties could help answer the previous questions.

- Arrival rates, as in QueryBot 5000 [4]. In this case, an assumption should be made that the distribution of sessions is roughly constant in a group, as it is not predicted.
- The probability distribution of executed queries. Nonetheless, by picking this property, information about arrival rates is not available. Thus, a quite strong restriction should be introduced, namely, the number of executed queries cannot change significantly. Obviously, it is not true in general.
- The frequency distribution of executed queries. By combining the two properties above, information can be obtained about both distribution and quantity of sessions. Hence, this property may be the most practical choice.

Presence of patterns in the values is essential for proper forecasting, e.g. increasing or decreasing trends, cyclic behavior, periodically slow increasing then fast decreasing, etc. [4], otherwise reasonable and meaningful predictions cannot be given. However, unfortunately a general method to recognize all kinds of patterns universally does not exist, an ensemble model can be constructed [3, 4], which is a combination of prediction models to recognize as many patterns as possible.

As an example, let us briefly consider how may a prediction model be constructed to identify and forecast cyclic patterns of arrival rates and frequency distributions. For this purpose, just like for language modeling, LSTM is an appropriate choice [4]. For simplicity, let us give a solution for arrival rate prediction at first, then it is going to be generalized for frequency distributions easily.

Training data can be actually interpreted as parts of time series, linearly mapped to [0..1] interval. For the normalization, a maximal arrival rate value is required, which may be the largest value observed during the learning phase (session and user recognition), or a theoretical value, received as a parameter. Then, the sequence is split to its k+1 long subsequences. Finding the proper value of k depends on the length of the context which affects the $k+1^{th}$ value. The model is then trained by the subsequences, by choosing an appropriate n_b batch size. For these, 3D input tensors are constructed with rank (n_b, k, n_f). In case of arrival rates, $n_f = 1$, but for frequency distributions n_f equals to the number of session clusters.

The architecture of the neural network is the following. As a hidden layer, two LSTMs are used. The activation function of the dense output layer is sigmoid,

therefore the output for a k-long sequence is a value, the predicted next element of the time series. In case of frequency distributions, this value is a vector, which represents the next, expected distribution of the following time unit. Since predictions should be usually made for multiple time units, the single forecasting method should be called continuously by utilizing the previously predicted value.

Note that groupwise model training can happen in parallel, even on separate processors, as these tasks do not depend on each other.

7 Live Phase

At this point, clustered users, their typical sessions, and a prediction for each user group are available. To generate helper structures, several already existing solution can be found [10, 11]. Nevertheless, it is practical to select significant queries that raise execution time the most as an input for the recommender system. Significant queries can be selected by defining a threshold, or by retrieving queries that run longer than the corresponding time slice. Before the final deployment, every index to be materialized should be checked whether it is consistent with predefined function priorities.

After model training, predictions and recommendations are generated periodically during the live phase of the system. Moreover, recognitions should be maintained based on continuously arriving query logs, as new queries, sessions, or even users may appear, or already existing ones disappear. They may change clusters, thus prediction models should be retrained.

8 Evaluation

A prototype has been created for demonstration and experimental purposes. The main goal is to compare and evaluate n-gram and neural network-based language models, by finding the minimal corpus size for training, and evaluating their session recognition ability. The advantage over already existing solutions is also demonstrated. Time complexity and the number of maintained models is specifically measured and compared to the most developed open source solution, Peloton [3]. Free parameters are also selected via example executions.

8.1 Experimental Setup

However, an appropriate workload history is indispensable for the evaluation, producing such history manually is quite infeasible. Thus, a program has been

created to generate workloads, by connecting to a test database, then executing transactions that follow a predefined pattern. Query logging should be enabled to retrieve the required data.

The concrete database schema, fitting data and transactions are received from the OLTPBenchmark project [12], which provides a framework for database performance testing. CH-benCHmark, which is built upon TPC-C standard, consists of OLTP and OLAP transactions, so it suits our needs. Furthermore, since a workload for several hours or days may be needed, rescaling should be done to transform the workload to a longer time interval. Generated workloads can be found in Table 1.

Table 1
Generated workloads for the evaluation

	test1.log	**test2.log**
Experiment 1	1 000 queries	a few seconds long history, 4 transactions
Experiment 2	10 000 queries	a few seconds long history, 4 transactions
Experiment 3	15 000 queries	a few seconds long history, 4 transactions
Experiment 4	43 000 queries	a few seconds long history, 4 transactions
Experiment 5	10 000 queries	4 hours long history with periodic patterns, 26 000 queries
Experiment 6	10 000 queries	6.5 hours long history with periodic patterns, 52 000 queries

Training queries for language models are in test1.log files, and test2.log files contain queries for session recognition and forecasting. With experiment 1, 2 and 3, the length of the corpus is determined, so that language models can identify session bounds sufficiently. With the other experiments, the execution time and dimensionality reduction are measured.

For the evaluation, a local PostgeSQL server was used, on an average PC configuration with 8 GB RAM and 4 cores. For logging, auto_explain option is set to true.

8.2 Determining Corpus Size

Firstly, let us identify the corpus size that is sufficient for the language models to recognize session bounds well. Beforehand, n for the n-gram model should be

defined. Let us choose n = 5, as training does not take too much time with this choice for large data, but a quite exact result is achievable.

At the first experiment, test1.log contains 1000 out of the 43000 queries, scaled to 1 day. Recognition of the neural language model is not exact, as initially 6 clusters were identified, because it has needlessly split sessions at multiple points. After the fine-tuning of the epoch number and the size of the LSTM layer, the number of clusters became 5, by less splitting. Although the 5-gram model identified 4 session clusters, it also split sessions at undesired points, just like the neural model.

During the second experiment, language models were trained with 10000 queries. The results were similar to the first experiment, but the loss function (mean squared error) of the neural model began to take its appropriate form.

Table 2
Summary of the evaluation results

	5-gram model	**Neural model**
Experiment 1	4 session clusters, but undesired splits	initially 6, then 5 session clusters, undesired splits
Experiment 2	4 session clusters, but undesired splits	5 session clusters, less undesired splits
Experiment 3	4 session clusters, proper session recognition	5 session clusters, identifying more correlations, form of loss function is correct
Experiment 4	4 session clusters, proper session recognition	4 session clusters, proper session recognition
Experiment 5	model training: 1637 ms recognition: 2907 ms clustering: 5825 ms 4527 recognizedsessions, 105 session classes, 23 clusters	model training: 58043 ms recognition: 38100 ms clustering: 10565 ms 3232 recognizedsessions, 82 session classes, 40 clusters
Experiment 6	model training: 1324 ms recognition: 3523 ms clustering: 5018ms 9122 recognizedsessions, 114 session classes, 21 clusters, 2 user groups	model training: 62927 ms recognition: 65481 ms clustering: 6867 ms 6447 recognizedsessions, 109 session classes, 41 clusters, 2 user groups

The 5-gram model recognized properly session bounds during the third experiment, hence a workload of 10000-15000 queries is necessary for corpus.

The neural model still identified 5 clusters, but the form of the loss function became correct. Based on the results of the fourth experiment, at least 40000-45000 queries are needed as a corpus to train the neural model, thus 4 clusters are identified. The fifth and sixth experiment cover two general cases, where particularly dimensionality reduction and execution time were inspected.

It is clearly visible that the 5-gram model often splits workloads for more sessions than the neural model, which can recognize more complex correlations. The neural model identifies more clusters by utilizing inter-transactional patterns, than the 5-gram model, which rather splits workloads along executed transactions.

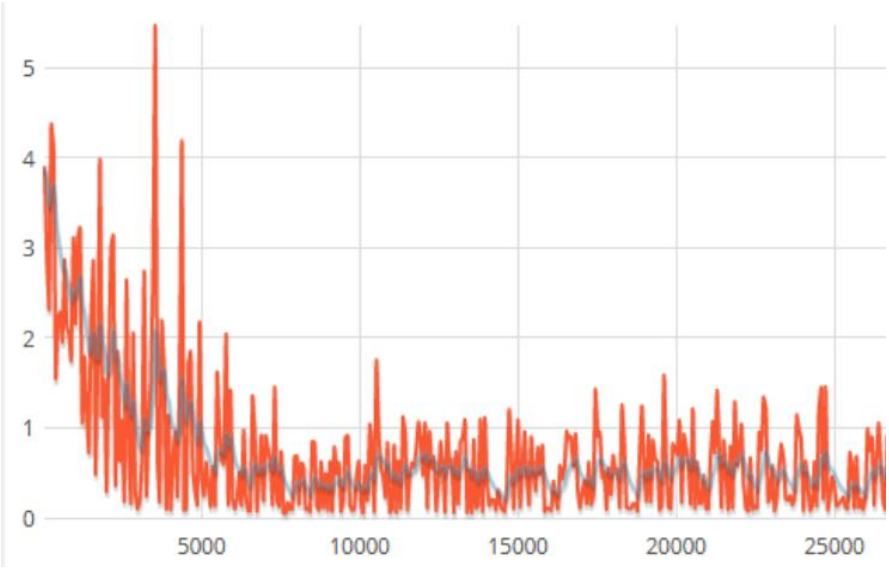


Figure 1
Loss function over the iteration count of experiment 3

8.3 Dimensionality Reduction and Execution Time

Workloads that have similar characteristics (e.g. cyclic behavior pattern, 2 user groups) and consist of 52000 queries, are split to 9122 sessions by the 5-gram model, and 6447 sessions by the neural model. QueryBot 5000 [4] templatizes incoming queries, then groups them by column access similarity, thus 334 templates are retrieved. The 5-gram and neural models construct 21 and 41 session clusters, respectively. Based on the clusters, user groups are created, then forecasting models are built. In contrast, QueryBot 5000 identifies 107 template clusters, and assumes that the top 5 largest clusters cover important patterns, therefore, prediction models are built only for them. It may be a strong

assumption, furthermore in this case, more models are created, thus space and time complexity is larger. Summarized results can be seen in Table 2.

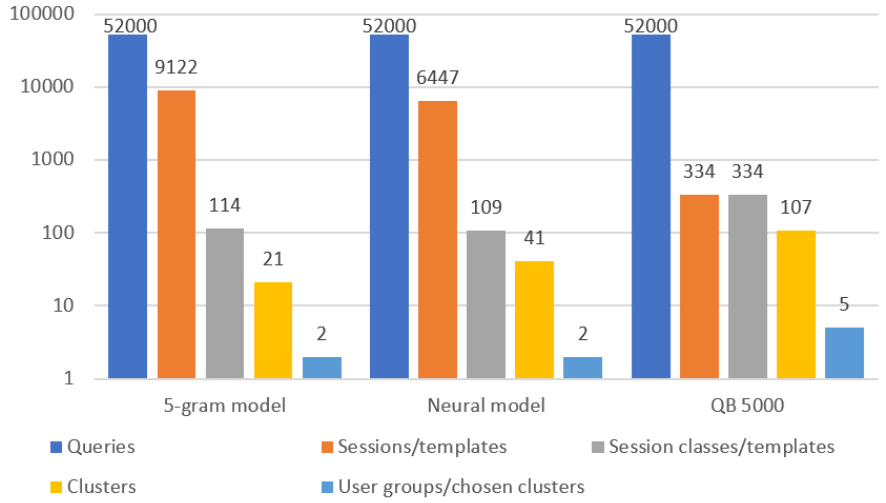


Figure 2
Comparing query number reduction on a logarithmic scale

Session recognition is the first subtask which may have significant execution time. As seen above in the presented approach, it is a two-phase process, which has the same goal as grouping query templates by the cosine similarity of their global arrival rate pattern histories, as seen in Peloton [3] and QueryBot 5000 [4]. Thus, the execution time of these subtasks should be compared.

Table 3
Comparison of execution times, 52 000 queries

	QueryBot 5000	**5-gram model**	**Neural model**
Preparing	8 162 ms (templatizing)	1 324 ms (model training)	62 927 ms (model training)
		3 523 ms (session recognition)	65 481 ms (session recognition)
Clustering	46 425 ms	5 018 ms (session)	6 867 ms (session)
		1 551 ms (user)	1 423 ms (user)
Total	≈ 54.5 sec	≈ 11.4 sec	≈ 136.5 sec

As we can see, the time complexity of QueryBot 5000 is larger than the 5-gram model, but the neural model performed much worse because of slower model training. As it can be seen in Table 2, the time cost of 5-gram model training and session recognition did not increase significantly because of larger data size. In the

case of the neural model, these phases have the most notable time complexity, but the time needed for clustering seems roughly constant.

By modifying the test data that is provided by the authors of QueryBot 5000 to contain similar sessions, it is verifiable that it does not recognize program logic similarities, and such sessions are going to be placed into different groups. On the other hand, by running experiment 5 and 6, it is shown that the demonstrated session-based method is able to recognize such similarities.

8.4 Evaluation Summary

Let us summarize the results of the evaluation above. It has been recognized that Peloton [3] and QueryBot 5000 [4] which are regarded as the most developed open-source, noncommercial self-tuning solutions available, cannot recognize similarity based on program logic. Thus, potentially similar user tasks are arranged to different clusters. This problem is eliminated by the session-based approach, and two language model implementations were compared. By comparing identified sessions and execution times, it became clear that neural models can also recognize inter-transactional patterns, but its cost is the notably slower training.

Several free parameters have been left undefined in the demonstrated method, which were selected empirically during the evaluation. The summary of these parameters and their suggested values can be found in Table 4.

Table 4

Suggested values of free parameters

Parameter	**Value**	**Description**
ε_1	0.3	The weight of Jaccard-similarity
ε_2	0.7	The weight of program logic similarity
ε_3	0.2	The maximal distance of a new element from a session group
ε_4	0.2	The threshold of Hellinger-distance for users in the same group

Conclusions

In this paper, a session-based approach was presented for autonomous database tuning, to let self-tuning systems take upper-level application requirements into consideration. In contrast with available solutions, the concept of session was made explicit, as it is an important notion for such systems. Because of this, sessions with minor differences due to program logic can be recognized as similar

much easier. For session recognition, details of a neural language model were specified, based on n-gram models. It was shown in practice that the neural model performs much slower due to the model training than the 5-gram model, but it did recognize inter-transactional, more complex patterns.

A method to characterize users and construct groups was elaborated, then for each user group a way to define appropriate forecasting models was presented, thus two properties of future workloads became predictable.

Further research would be beneficial to determine the correct training data size of the language models in general, as it can be different for each application. Experimenting with LSTM language models would be also practical. By building deep, more complex architectures, exciting results could be achieveable.

Actually, identifying sessions and clustering users is useful not solely for database tuning, but query recommender systems as well. When queries are manually written, it is desired to get automatic help, such as autocomplete or prediction of the next query, based on observed habits of the user group.

Acknowledgement

The project has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002).

References

[1] B. Dageville, D. Das, K. Dias, K. Yagoub, M. Zait, M. Ziauddin: AutomaticSQL Tuning in Oracle 10g, Proceedings of the 30th International Conferenceon Very Large Databases, 2004, pp. 1098-1109

[2] S. Agrawal, S. Chaudhuri, L. Kollar, A. Marathe, V. Narasayya, M. Syamala:Database Tuning Advisor for Microsoft SQL Server 2005, Proceedings of the 30th International Conference on Very Large Databases (VLDB) 2004, pp.1110-1121

[3] A. Pavlo, G. Angulo et al.: Self-Driving Database Management Systems, CIDR2017, Conference on Innovative Data Systems Research, Vol. 10, 2017, pp. 781-792

[4] L. Ma, D. Van Aken, A. Hefny, G. Mezerhane, A. Pavlo, G. J. Gordon: Query-based workload forecasting for self-driving database management systems, Proceedings of the 2018 International Conference on Management of Data, ACM, 2018, pp. 631-645

[5] G. Kul, D. T. A. Luong, T. Xie, V. Chandola, O. Kennedy, S. Upadhyaya:Similarity Metrics for SQL Query Clustering, IEEE Transactions on Knowledgeand Data Engineering, Vol. 30, 2018, pp. 2408-2420

[6] Q. Yao, A. An, X. Huang: Finding and Analyzing Database User Sessions,Database Systems for Advanced Applications 2005, 2005, pp. 851-862

[7] J. Aligon, M. Golfarelli, P. Marcel, S. Rizzi, E. Turricchia: Similarity Measuresfor OLAP Sessions, Knowledge and Information Systems (KAIS) Vol. 32, 2014, pp. 463-489

[8] M. Sundermeyer, R. Schlüter, H. Ney: LSTM neural networks for language modeling, Thirteenth annual conference of the international speech communicationassociation, 2012

[9] K. G. Derpanis: The Bhattacharyya Measure. Mendeley Computer, 1(4) 2008, pp. 1990-1992

[10] K.-U. Sattler, M. Luehring, K. Schmidt, E. Schallehn: Autonomous Management of Soft Indexes, Data Engineering Workshop, 2007 IEEE 23rd International Conference on. IEEE, 2007, pp. 450-458

[11] K. Schnaitter, S. Abiteboul, T. Milo, and N. Polyzotis: On-line index selection for shifting workloads. ICDEW ’07 Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop, 2007, pp. 459-468

[12] D. E. Difallah, A. Pavlo, C. Curino, P. Cudre-Mauroux: OLTP-Bench: Anextensible testbed for benchmarking relational databases, Proceedings of theVLDB Endowment, Vol. 7, 2013, pp. 277-288

Model Predictive Control of Hybrid Transformer with Matrix Converter

Paweł Szcześniak, Grzegorz Tadra, Zbigniew Fedyczak

Institute of Electrical Engineering, University of Zielona Góra
ul. prof. Z. Szafrana 2, 65-516 Zielona Góra, Poland
P.Szczesniak@iee.uz.zgora.pl, g.tadra@uesa.pl, Z.Fedyczak@iee.uz.zgora.pl

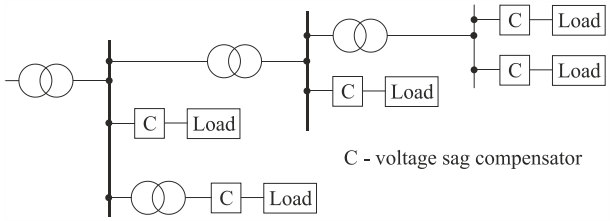
Abstract: This paper proposes a model predictive control for a hybrid transformer (HT) with matrix converter (MC). The proposed HT system is composed of a power transformer connected to a power electronic converter which controls the output voltage on the secondary side. The proposed solution of HT is used to resolve power quality issues and power flow control in the electrical power distribution system. The advantages of predictive control are its very simple realization and the possibility to optimize various predefined criteria. Presented in the paperr is a discussion of the conception of a model predictive control strategy to compensate for voltage sags, swells and harmonics on the grid side by providing continuous AC voltage regulation of the HT. Also, presented are the preliminary simulation results, which verify the high performance of the proposed predictive control.

Keywords: Model Predictive Control; Power Transformer; Matrix Converter; Voltage Compensator

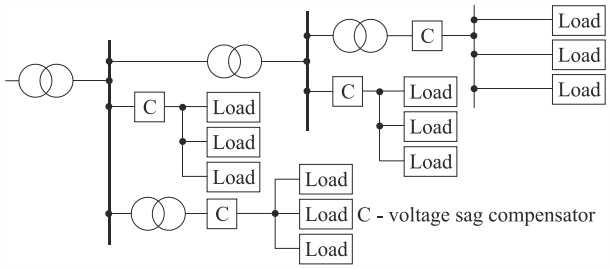
1 Introduction

Grid voltage sags and swells that propagate in the distribution grid may have a great impact on electric devices, and may consequently reduce their life cycles and generate additional economic costs [1], [2]. Modern systems of industrial plants are being established with high-tech equipment for increased productivity that requires a higher demand for quality power. Reducing the impact of supply voltage changes on the operation of electrical devices requires the use of various improvements throughout the power energy system. Two main compensation methods of voltage-disturbance in power systems occur: parallel – "current" (SVC – static VAR compensator and STATCOM – static synchronous compensator,) and series – "voltage" (DVR – Dynamic Voltage Restorer). In addition, conventional transformer with a tap changer are used to compensate seasonal voltage changes or low incidence occurances.

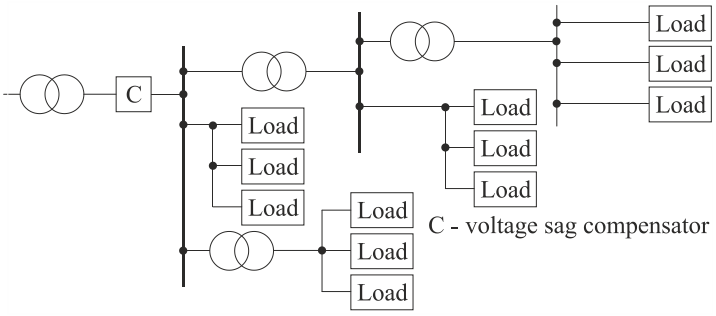
One of the solutions for power quality problems is the use of power electronic systems that enable various ways to compensate for existing disturbances [3]-[5]. The group of compensators which are based on power electric devices (STATCOM, SVC, and DVR) guarantees good dynamic properties of the compensation process. Depending on voltage compensators position in the electrical system we can distinguish: (i) separate compensator – local, (ii) group compensator and (iii) central compensator. The above methods of compensators arrangements are presented in Fig. 1. Separated compensation – applies to direct cooperation between compensator and load. The group and central compensation – allow the use compensators for a selected industrial plants, buildings or parts of the power network.



(a)



(b)



(c)

Figure 1
Methods of location of voltage compensators: local (a), group (b), central (c)

In addition, the concept of various hybrid transformers has been proposed as a compensator in AC voltage amplitude changes [6]-[10]. HT is devices where a conventional electromagnetic transformer cooperates with a pulse-width modulation (PWM) AC/AC converter. In the literature, one can find many different solutions of hybrid compensators [10]-[12]. In this publication, HT with a matrix converter will be considered [10].

Proposed in publication [10] is the concept of a HT with a MC, but without indication of any automatic regulation algorithm in a closed feedback loop. In this article, static parameters such as the voltage regulation range and its phase shift angle for Space Vector Modulation [13] control have been determined and the basic control characteristics have been shown. In addition, examples of compensation for symmetrical and asymmetrical AC voltage changes have been presented. For symmetrical voltage variations in each of the supply lines, good compensating properties were obtained. However, for compensating unbalanced voltage distortions, the obtained results were not satisfactory. In the output voltages after compensation there were additional low orders harmonics. The closed loop control strategy with PI regulators has been proposed in the reference [11] and [12]. In presented algorithm the detection of both input and output voltages and the determination of compensating signals is requires. The proposed control was quite complex, and it requires optimization of PI regulator settings, which also depended on the parameters of the transformer used. Additionally, in order to precisely shape compensating signals with unsymmetrical strained supply voltage, complex algorithms of PLL synchronization loops should be used [14, 15].The obtained results presented in [11] and [12] for the compensation of asymmetric voltage sag indicates a lack of full compensation. Similar as in [10] the lower orders of voltage harmonics occur in the output voltages.

In the scientific literature many methods have been proposed for controlling power converters. Among them, one can distinguish: hysteresis control, linear, sliding, predictive, non-linear and using artificial intelligence. The indicated methods are characterized by different complexity and properties. More complex methods require significantly higher computational power of dedicated control platforms (DSP, FPGA), but they significantly improve the properties of controlled systems. In recent years, a lot of work has been devoted to the development of non-linear methods and with the use of artificial intelligence in various fields of technology such as automation, process control, telecommunications, logistics, transport and optimization of networks and processes [16]-[24].

With the control methods developed in recent times, predictive control has gained a lot of interest especially in the applications of power electronic converters and electric drives [25]. Predictive control includes a very wide group of controllers with very different approaches to operation. However, a common approach for all types of predictive control is modeling of the controlled system and determining the future state of the system, including the optimization of control criteria [25].

In this paper the model predictive control of a hybrid transformer with matrix converter is proposed. Model-based predictive control has been commonly used as a method to control load current, motor speed and torque control by different types of power electronic converter [25]. Furthermore, using the MPC in converters control allows control of different converter parameters such as: the switching frequency, input power factor (PF) and common mode voltage. Those parameters are control depending on the assumed cost function. Additionally, the MPC is usually employed in parallel compensators working as STATCOM to shape the compensator currents [26]. The novelty of this paper is the use of MPC in a series voltage compensator to shape the compensator voltages. The working principle of the MPC in the proposed AC voltage conditioner is to generation of the compensation voltage signals at different voltage distortion conditions. In addition to the generation of compensation voltages, an additional objective function has been introduced related to the adjustment of the input power factor. The proposed methods for controlling the series AC voltage compensator have advantages to classical control methods. A significant advantage is the compensation of both changes in voltage amplitude as well as asymmetrical and harmonic distortions. In addition, better dynamic properties are achieved.

Because the transformer has unity voltage ratio and the grid voltages are measured on the transformer secondary side, only the matrix converter with input and output filters models are included in the proposed control algorithm. In the presented preliminary stage of research, the parameters of the electromagnetic transformer were not taken into account. The influence of transformer parameters on the prediction of grid currents must be analyzed in the future in more detailed studies. At the current stage of research, tests of the proposed prediction algorithm with symmetrical and unsymmetrical voltage distortions will be carried out.

The structure of the paper is organized as follows. Sections 2 introduce hybrid transformer witch matrix converter topology. The main idea of the proposed control technique based on Model Predictive Control is shown in Section 3. In this chapter, the input and output filter models were determined and the prediction algorithm based on the allowed combinations of matrix converter switch states was defined. In addition, the method of determining compensation voltages and minimizing the set target functions has been defined. The results of simulation verification of the system are presented in Section 4. Finally, Section 5 concludes the paper and shows possible directions for future research.

2 Proposed HT with MC

The main HT system with MC proposed in accordance with [10] is shown in Fig. 2. A fundamental component in the provision of reliable electricity to the end-user is the distribution transformer with two secondary windings. The output

voltage of the whole HT is the sum of the MC output voltage and the second secondary winding voltage of the transformer. The transformer voltage ratios between primary and each secondary winding of the transformer are 1:1. Such voltage ratios allow the MC to be switched-off via bypass switches, while working with nominal amplitude value of the grid voltage. The MC can adjust the amplitude of the output voltage in the range from 0 to 0.866 times the supply voltage and adjust the phase of the output voltage within the whole range of its voltage change [13].

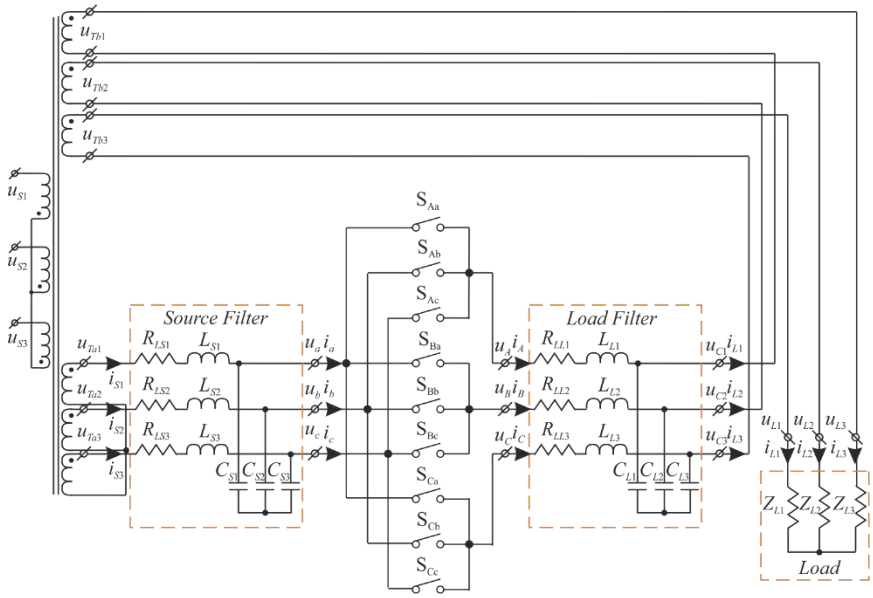


Figure 2
Proposed Hybrid Transformer (HT) with Matrix Converter (MC)

3 Model Predictive Control

The main characteristic of predictive control is the use of a model of the system for predicting the future behavior of controlled variables [25]. Obtained information is used by the control unit to obtain the optimal switch configuration, according to a predefined optimization criterion. Therefore, in the MPC, the controller model consists of: 1) a discrete input filter model; 2) the model of the converter determining the relations between the input and output terminals, 3) the discrete output filter model; 4) optimization criteria. On the basis of the mathematical model of the system, one state is determined for all allowed combinations of connectors, for which the optimization criterion has a minimum

value. The mathematical model of individual components and optimization criteria are described in this chapter.

We assume in this paper an idealized transformer with two taps both with a voltage ratio of 1:1. All three-phase currents and voltages are written in a complex form using Clark transformation:

$$\underline{x} = x_\alpha + jx_\beta \Leftrightarrow \begin{bmatrix} x_\alpha \\ x_\beta \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{1}$$

The continuous time system of source and load filter can be rewritten as follows:

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \Leftrightarrow s\mathbf{x}(s) = \mathbf{A}\mathbf{x}(s) + \mathbf{B}\mathbf{u}(s) \tag{2}$$

The input filter is described by the following equation:

$$\begin{bmatrix} \frac{d\underline{u}_{abc}}{dt} \\ \frac{d\underline{i}_S}{dt} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C_S} \\ -\frac{1}{L_S} & -\frac{R_{LS}}{L_S} \end{bmatrix}\begin{bmatrix} \underline{u}_{abc} \\ \underline{i}_S \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{C_S} \\ \frac{1}{L_S} & 0 \end{bmatrix}\begin{bmatrix} \underline{u}_S \\ \underline{i}_{abc} \end{bmatrix} \tag{3}$$

where: $\underline{u}_{abc}$ – the voltage at the source filter capacitors (MC input voltage), $\underline{i}_S$ – the source current, $\underline{u}_S$ – the source voltage ($\underline{u}_S = \underline{v}_{Ta}$), $\underline{i}_{abc}$ – the MC input current. All variables are described in complex form $x_\alpha + jx_\beta$. Similar, the output filter model has the following form:

$$\begin{bmatrix} \frac{d\underline{u}_C}{dt} \\ \frac{d\underline{i}_{ABC}}{dt} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C_L} \\ -\frac{1}{L_L} & -\frac{R_{LL}}{L_L} \end{bmatrix}\begin{bmatrix} \underline{u}_C \\ \underline{i}_{ABC} \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{C_L} \\ \frac{1}{L_L} & 0 \end{bmatrix}\begin{bmatrix} \underline{u}_{ABC} \\ \underline{i}_L \end{bmatrix} \tag{4}$$

where: $\underline{u}_{ABC}$ – the MC output voltage, $\underline{i}_L$ – the load current, $\underline{u}_C$ – the compensation voltage, $\underline{i}_{ABC}$ – the MC output current.

Using a bilinear transformation, we replace "*s*" with the following relationship:

$$s = \frac{2}{T_D}\frac{(z-1)}{(z+1)} \tag{5}$$

where T_D is the discretization time, we obtained digital space equations:

$$zx(z) = \mathbf{A}_\mathrm{d}x(z) + \mathbf{B}_\mathrm{d}u(z) \tag{6}$$

Individual matrices after digitization have the following form:

$$\mathbf{A}_\mathrm{d} = \left(\frac{2}{T_D}\mathbf{I} - \mathrm{A}\right)^{-1}\left(\frac{2}{T_D}\mathbf{I} + \mathbf{A}\right) \tag{7}$$

$$\mathbf{B}_\mathrm{d} = \left(\frac{2}{T_D}\mathbf{I} - \mathrm{A}\right)^{-1}\mathbf{B} \tag{8}$$

where **I** - unit matrix. Based on equation (6) we get the difference equations:

$$\begin{bmatrix} \underline{u}_{abc}(k+1) \\ \underline{i}_S(k+1) \end{bmatrix} = \mathbf{A}_{Sd}\begin{bmatrix} \underline{u}_{abc}(k) \\ \underline{i}_S(k) \end{bmatrix} + \mathbf{B}_{Sd}\begin{bmatrix} \underline{u}_S(k) \\ \underline{i}_{abc}(k) \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} \underline{u}_C(k+1) \\ \underline{i}_{ABC}(k+1) \end{bmatrix} = \mathbf{A}_{Ld} \begin{bmatrix} \underline{u}_C(k) \\ \underline{i}_{ABC}(k) \end{bmatrix} + \mathbf{B}_{Ld} \begin{bmatrix} \underline{u}_{ABC}(k) \\ \underline{i}_L(k) \end{bmatrix} \tag{10}$$

where $\mathbf{A}_{Sd}$, $\mathbf{B}_{Sd}$, $\mathbf{A}_{Ld}$, $\mathbf{B}_{Ld}$ – matrices of input and output filter space model after the digitization, respectively. Based on the models described in the recursive formulas (9) and (10), the values of the predictive source current $\underline{i}_S$ and the compensator output voltage $\underline{u}_C$ are determined as follows:

$$\begin{aligned} \underline{i}_S(k+1) &= a_{S21}\underline{u}_{abc}(k) + a_{S22}\underline{i}_S(k) + b_{S21}\underline{u}_S(k) + b_{S22}\underline{i}_{abc}(k) \\ \underline{u}_C(k+1) &= a_{L11}\underline{u}_C(k) + a_{L12}\underline{i}_{ABC}(k) + b_{L11}\underline{u}_{ABC}(k) + b_{S12}\underline{i}_L(k) \end{aligned} \tag{11}$$

where: a_{S21}, a_{S22}, b_{S21}, b_{S22} are the appropriate elements of the matrix $\mathbf{A}_S$ and $\mathbf{B}_S$; a_{L21}, a_{L22}, b_{L21}, b_{L22} are the appropriate elements of the matrix $\mathbf{A}_L$ and $\mathbf{B}_L$.

Other values in the model are measured directly in the system or calculated based on the following relationships:

$$i_{abc} = \mathbf{D}^{\mathrm{T}} i_{ABC} \tag{12}$$

$$u_{ABC} = \mathbf{D} u_{abc} \tag{13}$$

$$\mathbf{D} = \begin{bmatrix} s_{aA} & s_{bA} & s_{cA} \\ s_{aB} & s_{bB} & s_{cB} \\ s_{aC} & s_{bC} & s_{cC} \end{bmatrix} \tag{14}$$

In order to compensate for variations in the source voltages (voltage sag/swell, variations, harmonic distortion), the reference compensating voltage signals should be generated at the output of the MC (u_{Cref}), which signals are the differences between the expected value at the HT output (u_{Lref}) and the measured value of the supply network (u_S).

$$u_{Cref} = u_{Lref} - u_S \tag{15}$$

Hence, determining the compensation signal is relatively simple and results from the natural difference of the expected and measured signal. The HT output voltage is the sum of compensator and source voltages:

$$\underline{u}_L = \underline{u}_C + \underline{u}_S \tag{16}$$

In the prediction algorithm, in the next iteration step, 27 different solutions for 27 allowed switching combinations of MC switches are checked (Table 1).

Then, based on the obtained results, the cost function g is minimized [25], [27]. It is possible to minimize many functions of costs. In the presented preliminary research approach, the cost function regarding the accuracy of shaping the output voltage "g_1" and the shaping of the input power factor "g_2" are minimized. Most applications request a unity input power factor, therefore, the input reactive power Q_S value is minimized. The final cost function "g" is the sum of the individual cost functions with the selected weighting factors K_1, K_2.

Table 1
Switch configuration in the and corresponding output voltage and source current

No	a	b	c	S_{Aa}, S_{Ab}, S_{Ac}	S_{Ba}, S_{Bb}, S_{Bc}	S_{Ca}, S_{Cb}, S_{Cc}	u_{ab}, u_{bc}, u_{ca}	i_A, i_B, i_C
1	A	A	A	1 0 0	1 0 0	1 0 0	0 0 0	0 0 0
2	B	B	B	0 1 0	0 1 0	0 1 0	0 0 0	0 0 0
3	C	C	C	1 0 0	1 0 0	1 0 0	0 0 0	0 0 0
4	A	C	C	1 0 0	0 0 1	0 0 1	$-u_{CA}$ 0 u_{CA}	i_a 0 $-i_a$
5	B	C	C	0 1 0	0 0 1	0 0 1	u_{BC} 0 $-u_{BC}$	0 i_a $-i_a$
6	B	A	A	0 1 0	1 0 0	1 0 0	$-u_{AB}$ 0 u_{AB}	$-i_a i_a$ 0
7	C	A	A	0 0 1	1 0 0	1 0 0	u_{CA} 0 $-u_{CA}$	$-i_a$ 0 i_a
8	C	B	B	0 0 1	0 1 0	0 1 0	$-u_{BC}$ 0 u_{BC}	0 $-i_a i_a$
9	A	B	B	1 0 0	0 1 0	0 1 0	u_{AB} 0 $-u_{AB}$	i_a $-i_a$ 0
10	C	A	C	0 0 1	1 0 0	0 0 1	u_{CA} $-u_{CA}$ 0	i_b 0 $-i_b$
11	C	B	C	0 0 1	0 1 0	0 0 1	$-u_{BC}u_{BC}$ 0	0 i_b $-i_b$
12	A	B	A	1 0 0	0 1 0	1 0 0	u_{AB} $-u_{AB}$ 0	$-i_b i_b$ 0
13	A	C	A	1 0 0	0 0 1	1 0 0	$-u_{CA}u_{CA}$ 0	$-i_b$ 0 i_b
14	B	C	B	0 1 0	0 0 1	0 1 0	u_{BC} $-u_{BC}$ 0	0 $-i_b i_b$
15	B	A	B	0 1 0	1 0 0	0 1 0	$-u_{AB}u_{AB}$ 0	i_b $-i_b$ 0
16	C	C	A	0 0 1	0 0 1	1 0 0	0 u_{CA} $-u_{CA}$	i_c 0 $-i_c$
17	C	C	B	0 0 1	0 0 1	0 1 0	0 $-u_{BC}u_{BC}$	0 i_c $-i_c$
18	A	A	B	1 0 0	1 0 0	0 1 0	0 u_{AB} $-u_{AB}$	$-i_c i_c$ 0
19	A	A	C	1 0 0	1 0 0	0 0 1	0 $-u_{CA}u_{CA}$	$-i_c$ 0 i_c
20	B	B	C	0 1 0	0 1 0	0 0 1	0 u_{BC} $-u_{BC}$	0 $-i_c i_c$
21	B	B	A	0 1 0	0 1 0	1 0 0	0 $-u_{AB}u_{AB}$	i_c $-i_c$ 0
22	A	B	C	1 0 0	0 1 0	0 0 1	$u_{AB}u_{BC}u_{CA}$	$i_a i_b i_c$
23	A	C	B	1 0 0	0 0 1	0 1 0	$-u_{CA}$-u_{BC}-u_{AB}	$i_a i_c i_b$
24	B	A	C	0 1 0	1 0 0	0 0 1	$-u_{AB}$-u_{CA}-u_{BC}	$i_B i_a i_c$
25	B	C	A	0 1 0	0 0 1	1 0 0	$u_{BC}u_{CA}u_{AB}$	$i_c i_a i_b$
26	C	A	B	0 0 1	1 0 0	0 1 0	$u_{CA}u_{AB}u_{BC}$	$i_b i_c i_a$
27	C	B	A	0 0 1	0 1 0	1 0 0	$-u_{BC}$-u_{AB}-u_{CA}	$i_c i_b i_a$

$$g_1 = \left|u_{\alpha Lref}(k+1) - u_{\alpha L}(k+1)\right| + \left|u_{\beta Lref}(k+1) - u_{\beta L}(k+1)\right| \tag{17}$$

$$g_2 = |Q_S| = \left|u_{\beta S}(k+1) i_{\alpha S}(k+1) - u_{\alpha S}(k+1) i_{\beta S}(k+1)\right| \tag{18}$$

$$g = K_2 g_1 + K_2 g_2 \tag{19}$$

The selection of the switch combinations for the successive time interval is performed using a quality function "g" minimization technique. For the computation of "g", the output voltage and input reactive power on the next sampling interval are predicted.

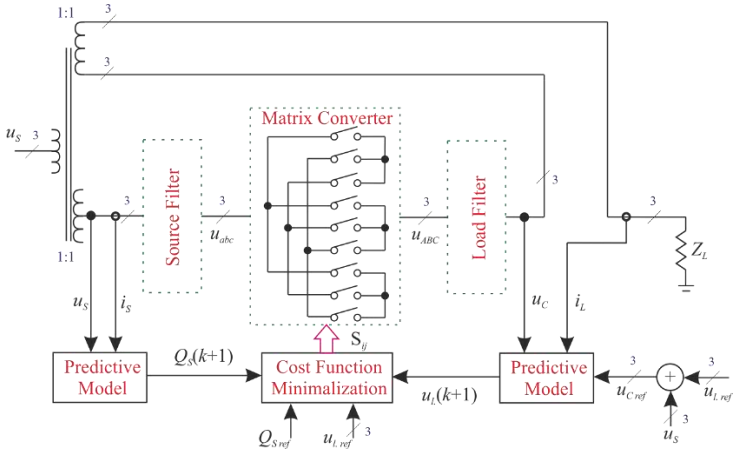
4 Simulation Results

The proposed HT control algorithm with model predictive control has been verified in simulation studies in Matlab Simulink. The simulation parameters are as follows: process discretization time T_D is equal 5 μs, generator voltage u_S=230 V/ 50 Hz, inductance of the input filter L_S=0.1 mH/0.2 Ω, capacitance of the input filter C_S=100 μF, inductance of the output filter L_L=3 mH/ 0.2 Ω, capacitance of the output filter C_L=25 μF. The tests were carried out for the *RL* load R_{ZL}=10 Ω and L_{ZL}=50 mH.

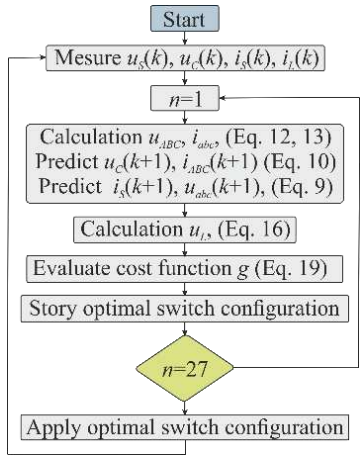
In simulation verification, a converter model built from idealized connectors was used. Simulation files are available from [28]. In addition, discrete models of input and output filters have been implemented. The value of compensation voltages was determined in accordance with the relation (15), by measuring the instantaneous value of the supply voltage. Then, in each sampling period t_D, the $\underline{i}_S$, $\underline{u}_{abc}$ and $\underline{u}_C$, $\underline{i}_{ABC}$ signals are calculated in accordance with (9), (10), respectively for each of the 27 allowed combinations of matrix converter switches. Out of these 27 results, one combination of switches is activated, for which the objective function "g" (19) has a minimum value. The simplified scheme of MPC for the proposed HT with MC is shown in Fig. 3(a), whereas control flowchart is shown in Fig. 3(b).

Examples of output voltage stabilization during the symmetrical 70% source voltage sag (observed within 45 ms to 80 ms) and 130% overvoltage (observed within 80 ms to 120 ms), using the HT with MC and model predictive control are presented in Fig. 4. As is visible in the presented time waveforms, the amplitude of output voltages is constant. Oscillations due to step voltage change are visible only at the moments of initiation and cessation of the voltage sag or overvoltage.

The use of predictive control is also a very simple way of enabling compensation of asymmetric voltage sag and overvoltage. An example of such compensation is shown in Fig. 5 for asymmetrical voltage dip with phase shift and overvoltage in one line. The asymmetrical voltage sag with phase shift in two lines is observed within 40 ms to 100 ms in Fig. 5. In addition, in the time from 100 ms asymmetrical overvoltage in the one line occurs. Without changing the control strategy, it is possible to fully compensate voltage for unbalanced changes in the supply voltage amplitude. We obtain much better results than those presented in reference [10] which uses the classic SVM method. A similar situation occurs during compensation of voltage harmonics [29]. An example of such a situation is shown in Fig. 6 for harmonic distortions at the nominal value of the supply voltage (observed within 40 ms to 100 ms) and 80% of the supply voltage sag and harmonic distortions (observed within 100 ms to 140 ms).



(a)



(b)

Figure 3

Blok diagram of the predictive control of HT with MC with minimalization of the output voltage tracking error and input reactive power (a), control flow chart (b)

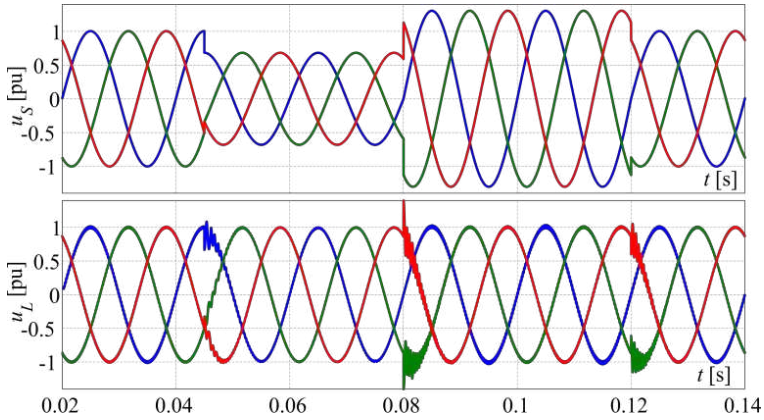


Figure 4

Simulation results of HT source and output voltage time waveforms during 70% symmetrical source voltage sag and 130% source symmetrical overvoltage

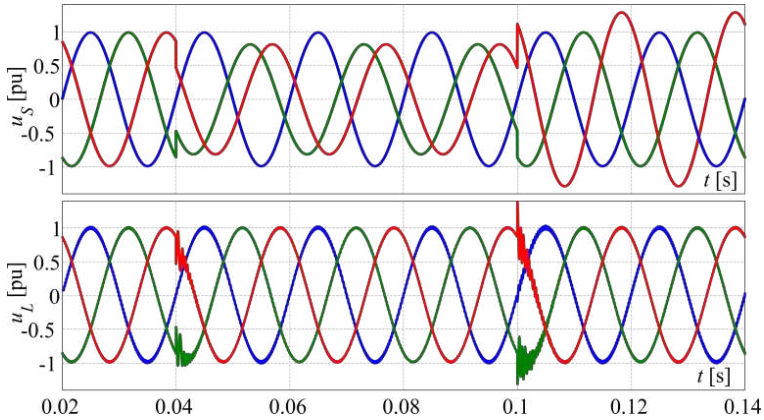


Figure 5

Simulation results of HT source and output voltage time waveforms during 70% unsymmetrical source voltage sag in two lines with voltage phase shift and 130% unsymmetrical overvoltage in one line

The obtained result confirms the possibility of compensation for this type of distortion, without additional expansion of the control algorithm, and the result is quite satisfactory. The presented preliminary results of research limited to simulation studies indicate the huge potential of using the predictive method. The next tests will have to take into account additional parameters such as those of the transformer or power network.

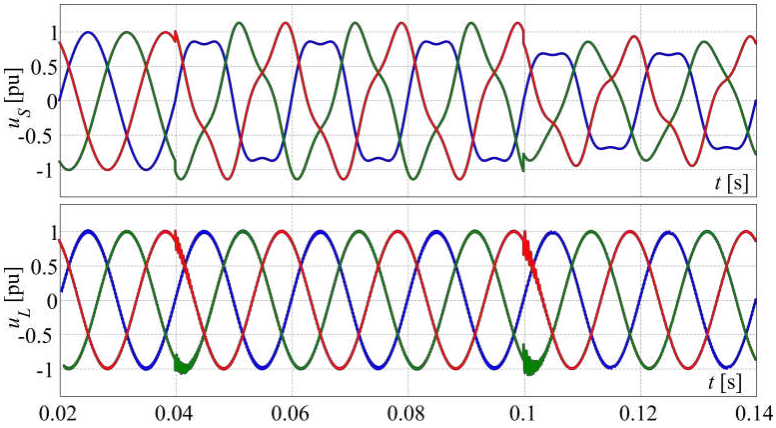


Figure 6

Simulation results of HT source and output voltage time waveforms with harmonic distortion and 80% voltage sag with harmonic distortion

5 Limitations and Scope

The proposed approach has several limitations. Firstly, the implementation of the control system will require a large number of A/D converters and voltage/current sensors to measure three-phase voltages and currents in the compensator circuit. Furthermore, the proposed voltage compensator has some drawbacks, related to the connections of the HT to the same system in which there is a grid fault. Because the system has no energy storage element, the HT needs a minimum network voltage to work properly and it may not be able to compensate very deep sags and large overvoltage.

The proposed model predictive control algorithm does not take into account the transformer parameters that will be important in the control of a real system. The preliminary research results presented in the letter indicate the significantly improved properties of this type of control compared with classical methods based on SVM algorithms.

It should be emphasized that the proposed system can already be easily implemented to protect consumers whose loads are sensitive to voltage fluctuations. The load power range for which a compensator is provided is equal to several dozen kVA. An important issue of future work will be to assess the benefits and costs of the AC voltage compensation system [30].

Conclusions

In this paper, a three-phase hybrid transformer with a matrix converter and model predictive control, for compensation of source voltage sag/swell and harmonics has been presented. The operation and circuit have been described previously in paper [10]. In this letter the main topic is related to the presentation of the proposed automatic control of the analyzed source voltage variations compensator using the model predictive control method. Initial investigation has given very satisfactory results. The preliminary research results presented in the article indicate significantly improved properties for this type of control compared with classical methods based on SVM algorithms [10].

The next step of the research will focus on the theoretical modeling of the electromagnetic transformer which is used in the HT and will take into account the parameters in the predictive control algorithm. In addition, the experimental verification of the properties of the proposed system will be the target of future research. Economic analyzes [31], construction and implementation of the proposed solution in the context of individual, group and central applications will also be considered.

References

[1] S. C. Vegunta and J. V. Milanovic: Estimation of cost of downtime of industrial process due to voltage sags, IEEE Transactions on Power Delivery, 2011, Vol. 26, No. 2, pp. 576-587

[2] Z. Djokic, J. Desment, G. Vanalme, J. Milanovic, K. Stockman: Sensitivity of personal computer to voltage sags and short interruptions, IEEE Transactions on Power Delivery, 2005, Vol. 20, No. 1, pp. 375-383

[3] O. P. Mahela, A. G. Shaik: Topological aspects of power quality improvement techniques: A comprehensive overview, Renewable and Sustainable Energy Reviews, 2016, Vol. 58, pp. 1129-1142

[4] D. Divan and P. Kandula: Distributed power electronics: An enabler for the future grid, CPSS Transactions on Power Electronics and Applications, 2016, Vol. 1, No. 1, pp. 57-65

[5] A. Tahri, H. M. Boulouiha, A. Allali, T. Fatima: A multi-variable LQG controller-based robust control strategy applied to an advanced static VAR compensator, Acta Polytechnica Hungarica, 2013, Vol. 10, No. 4

[6] J. Burkard, J. Biela: Evaluation of topologies and optimal design of a hybrid distribution transformer, Proceedings of European Conference on Power Electronics and Applications, EPE‘15 ECCE Europe, 2015, Geneva, Switzerland, pp. 1-10

[7] D. Das, R. P. Kandula, J. A. Muñoz, D. Divan, R. G. Harley, J. E. Schatz: An integrated controllable network transformer - hybrid active filter

system, IEEE Transactionon Industrial Application, 2015, Vol. 51, No. 2, pp. 1692-1701

[8] J. Kaniewski, Z. Fedyczak, G. Benysek: AC voltage sag/swell compensator based on three-phase hybrid transformer with buck-boost matrix-reactance chopper, IEEE Transaction on Industrial Electronics, 2014, Vol. 61, No. 8, pp. 3835-3846

[9] T. Kang, S. Choi, A.S. Morsy and P.N. Enjeti: Series voltage regulator for a distribution transformer to compensate voltage sag/swell, IEEE Trans. Ind. Electron, Vol. 64, No. 6, pp. 4501-4510, 2017

[10] P. Szcześniak, J. Kaniewski: Hybrid transformer with matrix converter, IEEE Transactions on Power Delivery, 2016, Vol. 31, No. 3, pp. 1388-1396

[11] P. Szcześniak: The compensator of voltage sage/swell installed in connection terminals of small industrial plant or selected loads,European Conference on Power Electronics and Applications (EPE'17 ECCE Europe) Warsaw, Poland (2017)

[12] P. Szcześniak, J. Kaniewski, P. Sanjeevikumar: Control algorithm concept for AC voltage stabilizer based on hybrid transformer with a matrix converter, Eds. Garg, A., Bhoi, A.K., Sanjeevikumar, P., Kamani, K. K. in: Advances in Power Systems and Energy Management, Singapore: Springer, 2018 (Lecture Notes in Electrical Engineering; Vol. 436) pp. 337-346

[13] J. Andreu, I. Kortabarria, E. Ormaetxea, E. Ibarra, J. L. Martin, S. Apiñaniz: A step forward towards the development of reliable matrix converters, IEEE Transaction on Industrial Electronics, 2012,Vol. 59, No. 1, pp. 167-183

[14] D. Zieliński, P. Lipnicki, W. Jarzyna: Synchronization of voltage frequency converters with the grid in the presence of notching, COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 2015, Vol. 34, pp. 657-673

[15] M. Bobrowska-Rafal, K. Rafal, M. Jasinski, M. P. Kazmierkowski: Grid synchronization and symmetrical components extraction with PLL algorithm for grid connected power electronic converters – a review, Bulletin of the Polish Academy of Sciences Technical Sciences, 2011, Vol. 59, No. 4, pp. 485-497

[16] J. A. Tenreiro Machado, B. Pátkai, I. J. Rudas: Intelligent Engineering Systems and Computational Cybernetics, Springer Netherlands 2009

[17] R. E Precup, S. Preitl, E. M. Petriu, J. K. Tar, M. L. Tomescu, C. Pozna: Generic two-degree-of-freedom linear and fuzzy controllers for integral processes, Journal of the Franklin Institute, 2009, Vol. 346, No. 10, pp. 980-1003

[18] S. Vrkalovic1, E. C. Lunca, I. D. Borlea: Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants, International Journal of Artificial Intelligence, 2018, Vol. 16, No. 2, pp. 208-222

[19] R. E. Precup, S. Preitl: Stability and sensitivity analysis of fuzzy control systems. Mechatronics applications, Acta Polytechnica Hungarica, 2006, Vol. 3, No. 1, pp. 61-76

[20] A. Ürmös, Z. Farkas, M. Farkas, T. Sándor, L. T. Kóczy, Á. Nemcsics: Application of self-organizing maps for technological support of droplet epitaxy, Acta Polytechnica Hungarica, 2017, Vol. 14, No. 4, pp. 207-224

[21] S. Padmanaban, F. J. L. Daya, F. Blaabjerg, P. W. Wheeler, P. Szcześniak, V. Oleschuk, A. H. Ertas, Wavelet-fuzzy speed indirect field oriented controller for three-phase AC motor drive – Investigation and implementation, Engineering Science and Technology, an International Journal, 2016, Vol. 19, No. 3, pp. 1099-1107

[22] B. Lantos, Z. Bodó, High level kinematic and low level nonlinear dynamic control of unmanned ground vehicles, Acta Polytechnica Hungarica, 2019, Vol. 16, No. 1, pp. 97-117

[23] M. Trojanová, A. Hošovský, Comparison of different neural networks models for identification of manipulator arm driven by fluidic muscles, Acta Polytechnica Hungarica, 2018, Vol. 15, No. 7, pp. 7-28

[24] E. Tóth-Laufer, M Takács, I. J. Rudas, Fuzzy logic-based risk assessment framework to evaluate physiological parameters, Acta Polytechnica Hungarica, 2015, Vol. 12, No. 2, pp. 159-178

[25] J. Rodríguez, P. Cortes: Predictive Control of Power Converters and Electrical Drives. John Wiley & Sons, 2012

[26] M. R. Nasiri, S. Farhangi, J. Rodríguez, Model predictive control of a multilevel CHB STATCOM in wind farm application using diophantine equations, IEEE Transaction on Industrial Electronics, 2019, Vo. 66, No. 2, pp. 1213-1223

[27] P. Wheeler, M. Rivera, S. Toledo: An indirect model predictive current control for a direct matrix converter with instantaneous reactive power minimization, Proceedings of IEEE Southern Power Electronics Conference (SPEC'2017) Puerto Varas, 2017, pp. 1-6

[28] https://www.mathworks.com/matlabcentral/fileexchange/71962-model-predictive-control-hybrid-converter-with-matrix-conver

[29] S. Yusoff, L. De Lillo, P. Zanchetta, P. Wheeler: Predictive control of a direct AC/AC matrix converter power supply under non-linear load conditions, Proceedings of 15th International Power Electronics and Motion Control Conference (EPE/PEMC'2012) Novi Sad, 2012, pp. DS3c.4-1-DS3c.4-6

[30] N. Katic: Performance Analysis of smart grid solutions in distribution power systems, Acta Polytechnica Hungarica, 2018, Vol. 15, No. 6

[31] W. E. Brumsickle, R. S. Schneider, G. A. Luckjiff, D. M. Divan, M. F. McGranaghan, Dynamic sag correctors: cost-effective industrial power line conditioning, IEEE Transactions on Industry Applications, 2001, Vol. 37, No. 1, pp. 212-217

Constrained Predictive Control of Three-PhaseBuck Rectifiers

László Richárd Neukirchner[1], Attila Magyar[1], Attila Fodor[1], Nimród Dénes Kutasi[2], András Kelemen[2]

[1] University of Pannonia, Department of Information Technology, Egyetem u. 10, 8200 Veszprém, Hungary, e-mails: neukirchner.laszlo@virt.uni-pannon.hu, magyar.attila@virt.uni-pannon.hu, foa@almos.uni-pannon.hu

[2] Sapientia Hungarian University of Transylvania, Department of Electrical Engineering, Corunca 540485, Romania, e-mail: kutasi@ms.sapientia.ro, kandras@ms.sapientia.ro

Abstract: In this paper, constrained optimal control of a current source rectifier (CSR) is presented, based on a mathematical model developed in Park's frame. To comply with the system constraints an explicit model-based predictive controller was established. To simplify the control design, and avoid linearization, a disjointed model was utilised due to the significant time constant differences between the AC and DC side dynamics. As a result, active damping was used on the AC side, and explicit Model Predictive Control (MPC) on the DC side, avoiding non-linear dynamics. The results are compared by simulation with the performance of a state feedback control.

Keywords: model predictive control; current source rectifier; space vector modulation; modeling; constrained control

1 Introduction

Current source rectifiers (CSR) are widely used in front-end power electronic converter for the uncontrollable or controllable DC-bus in industrial and commercial applications. They have maintained their position through many applications, with uses such as medium-voltage high-power drives [1], [2] STATCOMs [3] and renewable systems [4], [5]. They have a plain and reliable circuit structure, which makes them attractive for simple control design. The CSRs are traditionally controlled by classic cascaded linear control loops such as PI controllers. These simple control applications are suitable for induction motor control [6], and other electromechanical actuators [7], and unusual topologies [8]. Also, worth mentioning is self-tuning variants of PI controllers [9]. In the past, the modulation methods used were trapezoidal pulse width modulation techniques

(TPWM), or application of pulse patterns calculated offline for selective harmonic elimination (SHE). More recently, current space vector modulation (SVM) has been used for the synthesis of the transistor control signals [10]. Even so, AC-side harmonic elimination could still be an issue at lower switching frequencies where LCL filtering would be advised [11]. In order to keep switching frequencies low and to minimize switching losses, new topologies and hybrid modulations are used, mixing TPWM and SHE depending on the grid frequency [12].

In terms of the amplitude of the grid and DC-link voltages, CSRs exhibit a step-down conversion. When used as DC voltage source, the rectifier can output a lower DC voltage without the need of a grid-side transformer, as is usually employed in voltage source rectifiers (VSR). Because of their current source behaviour, CSRs can be easily paralleled and provide inherent short-circuit protection, representing an excellent potential in DC power supply applications [13], [14].

There are several control strategies in addition to classical PI control for applications in this domain. Self-adapting control methods are on the rise with more sophisticated algorithms in the field of fuzzy logic [15]. They are capable of handling increasingly more complicated models and systems with high dynamics and accuracy [16], [17], and even without establishing and validating classical state-space models [18]. The other filed is the sliding mode control, which can achieve good dynamic performance and handle non-linearity. Still, they might also introduce chattering, which can be very undesirable when applied to real-life systems like in [19] and [20].

In the linear domain implicit model predictive control (IMPC or just MPC) is a fair solution due its; effectiveness in power electronics, configurable cost function and such scalable nature [21], [22]. In this field also finite-state solutions are present which can be considered also predictive control, where the modulation scheme's defined states serve as optimization potential [23], [24]. As a further step adaptive application was established to tackle parameter estimation problems for better performance [25]

Recently, beside implicit, finite-state, and adaptive predictive control, explicit model predictive control has emerged in the field of power electronics [26]. Establishing the MPC cost function can range widely depending on the expected dynamics, degree of noise cancellation, and model complexity. Additionally, the current limitation can also be implemented introducing constraints in the modulation algorithm.

In [27] the validity of an MPC-based, digital pulse width modulation control strategy for single-phase voltage source rectifiers is discussed, further confirming the validity of this method in control systems.

1.1 New Contributions

In this paper a model predictive control method is developed for a classical current source buck-type rectifier (CSR). The contribution of the paper is to show how to design EMPC on a model of a CSR which has a complex model due to bilinearity. To overcome the burden of bilinearity a simple solution is shown which enables handling the model parts as linear disjointed systems of their own.

The structure of the paper is as follows. In Section 2 the topology is presented, followed by the mathematical model derived in the synchronous rotating coordinate frame. Next, the control structure is presented, followed by the detailed description of the DC-side Explicit Model Predictive Control (EMPC) and by the presentation of the AC-side active damping. In the fourth section, the current space vector modulation scheme is shown, with optimized switching pattern to reduce the switching frequency. Lastly, the simulation results are presented and the performance of the proposed control structure is compared with the performance of a state feedback controller, before the conclusions are finally drawn.

2 Mathematical Modeling of the CSR

The structure of the classical three phase buck-type current source rectifier (CSR) is presented in Fig. 1. In continuous current mode, the differential equations corresponding to the CRS's inductor currents and capacitor voltages are the following:

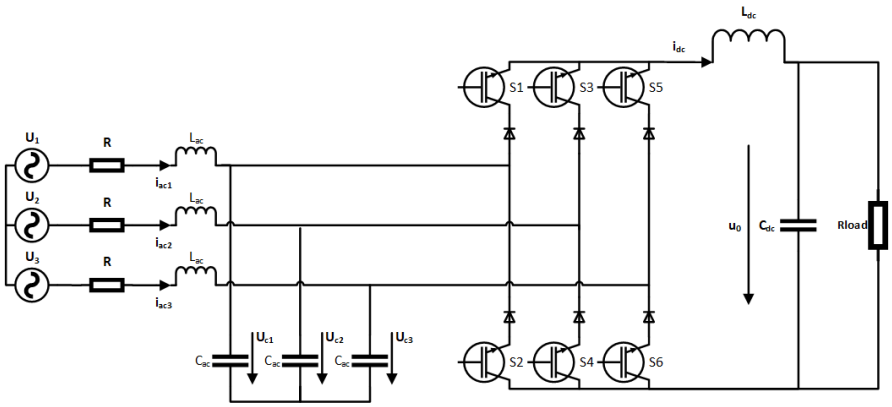


Figure 1
Circuit diagram of the three-phase buck-type rectifier with insulated gate bipolar transistors (IGBTs)

$$L_{ac}\dot{\imath_{ac_p}} = u_p - u_{c_p} - Ri_{ac_p} \tag{1}$$

$$C_{ac}\dot{u_{c_p}} = i_{ac_p} - \delta_p i_{dc}$$
$$L_{dc}\dot{\imath_{dc}} = \left(\sum_{p=1}^{3} \delta_p\, u_{c_p}\right) - u_0$$
$$C_{dc}\dot{u_0} = i_{dc} - \frac{u_0}{R_{load}}$$

where $p \in \{1,2,3\}$ is the index of three phases and δ_p describes the conduction state of the rectifier leg p (2).

$$\delta_p = \begin{cases} 1\ if\ the\ upper\ transitor\ is\ ON \\ -1\ if\ the\ lower\ transistor\ is\ ON \\ 0\ if\ both\ are\ ON\ or\ OFF \end{cases} \tag{2}$$

Using the components in the stationary frame of the space phasors of the three-phase quantities, from (1) it results:

$$\begin{aligned} L_{ac}\dot{\imath_{ac_\alpha}} &= u_\alpha - u_{c_\alpha} - Ri_{ac_\alpha} \\ L_{ac}\dot{\imath_{ac_\beta}} &= u_\beta - u_{c_\beta} - Ri_{ac_\beta} \\ C_{ac}\dot{u_{c_\alpha}} &= i_{ac_\alpha} - \delta_\alpha i_{dc} \\ C_{ac}\dot{u_{c_\beta}} &= i_{ac_\beta} - \delta_\beta i_{qc} \\ L_{dc}\dot{\imath_{dc}} &= 1.5\left(\delta_\alpha u_{c_\alpha} + \delta_\beta u_{c_\beta}\right) - u_0 \\ C_{dc}\dot{u_0} &= i_{dc} - \frac{u_0}{R_{load}} \end{aligned} \tag{3}$$

Equation (3) is transformed to the synchronous reference frame rotating with the u_{c_d} capacitor voltage space vector. The resulting mathematical model is thus:

$$\begin{aligned} L_{ac}\dot{\imath_{ac_d}} &= u_d - u_{c_d} - R\mathrm{i}_{ac_d} + \omega_s L_{ac}\mathrm{i}_{ac_q} \\ L_{ac}\dot{\imath_{ac_q}} &= u_q - u_{c_q} - R\mathrm{i}_{ac_q} - \omega_s L_{ac}\mathrm{i}_{ac_d} \\ C_{ac}\dot{u_{c_d}} &= i_{ac_d} - \delta_d i_{dc} + \omega_s C_{ac} u_{c_q} \\ C_{ac}\dot{u_{c_q}} &= i_{ac_q} - \delta_q i_{dc} - \omega_s C_{ac} u_{c_d} \\ L_{dc}\dot{\imath_{dc}} &= 1.5\left(\delta_d u_{c_d} + \delta_q u_{c_q}\right) - u_0 \\ C_{dc}\dot{u_0} &= i_{dc} - \frac{u_0}{R_{load}} \end{aligned} \tag{4}$$

where ω_s represents the network voltage vector's angular velocity.

2.1 Model Simplification

Notice, that the sixth-order ODE model (4) is bilinear in its states and inputs because of the product terms ($\delta_d i_{dc}$ for example). As such, using design methods for linear systems is not straightforward. The high complexity given by the system's order is another problem to tackle. For designing classic MPC, linear, low-order equation systems are favorable. Hence

simplification of the model would bring noteworthy benefits, making the MPC design more straightforward, when a linear system resulted.

Since the AC and DC side's time constants differ significantly (as in the AC: $\omega_{ac} = \frac{1}{\sqrt{L_{ac}C_{ac}}} \cong 5.7 \cdot 10^3 [rad/s]$, and on the DC:$\omega_{dc} = \frac{1}{\sqrt{L_{dc}C_{dc}}} \cong 2.8 \cdot 10^2 [rad/s]$, see Table 3. for reference). Thus, the differential equations can be separated into two sets, and the control of the AC and DC sides can be decoupled as described in [28]. The AC side model results as follows:

$$\begin{pmatrix} \dot{i}_{ac_d} \\ \dot{i}_{ac_q} \\ \dot{u}_{c_d} \\ \dot{u}_{c_q} \end{pmatrix} = \begin{pmatrix} -\frac{R}{L_{ac}} & \omega_s & -\frac{1}{L_{ac}} & 0 \\ -\omega_s & -\frac{R}{L_{ac}} & 0 & -\frac{1}{L_{ac}} \\ \frac{1}{C_{ac}} & 0 & 0 & \omega_s \\ 0 & \frac{1}{C_{ac}} & -\omega_s & 0 \end{pmatrix} \begin{pmatrix} i_{ac_d} \\ i_{ac_q} \\ u_{c_d} \\ u_{c_q} \end{pmatrix} + \begin{pmatrix} \frac{u_d}{L_{ac}} \\ \frac{u_q}{L_{ac}} \\ -\frac{\delta_d i_{dc}}{C_{ac}} \\ -\frac{\delta_q i_{dc}}{C_{ac}} \end{pmatrix}. \tag{5}$$

Looking at the state matrix it can be further stated that there are only weak couplings between the d and q components. This allows to handle them separately, and later to design separate control for each.

The equation system describing the DC side dynamics is the following:

$$\begin{pmatrix} \dot{i}_{dc} \\ \dot{u}_0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{-1}{L_{dc}} \\ \frac{1}{C_{dc}} & \frac{-1}{R_{load}C_{dc}} \end{pmatrix} \begin{pmatrix} i_{dc} \\ u_0 \end{pmatrix} + \begin{pmatrix} \frac{1.5}{L_{dc}} \left(\delta_d u_{c_d} + \delta_q u_{c_q} \right) \\ 0 \end{pmatrix}. \tag{6}$$

It can be noticed that, with the AC and DC model separation, bilinearity disappears, since the binding coefficients are present only in the input ($\boldsymbol{u}$) of the DC state space model. Consequently, all equations are linear and with a considerably lower order, making control design much easier and allowing for the application of linear design methods. For the DC side dynamics, the linear time invariant differential equation system's matrices can be identified for predictive control design purposes:

$$\boldsymbol{x} = \begin{pmatrix} i_{dc} \\ u_0 \end{pmatrix}, \boldsymbol{u} = \left(\delta_d u_{c_d} + \delta_q u_{c_q} \right), \boldsymbol{y} = u_0,$$

$$\boldsymbol{A} = \begin{pmatrix} 0 & \frac{-1}{L_{dc}} \\ \frac{1}{C_{dc}} & \frac{-1}{R_{load}C_{dc}} \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} \frac{1.5}{L_{dc}} \\ 0 \end{pmatrix}, \boldsymbol{C} = (0 \quad 1). \tag{7}$$

where $\boldsymbol{x}$, $\boldsymbol{u}$and $\boldsymbol{y}$are the state, input and output vectors of the DC-side system, and $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are the state, input and output matrices.

The circuit parameters used for the implementation of the control structure based on this model are presented in Table 3.

3 The Control Structure

Relying on the possibility of separation of the AC-side and DC-side controllers, the control structure from Fig. 2 is proposed.

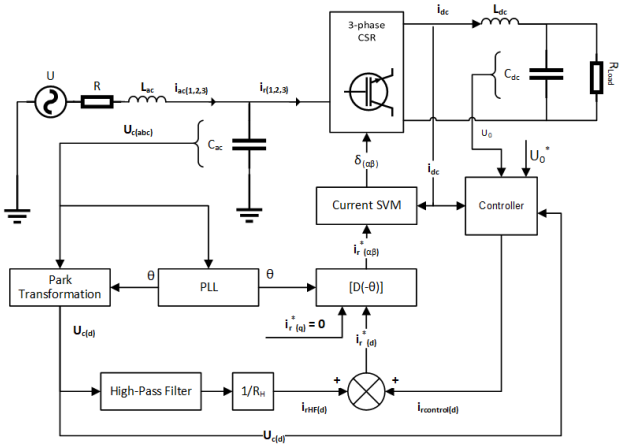


Figure 2

Block diagram of the control structure

The controllers operate in the synchronous frame of the AC filter capacitor voltages $u_{c_{(1,2,3)}}$, and the rectifier input currents $i_{r_{(1,2,3)}}$ are in phase with the capacitor voltages.

The current reference $i^*_{r(\alpha\beta)}$ supplied to the space vector modulation unit in the stationary frame, is obtained by coordinate transformation $[D(-\theta)]$ of the current reference (8) delivered by the current controllers in the synchronous frame.

$$\begin{cases} i^*_{r_d} = i_{rcontrol_d} + i_{rHF_d} \\ i^*_{r_q} = 0 \end{cases} \tag{8}$$

In (8), $i_{rcontrol_d}$ represents the output of the DC voltage controller, while i_{rHF_d} represents the damping current, proportional with the high frequency component of the filter capacitor voltage (the fundamental component of the capacitor voltage in the stationary frame becomes a DC component in the synchronous frame). The DC and AC side control units are explained in more detail in the following sections, and the performance of the control structure is evaluated.

3.1 DC-SideExplicit Model Predictive Control

Model predictive control (MPC) is an efficient and systematic method for solving complex multi-variable constrained optimal control problems [3]. The MPC control law is based on the "receding horizon formulation", where the model's assumed behavior is calculated for a number of N steps, where N stands for the

horizon's length. Only the first step of the computed optimal input is applied in each iteration. The remaining steps of the optimal control input are discarded and a new optimal control problem is solved at the next sample time. Using this approach, the receding horizon policy provides the controller with the desired feedback characteristics, although with high order systems the computational effort is considerably demanding since all the steps should be taken in to account on the specified horizon in every iteration. With Explicit MPC (EMPC), the discrete time constrained optimal control problem is reformulated as multi-parametric linear or quadratic programming. Using this approach, the problem of optimization can be solved offline, making it much more feasible from the perspective of the optimal control task. The optimum control law is a piecewise affine function of the states, and the resulting solution is stored in a pre-calculated lookup table. The parameter space, or the state-space is partitioned into critical regions. The real-time implementation consists in searching for the active critical region, where the measured state variables lie, and in applying the corresponding piecewise affine control law to achieve the desired dynamics.

In order to introduce the MPC implementation from this paper, let us consider a linear discrete time system (9) derived with the discretisation of system (6) with zero-order hold method, where control inputs are assumed piecewise constant over the simulation sample time $T_s = 1/f_s$:

$$\begin{aligned} \boldsymbol{x}(t+1) &= \boldsymbol{A}_d\boldsymbol{x}(t) + \boldsymbol{B}_d\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}_d\boldsymbol{x}(t) \end{aligned} \tag{9}$$

where $\boldsymbol{A}_d, \boldsymbol{B}_d, \boldsymbol{C}_d$ere the matrices of the discretised system derived from (7). With system (9) appears to be linear time invariant, MPC design can be followed. The following constraints have to be satisfied:

$$\boldsymbol{y}_{min} \le \boldsymbol{y}(t) \le \mathbf{y}_{max}, \boldsymbol{u}_{min} \le \boldsymbol{u}(t) \le \boldsymbol{u}_{max} \tag{10}$$

where $t > 0$, $\boldsymbol{x} \in R^n, \boldsymbol{u} \in R^m, \boldsymbol{y} \in R^p$. The MPC solves the following constrained optimization problem [23]:

$$\min_{U=\{\boldsymbol{u}_t,\dots,\boldsymbol{u}_t+N_u-1\}} J\big(\boldsymbol{u}, \boldsymbol{x}(t)\big) = \sum_{k=0}^{N_y-1} \left(\boldsymbol{x}_{t+N_y \vee t}^T Q_w \boldsymbol{x}_{t+N_y \vee t} + \boldsymbol{u}_{t+k}^T R_w \boldsymbol{u}_{t+k}\right) \tag{11}$$

subject to:

$$\begin{aligned} &\boldsymbol{x}_{min} \le \boldsymbol{x}_{t+k|t} \le \boldsymbol{x}_{max}, k = 1, \dots, N_c - 1 \\ &\boldsymbol{u}_{min} \le \boldsymbol{u}_{t+k|t} \le \boldsymbol{u}_{max}, k = 0,1, \dots, N_c - 1 \\ &\boldsymbol{x}_{t|t} = \boldsymbol{x}(t) \\ &\boldsymbol{x}_{t+k+1|t} = \boldsymbol{A}_d\boldsymbol{x}_{t+k|t} + \boldsymbol{B}_d\boldsymbol{u}_{t+k|u} \\ &\boldsymbol{y}_{t+k+1|t} = \boldsymbol{C}_d x_{t+k|t} \\ &\boldsymbol{u}_{t+k|t} = -K\boldsymbol{x}_{t+k|t} \end{aligned} \tag{12}$$

$k \geq 0$

This problem is solved at each time instant *t*, where $\boldsymbol{x}_{t+k|t}$ denotes the state vector predicted at time *t+k,* obtained by applying the input sequence $\boldsymbol{u}_{t|t\ldots}\boldsymbol{u}_{t+k-1|t}$ to model (15), starting from the state $\boldsymbol{x}_{t|t}$. Further, it is assumed that the weighting matrices Q_w and R_w, are symmetric positive semidefinite ($Q_w = Q_w^T \geq 0, R_w = R_W^T > 0$) and K is a feedback gain. Further, N_y, N_u, N_c, are the output, input and constraint horizons, respectively.

Using the model for predicting the future behavior of the system and with some appropriate substitution and variable manipulation, the problem (11), (12) can be transformed to the standard multi-parametric quadratic programming (mp-QP) form, as described in [29]:

$$V_z(\boldsymbol{x}) = \quad min \frac{1}{2} z^t H z \tag{13}$$

subject to:

$$Gz \quad \leq \quad W + S\boldsymbol{x}(t) \tag{14}$$

where the matrices H, G, W, S result directly from the coordinate transformations and variable manipulations. The solution of the mp-QP problem for each critical region has the form:

$$\boldsymbol{u}^* = \quad f_i \boldsymbol{x} + g_i \tag{15}$$

and the critical region is described by:

$$C_{reg_i} = \quad \{\boldsymbol{x} \in R^n \vee H_i \boldsymbol{x} \leq K_i\}. \tag{16}$$

Thus, the explicit MPC controller is completely characterized by the set of parameters:

$$\{f_i, g_i, H_i, K_i\}^{i=1\ldots N}. \tag{17}$$

In case of the discrete time system resulting from (7), for sampling time equal with the switching period $T_s = 50 \cdot 10^{-5}\, s$, the problem defined to be solved by MPC is the minimization of the quadratic cost function (11) for:

$$R_w = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_w = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \text{and } N_y = N_u = N_c = 2. \tag{18}$$

Since N_y, N_u, N_c take the same value, they will be substituted by N.

The constraints defined based on the rated power of the CSR $P_N = 2500\, W$, are:

$$\begin{aligned} 0 \leq \quad i_{dc} \quad &\leq 50A \\ 0 \leq \quad u_0 \quad &\leq 500V \end{aligned} \tag{19}$$

The state space partition resulting from this problem has 13 critical regions, which can be observed in Fig. 3.

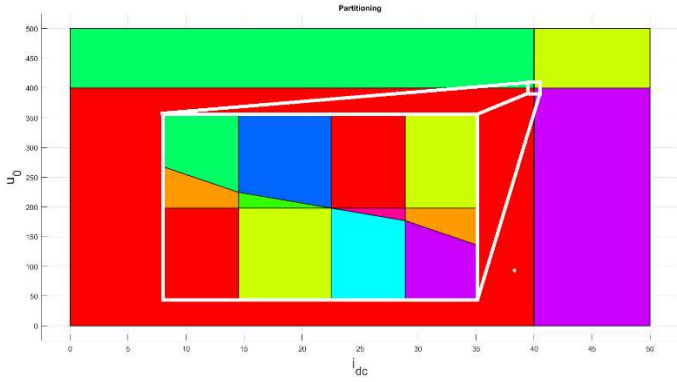


Figure 3
State space partitioning

From the basis of the discretized model (9), the given constraints, and horizon (19) the cost function (11) is established via the MPT toolbox [30] and used in the generated controller for the EMPC design [29], [31]. The controller is created as a compliable S-function in the Matlab/Simulink environment and its place in the control structure can be observed in Fig. 4 as the EMPC controller.

The output of the MPC controller is the control variable obtained via solving (12) $u_{MPC} = \left(\delta_d u_{c_d} + \delta_q u_{c_q}\right)$, from which the current reference can be calculated using (19). The quadrature component u_{c_q} is zero in the synchronous frame of the filter capacitor voltage.

$$i_{rMPC_d} = \frac{u_{MPC}}{u_{c_d}} \cdot i_{dc} \tag{20}$$

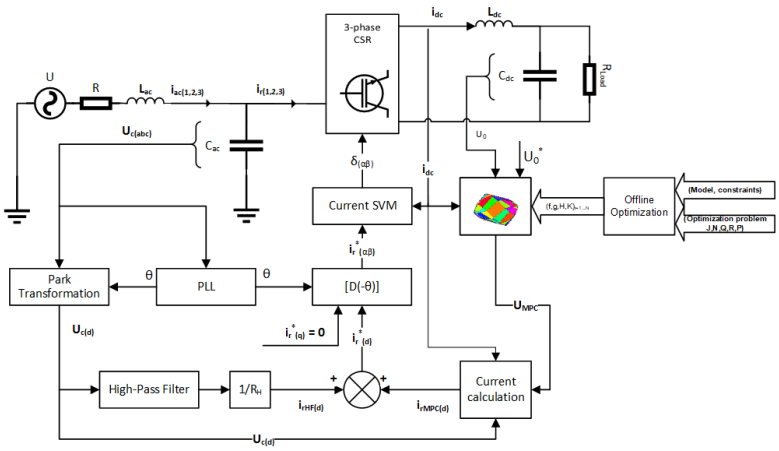


Figure 4

The control structure of the CSR, with MPC controller on the DC side

3.2 Active AC-Side Damping

The CSR requires a voltage supply on the AC side. Taking into consideration the inductive character of the mains, the presence of a three-phase capacitor tank at the input of the CSR is a must. The most convenient is to use three-phase LC filtering with inductors on the lines and star connected capacitors resembling those in Fig. 1, although the resonance phenomena between these components can still cause difficult problems. The simplest way to dampen the resonance on the AC side LC filter is to add a damping resistor across the capacitor [23]. Because these resistors result in high losses, active damping methods have been proposed, which emulate damping resistors by control. This makes the CSR bridge produce an additional high frequency current, equivalent to the presence of virtual damping resistors connected in parallel with the AC capacitors. The resonance of the AC side LC filter produces harmonics in the capacitor voltage with frequency close to $\omega_{ac} = \frac{1}{\sqrt{L_{ac}C_{ac}}}$, which appears as $\omega_{ac} - \omega_s$ component in u_{c_d}, where $\omega_s = 2\pi f$. The fundamental component of the capacitor voltage represents a DC component in the synchronous reference frame. Therefore, a high-pass filter (HPF) is applied to filter out this DC component, with the transfer function:

$$HPF(s) = \frac{s}{s + 0.1 * (\omega_{ac} - \omega_s)}. \tag{21}$$

A virtual damping resistance R_H has been defined for calculation of the damping current component i_{HPF} from the HF component of the capacitor voltage.

4 Space Vector Modulation Strategy

The chosen modulation strategy is developed in the "αβ" stationary reference frame. The structure requires simultaneous conduction of the upper and lower transistors of the bridge, since the current of the L_{dc}choke must not be interrupted. Additionally, the switching devices are considered as ideal.

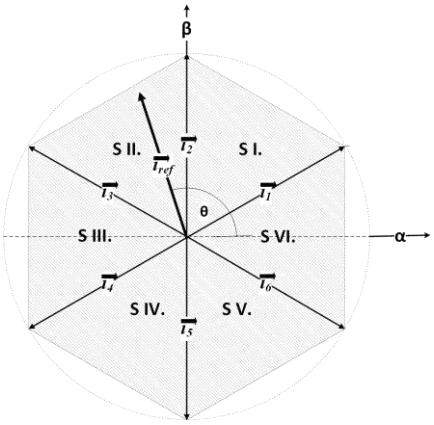


Figure 5

The fundamental input current vectors corresponding to the active switching states of the CSR

According to this, one of the upper and one of the lower switches must be closed at all times. This allows nine states, six of which are active. There are three "zero" vectors, corresponding to the switching states, when both devices of one of the bridge legs are in conduction, shown in Table 1.

Table 1

Switching states of the rectifier and the corresponding space phasors

Name	Switching State						Phase currents			Vector representation
	1	2	3	4	5	6	i_a	i_b	i_c	
$\vec{i_1}$	1	0	0	0	0	1	i_{dc}	0	-i_{dc}	$\left(2i_{dc}e^{\left(\frac{j\pi}{6}\right)}\right)/\sqrt{3}$
$\vec{i_2}$	0	0	1	0	0	1	0	i_{dc}	-i_{dc}	$\left(2i_{dc}e^{\left(\frac{j\pi}{2}\right)}\right)/\sqrt{3}$
$\vec{i_3}$	0	1	1	0	0	0	-i_{dc}	i_{dc}	0	$\left(2i_{dc}e^{\left(\frac{j5\pi}{6}\right)}\right)/\sqrt{3}$
$\vec{i_4}$	0	1	0	0	1	0	-i_{dc}	0	i_{dc}	$\left(2i_{dc}e^{\left(\frac{j7\pi}{6}\right)}\right)/\sqrt{3}$
$\vec{i_5}$	0	0	0	1	1	0	0	-i_{dc}	i_{dc}	$\left(2i_{dc}e^{\left(\frac{j3\pi}{2}\right)}\right)/\sqrt{3}$
$\vec{i_6}$	1	0	0	1	0	0	i_{dc}	-i_{dc}	0	$\left(2i_{dc}e^{\left(\frac{j11\pi}{6}\right)}\right)/\sqrt{3}$
$\vec{i_7}$	1	1	0	0	0	0	0	0	0	0
$\vec{i_8}$	0	0	1	1	0	0	0	0	0	0
$\vec{i_9}$	0	0	0	0	1	1	0	0	0	0

The neighboring space phasors can be formulated as:

$$\overrightarrow{i_n} = \frac{2}{\sqrt{3}} i_{dc} expj\left(\frac{n\pi}{3} - \frac{\pi}{6}\right)$$
$$\overrightarrow{i_{n+1}} = \frac{2}{\sqrt{3}} i_{dc} expj\left(\frac{n\pi}{3} + \frac{\pi}{6}\right) \quad (22)$$
$$n = 1,2,\dots 6$$

The reference current vector is sampled with fixed sampling period T_s. The sampled value of $\overrightarrow{i_{ref}}$ is synthesized as the time average of two neighbouring space phasors adjacent to the reference current:

$$T_n \overrightarrow{i_n} + T_{n+1} \overrightarrow{i_{n+1}} = T_s \overrightarrow{i_{ref}}. \quad (23)$$

T_n and T_{n+1} represent the individual durations of the switching states corresponding to the neighboring vectors. For example, in case of a current reference vector situated in the first sector, T1, T2 and T0 can be calculated using (24).

$$T_1 = T_s \frac{i_{ref_\alpha}}{i_{dc}}$$
$$T_2 = T_s \frac{\sqrt{3}}{2} \frac{1}{i_{dc}} \left(i_{ref_\beta} - \frac{1}{\sqrt{3}} i_{ref_\alpha}\right) \quad (24)$$
$$T_0 = T_s - T_n - T_{n-1} = T_{7,8,9}$$

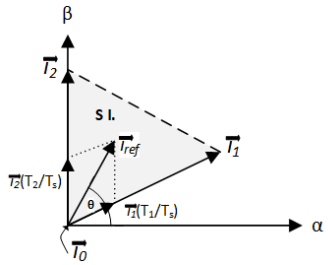


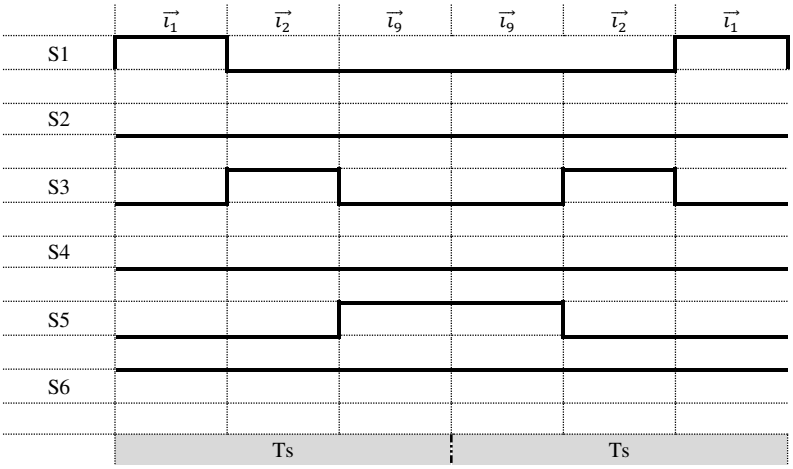
Figure 6
Synthesis of $\overrightarrow{i_{ref}}$ by $\overrightarrow{i_1}$, $\overrightarrow{i_2}$, and $\overrightarrow{i_0}$

The complex plane is naturally divided by the fundamental space vectors into six areas, named "sectors".

$$\frac{\pi}{6} + \frac{(n-1)\pi}{3} \leq \theta_n \leq \frac{\pi}{6} + \frac{n\pi}{3} \quad (25)$$
$$n = 1,2,\dots 6$$

The non-zero space vectors are selected based on the phase angle θ between $\overrightarrow{i_{ref}}$ and the real axis. Table 2 presents an example of switching pattern in case of a current reference vector situated in Sector I.

Table 2
Representation of switching sequences for SECTOR I



The switching scheme represented in Table 1 is aimed at reducing the number of commutations in a switching cycle, resulting in the reduction of the switching losses [32]. Additionally, the constraint (26) resulting from the available magnitudes of the current vectors, is applied to the current reference.

$$0 \leq \left|i_{ref}\right| \leq \frac{\sqrt{6} i_{dc}}{cos(\theta) + \sqrt{3} sin(\theta)} \tag{26}$$

5 Performance Evaluation

From the continuous AC (5), and DC (6) model equations described in Ch. 2, the controller is formulated form discretised system (9), and it is described via the cost function and control problem of (11), and (12) in Ch. 3. The evaluated model and control structure are shown on Fig. 4. In the following section said EMPC's computational requirements are evaluated, and the Matlab/Simulink simulation results are compared to a classic state feedback controller's dynamic performance.

5.1 Computational Effort

The binary search tree generated for the control problem presented in Fig. 7, and described in Ch. 3. The depth of the search tree is 5 and it has a total number of 29 nodes. It is utilized with the MPT toolbox [30], [31], [33] and it can be used for the computationally optimal real-time implementation of the proposed algorithm on low-cost hardware.

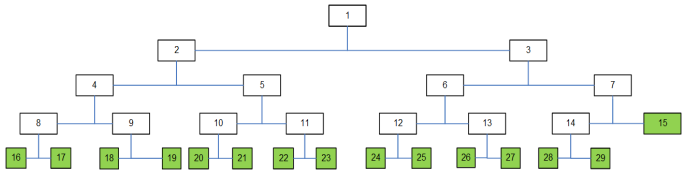


Figure 7

Binary search tree of the controller for a horizon of N=4. The leaf nodes are depicted with filled squares. The depth of the tree is 5.

The search for an active critical region starts from the first level and represents the evaluation in each adjacent node of an inequality of the form: $x \leq K$. Thus, in this case a maximum number of 4 inequalities have to be evaluated to reach the active critical region. Implementing the presented algorithm is straightforward on a DSP processor, for instance from the dsPIC33 family by Microchip. Using the *mac* (multiply and accumulate) instruction the inequality is evaluated for each node using 4 instructions, thus in 80 ns on a 50 MIPS processor (Fig. 8). The active critical region can be reached in a maximum of 400 ns. Compared to the typical sample rate of 10us in the case of a CSR, the real-time implementation on a DSP processor is possible.

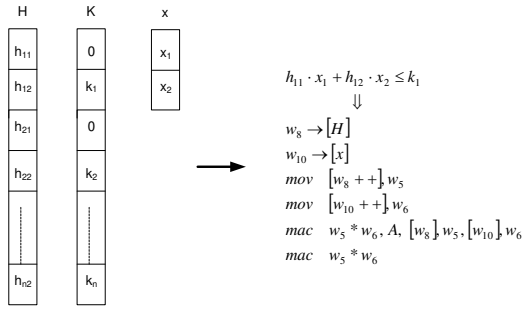


Figure 8

Data organization in the data memory of a single core DSP and the evaluation of a 2-dimensional inequality

5.2 Horizon Performance

With the cost function (11) employed using (18), changing the length of the horizon (*N*) affects the system's complexity illustrated by the partition in the state space shown in Fig. 3, and Fig. 11 presents the step response of the controlled system for different lengths of the horizon. It shows, that the response is not affected by the increase of the horizon above N=2, supporting the choice of this value for Matlab Simulink implementation.

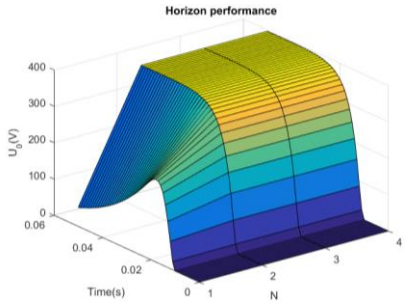


Figure 9
Step response of the system as a function of the horizon length (N)

5.3 Simulation Results

The simulation results are produced with Matlab/Simulik. The discrete model's (9) simulation frequency was $f_s = 10^6 Hz$, with the model parameters represented in Table 3, and with the control structure shown on Fig. 4. The EMPC performance is shown in Fig. 10 and Fig. 11.

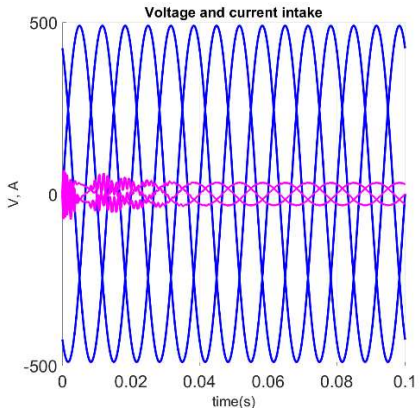


Figure 10
Three-phase voltage and current intake of the CSR with EMPC

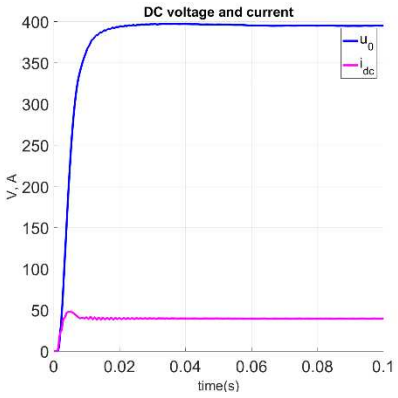


Figure 11
Resulting current and voltage trajectories of the CSR with (EMPC)

More details about the Matlab simulation are presented in [34].

5.4 Comparison with a State Feedback Control

On the DC side, not only the output voltage u_0 but also the inductor current i_{dc} needs to be controlled. Described in [28], a state feedback control with optimal parameters can be used as a reference based on the model properties listed in Table 3, with output voltage u_0 and DC bus current i_{dc} chosen as the state

variables. Since u_0 is a DC quantity in steady state, an integrator signal is introduced to diminish the steady-state error. The structure of the controller is represented in Fig. 12.

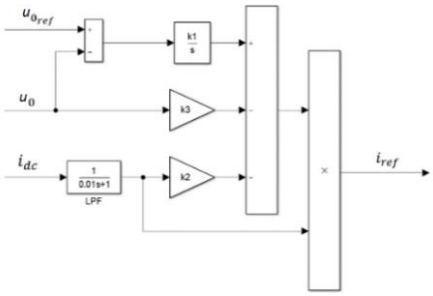


Figure 12
Simple DC side state feedback control structure

The tuning constants applied and calculated according to [24] are:

$$k1 = \frac{\omega_n^3}{1.5U_n\omega_{dc}^2},\ k2 = \frac{1.9\omega_n L_{dc}}{1.5U_n},\ k3 = \frac{2.2\omega_n^2}{1.5U_n(\omega_{dc}^2-1)}, \tag{26}$$

where $\omega_n = 1.1, \omega_{ac} = \frac{1}{\sqrt{L_{ac}C_{ac}}}$, and $\omega_{dc} = \frac{1}{\sqrt{L_{dc}C_{dc}}}$.

The state feedback controllers block on the diagram is taking the controller's place, shown on Fig. 2. The independent outputs are the high pass filter's output $i_{rHF(d)}$ and the controller's output $i_{rcontrol(d)}$. The sum of the independent current values is converted to Clarke frame to be able to govern the switching states of the IGBT's. This can be done because $i_{rHF(d)}$ has only high frequency components and $i_{rcontrol(d)}$ has low frequency components due to the differences in LC time constants, as discussed in the second section. Then, the control signal governing the switches is applied in the same manner, described at the start of Section 3. The state feedback control's performance in comparison with the EMPC is shown in Fig. 13.

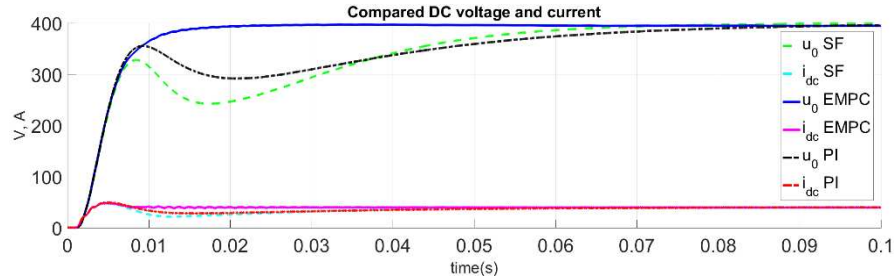


Figure 13
Resulting current and voltage trajectories of the CSR with explicit model predictive control (MPC) compared state feedback control (SF), and simple proportional-integral control (PI), where $P = 0.01$, and $I = 100$, with the respect of constraints described in (19)

Appendix

Table 3
The applied parameters in model and controller design

Parameter	Value	Description
R	0.3 Ω	Phase resistance
R_{load}	10 Ω	Load resistance
L_{ac}	1 mH	AC-side filter inductance
L_{dc}	30 mH	Choke inductance
C_{ac}	30 μF	AC-side filter capacitance
C_{dc}	400 μF	DC-side capacitance
f_s	$10^{-6} Hz$	Simulation frequency
f	50 Hz	Network frequency
f_{pwm}	20 kHz	Modulation frequency
U_n	400 V	Network line voltage
R_w	$\boldsymbol{I}_2$	State weighting matrix
Q_w	$10^{-6}\boldsymbol{I}_2$	Input weighting matrix
$\boldsymbol{N}$	2	Control horizon
$\boldsymbol{\omega}_{\mathrm{n}}$	1.1	undamped oscillation frequency

Conclusions

The constrained, model-based optimal control of a current source rectifier has been presented in this paper. The dynamic model of a three-phase current source rectifier has been developed in Park frame. The proposed model has been examined from the design and implementation points of view with the purpose of explicit model-based predictive control. It proved to be the case that the regular set of differential equations of the CSR appears to be too complex, and contains non-linearity for such a design approach. To address this issue the usage of separated AC and DC equation sets was suggested to avoid linearization and complexity reduction. This solution eliminates bilinearity and enables the application of linear control design techniques. Current-based SVPWM of the three-phase converter has been used with an emphasis on the reduction of switching losses. Throughout the article the explicit model predictive control method is described and the method's effectiveness compared to conventional state feedback control is show. The implementation and simulation experiments have been performed in Matlab/Simulink environment. Moreover, the proper implementation of the system in a modern DSP chip will result in real-time operation.

Acknowledgement

Attila Fodor acknowledges the financial support of Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015. Attila Magyar was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

[1] I. Vajda. Y. N. Dementyev, K. N. Negodin, N. V. Kojain, L. S. Udut, Irina. A. Chesnokova: Limiting Static and Dynamic Characteristics of an Induction Motor under Frequency Vector Control, Acta Polytechnica Hungarica, Vol. 14, No. 6, 2017

[2] B. Ghalem, B. Azeddine: Six-Phase Matrix Converter Fed Double Star Induction Motor, Acta Polytechnica Hungarica, Vol. 7, No. 3, 2010

[3] S. Gupta, R. K. Tripathi: Two-Area Power System Stability Improvement using a Robust Controller-based CSC-STATCOM, Acta Polytechnica Hungarica, Vol. 11, No. 7, 2014

[4] Y. Li, P. Li, Y. Chen, D. Zhang: Single-stage three-phase current-source rectifier for photovoltaic gridconnected system, Conference of Power Electronics and Applications (EPE'14-ECCE Europe), Finland, 2014

[5] B. Exposto, R. Rodrigues, J. G. Pinto, V. Monteiro, D. Pedrosa, J. L. Afonso: Predictive Control of a Current-Source Rectifier for Solar Photovoltaic Grid Interface, Compatibility and Power Electronics (CPE) Conference, Portugal, 2015

[6] M. Chebre, A. Meroufel, Y. Bendaha: Speed Control of Induction Motor Using Genetic Algorithm-based PI Controller, Acta Polytechnica Hungarica Vol. 8, No. 6, 2011

[7] R. Salloum: Robust PID Controller Design for a Real Electromechanical Actuator, Acta Polytechnica Hungarica, Vol. 11, No. 5, 2014

[8] L. Neukirchner, P. Görbe, A. Magyar: Voltage unbalance reduction in the domestic distribution area using asymmetric inverters, Journal of Cleaner Production Vol. 142, Part 4, 20 January 2017, pp. 1710-1720

[9] F. Tahri, A. Tahri, A. Allali and S. Flazi: The Digital Self-Tuning Control of Step a Down DC-DC Converter, Acta Polytechnica Hungarica, Vol. 9, No. 6, 2012

[10] H. Gao, D. Xu, B. Wu, N. R. Zargari: Model Predictive Control for Five-level Current Source Converter with DC Current Balancing Capability, Industrial Electronics Society , IECON 2017 - 43rd Annual Conference of the IEEE, China, 2017

[11] Y. Han, L. Xu, M. M. Khan, C. Chen: Control Strategies, Robustness Analysis, Digital Simulation and Practical Implementation for a Hybrid APF with a Resonant Ac-link, Acta Polytechnica Hungarica, Vol. 7, No. 5, 2010

[12] T. Venkatraman, S. Periasamy: Multilevel Rectifier Topology with Modified Pulse Width Modulation and Reduced Switch Count, Acta Polytechnica Hungarica, Vol. 15, No. 2, 2018

[13] H. Feroura, F. Krim, B. Tabli, A. Laib: Finite-Set Model Predictive Voltage Control for Islanded Three Phase Current Source Rectifier, Conference of Electrical Engineering - Boumerdes (ICEE-B) Algeria, 2017

[14] Z. Yan, X. Xu, Z. Yang, X. Wu: Study of Effective Vector Synthesis Sequence for Three-Phase Current Rectifier, Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC) IEEE, 2015, pp. 1065-1070

[15] A. Ürmös, Z. Farkas, M. Farkas, T. Sándor, L. T. Kóczy, Á. Nemecsics: Application of self-organizing maps for technological support of droplet epitaxy, Acta Polytechnica Hungarica, Vol. 14, No. 4, 2017, pp. 207-224

[16] A. Chatterjee, R. Chatterjee, F. Matsuno, T. Endo: Augmented stable fuzzy control for flexible robotic arm using LMI approach and neuro-fuzzy state space modeling, IEEE Transactions on Industrial Electronics, Vol. 55, No. 3, 2008, pp. 1256-1270

[17] T. Haidegger, L. Kovács, R. Precup, B. Benyó, Z. Benyó, S. Preitl: Simulation and control for telerobots in space medicine, Acta Astronautica, Vol. 181, No. 1, 2012, pp. 390-402

[18] S. Vrkalovic, E. Lunca, I. Borlea: Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants, International Journal of Artificial Intelligence, Vol. 16, No. 2, 2018, pp. 208-222

[19] C. B. Regaya, A. Zaafouri, A. Chaari: A New Sliding Mode Speed Observer of Electric Motor Drive Based on Fuzzy-Logic, Acta Polytechnica Hungarica, Vol. 11, No. 3, 2014

[20] K. Széll, P. Korondi: Mathematical Basis of Sliding Mode Control of an Uninterruptible Power Supply, Acta Polytechnica Hungarica, Vol. 11, No. 3, 2014

[21] A. Kelemen, N. Kutasi, M. Imecs, I. I.I ncze: Constrained Optimal Direct Power Control of Voltage-Source PWM Rectifiers, International Conference on Intelligent Engineering Systems (INES) IEEE International Conference, Las Palmas of Gran Canaria, Spain, 2010

[22] A. Tahri, H. M. Boulouiha, A. Allali and T. Fatima: Model Predictive Controller-based, Single Phase Pulse Width Modulation (PWM) Rectifier for UPS Systems, Acta Polytechnica Hungarica, Vol. 10, No. 4, 2013

[23] M. Rivera, S. Kouro, J. Rodriguez, B. Wu, V. Yaramasu, J. Espinoza and P. Melin: Predictive Current Control in a Current Source Rectifier Operating with Low Switching Frequency, 4th International Conference on Power Engineering, Energy and Electric Drives, Turkey, 2013

[24] A. Godlewska, A. Sikorski: Predictive control of current source rectifier, Selected Problems of Electrical Engineering and Electronics (WZEE), IEEE, 2015, pp. 1-6

[25] N. Muthukumar, S. Srinivasan, K. Ramkumar, K. Kannan, V. E. Balas: Adaptive Model Predictive Controller for Web Transport Systems, Acta Polytechnica Hungarica, Vol. 13, No. 3, 2016

[26] N. Kutasi, A. Kelemen, M. Imecs: Constrained Optimal Control Of Three-Phase AC-DC Boost Converters, Automation Quality and Testing Robotics (AQTR) IEEE International Conference, Cluj-Napoca, Romania, 2010

[27] S. F. Ahmed, Ch. F. Azim, H. Desa, A. T. Hussain: Model Predictive Controller-based, Single Phase Pulse Width Modulation (PWM) Rectifier for UPS Systems, Acta Polytechnica Hungarica, Vol. 11, No. 6, 2014

[28] Y. Zhang, Y. Yi, P. Dong, F. Liu, Yong Kang: Simplified Model and Control Strategy of Three-Phase PWM Current Source Rectifiers for DC Voltage Power Supply Applications, IEEE Journal of emerging and selected topics in power electronics, Vol. 3, No. 4, december 2015

[29] A. Bemporad, M. Morari, V. Dua, Efstratios N. Pistikopoulos: The explicit linear quadratic regulator for constrained systems. Automatica Vol. 38, 2002, pp. 3-20

[30] M. Herceg, M. Kvasnica, C. N. Jones, M. Morari: Multi-Parametric Toolbox 3.0, Proceedings of the European Control Conference, Zürich, Switzerland, 2013, pp. 502-510

[31] N. Kutasi, A. Kelemen, Sz. Matyasi, M. Imecs: Hardware implementation of explicit mode-predictive control of three phase PWM rectifiers, ICCC2010 Eger, Hungary, 2010, pp. 133-136

[32] L. Moussaoui, A. Moussi: An open loop space vector PWM control for CSR-fed field oriented induction motor drive with improved performances and reduced pulsating torque, Proceedings of the 6th WSEAS international conference on Automation & information, March 2005, Part 4. pp. 329-335

[33] M. Kvasnica, I. Rauová and M. Fikar: Automatic code generation for real-time implementation of Model Predictive Control, 2010 IEEE International Symposium on Computer-Aided Control System Design, Yokohama, 2010, pp. 993-998

[34] L. Neukirchner: Constrained Predictive Control of Three-Phase BuckRectifiers Simulation details, http://virt.uni-pannon.hu/ver/index.php/en/projects/30-empc-csr, 2019

Comparison of Gabor Filter Bank and Fuzzified Gabor Filter for License Plate Detection

Vladimir Tadic, Zoltan Kiraly, Peter Odry

University of Dunaújváros, Táncsics Mihály u. 1/A, 2401 Dunaújváros, Hungary, tadityv@uniduna.hu,kiru@uniduna.hu,podry@uniduna.hu

Zeljen Trpovski, Tatjana Loncar-Turukalo

University of Novi Sad, TrgDositejaObradovića 6, 21000 Novi Sad, Serbia, zeljen@uns.ac.rs, turukalo@uns.ac.rs

Abstract: This paper presents a comparative study of improvements to the algorithms for license plate extraction from images captured using conventional, modest-quality cameras. It compares the results and efficacy of two similar algorithms which primarily differ in the raw image pre-processing stage of the initial image. One algorithm uses the Gabor filter bank with distinctly crisp parameters; the other relies on the fuzzified Gabor filter bank with fuzzified parameters to facilitate their adjustment. Results indicate that the fuzzy reasoning introduced for Gabor filter parameter adjustment improves the detection of the components of interest in complex images and adds minimal deviation compared to the Gabor filter bank with distinctly crisp parameters.

Keywords: Gabor Filter; Filter Bank; Crisp Parameters; Parameter Fuzzification; License Plate Detection

1 Introduction

1.1 Background

One of the most common uses of object recognition/computer vision is an automatic detection and recognition of license plates. Many systems, such as those used in industry, parking services, airports, customs control, the military, and the police, contain on board software for license plate recognition. An additional application of license plate recognition is in control systems in areas with limited accessibility (embassies, factories, police and military facilities) and for identifying lost or stolen cars [1].

The recognition of license plate's characters falls under the more general and well explored problem of characters' segmentation and recognition. There are numerous algorithms with an excellent performance that can be easily adapted to any license plate recognition system [2]. However, the extraction of a license plate from an image still remains a highly researched issue within object recognition, with many papers still being published on the topic [1], [3], [4], [5].

The goal of this research is to compare and evaluate two robust algorithms that achieve good performance on the images taken with commercial cameras with all the problems inherent to this image acquisition method. In commercial applications, such as low enforcement, the images are acquired with special REG (special register) cameras that combine special and expensive optics with pre-processing software. The images captured with commercial digital cameras have numerous flaws such as reflections, shadows, and noise, resulting in a poor image quality. The compared algorithms would enable the use of affordable cameras and less strictly controlled scene for a license plate detection.

In this paper the performance and efficacy of two similar algorithms, with the main difference in the pre-processing step, are compared and contrasted. Well-conceived pre-processing step can compensate for flaws in an image acquisition, or ambient conditions, facilitating easier extraction of a license plate from an image in the later stages. The first algorithm used a Gabor filter bank with distinctly crisp parameters (denoted GFBank) [5], while the second algorithm relied on a fuzzified Gabor filter bank (denoted FuzzyGF) [1] with the parameters fuzzified to maximally adjust the filter's response. The problem of filtration using Gabor filters has been accessed by taking into account fuzzy logic, a powerful mathematical apparatus [6], [7], [8] that closely imitates human reasoning. Using this apparatus, the Gabor filter parameters were converted into fuzzy parameters to improve the filter's response [9], [10]. The results of this filtering contribute to more robust detection and extraction of license plates from complex vehicle images.

There are several relevant variables influencing the performance of a license plate detection systems: illumination, quality of the image acquisition system, capturing conditions (distance, angle), and variability in license plate dimensions. However, the main challenge with a plate extraction is the skewness and the presence of extra components of various dimensions similar to a license plate. In such cases, invalid detection can occur, and the wrong component may be extracted. Poor image capture conditions include a greater angle deviation from the regular 90°, and insufficient or excessive capture distance. Sufficient or excessive illumination is also of significant importance for a successful performance. The algorithms compared in this work are resistant to poor capture conditions if the relevant parameters are set within the acceptable limits. It is important to emphasize that the use of Gabor filters provides excellent results in both cases, as Gabor filters successfully handle noise, poor illumination, and different artefacts in the image. These features are of uttermost importance since our initial motivation was

improving the license plate extraction performance for a low-cost image acquisition regardless of illumination conditions.

The rest of this article is organized as follows. The literature overview on this topic is covered in Section 1.2. The Gabor filter and the fuzzy methods are briefly introduced in Section 2. The comparison of license plate detection algorithms, the GFBank and the FuzzyGF is described in Section 3. The validation of the algorithms is explained in Section 4 followed by the concluding remarks.

1.2 Recent Works

Many different procedures for number plate detection and extraction have been recently proposed, such as the fuzzified Gabor filter [1], the method based on the wavelet transform [11], attribute filtering [12], edge detection and noise reduction [13], and mathematical morphology [14]. Each of these approaches has its own limitations and advantages [15], [16], [17]. The quality of the cropped license plates resulting from these procedures is primarily influenced by the quality of the camera that captures the source image and by the quality of the algorithm used for the image processing.

Corneto et al. [18] used a Haar classifier for a plate detection and extraction followed by plate segmentation. The results indicate that the computational cost and accuracy rate considering this approach were acceptable to real time applications, with an execution time less than one second. Babu K et al. [19] applied a Sobel mask for plate region localization followed by morphological filtering and connected components analysis. The plate region was extracted using a bounding box and a template matching method relying on a correlation for recognition of each character in the number plate. Do et al. [2] developed an Android program that processes the image acquired by the built-in camera of a mobile device to capture the license plate number and save it into a database for subsequent applications. They imported open-source Open CV libraries into the project for selected image processing steps to reduce the programming time. The Tesser act engine and neural network were used for optical character recognition (OCR) to convert the license plate image into machine-encoded text. Qiu et al. [20] integrated color and edge detection methods to increase the success rate of locating license plates. They employed connected component analysis and vertical projection methods alternatively to improve the precision and efficiency of segmentation. For character recognition, they applied an improved K-Nearest Neighbours algorithm. Their experimental results indicate that the optimized system had a high license plate recognition rate with an accuracy of 96.75%. Roy et al. [21] developed an algorithm for detection and recognition of Bangladeshi license plates. They noted that the variations among the license plate patterns and complex background of Bangladeshi license plates made it difficult to use the existing algorithms. They selected the portions of green color for matching RGB intensity of the plate. A boundary-based contour algorithm, area, and aspect ratio

were used to locate the license plate in the vehicle region. License plate rows containing registration information were separated using horizontal projection with the threshold value. The characters and digits of the rows were segmented using vertical projection with the threshold value. Finally, template matching was used for recognizing [22], [23], [24] the characters and digits of the license plate. Wang et al. [25] used the Roberts operator for edge detection and morphological erosion, image clustering and region filling for plate detection from complex backgrounds. They used vertical projection for plate segmentation and a back propagation neural network for character recognition. They reported a 97.80% correct license plate recognition rate. Yuan et al. [26] presented a robust and efficient method for license plate detection with the purpose of accurately localizing vehicle license plates from complex scenes in real time. A simple effective image downscaling method was first proposed to substantially accelerate license plate localization without sacrificing detection performance compared with that achieved using the original image. Then, a novel line-density filter approach was proposed to extract the candidate regions, thereby significantly reducing the area to be analyzed for license plate localization. Moreover, a cascaded license plate classifier based on linear SVMs using color saliency features was introduced to identify the true license plate from among the candidate regions. Extensive experiments on the widely used Caltech license plate dataset and their introduced dataset demonstrate that the proposed approach substantially outperformed state-of-the-art methods in terms of both detection accuracy and run-time efficiency, increasing the detection ratio from 91.09% to 96.62% while decreasing the run time from 672 ms to 42 ms for processing an image with a resolution of 1082 × 728. Panchal et al. [27] considered feature extraction using the Harris corner detection algorithm. An aspect ratio limit was set to restrict the license plate viably subsequent to separating each single corner point and using a scanning window approach for plate extraction. Morphological operations were performed in the plate segmentation stage. The overall accuracy of the proposed method was 93.84%. Xie et al. [28] presented a novel convolutional neural network (CNN) – based method for high-accuracy real-time car license plate detection. They proposed a CNN-based MD-YOLO (Multi-Directional "You Only Look Once") framework for multi-directional car license plate detection. They also introduced the prepositive CNN model, ALMD-YOLO (attention-like MD-YOLO) method which performs better than the single MD-YOLO. These solutions can elegantly solve the problem of multi-directional car license plate detection, and can also be deployed easily in real-time circumstances, because of their reduced computational complexity.

Table 1 summarises the above mentioned approaches, main methodology used, database size and the rate of successful localizations. The direct comparison is hampered by the different databases used, and various assumptions and limitation introduced in the methodological steps. The universally applicable algorithm, irrespective of image quality, illumination, capturing settings and licence plate type is still a challenging task.

Table 1
Summary of the mentioned approaches

Related Work	Algorithm	Database size	Correctly localized
Tadic et al. [1]	FuzzyGF	718	97.9%
Yu et al. [11]	Wavelets	765	97.91%
Yimyam et al. [13]	Edge detection and noise reduction	50	82%
Jabar et al. [14]	Mathematical morphology	100	84.19%
Corneto et al. [18]	Haar classifier	1410	92.55%
Babu K et al. [19]	Sobel mask	45	93.33%
Do et al. [2]	Neural network	200	96%
Qiu et al. [20]	Integrated color and edge detection methods	400	96.75%
Roy et al. [21]	Boundary-based contour algorithm	180	93%
Yuan et al. [26]	Line-density filter	126	96.62%
Panchal et al. [27]	Harris corner detection	65	93.84%.
Xie et al. [28]	Convolutional neural network	2049	99.5%

In the majority of studies in this field, the initial license plate images were captured with special register (REG) cameras designed for license plate recognition or with high-quality cameras. These cameras are considerably more expensive than commercial automatic digital cameras. REG cameras combine special optics with software support that contributes significantly to the development of high-quality initial images [1], [5].

The fuzzy set theory was a cornerstone in numerous papers dealing with numerous application scenarios. In the following is a brief overview of the papers related to fuzzy logic. Várkonyi-Kóczy [29] and Moya-Albor et al. [30] successfully applied fuzzy logic to corner and edge detection. Cirneanu et al. [31] proposed a supervised technique for the detection and localization of the optic disc in retinal images. The proposed processing technique is based on Discrete Fourier Transform and Gabor filters. Devasenapati [32] et al. in their paper evaluates the use of fuzzy unordered rule induction algorithm (FURIA) with correlation-based feature selection embedded feature subset selection as a tool for misfire detection. From the results obtained the authors conclude that the combination of statistical features and FURIA algorithm is suitable for detection of misfire in spark ignition engines. Haidegger et al. [33] described the minimalistic models that were tested with linear and PID-fuzzy control options to provide a simple, universal and scalable solution for the challenges of tele-surgery over large distances. Jocic et al. [34] proposed an algorithm for discovering similar nodes in very large directed graphs, with millions of nodes with billions of connections, which is based on the

fuzzy set theory. Vrkalovic et al. [35] presented a model-free sliding mode controllers and Takagi-Sugeno fuzzy controllers for the flux and conductivity control of Reverse Osmosis Desalination Plants.

2 Methods

2.1 Gabor Filter

Gabor filters are extensively used for texture detection and extraction because of their optimal localization properties in both spatial and frequency domain. When a Gabor filter is applied to an image, it gives the highest response at points and at edges where texture changes. By virtue of these characteristics, algorithms based on Gabor filters are successfully applied in computer vision applications [31] such as license plate detection, face detection, texture extraction, character recognition and others [1], [5].

The complex Gabor function in a spatial domain is represented as follows:

$$g(x,y) = s(x,y)w_r(x,y) \tag{1}$$

where the function $s(x,y)$ represents the complex sine wave and the function $w_r(x,y)$ represents the 2D Gaussian.

The complex sine wave is defined as:

$$s(x,y) = e^{j(2\pi(u_0x+v_0y)+\varphi)} \tag{2}$$

where the (u_0, v_0) denotes spatial frequencies and the$\varphi$is the phase of the filter.

This complex sine wave can be split into the real and imaginary part as follows:

$$Re(s(x,y)) = \cos(2\pi\,(u_0x + v_0y) + \varphi) \tag{3}$$

$$Im(s(x,y)) = \sin(2\pi\,(u_0x + v_0y) + \varphi) \tag{4}$$

Parameters (u_0, v_0) represent spatial frequencies in Cartesian coordinates. These spatial frequencies can be represented in polar coordinates as follows:

$$f = \sqrt{{u_0}^2 + {v_0}^2} \tag{5}$$

$$\Phi = arctg^{-1}\left(\frac{v_0}{u_0}\right) \tag{6}$$

and the spatial coordinates are expressed as:

$$u_0 = f cos\,\Phi \tag{7}$$

$$v_0 = f sin\,\Phi \tag{8}$$

Using the previous equations, the complex sine wave is represented as:

$$s(x,y) = e^{j(2\pi f(xcos\Phi + ysin\Phi)+\varphi)} \tag{9}$$

If the 2D Gaussian is defined as:

$$w_r(x,y) = Ae^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2}+\frac{(y-y_0)^2}{2\sigma_y^2}\right)} \tag{10}$$

where A is the amplitude, (x_0, y_0) are the centre of the function and σ_x and σ_y denotes the deviation of the Gaussian by each of the spatial coordinates.

After preliminary consideration, a function $g(x,y,f,\Phi,\sigma)$ which represents the Gabor filter with the orientation of 0° where the f is the spatial frequency and the Φ denotes the filter orientation is obtained as follows [1], [5]:

$$g(x,y,f,\Phi,\sigma) = Ae^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2}+\frac{(y-y_0)^2}{2\sigma_y^2}\right)}e^{j(2\pi f(xcos\Phi + ysin\Phi)+\varphi)} \tag{11}$$

As it can be seen, the expression (11) is a developed form of the expression (1) [1], [5], [10].

The first exponential function represents the 2D Gaussian; the second exponential function represents the 2D complex sinusoid. Again, the A denotes the amplitude, (x_0, y_0) is the center of the function, $\sigma_x$andσ_y are the standard deviations, f is the spatial frequency, $\varphi$is the phase and Φ is the orientation of the Gabor filter.

It has been shown, that the deviation of the filter can be represented via the spatial frequency f as follows [1], [5]:

$$\sigma_x = \lambda k_x \; and \; \sigma_y = \lambda k_y \tag{12}$$

In the previous expressions k_x and k_y represent scaling factors for tuning the bandwidth of the filter. The wave length λ is in pixels and it is related with the spatial frequency as [1], [5]:

$$f = \frac{1}{\lambda} \tag{13}$$

All these parameters are essential in forming the Gabor filter for appropriate filtering [1], [5].

2.2 Gabor Filter Bank with Crisp Parameters

With further development, Gabor filter banks are created, which usually consist of Gabor filters with various scales, frequency and orientations. This concept makes it possible to use a larger number of values for each filter parameter in order to improve the quality of the Gabor filter's response [5].

In the initial research, the first developed Gabor filter bank has variable crisp parameters, wavelength and angle [5]. Both the wavelength and the angle have 17 values each, stored in two separate arrays. These parameters are obtained empirically [5]. Finally, the filter bank's final response is obtained by summing the results of all individual filter responses within the filter bank. The effect and the response of this filter bank is presented in the next sub-section on examples of license plate extraction.

Table 2

Gabor filter bank parameters

Parameters	**Values**
wavelength	9 9 9 9 9 9 9 15 15 15 15 15 15 15 15 15 15 15
angle	-3 -2 -1 0 1 2 3 -90 90 91 -91 92 -92 89 -89 88 -88
k_x	0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.7 0.7 0.7 0.7 0.7 0.7 0.7 0.7
k_y	0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.7 0.7 0.7 0.7 0.7 0.7 0.7 0.7

2.3 Fuzzified Gabor Filter

Another approach in the development of license plate detection algorithms was the inclusion of fuzzy reasoning [1]. The reason for this is the ability of fuzzy logic to express the degree of uncertainty in human thinking. Namely, it was noticed during the experiments [5] that the algorithm with the Gabor filter bank with crisp parameters has detection difficulties when the license plate is at a very skewed angle, is slanting or when the size of the plate excessively varies in the input image. The wavelength and the orientation of the Gabor filter are fuzzified. The values of the intervals for input and output membership functions were empirically obtained [1].

In the fuzzification process a Bell type membership function has been selected as the input membership function in interval [-10 10]. This membership function smoothly and equally describes the angles of interest.Namely, the angle of 0° is the central angle around which it is necessary to detect vertical components, but due to the imperfection of the input image, some adjacent angles come into consideration too, whose influence decreases from zero around both sides according to the Bell's function. This can be expressed with the formula [9], [10]:

$$\mu(z) = \begin{cases} S(z, \phi_1, \phi_2) & z < \phi_2 \\ S(2\phi_2 - z, \phi_1, \phi_2) & \phi_2 \le z \end{cases} \tag{14}$$

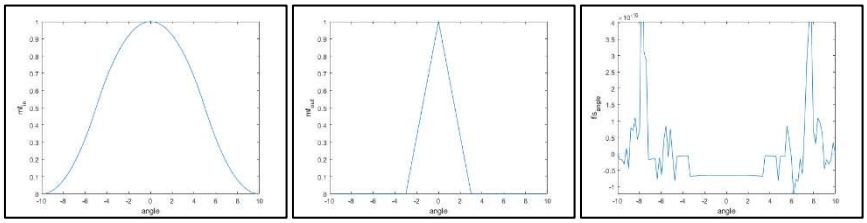
where the interval of the angles is $\phi \in [\phi_1, \phi_2]$. In this research, the end values of the interval were empirically chosen as -10 and 10, and the function *S* is given by the equation (14) [9], [10].

To improve the filtering, the interval of all possible orientations of the filter should be small, or a smaller number of adjacent angles should be used to facilitate the localization of the license plate in the later stages. For these purposes, the triangular shape membership function is used. In this case the interval of the output membership function is [-3 3], which corresponds to the need for small angle deviations from the 0° and it is obtained that $\mu(\phi_{crisp}) = 1$, where the ϕ_{crisp} is the maximum value of the triangular membership function for the angle of 0°. The start and end points of the interval are chosen as [1], [9], [10]:

$$\phi_{crisp} = \frac{\phi_i + \phi_f}{2} \tag{15}$$

and the corresponding interval is selected as $\phi_f - \phi_i = 6°$. The Fig. 1 (a) and Fig. 1 (b) shows both membership functions. It is noticeable that narrowing exists around the zero in Fig. 1 (b).

In this way, the appropriate input and output membership functions have been defined. In order to obtain the fuzzy system and complete the fuzzification process, the IF-THEN rules have to be defined. The complete procedure for fuzzification was simplified using the built in functions based on literature [9], [10]. The IF-THEN rules were generated automatically using the functions provided by Gonzales et al. [9], [10] (function *fuzzysysfcn*) without any additional interfaces. After the defuzzification, an approximation was applied using the built-in lookup table from the literature [9], [10]. The resulting fuzzy system for the angles is presented in Fig. 1 (c).



(a) (b) (c)

Figure 1

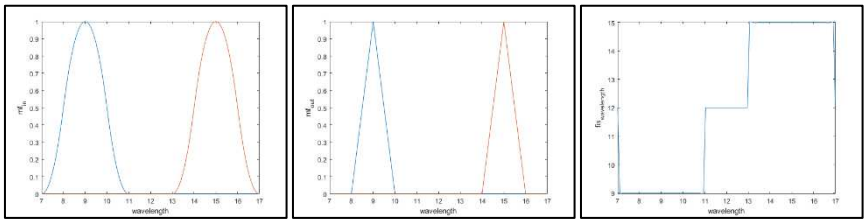
(a) Input membership function for the angles; (b) output membership function for the angles; (c) fuzzy system for the angles

For the fuzzification of the wavelengths the procedure was almost identical with some minor changes, which arise primarily from the nature of the problem.

It has been experimentally obtained that the spatial wavelength's range is from 9 and 15 pixels and the fuzzification has been carried out around these values. For input membership function a generalized Bell's function was used, with two inputs and two intervals that cover the input domain. The intervals centered around 9 and 15 are [7, 11] and [13, 17], respectively (Fig. 2(a)).

Two output segments and the triangular membership function were used to describe the output and achieve the smallest deviations from the central wavelength values. For the central input values of 9 and 15 pixels, the selected intervals were [8, 10] and [14, 16]. These functions have fully described the input and the output of the fuzzy system.

After defining the IF-THEN rules and the fuzzification of the system, a new fuzzy system was obtained that describes the wavelength changes. Fig. 2 (a) and Fig. 2 (b) present the input and output membership functions defined above certain intervals of interest. Finally, the Fig. 2 (c) shows the fuzzy system for the wavelengths.



(a) (b) (c)

Figure 2

(a) Input membership function for the wavelengths; (b) output membership function for the wavelengths; (c) fuzzy system for the wavelengths

Once the fuzzification of the Gabor filter had been completed, the new Gabor filter bank has been developed with fuzzified parameters. Having filtered the input images with the fuzzified Gabor filter, the detection rate has been increased [1]. Also, the effect and the response of the fuzzified Gabor filter bank is displayed at a later stage in the next chapter on an exemplary image.

3 Comparison of the License Plate Detection Algorithms

In the following section, the performance of two algorithms for license plate detection will be compared. The processing pipeline introducing the order and roughly indicating the effect of each algorithm step in an exemplary image is presented in Fig. 3. The only significant difference between the two algorithms is in the pre-processing block performing initial filtering with the specific Gabor filter [1], [5]: Gabor filter bank with the constant parameters (GFBank) and fuzzy Gabor filter (FuzzyGF). After the pre-processing step, both algorithms follow the same pipeline: a sequence of morphological operations, followed by edge detection, threshold application, detection of the frame, plate marking, and cropping. In the following the common algorithms steps will be presented (Fig. 3).

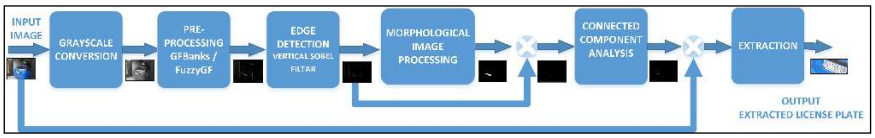


Figure 3

The processing pipeline of the algorithms

After the input image was read and rescaled, the image was grayscaled. This is followed by filtering using the GFBank or the FuzzyGF. The results and a comparison of the two algorithms, i.e., the influence of both versions of the Gabor filters on the initial image, are described in the example below.

The image of the vehicle in Fig. 4 (a) was captured at a rather sharp angle, at daylight, in sunny weather, without an artificial light source. The image was taken with a 2-Megapixel camera on an old mobile phone (Nokia C2). It is clearly visible that the image is slightly blurred due to the sunrays and weather conditions. Fig. 4 (b) displays the same image, in grayscale.

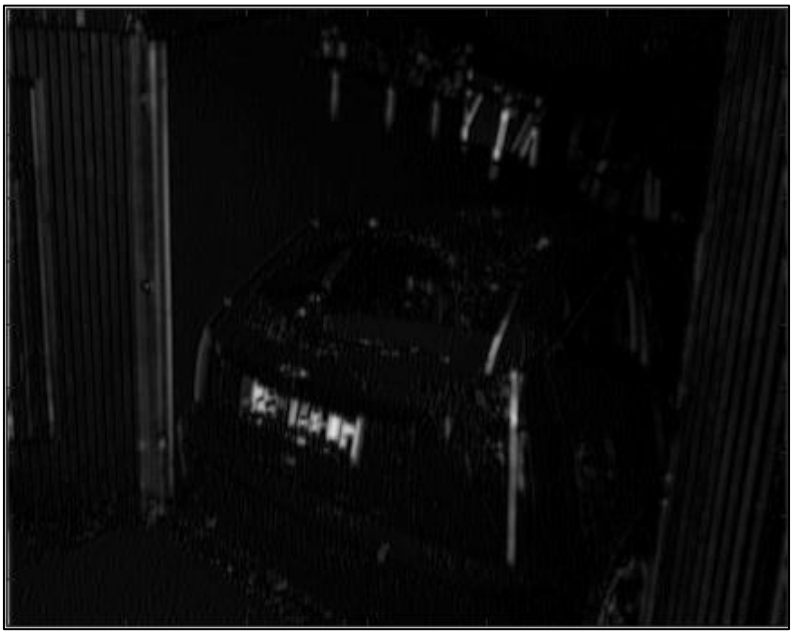
(a) (b)

Figure 4

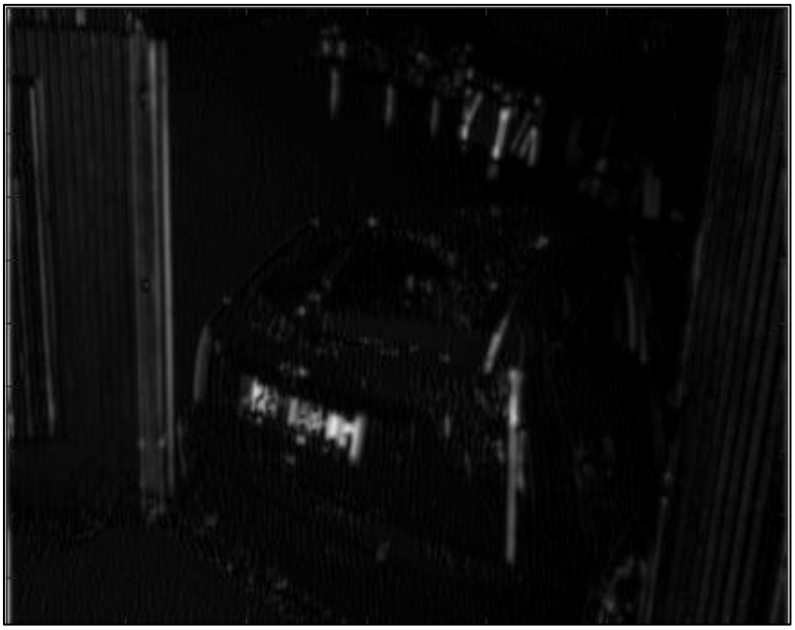
(a) Original colored image and (b) grayscaled image

Fig. 5 (a) displays the response of the GFBank with clearly distinct parameters [5] and (b) the response of the FuzzyGF [1]. As it can be observed, the response is the strongest in the area around the license plate itself; however, even in this region, some darker textures remain.

The response of the image to the fuzzified Gabor filter (Fig. 5 (b)) is the strongest in the area around the plate, however, with fewer dark textures immediately around the plate, which is an improvement achieved with the FuzzyGF.



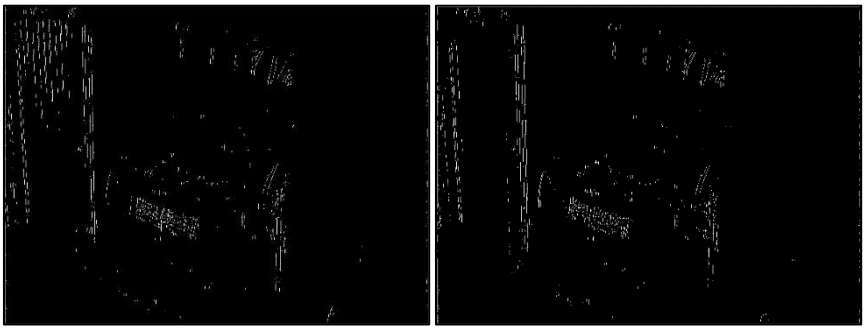
(a)



(b)

Figure 5
(a) GFBank response and (b) FuzzyGF response

After filtering, the next step was edge detection using a vertical Sobel operator [1], [5], [9]. Fig. 6 presents the images after the edge detection for two compared algorithms with different pre-processing step:(a) in case of the GFBank pre-processing, (b) case of the FuzzyGF pre-processing. In both results, there are tiny, dense components of interest in the area around the license plate. This is a promising result towards a more successful final plate detection.



(a) (b)

Figure 6

Result of edge detection to: (a) GFBank response and (b) FuzzyGF response

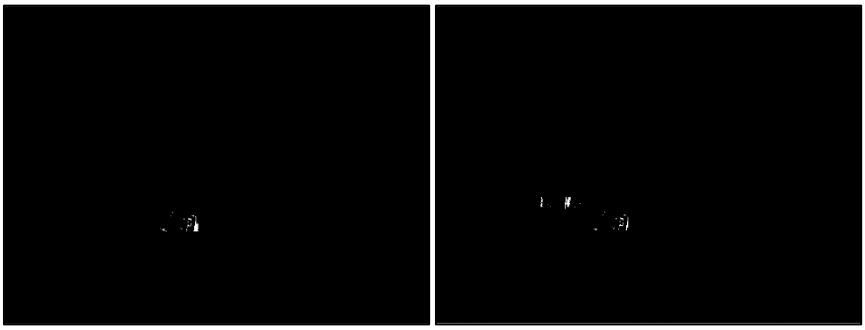
After edge detection, a sequence of several morphological operations was performed and an algorithm for the analysis of connected components was applied [1], [5], [9], [19]. The objective of this step is to remove the unnecessary parts of the image and to form a binary mask of the license plate region, which facilitates the extraction of the plate from the complex vehicle image and its surroundings [1], [5], [10].

(a) (b)

Figure 7

(a) Binary mask of the algorithm with the GFBank and (b) binary mask of the algorithm with the FuzzyGF

Fig. 7 displays the resulting binary masks of both approaches. It is clearly visible that the mask is somewhat narrower in the left image (GFBank) than on the right (FuzzyGF).



(a) (b)

Figure 8

(a) Extracted content based on the binary mask of the GFBank and (b) extracted content based on the binary mask of the FuzzyGF

Fig. 8 displays the binary content of the images in both examples, extracted with the appropriate binary masks [1], [5], [10].In Fig. 8 (a), where the filter bank with crisp parameters was used, the content has been marginally cropped on the left side of the plate and the content of the plate is not visible in the binary image. In Fig. 8 (b), the entire plate is extracted, but the content of the plate is not visible due to the poor quality of the input image.

Finally, after completion of the morphological filtering and connected component analysis [1], [5], [10] the license plate was extracted and cropped from the original colored image. As demonstrated on the provided example, the improvement of the algorithm based on the FuzzyGF is significant, as the content of the plate is entirely extracted and can easily be recognized. In Fig. 9 (a), the loss of two almost entire characters in this example renders the recognition of the license plate impossible. The reason for such a result is an extremely sharp angle, for which the fuzzified filter provides a superior response. This result proves the potential of the fuzzification of the Gabor filter parameters in the pre-processing step for the detection of the vicinity of the license plate.

(a) (b)

Figure 9

(a) Plate extracted with a GFBank and (b) plate extracted with a FuzzyGF

4 Validation of the Algorithms

Both algorithms were validated on a dataset of 718 vehicle images captured with a classical modest-quality digital camera [http://appl-dsp.com/wp-content/uploads/2019/06/test%20database2.zip] [1]. The images are taken both during a day, and during night, with poor illumination conditions. Also, the assumption is that the slant of the license plates is approximately in interval [-10 10] degrees. The distance between the vehicle and camera is between 2-10 m. The percentage of correctly localized license plates is a commonly used measure for the evaluation of algorithms for license plate detection. Table 3 presents a performance summary of the two compared procedures. The performance improvement, as measured by the percentage of precise detection and extraction of a plate from a complex image has been noted in a fuzzified approach used in FuzzyGF.

Table 3
Summary of the detection results for different pre-processing approaches in FuzzyGF and GFBank

Algorithm	Number of Images	Correctly localized	Error (Miss Rate)
GFBank	718	94%	6%
FuzzyGF	718	97.9%	2.1%

The performance of the two presented and compared algorithms is similar to the performance of other license plate detection algorithms [36], [37]. For example, the algorithm proposed by Tan et al. [36], uses the morphological operations for detection and separation of the license plate region similarly to the algorithms compared in this paper. The algorithm pipeline proposed in [36] starts with a contrast enhancement, followed by a difference of Gaussians to improve the edge visibility, and application of a vertical Sobel detector combined with morphological operations to extract the binary mask. As a final step, the geometric properties of the detected areas are extracted and forwarded to the rule-based candidate filtering as a validation step to identify the license plate candidate [36]. The need for a more complex procedure in [36] was imposed by a specific problem of license plate detection in Malaysia. Namely, the plates in this country differ by type, fonts, character ordering and fonts, with additional personalized possibilities. The performance of this algorithm, as evaluated on the proprietary Malaysian database, is 96.9% of successful detection [36]. It is worth noting that the database images were captured under predefined conditions with respect to the angle and distance to control the experiments [36]. The reported performance of 96,9% of successful localization is comparable to results presented in this paper.

The development of a universal algorithm providing robust results regardless of the illumination conditions, image quality, capturing settings, and license plate's type would be a challenging image processing task. Imposing some limitations to any of these variables are needed to produce an image processing algorithm with a stable performance. In the low enforcement applications, the controlled

illumination, capturing distance and angle, and image quality (by using a special REG cameras) ensure robust performance. With the commercial image qualities, as used in this work, the limitations have to be imposed at least on the capturing settings (distance and angle), with similar examples in [1], [2], [5], [11], [18], [25], [28], [36].The fair comparisons of the license plate algorithms developed under the similar assumptions and for the predefined settings would be possible only over the same image database. However, such an exhaustive comparison is beyond a scope of this paper.

The FuzzyGF used in this work has achieved a markedly better performance as compared to GFBank. A plausible explanation is an improved robustness to the image capturing conditions enabled by a fuzzy reasoning. In the applications where some parameters are imprecisely or incorrectly set (such as the input image for detection of license plates, where the precise angle of shooting, brightness, etc. are not known in advance), the use of the fuzzy logic is a good solution, since the fuzzy system successfully uses human reasoning in solving the problem [1], [9], [10], [37]. In the reality, the boundaries between the wavelength and the angle values may be ambiguous and imprecise, and it is difficult to determine whether an input wavelength or angle belongs completely to a certain interval. For that reason, using the concept of the fuzzy set theory [1], [6], the GFBank parameters are replaced with Bell's and triangular fuzzy membership functions within the Gabor filter's parameter calculation [1], [37], [38]. In practice, any membership function can be chosen to describe the similarities between two data sets. As there are no strict boundaries in the fuzzy membership functions and as the membership functions vary continuously and smoothly around chosen values, FuzzyGF performs robustly to slight changes in the wavelength and the angle values.

Conclusions

This paper evaluates the performance improvement in the algorithm for detection of the license plates in complex images when different filtering procedures in the pre-processing step are applied. The first validated pipeline in a pre-processing block relies on the Gabor filter bank with distinctly crisp parameters, while the other uses the Gabor filter with the fuzzified parameters. The fuzzy logic provides the possibility to advance the problems faced in signal and image processing closer to human reasoning. Both algorithms for the detection of license plates in complex vehicle images, captured with a commercial digital camera, were validated on the same dataset and presented using an example of license plate extraction. In the analysis of the intermediate results, it was evidenced that the GFBank with distinctly crisp parameters would detect and extract the plates with a certain loss, which would later render the full recognition of the characters on the plates impossible. Owing to the method used to capture the images and the modest characteristics of the commercial camera, the initial images included numerous artefacts, reflections, and shadows. The images in the database used in this paper are of commercial, thus lower quality, than the images captured in professional

law enforcement applications under controlled ambient conditions using a REG camera.

Future projects will focus on a development of an improved license plate detection procedure addressing different limitations that affect both plate segmentation and character recognition. Finally, it should be noted that the procedures described herein can be used to address many other problems in digital image processing and object recognition. Based on the above results, the main conclusion of this study is that owing to fuzzy reasoning, the FuzzyGF has better detected components of interest in complex images and displayed minimal deviation, compared to the GFBank with distinctly crisp parameters. This robustness to the capture setting was achieved using fuzzy logic, which flexibly adjusts the filter parameters to the current problem, which is not possible using classical mathematical reasoning.

Appendix

The pseudo-code of the algorithm is given in this appendix. The high level view on the development of the license plate extraction software is amended with the relevant references providing additional details and reasoning in the parameter selection.

Pseudo-code of the algorithm	
A=read (INPUT)	\\ reading the input color image
A1=grayscale(A)	\\ converting to grayscale
gabor_response=gabor_filter(A1)	\\ Gabor filter: GFBank [5, page 3] or FuzzyGF [1, pages 5-7]
edge=edge_Sobel(gabor_response)	\\ Sobel edge detector and binarization
morph=morphological_processing(edge, se)	\\ morphological image processing using operations and structural elements as in [1], [5] for extraction of the binary mask candidates
label=conn_comp(morph,8)	\\ calculating and analyzing the 8-connected components in order to extract the binary mask
mask=max(label)	\\ the candidate with the max. number of 8-connected components is the binary mask
bw=imbinarize(A1)	\\ binarization of the input grayscaled image
lp_area=bw& mask	\\ extraction of the license plate area
reg_prop=regionprops(lp_area)	\\ calculating the region properties of the license plate area
extracted_lp=crop(a, reg_prop)	\\ extracting the license plate from the input coloured image based on bounding box specified by the calculated region properties

Acknowledgement

This work is supported by the European Union, EFOP-3.6.1-16-2016-00003 project and in part supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, III43002, III47020.

References

[1] Vladimir Tadic, Miodrag Popovic, Peter Odry, "Fuzzified Gabor filter for license plate detection", Enginnering Applications of Artificial Intelligence, 48 (C), pp: 40-58, 2016, ISSN: 0952-1976

[2] Hung Ngoc Do, Minh-Thanh Vo, Bao Quoc Vuong, Huy Thanh Pham, An Hoang Nguyen, and Huy Quoc Luong, "Automatic License Plate Recognition Using Mobile Device", International Conference on Advanced Technologies for Communications (ATC), 978-1-5090-2710-1/16, 2016, IEEE

[3] Gurwinder Kaur, ReechaSharma,"A Systematic Performance Comparison of Artificial Intelligence Techniques used for ALNPR System", Research Cell: an International Journal of Engineering Sciences, 17(1), pp: 161-167, 2016, ISSN: 2229-6913

[4] Helli, B., Moghaddam, M. E., "A text-independent Persian writer identification based on feature relation graph (FRG)", Jornal of Pattern Recognition, 43(6), pp: 2199-2209, 2010, ISSN: 0031-3203

[5] Vladimir Tadić, Željen Trpovski, Peter Odry, "License Plate Detection using Gabor Filter Banks and Texture Analysis", Proc. 9th IEEE International Symposium on Intelligent Systems and Informatics, Subotica, pp. 381-386, September 2011, ISBN: 978-1-4577-1975-2

[6] Zadeh, L. A., "Fuzzy sets", Inf. Control, Vol. 8, pp:338-353, 1962

[7] József Menyhárt, Róbert Szabolcsi, "Support Vector Machine and Fuzzy Logic", Acta Polytechnica Hungarica, Vol. 13, No. 5, 2016

[8] László Horváth, Imre J. Rudas, "Information Content Driven Model for Virtual Engineering Space", Acta Polytechnica Hungarica, Vol. 15, No. 2, 2018

[9] Gonzales, R. C., Woods, R. E., "Digital Image Processing", 3rd Edition, New Jersey: Pearson Prentice Hall, 2008

[10] Gonzales, R. C., Woods, R. E., Steven L. Eddins, S. L. "Digital Image Processing Using MATLAB", 2nd Edition, Gatesmark, LLC, 2009

[11] Shouyuan Yu, Baopu Li, Qi Zhang, Changchun Liu, MaxQ.-H. Meng, "A novel license plate location method based on wavelet transform and EMD analysis", Pattern Recognition, 48(1), pp: 114-125, 2015, ISSN: 0031-3203

[12] Juan Climenta, Luiz S. Oliveira, "A new algorithm for number of holes attribute filtering of grey-level images", Pattern Recognition Letters, 53(C), pp. 24-30, 2015, ISSN: 0167-8655

[13] WorawutYimyam, MahasakKetcham, "The Automated Parking Fee Calculation Using License Plate Recognition System", IEEE International Conference on Digital Arts, Media and Technology (ICDAMT) 2017, DOI: 10.1109/ICDAMT.2017.7904985

[14] Khadija Ahmad Jabar, Mohammad FaidzulNasrudin, "Libyan Vehicle Plate Recognition Using Region-Based Features and Probabilistic Neural Network", Journal of Theoretical and Applied Information Technology, 94(1), pp: 104-114, 2016

[15] K. S. Raghunandan, PalaiahnakoteShivakumara, Hamid A. Jalab, Rabha W. Ibrahim, G. Hematha Kumar, Umapada Pal, Tong Lu, "Riesz Fractional Based Model for Enhancing License Plate Detection and Recognition", IEEE Transactions on Circuits and Systems for Video Technology, Volume: PP, Issue: 99, 2018, DOI: 10.1109/TCSVT.2017.2713806

[16] AsadollahShahbahrami, Babak Abad Foomani, AlirezaAkoushideh, "A Style-Free and High Speed Algorithm for License Plate Detection", 10th Iranian Conference on Machine Vision and Image Processing, 2017, DOI: 10.1109/IranianMVIP.2017.8342372

[17] Rafique, M. A., Pedrycz, W., Jeon, M., "Vehicle license plate detection using region-based convolutional neural networks", Soft Computing, Springer-Verlag GmbH Germany, 2017, doi:10.1007/s00500-017-2696-2

[18] G. L. Corneto, F. A. Silva, D. R. Pereira, L. L. Almeida, A. O. Artero, J. P. Papa, V. H. C. de Albuquerque and H. M. Sapia, "A New Method for Automatic Vehicle License Plate Detection", IEEE Latin America Transactions, Vol. 15, No. 1, Jan. 2017, DOI: 10.1109/TLA.2017.7827890

[19] Mahesh Babu K, M V Raghunadh,"Vehicle Number Plate Detection and Recognition using Bounding Box Method", International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), ISBN No.978-1-4673-9545-8. 2016

[20] Jianing Qiu, Naida Zhu, Yi Wei, XiaoQing Yu, "An Optimized License Plate Recognition System for Complex Situations", IEEE International Conference ICALIP, 2016, 978-1-5090-0654-0/16

[21] Animesh Chandra Roy, Muhammad Kamal Hossen, Debashis Nag, "License Plate Detection and Character Recognition System for Commercial Vehicles based on Morphological Approach and Template Matching", IEEE International Conference iCEEiCT, 2016, 978-1-5090-2906-8/16, 2016 IEEE

[22] WichaiPuarungroj, Narong Boonsirisumpun, "Thai License Plate Recognition Based on Deep Learning", Procedia Computer Science, Vol. 135, pp:214-221, 2018, DOI:10.1016/j.procs.2018.08.168

[23] Faming Shao, Xinqing Wang, Fanjie Meng, Ting Rui, Dong Wang, Jian Tang, "Real-Time Traffic Sign Detection and Recognition Method Based on Simplified Gabor Wavelets and CNNs", Sensors, 18(10). pii: E3192, 2018, doi: 10.3390/s18103192

[24] Youting Zhao, Zhi Yu, Xiying Li, "Evaluation Methodology for License Plate Recognition Systems and Experimental Results", IET Intelligent Transport Systems, 12(5) pp:375-385, 2018, DOI:10.1049/iet-its.2017.0138

[25] Wang Naiguo, Zhu Xiangwei, Zhang Jian, "License Plate Segmentation and Recognition of Chinese Vehicle Based on BPNN", 12th International Conference on Computational Intelligence and Security, 2016, DOI 10.1109/CIS.2016

[26] Yule Yuan, Wenbin Zou, Yong Zhao, Xinan Wang, Xuefeng Hu, and Nikos Komodakis, "A Robust and Efficient Approach to License Plate Detection" IEEE Transactions on Image Processing, pp:1102-1114, 2016, DOI: 10.1109/TIP.2016.2631901

[27] Tejendra Panchal, Hetal Patel, Ami Panchal, "License Plate Detection using Harris Corner and Character Segmentation by Integrated Approach from an Image", Procedia Computer Science, Vol. 79, pp:419-425, 2016, DOI: 10.1016/j.procs.2016.03.054

[28] LeleXie, Tasweer Ahmad, LianwenJin, Yuliang Liu, Sheng Zhang, " A New CNN-Based Method for Multi-Directional Car License Plate Detection", IEEE Transactions on Intelligent Transportation Systems, 19(2), pp:507-517, 2018, DOI: 10.1109/TITS.2017.2784093

[29] Annamaria R. Varkonyi-Koczy, "Fuzzy logic supported corner detection", Journal of Intelligent & Fuzzy Systems, 19(1), pp. 41-50, 2008

[30] Ernesto Moya-Albor, Hiram Ponce, Jorge Brieva, "An Edge Detection Method using a Fuzzy Ensemble Approach", Acta Polytechnica Hungarica, Vol. 14, No. 3, 2017

[31] Stefan Cirneanu, Loretta Ichim, Dan Popescu, "Improvement of Optic Disc Localization using Gabor Filters", 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, 2018, DOI: 10.1109/TSP.2018.8441261

[32] S. BabuDevasenapati and K. I. Ramachandran, "Hybrid Fuzzy Model Based Expert System for Misfire Detection in Automobile Engines", International Journal of Artificial Intelligence, Vol. 7, No. A11, pp. 47-62, 2011

[33] Tamas Haidegger, Levente Kovacs, Radu-Emil Precup, Balazs Benyo, Zoltan Benyo, Stefan Preitl, "Simulation and control for telerobots in space medicine", ActaAstronautica, Vol. 181, No. 1, pp. 390-402, 2012

[34] Marko Jocic, Endre Pap, Anikó Szakál, DjordjeObradovic, Zora Konjovic, "Managing Big Data Using Fuzzy Sets by Directed Graph Node Similarity", Acta Polytechnica Hungarica, Vol. 14, No. 2, pp. 183-200, 2017

[35] S. Vrkalovic, E.-C. Lunca and I.-D. Borlea, "Model-Free Sliding Mode and Fuzzy Controllers for Reverse Osmosis Desalination Plants", International Journal of Artificial Intelligence, Vol. 16, No. 2, pp. 208-222, 2018

[36] Jinn-Li Tan, Syed A. R. Abu-Bakar, Musa M. Mokji, "License plate localization based on edge-geometrical features using morphological approach", Image Processing (ICIP) 2013 20th IEEE International Conference

[37] Mosabber U. Ahmed, TheerasakChanwimalueang, SudhinThayyil, Danilo P. Mandic, "A Multivariate Multiscale Fuzzy Entropy Algorithm with Application to Uterine EMG Complexity Analysis", Entropy 2017, 19, 2; doi:10.3390/e19010002

[38] Cheng-Hung Lin, Yong-Sin Lin, and Wei-Chen Liu, "An Efficient License Plate Recognition System Using Convolution Neural Networks", Proceedings of IEEE International Conference on Applied System Innovation 2018, IEEE ICASI 2018- Meen, Prior & Lam (Eds), ISBN 978-1-5386-4342-6

[39] AyodejiOlalekanSalaua, Thomas KokumoYesufua, Babatunde Sunday Ogundareb, "Vehicle plate number localization using a modified GrabCut algorithm", Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.01.011

An Efficient Routing Algorithm for Wireless Sensor Networks based on Centrality Measures

Stephanie Mwika Mbiya, Gerhard P. Hancke, Bruno Silva

Advanced Sensor Networks Research Group, Department of Electrical, Electronic and Computer Engineering, University of Pretoria
Lynnwood Road, Pretoria 0002, Republic of South Africa
u14174406@tuks.co.za, gerhard.hancke@up.ac.za, bruno.silva@up.ac.za

Abstract: A routing algorithm for wireless sensor networks with a random distribution in a target observation area is proposed. In practice, selecting a path to route data from a source node to a destination node in a sensor network is very useful. An investigation is carried out on the combination of centrality measures and a routing algorithm to determine whether this can improve the route selected by the network's decision. Various measures of centrality are used and the network's response is evaluated with regards to the route selected by the network's decision when some nodes fail. It is demonstrated through simulations that controlling sensor nodes efficiently with a high measure of centrality gives a network the ability to resist node failures or attacks. Furthermore, this provides the network with high failure tolerance. In this paper, a routing algorithm that uses centrality measures to select the shortest path (a low-energy path between the source and destination node) is implemented.

Keywords: Graph Theory; Routing algorithm; Shortest path; low cost; Wireless Sensor Network

1 Introduction

The routing problem in wireless sensor networks is to select routing paths between nodes in the network so that data can be forwarded to the nearest node with the smallest distance (e.g. number of hops) within a random network. This also includes establishing connections from source nodes to the base station (BS).

Routing is required in higher level decision making when packets in the network have to be forwarded from their source to their destination via intermediate nodes by using various mechanisms to compute the distances between nodes within the network.

The main challenge of routing is to reduce the energy consumption without compromising the network's reliability, whilst keeping fault tolerance high.

One of the limitations of existing routing algorithms is fault tolerance. Fault tolerance is important as it enables the network to still function even if some nodes are disconnected. In some algorithms, there is considerable overhead involved as messages have to be exchanged between a large number of nodes when selecting routing paths. Algorithms which can select paths with a lower number of messages are preferable.

In literature, when routing algorithms choose a path, the data packet is collected through the network, and the source node sends it to the nearest nodes in the network, randomly. In this paper, we study the case where the data packet has been collected through the routing network but the source node does not send it to all other nodes at random; rather, it chooses which node to send the data packet to. We propose an algorithm where the closest node to the source node is chosen for data transmission.

In the proposed algorithm, the network has to be reliable regardless of connectivity.

This also includes investigating whether routing algorithms based on centrality measures can outperform routing algorithms such as Dijkstra's in wireless sensor networks.

The proposed algorithm computes the smallest number (Snbr) of nodes between the source and the nearest node. The nearest node with Snbr is treated as if it is a source node, and again the process is repeated to find the nearest node that has Snbr. The cost of this new found node becomes the total cost calculated from the original source node. This process is repeated until the destination node is reached. What distinguishes this dynamic approach to routing from the greedy approach is that the former will always lead to the optimal solution, while in the latter case, one is not assured of obtaining the optimal solution although the solution might be satisfactory.

The proposed algorithm uses a dynamic approach as opposed to Dijkstra's algorithm which uses a greedy approach.

- **Contribution:** In this paper, a shortest path routing algorithm is proposed pertaining to the calculation of the shortest distance among the connected nodes in a network. This proposed routing algorithm addresses key issues existent in routing algorithms such as Dijkstra.

  The research investigates choosing Snbr where 0 is the source or a permanent node denoted as S sending information to the connected node, which we call the tentative node and is the shortest distance away. This procedure is repeated for the previous node (which is denoted as permanent node) until the process reaches the destination node.

- The existing routing algorithm such as Dijkstra Algorithm is compared to the proposed algorithm based on centrality measures. However the result

shows in the implementation has beter performance in term of releability of nodes within the network can further connect nodes for WSNs and reducing the energy consumption.

2 Network Modelling

2.1 Overview of Centrality Metrics

This subsection starts with some notions required to understand centrality measures. Centrality measures (or metrics) depend on the shortest paths between two nodes. In Figure 1, we observe that although 8 and 5 have the largest measure of betweenness, 0 is top in terms of closeness. This is intuitive, as 6 seems to fill a more central position. On the other hand, 8, whose ranking is above 2 with regards to closeness and degree, takes a second position as far as betweenness is concerned. A routing path to 6 or 2 has to go through 8. This is illustrated in Figure 2.

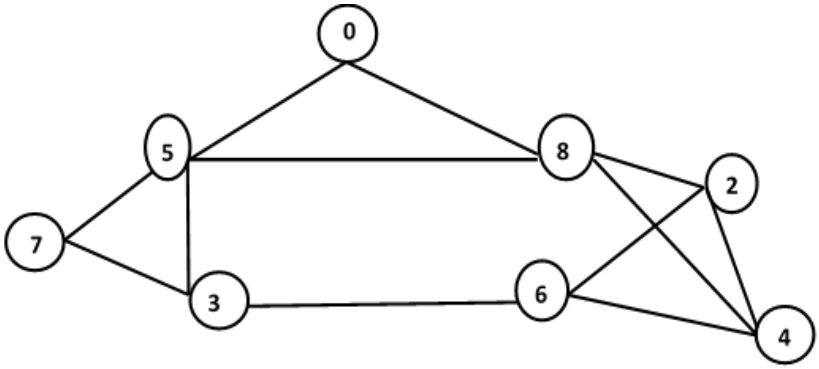


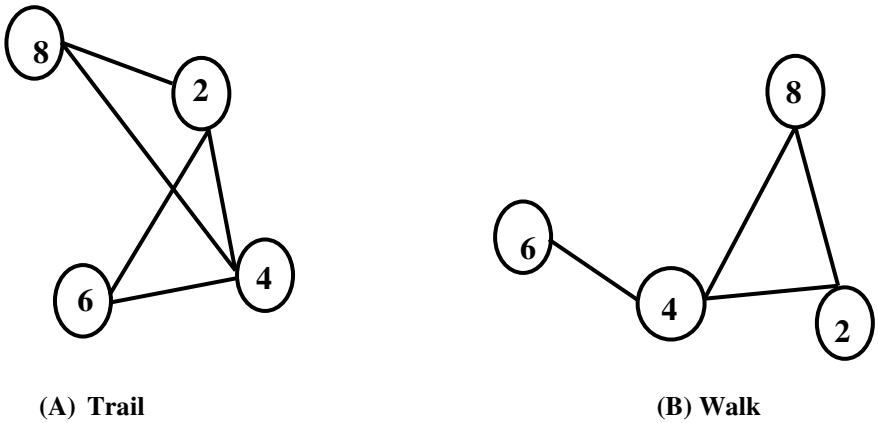
Figure 1
A simple network

Some necessary notions to understand centrality measures include:

- Walk: A walk is a sequence of edges and vertices, where each of the end edges are two vertices which are adjacent.
- Trail: A trail is a walk with no repeated edges.

- Gossip: When gossiping is used as a routing method, data from a node is forwarded to a randomly selected neighbor node sequentially until a packet reaches the destination node

For instance, Figure 2 indicates that route 8, 2, 4, 6, 2 is neither a path nor a trail. Some information, e.g., a gossip, usually accelerates on a trail. Thus, 4 may hear a gossip from 2 as well as 6, where 4 could hear it from 6, but the chance that 6 and 4 can gossip back to each other is not high. Figure 2 (B) indicates that route 8, 2, 4, 6 forms a walk. It is neither a trail nor a path. Communication between node A and node B can be reciprocal; 6 can write a bill to 4 in a single transaction, 4 can return it to 6 in another. Figure 2 (A) and 2 (B) show nodes that have high closeness and betweenness.



(A) Trail **(B) Walk**

Figure 2

Different nodes with different degrees of closeness

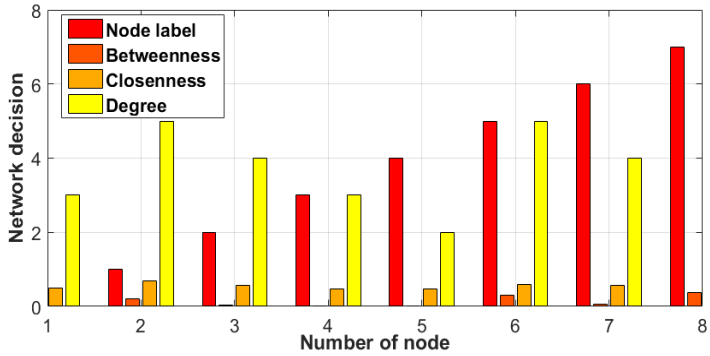


Figure 3

Centrality measures for the network

2.2 Modelling of Centrality Measures

It is possible to have a mathematical representation of a network. In this representation, a network is visualised as a graph of nodes and links. The links are edges that connect the nodes. Therefore, a network can be defined as $G(V, E)$:

$$G(V, E) = \{(v, e): v \in V, e \in E\} \quad (1)$$

in which V is the set of nodes and E the set of links (or edges). The links in the network represent paths of communication. Alternatively, instead of writing $G(V, E)$, we simply use the notation (V, E) [1], [2], [3].

The connection between sensors can be represented by a matrix, called the adjacency matrix. Denote this adjacency matrix by A. The matrix A has order $n * n$ where n is the total number of sensors. An entry of A, a_{ij} is given by:

$$\sigma(p_j, p_i) = \begin{cases} 1, & if\ p_{i\ is\ connected\ to\ p_j} \\ 0, & Otherwise\ . \end{cases} \quad (2)$$

Here, sensors are labelled 1,2,3, ... up to n and sensor p_i is represented by the label i. Note that A is symmetric if sensor p_i is connected to sensor p_j, then p_i is also connected to p_i and we assume that a given sensor p_i is not connected to itself. Therefore $a_{ii} = 0$.

In this section, we denote the distance between sensors p_i and p_j by $w(i, j)$. The shortest path between sensors is the path whose number of links that connect sensor nodes is minimal. If there is no path between p_i and p_j then $w(i, j) = \infty$. The network's diameter is calculated by taking the average of the shortest paths between two pairs of nodes [4], [5], [6], [7].

Extensive research has been done on probabilistic graphs. In this study, we concentrate on sensors that are static where they acquire information, execute decisions and exchange data with neighbours. The data must be communicated to a sink node.

Utilizing the mathematical framework in a network, many algorithms have been developed to study networks and measures have been defined. Centrality measures provide information about network features and how data is spread over the network.

- **Connectivity centrality metric**

The degree of centrality of node p_i is simply the number of links connected to it [3], [2], [4]. It is computed from the formula:

$$\deg(p_i) = \sum_{p_j \in V} \sigma(p_j, p_i), \quad (3)$$

where,

$$\sigma(p_j, p_i) = \begin{cases} 1, & if\ p_{i\ is\ connected\ to\ p_j} \\ 0, & Otherwise\ . \end{cases} \quad (4)$$

With regards to the matrix A, we can alternatively define $\deg(p_i)$ to be the sum of all entries in row i or column i, i.e.,

$$\deg(p_i) \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ji}. \quad (5)$$

For a sensor p_i, $\deg(p_i)$ gives an indication of the sensor's impact on how information is communicated in a network. This occurs in such a way that the more connections a sensor has, the higher the probability of making a high contribution to the communication of information. Hence, such sensors are crucial for information transfer. Given a network G with n nodes, one defines the normalized degree of centrality $CD(p)$ for node p as:

$$C_D(p) = \frac{\deg(p)}{n-1}. \quad (6)$$

where one can extend $C_D(p)$ to the entire network.

Let a node p^* be such that $\deg(p^*)$ is the greatest. Furthermore, let X be the component (connected) of G that maximizes the quantity H given by,

$$H = m \sum_{j=1}^{n_x} \left[C_D(y^*) - C_D(y_j)\right]. \quad (7)$$

where n_x enumerates nodes in X, y^* the node with highest degree centrality in X, and y_j is a node in X. Hence, the centrality degree of G is the following Quantity C_D,

$$C_D = \frac{\sum_{i=1}^{n} C_D(p^*) - C_D(p_i)}{H}. \quad (8)$$

Note that when G is connected, then H is maximal. A connected graph is a graph where every node is linked to all nodes. In that case, $H = (n-1)(n-2)$, and C_D becomes [2], [3], [8], [9].

$$C_D = \frac{\sum_{i=1}^{n} C_D(p^*) - C_D(p_i)}{(n-1)(n-2)} \quad (9)$$

Another assumption is that $n \geq 3$.

- **Closeness centrality metric**

The closeness centrality metric of a node p_i is the reciprocal of the total number of path lengths that are the shortest distance from the rest of the nodes [3], [4], [5]. In a network that is connected, the centrality of a node p_i is calculated from the status of p_i and the average distance to all other nodes. We denote the status of a node p_i by S_{p_i}. This implies that the status of a node p_i is the ratio of the sum $w(i,j)$ for all nodes p_j to the total number of such possible paths, i.e., $n(n-1)$. Hence, the status of a node p_i is given by [10], [11]:

$$S_{p_i} = \frac{1}{n=(n-1)} \sum_{j=1}^{n} w(ij), \quad (10)$$

The closeness centrality metric Cp_i of a node P_i is the reciprocal of its status. It represents the extent to which the node is able to acquire information through

other nodes and relay it. The central node has a high closeness metric because it is the sink. On average, the nodes that have the closest proximity are those which are positioned at a smaller number of hops to other nodes, and form a group of nodes which enable higher information transfer in the network. The capacity of the network is given by:

$$G = \{p_i \in V : C(p_i) = max_{p_i} \in VC(p_j)\}. \quad (11)$$

- **Betweenness centrality metric**

Betweenness centrality measures the ability of a node to take the shortest paths [4]. Nodes that appear on several shortest paths possess higher betweenness centrality. This metric reflects the impact that a node has on others in the network and also shows its importance for information transfer [3], [12], [13], [14]. We denote the betweenness centrality measure of a node *v* by $B(v)$, given by:

$$B(v) = \sum_{(p,t)\in V_v} \frac{\sigma_{pt}(v)}{\sigma_{pt}} \quad (12)$$

where $\sigma_{pt}(v)$ enumerates the shortest paths from node p to node t passing through v, and $\sigma_{pt}(v)$ enumerates shortest paths from p to t. Furthermore, V_v is given by:

$$V_v = \{(p,t) \in V^2 : p \neq v \neq t, p \neq t\}. \quad (13)$$

The normalized betweenness of a node *v* is given by:

$$B_N(v) = \frac{1}{(n-1)(n-2)} \sum_{(pt)\in V_v} \frac{\sigma_{pt}(v)}{\sigma_{pt}}, \quad (14)$$

It follows that nodes with high B_v have a high capability of bridging or disconnecting the network. Therefore, such nodes are crucial for robustness, integrity and continued communication in the network [5], [11].

3 Discussion of the Proposed Algorithm

Dijkstra's algorithm lies in Bellman's Principle of Optimality, and both algorithms are based on an optimization method called dynamic programming. Dijkstra's algorithm is used to compute the minimum cost or shortest path between one node and all other nodes when the vertices of the graph represent nodes and path costs are represented by path distances between pairs of nodes connected by a direct link. The proposed algorithm computes the shortest path between two nodes. It uses a graph consisting of nodes and edges. The cost of each node is calculated from the source by summing up the cost up to that node. For a given source vertex (i.e. node), the algorithm works out all paths and computes the path with the lowest cost, (i.e. the shortest distance between the source node and any other node) [15], [16].

This algorithm offers another method of computing the costs of the shortest paths from a single source node to a single destination node. The goal of both algorithms, the proposed algorithm and Dijkstra's algorithm, is to select the nodes in the shortest path problem. The main difference in the algorithms is that Dijkstra carries the overall information of the network and every node is involved, while the proposed algorithm deals only with the nearest node, and not all nodes are involved as in Djikstra's algorithm [17], [18], [19].

Figure 4 shows a flowchart that illustrates the steps performed by the proposed routing algorithm from start to end. Firstly, the algorithm identifies the source and destination nodes, denoted R1 and R2, respectively, and a cost of 0 is assigned to the source node. Subsequently, another node is selected and labeled as P (or permanent node). This node then becomes a tentative node and another node close to it is selected and labeled as P. This procedure is repeated until the destination node is reached [20].

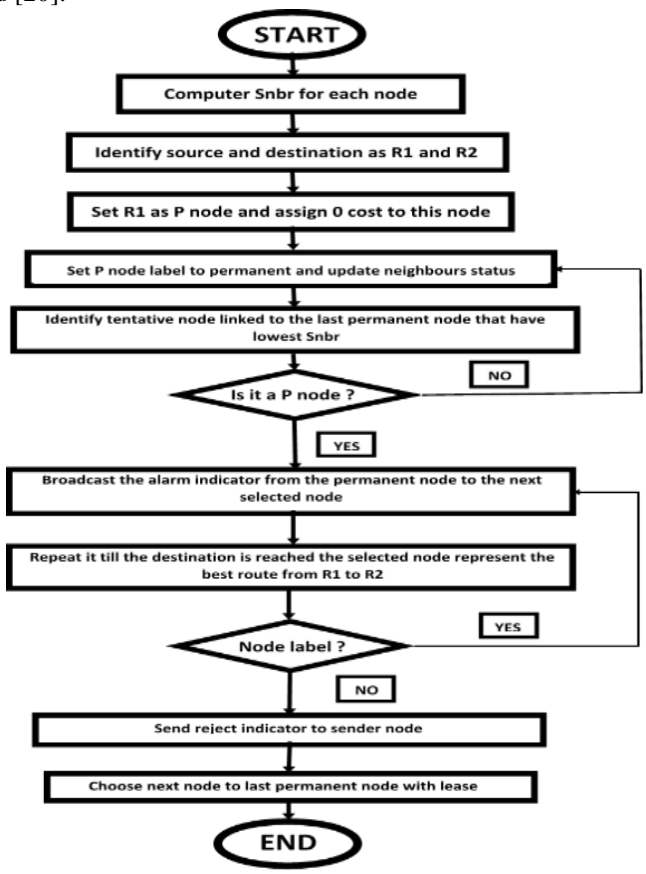


Figure 4
Flow chart of the proposed algorithm

4 Implementation and Results

4.1 Simulation

We investigate the behaviour of a network that finds a route between source and destination nodes, with particular focus on the routing algorithm. We confine the analysis to the case where some nodes fail in the network. The impact of random node losses on network performance is also measured.

A random deployment of n nodes is generated in an $80m \times 80m$ network field. The origin $(0,0)$ is located at the bottom left corner of the square. This square, shown in Figure 5, is in the first quadrant with the x and y-axes. We assume that the node collecting the information from the source node is in location $(4.0, 48.0)$. The node nearest to the sink is the node that connects the sensor network to the sink. All data must be transmitted through this node to reach the sink. Within the network, two nodes are said to be connected as long as the distance between them is smaller than the length between the communicating nodes and the BS (or sink). The state of the node is either 0 or 1. When a node is connected to another node, its state is 1. Otherwise, it is 0. Nodes are randomly distributed. In our simulation, to generate 100 different states, different random seeds were applied. Every simulation is executed 3 times for different network sizes $(i.e.\ 30\ nodes)$, in which the state samples are independent. For each network size, the centrality measures are computed. Table I shows details of the simulation setup.

Table 1
Simulation Setup

Parameter	Value
Deployment area	80m x 80m
Maximum number of nodes	100
Source node	1
Sink node	1
Number of topologies (runs) per experiment	100
The origin	(0,0)
Node distribution	Uniform
Position coordinate	(4.0, 48.0)
Number of simulation runs	3 times for each network size

4.2 Decision Making in the Network

Decision making in the network is the number of failed sensors. The network can be regarded as a collection of sensor nodes executing decisions. A decision that a sensor executes depends on its own decision at a particular time as well as the decisions of its closest neighbors. The formula for the decision of the node p_i at time k is:

$$\text{Deg } p_i = \frac{\sum_{i=1}^{N_i} d_i}{N_i} \tag{15}$$

where $Ni = deg(p_i) + 1$, d_i is node i, and 1 is added because the "neighbourhood" includes node p_i itself.

We assume that node i detects an event that it uses in executing its unilateral decision. The methods employed in making decisions that are followed by the nodes are those of decentralized data fusion systems such as those shown in [6]. A few requirements are introduced in the network. Firstly, the data should eventually reach the sink node. The distance from one node to other nodes is determined based on the Snbr, where the Snbr is used to select nodes as relays to forward data.

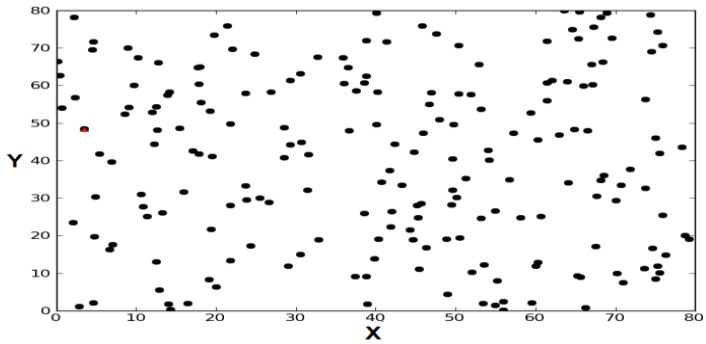


Figure 5
Random network

In the context of a network where there are human agents, it may be the decision maker who decides the next possible action for the group. In this situation, one node or point must decide to stimulate the network into action. Eventually, the decisions made by the individual nodes arrive at this central point. In the simulation, all nodes apart from the sink are displayed. In each case, the node nearest to the sink is the grey node (shown in Figure 5), which takes the role of decision-maker as well as the final link between the sensor network and the action command center.

4.3 Test for Degree, Closeness and Betweenness Metrics

A network of 30 nodes is shown in Figure 6. The gray-colored node is the one nearest to the BS. This reveals that as the sensors tend to be near each other, (observe the betweenness and closeness centrality measures in Figure 7), the sensor disruption has a negative impact on the network. Eventually, this will lead to a higher energy consumption on the entire network as the distance between nodes increases. This shows that with regards to the network's topology, the network may be tolerant to sensor failures, however, higher overall energy consumption can still be seen as the network tries to maintain network decisions.

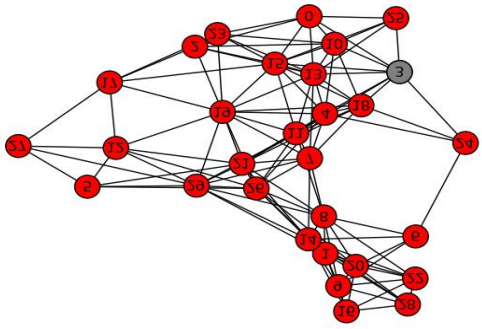


Figure 6
Network of 30 nodes

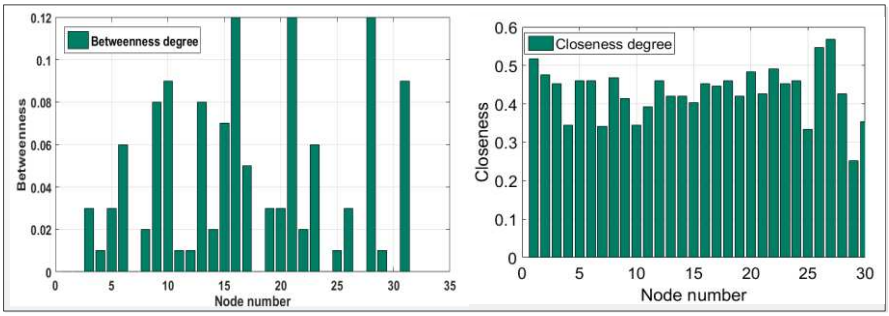


Figure 7
The impact of betweenness and closeness centrality measures of sensors on the network decisions of 30 nodes

The network illustrated in Figure 8 is tested under varying random failure situations. To simulate this, a number of nodes are excluded randomly. For instance, the network decision is still well-respected when a node is excluded. The single node omission is repeated in the network and the average network measures are determined. The goal of exclusions is to execute a network test where some of the sensors are left out. The exclusion of a node is randomly done for a number of sensors that are eliminated per execution.

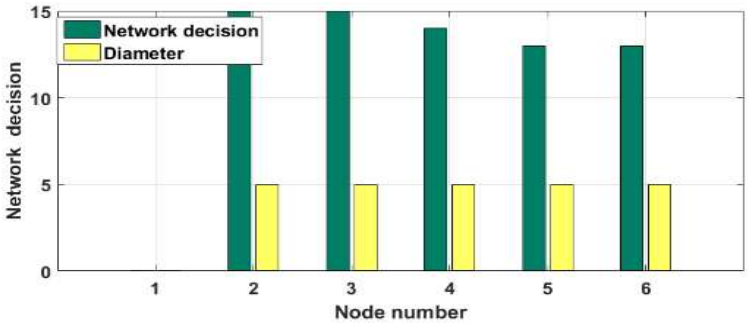


Figure 8

Network decision in terms of the percentage of failed sensor nodes for 30 nodes network test

Note that the objective is to determine the impact of random failures (or attacks) on the network and to determine the stability of the network's decision-making in the event that nodes are excluded. Finally, for every network, the impact of the exclusions is determined.

4.4 Results

This section evaluates the behaviour of the proposed algorithm in a network and presents the results. In this experiment, the proposed algorithm based on centrality measures is used to find the shortest path (i.e. path with the lowest energy consumption) in a network. The algorithm used to determine the shortest path is implemented in MATLAB.

The first step is to calculate the time it takes to find a route between source and destination nodes. The individual paths between nodes are traced until the overall path reaches the target (i.e. destination) node. In the proposed algorithm, the execution period is defined as time elapsed between event detection at the source node and data delivery to the destination node. However, the deployment of the specific node closest to the sink node has to be considered and it is because of this requirement that a multitude of routing algorithms exist. Centrality measures help reduce the time taken to select intermediate routing nodes between the source node and destination nodes, which is equivalent to minimizing the shortest distance from one node to another. The algorithm expands the first nodes connected to the intermediate node with small number (Snbr) to ensure that energy consumption and congestion are reduced when data is transmitted.

The general steps to select a forwarding node in the proposed algorithm are as follows. In the network graph representation, a path can be found between a starting point and end point in the graph. Firstly, a vertex from the graph is chosen as the starting point. The degree of an edge (i.e node) is the number of vertices connecting said edge to the adjacent edges. The current node attempts to find an adjacent node with a small number. This is repeated for every adjacent node, until

the node with the smallest number is selected, and that node becomes the current node. This procedure is repeated until the destination node is reached. Then, all paths from the beginning to the end vertex in the graph are found and lastly it is determined whether the graph is connected. If one considers the starting point (i.e. source node) to be the same for both the proposed algorithm and Dijkstra's, Dijkstra's algorithm needs to visit all nodes in the network before it is able to select a path, whilst the proposed algorithm visits only the nearest nodes when selecting the shortest path from source to destination.

A comparison between the running (i.e. execution) times of Dijsktra and the proposed algorithm is shown in Table 2. Each algorithm is executed ten times. The algorithm was also executed for 100 times and the results were found to be statistically similar to 10 times. Only the results for 10 executions are shown here because it's easier to visualize. The execution times in seconds for both algorithms are shown in Table 2, where the execution time for the proposed algorithm (PA) is 0.000118017196655 and the execution time for Dijkstra is 0.000144004821777 for the first execution. The results show that there is a difference in the execution times and the proposed algorithm is faster with regards to the average execution time.

Table 2
Running time for simulation in seconds

Execution number	**Proposed algorithm (10^{-4})**	**Dijkstra's algorithm (10^{-4})**
1	1.2	1.4
2	1.1	2.6
3	2.2	1.5
4	2.7	1.1
5	1.3	4.4
6	1.1	1.7
7	1.3	2.7
8	1.6	1.5
9	1.3	1.3
10	1.9	4.9
Average	1.6	2.3

Figure 9 shows a more clear representation of the data from Table 2 using a bar-stacked graph. It clearly shows that overall, the proposed algorithm executes faster than Dijkstra. Figures 10 shows the execution times for both algorithms over 10 executions. It is seen that the worst-case execution time for Dijkstra's algorithm is 4.67 and for the proposed algorithm it is 2.6, showing that the proposed algorithm is indeed faster.

It has also been shown that the proposed algorithm consistently chooses shorter paths between nodes than Dijsktra. This is attributed to the use of centrality measures.

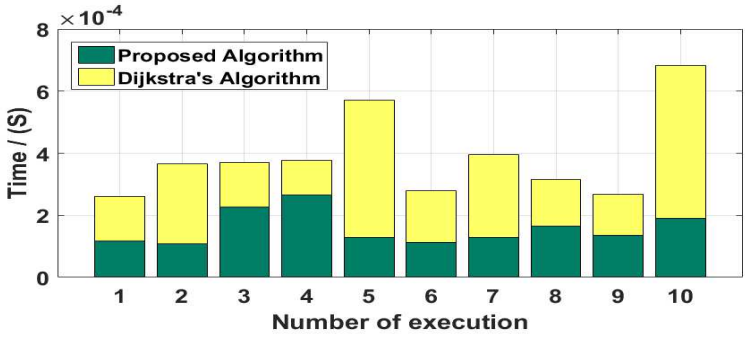


Figure 9

Time response comparison for the bar-stacked chart

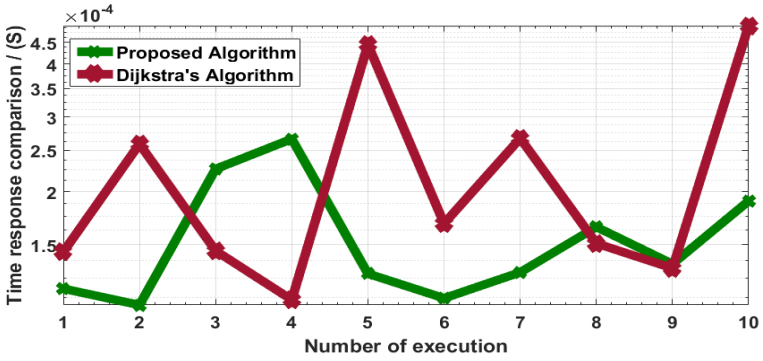


Figure10

Time response comparison for both algorithms

Conclusion

A study on routing using centrality measures has been conducted. The study enabled us to analyze how information is disseminated over a network, from source node to nearest nodes and finally to the base station via the shortest path. The network's performance can be evaluated by utilising the degree, closeness and betweenness centrality measures, as well as the shortest path carrying data from one node to another until the sink nodenis reached. We conducted an analysis on certain characteristics of sensor networks to address different challenges encountered in a network. These include reliability, failure tolerance and robustness. These three characteristics enable the network to perform better. Each of the characteristics has a role to play, especially failure tolerance. Failure tolerance is crucial if a network has to keep making network decisions that are stable even when some nodes are disconnected. Currently, centrality measures are

used for many applications in sensor networks. One of these applications is finding a routing path within a sensor network. In this study, an algorithm that finds a routing path with the shortest distance between nodes is implemented. The proposed algorithm lowers the energy consumption and thus increases the network's lifetime. When considering a routing algorithm for sensor networks, resilience against attacks or failure is very important. Using simulations, the impact of nodes on the network's performance via centrality measures has been shown. It was observed that the proposed algorithm is faster than Dijkstra's algorithm in terms of execution time. The results show that the proposed routing algorithm based on centrality measures outperforms Dijkstra's algorithm in terms of connectivity, and centrality measures (i.e. Degree, Closeness and Betweenness) can be used for routing in sensor networks, with the Betweenness centrality measure outperforming the centrality measures. The work presented in this paper can be extended by changing some simulation parameters such as the number of simulation repetitions in order to extend the results. Additionally, this work can be extended by considering other metrics such as packet delivery ratio and congestion level in the network.

References

[1] S. M. Mwika and S. W. Utete, "Sensor networks for detecting events," 2011, unpublished Essay

[2] D. J. Higham, P. Grindrod, and E. Estrada, "People who read this article also read...: Part 1," *SIAM News*, Vol. 44, No. 1, January/February 2011

[3] B. Ruhnau:, "Eigenvector-centrality a node-centrality?" *Social Networks*, Vol. 22, pp. 357-365, 2000

[4] L. Freeman, "A set of measures of centrality based on betweenness, sociometry," *Sociometry*, Vol. 40, pp. 35-41, March, 1977

[5] C. E. Shannon, A mathematical theory of Communication. Urbana: University of Illinois Press, 1948

[6] S. Grime and H. F. Durrant-Whyte, "Data fusion in decentralized sensor networks," Control Engineering Practice, Vol. 2, No. 5, pp. 849-863, October, 1994

[7] A. Jain and B. Reddy, "Node centrality in wireless sensor networks: Importance, applications and advances," in Advance Computing Conference (IACC), *2013 IEEE 3rd International.* IEEE, 2013, pp. 127-131

[8] E. J. Segovia and P. Vila, "New applications of the betweenness centrality concept to reliability driven routing," in VIII Workshop in G/MPLS Networks, 2009

[9] L. d. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, "Characterization of complex networks: A survey of measurements," *Advances in physics*, Vol. 56, No. 1, pp. 167-242, 2007

[10] C. E. Shannon, A mathematical theory of Communication. Urbana: University of Illinois Press, 1948

[11] Z. Chen and H. Qi, "A distributed and shortest-path-based algorithm for maximum cover set problem in wireless sensor networks," *IEEE International Conference on Trust, Security and Privacy in Computing and Communications,* pp. 1224-1228, November 16, 2011

[12] M. Z. Siam, M. Krunz, A. Muqattash, and S. Cui, "Adaptive multi-antenna power control in wireless networks," in Proceedings of the 2006 international conference on Wireless communications and mobile computing. ACM, 2006, pp. 875-880

[13] M. A. Othman and H. A. Sulaiman, "An analysis of least-cost routing using bellman ford and dijkstra algorithms in wireless routing network," International Journal of Advancements in Computing Technology, Vol. 5, No. 10, June, 2013

[14] L. Sitanayah, K. N. Brown, and C. J. Sreenan, "Fault-tolerant relay deployment based on lengthconstrained connectivity and rerouting centrality in wireless sensor networks," in European Conference on Wireless Sensor Networks. Springer, 2012, pp. 115-130

[15] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad-hoc networking environment," Proceedings of IEEE INFOCOM , Anchorage AK, 2001

[16] S.-C. Wang, D. S. Wei, and S.-Y. Kuo, "An spt-based topology control algorithm for wireless ad hoc networks," Computer communications, Vol. 29, No. 16, pp. 3092-3103, 2006

[17] J. Wu and M. Gao, "Stojmenovic. on calculating power-aware connected dominating sets for efficient routing in ad-hoc wireless networks," Proceedings of the 30th Annual International Conference on Parallel Processing, September, 2001

[18] M. Dramski, "A comparison between dijkstra algorithm and simplified ant colony optimization in navigation," Zeszyty Naukowe/Akademia Morska w Szczecinie, pp. 25-29, 2012

[19] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," IEEE wireless communications, Vol. 11, No. 6, pp. 6-28, 2004

[20] V. Rodoplu and T. Meng, "Minimum-energy mobile wireless networks revisited," IEEE Journal of Selected Areas in Communications, Vol. 17, No. 8, pp. 1333-1344, 1999

[21] YJ. Jang, SY. Bae, SK. Lee. "An energy-Efficient routing algorithm in wireless sensor networks". In : Kim T. et al. (eds) Future Generation

information thechnology, Lecture notes in computer Science, Springer, Berlin, Heidelberg. Vol. 7105, pp. 183-189, FGIT 2011

[22] V. Jose, Deepa and G, Sadashivappa. "A novel Energy efficient Routing algoritm for wireless sensor networks using sink mobility." Internation Journal of Wireless and Mobile Networks. No. 6, pp. 15-25, 10.5121/ ijwmn, 2014

Quality of Institutions in the European Union countries. Application of TOPSIS Based on Entropy Measure for Objective Weighting

Adam P. Balcerzak

Nicolaus Copernicus University in Toruń, Department of Economics, ul. Gagarina 13a, 87-100 Toruń, Poland, e-mail: adam.balcerzak@umk.pl

Abstract: The relation between the quality of institutions and long term growth is currently an accepted stylized fact, which has good support in empirical literature with regard to both developing and developed countries. In the case of the second group, the quality of intuitions is crucial for keeping their abilities to face quick technological changes, which is a condition for maintaining their international competitiveness. The objective of the article is to analyse the quality of institutions in the European Union countries in the years 2000-2015 from the perspective of its influence on the speed of resources reallocation. The research touches upon the question of the abilities of the countries to take advantage of the knowledge based-economy potential. The concept of the institutional quality is rooted in the transaction costs theory framework. The phenomenon under evaluation is treated as a multiple-criteria problem. Thus, modified TOPSIS method based on entropy measure for objective weighting is applied here. The research shows institutional disparities between Northern and Southern European countries and confirms significant institutional reforms implemented in the Central European economies, which resulted in an improvement of their institutional quality with regard to knowledge-based economy requirements.

Keywords: quality of institutions; European Union; entropy measure; multiple-criteria analisis; TOPSIS

1 Introduction

Development of new institutional economics, with its empirical contribution to the growth theory, has significantly influenced mainstream economics for the last twenty years. The standard assumption of exogenic character of institutions, which is typical for both exogenous and endogenous growth models, is usually not accepted any more, especially in the case of policy-making implications and objectives. It is currently commonly assumed that the quality of institutions with their ability to transform under the influence of fundamental changes of economic conditions is crucial for long-term growth. Thus, both formal and informal

institutions affecting not only legal, but also the political and social environment, have a significant influence on the well-being of societies. However, this commonly accepted idea is constantly faced with both theoretical and practical problems. From both perspectives, the fundamental dilemmas of institutional economics relate to problems of definition and measurement of institutional quality, which have not only a multivariate but often also a latent character. From the policy perspective, the variety of institutional arrangements, which are the result of historical and culture long term formation process, indicate that many institutional factors of fundamental importance are very difficult to change or copy from one society to another [1, 2, 3, 4]. Therefore, from the policy guidelines perspective, special attention should be given to formal institutional factors influencing growth incentives, which can be modified in a medium or shorter term. These institutions are in the main scope of the article.

The objective of the current article is to analyse the quality of institutions in the European Union countries in the years 2000-2015. The research concentrates on the institutional aspects that *are of crucial importance for the process of reallocation of resources in the case of relatively developed countries in the reality of knowledge-based economy, and which can be modified in a relatively short time*. *The main common characteristic of the analysed institutional aspects is their influence on the speed of reallocation of the economic resources to potential new applications*. The definition of quality of institutions is based on the postulates of the transaction cost theory. The current research is a continuation of the Author's previous studies [5], which aimed at evaluation of the progress obtained by new member states of the European Union in reforming their institutions under the pressure of global financial crisis. In that research, a modified TOPSIS method was applied, based on the assumption of equal importance of all institutional aspects which form the phenomenon under evaluation. On the other hand, in the current article not only the times-pan of the study was extended, but based on recent research the specifics of the applied aspects were redefined. Additionally, the assumption of equal importance of all institutional aspects was removed.

The paper is structured as follows: first of all, a theoretical background and literature review for the research is given. In that part, special attention was paid to previous studies on institutional growth determinants, which are related to developed countries. Then, the measurement problems with regard to the quality of institutions and their implications for the current research are presented. In the next section, a specific description of methodological approach is done, which enables to replicate its results. Finally, the empirical results and their discussion are given. The article ends up with conclusions, ideas on possible future research and applications of the obtained results.

2 Literature Review

2.1 Crucial Institutional Aspects in Developed Countries under Knowledge-based Economy

In recent years institutional researchers have concentrated on a wide spectrum of factors such as institutions of governance (which guarantee contract enforcement), formal and informal institutions that can be transformed in different time-spans, and finally, many methodological approaches applied in the field [6, 7, 8, and 9]. Based on these directions of studies, it can be stated that current institutional research must be based on precise selection of institutional aspects, and the scope of the research must be narrowed to provide both theoretical results and policy implications.

The objective of this section is to provide theoretical framework and empirical background for selection of institutional aspects that are crucial for developed countries with regard to supporting their growth potential in the reality of an opened highly competitive knowledge-based economy.

Based on the current state of the art with regard to new institutional economics, it is commonly accepted that the influence of institutions is attributed to both supply and demand side of economy. In the first case, it is mostly related to supporting growth of capital, then improving conditions for technological progress, which finally can influence labour and productivity growth. On the demand side, effective institutions can support a long-term level of investments. In the reality of knowledge-based economy, the factors influencing the speed of technological progress and technological change are of the highest importance. As a result, in the case of developed countries the crucial role should be given to all institutional factors that can affect the speed of transformation of resources to high productivity allocations. The importance of this factor is of universal character and can be attributed to economies at any level of development and any time perspective. However, in the case of developed economies in the reality of quick technological changes its role is unprecedented. Therefore, it will be the core of further analysis.

Based on the previous studies of the Author, which have been concentrated on the determinants of productivity changes in developed countries since the mid-1990s [5], there are five crucial institutional areas which can have a dominant role for developed economies in the process of reallocation of resources in the reality of knowledge-based economy:

1) Formal regulations influencing entrepreneurship;
2) Effectiveness of juridical system in keeping a low level of transaction costs and supporting the effectiveness of market mechanism;

3) Regulations affecting competitive pressure;

4) Effectiveness of labour markets;

5) Financial markets institutions as a stimulation of development of enterprises with high growth potential.

The role of entrepreneurial activity and its influence on supply of high–growth firms, which in the cumulative process contribute to the macroeconomic productivity dynamics have been the object of intensive research since the mid-1990s [10]. Within this framework, the significant role of institutional arrangements on entrepreneurial activity with regard to innovative start-ups and its macroeconomic consequences based on the Swedish case was analysed by Davidsson and Henrekson [11], where the authors argue that the basic prerequisite for the emergence of a sizable number of high growth enterprises, which in turn brings positive macroeconomic results, is that the key industries and sectors of the economy are available for entrepreneurial exploitation. Therefore, the minimum number of entry barriers and low entrepreneurial transaction costs are of basic importance. In the recent research, Bosma *et al.* [12] have tested econometrically the relations between institutions, productive entrepreneurship and economic growth for 25 European economies. Their paper confirms that in the case of the EU countries, productive entrepreneurship contributes to economic growth, where proxies for quality of institutions, financial stability, small government, and perceived start-up skills were the most important predictors of productive entrepreneurship. The strong interdependence between entrepreneurial innovation and regulatory efficiency for the European Union economic growth determined by institutional factors has also been recently confirmed by Ignatov [13]. However, these results can be expanded internationally further. Aparicio *et al*. [14] analysed panel data with 43 countries in the years 2004-2012 with regard to institutional factors that encourage opportunity entrepreneurship in order to achieve higher rates of economic growth. They showed that informal institutions may have a higher impact on opportunity entrepreneurship than formal institutions.

The relation between the effectiveness of judicial system and long-term economic growth is currently undisputed. The quality of judicial system determines protection of property rights and low transaction cost in the process of contracts enforcement. Thus, it encourages savings and investments [15]. Based on the case of Italy, Lorizio and Currieri [16] show how the quality of legal system impacts fundamental dimensions of development, such as optimal allocation of resources, which finally influences productivity. They argue that economic disincentives arising from the failure of justice result in contraction of start-ups, the preference for keeping small-size enterprises, which affects negatively the competitiveness of the national production system and its low productivity, finally a lower level of investments and the phenomenon of freezing capital.

The next institutional aspect under consideration relates to internal competitive pressure in the economy, which also further affects its international

competitiveness. The internal competitive intensity is highly dependent on micro- and macroeconomic regulations influencing transaction costs and barriers to entry. Both sectorial and macroeconomic comparative research, which have been done since the mid-1990s, have confirmed that competitive pressure affected by national regulations can significantly form incentives to invest in new, more productive technological and organizational solutions. Thus, it affects both labour productivity and the total factor productivity growth [17].

Many research studies confirm that regulations affecting competitive pressure in the economy are often related to regulations influencing effectiveness of labour markets in the reallocation of human capital. Gust and Marquez [18] analyzed a panel of 13 industrial countries in the 1990s to find the sources of divergence in productivity growth. These authors argue that the burdensome regulatory environment and in particular regulations affecting labour markets were responsible for lower productivity growth in many industrial countries. Kuder [19] analysed the role of institutional factors in determining the growth rate of the United States economy in the years 1979-2001, where the main objective of the research was pointing out the institutions which were affecting growth to the highest extent. Under the influence of Amble [20], she concentrated on the following institutional areas: labour market, financial market, education and R&D. Her results indicated the most significant role of transformation of labour market arrangements, with special attention to institutional factors affecting the employment cost and the level of unionization in influencing long-term growth in the United States. The role of job protection regulations on productivity growth was also confirmed by Bassanini *et al.* [21]. They confirmed negative effects of mandatory dismissal regulations on productivity growth in the industries where layoff restrictions were more likely to be binding.

Finally, the role of financial system and its effectiveness in influencing growth has been one of the most explored research areas for last three decades [22, 23, and 24]. The direct relationship between financial development and productivity growth is related to efficient capital reallocation process that should enable to shift quickly capital from declining industries to those with high growth perspectives. Arizala *et al.* [25], in one of the bigger studies on 77 countries with data for 26 manufacturing industries in the years 1963-2002 confirmed a significant relationship between financial development and industry level TFP growth.

The next section is devoted to the problems and dilemmas of operationalizing and measuring framework for five institutional aspects described above.

2.2 Quality of Institutions and its Measurement

The concept of quality of institutions is currently considered as the core of institutional economics. Though, it is not easy to define and operationalize this term. Also, a vigorous discussion on the problem of distinction between *policies*

and *intuitions* shows fundamental problems of institutional theory [26, 27]. In this context, it is very hard to question Voigt's statement that "only if institutions can be measured with a minimum degree of confidence are empirical statements such as 'institutions matter for y' credible" [27].

In order to define institutions, most of the new institutional economists accept the North's proposal, where "Institutions are humanly devised constraints that structure human interaction. They are made up of formal constraints (e.g., rules, laws, constitutions), informal constraints (e.g., norms of behaviour, convention, self-imposed codes of conduct), and their enforcement characteristics. Together they define the incentive structure of societies and specifically economies." [28].

The current paper concentrates on the quality of institutions that influence the speed of resources reallocation that are crucial for taking advantage of the global knowledge-based economy. The definition of quality itself is based on the framework of transaction costs theory in complex economies, which is currently considered as one of the core concepts in new institutional economics [see 1; 29]. To remind shortly the definition of transaction costs, with regard to economic consequence of technological changes the transaction costs are not only the costs of coordinating the increasingly complex interdependent parts of an economy, but they are also the costs of enforcing agreements and making credible commitments across time and space, necessary to realize the potential of technology [29].

From the empirical perspective measuring quality of institutions, often also within transaction cost theory framework, is based on application of approximate measures, which are often based on the feelings of a given economic entity or group of economic actors. This naturally gives a high level of subjectivity to the process and creates features of institutions difficult to quantify [19]. It makes the issue of quality of institutions a multiple-criteria analysis problem, and in that way it is treated in the current article.

Practically in the case of international comparative empirical research, especially for application of quantitative methods, authors usually use indexes of institutional quality, which are based on the databases provided by international institutions and organizations such as World Bank, Fraser Institute, Heritage Foundation or World Economic Form [9; 30]. More specific discussion in this regard is given in Section 3.3.

Finally, before moving to the presentation of applied research methodology, it must be stressed that the empirical approach presented in this article, with the specific definition of the institutional aspects, is aimed at fulfilling the first institutions measurement objective given by Voigt [27], who argues that "measures of institutions should refer to specific institutions because aggregate measures such as 'the rule of law' are too bread and fuzzy to contain meaningful information". This objective resulted in application of two stage analytical levels in the current research and choice of modified TOPSIS method.

3 Research Methodology

The objective of the current research is to provide multiple criteria analysis of the quality of institutions, where the assumption is made that not all the specific institutional factors have equal importance for the effectiveness of the whole system. In that case, two approaches are possible. First of all, application of weights based on the knowledge and previous expertise of the researcher is possible. The basic criticism of that approach relates to the consequences of arbitral choices and preferences of the researchers, which can result in possibilities of results manipulation. Then, one can apply formal statistical objective criteria, which are based on statistical characteristics of the variables used in the research. In the current research, the second option was applied. As a result, TOPSIS method based on entropy measure for objective weighting was applied in the current research.

In the next subsection of the paper specific information on data preparation and formal information criteria for variables selection was presented. Next, the applied method and data were described in details.

3.1 Data Preparation and Formal Criteria for Variables Selection

The first stage of every multiple-criteria analysis relates to selection of set of object O_i and the economic phenomena under consideration. Next, a subset of specific aspects (Y_1, Y_2,...,Y_l) for given economic phenomenon and a set of attributes (diagnostic variables) X_1^s, X_2^s,..., X_k^s for every aspect Y_s is determined. After selection of attributes X_j^s their character must be specified. The diagnostic variables are categorized as benefit variables or negative variables. For benefit variables X_j^s for every two values $x_{i,j}^s$, $x_{k,j}^s$ that refer to objects O_i, O_k, the relation $x_{i,j}^s > x_{k,j}^s \rightarrow O_i \succ O_k$ is fulfilled, where $\succ$ means that object O_i is preferred to O_k. Thus, a maximum value of variable is preferred. For negative variable X_j^s for every two values $x_{i,j}^s$, $x_{k,j}^s$ that refer to objects O_i, O_k the relation $x_{i,j}^s < x_{k,j}^s \rightarrow O_i \prec O_k$ is fulfilled, where $\prec$ means that object O_k is preferred to object O_i. Thus, a minimum value of variable is preferred [5].

After preliminary selection of the diagnostic variables, which is based on the content related criteria and experience of a researcher, the second step of multiple-criteria analysis concentrates on a formal verification of variables information value based on formal/statistical criteria. For this purpose, taxonomic criteria of

information value can be used [31]. The final diagnostic variables should fulfil three criteria, of which the first two ones are the most commonly applied [32; 33]:

a) High level of variation – the final diagnostic variables should not be similar to each other in the sense of information on the objects, so the diagnostic variables should obtain at least a minimum level of variation. To evaluate the level of variation, coefficient of variation can be applied, which is given with Formula 1.

$$V_j = \frac{s_j}{\overline{x_j}} \qquad i = 1,2,...n, \quad (1)$$

Where s_j is the standard deviation of j variable; $\overline{x_j}$ is arithmetic mean of j variable. The variables which do not fulfil a formal criterion, for example such as the minimum accepted value of $V_j = 0{,}1$, are eliminated.

b) Level of correlation – the diagnostic variables should not be too highly correlated. High correlation of the variables leads to overlapping of information, thus some of the variables are redundant. For highly correlated diagnostic variables a Hellwig's parametric method can be applied, where the maximum acceptable value of correlation coefficient for potential variables can be set as $r = 0{,}85$. This value can be set with application of formal criteria or is usually given at the researcher's discretion [31, 33].

c) High information value – the final diagnostic variables should reach high values with difficulty. To evaluate the information values of the variables the skewness coefficient can be used. In the case of benefit variables, the distribution of the variable should be right-skewed. When it is left-skewed, it means that most of the objects relatively easily reach high values of the variable, thus it does not differentiate the objects significantly.

In the final step, the selected diagnostic variables should be normalised, and the negative variables should be transferred into benefit variables. As common practice in this process, the unitarisation method or standardization are applied [34, 35, 36]. As all the final diagnostic variables applied in the research were classified as benefit variables, a classic standardization procedure given with Formula 2 was used here. It allowed to obtain variables characterized with mean at the level 0 and variance that is equal to 1.

$$x_{ijt} := \frac{x_{ijt} - \overline{x_{jt}}}{s_{jt}} \qquad i = 1,2,...n,\ j = 1,2,...,\ p,\ t = 1,2,..l \qquad (2)$$

where $\overline{x_{jt}}$ and s_{jt} are given with formulas 3.

$$\overline{x_{jt}} = \frac{1}{n}\sum_{i=1}^{n} x_{ijt}\ , \qquad s_{jt} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ijt} - \overline{x_{ijt}})^2}$$
$$i = 1,2,...n\ ,\ j = 1,2,...,\ p\ ,t = 1,2,..l \tag{3}$$

3.2 TOPSIS Based on Entropy Measure for Objective Weighting

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is a tool which is commonly applied for solving multiple criteria decision making problems (MCDM) [37, 38]. The method is based on the concept of ordering alternatives by similarity to an ideal solution. The TOPSIS methodology can be also effectively applied for description and evaluation of economic objects under consideration, not as it is in the case of MCDM for the procedure of choosing the best alternative [5; 39, 40, 41].

After selection of a final set of diagnostic variables, which was described in the previous subsection, a synthetic measure of development of the phenomenon under research can be assessed. In the case of TOPSIS method it is defined as the similarity or relative closeness to the positive ideal solution. The synthetic measure enables to assess the development level of phenomenon for every object. The method allows assessing and comparing the current situation of the objects.

In the research, the modified TOPSIS method is proposed. This approach can be especially useful when there is a possibility to extract complex economic aspects which refer to a specific feature of the object. In the case of every aspect Y_s for variables X_{jt}^{s} the positive ideal solution $I_{s,jt}^{P}$ and negative ideal solution $I_{s,jt}^{N}$ are specified with equations 3:

$$\boldsymbol{I}_{s,jt}^{P} = \max_{i} x_{ijt} \quad , \quad \boldsymbol{I}_{s,jt}^{N} = \min_{i} x_{ijt} \tag{3}$$

In the case of every distinct aspect Y_s separation measures $D_{s,it}^{P}$ from the positive ideal solution and $D_{s,it}^{N}$ from negative ideal solution for every object O_i are calculated with equations 4:

$$\boldsymbol{D}_{s,it}^{P} = \sqrt{\sum_{j=1}^{n}(x_{ijt} - I_{s,jt}^{P})^2} \quad \boldsymbol{D}_{s,it}^{N} = \sqrt{\sum_{j=1}^{n}(x_{ijt} - I_{s,jt}^{N})^2} \tag{4}$$

The value of synthetic sub-index R_{it}^{s} that describes every chosen aspect Y_s for every O_i is obtained by combining the proximity to the positive ideal solution and the remoteness from the negative ideal solution. The sub-index is given with the equation (5).

$$R_{it}^{S} = 1 - \frac{D_{s,it}^{P}}{D_{s,it}^{P} + D_{s,it}^{N}} \quad (5)$$

Relative closeness to the positive ideal solution is a normalized measure usually on scale of 0-1, where its higher values indicate a higher level of development of the aspect Y_s.

Final synthetic index R_{it} is calculated as the weighted arithmetic mean given with the equation (6).

$$R_{it} = \sum_{s=1}^{k} w_s R_{it}^{s} \quad (6)$$

where w_s means weights for every aspect Y_s and the sum of weights equals to 1.

In order to avoid the problem of arbitral choice of the values of weights or simplified assumption of equal weights, the idea of objective weights, which are based on the entropy weights, can be used [42, 43, 44]. In that case, the definition of Shannon entropy, which was developed and implemented in the information theory, is applied [45].

To assess a set of entropy weights w_{st} for variables R_{it}^{s} based on the entropy value in the first step the entropy value e_{st} is calculated (7) [42].

$$e_{st} = -\frac{1}{\ln(k)} \sum_{i=1}^{n} p_{ist} \ln(p_{ist}) \quad (7); \text{ where: } p_{ist} = \frac{R_{it}^{s}}{\sum_{i=1}^{n} R_{it}^{s}}$$

Finally, entropy weights w_{st} with the equation 8 are obtained.

$$w_{st} = \frac{1 - e_{st}}{n - \sum_{s=1}^{k} e_{st}} \quad (8)$$

where the sum of entropy weights w_{st} is equal to 1.

The assessing of the synthetic measure of development with TOPSIS method enables not only to order the objects, but it also allows to divide the objects into homogenous subsets (classes), which significantly simplifies the interpretation of the obtained results. For this purpose, the method of natural breaks (Jenks method) can be used. The natural breaks method is based on the idea of minimization of variance for objects from the chosen subsets and maximization of variance between the subsets [46, 5].

3.3 Empirical Data

In order to choose the data bases, the following objectives were formulated: a) international comparative coverage; b) long-term data availability and its comparability in time, c) processing of the data with one methodology, d) finally, international recognition of the quality of the obtained data and the mythology of its processing. Based on these criteria, the data from detailed base provided by Fraser Institute, which is used for publications of the Economic Freedom of the World Annual Reports, was chosen [47].

Based on the first condition, the possibility of application of country-specific data from national sources has been excluded. Beside Fraser Institute database, there are also other international bases that fulfill the first, third and fourth criterion, such as Doing Business database prepared by the World Bank, Heritage Foundation data base for Index of Economic Freedom or World Economic Form data base for Global Competitiveness Report. All of them cover shorter timespans than the Fraser Institute database. Additionally, the big advantage of the Economic Freedom of the World data base is its breadth of its coverage and the multidimensionality of its definition.

4 Results

The current research was done for 25 European Union countries in the following years: 2000, 2005, 2010, 2015. Cyprus, Luxemburg and Malta were excluded from the research due to the specifics of their economies and lack of empirical data for some institutional variables. The period 2000-2015 was restricted by the availability of data. The five-year sub-periods were considered as long enough for implementation of formal changes, which can affect institutional variables. The institutional aspects with a sub-set of preliminary variables are given in Table 1.

Table 1
The vector of attributes and preliminary diagnostic variables

Y_1 **– formal regulations influencing entrepreneurship**
X_1^1 – Administrative requirements for entrepreneurs
X_2^1 – Bureaucracy costs
X_3^1 – The cost of starting business
X_4^1 – Extra payments/bribes/favouritism
X_5^1 – Licensing restrictions
X_6^1 – Tax compliance

Y_2 – **effectiveness of juridical system in keeping**
X_1^2 – Judicial independence
X_2^2 – Impartial courts
X_3^2 – Protection of property rights
X_4^2 – Integrity of the legal system
X_5^2 – Legal enforcement of contracts
X_6^2 – Regulatory restrictions on the sale of real property
X_7^2 – Reliability of police
X_8^2 – Business costs of crime
Y_3 – **regulations affecting competitive pressure**
X_1^3 – Mean tariff rate
X_2^3 – Standard deviation of tariff rates
X_3^3 – Non-tariff trade barriers
X_4^3 – Compliance costs of importing and exporting
X_5^3 – Foreign ownership/investment restrictions
X_6^3 – Capital controls
Y_4 – **effectiveness of labour markets**
X_1^4 – Hiring regulations and minimum wage
X_2^4 – Hiring and firing regulations
X_3^4 – Centralized collective bargaining
X_5^4 – Hours regulations
X_6^4 – Mandated cost of worker dismissal
Y_5 – **financial markets institutions as a stimulator**
X_1^5 – Ownership of banks
X_2^5 – Private sector credit

At the first stage, completeness of data for the preliminary variables was verified. The data on *Reliability of police* and *Business costs of crime* are not available for the years 2000-2004. The data on variable *Licensing restrictions* are not available in the years 2000-2003. The three variables were removed from the study.

Table 2 presents the result for the quality information criteria analysis, which is based on the required minimum level of variable variation. The criterion of minimum value of $V_{jt} = 0.1$ was applied for all the years 2000-2015. Based on the obtained results, X_3^1, X_1^3, X_2^3 were removed from the set of diagnostic variables. The value of V_j for X_3^3 in the year 2013 was close to 0.09; for X_4^3 in the years 2000-2003 and 2010-2013 it was higher than 0,07, and for X_2^5 in the years 2007-

2009 it was higher than 0,08. As these variables do not reach the minimum criteria only in the case of the chosen years, and additionally the given values of V_{jt} can be considered as close to the benchmark value, the variables were left in set of diagnostic variables, which were analyzed with correlation criteria.

The conducted analysis of variables variation provided interesting information with regard to X_3^1. In the analyzed period, there is a tendency to unify and reduce the cost of starting business in the EU countries. In the year 2000 the average value for this variable was equal to 8,63, in the year 2005 it was 9.03, in 2010 it was 9.46, and finally in the year 2015 it was 9,62, where 10 is the highest possible value of the variable. In the years 2000-2003 the value of V_{jt} was higher than 0,09, then in the years 2004 to 2015 it was reduced from the level 0,06 to about 0,02. These factors confirm a common understanding of negative consequences of higher transaction costs with regard to the cost of starting business.

Table 2 provides specific information on variables that do not fulfil the conditions for minimum level of variation.

Table 2

Variables that do not fulfil the conditions for minimum level of variation

Variable	The years where V<0.1
X_3^1 – The cost of starting business	2000-2015
X_3^3 – Non-tariff trade barriers	2013
X_1^3 – Mean tariff rate	2000-2001, 2004-2015
X_2^3 – Standard deviation of tariff rates	2005-2015
X_4^3 – Compliance costs of importing and exporting	2000-2003, 2010-2015
X_2^5 – Private sector credit	2007-2009

Source: own estimation based on data base from Fraser Institute

At the next stage, the analysis of correlation between variables was conducted. After application of Hellwig's parametric method, three variables were removed from the final set of diagnostic variables X_4^1, X_1^2 and X_2^2.

The current Fraser Institute data base is created for 159 countries. From that perspective, the European Union countries can be considered as a relatively homogenous group of developed countries, which tend to obtain relatively good results in many aspects. It is also actual with regard to quality of institutional factors. It is natural that due to the methodology of data base preparation, in the case of many aspects the condition of reaching right-skewed distribution for befit variables is not fulfilled. Due to the dominance of the merit criteria over formal statistical requirements this condition is not applied in the current research. The final set of diagnostic variables applied for the research is given in Table 3.

Table 3

The vector of attributes and final diagnostic variables applied for TOPSIS method

Y_1 – **formal regulations influencing entrepreneurship** $X_1^1,\ X_2^1,\ X_6^1$
Y_2 – **effectiveness of juridical system in keeping low level of transaction costs and supporting effectiveness of market mechanism** $X_3^2,\ X_4^2,\ X_5^2,\ X_6^2$
Y_3 – **regulations affecting competitive pressure** $X_3^3,\ X_4^3,\ X_5^3,\ X_6^3$
Y_4 – **effectiveness of labour markets** $X_1^4,\ X_2^4,\ X_3^4,\ X_5^4,\ X_6^4$
Y_5 – **financial markets institutions as a stimulator of development of enterprises with high growth potential** $X_1^5,\ X_2^5$

Source: own estimation based on data base from Fraser Institute

Table 4

The waights for institutional aspects in the years 2000, 2005, 2010, 2015

Institutional Aspect	**2000**	**2005**	**2010**	**2015**
Y_1	9.54%	23.15%	8.70%	18.96%
Y_2	22.73%	18.42%	19.34%	30.51%
Y_3	24.47%	26.70%	26.29%	14.31%
Y_4	29.56%	25.09%	25.23%	27.49%
Y_5	13.70%	6.63%	20.45%	8.73%

Source: own estimation based on data base from Fraser Institute

The values of weights obtained with application of entropy measure for objective weighting are given in Table 4. The average values of obtained weights for all the attributes in the analysed period indicate that there are two subsets here. The highest average value for weights for all four years was assessed for effectiveness of labour markets 26,84%, then the quality of juridical system and regulations affecting competitive pressure obtained 22,75% and 22,94%, whereas the regulations on entrepreneurship and financial markets obtained 15,09% and 12.38% respectively. That obtained outcome is to a high extent consistent with the previous studies on the importance of institutional factors in forming growth incentives, which has been also already given in the literature review.

The final results with ranking and grouping of the countries into four classes based on natural breaks method are given in Table 5 and Figure 1. In the first class, the countries with the highest value of institutional quality synthetic measure are placed, whereas in the fourth one the economies with its lowest value are found. The obtained ranking and grouping confirm the long-term structural institutional divergence in the European Union, where the historically determined differences between Eastern and Western European countries are not so valid anymore. On the other hand, the research confirms long term institutional disparities between Northern and Southern European economies, which also significantly contributed to economic instability of the European Union during last global financial crisis.

Table 5
Ranking and grouping of the EU countries based on quality of institutions

	2000			2005			2010			2015		
Countries	R_{it}	Rank	Class	R_{it}	Rank	Class	R_{it}	Rank	Class	R_{it}	Rank	Class
Denmark	0.7759	2	1	0.7657	1	1	0.7600	1	1	0.7441	1	1
Ireland	0.6335	8	2	0.6773	4	1	0.6170	11	2	0.7300	2	1
United Kingdom	0.7930	1	1	0.7105	2	1	0.6859	4	1	0.6988	3	1
Netherlands	0.6900	5	2	0.6130	9	2	0.6532	6	2	0.6945	4	1
Estonia	0.6522	6	2	0.6545	5	2	0.7160	2	1	0.6901	5	1
Finland	0.7065	3	1	0.6940	3	1	0.7091	3	1	0.6766	6	2
Sweden	0.6404	7	2	0.6195	7	2	0.6600	5	2	0.6718	7	2
Latvia	0.5982	13	2	0.6026	10	2	0.6373	8	2	0.6668	8	2
Austria	0.6939	4	1	0.6247	6	2	0.6421	7	2	0.6420	9	2
Lithuania	0.5504	16	3	0.5562	12	3	0.5223	20	3	0.6403	10	2
Germany	0.6004	12	2	0.5478	13	3	0.5611	16	3	0.6213	11	2
Czech Rep.	0.5345	19	3	0.4721	20	4	0.5730	15	3	0.6112	12	3
Belgium	0.6328	9	2	0.5829	11	2	0.6279	9	2	0.5961	13	3
Romania	0.5399	17	3	0.5361	15	3	0.5791	14	3	0.5879	14	3
France	0.6295	10	2	0.5468	14	3	0.5950	12	2	0.5826	15	3
Spain	0.6034	11	2	0.5321	16	3	0.5254	19	3	0.5825	16	3
Portugal	0.5347	18	3	0.4528	21	4	0.4783	23	4	0.5781	17	3
Poland	0.4256	25	4	0.4309	24	4	0.4978	22	3	0.5714	18	3
Hungary	0.5870	14	2	0.5266	17	3	0.6203	10	2	0.5679	19	3
Slovak Rep	0.5088	21	3	0.6180	8	2	0.5832	13	3	0.5579	20	3
Slovenia	0.4798	23	4	0.4488	22	4	0.4551	24	4	0.5159	21	4
Bulgaria	0.5220	20	3	0.4180	25	4	0.5595	17	3	0.5132	22	4
Croatia	0.4418	24	4	0.4722	19	4	0.5261	18	3	0.5122	23	4
Italy	0.5591	15	3	0.4747	18	4	0.5206	21	3	0.5119	24	4
Greece	0.4925	22	3	0.4341	23	4	0.3834	25	4	0.4729	25	4

Source: own estimation based on data base from Fraser Institute

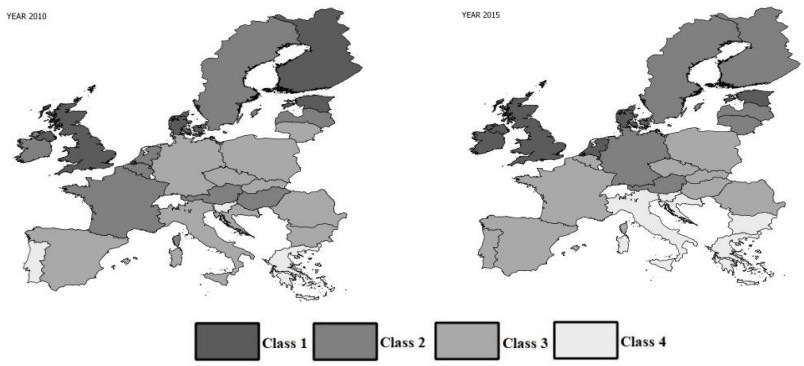


Figure 1
Grouping of the EU countries based on quality of institutions

Source: own estimation based on data base from Fraser Institute

So, throughout the whole period under research Scandinavian economies with Ireland and the United Kingdom are the leaders. Then, there is the group of the Central European member states, which can be compared with regard to obtained relative value of measure of institutional quality to such economic powers as Germany and France. This can be especially seen in the last year of the study. Finally, with the exception of Romania, which obtains relatively good results, we have the "latest new member" states such as Bulgaria, Croatia and, finally, Southern European countries, such as Italy and Greece.

5 Discussion

The current research is consistent to a high extent with previous studies of the Author [5], where the assumption on equal weights with regard to institutional aspects was taken, and contributions of other researchers in the field, which were devoted to multiple-criteria analysis of institutional factors in the EU countries. But it must be stressed that most of these other studies were concentrated on more widely defined institutional factors not always directly relating to the knowledge-based economy and the level of international competitiveness [48, 49, 50, 30, 33]. That result can be interpreted in favour of the new institutional economics postulates, in spite of the variety of approaches to defining institutions and variety of methodologies for their quantifying there is still one big picture with a common message, where the economic success is directly related not only to good macroeconomic policies, but also effective reforming of institutional economic order. What is important here, this factor cannot be only attributed to developing countries, but it is also actual in regard to developed economies that currently must face the challenges of the global knowledge-based economy.

From the policy guidelines perspective, when one looks in detail at the obtained rankings, especially interesting results can be seen in the case of the Baltic countries and other Central European economies, such as the Czech Republic, Hungary and Poland. During the 15 years under investigation, the region has significantly improved its relative institutional quality. In that group, the case of Estonia is especially interesting, as in the year 2010 and 2015 the country could be found in the first class of economies with the highest value of synthetic measure of institutional quality. What is more, in the last year of the research it was even rated higher than Finland and Sweden. The improvements of relative positions were also obtained by Latvia and Lithuania. The success of Estonia is often given as a case of small and opened economy that on the one hand, is easy to reform due to its size, but on the other hand, from the perspective of informal institutions is quite close to Scandinavian economies that additionally supports the reforming process. However, an interesting example of institutional changes for a relatively big economy can be seen in the case of Poland, which started with the lowest value of synthetic measure in the year 2000. Since the year 2010, the country has been placed in the third class, with the value of institutional quality measure close to such countries as France, Spain and Portugal. Also, relatively high positions of Romania should be stressed here.

On the other hand, one can point to two negative benchmark countries: Greece and Italy, where the case of the former one is especially significant. In the year 2000, Italy was rated in the third class at 15th position in the ranking. During the next years the relative position of the country was worse, which indicates that other economies have been much more effective at improving their institutional systems with regard to the knowledge-based economy requirements.

Conclusions

The objective of the article was to conduct the research on the quality of institutions in the context of knowledge-based economy requirements, where the phenomenon was considered as a multiple-criteria subject of the analysis. In the study the common assumption on equal importance of all institutional aspects was abandoned and the modified TOPSIS method based on entropy measure for objective weighting was applied.

From a policy perspective the current research confirms the fundamental problems of the European Union in regard to institutional disparities between Northern and Southern European countries, which have threatened the stability of the European economy during last global financial crisis. On the other hand, from the perspective of Central Europe it shows the cases of good institutional reforms in the Baltic countries with Estonia as a leader of the group, which can be considered as a positive benchmark for the whole region.

The current study is a novel approach, which contributes to institutional economics and can be used in research in growth determinants in the European Union economies. Its most important value added to the field is related to:

1) Theoretical contribution – in recent years, most of the research related to institutional growth determinants have concentrated on the developing economies, where a relatively wide range of formal and informal institutions are analysed. On the other hand, current research concentrates on relatively developed countries and institutions that play a crucial role in the process of quick and effective reallocation of resources under conditions of the knowledge-based economy. For this purpose, a multivariate measure of the quality of institutions for developed countries was proposed.
2) Methodological approach – in the research, an application of multiple-criteria analysis approach is proposed, which is currently considered as an effective method with regard to measurement and description of phenomena which are multivariate, latent and difficult to measure. In international comparative research, in order to keep objectivity of the research, a standard assumption tends to be made that all the factors (variables) taken into consideration have equal importance and influence equally the phenomenon under investigation. In the case of the current research this assumption was abandoned. In order to avoid arbitral decisions in the evaluation process, objective weights were applied based on entropy measure for objective weighting. If it is not the first application of the method in institutional economics comparative research, it can be stated that the methodology is not applied as a standard quantitative tools in the field.
3) Practical value – the current research has a great importance for policy makers in the European Union countries in general, but it is of special value for decision makers in Central European economies, who face the problem of closing the development gap and building conditions, which will support these economies in the process of avoiding the middle income trap. Based on the modernisation experiences of the countries that avoided the middle income trap, the role of an institutional system is of the highest importance in this process.

The main limitation of the current research relates to the selection of diagnostic variables and data availability. In the case of every multiple-criteria analysis, the first step of the research – the decomposition of the phenomenon on given aspects and the selection of preliminary variables is always based on an arbitral expertise decision of a researcher. Then, in the case of longer-term research, quantitative analysis is directly restricted by the availability of data and their quality with regard to description of a given phenomenon.

The current article is the first step of the research which concentrates on the role of institutional factors in supporting long-term growth in Central European countries and their significance in avoiding the middle income trap. As a result, in the next step the research will be done for all the years of 2000-2015 with different assumptions with regard to weights for given institutional aspects, which will enable to obtain time series applicable for econometric analysis. Then, the time series will be used in the research on the institutional growth determinants in

Central European economies and their possible contribution to avoid the middle income trap.

References

[1] Williamson, O. E. (1985) *The Economic Institutions of Capitalism. Firms, Markets, Relational Contracting*. Macmilllan

[2] North, D. C. (1990) *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press

[3] Rodrik, D (2008) *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*. Princeton, NJ: Princeton University Press

[4] Rodrik, D, McMillan, M, & Sepulveda, C (Eds) (2017) *Structural Change, Fundamentals, and Growth: A Framework and Case Studies*. Washington, DC: International Food Policy Research Institute

[5] Balcerzak, A. P., & Pietrzak, M. B. (2016) Quality of Institutions for Knowledge-based Economy within New Institutional Economics Framework. Multiple Criteria Decision Analysis for European Countries in the Years 2000–2013. *Economics & Sociology*, 9(4), pp. 66-81, doi: 10.14254/2071-789X.2016/9-4/4

[6] Acemoglu, D., Johnson, S., & Robinson, J. A. (2001) The Colonial Origins Of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5), 1369-1401

[7] Acemoglu, D., Johnson, S., & Robinson, J. A. (2002) Reversal Of Fortune: Geography And Institutions In The Making Of The Modern World Income Distribution. *Quarterly Journal of Economics*, 107(4), pp. 1231-1294

[8] Rodrik, D. (1999) Where Did All The Growth Go? *Journal of Economic Growth*, 4(4), pp. 385-412

[9] Carmignani, F. (2009) The Distributive Effects of Institutional Quality When Government Stability is Endogenous. *European Journal of Political Economy*, 25(4) 409-421, doi: 10.1016/j.ejpoleco.2009.03.003

[10] OECD (2001) *The New Economy. Beyond the Hype*. Paris: OECD

[11] Davidsson, P. & Henrekson, M. (2002) Determinants of the Prevalence of Start-ups and High-Growth Firms Magnus Henrekson. *Small Business Economics*, 19, pp. 81-104

[12] Bosma, N., Content, J., Sanders, M., &, Stam, E. (2018) Institutions, entrepreneurship, and economic growth in Europe. *Small Business Economics*, 51(2) pp. 483-499, doi: 10.1007/s11187-018-0012-x

[13] Ignatov, A. (2018) Entrepreneurial Innovation: the European Union Perspective. *Review of Economic Perspectives*, 18(2), pp. 137-154, doi: 10.2478/revecp-2018-0008

[14] Aparicio, S., Urbano, D., & Audretsch, D. (2016) Institutional Factors, Opportunity Entrepreneurship and Economic Growth: Panel Data Evidence. *Technological Forecasting and Social Change*, 102, pp. 45-61, doi: 10.1016/j.techfore.2015.04.006

[15] Cusatelli, C., & Giacalone, C. (2014) Evaluation Indices of the Judicial System and ICT Developments in Civil Procedure. *Proceedia Economics and Finance*, 17, pp. 113-120, doi: 10.1016/S2212-5671(14)00885-5

[16] Lorizio, M., & Gurrieri, A. R. (2014) Efficiency of Justice and Economic Systems. *Proceedia Economics and Finance*, 17, pp. 104-112, doi: 10.1016/S2212-5671(14)00885-3

[17] Bassanini, A., & Scarpetta, S. (2002) Growth, Technological Change, and ICT Diffusion: Recent Evidence from OECD Countries. *Oxford Review of Economic Policy*, 18(3), pp. 324-344, doi: 10.1093/oxrep/18.3.324

[18] Gust, C., & Marquez, J. (2002) International Comparisons of Productivity Growth: The Role of Information Technology and Regulatory Practices. *International Finance Discussion Papers, Board of Governors of the Federal Reserve Systems*, No. 727, doi: 10.2139/ssrn.314200

[19] Kuder, D. (2015) Impact of Institutional Factors on Economic Growth in the United States in the Years 1979–2007, *Oeconomia Copernicana*, 6(1) pp. 137-159, doi: 10.12775/OeC. 2015.008

[20] Amable, B. (2003) *The Diversity of Modern Capitalism.* Oxford–New York: Oxford University Press

[21] Bassanini, A., Nunziata, L., & Danielle Venn (2009) Job Protection Legislation and Productivity Growth in OECD Countries. *Economic Policy*, 24(58) pp. 349-402, doi: 10.1111/j.1468-0327.2009.00221.x

[22] Pagano, M. (1993) Financial Markes and Growth. *European Economic Review*, 37, pp. 613-622

[23] Brzozowski, M. (2018) Credit Volatility and Productivity Growth. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *13*(2) pp. 215-232, doi: 10.24136/eq.2018.011

[24] Pietrucha, J., & Acedański, J. (2017) Financial Depth and Post-2008 Change of GDP. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *12*(3) pp. 469-482, doi: 10.24136/eq.v12i3.25

[25] Arizala, F., Cavallo, E., & Galindo, A. (2009) Financial Development and TFP Growth: Cross-country and Indsutry-level Evidence. *Inter-American Devleopment Bank. Working Paper*, 682, doi: /10.2139/ssrn.1395716

[26] Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A., (2004) Do Instituions Cause Growth? *Journal of Economic Growth,* 9(3) pp. 271-303, doi: 10.1023/B:JOEG.0000038933.16398.ed

[27] Voight, S. (2013) How (not) to Measure Instituons. *Journal of Institutional Economics*, 9(1) pp. 1-26, doi: 10.1017/S1744137412000148

[28] North, D. C. (1994) Economic Performance Through Time. *American Economic Review*, 84(3) pp. 359-368

[29] North, D.C. (1993) Institutions, Transaction Costs and Productivity in the Long Run. Economic History, 9309004, EconWPA, Retrieved form http://econwpa.repec.org/eps/eh/papers/9309/9309004.pdf

[30] Żelazny, R., & Pietrucha, J. (2017) Measuring Innovation and Institution: The Creative Economy Index. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(1) pp. 43-62, doi: 10.24136/eq.v12i1.3

[31] Hellwig, Z. (1972) On the Optimal Choice of Predictors. In Z. Gostkowski (Ed.). *Towards a System of Human Capital Resources Indicators for Less Developed Countries. Papers Prepared for a UNESCO Research Project.* Wrocław: Ossolineum, Polish Academy of Sciences Press, pp. 69-90

[32] Balcerzak, A. P. (2016) Technological Potential of European Economy. Proposition of Measurement with Application of Multiple Criteria Decision Analysis. *Montenegrin Journal of Economics*, 12(3) pp. 7-17, doi: 10.14254/1800-5845.2016/12-3/1

[33] Cheba, K., & Szopik-Depczyńska, K. (2017) Multidimensional Comparative Analysis of the Competitive Capacity of the European Union Countries and Geographical Regions. *Oeconomia Copernicana*, 8(4) pp. 487-504, doi: 10.24136/oc.v8i4.3

[34] Kukuła, K., & Bogocz, D. (2014) Zero Unitarization Method and Its Application in Ranking Research in Agriculture. *Economic and Regional Studies*, 7(3) pp. 5-13

[35] Bartkowiak-Bakun, N. (2017) The Diversity of Socioeconomic Development of Rural Areas in Poland in The Western Borderland and the Problem of Post-state Farm Localities. *Oeconomia Copernicana*, 8(3), pp. 417-432, doi: 10.24136/oc.v8i3.26

[36] Wierzbicka, W. (2018) Information Infrastructure as a Pillar of the Knowledge-Based Economy — an Analysis of Regional Differentiation in Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 13(1) pp. 123-139, doi: 10.24136/eq.2018.007

[37] Hwang, C. L., & Yoon, K. (1981) *Multiple Attribute Decision Making: Methods and Applications*. Heidelberg: Springer. doi: 10.1007/978-3-642-48318-9

[38] Yoon, K. P., & Hwang, C. L. (1995) *Multiple Attribute Decision Making: An Introduction*. Thousand Oaks: CA Sage Pub

[39] Pietrzak, M. B. (2016) The Problem of the Inclusion of Spatial Dependence Within the TOPSIS Method. *Montenegrin Journal of Economics*, 12(3) pp. 69-86, doi: 10.14254/1800-5845.2016/12-3/5

[40] Pietrzak, M. B, & Ziemkiewicz, B. (2018) Multiple Criteria Analysis of Digital Economy in the European Union Countries. In M. Reiff & P. Gezik (Eds.) *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XIX*. Trenčianske Teplice: Letra Edu, pp. 283-290

[41] Pietrzak, M. B, & Ziemkiewicz, B. (2018) Multiple Criteria Analysis of Digital Economy in the European Union Countries. In M. Reiff & P. Gezik (Eds.). *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XIX*. Trenčianske Teplice: Letra Edu, pp. 283-290

[42] Wang, T. C., & Lee, H. D. (2009) Developing a Fuzzy TOPSIS Approach Based On Subjective Weights and Objective Eeights. *Expert Systems with Applications*, 36(5) pp. 8980-8985

[43] Shemshadi, A., Shirazi, H., Toreihi, M.,. & Tarokh, M. J. (2011) A Fuzzy VIKOR Method for Supplier Selection Based on Entropy Measure for Objective Weighting. *Expert Systems with Applications*, 38(10) pp. 12160-12167

[44] Balcerzak, A. P. (2017) Digital Economy in Czech Republic, Slovakia and Hungary. Measurement with TOPSIS Based on Entropy Measure for Objective Weighting. In T. Loster & T. Pavelka (Eds.) *The 11th International Days of Statistics and Economics. Conference Proceedings. September 8-10, 2016*, Prague: Libuse Macakova, Melandrium, 49-57

[45] Shannon, C. E., & Weaver, W. (1947) *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press

[46] Jenks, G. F. (1967) The Data Model Concept in Statistical Mapping, *International Yearbook of Cartography*, 7, pp. 186-190

[47] Gwartney, J., Lawson, R., & Hall, J. (2017) Economic Freedom of the World: 2017 Annual Report. Fraser Institute. Retrieved form https://www.fraserinstitute.org/studies/economic-freedom

[48] Jantoń-Drozdowska, E., & Majewska, M. (2016) Investment attractiveness of Central and Eastern European countries in the light of new locational advantages development. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 11(1), 97-119, doi: 10.12775/EQUIL.2016.005

[49] Roszko-Wójtowicz, E., & Białek, J. (2017) Evaluation of the EU countries' innovative potential – multivariate approach. *Statistics in Transition new series*, 18(1) pp. 167-180

[50] Kruk, H., & Waśniewska, A. (2017) Application of the Perkal Method for Assessing Competitiveness of the Countries of Central and Eastern Europe. *Oeconomia Copernicana*, *8*(3) pp. 337-352, doi: 10.24136/oc.v8i3.21

The Development of Hybrid IP Architecture for Solving the Problems of Heating Networks (using pipeline-parallel data processing technology)

Alexander Baklanov, Olga Baklanova, Svetlana Grigoryeva, Saule Kumargazhanova, Indira Sagynganova, Yuriy Vais

D. Serikbayev East Kazakhstan State Technical University
Faculty of Information Technology and Power Engineering
A. K. Protazanov Str. 69, 070004, Ust-Kamenogorsk, Kazakhstan
e-mail: {ABaklanov, OBaklanova, SGrigorieva, Saule.Kumargazhanova, ISagynganova, YuVais}@ektu.kz

György Györök

Óbuda University, Alba Regia Technical Faculty
Budai út 45, H-8000 Székesfehérvár, Hungary
gyorok.gyorgy@amk.uni-obuda.hu

Abstract: The paper covers organisation of pipelined processing of data received from heat stations. New software architecture was developed for processing data received from the heating network. Development of a new analytical information system based on pipeline data processing has allowed increased efficiency of work in heat supply systems. The mechanisms for data storage have been established, as well as work sequence and monitoring of heating networks. This architecture is based on a parallel-pipeline data processing system. The idea of such an approach of data processing was transferred from the system of organization of work of the central processor of a personal computer with processes and streams. A distinctive feature of our system is the ability to work with different databases. It can be adapted to various modern systems of data storage and communication in heating systems. The timely operational monitoring described in the article made it possible to change the modes of operation of heat points and the central heat point in real time, which affects the increase in the reduction of heat consumption. The real reduction made was approximately 9 percent. Hopefully, this allows that with the implementation of this approach at the industrial level heat gain and, respectively, energy will increase even more.

Keywords: pipelined data processing; heat stations; heating networks; database; information system

1 Introduction

In 2011, Hungary declared a national energy strategy until 2030 [1]. It included several targets for district heating, such as the connection of public institutions to district heating and an increasing share of renewable energy and waste in heat production. The strategy forecasts the decrease of the share of district heat consumption within the residential and tertiary sector from 12% to 10% by 2020, due to renovation and insulation of buildings.

Despite the fact that the use of renewable energy is in great demand [2], natural gas remains the main fuel used in the energy sector in Hungary. It plays a very important role in electricity production, particularly in cogeneration, as well as in district heating, as 78% of district heating is produced by natural gas. The majority of residential areas are connected to the natural gas network. Out of 4.3 million dwellings, 3.3 million are connected to the natural gas supply and 2.7 million (63%) are heated by natural gas through central or individual heating. As seen in the graph below, the share of district heating in the residential sector is 12% (Figure 1).

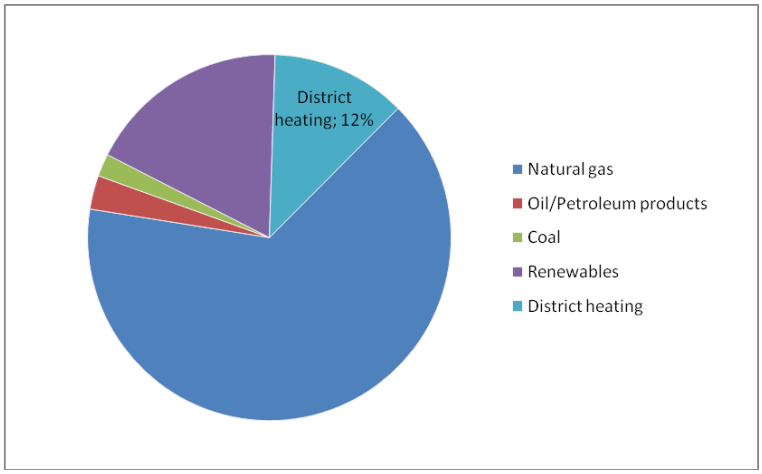


Figure 1
Share of energy source to satisfy heat demand in the residential sector

Before 2010, up to 60% of natural gas related to district heating was used in cogeneration, with the remaining part in heat-only boilers. Afterwards, CHP decreased by 35% until 2012 due to the end of the feed-in tariff in Hungary [3].

For the period 2014-2020, approximately 140 million EUR is to be spent to support the Environmental and Energy Efficiency Operational Programme for energy efficiency and renewable energy projects in district heating systems. According to the National Energy Strategy 2030, the share of energy use from

renewables for residential and institutional heating will increase to 32% and the use of renewable energy will double by 2030. The best framework for renewable heat is district heating.

In the world's most developed countries the main part of the produced fuel and energy resources is spent on the production of electricity and heat, of low and medium potential. Therefore, an important task is to improve the schemes and equipment of energy-consuming plants of industrial enterprises. Each percentage of reduction in energy consumption (electricity and heat) in industry is currently equivalent to a national coal equivalent of fuel economy of about 4 million tons per year.

Heating is Europe's largest end-use of energy. This accounts for approximately 50% of total final energy consumption. District heating (DH) systems provide heating for a wide range of customers, from residential to agricultural, including commercial, public and industrial customers.

There are about 7000 DH systems in Europe, which are currently providing more than 10% of total European needs in heat energy with an annual turnover of 25-30 billion EUR (556 TW*h). Market penetration in district heating is distributed unevenly, in some countries it's nearing zero, yet in others it's up to 70%.

The advantages of centralised heating and centralised cooling are most obvious in regions with high energy requirements. In the European Union about 73% of the population lives in cities, expecting growth to about 80% by 2030. Currently, 69% of all primary energy requirements are concentrated in urban regions [4].

At present, substations for domestic buildings and offices are mostly homemade. By agreeing with coordination of a number of functions at heating substations the district heating sector could produce standard heating stations with significantly reduced homemade portions. By reaching agreement on central heating substations, the industry will be able to produce safer, more environmentally friendly, cost-competitive equipment.

For the organization of effective monitoring and management of modern heating systems, we have developed a conveyor - parallel processing of these heat points. The implementation of such technology was carried out with the help of an information-analytical system that takes into account the experience of software development described in papers [5]. A standard IP network was used to connect all the heating units to the central heating unit.

2 Organization of an Informational and Analytical System

Some of the priority problems are the development and implementation of control and management primarily in large thermal networks; improvement of auto-regulation and protection devices, development of methods and devices for determining the places of coolant leakage before opening the channel. An important task is to improve the schemes and equipment of industrial heat-consuming plants from the point of view of the most rational combination of technological and energy processes and optimization of energy consumption.

Currently, the processing of these heat points is carried out by surveying the heat points in real-time. Because the service is using the server [6], the performance of which is often unable to cope with the amount of information, the incoming data is not fully accountable and does not allow proper control of the heat supply.

Consequently, the use of modern management technologies of heat stations united into a single network will significantly save electricity energy and more accurately distribute heating in residential and industrial premises.

A heat station (HS) is a set of equipment located in a separate, or in the same room with consumers which includes the elements of thermal power plants and makes it possible to connect these installations to a heating system, to control heat consumption modes, to convert and control heating agent parameters as well as heating agent distribution by the type of consumption.

Heating equipment capacity, heat consumption control, the distribution of heating agent by consumption type (heating, hot water, ventilation and air conditioning) is performed through a heat station; the parameters of the heating agent are adjusted and changed.

Heat stations are mandatory both in residential, industrial premises and warehouse facilities. Maintenance of heat stations depends on their type.

Heat station functionality:

- heating agent parameters control and optimization;
- converting heating agent type;
- heating systems protection, reducing the risk of an emergency;
- heating agent distribution in heating systems, water supply and ventilation systems;
- control over the heating agent and heat consumption, as well as providing the necessary consumption of heating agent (this trait is impacted by heat loss, characteristics of the object in conjunction with the specified parameters);
- turning heating systems on and off;
- reduction of heat loss.

There are signs of transition, resulting from the reform of housing and utility services, to a payment system where an owner or a tenant of property will pay for actually consumed heat. Prices for heat, provided by heating enterprises, are currently estimated for the maximum possible consumption. To solve this problem, we need objective testing methods of the amount of heat consumed. Thus, we need a system combining the metering functionality of heat consumption and heat supply regulation, in order to save energy and, at the same time, ensure comfortable conditions in the premises [7]. Based on what is mentioned above, we can conclude that we need such data integration, which includes combining data from different sources and providing users with data in a uniform manner. This process is essential for commercial issues (when two similar companies need to merge their databases), as well as scientific ones (e.g.: combining research results from various bioinformation repositories). The role of data integration increases with the increase in the amount and need for data sharing.

In contrast to the common approaches [8], where data from different databases is converted into a single database, we propose parallel work with data from different databases [9], in order to solve the problems of heat network information resources integration.

The proposed integrated information system architecture is a hybrid model that combines elements of "client-server" architecture and using a computing cluster with parallel distributed heterogeneous information processing (Figure 2).

The information system architecture has four levels.

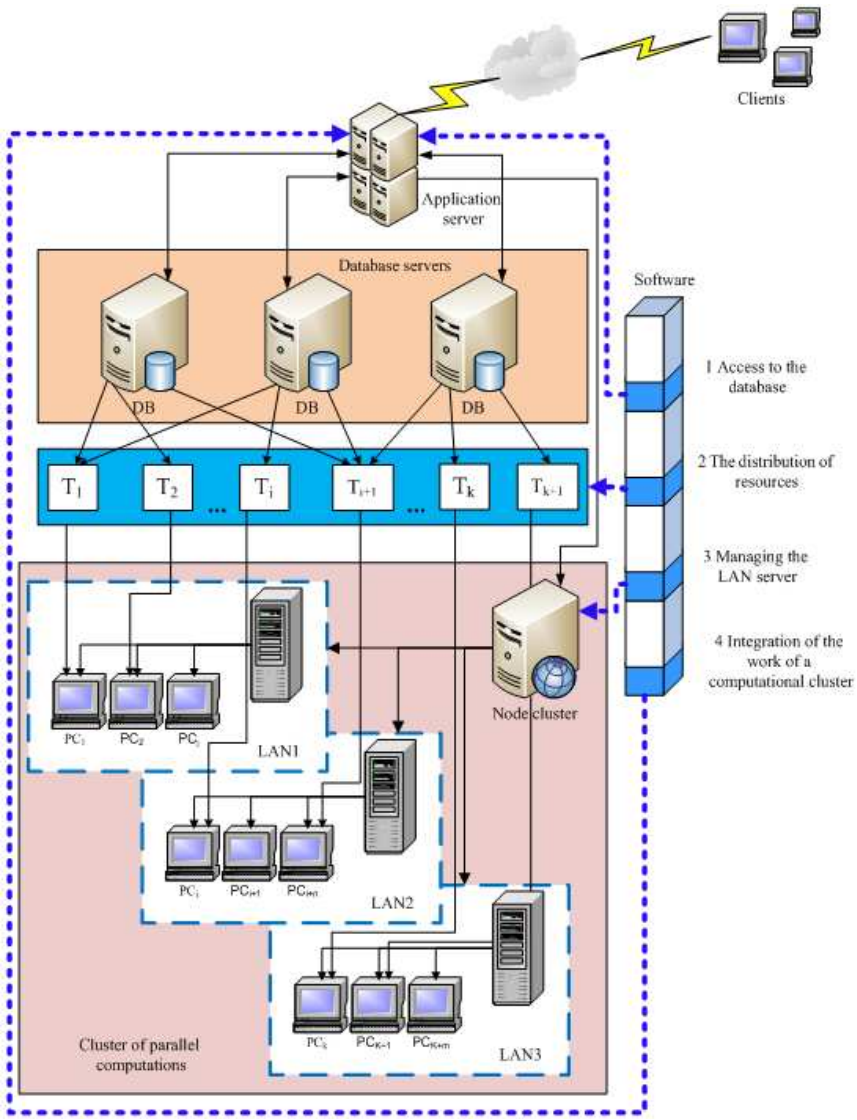
The first level is a "client-server" information system foundation, ensuring parallel operation management of a cluster of computers as well as interaction with the client (user) stations.

The second level is the level of work with the data. Key elements of this level are database servers that operate with databases having different data formats or different platforms based on different storage mechanisms such as relational databases and object-oriented databases.

The third level is the level of solved informational tasks, which specifies the algorithm and sorts out records used in the information system from the databases available at the second level.

The fourth level is the level of implementation of parallel data processing; it is based on a computer cluster which, in our case, allows parallel computing coordinated by the cluster node server which performs dispatch functions for distribution of tasks among the workstations in the cluster in accordance with the instructions of the application server of the information system.

In recent years the world experiences a rapid introduction of computational clusters – local networks, with nodes of workstations or personal computers specially collected to be used as a multiprocessor computing system (supercomputer).



LAN – local area network; DB – databases; Ti – The i-th task

Figure 2
Architecture of an integrated information system

World experience of computational cluster development represents a considerable number of examples from a modest 20-30 node cluster in academic or scientific laboratories at universities to giant computer systems consisting of 1000 – 2000 workstations created in the framework of special projects.

To build computational clusters the following are usually used: public computers based on Intel or AMD processors, a standard Ethernet network technology or the Fast Ethernet, open source Windows operating system [10] and the communication library MPI [11] which implements the connection between the branches of the parallel computing process. Thus, today computing clusters have become a public and relatively cheap alternative to traditional supercomputers. In many classes of tasks and with a sufficiently large number of nodes these cluster systems achieve performance comparable to supercomputer [8, 10].

A set of required software is determined by the objectives of the cluster: a stable multi-user and multi-tasking mode and support for parallel programming techniques.

Schematically, the interaction of software implementation of the information system with the hardware component is also shown in Figure 2.

Initially, information on IP-addresses of database servers, a cluster node, and local area network servers included in a parallel computing cluster is loaded in information system memory modules. These addresses are required to access the database and coordinate parallel operation of a computer cluster workstations. Access to databases is the next step; at that, the information system provides the possibility to use multiple technologies of interaction with databases, namely, ODBC, ADONet and Microsoft Jet, which allows working with relational databases. To work with object-oriented databases (OODB) the information system program code contains a module of access code to databases with a specially designed class, its objects are records in the used OODB.

The module of the information system associated with database interaction is shown in Figure 3.

The module links to the NET Framework software platform. NET Framework provides for a variety of ways to operate databases. The .NET Framework platform has its own technology for data access - ADO.NET (ActiveX Data Object for .NET). ADO.NET includes managed classes allowing .NET application to connect to databases, as well as operate data and control standalone data. ADO.NET technology allows to operate data with Microsoft SQL Server, Microsoft Access, Microsoft Excel, Microsoft Outlook, Microsoft Exchange, Oracle, OLE DB, ODBC, XML in a standalone mode with DataSet objects [12]. DataSet objects allow to extract copies of interconnected local data tables from MS Access. Afterwards, the module operates on DataSet contents, without a need for an active connection to the data source, while also allowing to send modified data back for processing with a corresponding data adapter.

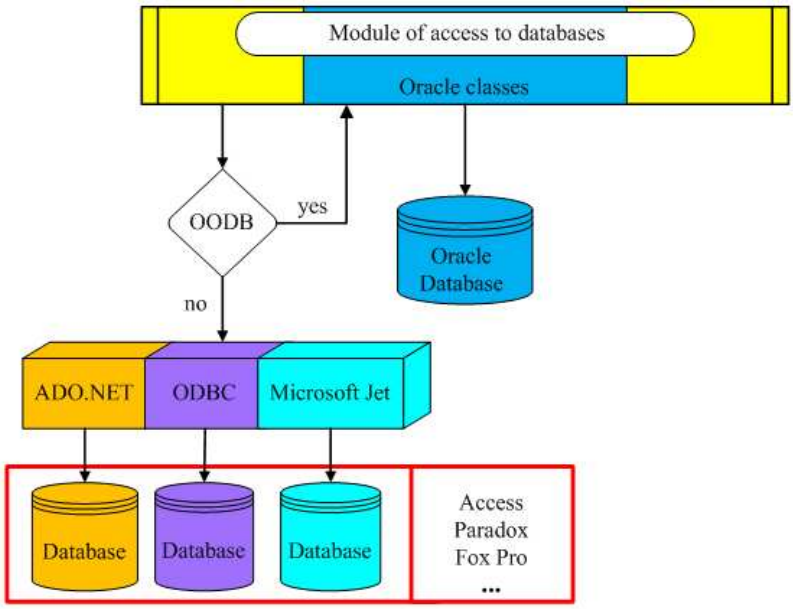


Figure 3
Scheme of the module of interaction with databases

The outline of interaction with the database through ADO.NET is provided in Figure 4.

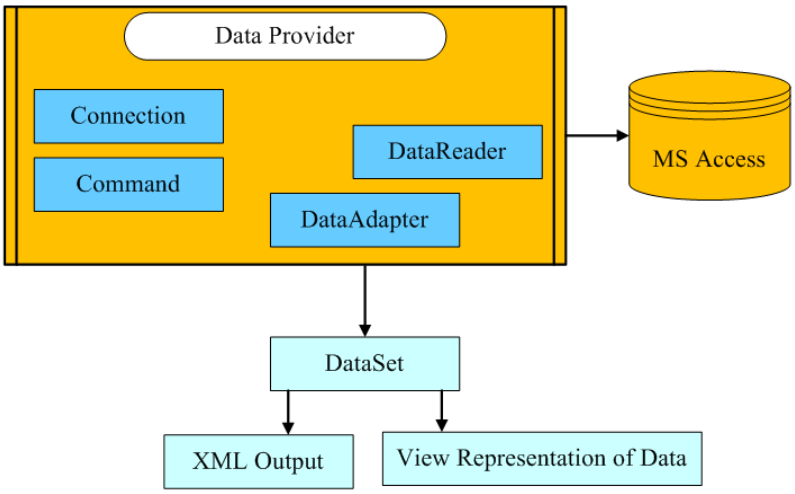


Figure 4
The outline of interaction with the database through ADO.NET

For database operation, including working with Paradox, the module implements an open interface for database access ODBC (Open DataBase Connectivity) [13]. ODBC allows our module to interact with various databases with no need to worry about the intricacies of interacting with multiple sources.

Database operation through ODBC API is carried out in the following manner. First, the connection with database is established. The scheme of access to Paradox database through ODBC interface is shown in Figure 5. In the system the Microsoft Driver Manager (odbc32.dll) interacts with odbcint.ini and odbc.ini. For the operation of Microsoft Driver Manager, which allows loading drivers, ODBC Administrator (odbcad32.exe, odbccp32.dll and odbccp32.cpl) is used. After ODBC Administrator is loaded, we can set the database name, load drivers, modify data, etc. The system has a corresponding driver for operating the databse. Afterwards, the query is run and, after getting the data, the connection is closed.

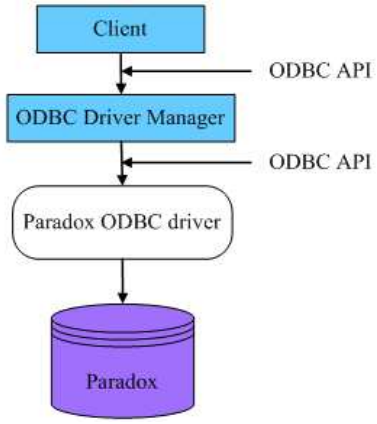


Figure 5
The scheme of access to Paradox database through ODBC interface

For operating Visual FoxPro databases the module makes use of the Microsoft JET Engine technology [14]. One of the three modules of Microsoft Jet Database Engine contains the ISAM drivers, DLL libraries allowing connections to ISAM [15] databases including Visual FoxPro. Another one of DAO modules implements the API. API allows to access JET databases through an arbitrary programming language, which the module of informational system uses for interaction with Visual FoxPro databases.

A part of module of an informational system for interaction with Oracle database was developed in an integrated development environment for database applications, PowerBuilder [16]. PowerBuilder was chosen because it uses the native interfaces for connection to Oracle and a patented technology for data operations – DataWindow. PBNI technology in PowerBuilder eliminates the flaw of long compilation time for analysis.

Switching of access modes of operating heat stations in the cluster is done by the program in accordance with the database format. At that, the module of work with relational databases uses standard technologies for accessing databases; when it works with an object-oriented database (DBMS Oracle in this case) it uses a special access code which creates mirror classes with the used OODB.

Such a mechanism of interaction with databases ensures independent operation of heat stations in the cluster.

The distribution of resources is as follows: the application server consistently produces the analysis of the forthcoming tasks (Figure 6), and then selects task fields associated with the use of a single database. The application server redirects this group of tasks to workstations in one of local area networks using the cluster node. Further, other groups of tasks are sequentially formed. The final distribution of parallel computation cluster resources is performed by the application server up to the last task. Moreover, if the task uses data stored in various databases that task is broken down into sub-tasks, and the distribution of these sub-tasks for workstations in the local area network is similar to the distribution of the tasks themselves.

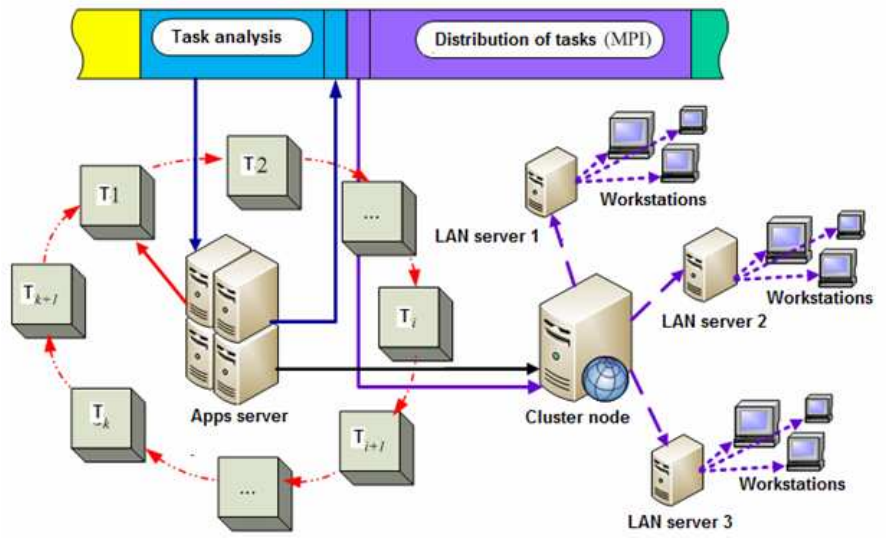


Figure 6
Scheme of distribution of information system resources

At the final stage of work, the program, using the MPI communication library [17], transfers control in accordance with the distributed tasks to the cluster workstations. Workstations independently process data from the appropriate storage and perform the calculations necessary to solve the corresponding task of the workstation (Figure 7).

The main application performs querying of all operating heat substations involved in the solution of the problem at specified intervals. The querying is carried out cyclically by searching all the working heat points; and in the case all the problems are solved, the last module is launched, integrating the results and providing documentation for the projects of construction, repair or reconstruction of heat networks.

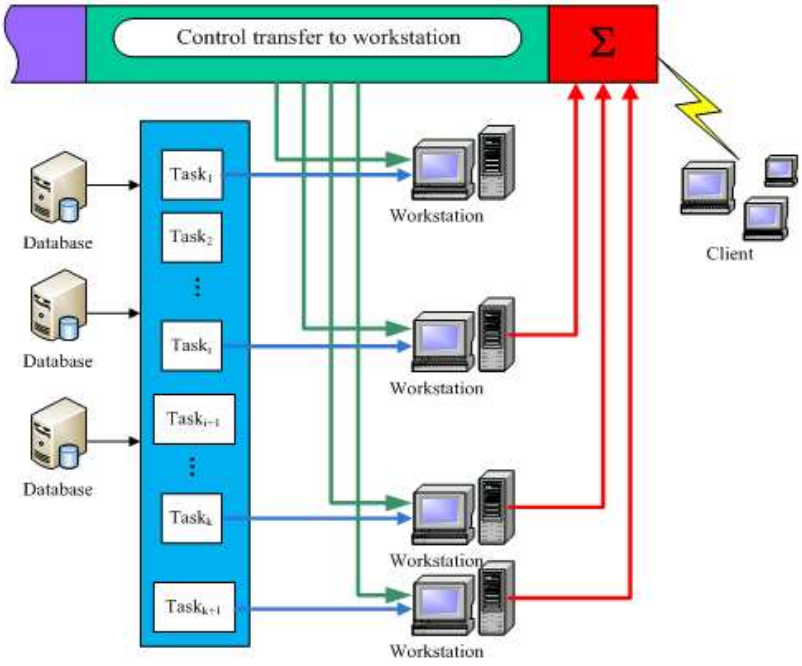


Figure 7
Scheme of IP management of workstations in a computational cluster

A client application that performs the functions of the automation unit on management of heating networks, uses received at the application server results ready for the formation of documentation on work of heating networks. In our case, it will allow to quickly process the data on thermal stations stored in different databases and at the same time to operate objects with remote access.

The organization of data protection was carried out taking into account the material of article [18, 19].

Based on the above approach, the authors developed an information system for servicing a typical heating station.

3 Implementation of Pipelined Data Processing in Heating Networks

Heat supply systems implement central, local and individual regulation.

Local regulation is applied at consumer entrance points and heat stations and and aims to adjust the mode of central regulation of heat consumption. The main factors causing the need for regulation of heat consumption for heating at entrance points can be categorised according to Table 1.

Table 1

The main factors causing the need for local regulation

Name of the factor	Description
1 Mismatch between the static characteristics of the heating system and the mode of central regulation of heat consumption.	Different values of calculated air temperatures in heated buildings (dwelling houses, schools, kindergartens, etc). Different values of calculated air temperatures of outside air (buildings of varying complexity). Mismatch of the heating surface of heating equipment installed in a building to the temperature regime of the heating network. Uneven cooling of the water in the pipes during the transport of heat carrier to various buildings.
2 Unequal dynamic characteristics of heated buildings, consumer heating systems and heating network sections from heat generator to the building.	Different thermal stability of buildings. Different dynamic characteristics of heating systems (radiator, panel, direct heating). Varied values of transport lag in the heating system (up to the building).
3 The influence on the regulating value of perturbations (temperature and heat carrier flow) acting between the heat source and the input into the building.	Operation of hot water supply installations. Inclusion of forced ventilation setups. Switching in the heating network.
4 Nonuniform nature of heat consumption.	Impossibility of implementation of central regulation according to the heating schedule in the entire range of heat demands.

The results of examination of the stated facts have shown that in the absence of local regulation in some cases there may be serious violations of the thermal regime of buildings.

Proceeding from the above, we have developed and introduced “TSmonitor”, an original software program (Analysis of parameters of a heat station), allowing to manage the system by heating network objects, calculate and accomplish other actions. The specified technique uses an informational model based on automation of a control system of heat stations with use of pipelined data processing [20, 21].

The goal of given software solution is to increase the effectiveness of control and metering of heat energy [22].

The software product represents a unified informational system in which the user can accomplish management in the optimal conditions. Informational system uses the ZULU geoinformation system, which was developed in Politerm Company (St. Petersburg, Russia), for map display [23]. The given geoinformational system gives an opportunity to draw maps with layers and use one's own object model (Activ X). The interface of the main working windows is shown in Figure 6.

For the experiment we have chosen 11 heat stations in Zyryanovsk in East Kasakhstan Oblast. A standard IP network was used to connect all the heating units to the central heating unit. In each heat station 11 parameters were registered with their minimum and maximum values, as well as the interval boundaries.

These are the following parameters (see Figure 8): the network water temperature in the supply pipeline, ^{o}C; the network water temperature in the return pipeline, ^{o}C; circulation temperature, ^{o}C; the network water flow in the supply pipeline, tons/hour; the network water flow in the return pipeline, tons/hour; the network water pressure in the supply pipeline, kgf/cm; the network water pressure in the return pipeline, kgf/cm; indoor temperature, ^{o}C; heat released, Gcal; voltage 1, V; voltage 2, V.

Besides that the following are displayed in real-time: date and time of connection and analysis; registered parameters and their value at the moment of connection.

In the software an approach using pipelined data processing was implemented. Two threads are created. The first is responsible for the process of downloading data from the "device", that is, the cycle goes through the recorded thermal nodes (TN) and downloads the previously generated parameters to the database (generated by the DataGenerate module.exe). Thus, the process of downloading data from the devices is simulated, when the server alternately surveys each node and reads data from their database.

After the survey of each node there is a delay of several milliseconds, so that to delay the process of reading. The delay can be configured in the "Connection parameters – Settings" menu. The amount of records kept in the downloaded data storage table can also be configured; after reaching the maximum the recording will be done as if into "stack". Launching and stopping a thread takes place by pressing the buttons placed on the panel. The larger the number of nodes from which the reading is taken, the slower the update process for every node.

Figure 8

The main program window

The second thread is responsible for updating the parameters of the selected node on the map. It starts automatically when the program starts. When you select a TN on the map this stream in continuous mode refers to the database and downloads the latest records on this TN and updates the appropriate field.

Using the program allows to considerably increase the reliability, longevity, as well as work efficiency both of the heat stations and of the heating system in general. Thanks to the constant monitoring of the operation of heat stations in the pipeline mode, the importance of task pipelining techniques in automated control systems was proven.

This conclusion follows from the production development theory and data transformation theory: it makes sense to organize all the same type of task sequences in the pipeline plan; the presence of task pipelining mechanisms in an automated control system leads to the increase in the work efficiency of the system.

We have charted the heat consumption in residential area in Figure 9.

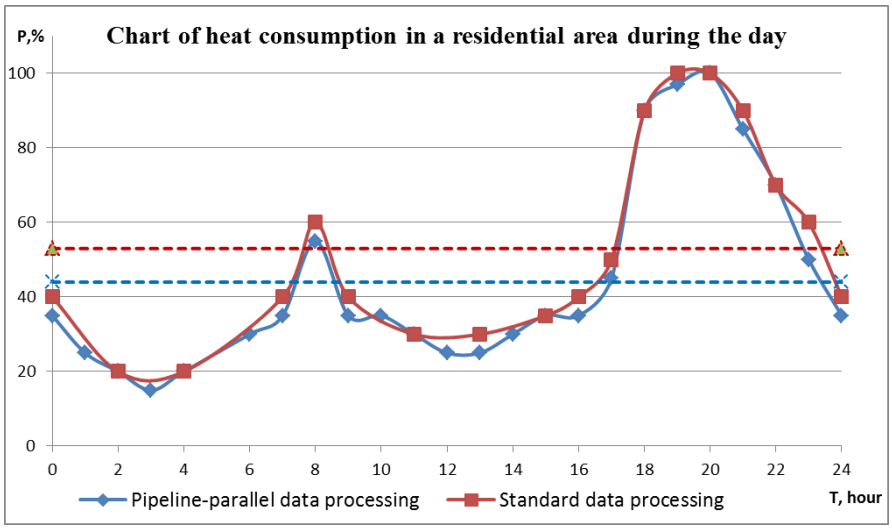


Figure 9
Heat consumption in residential area

We have made measurements of heat consumption during the day for the spring quarter. The data of heat consumption during the day was recorded while using pipeleined system (Chart 1) and while using a typical data processing system (Chart 2). It is seen that while using an informational and analytical system the heat consumption decreases by about 9%. So the average heat consumption when using pipelining data was 44% in relative units (red dotted line), whereas with the usual system it was 53% (blue dotted line).

Conclusion

The analysis of the measurements (Figure 9) showed that the use of parallel-conveyor data processing technology can significantly reduce heat and electricity consumption in the period when the heat consumption is close to or below the average level of heat consumption per day. So, in the period of time between 12-1400 hours the heat saving is 15%, and around 23 hours = 20%. It shows that the efficiency of our system increases in the period of time when there are no peak loads.

Thus, the hybrid information system that has been developed allows timely solving of complex problems of controlling heating networks operation, and regional heat stations in particular, without involving supercomputers but using a computer cluster developed on the basis of available technical support at JSC "Heating Networks". Also, this approach can be used in image processing of minerals [24, 25].

In the future, it is planned to use intelligent mechanisms in the system [26] to improve data analysis, as we believe that this will eventually further reduce excessive heat consumption.

References

[1] Hungary National Energy Strategy, Publisher: Ministry of National Development, 2012, 131 p. ISBN 978-963-89328-3-9

[2] Brożyna, J., Mentel, G., Szetela, B: Renevable Energy and Economic Development in the European Union, Acta Polytechnica Hungarica, Vol. 14, No. 7, 2017, pp. 11-34, DOI:10.12700/APH.14.7.2017.7.2

[3] Euroheat & Power, District Energy in Hungary. Link: https://www.euroheat.org/ knowledge-centre/district-energy-hungary

[4] European Committee for Standardization. Link: https://www.cen.eu/work/ areas/energy/heatingSubstations/Pages/default.aspx

[5] Györök, Gy., Kumargazhanova, S., Baklanov, A., Uvaliyeva, I., et al.: Development of the Information and Analytical System in the Control of Management of University Scientific and Educational Activities, Acta Polytechnica Hungarica, Vol. 15, No. 4, 2018, pp. 27-44, DOI: 10.12700/APH.15.4.2018.4.2

[6] Automated process control systems of heat points. Link: https://www.krug2000.ru/decisions/solutions_zkx/teplopunkt.html

[7] Kutateladze, S. S.: Research and development of the Siberian Branch of the Russian Academy of Sciences in the field of energy efficient technologies, SBRAS Institute of Thermophysics Novosibirsk, 2009, 405 p.

[8] Ádám, N., Balaz, A., Pietriková, E., Chovancová, E. and Feciľak, P.: The Impact of Data Representationson Hardware Based MLP Network Implementation, Acta Polytechnica Hungarica, Vol. 15, No. 2, 2018, pp. 69-88, DOI: 10.12700/APH.15.1.2018.2.4

[9] Rybakova, D., Kvasov, A., Baklanov, A.: Using pipeline data processing system to regulate operation of heat stations, Vestnik EKSTU, N.3, 2014, pp. 101-105

[10] Voevodin, V., Voevodin, Vl.: Parallel computing, BHV-St Petersburg, 2002, 609 p.

[11] Shpakovsky, G. I., Serikov, N. V.: Programming for multiprocessor systems in MPI standard. Mn .: BSU, 2002, 323 p.

[12] ADO.NET. Link:https://docs.microsoft.com/ru-ru/dotnet/framework/data/adonet/index+

[13] Ganeev, R. M.: Web-interface of ODBC databases, Publisher: Hotline - Telecom, 2003, 202 p.

[14] Haught, D., Ferguson, J.: Jet Database Engine Programmers Guide (Microsoft Professional Editions) 2nd edition, Publisher: Microsoft Press; 1997, 756 p.

[15] Informational portal Bauman National Library. Link: https://ru.bmstu.wiki/Microsoft_Jet_Database_Engine

[16] Brown, M., Armstrong, F.: PowerBuilder 9: Advanced Client/Server Development. Publisher: Sams, 2003, 840 p.

[17] Pacheco, P.: Parallel Programming with MPI. Publisher: Morgan Kaufmann; 1996, 500 p.

[18] Ádám, N., Madoš, B., Baláž, A. and Pavlik, T.: Artificial neural network based IDS, 15th International Symposium on Applied Machine Intelligence and Informatics, Herl'any, Slovakia, 2017, pp. 159-164

[19] Vokorokos, L., Baláž, A., Ádám, N.: Secure Web Server System Resources Utilization. Acta Polytechnica Hungarica, Vol. 12, No. 2, 2015, pp. 5-19, DOI: 10.12700/APH.12.2.2015.2.1

[20] Rybakova, D., Baklanov, A., Kvasov, A.: The conveyor organization of the tasks in the work of the ACS heat point, Energy and fuel resources of Kazakhstan, 2010, N3, pp. 87-92

[21] Grigoryeva, S., Baklanov, A., Gyorok, Gy.: Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica, Vol. 13, No. 5, 2016, pp. 105-119

[22] Rybakova, D., Sygynganova, I., Kumargazhanova, S., Baklanov, A., Shvets, O.: Application of a CPU Streaming Technology to Work of the Computer with Data Coming from the Network on the Example of a Heating Station. 18th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM), Erlagol, Russia, 2017, pp. 128-130

[23] Company Politerm. Link: https://www.politerm.com/en

[24] Baklanova, O., Baklanov, M.: Methods and Algorithms of Image Recognition for Mineral Rocks in the Mining Industry, Advances in Swarm Intelligence. ICSI 2016. Lecture Notes in Computer Science, Vol. 9713, Springer, 2016, pp. 253-263, DOI:10.1007/978-3-319-41009-8_27

[25] Baklanova, O., Shvets, O.: Cluster analysis methods for recognition of mineral rocks in the mining industry. 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, France, 2014, pp. 273-277, DOI:10.1109/IPTA.2014.7001972

[26] Jocic, M., Pap, E., Szakál, A., Obradovic, D., Konjovic, Z.: Managing Big Data Using Fuzzy Sets by Directed Graph Node Similarity, Acta Polytechnica Hungarica, Vol. 14, No. 2, 2017, pp. 183-200, DOI: 10.12700/APH.14.2.2017.2.10

Development of a Novel Automotive Cybersecurity, Integrity Level, Framework

Árpád Török, Zsolt Szalay, Balázs Sághi

Budapest University of Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Department of Automotive Technologies, Műegyetem rkp. 3, 1111 Budapest, Hungary; arpad.torok@auto.bme.hu; zsolt.szalay@auto.bme.hu; saghi.balazs@mail.bme.hu

Abstract: As automated driver assistance functions are getting more and more popular, they will surly have a significant impact on our life especially considering security and the expected serious effects of malicious interventions. In light of the introduced aspects, the Test Field of Zalaegerszeg has started a research to evaluate the required professional and scientific framework to prevent transport systems from malicious external intervention. With regard to this aspect, the homologation system has been analyzed. In accordance with this the main objective of the article is to develop integrity level structure related to transport systems especially focusing on security issues. In this context the paper reconsiders the structure of ASIL to provide a proper framework focusing on the outstandingly important issues of cybersecurity. To develop a novel SIL framework a tailor-made method is identified to define risk parameter values related to certain SIL categories. At the end of the research, the investigation has determined the acceptable hazard rates related to the S&SIL architecture.

Keywords: cybersecurity integrity level; automotive safety integrity level; clustering model; tolerable hazard rate; hazard classification category

1 Introduction

Currently the significance of the secured and well-protected cyberspace is continuously increasing. By this time, online connections have become as important as personal relationships, everybody uses the internet to keep in touch with the surrounding environment. On the one hand, everybody feels that the digital development facilitates the formation and the maintenance of social networks. Everything is getting closer, goods can be easily purchased from the other side of the world and people from different continents can easily meet each other on the internet. These new possibilities extremely increase the effectiveness of communication related processes. New methods like fifth generation cellular network technology [1] and intelligence demonstrated by machines [2] will

probably revolutionize the recent world. However, on the other hand our online dependency is going to strongly influence the security of the society. In light of the introduced trends, data integrity, privacy, data security, individual safety and even public safety can be threatened. To summarize, the spread of connected devices and the growing influence of the cyberspace on our life will make it necessary to improve the protection of safety and security.

Beside this, the spread of self-driving cars can provide additional advantages through the concept of connected vehicles by utilizing the possibilities of real time information exchange. However, this new mode of transportation will surly generate new vulnerabilities related to mobility processes influencing the safety of our everyday life. In light of this, attack detection is going to be one of the main objectives to provide the proper safety and security of the transportation system [3]. Since the increasing number of connecting devices in highly automated cars can lead to serious vulnerabilities, which can be exploited by malicious intruders influencing safety and security of everyday mobility processes. On the other hand, moral considerations can also not be forgotten, which in many cases make the cybersecurity related development processes even more difficult [4].

In accordance with this, the proper and acceptable security level of highly automated vehicles has to substantially be evaluated. For this purpose, it is crucial to perceive and remove those factors in the system which increase security risk. To consider safety and security together, even during the earliest stages of the design procedure, researchers have developed a new method for the automotive industry [5]. In light of the introduced development processes, the Test Field of Zalaegerszeg [6] has started a research to evaluate the required professional and scientific framework to prevent transport systems from malicious external intervention. Since nowadays it seems to be an outstandingly important research objective to identify the homologation framework of automotive cybersecurity, the gap between the classical methods of automotive standardization and validation and the quickly changing informatics field has to be bridged [7]. Considering that the methods followed by the two mentioned segments are considerably different, this purpose seems to be a difficult task. Since if classical functional safety related audit methods and the high-tech cybersecurity related analytical background are compared, we can find significant differences.

In accordance with this, the main objective of the evaluation coordinated by the Test Field of Zalaegerszeg Research Committee is to identify the foundations of a novel integrity level architecture taking into account both safety and security related aspects. In the first step related works have been reviewed to get a clear picture about the recently applied most up-to-date methods and models, which can support the development of a novel complex integrity level architecture. Even the functional safety characteristics of microchips and circuits can have a serious impact on security of highly automated vehicles, hence the work of Belotti et al. has to be thoroughly investigated to build up a comprehensive methodological framework [8]. Their paper introduces in details the integrity level frame work

applied by the automobile industry, on the other hand the paper does not do further steps to propose development orientations related to the recently applied integrity level architecture. Chang and other researchers have also analyzed the effect of automotive functional safety standards on the applied microchips [9]. Their paper provides a comprehensive safety investigation in case of the analyzed automobile elements. However, it has to be emphasized that the paper does not focus on security related problems or issues [10]. Other important automotive industry related forums like SAE proposes different approaches to evaluate the security level of highly automated cars. These proposals are in accordance with the most important automotive industry related standards (e.g. ISO 26262). One of the most known methodological framework is HEAVENS, beside this, EVITA is also a well-structured model environment [11]. On the other hand, it needs to be considered that theses frameworks do not combine the aspects of safety and security, which should be a key issue in the automotive sector. Beyond this, the mentioned models place much less emphasis on the investigation of acceptable level of cybersecurity risk than it would be reasonable. According to the results of the relevant researches related to the combined evaluation of safety and security in the automotive sector [12], it is now obvious that accident risk of highly automated vehicles cannot be analyzed without the comprehensive evaluation of automotive cybersecurity [13]. In accordance with this, when safety integrity level of automotive systems [14] are investigated, security vulnerabilities and the related preventive interventions are also necessary to be analyzed in details. To present the newly developed S&SIL (combined safety and security integrity level) architecture, the relevant risk parameters have been evaluated, especially considering severity, probability and controllability [15]. In light of the introduced considerations, this work focuses on the identification of the novel integrity level framework, especially considering the relevant risk parameters, which need to be taken into account. In accordance with this, the paper covers the following main topics: interpretation of the used models, introduction of the main findings and the explanation of the most important outputs of our investigation.

2 Applied Models

To identify the novel integrity level framework, the proper terms should be identified and well described. For this purpose, the collection and description of the most important definitions related to the integrated model of safety and security should be the first and foremost step. In accordance with the introduced field, the most relevant terms are necessary to be defined. During our model development process hazard is defined as the alignment of the possible conditions and circumstances, which can lead to harm. Therefore, the investigation of potential hazards has to be followed by further analysis. In the following phase, it has to be evaluated whether the identified hazards affect system vulnerabilities or

not. In light of this, vulnerability can be defined as the ability to be affected by an investigated hazard. The term of impact reflects to relationship between vulnerability and hazard. If a hazard is getting activated, its impact on the system depends strongly on the system vulnerability with regard to the given hazard. Furthermore, the definition of risk can be derived by the terms of probability and its possible impact through multiplying the two factors by each other. Beside this, the term of threat represents the alignment of the above mentioned negative factors. Accordingly, if there is a relevant hazard, a strongly related serious vulnerability and the expected impact seems to be significant, then the investigated system is considerably threatened. Accordingly, the figure below (Figure 1) represent the introduced processes.

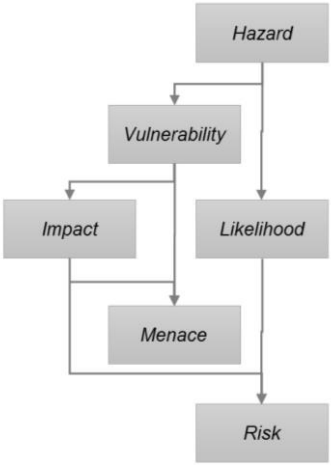


Figure 1
The connection among the introduced terms in the integrated fields of safety and security

Safety hazards are described by many popular models, such as GAMAB, ALARP and MEM. The above-introduced methods investigate the level of acceptable and tolerable hazards. Their aim is to describe if certain hazards in relation to their derived risk can or cannot be accepted on the level of the society. The main approach of the paper is in accordance with the GAMAB model, applying the main concept of integrity level framework to develop a novel architecture, which fits to the requirements of the integrated fields of safety and security. Tithe outcomes of this concept can lead to complex and holistic representation with regard to the analyzed functions. This added value can be outstandingly important from the viewpoint of safety and security, since the newly generated information on system characteristics can strongly influence the applied analytical methods and the efficiency of the investigation during the product lifecycle. [16]. Integrity level framework can support the investigation in describing the ability of the system to stay in a safe and secured operation mode. In light of this, the goal of the analysis is to determine if the derived risk level in case of a certain hazard is

tolerable or not. Accordingly, Kreiner et al. have shown that safety and security of highly automated systems have an outstanding importance especially if they have been built-up from strongly interrelated hardware and software components [17].

The acceptable risk level in case of safety critical system can be determined considering the related risk category. In this regard, the applied model - as mentioned above – takes into account the severity, probability and controllability. Accordingly, the demanded safety and security characteristic of the investigated safety critical system is determined based on the acceptable risk level of the society. In light of this, the demanded safety and security characteristic of a safety critical system is determined as the level of the acceptable risk [18]. It has to be kept in mind that the article aims to discuss safety and security related issues of the automotive segment, in accordance with this, the new integrity level framework is derived from the ASIL concept [19]. The expected likelihood of an effective malicious intervention, is difficult to be estimated, especially with acceptable prediction accuracy. In accordance with this, during our investigation the method introduced by Dudrov et al. is used to determine the probability of the investigated malicious intervention types [20].

3 Results

The primary outcome of the completed investigation is the understanding, the interpretation and the demonstration of the timeliness of a novel model, in case of the combined fields of safety and security. Firstly, it has to be understood that the set of consequences are not the same in case of safety and cybersecurity. While safety mainly focuses on human severity, cybersecurity has to consider also information security related aspects as well. On the other hand, classical cybersecurity risk classification models primarily investigate privacy, data protection and national security related issues. Accordingly, if the aim is the development of a common framework, safety and security related consequences have to be reduced to a common denominator. Furthermore, the applicability of controllability as a risk parameter should be reconsidered, since this ability of the system fundamentally depends on the relationship of the system and the investigated attack, namely, how efficiently can the given system detect and then treat the investigated malicious intervention. Accordingly, detectability (D1, D2, and D3) and treatability (T1, T2, and T3) should be among the most important risk parameters in the newly developed model focusing on the integrated fields of safety and security. In light of the introduced findings a novel S&SIL architecture is developed, which covers private data security (PR-S), public data security (PU-S), national security (NS) and also safety integrity. The following table describes the new integrity level framework representing the field of safety and security (Table 1). In Table 1, QM refers to issues, which can be solved by the tools of

quality management and S&SIL A..D refer to issues, which have to be investigated by the tools of safety and security assessment methodologies (Tab. 1).

Table 1
Representation of S&SIL framework

		T1			T2			T3		
		D1	D2	D3	D1	D2	D3	D1	D2	D3
B-S-PR	E1	QM	QM	QM	QM	QM	QM	QM	QM	QM
	E2	QM	QM	QM	QM	QM	QM	QM	QM	S&SIL A0
	E3	QM	QM	QM	QM	S&SIL A0	QM	QM	S&SIL A0	S&SIL A1
M-S-PR	E1	QM	QM	QM	QM	S&SIL A0	QM	QM	S&SIL A0	S&SIL A1
	E2	QM	QM	QM	S&SIL A0	S&SIL A1	QM	S&SIL A0	S&SIL A1	S&SIL A2
	E3	QM	QM	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL B0
B-S-PU	E1	QM	QM	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL B0
	E2	QM	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL B0	S&SIL A1	S&SIL A2	S&SIL B0	S&SIL B1
	E3	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL B2
M-S-PU	E1	S&SIL A0	S&SIL A1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL B2
	E2	S&SIL A1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL B2	S&SIL B0	S&SIL B1	S&SIL B2	S&SIL C0
	E3	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL C1
M-NS	E1	S&SIL A2	S&SIL B0	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL C1
	E2	S&SIL B0	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL C1	S&SIL B2	S&SIL C0	S&SIL C1	S&SIL C2
	E3	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL D0
C-NS	E1	S&SIL B1	S&SIL B2	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL D0
	E2	S&SIL B2	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL D0	S&SIL C1	S&SIL C2	S&SIL D0	S&SIL D1
	E3	S&SIL C0	S&SIL C1	S&SIL C2	S&SIL D0	S&SIL D1	S&SIL C2	S&SIL D0	S&SIL D1	S&SIL D2

The following investigation focuses on the definition of the currently acceptable risk level related to the combined fields of safety and security. In the next step, the defined risk level is going to be implemented in case of the newly introduced architecture. The acceptable risk level is derived from the estimated number of connected tools and instruments, from the expected number of malicious interventions arriving from the cyberspace and from the expected likelihood of efficient interventions [20].

$$L = \frac{N \cdot \sum_{i=1}^{n} r_i \cdot l_i}{M} \qquad (1)$$

L: expected likelihood of efficient interventions,

N: expected number of malicious interventions arriving from the cyberspace in 2020,

M: estimated number of connected tools and instruments in 2020,

li: expected likelihood of efficient interventions regarding the i^{th} attack type,

ri: is the ratio of the i^{th} intervention type, and

n: number of intervention types.

Based on the ASIL framework, in light of the derived acceptable risk level, the novel S&SIL rating scale can be introduced as follows (Table 2):

Table 2
S&SIL rating

	S&SILs	Probability of succ. incident
1	S&SIL A	<10-7
2	S&SIL B	<10-8
3	S&SIL C	<10-8
4	S&SIL D	<10-9

4 Discussion

The development of the combined architecture of safety and security integrity levels should start with identification of the main consequence classes. At this stage it is a basic aspect with regard to the novel architecture to integrate the classes considered by ASIL, ENISA [21] NCCIC framework. Taking into account the mentioned models, the primary goal of the analysis is to combine the introduced risk evaluation methods. ASIL [8] actually includes three different types of possible consequences, ENISA [21] classifies the investigated functions and processes based on three different evaluation assurance level (EAL), and NCCIC categorizes the malicious interventions in seven groups. The framework applied by the automotive sector considers primarily those kind of hazards, which can be dangerous to life, in accordance with the applied risk categories refer to the severity of the expected injury. On the contrary, ENISA uses a much simpler approach and hazards are classified into basic, substantial and high risk groups. This simpler model primarily considers the vulnerability of the evaluated function or process. In case of the ranking system used by NCCIC the classification of the investigated hazard is primarily influenced by the assumed attack vector. According to the applied methodology, the threat class of a given malicious

intervention is determined as legible if the expected effect of the given attack is marginal. The priority of an attack is defined as minor if the hazard related to the given incident to have impact on public health or national security is reasonably low. An attack is classified into the low priority group if it cannot be assumed to significantly influence public health or national security. When the evaluated attack can have an effect on public health or national security it should be classified into the medium-risk group. If the investigated incident is likely to influence public health or national security it should be assigned to the high-risk group. When a malicious intervention can have significant impact on public health or safety and national security it is defined as a severe priority incident. When a malicious intervention menaces critical infrastructure, national security or human lives it is defined as emergency. In accordance with the introduced risk ranking methods it is possible to develop an overall evaluation concept covering the most important considerations of the mentioned models. It has been introduced that a fundamental criteria of the development process of the new integrity architecture combining the fields of safety and security to be reconcilable with the automotive safety concept. In case of this criterion, the most important goal has been to provide the mutual interoperability between the integrity level frameworks of the classical automotive sector and the newly developed architecture of safety and security. In order to ensure interoperability between the classical automotive approach and the newly developed integrity level framework, it is expected to have the same structure with regard to their lines and columns. In light of this, the numbers of lines and columns of the novel architecture should be the integer multiple of the numbers of columns and lines used in case of the integrity level framework of the automotive sector. Based on the introduced considerations, the developed risk evaluation framework includes six different risk groups, which are mainly derived from the incident scoring system of NCCIC. The number of baseline risk levels are contracted to one group. At the same time, the risk groups used in the novel concept are also in accordance with the severity categories of the automotive safety integrity levels. In accordance with this the definitions of the risk groups contain descriptions related to effects possibly resulting personal injury as well as information damage or theft. Due to the introduced considerations the following levels can describe the risk group structure related to the developed cybersecurity integrity level architecture (Table 3).

In light of Table 3 it is obvious that beyond the mentioned cybersecurity related approach of NCCIC and the introduced automotive sector related framework; the newly developed concept also refers to the presented evaluation assurance level applied by the ENISA. In accordance with this, the first two risk groups of S&SILs are related the basic assurance level of ENISA, the second two groups of S&SILs are related to the substantial assurance level of ENISA, while the third two risk groups are related to the high assurance level of ENISA. With regard to the likelihood of the investigated attacks, classical likelihood approach based groups are re-clustered into three groups.

Table 3
Representation of cybersecurity significance level

Risk group	Description
Threatening either basic safety or security of private data (B-S-PR)	The attacks in this class cause less relevant menace to private data or slight personal injuries.
Threatening either moderate safety or security of private data (M-S-PR)	Malicious interventions in this group can lead to moderately relevant menace to private data or moderate injuries.
Threatening either basic public safety or public data security (B-S-PU)	Attacks in this class cause less relevant menace to public data or severe personal injuries.
Threatening either moderate public safety or public data Security (M-S-PU)	Malicious interventions in this group can lead to moderately important menace to public data or life-threating personal injuries.
Threatening national security (M-NS)	Attacks in this class cause moderately relevant menace to national security or fatal injury.
Critical threatening national security (C-NS)	Malicious interventions in this group lead to critical menace to national security or numerous fatal injuries.

Based on this classification, incidents can be less (L1), moderately (L2) or critically (L3) likely to cause threatening. In case of the safety and security related combined integrity level framework in light of the classical automotive sector related architecture, the factor of controllability becomes more complex, since the controllability of a security incident is strongly affected by its detectability and treatability. [22]. Actually, in case of the combined field of safety and security, with the detection and the treatment of a given attack the whole control process can be performed. Accordingly, in the first step the investigated attack should be recognized and only then that attack can be averted. Based on this, with regard to detectability as well as treatability, attacks can be assigned to three classes (Table 4).

Table 4
S&SIL risk parameters

	Detectability levels	Definition		treatability levels	Definition
1	D1	Basic level of detectability	1	T1	Basic level of treatability
2	D2	Moderate level of detectability	2	T2	Moderate level of treatability
3	D3	Critical level of detectability	3	T3	Critical level of treatability

Based on the application of the previously interpreted correspondences, the novel integrity level framework combining the aspects of safety and security can be built up. (table 1). In this matrix lines include the levels of the described risk groups (B-S-PR, M-S-PR, B-S-PU, M-S-PU, M-NS, C-NS). The likelihood to detect or treat incidents is described by the columns of the investigated matrix. Based on the classical safety integrity level model applied by the automotive sector, if the occurrence probability and the expected effect of the malicious interventions related to the analyzed function or module are low and its detectability and treatability are high then the related issues can be solved by quality management related processes. The objective of the next development step is to fit the novel framework to the classical automotive safety integrity architecture. In light of this, the assignment process between hazard classes (ASIL QM, A, B, C, D) and risk parameter groups (the combinations of severity, probability and controllability) are represented by a specific optimization problem. The aim of the method is to assign the proper hazard class to a certain combination of risk parameters. To do so the adequate estimation function has been selected, which can estimate the proper hazard class based on the risk parameters as input variables. Accordingly, the objective of the optimization problem is to define the coefficients of the estimation functions by minimizing the differences between the values of the classical automotive framework and the newly developed architecture. Input data has been generated by using risk parameter levels as integer input values (e.g. if the level of severity is S1 S=1). To identify the considered factors of the estimation function, the matrix of the safety integrity level framework has been investigated. The mentioned matrix includes the related parameters of severity, probability and controllability in two dimension, so practically both severity and probability are included by the lines of the table. This has inspired the contraction of the introduced two parameters in the estimation function. In light of this, during the estimation process severity (S) and probability (P) are taken into account in an

integrated form by generating risk values (R) through multiplying their scale parameters.

$$R = S \cdot P \tag{2}$$

With this, the further clustering method has used two fundamental classification parameters risk (R) and controllability (C). To conclude the main goals of the developed model, the primary objective has been to build up a framework which applies the parameters of the classical automotive integrity level architecture and which can generate a cluster structure considering the aspects of automotive safety and security in an integrated way. During the development process three different models have been developed including linear, exponential and a mixed model. The objective of the evaluation has been to analyze the efficiency of the classification models (Class) and to identify safety integrity levels unequivocally based on the function values of the classification models.

$$Class_{lin} = a \cdot R + b \cdot C \tag{3}$$

$$Class_{exp} = R^{\alpha} \cdot C^{\beta} \tag{4}$$

$$Class_{mix} = R^{\omega} + C^{\gamma} \tag{5}$$

where

a, b: linear coefficients,

α, β, ω, γ: are non-linear coefficients.

The next step to define the most efficient clustering method, the table of classical safety integrity framework (Table 5) is represented by the clustering model function in light of the interrelated input values (Table 6).

It can be seen in the presented matrix (Table 6) that in case of the applied clustering models, the input variables and the possible combination of the input parameters have discrete values. In light of this, using the presented input parameter combinations, the results of the clustering models can be systematically connected to a certain integrity level class. Accordingly, integrity levels are represented in Table 6 by the characters in the lower indexes.

To select the adequate approach for the clustering process the below presented optimization model was applied. The aim of the introduced method has been to identify the best model which results the minimum values of overlapping in comparison of the resulted cluster boundaries.

$$max: \sum_{k} Min\left(Class(R_{k+1}, C_{k+1})_{ASIL_{k+1}}\right) - \left(Max\left(Class(R_k, C_k)_{ASIL_k}\right)\right) \tag{6}$$

Table 5
Basic ASIL framework

		Level of Controllability		
Level of Severity	Likelihood of the incident	C1	C2	C3
S1	P1	QM	QM	QM
	P2	QM	QM	QM
	P3	QM	QM	ASIL A
	P4	QM	ASIL A	ASIL B
S2	P1	QM	QM	QM
	P2	QM	QM	ASIL A
	P3	QM	ASIL A	ASIL B
	P4	ASIL A	ASIL B	ASIL C
S3	P1	QM	QM	ASIL A
	P2	QM	ASIL A	ASIL B
	P3	ASIL A	ASIL B	ASIL C
	P4	ASIL B	ASIL C	ASIL D

The solution of this optimization problem provides the minimum overlapping between the sum of upper boundary of the k-th and the lower boundary of the k+1-th neighboring ASIL clusters. Based on the introduced model, it has become obvious that the mixed model can provide the best results (5) In case of this formula coefficients have been chosen ω=0,7 and γ=1,075. Based on the introduced clustering model S&SIL architecture has been be identified, and its compatibility with classical automotive integrity architecture is ensured. To define the necessary clustering parameters, first parameter is derived based on formula (2) and the second clustering parameter is identified according to formula (7).

$$C' = T \cdot D \tag{7}$$

Table 6
Result of the estimation function depending on the input variables EST (D,C)

			2nd risk parameter group (C)		
		1st risk parameter group (R)	1	2	3
S1	P1	1	CLASS (1,1)QM	CLASS (1,2)QM	CLASS (1,3)QM
	P2	2	CLASS (2,1) QM	CLASS (2,2) QM	CLASS (2,3) QM
	P3	3	CLASS (3,1) QM	CLASS (3,2) QM	CLASS (3,3) A
	P4	4	CLASS (4,1) QM	CLASS (4,2) A	CLASS (4,3) B
S2	P1	2	CLASS (2,1) QM	CLASS (2,2) QM	CLASS (2,3) QM
	P2	4	CLASS (4,1) A	CLASS (4,2) QM	CLASS (4,3) A
	P3	6	CLASS (6,1) QM	CLASS (6,2) A	CLASS (6,3) B
	P4	8	CLASS (8,1) A	CLASS (8,2) B	CLASS (8,3) C
S3	P1	3	CLASS (3,1) QM	CLASS (3,2) QM	CLASS (3,3) A
	P2	6	CLASS (6,1) QM	CLASS (6,2) A	CLASS (6,3) B
	P3	9	CLASS (9,1) A	CLASS (9,2) B	CLASS (9,3) C
	P4	12	CLASS (12,1) B	CLASS (12,2) C	CLASS (12,3) D

Formula (7) describes the integration of detectability and treatability in one unique parameter. The deliberate malicious attack or unintentional failures of automotive functions and components that has been assigned to a white colored field of the following matrix can be handled by quality management. Spotted cells describe A level functions or components, striped cells describe B level functions or components, checked cells refer to C level functions or components, while grey colored fields refer to D class functions or components (Table 7).

Table 7
Representation of S&SIL architecture

		T1			T2			T3		
		D1	D2	D3	D1	D2	D3	D1	D2	D3
B-S-PR	E1									
	E2									
	E3									
M-S-PR	E1									
	E2									
	E3									
B-S-PU	E1									
	E2									
	E3									
M-S-PU	E1									
	E2									
	E3									
M-NS	E1									
	E2									
	E3									
C-NS	E1									
	E2									
	E3									

For the following phase of the architecture development task, the aim of the investigation has been to define the accepted risk level related to the combined field of safety and security. Based on the introduced methodology, if expected number of malicious interventions and the expected likelihood of efficient interventions and estimated number of connected tools and instruments are definable, then the expected likelihood of efficient interventions can be estimated (8).

$$L = \frac{N \cdot \sum_{i=1}^{n} r_i \cdot l_i}{M} \tag{8}$$

The next step of the investigation is to identify the expected distribution of the certain intervention types. In case of our investigation, the distribution of the different intervention types is estimated to be uniform. In accordance with the mentioned expectation, the previously interpreted formula can be modified as follows:

$$L = \frac{N \cdot \sum_{i=1}^{n} \frac{l_i}{n}}{M} \tag{9}$$

In the following phase of the analysis, the estimated likelihoods of effective malicious interventions - introduced by the colleagues City University London - are applied [20]. Through estimating the likelihood of effective interventions, the expected value of effectively attacked objects can be defined. In the evaluation phase the most relevant attack surfaces of the vehicle industry are taken into account (direct intervention, intervention performed via Bluetooth connection, interventions performed via cellular network, connection and messaging and via phising and planting) [20].The number of malicious interventions is assumed to reach the eight hundred thousand in the United States of America according to the data published by Brooks statistics, at the same time, the number of online objects is expected to reach the fifty billion in the world [21]. The number of malicious interventions has been modified in accordance with the ratio of the world and the US population. In light of the interpreted aspects, it is possible to identify the annual accepted risk level of effectively implemented malicious interventions targeting online objects based on the below presented equation.

$$L = \frac{N \cdot \sum_{i=1}^{n} \frac{l_i}{n}}{M} = (8/10^5) \cdot (23/5) \cdot (0.63 + 0.52 + 0.33 + 0.76 + 0.6)/50/10^9 \tag{10}$$

In light of the introduced data describing the actual trend of the world's cybersecurity, based on the previously defined annual accepted risk level of effectively implemented malicious interventions, the hourly accepted risk level related to cyber incidents is assumed to be under 10 effective attacks. The identified hourly risk level has to be primarily interpreted as a mean value describing the society as a whole. In light of this, of course attack frequencies and their acceptance level can vary over the world. Hence, according to our interpretation the identified average value should be the central element of the developed rating scale. In accordance with this rating scale of the newly developed S&SIL architecture can be completed, and it is going to be correspond well to the classical framework of the automotive industry (Table 8):

Table 8
Saftey and security integrity levelrating scale

	ASILs	Likelihood of failure		S&SILs	Likelihood of succ. incident
1	ASIL A	$<10^{-6}$	1	S&SIL A	$<10^{-7}$
2	ASIL B	$<10^{-7}$	2	S&SIL B	$<10^{-8}$
3	ASIL C	$<10^{-7}$	3	S&SIL C	$<10^{-8}$
4	ASIL D	$<10^{-8}$	4	S&SIL D	$<10^{-9}$

On the other hand, the defined level of risk clusters shall be handled cautiously, because due to the expected quick development of the field, the number of attacks related to automated driver assistance systems is estimated to increase dynamically. The introduced trends are expected to completely modify the attitude of society to the risk of cyber threats. In accordance with this the frequency of malicious interventions, the ratio of remote interventions on vehicles are expected to increase dynamically in the close future. In light of the above-mentioned considerations the likelihood of efficient malicious interventions can achieve critical levels, which is likely to influence the sensitivity of society with regard to the combined field of safety and security in the automotive industry. Then, the operation efficiency of vehicle systems dealing with cybersecurity will need to be considerably improved. Accordingly, in case of formula (10), ten times more efficiently performed cyberattack can raise the hourly risk level to the value of 10^{-7}. This would definitely make necessary intelligent, targeted and strong international actions in the security field to make the average risk decrease under the acceptable value.

5 Summary

The aim of the article is to develop a novel integrity level framework in the combined field of automotive safety and security. In light of this, the paper discusses the most relevant terms related to the integrated field of safety and security, especially considering clustering parameters, and integrity levels. The investigation utilizes the approach of Safety Integrity Levels (SIL) to develop the novel integrity level framework suiting the combined field of automotive safety and security. In the article integrity levels are built up primarily according to the approach of the GAMAB model. It is important to emphasize that the study considers the aspects of automotive sector in an accentuated way, so the novel frame work is compatible to classical integrity level architecture of the automotive

industry [9]. To define the summarized likelihood of the possible remote interventions, the approach of Dudorov et al. is adopted [20]. This novel model takes controllability into account, as covering detectability and treatability. The novel S&SILs comprise a wide range of safety and security aspects. The last phase of this investigation was focused on current accepted risk levels of society, in the field of cybersecurity, which makes it possible to identify rating scale values, related to this newly developed S&SIL framework.

Acknowledgement

The research reported in this paper was supported by the Higher Education Excellence Program of the Ministry of Human Capacities in the frame of Artificial Intelligence research area of Budapest University of Technology and Economics (BME FIKP-MI/FM).

References

[1] Alzenad, M., Shakir, M. Z., Yanikomeroglu, H., & Alouini, M. S. (2018) FSO-based vertical backhaul/fronthaul framework for 5G+ wireless networks. IEEE Communications Magazine, 56(1), 218-224

[2] Briand, L. C., Labiche, Y., and Liu, X., "Using machine learning to support debugging with tarantula," In Proceedings of the 18th IEEE International Symposium on Software Reliability, Washington DC, USA, 2007, pp. 137-146

[3] Rosenberg, I., Shabtai, A., Elovici, Y., & Rokach, L. (2018) Low Resource Black-Box End-to-End Attack Against State of the Art API Call Based Malware Classifiers. arXiv preprint arXiv:1804.08778

[4] Baldini, G., Botterman, M., Neisse, R., & Tallacchini, M. (2018) Ethical design in the internet of things. Science and engineering ethics, 24(3), 905-925

[5] Macher, G., Armengaud, E., Kreiner, C., Brenner, E., Schmittner, C., Ma, Z., Martin, H., Krammer, M. (2018) Integration of security in the development lifecycle of dependable automotive CPS. In Solutions for Cyber-Physical Systems Ubiquity (pp. 383-423). IGI Global

[6] Szalay, Z., Tettamanti, T., Esztergár-Kiss, D., Varga, I., & Bartolini, C. (2018) Development of a Test Track for Driverless Cars: Vehicle Design, Track Configuration, and Liability Considerations. Periodica Polytechnica Transportation Engineering, 46(1), 29-35

[7] Zöldy M. (2018) Investigation of Autonomous Vehicles fit into Traditional Type Approval Process, Procedings of ICTTE 2018 Beograde, pp. 428-432

[8] Bellotti, M., & Mariani, R. (2010) How future automotive functional safety requirements will impact microprocessors design. Microelectronics Reliability, 50(9-11), 1320-1326

[9] Chang, Y. C., Huang, L. R., Liu, H. C., Yang, C. J., & Chiu, C. T. (2014, April) Assessing automotive functional safety microprocessor with ISO 26262 hardware requirements. In VLSI Design, Automation and Test (VLSI-DAT), 2014 International Symposium on (pp. 1-4) IEEE

[10] Jaskolka, J., & Khedri, R. (2016) Mitigating covert channels based on analysis of the potential for communication. Theoretical Computer Science, 643, 1-37

[11] Cheah, M., Shaikh, S. A., Bryans, J., & Wooderson, P. (2018) Building an automotive security assurance case using systematic security evaluations. Computers & Security, 77, 360-379

[12] Gheraibia, Y., Djafri, K., & Krimou, H. (2018) Ant colony algorithm for automotive safety integrity level allocation. Applied Intelligence, 48(3), 555-569

[13] Guiochet, J., Machin, M., & Waeselynck, H. (2017) Safety-critical advanced robots: A survey. Robotics and Autonomous Systems, 94, 43-52

[14] Miller, C., Kassie, J., & Poston, D. (2017) Assessing and Computing the Safety Integrity Level (SIL) for Turbo Machinery Protection. In Proceedings of the 46th Turbomachinery Symposium. Turbomachinery Laboratory, Texas A&M Engineering Experiment Station

[15] Kaczor, G., Młynarski, S., & Szkoda, M. (2016) Verification of safety integrity level with the application of Monte Carlo simulation and reliability block diagrams. Journal of Loss Prevention in the Process Industries, 41, 31-39

[16] Sebron, W., Tschürtz, H., & Krebs, P. (2018, September) The Shell Model–A Method for System Boundary Analysis. In European Conference on Software Process Improvement (pp. 68-79) Springer, Cham

[17] Kreiner, C., & Messnarz, R. (2018) Effective Approaches to Training CPS Knowledge and Skills. In Solutions for Cyber-Physical Systems Ubiquity (pp. 111-135). IGI Global

[18] Fülep, T. (2007) Design methods of safety-critical systems and their application in electronic brake systems

[19] Chuck Brooks (2017) Keep Calm and… Here Is a List of Alarming Cybersecurity Statistics. URL: https://www.itspmagazine.com/from-the-newsroom/keep-calm-and-here-is-a-list-of-alarming-cybersecurity-statistics

[20] Dudorov, D., Stupples, D., & Newby, M. (2013, August) Probability analysis of cyber attack paths against business and commercial enterprise systems. In Intelligence and Security Informatics Conference (EISIC), 2013 European (pp. 38-44) IEEE

[21] E. Commission, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on ENISA, the ”EU

Cybersecurity Agency", and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification ("Cybersecurity Act"), 2017

[22] Guerrero-Higueras, Á. M., DeCastro-García, N., Rodríguez-Lera, F. J., & Matellán, V. (2017) Empirical analysis of cyber-attacks to an indoor real time localization system for autonomous robots. Computers & Security, 70, 422-435

[23] Tarnai G., Sághi B.: Hazard and Risk Analysis of Human-Machine Interfaces of Railway Interlocking Systems. 7th World Congress on Railway Research. 4-8.June, 2006, Montréal, Canada

Active and Reactive Power Losses in Distribution Transformers

Michal Kolcun[1], Anna Gawlak[2], Miroslaw Kornatka[2], Zsolt Čonka[1]

[1] Technical University of Košice, Department of Electric Power Engineering, Mäsiarska 74, 040 01 Košice, Slovakia; michal.kolcun@tuke.sk; zsolt.conka@tuke.sk

[2] Technical University of Częstochowa, Faculty of Electrical Engineering, Al. Armii Krajowej 17, 42-200 Częstochowa, Poland; gawlak@el.pcz.czest.pl; kornatka@el.pcz.czest.pl

Abstract: The problem of energy quality has recently gained much recognition, one of the reasons being, that there is an increasing number of devices which require energy that meets high quality standards. An improvement in this respect, can be achieved by effective management of reactive power flow in the power system. Maintaining balance in active and reactive power is of key importance for the flawless functioning of the power system. This paper discusses theoretical issues underpinning calculations of active, reactive power and of energy loss in MV/LV transformers. Based on the parameters of transformers and data from consumer meters on active and reactive power, active and reactive power and energy loss was obtained, with the view to assess the efficiency of active power and energy transfer through MV/LV transformers.

Keywords: modelling; energy losses; active and reactive energy

1 Introduction

The flow of current through the elements of a power network, is accompanied by power and energy losses. Loss occurring at resistive elements is known as active loss, whereas the loss occurring at reactance elements is known as reactive loss. Both kinds of loss are detrimental due to a number of reasons: the extra amount of energy has to be generated in power plants, which requires additional devices and resources; the extra energy has to be transferred through all of the network elements, which requires increasing their transfer capacity; on Joule-Lenz law, energy turns into heat, which affects the dimensions of the network elements [1-4]. The consequences of losses vary depending on their type and place of their occurrence. As regards transformers, the key element is the core, which is

magnetized by means of inductive power, constituting reactive idle loss. This kind of loss can be greater than load loss and in the case of low power transformers, the ratio of idle to load loss can be about 4-5. Reactive idle loss in transformer cores is compensated for by means of capacitor batteries connected directly to the transformer [5-9]. The power of such capacitors should be equal to the power of the rated idle loss. Even though no energy resources are consumed for the generation of reactive power, which does not yield any work, it flows through the power system increasing its load. By this token, the detrimental effect of reactive power is multiplied. It is typically assessed by an equivalence coefficient of reactive power, by means of which it is possible to calculate the amount of active loss per unit of reactive power. The value of the coefficient depends on the location of the occurrence of reactive power. According to The Energy Efficiency Act [10], reducing reactive energy loss (art.19 par.1 pt. 5a) is one of the moves that should be implemented towards improving energy efficiency. In this respect, special attention should be paid to power transformers, which are one of the key elements of the power system [4, 7, 11]. Energy losses occurring in MV/LV transformers constitute from about 30 to 60% of total losses in LV networks [12-14]. At present, MV/LV transformers are typically not metered, but since the number of smart meters installed at end-consumers is continuously increasing, it is possible to measure the amount of energy flowing through MV/LV transformers with large precision [15-17].

2 Measurements of Energy Losses in a Two-Winding Transformer

Fig. 1 presents the measurement system, including an electronic meter of electrical energy, installed at the lower-voltage side of a two-winding transformer.

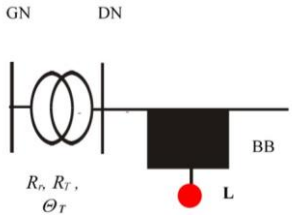


Figure 1
Measurement at the lower voltage side of a transformer

The following symbols are used in Fig. 1:

– GN – higher-voltage side

– DN – lower-voltage side

- BB – "black box" (voltage and current transformers); the following quantities are given for the transformers:
  - ϑ_U – winding ratio of the voltage transformer
  - ϑ_I – winding ratio of the current transformer
- **L** – electronic meter
- R_r – idle resistance of the transformer
- R_T – rated load resistance of the transformer
- Θ_T –winding ratio of the transformer, equal to:

$$\Theta_T = \frac{U_{GN}}{U_{DN}}(1+0{,}01\cdot\Delta z) \tag{1}$$

where: U_{GN} – rated voltage of the transformer primary winding, U_{DN} – rated voltage of the transformer secondary winding, Δz – percentage correction of the transformer winding ratio.

The meter records [19] the following data:

- *power profile*, including active consumed power $\{P+\}_{Tn}$, reactive consumed (inductive) power $\{Q_L+\}_{Tn}$, reactive returned (capacitive) power $\{Q_C-\}_{Tn}$ at the n-th calculation period T_n, and possibly also active returned power $\{P-\}_{Tn}$, reactive returned (inductive) power$\{Q_L-\}_{Tn}$, and reactive consumed (capacitive) power $\{Q_C+\}_{Tn}$),
- *energy losses* at the end of the n-th calculation period for the z-th zone, including losses of the active consumed energy SP_{nz}, of reactive consumed (inductive) energy $SQ_{L,nz}$, reactive returned (capacitive) energy $SQ_{C,nz}$, and possibly also energy states corresponding to the power profile, as enumerated above; n – index of the calculation period, z = 1, 2, …, Z, Z – the number of zones within 24 hours,
- *indications of the loss counters*: $S(V^2h)_n$, $S(I^2h)_n$ at the end of a n-th calculation period without dividing it into time zones. The values are given for the whole 24-h period and the meter does not register, unfortunately, the indications of the loss counters for each measuring cycle.

On the basis of the *indications of the loss counters,* the following values are obtained for the current n-th calculation period:

$$\Delta S(V) = S(V^2h)_n - S(V^2h)_{n-1} \tag{2}$$

$$\Delta S(I) = S(I^2h)_n - S(I^2h)_{n-1} \tag{3}$$

Let us turn to active energy. At the resistance R_r there are voltage energy losses and at the resistance R_{TN} there are load energy losses. The two resistances are obtained from the formulas:

$$R_r = \frac{U_{GN}^2 \cdot 10^3}{\Delta P_{FeN}} \tag{4}$$

$$R_T = \Delta P_{CuN} \frac{U_{GN}^2}{S_N^2} \cdot 10^3 \tag{5}$$

where: ΔP_{FeN} – rated loss of the active power in the transformer core, kW, ΔP_{CuN} – rated loss of the active power in the transformer windings, kW, S_N – rated power of the transformer, kVA, U_{GN} – rated voltage (see above), kV.

The voltage loss of active energy occurs in accordance with the formula (2):

$$U_{D(h)}^2 = \Delta S(V) \vartheta_U^2 \text{ , [V}^2\text{ h]} \tag{6}$$

$$U_{G(h)}^2 = \Delta S(V) \vartheta_U^2 \Theta_T^2 \text{ , [V}^2\text{ h]} \tag{7}$$

As follows from the formula $\Delta P_{FeN} = \frac{U_{GN}^2 \cdot 10^3}{R_r}$, the voltage energy loss ΔE_U is equal to (cf. (7))

$$\Delta E_U = \frac{\Delta S(V) \vartheta_U^2 \Theta_T^2}{R_r \cdot 10^3} \text{ , [kWh]} \tag{8}$$

After substituting (4) into (8)

$$\Delta E_U = \frac{\Delta S(V) \vartheta_U^2 \Theta_T^2 \Delta P_{FeN}}{U_{GN}^2 10^6} \text{ , [kWh]} \tag{9}$$

The load (current) loss of the active energy occurs in accordance with (cf. (3))

$$I_{D(h)}^2 = \Delta S(I) \vartheta_I^2 \text{ , [A}^2\text{h].} \tag{10}$$

On the basis of (10) and by analogy to $\Delta P_{CuN} = 3 I_N^2 R_T \cdot 10^{-3}$, the current energy loss ΔE_I equals

$$\Delta E_I = \Delta S(I) \vartheta_I^2 R_{T(DN)} \cdot 10^{-3} \text{ , [kWh]} \tag{11}$$

where: $R_{T(DN)}$ – load resistance of the transformer at the voltage DN, equal to (cf. (1), (5))

$$R_{T(DN)} = R_T \cdot \frac{1}{\Theta_T^2} = \Delta P_{CuN} \frac{U_{GN}^2}{S_N^2} \cdot \frac{10^3}{\Theta_T^2} \tag{12}$$

After substituting (12) into (11)

$$\Delta E_I = \Delta S(I) \vartheta_I^2 \Delta P_{CuN} \frac{U_{GN}^2}{S_N^2} \cdot \frac{1}{\Theta_T^2}, \text{ [kWh]} \tag{13}$$

Active energy in the z-th zone equals

$$E(P)_z = (SP_{nz} - SP_{n-1,z})M, \text{ [kWh]} \tag{14}$$

where: M – constant (multiplier) of the meter, depending on the winding ratio: ϑ_U, ϑ_I.

The total energy is a summation (cf. (14))

$$E(P) = \sum_{z=1}^{Z} E(P)_z \tag{15}$$

Electronic meters equipped with loss recording modules are typically installed at a large energy user's, with measurements performed at the DN (lower voltage) circuit of a transformer. The consumer pays, among other things, for energy consumed and energy loss occurring in the transformer, the tariffs being different for the particular zones. Therefore, the problem arises how to assign voltage and current energy losses to the particular zones. The total loss should be distributed correctly over the time zones, in order to add an appropriate value to the energy obtained (14).

3 Calculating Power/Energy Losses in a Transformer on the Basis of Load

A meter installed at the lower-voltage side of the transformer takes measurements at 15-minute intervals. On the basis of these measurements, mean active and reactive power, i.e. the power profile is energy QL(+)t, reactive capacitive returned energy QC(–)t, reactive inductive returned energy QL(–)t, reactive capacitive consumed energy QC(–)t can be obtained [15]:

$$P_t = P(+)_t - P(-)_t;\ Q_t = Q_L(+)_t - Q_C(-)_t - Q_L(-)_t + Q_C(+)_t \tag{16}$$

Active energy in the h-th hour is equal to the *active power* $\overline{P}_h$ averaged over an hour

$$\overline{P}_h = \sum_{t(h)=k(h)-3}^{k(h)} P_{t(h)} \tag{17}$$

and reactive energy in the *h*-th hour is equal to the *reactive power* $\overline{Q}_h$ averaged over an hour

$$\overline{Q}_h = \sum_{t(h)=k(h)-3}^{k(h)} Q_{t(h)} \tag{18}$$

where: *k(h)* – number of the last quarter of the *h*-th hour, *h* – number of the hour, *h* = 1, 2,…, *H*, *H* – number of hours analyzed in the period under scrutiny.

Active power P_t and reactive power Q_t for the *t*-th 15 minute cycle are, respectively:

$$P_t = 3U_{ft} I_t \cos\varphi_t = 3U_{ft} I_{ct} \qquad Q_t = 3U_{ft} I_t \sin\varphi_t = 3U_{ft} I_{bt} \tag{19}$$

where: U_{ft} –phase voltage for the *t*-th cycle, I_{ct}, I_{bt} active and reactive current components for the *t*-th cycle, respectively, equal to:

$$I_{ct} = \frac{P_t}{3U_{ft}} \qquad I_{bt} = \frac{Q_t}{3U_{ft}} \tag{20}$$

In a 60-minute period consisting of four 15-minute cycles, the active component I_c and the reactive component I_b of the current are (cf. (20)):

$$I_c = \sum_t I_{ct} = \frac{1}{3}\sum_t \frac{P_t}{U_{ft}} \qquad I_b = \sum_t I_{bt} = \frac{1}{3}\sum_t \frac{Q_t}{U_{ft}} \tag{21}$$

and their resultant sum I is

$$I = \sqrt{I_c^2 + I_b^2} \tag{22}$$

Substituting (21) into (22)

$$I = \frac{1}{3}\sqrt{(\sum_t \frac{P_t}{U_{ft}})^2 + (\sum_t \frac{Q_t}{U_{ft}})^2} \tag{23}$$

On the basis of mean power/hour (por. (17), (18); the index *h* was dropped), the apparent power is obtained

$$S = \sqrt{\overline{P}^2 + \overline{Q}^2} \tag{24}$$

On the basis of (24) and the formula $S = \sqrt{3}\overline{U}I$, the *mean line-to-line voltage in an hour* is obtained

$$\overline{U} = \frac{\sqrt{3[(\sum_t P_t)^2 + (\sum_t Q_t)^2]}}{3I} \tag{25}$$

which, after (23) is taken into account, is equal to

$$\overline{U} = \sqrt{\frac{3[(\sum_t P_t)^2 + (\sum_t Q_t)^2]}{(\sum_t \frac{P_t}{U_{ft}})^2 + (\sum_t \frac{Q_t}{U_{ft}})^2}} \tag{26}$$

For each MV/LV transformer the following quantities are known:

1) Rated values: S_N, ΔP_{FeN}, ΔP_{CuN}, $\Delta u_{z\%}$, $I_{0\%}$, U_{GN}, U_{DN} and the windings ratio correction $\Delta z_{\%}$, performed by means of a tap changer,

2) DT_h (date and time, i.e. the timestamp), $\overline{P}_h$, $\overline{Q}_h$, $\overline{U}_h$ (B.11), $h = 1, 2, ., H$.

Below a method of calculating power/energy losses in an hour-period is presented for an r-th transformer (r = 1, 2, …, R, R – number of transformers under scrutiny; a transformer index is omitted in further discussion). The following losses are obtained:

ΔP_{Uh} – voltage loss of active power in the transformer core

ΔP_{Ih} – current loss of active power in the transformer windings

ΔQ_{Uh} – voltage loss of reactive power in the transformer core

ΔQ_{Ih} – current loss of reactive power in the transformer windings

The voltage loss of active power in the transformer core is obtained from the formula:

$$\Delta P_{Uh} = \Delta P_{FeN} \frac{\overline{U}_{Gh}^2}{U_{GN}^2} \tag{27}$$

where: $\overline{U}_{Gh}$ –mean voltage in the *h*-th hour at the primary side of the transformer, obtained from the formula:

$$\overline{U}_{Gh} = \overline{U}_h(1 + 0{,}01 \cdot \Delta u_{T\%h})\Theta \tag{28}$$

where: $\overline{U}_h$ – voltage in the *h*-th hour at the secondary side of the transformer, $\Delta u_{T\%h}$ – relative voltage drop in the transformer for the *h*-th hour, Θ – real voltage winding ratio of the transformer (constant throughout the analysis), equal to:

$$\Theta = \frac{U_{GN}}{U_{DN}}(1+0{,}01\cdot\Delta z_{\%})$$

The relative voltage drop in the transformer for the *h*-th hour is obtained from the formula:

$$\Delta u_{T\%h} = (\Delta u_{RN\%}\cos\varphi_h + \Delta u_{XN\%}\sin\varphi_h)\beta_h \tag{29}$$

where:

$$\beta_h = \frac{\sqrt{\overline{P}_h^2 + \overline{Q}_h^2}}{S_N} \tag{30}$$

is the load coefficient of the transformer.

The current loss of active power in the transformer windings ΔP_{Ih} is obtained from the formula:

$$\Delta P_{Ih} = \Delta P_{CuN}\beta_h^2 k_{Th} \tag{31}$$

where: k_{Th} – *temperature coefficient,* allowing for the change in the resistance in the transformer windings depending on the load, equal [3] to:

$$k_{Th} = 0.3179\cdot\beta_h^2 - 0.082\cdot\beta_h + 0.77 \tag{32}$$

The voltage loss of reactive power in the transformer core ΔQ_{Uh} is obtained from the formula:

$$\Delta Q_{Uh} = \sqrt{(0.01\cdot I_{0\%}S_N)^2 - \Delta P_{FeN}^2}\cdot\frac{\overline{U}_{Gh}^2}{U_{GN}^2} \tag{33}$$

The current loss of reactive power in the transformer windings ΔQ_{Ih} is obtained from the formula:

$$\Delta Q_{Ih} = 0.01\cdot\Delta u_{XN\%}S_N\beta_h^2 \tag{34}$$

4 Results

The analysis was carried out on the basis of data from the year 2016, obtained from a distribution company in Poland. The data from consumer meters with active and reactive load for one year periods and transformers MV / LV were collected. The MV / LV transformer load data is shown in Table 1. Using the data on recipients' loads, the active energy load in transformers was calculated for every month of the year on the basis of formula (17) and reactive – on the basis of formula (18).

Table 1. Active and reactive energy flowing through MV/LV transformers for every month of the year.

Table 1
Active and reactive energy flowing through MV/LV transformers for every month of the year

	Active energy [MWh]			Reactive energy [Mvarh]		
	160 [kVA]	250 [kVA]	400 [kVA]	160 [kVA]	250 [kVA]	400 [kVA]
I	341	869	1847	74	145	400
II	294	739	1601	66	129	373
III	294	758	1610	73	140	397
V	272	678	1466	75	130	389
V	271	675	1419	82	145	409
VI	256	624	1304	83	143	390
VII	242	640	1361	80	183	470
VIII	246	640	1358	81	182	463
IX	250	674	1425	68	147	417
X	272	744	1625	69	154	469
XI	286	779	1681	67	148	424
XII	340	875	1824	75	148	404

On the basis of the data, voltage and current losses of active and reactive energy were calculated. Results for the particular months are shown in Table 2. Voltage loss of active energy is obtained from formula (27) and reactive from formula (33). Current loss of active energy is obtained from formula (31) and reactive from formula (34).

As can be seen in Table 2 reactive energy loss in transformers is significantly higher than active energy loss. Besides, voltage losses, both of active and reactive energy are higher than current losses. The low amount of current losses is due to relatively low load coefficients of MV/LV transformers, varying from 0.138 for the group 160 kVA to 0.148 for the group 400 kVA.

Table 2
Active and reactive energy losses flowing through MV/LV transformers

	Power kVA]	Number of transf.	β	Active energy losses		Reactive energy losses	
				Voltage [MWh]	Current [MWh]	Voltage [Mvarh]	Current [Mvarh]
I	160	18	0.163	4.8	0.90	48.3	3.2
	250	30	0.158	15.5	2.17	111.0	8.8
	400	47	0.172	23.3	5.90	194.7	22.9
II	160	18	0.156	4.4	0.75	43.7	2.7
	250	30	0.149	14.0	1.73	100.2	7.1
	400	47	0.165	21.1	4.86	176.2	19.0
III	160	18	0.142	4.8	0.70	48.3	2.5
	250	30	0.138	15.4	1.69	110.8	6.9
	400	47	0.151	23.2	4.52	194.2	17.8
IV	160	18	0.136	4.7	0.65	46.9	2.3
	250	30	0.128	14.8	1.44	106.2	5.9
	400	47	0.142	22.6	3.86	188.8	15.3
V	160	18	0.132	4.9	0.64	48.6	2.3
	250	30	0.124	15.7	1.39	112.3	5.7
	400	47	0.134	23.7	3.52	198.8	14.0
VI	160	18	0.130	4.7	0.60	47.0	2.2
	250	30	0.119	15.2	1.28	108.7	5.2
	400	47	0.128	23.0	3.14	192.5	12.5
VII	160	18	0.119	4.8	0.56	48.3	2.0
	250	30	0.119	15.5	1.35	111.1	5.5
	400	47	0.131	23.3	3.27	195.2	13.1
VIII	160	18	0.121	4.9	0.59	48.3	2.1
	250	30	0.119	15.5	1.35	111.1	5.5
	400	47	0.130	23.3	3.31	195.1	13.2
IX	160	18	0.125	4.7	0.58	46.9	2.1
	250	30	0.128	15.0	1.43	107.7	5.8
	400	47	0.139	22.6	3.60	189.2	14.3
X	160	18	0.131	4.9	0.64	48.4	2.3
	250	30	0.136	15.5	1.66	111.1	6.8
	400	47	0.153	23.4	4.57	195.5	18.0
XI	160	18	0.142	4.7	0.71	46.7	2.5
	250	30	0.147	15.0	1.85	107.4	7.6
	400	47	0.163	22.6	5.09	189.0	19.9
	160	18	0.163	4.8	0.92	48.3	3.3

XII	250	30	0.159	15.4	2.25	110.6	9.1
	400	47	0.170	23.2	5.77	193.8	22.5
Year			**0.142**	**516**	**79**	**4182**	**313**

Using the formula:

$$\eta_E = \frac{E_c}{E_c + \Delta E_{cU} + \Delta E_{cI}}$$

the average efficiency of active energy transfer was obtained. The results are presented in Fig. 2.

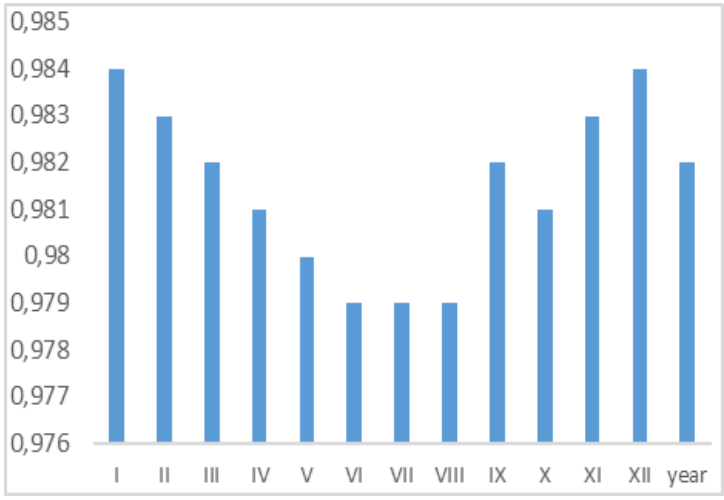


Figure 2
Average efficiency of energy transfer

The highest efficiency of energy transfer in MV/LV transformers occurs in winter and the lowest in summer. This is connected, among other factors, with voltage losses, which are practically constant throughout the year, so when the amount of energy flowing through transformers is lower as it is in summer, the network efficiency is also lower.

Extremely low transformer load coefficients generate high active and reactive energy loss. Taking this into consideration, a simulation experiment was carried out for the period of a year, in which the transformer load was adjusted to the amount of energy flowing through them. The results are presented in Table 3.

Table 3
Active and reactive energy loss after the transformer power has been adjusted to load

Current condition						
Power [kVA]	Number of transf.	β	Active energy losses		Reactive energy losses	
			Voltage [MWh]	Current [MWh]	Voltage [Mvarh]	Current [Mvarh]
160	18	0.138	57.723	8.291	570.299	29.972
250	30	0.135	183.208	19.642	1308.847	80.583
400	47	0.148	275.788	51.450	2303.423	203.312
total	516.719	79.383	4182.569	313.867	516.719	79.383
Transformer power adjusted to load						
Power [kVA]	Number of transf.	β	Active energy losses		Reactive energy losses	
			Voltage [MWh]	Current [MWh]	Voltage [Mvarh]	Current [Mvarh]
63	22	0.351	31.910	9.2199	352.674	19.977
75	31	0.451	95.76	75.997	580.455	191.412
160	42	0.37	118.652	106.546	1172.281	385.164
total	246.323	191.762	2105.410	596.552	246.323	191.762

It follows that when the transformer power was adjusted to the load, the active energy loss was reduced by 26.5%, and the reactive energy loss by 40%.

Conclusions

It has been demonstrated that transformers are characterized by a low load coefficient, which is evidenced by high voltage energy loss, significantly exceeding current energy loss.

The highest value of the power coefficient ($cos\varphi$), equal to 0.952 ($tg\varphi = 0.323$) was attested in December and the lowest value equal to 0.916 ($tg\varphi = 0.437$) in June. For the whole year, these values for the primary transformer winding are (respectively) $cos\varphi = 0.932$, $tg\varphi = 0.390$. The relatively high power coefficient is affected by the batteries of parallel capacitors connected to the low voltage side of MV/LV transformers as compensation for the idle state.

The study has found that the low efficiency of energy distribution is caused by low load in MV/LV transformers. The extremely low load coefficient contributes to high reactive energy loss, thereby lowering the network efficiency. It follows that transformer power should be adequately adjusted to the load, which will significantly reduce the energy losses.

Acknowledgement

The Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences under the contract no. VEGA 1/0372/18 supported this work.

References

[1] Nawrowski, R., Stein, Z., and Zielińska, M.: Analiza wpływu przekraczania dopuszczalnych wartości współczynnika mocy w sieci nn na pracę systemu elektroenergetycznego, Electrical Engineering, No. 74, 2013, pp. 111-117

[2] Pozna, ., Precup, R., Tar, J., Škrjanc, I., Preitl, S.: New results in modelling derived from Bayesian filtering, Knowledge-Based Systems, Vol. 23, No. 2, 2010, pp. 182-194

[3] Niewiedział, E., and Niewiedział, R.: Problematyka strat mocy i energii w transformatorach rozdzielczych SN/nn, Electro.info 10/2017

[4] Gawlak, A.: Technological aspects of electrical energy distribution. 14th International Scientific Conferenc Electric Power Engineering 2014, May 2014, Brno Czech Republic ISBN 978-80-214-4514-7

[5] Zajkowski, K.: Analiza szacunkowa w audycie energetycznym rozliczająca działania zmniejszające przepływy mocy biernej w sieci energetycznej, Logistyka 6/2014

[6] Ürmös, A., Farkas, Z., Farkas, M., Sándor, T., Kóczy, L.T., and Nemcsics, A.: Application of self-organizing maps for technological support of droplet epitaxy, Acta Polytechnica Hungarica, Vol. 14, No. 4, 2017, pp. 207-224

[7] Kolcun, M., Kornatka, M., Gawlak, A., and Čonka, Z.: Benchmarking the reliability of medium-voltage lines, Journal of Electrical Engineering, Vol. 68 (3), 2017 r, pp. 212-215

[8] Ürmös, A., Farkas, Z., Farkas, M., Sándor, T., Kóczy, L. T., and Nemcsics, A.: Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants, International Journal of Artificial Intelligence, Vol. 16, No. 2, 2018, pp. 208-222

[9] Bielecki, S.:Ookreślanie strat powodowanych obciążeniem mocą bierną – metoda nie wykorzystująca pojęcia energetycznego równoważnika mocy biernej, Przegląd Elektrotechniczny, issn 0033-2097, r. 94 nr 9/2018

[10] Ustawa z dnia 20 maja 2016r. o efektywności energetycznej. Dz.U. z 2016 r. poz. 831

[11] Precup, R., and Preitl, S.: Stability and sensitivity analysis of fuzzy control systems. Mechatronics applications, Acta Polytechnica Hungarica, Vol. 3, No. 1, pp. 61-76, 2006

[12] Gawlak, A.: Analysis of technical losses in the low and medium voltage power network. 11th International Scientific Conference Electrical Power Engineering - EPE 2010, Brno Czech Republic, ISBN 978-80-214-4094-4 pp. 119-123)

[13] Gawlak, A.: The Influence of Investment on Reducing Energy Losses in Distribution Networks, in Proc. 16th International Scientific Conference on Electric Power Engineering, 2015r, pp. 315-319

[14] Gawlak, A.: Podział środków inwestycyjnych na rozwój sieci rozdzielczych przy zastosowaniu metody taksonomicznej, Przegląd Elektrotechniczny, R.85 nr 3, 2009 r, pp. 57-160

[15] Kornatka, M., and Gawlak, A.: Comparative Analysis of Operating Conditions in Polish Medium-voltage and 110 kV Networks, in Proc. 8th International Scientific Symposium on Electrical Power Engineering, 2015 r, pp. 57-60

[16] Gono, M., Knycl, M., Gono, R., and Kłosok-Bazan, I.: Experience with the production of electricity from biogas at sewage treatment plant in the Czech Republic, Przeglad Elektrotechniczny, Volume 89, Issue 11, 2013 r, pp. 12-15

[17] Gawlak, A.: Noninvestment Forms of Reducing Energy Losses in Distribution Networks, in Proc. 8th International Scientific Symposium on Electrical Power Engineering, 2015, pp. 61-64

[18] Czepiel, S.: Transfiguration of Supplied-by-HV/MV Transformer Network to the Supply Radius, Electrical Power Quality and Utilization, Journal, ISSN 1234-6799, Volume XIV – Number 1 – 2008

A Repulsive Interaction in Classical Electrodynamics

Katalin Gambár[1], Mario C. Rocca[2], Ferenc Márkus[3]

[1]Institute of Microelectronics and Technology, Kálmán Kandó Faculty of Electrical Engineering, Óbuda University, Tavaszmező u. 17, H-1084 Budapest, Hungary
[2]Departamento de Física, Faculdad de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 67, 1900 La Plata, Argentina
[3]Deparment of Physics, Budapest University of Technology and Economics, Budafoki út 8, H-1521 Budapest, Hungary

gambar.katalin@kvk.uni-obuda.hu, rocca@fisica.unlp.edu.ar, markus@phy.bme.hu

Abstract: Herein, we introduce an additional term into the induction equation (one of the Maxwell's equation). The related Lagrangian formalism applying the scalar and vector potentials is fitted to this modified Maxwell's equations. In the framework of Hamiltons's principle we are able to deduce Klein-Gordon equations with negative "mass term" for the field variables electric field ***E*** *and magnetic induction* ***B****. We can conclude from the mathematical structure of the equations that a repulsive interaction appears. The Wheeler propagator can be calculated for the present case by which the time evolution of the field can be discussed. In spite of the situation that these equations have tachyon solutions, the results are in line with the causality principle. As a consequence of the theory, a spontaneous charge disjunction process may rise in the field.*

Keywords: Maxwell's equations; Klein-Gordon equation with negative "mass term"; Lagrangian, Wheeler propagator; charge distribution

1 Introduction

Mechanical [1, 2], thermodynamic [2-4] and further field theoretical examples [5-7] for the Klein-Gordon equation with negative "mass term" involve the same dynamical phase transition that operates between the diffusive and the wave type dynamics. Studying the existence of these kinds of phenomena we may assume that the occurrence of these are more general and not restricted exclusively to a certain part of physics.

In the last decades a wider study of the negative mass term Klein-Gordon equation has been accomplished to get a detailed impression about the efficiency and the validity of this formulation [8-14]. This kind of examination is not a gratuitous mathematical whim at all, because there exist more realistic processes which are described by such equations. Thus, there is no reason to doubt their reality [15].

The mechanical example [1, 2] is a stretched string lying on the diameter of a rotating disk when the appearing centrifugal force behaves as a repulsive interaction as Fig. 1 shows.

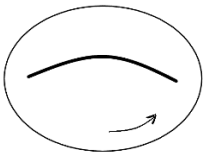


Fig. 1
Stretched string of a rotating disk

The model is covered by the equation:

$$\frac{\partial^2\Psi}{\partial t^2} - \frac{F}{\rho A}\frac{\partial^2\Psi}{\partial x^2} - \omega_0^2\Psi = 0 \tag{1}$$

where Ψ is the displacement from the equilibrium position, ρ is the mass density of the string, F is the stretching force, A is the cross section of the string and ω_0 is the angular velocity of the disk. The second term pertains to the spring force, i.e., it is an attractive interaction. The so-called negative "mass term" is the third term of this equation due to the negative sign of the term. (This term is positive in the “well behaved” Klein-Gordon equation [16, 17]). The change of dynamics can be understood form the following physical picture. If the angular velocity ω_0 is small enough, the spring vibrates around its equilibrium position, but above a certain threshold angular velocity the centrifugal force elongates the spring towards the bigger radius without vibration. This is a transition between the vibrating and the dissipative state, i.e., this a dynamic phase transition. The detailed studies can be done by the examination of dispersion relation [1, 2]

$$\omega(k,\omega_0) = \sqrt{\frac{F}{\rho A}k^2 - \omega_0^2}$$

Waves modes exist if

$$\frac{F}{\rho A} > \frac{\omega_0^2}{k^2}$$

and there are no wave modes if

$$\frac{F}{\rho A} < \frac{\omega_0^2}{k^2}$$

It is clear that the third term in Eq. (1) behaves as a repulsive interaction.

A much more interesting dynamic transition can be found in between wave like thermal propagation and the Fourier heat conduction. To achieve the aim, first, the equation of motion of Lorentz invariant thermal energy propagation [2-7] must be formulated:

$$\frac{1}{c^2}\frac{\partial^2 \mathrm{T}}{\partial t^2} - \frac{\partial^2 \mathrm{T}}{\partial x^2} - \frac{c^2 c_v^2}{4\lambda^2} T = 0 \tag{2}$$

It is obvious that the structure of this equation – aside from the meaning of the parameters – is the same as in Eq. (1). Here, c is the speed of light; c_v is the heat capacity and λ is the heat conductivity, both of them are constant parameters, now. This equation implies the case of classical Fourier heat conduction [18]

$$c_v \frac{\partial T}{\partial t} - \lambda \frac{\partial^2 T}{\partial x^2} = 0$$

which solution is separated from the wave-like propagation by a dynamic phase transition via a spinodal instability [19, 20]. This transition is also describable via the dispersion relation [3, 4]

$$\omega(k) = \sqrt{c^2 k^2 - \frac{c^4 c_v^2}{4\lambda^2}}$$

The propagation is wave-like if:

$$k > \frac{c\, c_v}{2\lambda}$$

and dissipative if:

$$k > \frac{c\, c_v}{2\lambda}$$

Similarly, to the mechanical example, the third term introduces a certain repulsive interaction in the thermal energy propagation.

The present work is based on the idea of the previous examples. It seems interesting to formulate a Klein-Gordon type equation with the above mathematical structure – with negative "mass term" – to electrodynamics problems and to examine what kind of physical process may be governed by this way. Now, the whole description remains within the framework of the classical electrodynamics [21, 22]. It will be shown that to achieve this aim we should add an extra term to one of the Maxwell's equations (see later Eq. (3b). The Lorentz invariance of the theory can be completed by the careful choice of this term. As the examples show, the meaning of the appeared term in the Klein-Gordon type equation should be a repulsive interaction. In the light of our knowledge by calculating the propagator of the process [8, 9] we point out that this repulsive interaction may cause a large spontaneous charge disjunction in the electric conductive medium.

We mention that in several areas of modelling of natural processes brave steps are needed, since the phenomena include not negligible – but not obviously deducible – additional interactions [23-26].

The structure of the paper is as follows. In Chapter 2 the mathematical construction is elaborated that leads to a repulsive interaction in the electrodynamics. The method is based on this new idea to add an extra term to one of the Maxwell's equations. Finally, a negative mass term Klein-Gordon equation is obtained. As the propagators can generate from an initial state to a final state of a system in time, to achieve to time evolution of the electric and magnetic fields, and mainly the charge density, the so-called Wheeler propagator is needed calculate for this kind of Klein-Gordon equation. Applying the result of previous studies [8-14] this calculation is shown in Chapter 3. The evolution of electric field and the charge density is calculated analytically in Chapter 4, while a numerical calculation is presented for a simple data set for the charge densities at different time in Chapter 5. The conclusions and some final remarks are presented in Chapter 6.

2 A Repulsive Force in the Electrodynamics

Our present aim is to formulate those kind of equations of motion that preserve the Lorentz invariance of the theory providing the Klein-Gordon type equation with the negative "mass term". At this stage it is not seen immediately what kind of physical process is generated really, but we know that the physical meaning of this negative term is a repulsive interaction. We restrict our consideration to only a pulse-like impact within the system, thus the additional term in the Maxwell's equation should be a process starting at the initial t_0 and ending at the final t, i.e., assuming that the elapsed time $\tau \ \ of \sim 10^{-18} \ to \ 10^{-12}$ seconds, is very short.

To achieve our previously summarized goals we start from the regular form of the Maxwell's equations [21, 22] modifying the second one (Eq. (3b) with an additional term

$$\alpha^2 \int_0^t \boldsymbol{B}(r,t')dt' \tag{3}$$

that brings a short time $(0 < t < \tau)$ interaction in the theory. Thus the four equations take the form:

$$\frac{1}{\mu_0} rot\boldsymbol{B} = \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{J} \tag{3a}$$

$$rot\boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} + \alpha^2 \int_0^t \boldsymbol{B}(r,t')dt' \tag{3b}$$

$$\varepsilon_0 div\boldsymbol{E} = \rho \tag{3c}$$

$$div\boldsymbol{B} = 0 \tag{3d}$$

Here, $\boldsymbol{B}$ denotes the magnetic field, $\boldsymbol{E}$ is the electric field, $\boldsymbol{J}$ is the current density, ρ is the charge density, ε_0 is the vacuum permittivity and μ_0 is the vacuum permeability. The parameter α pertains to the assumed interaction. (The square is used for the later calculation convenience.)

Now, we show the calculations to understand the influence of the above extra term (Eq. (3)) in the theory. In order to solve these equations it is usual to introduce the vector potential $\boldsymbol{A}$ [21, 22] by the help of Eq. (3d) as:

$$\boldsymbol{B} = rot\boldsymbol{A} \tag{4}$$

Substituting this into Eq. (3b) and rearranging the obtained formula we get

$$rot\left(\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} - \alpha^2 \int_0^t \boldsymbol{A}(r,t')dt'\right) = 0 \tag{5}$$

We can express the electric field $\boldsymbol{E}$ from the above equation

$$\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t} - grad\varphi + \alpha^2 \int_0^t \boldsymbol{A}(r,t')dt' \tag{6}$$

where the scalar potential φ is introduced too. Since there is a free degree of freedom in the connection of the scalar and the vector potentials, we are allowed to take the condition:

$$\frac{\partial \varphi}{\partial t} + div\boldsymbol{A} = 0 \tag{7}$$

with the choice

$$\varepsilon_0\mu_0 = 1$$

for these universal parameters. Now, we take Eq. (3c) and we replace $\boldsymbol{E}$ into it, thus we can write

$$div\boldsymbol{E} = -\frac{\partial(div\boldsymbol{A})}{\partial t} - \Delta\varphi + \alpha^2 \int_0^t div\boldsymbol{A}(r,t')dt' = \frac{\rho}{\varepsilon_0} \tag{8}$$

Eliminating the vector potential by the help of Eq. (7), thus we obtain

$$\frac{\partial^2 \varphi}{\partial t^2} - \Delta\varphi - \alpha^2\varphi = \frac{\rho}{\varepsilon_0} \tag{9}$$

The third term is a Lorentz invariant negative mass term Klein-Gordon equation. We remember that the term $-\alpha^2\varphi$ acts just in the time interval τ. The structure of equation is similar to the equations (1) and (2) (such as in Refs. [1-7]) that includes a dynamical phase transition depending on the parameter α as a consequence of a spinodal instability [19, 20]. The equation for the vector potential can be also formulated, starting from Eq. (3a) and substituting the form of electric field from Eq. (6)

$$rotrot\boldsymbol{A} = -\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \frac{\partial grad\varphi}{\partial t} + \alpha^2\boldsymbol{A} + \mu_0\boldsymbol{J} \tag{10}$$

Applying the vector identity:

$$rotrot = graddiv - \Delta$$

and the condition in Eq. (7), we can rewrite the above equation in a more expressive form

$$\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \Delta \boldsymbol{A} - \alpha^2 \boldsymbol{A} = \mu_0 \boldsymbol{J} \tag{11}$$

which is also a Lorentz invariant expression, and we can recognize that this is also a negative mass term Klein-Gordon type equation for the vector potential $\boldsymbol{A}$. Similarly to the previous remark, the term $-\alpha^2 \boldsymbol{A}$ is active in the time range τ. We can conclude that both the scalar and the vector potentials as basic fields – the components of a four vector

$$A_\mu = (-\varphi, \boldsymbol{A})$$

– fulfill the Lorentz invariant Klein-Gordon type equations with the same mathematical structure, the field variables propagate with the same speed, the whole description is Lorentz invariant.

It is important to emphasize that – from the viewpoint of the physical process description – the Lorentz invariant field equations (equations 9 and 11) for the scalar and the vector potentials have central roles. All of the other field variables can be deduced from these potentials [21, 22, 27, 28]. This fact can be obviously seen from the Lagrangian of the theory formulated [29, 30] as

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} \lambda \left(\partial_\mu A^\mu\right)^2 + \frac{1}{2} \alpha^2 A_\mu A^\mu + j_\mu A^\mu \tag{12}$$

where the $F_{\mu\nu}$ is the electromagnetic tensor field

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{13}$$

Selecting $\lambda = 1$ (Feynman's gauge) [31, 32] in the Lagrangian the movement equations for the four-potential can be obtained for the present case. It is clear that the electric and magnetic fields are not observables and are not components of neither the electromagnetic tensor field nor the Lagrangian.

Now, we should write the equations for the field variables, $\boldsymbol{E}$ and $\boldsymbol{B}$. Thus, we take the time derivative of Eq. (3a)

$$rot \frac{\partial \boldsymbol{B}}{\partial t} = \frac{\partial^2 \boldsymbol{E}}{\partial t^2} + \mu_0 \frac{\partial \boldsymbol{J}}{\partial t} \tag{14}$$

The term on the left hand side can be substituted after taking the rotation of Eq. (3b) by which we write

$$rot \left(-rot \boldsymbol{E} + \alpha^2 \int_0^t \boldsymbol{B}(\boldsymbol{r}, t') dt'\right) = \frac{\partial^2 \boldsymbol{E}}{\partial t^2} + \mu_0 \frac{\partial \boldsymbol{J}}{\partial t} \tag{15}$$

We can eliminate the field $\boldsymbol{B}$ applying again Eq. (3a), and finally we obtain

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \Delta \boldsymbol{E} - \alpha^2 \boldsymbol{E} = -\frac{1}{\varepsilon_0} grad\rho - \mu_0 \frac{\partial \boldsymbol{J}}{\partial t} + \mu_0 \alpha^2 \int_0^t \boldsymbol{J}(\boldsymbol{r}, t') dt' \tag{16}$$

Similarly, for the field $\boldsymbol{B}$, we take the time derivative of Eq. (3b)

$$rot\,\frac{\partial \boldsymbol{E}}{\partial t} = -\frac{\partial^2 \boldsymbol{B}}{\partial t^2} + \alpha^2 \boldsymbol{B} \tag{17}$$

and eliminating the field $\boldsymbol{E}$ by the help of the rotation of Eq. (3a) we obtain the equation for the magnetic field

$$\frac{\partial^2 \boldsymbol{B}}{\partial t^2} - \Delta \boldsymbol{B} - \alpha^2 \boldsymbol{B} = \mu_0 rot \boldsymbol{J} \tag{18}$$

It can be seen that for all of the field equations – equations (9), (11), (16) and (18) – have the same structure. We know from the former studies [1-7, 15] that these Klein-Gordon equations with a negative "mass term" are resulted from repulsive interactions. Thus, it seems to us that the interaction in the present case is a repulsive-like force which appears mathematically in the second Maxwell's equation, in Eq. (3b).

3 The Wheeler Propagator and the Time-Evolution of the Electric Field

The following physical description of the Wheeler propagator is based on Feynman's and Wheeler's original idea [33, 34]. The clear mathematical deduction of the Wheeler propagator is developed by Bollini, Rocca, Giambiagi and Oxman [8-14]. The difficult and complicated mathematical method to evaluate the calculations needs to apply the Bochner's theorem [35, 36] taking into account further complicated mathematical formulations [37].

In the knowledge of the time evolution equations we can study the processes evolving in the electric conductive media. We focus on the connection between the appearing electrical field and the charge distribution given by Eq. (16). Since the last two terms of this equation make rather complicated the solution and assuming that the contribution of the current and the time derivative of the current can be negligible at the initial time, we can simplify the problem to:

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} - \Delta \boldsymbol{E} - \alpha^2 \boldsymbol{E} = -\frac{1}{\varepsilon_0} grad\rho \tag{19}$$

This equation can be solved applying the Green function method. On the basis of it, the electric field $\boldsymbol{E}$ as the solution of this partial differential equation can be expressed by the following integral:

$$\boldsymbol{E}(r,t) = \int -\frac{1}{\varepsilon_0} grad\rho(x') \left[\frac{1}{(2\pi)^4} \int d^4k \frac{e^{ik(x-x')}}{k^2-\alpha^2}\right] dV' \tag{20}$$

Here, the four-vector:

$$x = \left(t = x_0, \boldsymbol{r} = (x_1, x_2, x_3)\right)$$

involves both the space and time coordinates:

$k = (\omega = p_0, k)$

denotes the four-momentum

$dV' = dx_1' dx_2' dx_3'$

is the volume element. (We follow the notations of Refs. [8-14] in the calculations of the Wheeler propagator.) The expression

$$G(x, x') = \frac{1}{(2\pi)^4} \int d^4k \frac{e^{ik(x-x')}}{k^2 - \alpha^2} \tag{21}$$

in the [...] bracket is the Green function generating the evolution of the process in the space-time from the initial $x'$to the final x.

In order to evaluate this integral, we find the zero points of the denominator

$$k^2 - \alpha^2 = p^2 - p_0^2 - \alpha^2 = 0 \tag{22}$$

from which we obtain

$$p_0 = \pm\sqrt{p^2 - \alpha^2} \tag{23}$$

To obtain the propagator, first, we need to calculate the integral in Eq. (21) with the

$$G_{adv}(x) = \frac{1}{(2\pi)^4} \int d^3p e^{ipr} \int_{adv} dp_0 \frac{e^{-ip_0 x_0}}{p^2 - p_0^2 - \alpha^2} \tag{24}$$

where Eq. (22) is used for the separation. Then the integration can be evaluated applying the residue theorem. We have two cases. We obtain the advanced propagator if the path of integration runs parallel to the real axis and below both the poles. (In the case of the retarded propagator the path runs above the poles.) Thus, considering the integration $x_0 > 0$ the path is closed on the lower half plane giving null result. In the opposite case, when $x_0 < 0$, there is a non-zero finite contribution of the residues at the poles

$$p_0 = \pm\omega = \sqrt{\boldsymbol{p}^2 - \alpha^2} \quad if \quad \boldsymbol{p}^2 \geq \alpha^2 \tag{25}$$

and

$$p_0 = \pm i\omega' = \sqrt{\boldsymbol{p}^2 - \alpha^2} \quad if \quad \boldsymbol{p}^2 \leq \alpha^2 \tag{26}$$

After all we can apply the Cauchy's residue theorem for the integration (for the internal integral) with respect to p_0. We take case if $\boldsymbol{p}^2 \geq \alpha^2$ and $x_0 < 0$. We have two poles (see Eq. (25)), thus the following integral:

$$\int_{adv} dp_0 \frac{e^{-ip_0 x_0}}{p^2 - p_0^2 - \alpha^2} = 2\pi i \frac{e^{-i\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}\, x_0}}{-2\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}} + 2\pi i \frac{e^{+i\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}\, x_0}}{2\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}}$$

$$= -2\pi \frac{\sin\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}\, x_0}{\sqrt{\boldsymbol{p}^2 - \alpha^2 + i0}}$$

In the other case, if $\boldsymbol{p}^2 \leq \alpha^2$ and $x_0 < 0$, the calculation is similar with the poles in Eq. (26). It is easy to check that formally the result is the same with the condition $\boldsymbol{p}^2 \leq \alpha^2$. The two cases ($\boldsymbol{p}^2 \geq \alpha^2$ and $\boldsymbol{p}^2 \leq \alpha^2$) can be summarized in one expression, i.e., we obtain a 3rd order integral for the advanced propagator:

$$G_{adv}(x) = -\frac{H(-x_0)}{(2\pi)^3}\int d^3\, p e^{ipr}\frac{sin\left[(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}x_0\right]}{(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}} \tag{27}$$

where $H(x)$ is the Heaviside's function which ensures the validity just for the retarded case $x_0 < 0$. Finally, we conclude that this formula is valid for $x_0 < 0$ and for any $\boldsymbol{p}$.

Reversing the previous procedure the retarded propagator ($x_0 > 0$) can be also calculated similarly for any $\boldsymbol{p}$

$$G_{ret}(x) = \frac{H(x_0)}{(2\pi)^3}\int d^3\, p e^{ipr}\frac{sin\left[(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}x_0\right]}{(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}} \tag{28}$$

Following Feynman's and Wheeler's idea [33, 34], i.e., considering that the propagator is the sum of the half advanced and the half retarded propagator we obtain the propagator

$$G(x) = \frac{Sgn(x_0)}{2(2\pi)^3}\int d^3\, p e^{ipr}\frac{sin\left[(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}x_0\right]}{(\boldsymbol{p}^2-\alpha^2+i0)^{\frac{1}{2}}} \tag{29}$$

Now, this formula is valid for any x_0 and for any $\boldsymbol{p}$. Here, the integrals can be rewritten by the Hankel transformation based on Bochner's theorem [35, 36] by which the propagator can be expressed analytically and denoted as

$$W^{(4)}(x) = \frac{\alpha}{8\pi}(x_0^2 - r^2)_+^{-\frac{1}{2}} I_{-1}\left(\alpha(x_0^2 - r^2)_+^{\frac{1}{2}}\right) \tag{30}$$

$I_{-1}(x)$ is the modified Bessel function – taking into account the notations:

$x_+^\beta = x^\beta \;\; for \;\; x > 0$

$x_+^\beta = 0 \;\; for \;\; x < 0$

This propagator is often called Wheeler propagator, when the negative sign is in the denominator of the Green function in Eq. (21), i.e., when the denominator is $k^2 - \alpha^2$. As a remark, it is important to emphasize, that the above propagator meets the requirement of causality.

Finally, we can express the resulted electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ generated from an initial charge distribution $\rho(x')$ during the elapsed time $0 \rightarrow t$ by the application of the calculated propagator:

$$\boldsymbol{E}(r,t) = \int -\frac{1}{\varepsilon_0} grad\rho(x')W^{(4)}(x - x')dV' \tag{31}$$

4 Evolution of an Initially Nearly Flat Gaussian Charge Distribution

On the basis of the previous calculations we can calculate the electric field in view of the initial charge distribution. The question is what kind of process is going within the system governed by the propagator. Now, we imagine a nearly Gaussian charge distribution:

$$\rho(r',0) = \rho_0 e^{-ar'^2} \tag{32}$$

at the initial time 0 in the space coordinate $r_0 = 0$. If the parameter

$a \sim 0$

the charge distribution can be considered practically homogeneous, since we can take that

$e^{-ar^2} \sim 1$

So, if we consider the charge gradient for small values of a we approximate

$$grad\rho(r') = -2\rho_0 a r' e^{-ar'^2} \sim -2\rho_0 a r' \tag{33}$$

Here, we apply the form of the Wheeler propagator from Eq. (30) for the present analytical calculations. Substituting the calculated charge gradient from Eq. (33) and the Wheeler propagator from Eq. (29) into the expression of $\boldsymbol{E}(r,t)$ in Eq. (31) then we obtain the time evolution of the electric field:

$$\boldsymbol{E}(r,t) = \frac{2\rho_0 a}{16\pi^3 \varepsilon_0} Sgn(t) \int_{V\,(all)} \int_{-\infty}^{\infty} \frac{sin\left[\left(\boldsymbol{p}^2 - \alpha^2 + i0\right)^{\frac{1}{2}} t\right]}{\left(\boldsymbol{p}^2 - \alpha^2 + i0\right)^{\frac{1}{2}}} e^{ip(r-r')} \boldsymbol{r}' e^{-ar'^2} d^3 p dV' \tag{34}$$

After the evaluation of the integral and simplifying the mathematical expression the electric field can be analytically expressed as:

$$\boldsymbol{E}(r,t) = \frac{\rho_0 a}{\varepsilon_0 \alpha} \boldsymbol{r} e^{-ar^2} Sgn(t)\, sinh(\alpha|t|) \tag{35}$$

It can be read out easily from this exact result that the magnitude of the electric field $\boldsymbol{E}$ follows an exponential behavior. The source of the huge electric field is the enormously growing charge distribution

$$\rho(r,t) = \frac{\rho_0 a}{\alpha}(3 - 2ar^2) e^{-ar^2} Sgn(t)\, sinh(\alpha|t|) \tag{36}$$

which can be obtained by the Maxwell's equation:

$$div\boldsymbol{E} = \frac{\rho}{\varepsilon_0}$$

It is interesting to see that if we integrate this charge density for the whole space the result is always zero for all positive values of the parameter $a > 0$

$$\int_0^\infty \rho(r,t)dV = 4\pi \int_0^\infty \frac{\rho_0 a}{\alpha}(3-2ar^2)e^{-ar^2} Sgn(t)\, sinh(\alpha|t|)dr = 0 \tag{37}$$

i.e., the conservation law of electric charge is completed, there is only internal movement of the charges – charge disjunction. Applying the continuity relation

$$\frac{\partial \rho}{\partial t} + div\boldsymbol{J} = 0 \tag{38}$$

and considering equations (3c) and (35) we obtain the current $\boldsymbol{J}(r,t)$

$$\boldsymbol{J}(r,t) = -\rho_0\, a\, \boldsymbol{r} e^{-ar^2} Sgn(t)\, cosh(\alpha|t|) \tag{39}$$

Here, we note that calculating the dropped part of Eq. (16)

$$-\mu_0 \frac{\partial \boldsymbol{J}}{\partial t} + \mu_0 \alpha^2 \int_0^t \boldsymbol{J}(r,t')dt' \tag{40}$$

with the above solution of the current in Eq. (39), we obtain zero. (This is in line that the current is zero at time 0.) Thus, we can say that the obtained solution for the electric field in Eq. (35) and for the charge density in Eq. (36) from the cut Klein-Gordon type equation in Eq. (19) can be considered as exact results.

5 Calculation Result

The time evolution of the charge density can be also calculated numerically by the propagator form given by Eq. (36). The charge density is homogeneous at the initial time $t_0 = 0$, the process ends in short time t. The resulted graphs are shown in Figs. 2 – 4 pertaining to time: 0.3, 0.7 and 1.0. Since the physical situation is spherically symmetric, it is enough to demonstrate the increase of the charge density, as a function of the radius, in different time segments.

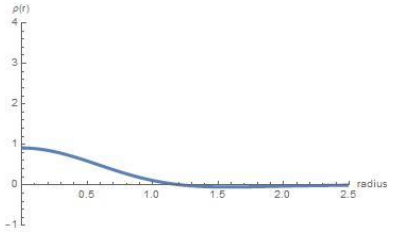


Figure 2

Charge density as a function of the radius r at time $t = 0.3$. The charge density and time are considered in natural units.

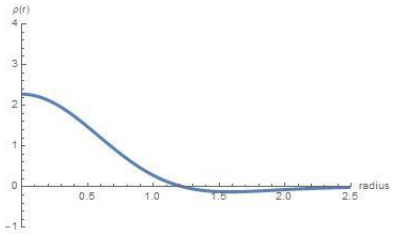


Figure 3

Charge density as a function of the radius r at time $t = 0.7$. The charge density and time are considered in natural units.

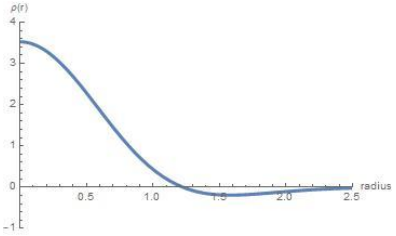


Figure 4

Charge density as a function of the radius r at time $t = 1$. The charge density and time are considered in natural units.

This figure demonstrates spectacularly how fast the charge density increases in time. Here, the applied parameters can be taken optionally at the present stage, thus

$\rho_0 = 1$

$a = 1$

and

$\alpha = 1$

are chosen. This means that the scales are in *natural units* in the figure. (This a similar assumption when the speed of light is taken $c = 1$ or the Planck constant is also $h = 1$ in other theories. The tendency does not depend on this choice.) We can see that at the beginning the charge density increases rather slowly comparing the later time, and in a certain time it can grow up in a giant form. During the elapsing time a negative spherically symmetric charge density is collecting with a maximal value at radius $\alpha = 1.5$. The process stops at the very short time τ , and it turns back, so finally the system reaches its originally homogeneous charge distribution. It seems from physical reasons natural that the process must be restricted to nano/micro-distances and for short time.

Conclusions

As a *thought experiment,* in the present work, it is shown how a negative mass term Klein-Gordon equation can be deduced in the electrodynamics. This aim could be achieved by adding an appropriate term to one of the Maxwell's equations. It is clear from other experiences of previous studies of mechanical, thermodynamic and field theoretical problems that the appeared term pertains to a repulsive interaction. As a result of the paper, on the one hand, the Wheeler propagator of the process is expressed. We pointed out, on the other hand, that this repulsive force causes a giant charge disjunction on a short range within a short time.

As a final consequence, we can predict that if this short time extreme intensive process (laser light or X-ray, gamma radiation) happens, it can contribute effectively, to the physical behavior of the entire or some small parts of systems, e.g., perhaps on nano/micro-scale, to the electric properties or to other transport phenomena of the few-body systems.

The present calculations do not involve the possibilities of charge oscillations, but this process should remain realistic. This examination and discussion are great challenges for future work.

Acknowledgement

Support by the Hungarian National Research, Development and Innovation Office of Hungary (NKFIH) Grant Nr. K119442 is acknowledged.

References

[1] K. Gambár, F. Márkus: A Simple Mechanical Model to Demonstrate a Dynamical Phase Transition, *Rep. Math. Phys.* **62** (2008) 219

[2] K. Gambár: Change of Dynamics of the Systems: Dissipative – Non-Dissipative Transition, *Informatika* **12** (2010) 23

[3] F. Márkus, K. Gambár, Quasiparticles in a Thermal Process, *Phys. Rev. E* **71** (2005) 066117

[4] K. Gambár, F. Márkus: A Possible Dynamical Phase Transition between the Dissipative and the Non-Dissipative Solutions of a Thermal Process, *Phys. Lett. A* **361** (2007) 283

[5] F. Márkus, F. Vázquez, K. Gambár: Time Evolution of Thermodynamic Temperature in the Early Stage of Universe, *Physica A* **388** (2009) 2122

[6] F. Márkus, K. Gambár: Wheeler Propagator of the Lorentz Invariant Thermal Energy Propagation, *Int. J. Theor. Phys.* **49**, (2010) 2065

[7] F. Márkus: *"Can a Lorentz Invariant Equation Describe Thermal Energy Propagation Problems?"* in Heat Concution – Basic Research (ed. V. S. Vikhrenko), Rijeka: InTech (2011) pp. 155-176

[8] C. G. Bollini, L. E. Oxman, M. C. Rocca: Coupling of Tachyons to Electromagnetism, *Int. J. Theor. Phys.* **38** (1999) 777

[9] C. G. Bollini, M. C. Rocca: Wheeler Propagator, *Int. J. Theor. Phys*. **37** (1998) 2877

[10] C. G. Bollini, J. J. Giambiagi: Dimensional Regularization in Configuration Space, *Phys. Rev. D* **53** (1996) 5761

[11] C. G. Bollini, M. C. Rocca: Convolution of Lorentz Invariant Ultradistributions and Field Theory, *Int. J. Theor. Phys.* **43** (2004) 1019

[12] C. G. Bollini, M. C. Rocca: Vacuum State of the Quantum String without Anomalies in any Number of Dimensions, *Nuovo Cimento A* **110** (1997) 353

[13] C. G. Bollini, M. C. Rocca: Is the Higgs a visible particle? *Nuovo Cimento A* **110** (1997) 363

[14] C. G. Bollini, L. E. Oxman, M. C. Rocca: Equivalence Theorem for Higher Order Equations, *Int. J. Theor. Phys.* **37** (1998) 2857

[15] T. Szőllősi, F. Márkus: Searching the laws of Thermodynamics in the Lorentz Invariant Thermal Energy Propagation Equation, *Phys. Lett. A* **379** (2015) 1960

[16] Ph. M. Morse, H. Feschbach: *Methods in Theoretical Physics I* McGraw-Hill, New York, 1953

[17] W. Greiner: *Relativistic Quantum Mechanics* Berlin, Heidelberg, New York: Springer (2000)

[18] S. R. de Groot, P. Mazur: *Non-Equilibrium Thermodynamics* North-Holland, Amsterdam, 1962

[19] Sz. Borsányi, A. Patkós, D. Sexty: Phys. Rev. D. **66** (2002) 025014

[20] Sz. Borsányi, A. Patkós, D. Sexty: Phys. Rev. D. **68** (2003) 063512

[21] J. A. Stratton: *Electromagnetic Theory* New York, London: McGraw-Hill (1941)

[22] J. D. Jackson: *Classical Electrodynamics* J. Wiley and Sons, New York, 1999

[23] J. Zheng, W. Zhuang, N. Yan, G. Kou, H. Peng, C. McNally, D. Erichsen, A. Cheloha, S. Herek, C. Shi, Y. Shi: Classification of HIV-I-mediated neuronal dendritic and synaptic damage using multiple criteria linear programming, Neuroinformatics **2** (2004) 303

[24] C. Pozna a , R.-E. Precup, J. K. Tar, I. Škrjanc, S. Preitl: New results in modelling derived from Bayesian filtering, Knowledge-Based Systems **23** (2010) 182

[25] A. Ürmös, Z. Farkas, M. Farkas, T. Sándor, L. T. Kóczy, Á. Nemcsics: Application of self-organizing maps for technological support of droplet epitaxy, Acta Polytechnica Hungarica **14** (2017) 207

[26] S. Vrkalovic, E.-C. Lunca, I.-D. Borlea: Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants, International Journal of Artificial Intelligence **16** (2018) 208

[27] R. Courant, D. Hilbert: *Methods of Mathematical Physics*, Vol. II. New York, London: Interscience (1962)

[28] K. Gambár, M. Lendvay, R. Lovassy, J. Bugyjás: Application of Potentials in the Description of Transport Processes, *Acta Polytech. Hung.* **13** (2016) 173

[29] S. Weinberg: *The Quantum Theory of Fields* Cambridge: Cambridge Univ. Press. (1995)

[30] M. Srednicki: *Quantum field theory* Cambridge: Cambridge Univ. Press. (2007)

[31] G. S. Adkins: Phys. Rev. D. **36** (1987) 1929

[32] J. D. Jackson: Am. J. Phys. **70** (2002) 917

[33] J. A. Wheeler, R. P. Feynman: Interaction with the Absorber as the Mechanism of Radiation, *Rev. Mod. Phys.* **17** (1945) 157

[34] J. A. Wheeler, R. P. Feynman: Classical Electrodynamics in Terms of Direct Interparticle Action, *Rev. Mod. Phys.* **21** (1949) 425

[35] S. Bochner: *Lectures on Fourier Integrals* New Jersey: Princeton Univ. Press (1959), pp. 224-230

[36] A. J. Jerri: *The Gibbs Phenomenon in Fourier Analysis, splines, and wavelet approximations* Dordrecht: Kluwer (1998)

[37] S. Gradshteyn, I. M. Ryzhik: *Tables of Integrals, Series, and Products* New York: Academic Press (1994)

Application of Compensation Algorithms to Control the Movement of a Robot Manipulator

Gennady K. Shadrin[1], Darya L. Alontseva[1], Assel T. Kussaiyn-Murat[1], Albina T. Kadyroldina[1], Olzhas B. Ospanov[1], Tamás Haidegger[2]

[1] D. Serikbayev East Kazakhstan State Technical University, Protozanov street 69, 070004, Ust-Kamenogorsk, Kazakhstan
shadrin.g.k@yandex.ru

[2] Antal Bejczy Center for Intelligent Robotics (IROB), EKIK, Óbuda University, Bécsi út 96/b, H-1034 Budapest, Hungary
haidegger@irob.uni-obuda.hu

Abstract: This article presents an application-oriented method for the structural synthesis of control systems for multichannel linear objects. It provides a general form, based on the compensation for object dynamics and disturbances. The algorithm is based onalgebraic transformations of the mathematical model of the object and reference systems. The general procedure for the synthesis of a control algorithm is presented by the example of a SISO first-order object. Parametric robustness and correspondence of the system behavior to its own reference filters were derived. The possibility of applying this method to control non-linear objects presented in the form of "State Dependent Coefficient (SDC)" is ascertained. A simplified example is given by the synthesis of the motion control of a one-link manipulator with a drive, described by a second-order nonlinear equation. Control of a two-link manipulator represented by its linearized equation is demonstrated. The convenience of the proposed synthesis method for controlling multi-channel objects under certain simplifications allowed by the specifics of the use of objects is shown. The numerical example shows the independence of the system through individual channels, a zero static error in all the modes of operation and the correspondence of the system behavior to given dynamics.

Keywords: robot manipulator; control algorithm; multi-dimensional control system; reverse model; reference filter; structural synthesis

1 Introduction

The application of robots in the area of manufacturing is rapidly growing [1–3]. Robot manipulators are widely used in manufacturing processes, such as plasma cutting, plasma surface treatment in the form of plasma spraying of powder or

wire coatings [4, 5]. Plasma processing requires the accurate control of a number of parameters during the entire procedure, including the distance from the plasma system's nozzle to the surface of the workpiece, the nozzle movement speed and the angle between the plasma jet and the surface being treated [4]. Exceeding these parameters beyond the critical limits can lead not only to defective products, but also to an accident (e.g., a short circuit). In cases when the robot program is generated according to a given geometrical model of a processed workpiece or part, very often, deviation of the shape of the real object from the model leads to the violation of process parameters, with all its undesirable consequences. This problem is particularly acute in the case of objects with complex shape, when small relative errors of geometric parameters and object positioning may lead to large deviations regarding the distance between the tools mounted on the manipulator and the object's surface. Thus, it is necessary to solve the problem by adequately controlling the motion of a robotic arm and planning its trajectory. A number of research publications has addressed this issue recently [6–15]. A manipulator is a machine, in which the mechanism usually consists of a series of segments, jointed or sliding relative to one another, for the purpose of grasping and/or moving objects (pieces or tools) usually in several degrees of freedom [16]. Each joint (link) or element is equipped with its own drive, controlled by a microcontroller, which ensures the movement of the tool or the entire robot to any point of the workspace. The Kawasaki RS10L industrial robot (Kawasaki Heavy Industries, Ltd., Japan) (Fig. 1) can serve as a demonstrator, where rotational joints are used, but the motion of the tool in a preset direction is prescribed. Typically, performing robotic plasma cutting or plasma spraying requires moving the robot arm with the equipment installed on it along a given flat path with a relatively low speed, for example, about 50 mm/sec for plasma spraying of coatings.

a *b* *c*

Figure 1

a) The Kawasaki RS10L industrial robot employed in the experiments; b) the robot with a plasma cutting device installed; c) a plasma cutting process in progress.

However, the constraint motion of joints or nodes, taking into account the inertia and limited power of the drives, is a rather complicated control task [8–14]. It is often required to provide the dynamics of movements' characteristic of complex robot application.

The dynamics of the mechanical system of a manipulator is well studied, and can be described by the Lagrange equations [17–20]:

$$\frac{d}{dt}\frac{\partial \boldsymbol{L}}{\partial \dot{\boldsymbol{q}}} - \frac{\partial \boldsymbol{L}}{\partial \boldsymbol{q}} = \boldsymbol{\tau},\ L = K - P, \tag{1}$$

where $\boldsymbol{q} \in R^n$ denotes the vector of generalized coordinates, $K \in R^1$, $P \in R^1$ the kinetic and potential energy of the manipulator, respectively, $\boldsymbol{\tau} \in R^n$ the vector of driving moments in manipulator links and n being the number of manipulator joints. After determining $\boldsymbol{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and calculating the derivatives, from eq. (1) for "hard" robots, the "standard" second-order vector–matrix nonlinear differential equations follow in the form:

$$\boldsymbol{M(q)\ddot{q} + N(q,\dot{q}) + G(q) = \tau}, \tag{2}$$

where $\boldsymbol{M(q)} \in R^{n \times n}$ denotes the inertia matrix, $\boldsymbol{N(q,\dot{q})} \in R^{n \times 1}$ the vector of Coriolis and centrifugal forces and $\boldsymbol{G(q)}$ being the gravity vector.

For a robot with rotational joints, $\boldsymbol{q}$ is the vector of rotation angles of its joints, therefore, to transform rotation angles to Cartesian coordinates, coordinate transformation should be added:

$$\boldsymbol{y = f(q)}, \tag{3}$$

where y denotes the vector of the Cartesian coordinates of the robot tool and $\boldsymbol{f}(\boldsymbol{q})$ being the vector function.

It should be noted that the functions included in (2), (3) are smooth nonlinear functions in the domains of their definition (trigonometric functions, squaring and certain others). In control theory, there are no universal engineering methods for the synthesis of control systems for nonlinear objects; only applications for certain classes of objects have been developed. A common technique for the design of control systems is the linearization of the object in the area of the operating point and the use of well-developed linear synthesis methods. Another approach for converting nonlinear objects to linear form in a wide range of parameter changes is feedback linearization [21, 22]; this method is widely used to synthesize control algorithms for affine systems, which includes robots, i.e., "Computed Torque Control" (CTC) [18, 22]. It should be noted the complexity of the control system obtained by CTC should be noted, given the need to calculate the second derivative of the control signal. The State Dependent Coefficient (SDC) technique [23–25] is also widespread, according to which a nonlinear object is represented as a linear model with state-dependent parameters. After linearization by any technique, the object is controlled using classical or modern linear methods of control theory. It should be noted that classical methods are poorly adapted to

solve the problem of controlling multi-channel objects, i.e. Multiple-Input Multiple-Output (MIMO) objects.

In all cases, the task is to create a MIMO motion control system for the manipulator. Classical approaches of control algorithm synthesis are ill-posed for the task, when it includes numerous dynamic constraints, as discussed above. Most of the classical approaches are intended for the analysis of control systems for a single-channel, single-input single-output (SISO) system [26, 27]. It is necessary to analyze a number of prior defined algorithms before a given dynamic motion system is achieved. Advanced techniques, such as modal control and LQ-optimization "in pure form" are means of correcting the inherent dynamics of the control object [28, 29]. Modern deep learning techniques may also be used for deriving such complex robot controllers, if enough teaching data is available [30]. For the proper testing of external disturbances, compensators are additionally included in the system, and again, many times a trial-and-error method is used in practice. To develop an algorithm to control the robot tool motion, this article proposes a method based on the object's dynamics and perturbations compensation, offering a number of benefits [31, 32]. This is an analytical method, proving the correction of the inherent dynamics of the controlled object in a single algorithm, as well as compensation for external influences, testing with zero static error. The source data is set in the form of reference filters of a closed-loop system. It is easy to apply to various manipulators and tasks. The control algorithm is obtained by performing a finite number of algebraic operations over matrices. The method is based on the compensation of all external additive effects, while the precision control of the object's dynamics is achieved with reference-filters by means of inverse mathematical modeling of the object and immediately deriving the required dynamics of the system. This method has not yet been sufficiently developed and studied; there are some results for linear systems with constant parameters [30, 31]. However, this technique may well be "adapted" for objects in the form of SDC, but this issue has yet to be considered.

The disadvantage of the method is its need for information on the state variables of the object, in case the correction of the own dynamics is required. The method has been developed for linear plants; however, research is currently being conducted on further applications for nonlinear objects in a generalized form.

Our motivation is the development of a robotic system for plasma processing of objects with complex shapes. The goal of this study is to design a new method to derive the robot manipulator's motion control. This article presents a brief description of the method of compensating the dynamics of an object and disturbances, describes the synthesis of a control algorithm using an example of a first-order object, describes the synthesis of a motion control algorithm for one link of a manipulator, when nonlinear equations of a link are presented in the form of SDC, and demonstrates the applicability of the proposed method for the synthesis of MIMO control objects using the example of a linearized model of a two-link manipulator.

2 Features of the Object Dynamics and Perturbations Compensation Method

When developing the so-called "*dynamics and perturbations compensation method*", the control feedback is not postulated, but appears after structure's equivalent transformations obtained on the basis of the compensation method. Through this approach, it is possible to obtain a method of structural synthesis of control systems, directly guarantying the specified quality of control. The general principle of the method is to put the controlled object in series with its inverse mathematical model, and subtract all additive external disturbances from the corresponding variables of the inverse model. Disturbing effects not possible to be compensated are evaluated based on the mismatch of the corresponding object variables and the inverse model. To implement a control device in the circuit of external actions of the inverse model, the reference filters are involved, determining the behavior dynamics of a closed-loop control system. The method has an obvious advantage over the known ones regarding the clarity of the source data setting, and the simplicity of the synthesis of the control algorithm. The algorithms obtained by this method provide the compensation of disturbances with zero static error and a given dynamics changing of the controlled object. The disadvantage of the method is the need for information on the state variables of the object in case the correction of the own dynamics of this object is required. When it is not available, a state observer should be used, which is still part of our future work. A multichannel control object, represented in the state space, is considered.

$$\begin{aligned} \dot{\boldsymbol{x}}_0' &= \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{B}\boldsymbol{u}, \\ \boldsymbol{y}_0' &= \boldsymbol{C}\boldsymbol{x}_0, \\ \dot{\boldsymbol{x}}_0 &= \dot{\boldsymbol{x}}_0' + \boldsymbol{f}_{\boldsymbol{x}}, \\ \boldsymbol{y}_0 &= \boldsymbol{y}_0' + \boldsymbol{f}_{\boldsymbol{y}}, \end{aligned} \tag{4}$$

where $\boldsymbol{x}_0 \in \mathrm{R}^n$ denotes the vector of the state variables, $\boldsymbol{y}_0 \in \mathrm{R}^m$ the vector of the output variables, $\boldsymbol{u} \in \mathrm{R}^m$ the vector of control input, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$are the given numerical matrices of the controlled object of corresponding dimensions, $\boldsymbol{f}_{\boldsymbol{y}} \in \mathrm{R}^m$ vector denoting the perturbation added to the outcome variables and $\boldsymbol{f}_{\boldsymbol{x}} \in \mathrm{R}^n$ being the perturbation on the state variables. The linear system according to (4) must be kept under stable control. The method requires that the matrix composed of the first *m* rows of the **B** matrix is a non-degenerate one. The control objective is to achieve the desired *y* output of the system via perturbation compensation of the $\boldsymbol{f}_x, \boldsymbol{f}_y$ within the accuracy of the given sample dynamic systems (Fig. 2).

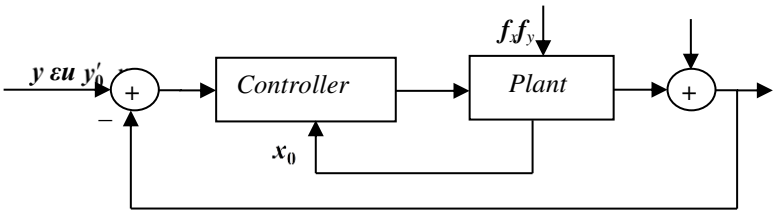


Figure 2
The block diagram of the control system

The control algorithm of the system (4) according to the method of compensation of the object's dynamics and perturbations [31, 32] has the form:

$$\begin{aligned}
&\dot{\boldsymbol{x}}_{\boldsymbol{\phi}} = \boldsymbol{R}_1 \boldsymbol{x}_{\boldsymbol{\phi}} + \boldsymbol{\Phi}_2 \boldsymbol{\varepsilon},\\
&\boldsymbol{u} = \boldsymbol{R}_3 \boldsymbol{x}_{\boldsymbol{\phi}} + \boldsymbol{N}_1 \cdot \tilde{\boldsymbol{x}}_0 + \boldsymbol{P}\boldsymbol{\varepsilon},\\
&\tilde{\boldsymbol{x}}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_{\boldsymbol{\phi}},\\
&\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{y}_0,
\end{aligned} \tag{5}$$

$$\begin{aligned}
&\boldsymbol{R}_1 = \boldsymbol{\Phi}_1 + \boldsymbol{\Phi}_2 \boldsymbol{\Phi}_3;\ \boldsymbol{R}_3 = \boldsymbol{N} + \boldsymbol{P}\boldsymbol{\Phi}_3;\ \boldsymbol{P} = \boldsymbol{G}\boldsymbol{\Phi}_2;\\
&\boldsymbol{N} = \boldsymbol{G}\boldsymbol{\Phi}_1 + \boldsymbol{H}\boldsymbol{\Phi}_3;\ \boldsymbol{N}_1 = \boldsymbol{G}\boldsymbol{\Phi}_{r1} + \boldsymbol{H}\boldsymbol{\Phi}_{r3}
\end{aligned} \tag{6}$$

where $x_\phi \in \mathrm{R}^n$ is the vector of variables of regulator state; $\boldsymbol{y} \in \mathrm{R}^m$ being the control vector. Input variables to be controlled are $\boldsymbol{y}_0, \boldsymbol{x}_0$; $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, $\boldsymbol{R}_3$, $\boldsymbol{N}$, $\boldsymbol{N}_1$ are intermediate matrices.

Block matrices $\boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}$ are matrices of the inverse model of the controlled system. Matrices $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$, $\boldsymbol{\Phi}_{r1}$, $\boldsymbol{\Phi}_{r3}$ determine sample dynamic systems, referred to as *reference filters* hereafter. A more detailed consideration of reference filters will be presented below. Matrices $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$, $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$, $\boldsymbol{\Phi}_{r1}$ and $\boldsymbol{\Phi}_{r3}$ are interconnected with $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ by the ratio:

$$\begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{G} & \boldsymbol{H} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{0} \end{bmatrix}^{-1}, \tag{7}$$

$$\begin{cases}
\boldsymbol{E}\boldsymbol{\Phi}_1 + \boldsymbol{F}\boldsymbol{\Phi}_3 = \boldsymbol{1},\\
\boldsymbol{E}\boldsymbol{\Phi}_2 = \boldsymbol{0},\\
\boldsymbol{\Phi}_3\boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2 = -\boldsymbol{1},\\
\boldsymbol{E}\boldsymbol{\Phi}_{r1} + \boldsymbol{F}\boldsymbol{\Phi}_{r3} = \boldsymbol{1}.
\end{cases} \tag{8}$$

In eqs. (7), (8) and further on in matrix expressions, **0** denotes the zero matrix, and **1** the unit (or identity) matrix with corresponding dimensions. Eq. (7)

unambiguously determines the matrices **E, F, G, H**. If the block matrix on the right-hand side of (7) is degenerate, then the system (4) is considered uncontrolled in output and our method of synthesizing a controller shall not be used. On the contrary, ratios (8) allow for the choice of the coefficients of the reference filters. Setting these coefficients is usually intuitive when synthesizing a control algorithm particularly for a specific task for the system. The effect of the coefficients free choice, in general, has not yet been studied.

The dynamic properties of the obtained closed-loop control system are represented by the equations:

$$\begin{aligned}
&\dot{\boldsymbol{x}}_{\boldsymbol{\phi}} = \boldsymbol{\Phi}_1 \boldsymbol{x}_{\boldsymbol{\phi}} - \boldsymbol{\Phi}_2 \boldsymbol{C}\tilde{\boldsymbol{x}}_0 + \boldsymbol{\Phi}_2(\boldsymbol{y} - \boldsymbol{f}_y),\\
&\dot{\tilde{\boldsymbol{x}}}_0 = \boldsymbol{\Phi}_{r1}\tilde{\boldsymbol{x}}_0 + \boldsymbol{f}_x,\\
&\boldsymbol{\varepsilon} = -\boldsymbol{\Phi}_3 \boldsymbol{x}_{\boldsymbol{\phi}} - \boldsymbol{C}\tilde{\boldsymbol{x}}_0 + (\boldsymbol{y} - \boldsymbol{f}_y),\\
&\boldsymbol{u} = \boldsymbol{N}\boldsymbol{x}_{\boldsymbol{\phi}} + (\boldsymbol{N}_1 - \boldsymbol{P}\boldsymbol{C})\tilde{\boldsymbol{x}}_0 + \boldsymbol{P}(\boldsymbol{y} - \boldsymbol{f}_y),\\
&\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{y}_0, \boldsymbol{P} = \boldsymbol{G}\boldsymbol{\Phi}_2
\end{aligned} \quad . \tag{9}$$

From these expressions it is easy to get the transfer function for:

$$\begin{aligned}
\boldsymbol{\varepsilon}(s) = &\left(\boldsymbol{\Phi}_3\left(s\cdot\mathbf{1} - \boldsymbol{\Phi}_1\right)^{-1}\boldsymbol{\Phi}_2 - \mathbf{1}\right)\boldsymbol{C}\left(s\cdot\mathbf{1} - \boldsymbol{\Phi}_{r1}\right)^{-1}\boldsymbol{f}_x(s) -\\
&-\left(\boldsymbol{\Phi}_3\left(s\cdot\mathbf{1} - \boldsymbol{\Phi}_1\right)^{-1}\boldsymbol{\Phi}_2 - \mathbf{1}\right)\left(\boldsymbol{y}(s) - \boldsymbol{f}_y(s)\right)
\end{aligned}, \tag{10}$$

where s is the scalar complex variable of the Laplace transform. Note that in static mode $\boldsymbol{\varepsilon}(\mathbf{0}) = \mathbf{0}$. This follows from (9), if we consider that according to (8) $\boldsymbol{\Phi}_3\boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2 = -\mathbf{1}$. Thus, the static control error is zero. From (10), it can be seen that the reference filter with matrices $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$ determine the behavior of the closed-loop system through the channels of setting job $\boldsymbol{y}$ and compensation $\boldsymbol{f}_y$ and is an m-channel low-pass filter.

A reference filter with the $\boldsymbol{\Phi}_{r1}$, $\boldsymbol{\Phi}_{r3}$ matrices can also be viewed as an *n*-channel low-pass filters through the $\boldsymbol{f}_x$ compensation channels. The $\boldsymbol{f}_x$ filter, according to (9), is connected in series with the $\boldsymbol{y}$ and $\boldsymbol{f}_y$ filter.

By specifying the free coefficients of the reference filters, one can determine the frequency bandwidths of the closed-loop system and cross-links between its control channels. According to [32], for an *m*-channel controlled object with *n* state variables, the coefficients of $m \times n$ blocks shall be free (i.e., *n* pieces for each of the *m* control channels) of the type:

$$\begin{bmatrix} \phi_{11} & \cdots & \phi_{1n} \\ \cdots & \cdots & \cdots \\ \phi_{m1} & \cdots & \phi_{mn} \end{bmatrix}; \begin{bmatrix} \phi_{r11} & \cdots & \phi_{r1n} \\ \cdots & \cdots & \cdots \\ \phi_{rm1} & \cdots & \phi_{rmn} \end{bmatrix}. \tag{11}$$

The blocks in eq. (11) occupy the $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_{r1}$ top rows, provided that the matrix composed of the first *m* rows of the matrix **B** is non-degenerate. For SISO control systems, it is advisable to choose the free coefficients so that they correspond to the standard forms of the low-pass filter. As for MIMO systems, the issues of selecting the free coefficients and determining the remaining coefficients of reference filters, in general, have not yet been worked out. These tasks are being solved by considering the object properties and system requirements, as well as by solving the system (8) in each individual case. From the second equation (5), it follows that along with the main feedback in the control system, there is a local feedback on the state variables of the plant with the $\boldsymbol{N_1}$ matrix. This feedback adjusts the dynamic properties of the object in accordance with the *fx* reference filter. If the correction is not needed, then the parameters of this filter should be chosen, so that $\boldsymbol{N_1}$ equals zero. Then, there is no need to measure or evaluate the state variables of the control object. Partial correction of the object is possible. Thus, according to the given mathematical model of the control plant and the specified parameters of the reference filters, the algorithm of the control device is calculated, which immediately guarantees the specified quality of the control processes. Therefore, it is of interest to consider the synthesis of algorithms and the analysis of the properties of the obtained systems specific to robot manipulators.

According to (5)–(8) by this technique, the control algorithm is obtained as a result of algebraic transformations over the matrices of the mathematical model of the control object and the matrices of filter standards. These transformations remain valid when replacing the matrix coefficients with functions of the state variables of this object, which makes it possible to use the proposed method for nonlinear objects presented in the form of SDC. Consequently, if the matrices of the filter-standard depend on state variables, then (10) is not applicable.

3 Control Algorithm Elaboration for the Object of the First Order

The single-channel Linear Time-Invariant (LTI)control object is represented as follows

$$\begin{aligned} \dot{x}_0' &= ax_0 + bu, \qquad \dot{x}_0 = \dot{x}_0' + f_x, \\ y_0' &= cx_0, \qquad\quad y_0 = y_0' + f_y, \end{aligned} \tag{12}$$

where *a, b, c* are coefficients. The objective is to reproduce the reference signal *y* at the output of the object and to compensate the f_x, f_y perturbations within an accuracy of the given reference dynamic system. This example is handy because the structure of system (12) replicates the structure of (4), but all the block matrices in (4) are replaced by coefficients. Therefore, the solution of this

problem, in general, is of a simple form, and at the same time allows us to follow the synthesis of the control algorithm, as well as to reveal the general properties of the obtained control system.

The synthesis of the control algorithm is performed in the following order. By (7) we calculate:

$$\begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{G} & \boldsymbol{H} \end{bmatrix} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a & b \\ c & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1/c \\ 1/b & -a/cb \end{bmatrix}, \tag{13}$$

where e, f, g, h are coefficients of the inverse model of the object (4). To determine the coefficients of reference filters, we record the equations (8):

$$\frac{1}{c}\phi_3 = 1, \phi_3 \frac{1}{\phi_1}\phi_2 = -1 \ , \frac{1}{c}\phi_{r3} = 1. \tag{14}$$

According to (14), we have three equations to determine the five coefficients of the reference filters ϕ_1 , ϕ_2 , ϕ_3 , ϕ_{r1} , ϕ_{r3} , hence, two coefficients from among them can be set arbitrarily. Considering (11), we assign as free ϕ_1 and ϕ_{r1}. Solving (14), we get the matrices of reference filters:

$$\begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 \\ \boldsymbol{\Phi}_3 & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ \phi_3 & 0 \end{bmatrix} = \begin{bmatrix} \phi_1 & -\frac{\phi_1}{c} \\ c & 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Phi}_{r1} \\ \boldsymbol{\Phi}_{r3} \end{bmatrix} = \begin{bmatrix} \phi_{r1} \\ \phi_{r3} \end{bmatrix} = \begin{bmatrix} \phi_{r1} \\ c \end{bmatrix}. \tag{15}$$

Substituting the coefficients from (14), (15) into (6), we find

$$\begin{gathered} r_1 = \phi_1 - \frac{\phi_1 c}{c} = 0 \, , p = g \cdot f_2 = -\frac{\phi_1}{bc} \, , n = g \cdot \phi_1 + h\phi_3 = \frac{\phi_1}{b} - \frac{a}{b}, \\ n_1 = g \cdot \phi_{r1} + h\phi_3 = \frac{\phi_{r1} - a}{b} \, , \ r_3 = n + p\phi_3 = \frac{\phi_1}{b} - \frac{a}{b} - \frac{\phi_1 c}{bc} = -\frac{a}{b} \, . \end{gathered} \tag{16}$$

Substituting (16) into (5), we record the object control algorithm (12):

$$\begin{gathered} \dot{x}_\phi = -\frac{\phi_1}{c}\varepsilon, \qquad \tilde{x}_0 = x_0 - x_\phi, \quad \varepsilon = y - y_0, \\ u = -\frac{a}{b}x_\phi + \frac{\phi_{r1} - a}{b}\tilde{x}_0 - \frac{\phi_1}{bc}\varepsilon. \end{gathered} \tag{17}$$

Thus, a proportional–integral (PI) control law is obtained with local feedback through the variable $\tilde{x}_0 = x_0 - x_\phi$.

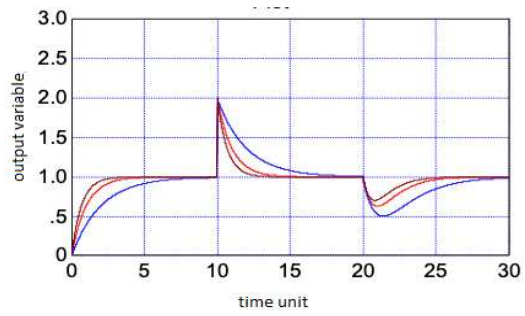
Processes in a closed loop control system consisting of the object (12) and the regulator (17) after substitution (13) and (15) in (9) take the form:

$$\begin{aligned}
\dot{x}_\phi &= \phi_1 x_\phi + \phi_1 \tilde{x}_0 - \frac{\phi_1}{c}(y - f_y), \\
\dot{\tilde{x}}_0 &= \phi_{r1} \tilde{x}_0 + f_x, \\
\varepsilon &= -c x_\phi - c\tilde{x}_0 + (y - f_y), \\
u &= \frac{\phi_1 - a}{b} x_\phi + \frac{\phi_{r1} - a + \phi_1}{b} \tilde{x}_0 - \frac{\phi_1}{c}(y - f_y)
\end{aligned} \quad . \tag{18}$$

As seen from (18), the free coefficients of the reference filter ϕ_1 and ϕ_{r1} determine the dynamic properties of the control system. For the sustainability of the processes in (18), they must be negative. The study of the influence of reference filters coefficients on the control process was performed by computer-based simulation method and was about to analyze the transients in a system composed of (12) and (17) for a numerical example when $a = -1$, $b = c = 1$. Figure 3 shows the transient response of the system for the values $\phi_1 = -0.5$ (blue), $\phi_1 = -1$ (red) and $\phi_1 = -1.5$ (brown). In Fig. 3and other figures, at zero time, the input was changed according to a unit step, after 10 seconds, a single step perturbation f_y, appeared, and after 20 seconds appeared a single step perturbation f_x of a negative sign.

As seen, the approach to increase the system performance by the increase of ϕ_1 leads to significant amplitude brightening of the control input. Apparently, it is advisable for this example when $\phi_1 = 1$, then the amplitude of the control action is moderate and characterized by a high speed of operation. Fig. 4 shows transition functions similar to Figure 3, but here the value of ϕ_{r1} changed when $\phi_1 = 1$. Expectedly from (18), this parameter affects only the nature of compensation of the perturbation of f_x. As seen in Fig. 4, it is possible to increase ϕ_{r1} without increasing the amplitude of the control input.

a



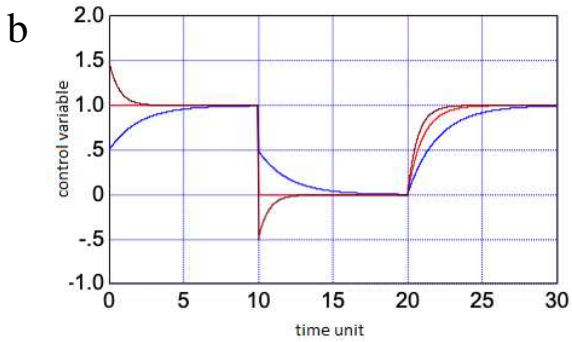


Figure 3
Input signal to test the system (12), with controller (17) when changing ϕ_1 : a) output variable, b) control variable

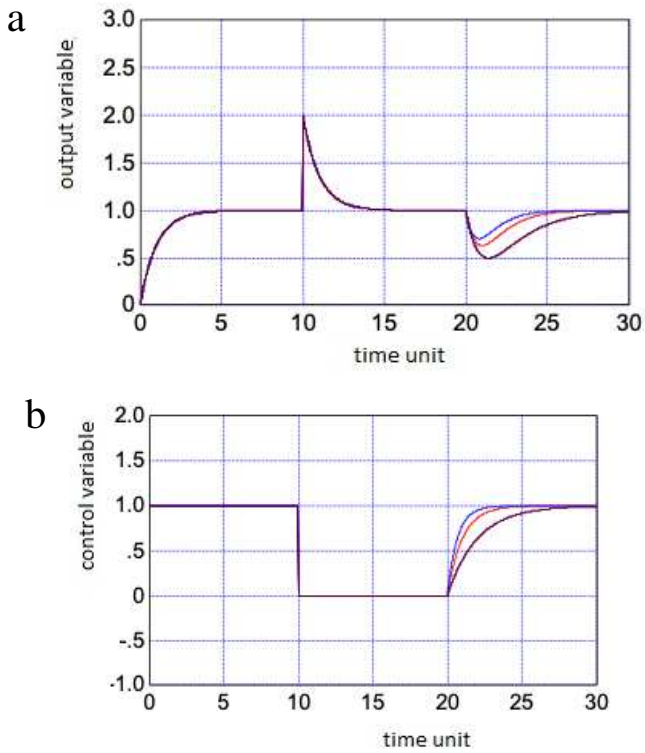


Figure 4
Input signal to test the system (12), with controller (17) when changing ϕ_{r1} a) output variable, b) control variable

The study of robustness for a given set of coefficients was to analyze transition functions while a significant change of the coefficients of the control object and the regulator being calculated for the nominal values of these coefficients. Figure 5 demonstrates these functions. Here, the coefficients *a, b, c* of the system (12) were successively changed by 50% from their nominal values.

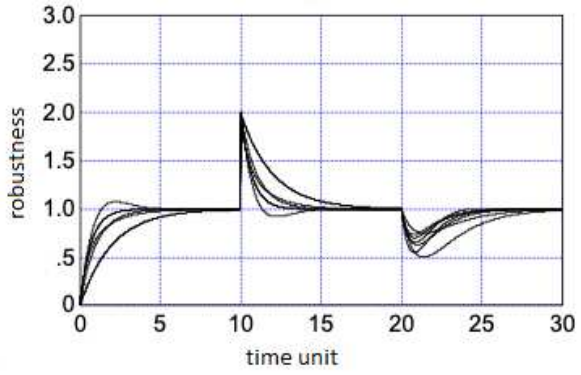


Figure 5

The study of the control system robustness (12), (17)

As seen, the system remains stable and the control quality is adequate. Simulation with a number of other sets of object coefficients (12), including for an unstable object, revealed similar results.

4 Motion Control of One Manipulator Joint

Let us consider the algorithm synthesis of the rotation angle control of one link of a rotary type manipulator. For simplicity, we neglect the dynamics of the electromagnetic processes of the drive and friction in the joints. In addition, we believe that the mass of the link is concentrated at its end point. Eq. (2) for one link, the axis of rotation of which is fixed at the origin, has the form:

$$(J_p + ml^2)\ddot{q} + mgl\,cos(q) = ku, \tag{19}$$

where m denotes the link mass, l –the link length, q the link rotation angle relative to the horizontal axis, g the acceleration of gravity and J_p being the drive moment of inertia. It was assumed in (19) that the driving moment created by the drive is connected with the control signal to this drive u by the formula $\tau = ku$, where k is the proportionality coefficient.

To synthesize a control algorithm, these equations must be represented in the state space. Dividing the right and left sides of (19) into $J_p + ml^2$ and denoting $x_{01} = \dot{q}$, $x_{02} = q$, $a_1 = \dfrac{mgl}{J_p + ml^2}$, $b = \dfrac{k}{J_p + ml^2}$, where k is a coefficient of proportionality, we obtain the mathematical model of the control object in the form:

$$\begin{aligned} \dot{x}_{01} &= -a_1 \cos x_{02} + bu, \\ \dot{x}_{02} &= x_{01}, \\ y_0 &= x_{02}. \end{aligned} \tag{20}$$

We represent (20) in the form of SDC; for this, we divide and multiply the first term in (20) by x_{02}, in the end we get:

$$\begin{aligned} \dot{x}_{01} &= a(\cdot) \cdot x_{02} + bu, \\ \dot{x}_{02} &= x_{01}, \\ y_0 &= x_{02}, \end{aligned} \tag{21}$$

where

$$a(\cdot) = a(x_{02}) = -\frac{a_1}{x_{02}} \cos x_{02}, \; x_{02} \neq 0. \tag{22}$$

Now in notation (4) ,we have:

$$\boldsymbol{A} = \begin{bmatrix} 0 & a(\cdot) \\ 1 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}. \tag{23}$$

We carry out the synthesis of the control algorithm according to (5)–(8) taking into account (11). By (7) we calculate:

$$\boldsymbol{E} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{F} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} \frac{1}{b} & 0 \end{bmatrix}, H = -\frac{a(\cdot)}{b}. \tag{24}$$

According to (8), we compose equations for determining filter standards:

$$\boldsymbol{E} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{F} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} \frac{1}{b} & 0 \end{bmatrix}, H = -\frac{a(\cdot)}{b}. \; \boldsymbol{E\Phi_1} + \boldsymbol{F\Phi_3} =$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \\ \phi_{21} & \phi_{22} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} \phi_{31} & \phi_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{25}$$

$$\boldsymbol{E\Phi_2} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{13} \\ \phi_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\boldsymbol{\Phi_3\Phi_1^{-1}\Phi_2} = \begin{bmatrix}\phi_{31} & \phi_{32}\end{bmatrix}\begin{bmatrix}\phi_1 & \phi_2\\ \phi_{21} & \phi_{12}\end{bmatrix}^{-1}\begin{bmatrix}\phi_{13}\\ \phi_{23}\end{bmatrix} = -1,$$

$$\boldsymbol{E\Phi_{r1} + F\Phi_{r3}} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\phi_{r1} & \phi_{r2}\\ \phi_{r21} & \phi_{r12}\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}\begin{bmatrix}\phi_{r31} & \phi_{r32}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

According to (11), we assign the following coefficients as free ones: ϕ_1, ϕ_2, ϕ_{r1}, ϕ_{r2}. Then, solving (25), we obtain:

$$\boldsymbol{\Phi_1} = \begin{bmatrix}\phi_1 & \phi_2\\ 1 & 0\end{bmatrix},\ \boldsymbol{\Phi_2} = \begin{bmatrix}-\phi_2\\ 0\end{bmatrix},\ \boldsymbol{\Phi_3} = \boldsymbol{\Phi_{r3}} = \begin{bmatrix}0 & 1\end{bmatrix},\ \boldsymbol{\Phi_{r1}} = \begin{bmatrix}\phi_{r1} & \phi_{r2}\\ 1 & 0\end{bmatrix}. \tag{26}$$

Substituting (24), (26) into (5) and (6) and taking into account (22) we obtain the object control algorithm (21):

$$\begin{aligned}
&\dot{x}_{\phi 1} = \phi_1 x_{\phi 1} - \phi_2 \varepsilon,\\
&\dot{x}_{\phi 2} = x_{\phi 1},\\
&u = \frac{\phi_1}{b} x_{\phi 1} + \left(\frac{a_1}{b x_{02}}\cos x_{02}\right) x_{\phi 2} + \frac{\phi_{r1}}{b}\tilde{x}_{01} + \frac{1}{b}\left(\phi_{r2} + \frac{a_1}{x_{02}}\cos x_{02}\right)\tilde{x}_{02} - \frac{\phi_2}{b}\varepsilon,\\
&\tilde{x}_{01} = x_{01} - x_{\phi 1},\\
&\tilde{x}_{02} = x_{02} - x_{\phi 2},\\
&\varepsilon = y - y_0.
\end{aligned} \tag{27}$$

Substituting (24), (26) into (9) derives equations defining the processes in a closed control system:

$$\begin{aligned}
&\dot{x}_{\phi 1} = \phi_1 x_{\phi 1} + \phi_2 x_{\phi 2} + \phi_2 \tilde{x}_{02} - \phi_2 (y - f_y),\\
&\dot{x}_{\phi 2} = x_{\phi 1},\\
&\dot{\tilde{x}}_{01} = \phi_{r1}\tilde{x}_{01} + \phi_{r2}\tilde{x}_{02} + f_{x1},\\
&\dot{\tilde{x}}_{02} = \tilde{x}_{01} + f_{x2},\\
&\varepsilon = -x_{\phi 2} - \tilde{x}_{02} + (y - f_y).
\end{aligned} \tag{28}$$

As it can be seen, despite the nonlinearity of the object, processes in a closed system are represented by linear equations of filter standards. Let us consider free coefficients of reference filters. It follows from (28) that simulating the $(y - f_y)$ signal by a closed loop system is characterized by the transfer function of a low-pass filter of the second order:

$$W(s) = \frac{\phi_2}{s^2 + \phi_1 s + \phi_2} = \frac{\omega_\phi^2}{s^2 + \varsigma_\phi \omega_\phi s + \omega_\phi^2},\ \phi_1 = -\varsigma_\phi \omega_\phi,\ \phi_2 = -\omega_\phi^2, \tag{29}$$

where ω_ϕ denotes the natural frequency of the filter, ς_ϕ the coefficient of damping and s being the Laplace transform variable. Record in the (23) form is favorable for reference filters forms applying. For example, for the second-order Bessel filter $\varsigma_\phi \approx 1{,}73$. Changing ω_ϕ, it is possible at the given damping of ς_ϕ to select the desired speed of the filter and the entire closed-loop control system. The above reasoning is also valid for the filter through $\boldsymbol{f}_x$ simulating channels. Computer simulation of the system (20), (27) in the case of substituting a number of coefficients (19) demonstrated the correspondence of transient processes to the equations (28).

5 Motion Control of a Two-Link Manipulator

Consider the problem of a given movement of the operating point of a flat two-link manipulator of rotational type. The axis of rotation of the first link is fixed relative to the coordinate system; the axis of the second link is located on the first link.

The dynamics of this manipulator can be described in the form [18]:

$$\begin{bmatrix} \alpha+\beta+2\eta\cos q_2 & \beta+\eta\cos q_2 \\ \beta+\eta\cos q_2 & \beta \end{bmatrix}\begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix}+\begin{bmatrix} -\eta(2\dot{q}_1\dot{q}_2+\dot{q}_2^2)\sin q_2 \\ \eta q_1^2 \sin q_2 \end{bmatrix}+$$
$$\begin{bmatrix} \alpha e_1 \cos q_2+\eta e_1\cos(q_1+q_2) \\ \eta e_1\cos(q_1+q_2) \end{bmatrix}=\begin{bmatrix} k_1u_1 \\ k_2u_2 \end{bmatrix}, \qquad (30)$$
$$y_{01}=l_1\cos q_1+l_2\cos q_2,$$
$$y_{02}=l_1\sin q_1+l_2\sin q_2,$$

where α, β, η, e_1, k_1, k_2, l_1, l_2 denote coefficients depending on the parameters of links and drives and y_{01}, y_{01} being Cartesian coordinates of the working point on the plane. Eq. (29) can be represented in the state space, they can be transformed exactly or approximately to the form (4) with state-dependent coefficients (SDC form), and then the control algorithm can be synthesized according to (5)–(8). But this difficult task has not yet been solved and may be the object of future research. In this section, a simpler problem was posed, namely, using the linearized model of a two-link manipulator as an example to show the ease of synthesis of control algorithms for a two-channel LTI object by the proposed technique, which immediately provides the specified indicators of control quality. This refers to the usual linearization in the area of the working point. Introducing the notation $x_{01}=\dot{q}_1$, $x_{02}=\dot{q}_2$, $x_{03}=q_1$, $x_{04}=q_2$, and performing such linearization, we obtain the following equations:

$$\begin{aligned}
\dot{x}_{01} &= a_{11}x_{01} + a_{12}x_{02} + a_{13}x_{03} + a_{14}x_{04} + b_1u_1,\\
\dot{x}_{02} &= a_{21}x_{01} + a_{22}x_{02} + a_{23}x_{03} + a_{24}x_{04} + b_2u_2,\\
\dot{x}_{03} &= x_{01},\\
\dot{x}_{04} &= x_{02},\\
y_{01} &= c_{13}x_{03} + c_{14}x_{04},\\
y_{02} &= c_{23}x_{03} + c_{24}x_{04},
\end{aligned} \tag{31}$$

In (31), for brevity, the initial notation of variables for increments is preserved.

The general formulas being clumsy, the solution is made for a numerical illustration. Matrix for a numerical illustration we assign in the form (32):

$$\boldsymbol{A} = \begin{bmatrix} -5 & 0 & 0 & 0\\ 0 & -5 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 10 & 0\\ 0 & 10\\ 0 & 0\\ 0 & 0 \end{bmatrix}, \boldsymbol{C} = = \begin{bmatrix} 0 & 0 & -1 & -2\\ 0 & 0 & 3 & 4 \end{bmatrix}. \tag{32}$$

We take $b_1 = 10$, $b_2 = 10$. The zero elements in the first two lines of A are accepted for ease of calculation. Non-zero coefficients of the matrix $\boldsymbol{A}$ in (30) provide drives of links with a sufficiently high speed. In matrix $\boldsymbol{C}$, these coefficients are assigned arbitrarily. Substituting (31) into (7) we get>

$$\boldsymbol{E} = \begin{bmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{F} = \begin{bmatrix} 0 & 0\\ 0 & 0\\ 2 & 1\\ -1{,}5 & -0{,}5 \end{bmatrix}, \tag{33}$$

$$\boldsymbol{G} = = \begin{bmatrix} 0{,}1 & 0 & 0{,}5 & 0\\ 0 & 0{,}1 & 0 & 0{,}5 \end{bmatrix}, \boldsymbol{H} = = \begin{bmatrix} 0 & 0\\ 0 & 0 \end{bmatrix}.$$

By (11) blocks of freely assigned coefficients are:

$$\begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14}\\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} \end{bmatrix}; \begin{bmatrix} \phi_{r11} & \phi_{r12} & \phi_{r13} & \phi_{r14}\\ \phi_{r21} & \phi_{r22} & \phi_{r23} & \phi_{r24} \end{bmatrix}. \tag{34}$$

The analysis (9) proves that the variables x_ϕ and $\tilde{x}_0$ in their physical sense correspond to the state variables in (4). It follows that some of the coefficients in (34) denote interconnections between state variables. Since we have to ensure the independence of control channels of the reference system, the coefficients of interconnections should be taken as zero. As a result, we have the following blocks of assigned coefficients:

$$\begin{bmatrix} \phi_{11} & 0 & \phi_{13} & 0 \\ 0 & \phi_{22} & 0 & \phi_{24} \end{bmatrix}; \begin{bmatrix} \phi_{r11} & 0 & \phi_{r13} & 0 \\ 0 & \phi_{r22} & 0 & \phi_{r24} \end{bmatrix}. \tag{35}$$

Now, each pair of coefficients in any row (33) corresponds to a filter of a transfer function (23). Assigning $\omega_\phi = 2$, $\omega_{r\phi} = 4$, $\varsigma_\phi = \varsigma_{r\phi} = 1{,}73$, we get for>

$$\begin{bmatrix} \phi_{11} & 0 & \phi_{13} & 0 \\ 0 & \phi_{22} & 0 & \phi_{24} \end{bmatrix} = \begin{bmatrix} -3{,}46 & 0 & -4 & 0 \\ 0 & -3{,}46 & 0 & -4 \end{bmatrix},$$

$$\begin{bmatrix} \phi_{r11} & 0 & \phi_{r13} & 0 \\ 0 & \phi_{r22} & 0 & \phi_{r24} \end{bmatrix} = \begin{bmatrix} -6{,}92 & 0 & -16 & 0 \\ 0 & -6{,}92 & 0 & -16 \end{bmatrix}. \tag{36}$$

The remaining elements of the reference filters are calculated by (8). The resulting matrices of a reference filter are recorded as:

$$\boldsymbol{\Phi_1} = \begin{bmatrix} -3{,}46 & 0 & -4 & 0 \\ 0 & -3{,}46 & 0 & -4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \boldsymbol{\Phi_2} = \begin{bmatrix} 8 & 4 \\ -6 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{\Phi_3} = \begin{bmatrix} 0 & 0 & -1 & -2 \\ 0 & 0 & 3 & 4 \end{bmatrix},$$

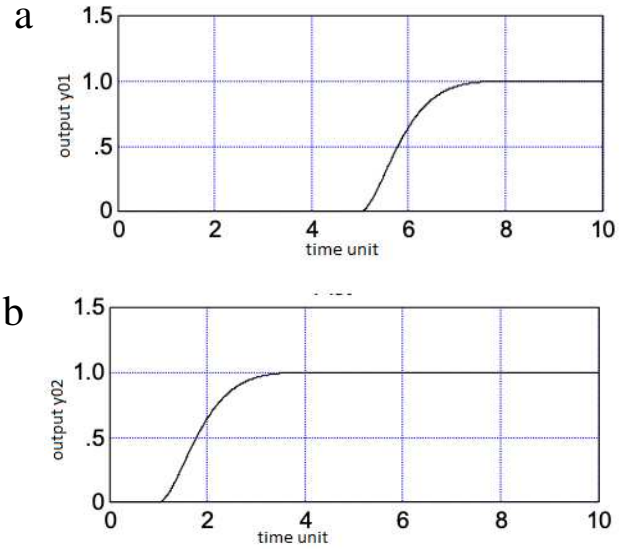
$$\boldsymbol{\Phi_{r1}} = \begin{bmatrix} -6{,}92 & 0 & -16 & 0 \\ 0 & -6{,}92 & 0 & -16 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \boldsymbol{\Phi_{r3}} = \begin{bmatrix} 0 & 0 & -1 & -2 \\ 0 & 0 & 3 & 4 \end{bmatrix}. \tag{37}$$

Solving the system of algebraic equations (8) when the given coefficient blocks are known, may be of some difficulty. Substituting (33), (37) into (5) and (6), we obtain the equations of the object control algorithm (31), (32) in the form>

$$
\begin{aligned}
&\dot{x}_{\phi 1} = -3{,}46x_{\phi 1} + 8\varepsilon_1 + 4\varepsilon_2,\\
&\dot{x}_{\phi 2} = -3{,}46x_{\phi 2} - 6\varepsilon_1 - 2\varepsilon_2,\\
&\dot{x}_{\phi 3} = x_{\phi 1},\\
&\dot{x}_{\phi 4} = x_{\phi 2},\\
&u_1 = 0{,}154x_{\phi 1} - 0{,}192\tilde{x}_{01} - 1.6\tilde{x}_{03} + 0{,}8\varepsilon_1 + 0{,}4\varepsilon_2,\\
&u_2 = 0{,}154x_{\phi 2} - 0{,}192\tilde{x}_{02} - 1{,}6\tilde{x}_{04} - 0{,}6\varepsilon_1 - 0{,}2\varepsilon_2,\\
&\tilde{x}_{01} = x_{01} - x_{\phi 1}, \qquad \tilde{x}_{02} = x_{02} - x_{\phi 2},\\
&\tilde{x}_{03} = x_{03} - x_{\phi 3}, \qquad \tilde{x}_{04} = x_{04} - x_{\phi 4},\\
&\varepsilon_1 = y_1 - y_{01}, \qquad \varepsilon_2 = y_2 - y_{02}.
\end{aligned}
\tag{38}
$$

Fig. 6 depicts the transition functions obtained on a computer model of a system (31), (32), (38) where coordinates of a given position of the operating point y_1, y_2 were changed in a stepwise way. In one second, the reference y_1 was changed, in 5 seconds – the reference y_2 was changed.

Figs. 6*a* and 6*b* demonstrate the changing of the actual coordinates of y_{01} and y_{02}, and Figures 6*c* and 6*d* demonstrate the changes in the control actions on the first and second drives during these tasks simulation.



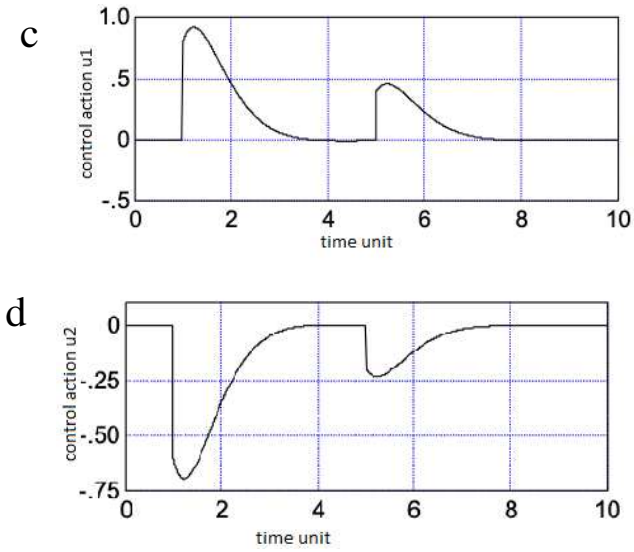


Figure 6

Simulating a unit step change of reference signals y_1, y_2 by the control system(31), (32), (38): a–b) changing y_{01}andy_{02}, c–d) changing the control actions $u_1$andu_2.

Fig. 7 shows the simulation by the system of unit step changes of f_{x1}, f_{x2} perturbations. These perturbations simulate changes in the load on the drives.

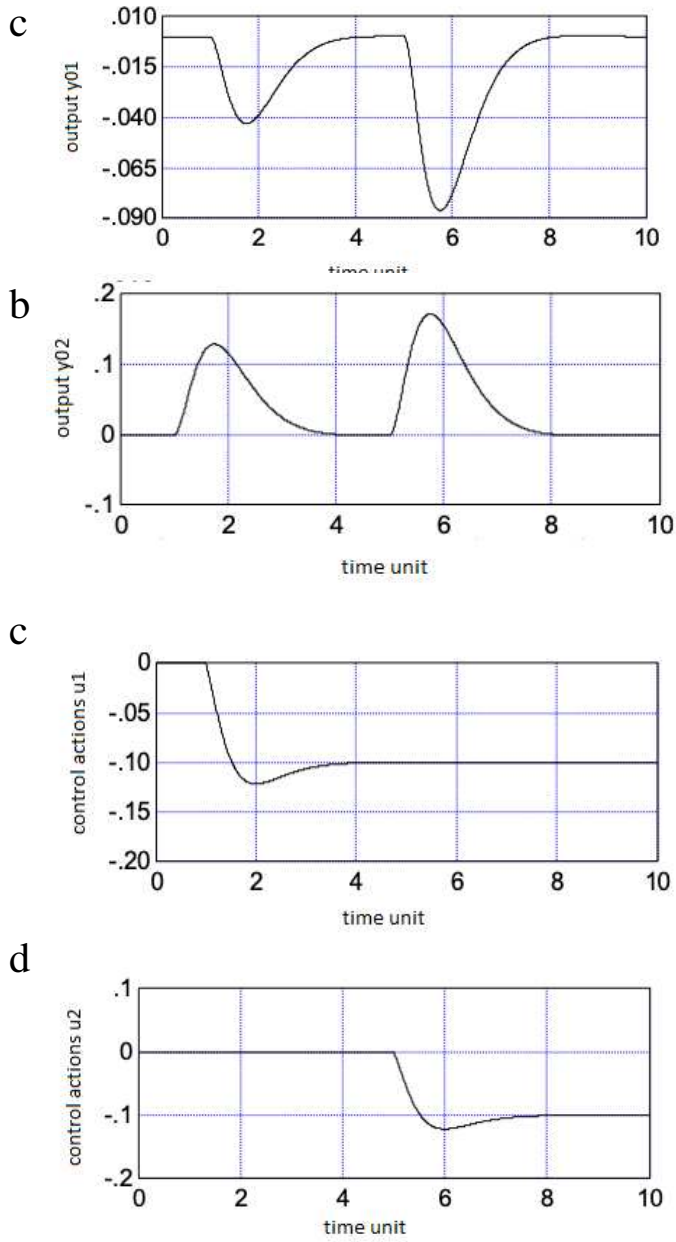


Figure 7
Simulating a unit step change of reference signals y_1, y_2 by the control system (31), (32), (38): a–b) changing y_{01} and y_{02}, c–d) changing the control actions u_1 and u_2

Based on Figs. 6 and 7 it is seen that the complete self-regulation through channels has been achieved: references y_1, y_2, outputs y_{01}, y_{02} and zero static control error. The control processes dynamics satisfies a given dynamics of reference filters. Obviously, autonomy is violated if the actual parameters of the control object do not match their calculated values, but the control error is always zero. This has been revealed by research methods on the computer model of the system.

Thus, a new method for compensating dynamics and perturbations of the object is considered, which allows developing the structure and parameters of a control algorithm for a multi-channel linear object. Here, the standard forms of low-frequency filters are proposed to be used and the effectiveness of their application is demonstrated. This synthesis technique application for multichannel objects is considered on the example of a two-link manipulator represented by its linearized equation. The calculation of the regulator for a numerical example of such an object has highlighted the complexity of solving a problem of selecting the remaining coefficients of reference filters when the blocks of assigned coefficients are given. Analytical research and computer modeling of the considered examples proved the correspondence of transients in control systems to a given dynamics in case of external actions, the autonomy of the SISO systems through separate control channels, zero static error and the robustness of the obtained systems. It is revealed that these properties are achieved immediately, without applying the trial and error method. The ease of the initial data assigning for the synthesis of systems should be noted. The drawbacks of the algorithms are the need for information on the state variables or higher complexity when using state observers. Generally, the issues of free coefficients determining and techniques for non-free coefficients of reference filters identifying have not been resolved yet. Once completely developed, the controller shall be applicable to robotic application with complex geometry constraints, such as plasma processing, or even for medical applications, such as automated bone surface milling for implants [33].

Conclusions

The new method of synthesis of a robot manipulator motion-control system by dynamics and perturbations compensation has been proposed. Under certain assumption, this method allows to develop the structure and parameters of a control algorithm for a multi-channel linear object. The possibility of applying the proposed algorithm to control nonlinear objects represented in the form of SDC is demonstrated. On the example of a simple SISO system of first order, duplicating the general structure, but giving a simple analytical solution, the sequence of synthesis of the control algorithm, and the principles of choosing free coefficients of reference filters for the general case are considered. On the example of the task of controlling the rotation of a link of a manipulator, the synthesis of a control algorithm for a nonlinear object is considered and the method of selecting free coefficients for a link of the second order is described. The obtained control

algorithm can be applied to automate a wide class of industrial facilities. The proposed method of structural synthesis can be used to develop effective motion control systems for robot manipulators. The results of the research are of significance for a range of application domains, developing control algorithms for production sites with mechatronic systems.

Acknowledgement

This work was supported by the Ministry of Education and Science of the Republic of Kazakhstan under the target financing program for the 2017-2019 years by the sub-program 0006/ PCF-17 "Manufacture of titanium products for further use in medicine". T. Haidegger is supported through the New National Excellence Program of the Hungarian Ministry of Human Capacities. T. Haidegger is a Bolyai Fellow of the Hungarian Academy of Sciences. The final grammatical improvements of the article were supported by the V4+ACARDC - CLOUD AND VIRTUAL SPACES grant.

References

[1] T. Haidegger, G. S. Virk, C. Herman, R. Bostelman, P. Galambos, G. Györök, I. J. Rudas: Industrial and Medical Cyber-Physical Systems: Tackling User Requirements and Challenges in Robotics, In Recent Advances in Intelligent Engineering, Springer, Cham, 2020, pp. 253-277

[2] J. Somló, G. Dávid Varga, M. Zenkl, B. Mikó: The „Phantom" Delta Robot A New Device for Parallel Robot Investigations, Acta Polytechnica Hungarica 15, No. 4, 2018, pp. 143-160

[3] G. Haidegger, I. Paniti: Issues in Manufacturing Automation & Robotics within the Past 4 Decades and Vision for the Next, Proceedings of IEEE International Conference on Intelligent Engineering Systems (INES), 2019, pp. 23-28

[4] R. C. Tucker, Jr., editor: Introduction to Coating Design and Processing, Thermal Spray Technology, Vol. 5A, 2013, pp. 76-88

[5] D. L. Alontseva, A. L. Krasavin, A. T. Kadyroldina, A. T. Kussaiyn-Murat, D. M. Nurekenov, Ye. T. Zhanuzakov, N. V. Prokhorenkova: Development of the Robotic Microplasma Spraying Technology for Applying Biocompatible Coatings in the Manufacture of Medical Products, Proceedings of AIS 2017 – 12th International Symposium on Applied Informatics and Related Areas, Székesfehérvár, Hungary, 2017, pp. 45-48

[6] L. Márton, Z. Szántó, T. Haidegger, P. Galambos, J. Kövecses: Internet-based bilateral teleoperation using a revised time-domain passivity controller, Acta Polytechnica Hungarica, Vol. 14, No. 8, 2017, pp. 27-45

[7] E. Horvátha, C. Pozna, R.-E. Precup: Robot Coverage Path Planning Based on Iterative Structured Orientation, Acta Polytechnica Hungarica 15, No. 2, 2018, pp. 231-249

[8] R. Zhao: Trajectory planning and control for robot manipulations. Robotics [cs.RO], thesis - Université Paul Sabatier - Toulouse III, 2015, pp. 158, URL: https://tel.archives-ouvertes.fr/tel-01285383v2/document

[9] R. Campa, C. Ramirez, K. Camarillo, V. Santibanez, I. Soto: Motion Control of Industrial Robots in Operational Space: Analysis and Experiments with the PA10 Arm, Advances in Robot Manipulators, Ernest Hall (Ed.), 2010, pp. 417-442, ISBN: 978-953-307-070-4

[10] K. M. Lynch, F. C. Park: Modern Robotics Mechanics, Planning and Control, Cambridge U. Press, 2017, pp. 642, ISBN: 978-110-715-630-2

[11] A. Nemeikšis: Trajectory modulation of 2.5 degree of freedom robot arm, Scientific researches and their practical application, Proceedings of Modern state and ways of development, Vol. 5, № 3, 2014, pp. 3-11, URL: https://elibrary.ru/item.asp?id=22270213

[12] S. Yurish, editor: Advances in Robotics and Automatic Control: Reviews, Book Series. Vol. 1, 2018, pp. 404, ISBN: 978-84-09-02449-0

[13] Zh.-G Liu., J.-M. Huang: A New Adaptive Tracking Control Approach for Uncertain Flexible Joint Robot System, Proceedings of International Journal of Automation and Computing, Vol. 12(5), 2015, pp. 559-566, DOI: 10.1007/s11633-015-0898-6

[14] Zh. Wang, W. Liu, B. Cui, J. He, Zh. Li, Yo. Zhao: Research on Robot Surface Tracking Motion Based on Force Control of Six-Axis Wrist Force Sensors, Advances in Mechanical Engineering, ID: 249696, 2014, pp.1-9

[15] Gy. Eigner: Control of Physiological Systems through Linear Parameter Varying Framework, Acta Polytechnica Hungarica, Vol. 14, No. 6, 2017, pp. 185-212

[16] ISO 8373:2015 ISO. 8373:2015 robot vocabulary standards. Manipulating industrial robots

[17] M. Vukobratovic, V. Potkonjak: Scientific Fundamentals of Robotics 1, Dynamics of Manipulation Robots: Theory and Application, Springer-Verlag, Vol. 1, 1982, ISBN 978-3-642-81854-7

[18] F. L. Lewis, S. Jagannathan, A. Yesildirek: Neural Network Control of Robot Manipulators and Nonlinear Systems, Taylor & Francis, 1999, ISBN 0-7484-0596-8

[19] B. Armstrong, O. Khatib, J. Burdick: The explicit dynamic model and internal parameters of the PUMA 560 arm, Proc. IEEE Conf. On Robotics and Automation, 1986, pp. 510-518

[20] J. Somló, B. Lantos, P. T. Cát: Advanced Robot Control, Akadémiai Kiadó, Budapest, 2002, p. 428, ISBN10 9630573504

[21] A. Isidori: Nonlinear Control Systems, Berlin: Springer, 1995, pp. 549, DOI 10.1007/978-1-84628-615-5

[22] N. K. Khalil: Nonlinear Systems, 3rd ed. New Jersey: Prentice Hall, 2002, pp. 575, DOI: 10.3969/j.issn.1006-7043. 201112013

[23] B. Friedland: Quasi Optimal Control and the SDRE Method, Proc. 17'IFAC Symp. on Automatic Control in Aerospace. Toulouse, France, 2007

[24] S. Skogestad, I. Postlethwaite: Multivariable Feedback Control, 2nded. Wiley-Interscience, 2005, p. 592, ISBN: 0470011688, 9780470011683

[25] J. S. Shamma, J. R. Cloutier:Existence of SDRE Stabilizing Feedback, IEEE Transactions on Automatic Control, No. 6 (3), 2003, pp. 513-517 DOI: 10.1109/TAC.2002.808473

[26] C. Dorf Richard, H. Bishop Robert: Modern Control Systems, 12th ed.Prentice Hall, 2011, p. 1111

[27] A. G. Aleksandrov: K analiticheskomu sintezu regulyatorov [To analytical synthesis of regulators], Automatics and telemechanics, No. 6, 2010, pp. 3-19 (in Russian)

[28] A. M. Letov: Analiticheskoe konstruirovanie regulyatorov I, II, III [Analytical design of regulators I, II, III], Automatics and telemechanics, No. 4, 1960, pp. 406-411; No. 5, 1960, pp. 561-568; No. 6, 1960, pp. 661-665 (in Russian)

[29] R. Isermann: Digital Control Systems, Springer Science & Business Media, 2013, p. 566, ISBN 3662023199, 978366202319

[30] A. I. Károly, R. Fullér, P. Galambos: Unsupervised Clustering for Deep Learning: A tutorial survey, Acta Polytechnica Hungarica, Vol. 15, No. 8, 2018, pp. 29-53

[31] G. K. Shadrin: A Physics-Based Approach to Control Systems Design Using Compensation of Controlled Plant Dynamics and Perturbations, Proceedings of Automation and Remote Control, Vol. 77, No. 7, 2016, pp. 1151-1162

[32] G. K. Shadrin, D. A. Porubov, M. G. Shadrin: Sintez algoritma upravleniya dvizheniem dvuhkolesnogo robota metodom kompensacii dinamiki ob"ekta i vozmushchenij [Synthesis of two-wheeled robot motion control algorithm by compensating object dynamics and disturbances], Proceedings of Automatics & software enginery, No. 4 (22), 2017, pp. 10-17 (in Russian)

[33] T. Haidegger: Autonomy for Surgical Robots: Concepts and Paradigms, Proceedings of IEEE Transactions on Medical Robotics and Bionics 1, No. 2, 2019, pp. 65-76, DOI 10.1109/TMRB.2019.2913282

Advantages of Simulating Logistics Processes

Antal Balogh[1], Balázs Gyenge[1], Ágnes Szeghegyi[2], Tímea Kozma[1]

[1]Department of Operations Management and Logistics, Szent István University, Páter Károly út 1, H-2100 Gödöllő, Hungary, balogh.antal@diego.hu, bgyenge@interm.gtk.gau.hu, kozma.timea@gtk.szie.hu

[2]Keleti Faculty of Business and Management, Óbuda University, Tavaszmező utca 15, H-1084 Budapest, Hungary, szeghegyi.agnes@kgk.uni-obuda.hu

Abstract: Manufacturing and service providing companies perform or make others perform logistics tasks. This area is of great importance in everyday operations, and managers do their best in choosing operations and processes. Appropriate logistics processes and decisions mostly determine the companies' successful operation, profitability, profit-making ability, and their costs. However, an important step before defining them is missing in most cases, namely getting to know the processes to execute, and the best method for that purpose is to get a preliminary simulation of the processes. The goal of the study is to highlight the advantages of modeling, what unnecessary introductory processes and financial losses can be avoided through simulation, and underline how deeper knowledge we can get with the help of this method.

Keywords: Lean; simulation; SME; logistics processes; loss sources

1 Introduction

This study forms a part of the research series "Logistics Problems of SMEs, Simulation of Economic Processes" focusing on the difficulties and challenges of the SME sector. The intense growth of market competition in the logistics areas requires a fast response and a flexible attitude if a company wants to be one of the best or to maintain its position. R. A. Thietart and R. Vivas said their study "One dimension which measures the quality of a strategic response is its speed of implementation. Recognizing this, to ensure a good (fast) strategic response, fast communication of information is imperative" [1]. Many companies have developed strategies that include investing heavily in information technology (IT) in order to enhance their performance [2].

For us our study inventory management represents a very important segment of business conducted by modern enterprises and as such, it is crucial for the success

of an enterprise's business operations [3]. Based on the Lean approach and a very advanced process analysis method (simulation) this research makes suggestions for a deeper reflection on warehouse processes and their change targeted through the example of a well-known Hungarian company concerning the commercial and service sectors.

According to Sarkis et al. (2010), the pressure of stakeholders companies usually apply three types of measures or requirements: eco-design procedures, resource reduction procedures, and implementation of environmental management system procedures [4]. Our invention efforts mainly belongs to the second one.

The actuality of the topic is confirmed by the fact that development is so intensive that nowadays it is more and more frequent that demands are not always expressed by the customers but the service providers have to recognize them proactively and provide the customers with some solution packages. We require a wide range of related services including comfort functions (IT transparency solutions) for our customers to reduce and recognize inflexibility, bottlenecks and halts in service processes. We also need analytical tools that extend the decision-maker's insight and decision dimensions to both the level of collaboration and the spatial and temporal progress of functions dynamically and in more dimensions at a much higher level of human sense. According to FatmirAzemi et al. (2019) the digital transformation of manufacturing and such trends like IoT, Industry 4.0, new way of data analytics, change the methods of data collection, performing skilled work or predicting customer behavior. Smart factories using integrated IT systems can provide sufficient data for both sides of the supply chain in a much more convenient way, which can significantly increase production capacity [5].

According to the data by the Hungarian Central Statistical Office, as of 31 December 2017, in Hungary 1,719,601 small and medium-sized enterprises were in operation, of which 529,608 were partnerships. This means a decrease of 3 per cent compared to a year earlier [6], however, almost 100% of active companies have no experience how to simulate economic processes while some foreign competitors have already been using this techniqueactively.

2 Methods and Resources

The present study is based on the book *AnyLogic 7 in Three Days* by Ilya Grigoryev and the general principles of Lean methods which are presented in the next chapter. The equivalence of simulation and warehouse processes may prove that the results achieved can not only be positively assessed (or not) after a practice"test" but the hypotheses formulated can also be verified by modeling without introducing any new processes. In addition to the results obtained (with minor modifications) the method can be put into practice in other areas regardless its applicationis limited to either a particular company or even to a specific storage

activity. Thanks to the actual message of the topic it could give useful advice to any company providing service or performing manufacturing activity. Efficient and optimal operation is actually available in several ways, however, according to the authors, the goal of this study is to present the Lean concept and the integration of the advantages of the simulation method in the supply chain in a narrower and a wider sense.

3 Lean as a Main Actor in Improving Physical Processes

With the development of logistics and applying management tools and their approaches, processes can be optimized which can have effects on the workforce, costs and competitiveness. In addition to these they support the maximization of the ownership and customer value besides several additional competitive advantages [7]. As Cifra, et. al. (2019) say "the concept of lean is based on the production of a flexible response to customer and demand" [8]. By applying the Lean principles efficient and economical operation can be realized.

Principles of Lean [9]:

- Value – specification
- Value stream (value process) – analysis
- Flow – realization
- Principle of pull – productionplanning
- Perfection – introducing Kaizen

The source and base of the Lean approach is the Toyota Production System (TPS). One of its primary philosophical cornerstones (among many other methodologies) is constant development and the Kaizen principle. Used chiefly in the area of manufacturing (nowadays a complex methodology on its own including standardization, idea realization, motivation, and automation) it aims to reduce or eliminate losses via continuous improvement in small steps, especially through modifications capable of increasing lead times, and the system's efficiency, flexibility and responsiveness. [10]

A summary in four points [11]:

- All employees should be authorised to develop their company according to their ability and responsibility.
- Toyota's production system is based on its constant development and respect to people as the lead philosophy.

- Lean management is a loss prevention strategy and not a cost reduction strategy.
- Lean practices should be closely linked to the supply chain process of the company.

Lean is a process management approach whose most important goal is to create customer value and eliminate waste. We call wastage all the activities, operations and processes the buyer refuses to pay for. Lean separates processes as:

- value adding (to maximize);
- necessary, but non value adding (to minimize);
- unnecessary and non value adding so to say wasteful activity (to eliminate).

For many people it is very difficult to understand the Lean philosophy, and what is even more difficult is to follow it. For example it is hard to apply the three ways of thinking mentioned above at the same time but in different terms; or maybe more difficult to put the customer's approach ahead of ours; and it is difficult to find the sources of loss even if we are convinced that our own system is good. As a matter of fact, in Western European thinking, localization often fails at this point and that is why many people do not understand what the real essence of Lean systems is. It is built upon modesty, perseverance, and above all, the will to give; this attitude is generally hard to achieve with the necessary morale to provide so as to gain the basics "a philosophy is more than a set of methods". If the necessary mental bases are supported by learning and knowledge then there will certainly be constant improvements making the goals of Lean available.

"Using fewer employees, fewer tools, less time and less space, so less resources the Lean system can provide the customer with (more) value."[12]

For this purpose it is very important to investigate thoroughly the most common sources of loss. By studying literature it can be stated that certain losses are more crucial than others and can be classified under specific categories. Taiichi Ohno and Womack set up the following categories (Based on Womack – Jones, 1996, supplemented and significantly paraphrased):

Lean's 7 major loss sources:

1. Stocks (overstocking): stocks, especially too high stocks, do not add value.
2. Shipping loss: the customer is not willing to pay for the delivery in all cases, this significantly increases the lead time of value adding processes and can make a lot of surpluses (e.g.journey, time).

3. Failure, waste: all forms of waste are loss and of Muda Type 2, their consequences can be additional sources of loss, not to mention the dissatisfied customer.
4. Waiting: all the time spent in any manufacturing system not adding value while the resources are getting less and less in other words if the capacities are staying unused which is also a loss.
5. Unnecessary activities, over-processing: all work completed for which the customer is unwilling to pay for, but somehow it is necessary for some reason, for instance the price will be the same no matter how many times a product should be checked or taken.
6. Movement loss, wasteful operation: any movement someone has to do, but not necessarily or even all the movements are unncessary.
7. Overproduction (over-ordering): any surplus not related to a registered external/internal need is more than the original need.

4 Presentation of the Processes Analysed

The company involved in the research has several warehouses; out of them was selected for analysis the one with an intense material flow requiring manual handling of goods. The processes used have been unchanged for a long time and the company does not currently have Lean solutions. The analysis focused on what and how much can be saved in terms of resources if we propose changes to process development.

The study covers the main processes with a dual purpose. On the one hand the optimal use of resources, and on the other hand the reduction of the time spent on performing the processes, i.e. time optimization.

Among the warehouse workers, 15 manual workers perform daily tasks. According to the daily routine the loading tasks are carried out by 5 people, the goods are picked by 9 people who also carry out the reloading of goods and 1 person who participates in loading the trucks as well as checking.

The main warehouse processes are the following:

- Loading–receiving the incoming goods and placing them to storage.
- Reloading– changing storage (in case of product type change or fashion change).
- Order picking–commission.
- Shipping–delivery control, loading of trucks.

4.1 Loading (and reloading)

A truck of incoming goods means approximately 4,000 units of goods (varying between 3,000 and 5,000). This amount should be moved manually during loading and placed in the designated storage by variety, size and type. In the model this was formalized in a way that some 50 (30-50) types x 5 size categories ≈ 250 carpets 'varieties'and 10-13 items are in the shipment.

Our model ignored the rare types and unusual sizes so the number of the items was slightly greater than in reality. (As for cca. 10, we compensate that by the picking tool's higher logistics capacity so there will be no significant deviation in the run. In the model the varieties (250-400), the types (30-50) and the items (10 - 13) will fluctuate to describe reality better, so will the amount of shipment between 3,000 and 5,000 items:

Table 1

Seasonal fluctuation (empirical data) (Source: own edition based on the sales data of the investigated company)

period of time	**frequency of shipments**	**size distribution of items**	**distribution of shipments (carpet types)**
month 1 "beginning period"	2-3 trucks/week	**the same (months 5-8)**	(30-50) uniform distribution
months 2-4 "introductory period"	3-4 trucks/week	60/80 – 5% 80/120 – 60% 120/170 – 5% 160/230 – 28%200/300 – 2%	(30-50) uniform distribution
months 5-8 "preparation period"	2-3 trucks/week	60/80 – 11% 80/120 – 13% 120/170 – 11% 160/230 – 60%200/300 – 5%	(30-50) uniform distribution
months 9-12 "season period"	5-10 trucks/week	**the same(months 5-8)**	(30-50) uniform distribution

Loading means receiving and checking goods when the physical quantities are also compared on the basis of the delivery documents and the amounts of order. The receiving process can be divided into sub-processes having their own time requirements:

- loading from truck to trolley and sorting,
- moving the trolley to the assigned storage,
- loading into storage.

After 25 measurement experiments, we determined the loading time of a complete shipment that it takes 8 hours for 5 workers on average. 2 people load the goods from the transport vehicle to the trolley, 1 person moves the trolley to the storage and brings the empty trolley back to the truck, and 2 people load them to the final

storage. From the transport vehicle to the storage the products travel 45 meters on the trolley on average.

Applying the Lean approach we examined the movement loss and wasteful operations during exploring the loss sources. We observed the movement from the transport vehicle to the trolley and then to the storage asa wasteful activity meaning handling of the products twice which is loss in both in its parts and entirely (movement loss) since it contains unnecessary movements compared to the strategy of "arriving at the destination"in Lean.

Each trolley can move 50 items (on average) and the same amount can be placed in the storage. On average 2 workers place 50 items into the trolley in 2 minutes, then the time required for moving to the storage is some 2 minutes and 15 seconds on average and the loading time to the storage is 1 minute and 45 seconds.

It takes 6 minutes for a trolley of goods to get into the storage. So for 5 workers loading 4,000 product items means 80 (rounds) x 6 minutes = 480 minutes, altogether 8 hours to move to the storage.

In the first round of our development we used framed pallets which we can put directly into storage so at least one step became superfluous and others were shortened. Using these the following results were recorded:

The time for loading onto a trolley is still 2 minutes, however, moving to the storage with a forklift together with piling the framed pallets takes 1 minute and 15 seconds and there is no need for further movement (totally 3 minutes and 15 seconds).

For 4,000 carpets 80x 3 min 15 sec = 260 sec adding up 4 hours and 20 minutes are required for loading. Additional saving occurs namely 2 workers less were needed for unloading since the manual loading of the storage can be saved! (Although the cost of the forklift as well as the cost of investing framed pallets arise.) The loading process is shown in Figure 1.

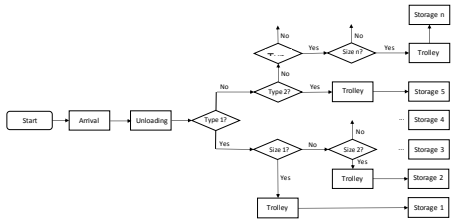


Figure 1

The loading process (flowchart) (Source: own edition)

In the later model, the (un)loading transaction times presented here are based on the second 'developed method' in a way that the average times statically specified earlier fluctuate by 10-15% around the averages above.

4.2 Order Picking – Commission

The warehouse in the analysis picks 8-10,000 items per day. This is done by 9 people in gross 8 hours daily. Performing this task takes net 7 hours. The processes of picking were examined within 10 working days and the analysis included the time of picking goods for each customer and measuring the time required for this. In this case. the length of time is hard to standardize because smaller customers order 50-100 items on delivery while larger ones do up to 5-600 units.

The picking process is usually performed by 2 workers, in some cases by 3 workers simultaneously. The physical allocation of the warehouse is one of the main sufferers of labor migration caused by the economic situation. Due to fluctuation, the newcomers are performing their tasks slowly with more mistakes thanks to the lack of defining products making the time of picking longer (Tthis is a waste failure which also has additional harmful and costly consequences.) That is why there is a lot of overtime typically in the last months of the year.

The trolley is considered the standardized unit capable of holding around 50 products. If 10,000 items are collected a day, 4 pairs of workers collect 2,500 items on 200 trolleys respectively. So picking an order on one trolley takes about 6 minutes on average.

The subject of the Lean study is the optimization of time and workforce concerning the picking process. Out of the 7 sources of loss, waiting is determined, as unused capacity (i.e. unused resources), or hidden waiting.

If a task is performed by more workers than necessary, the second or indeed the third person is a hidden unused resource. This assumption was also confirmed by restructuring the process carried out as follows.

In another case when picking was accomplished by 1 worker instead of 2, a similar process was measured 5 times. The benchmark was the standard time of picking performed by 2 people namely 6 minutes. When 1 worker did the activity of commission, the specific picking time of one trolley exceeded by 3 minutes meaning 60 per cent. However, if the second worker picks an order on a trolley during the same amount of time, then the two people will pick the order on a trolley together in 9 minutes instead of 12 minutes. It means saving 25% of time so to say saving 1 hour 45 minutes, 315 minutes instead of 420 minutes, for 200 trolleys per day.

In case of total workforce, the daily amount (tasks) can be performed in 7 hours instead of 9 even by 5 people. The picking process is illustrated in Figure 2.

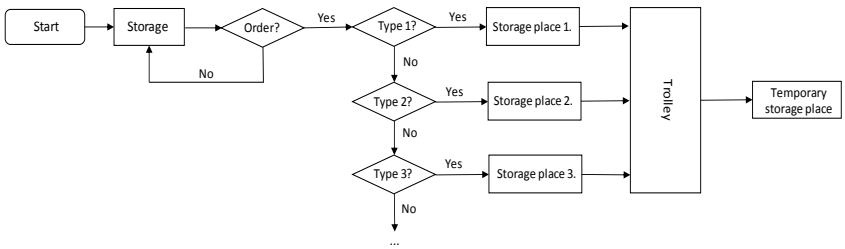


Figure 2
The flowchart of picking (Source: own edition)

4.3 Shipping–control

Both loading and goods control area common task of 1 manual worker and 1 office worker. Drivers arriving to load should wait to get controlled goods. As a result there is actual waiting time as a loss factor. In this case, not only the working hours of the warehouse workers have to be taken into account but also the drivers' who work under AETR (European Agreement Concerning the Work of Crews of Vehicles Engaged in International Road Transport)rules.

The control of 10,000 items takes about 6 hours and 40 minutes meaning 2 minutes per trolley (200 trolleys x 2 minutes = 400 minutes = 6 hours 40 minutes per 1+1 workers). If we double the resources spent on checking and shipping resulting in 2 more people dealing with checking, then the daily control time will be reduced to 3 hours and 20 minutes. This way the drivers cease to wait and the physical workload of the worker checking the items is reduced. We also note here that by extending the strategic partnership and shifting the control system to the supplier (based on Lean principles, introducing Poka Yoke systems and creating the proper motivation, but not described here), the whole process can be avoided and saved which falls into the category of unnecessary activity or over-processing. This can be a subject of further development.

In the study we stated that the amount of stocks in the modelled warehouse is overcalculated in current circumstances, and this case belongs to the loss category of stocking or rather overstocking. Measuring the rotation speed of the product line (28-32 days) the stock quantity can be significantly reduced and the overstocking experienced can be eliminated. Currently, one of the reasons for overstocking is supplier uncertainty, and the demand fluctuating and undetermined precisely. According to Lean principles a comprehensive supplier development can significantly reduce the uncertainties experienced in the first category. This option was not investigated in this development, although the amount of capital employed in stocking can be significantly reduced and the capacity of the warehouse with almost maximum storage utilization can be partly released, the internal reloading demand can also be decreased (unnecessary activity).

We could talk about shipping losses if certain shipments could be routed directly to individual partners (no chance for that), or due to bad communication the delivery covers unnecessary distances or is waiting unnecessarily. The latter can be analysed, but it rarely occurs in our case. The movement of goods inside the warehouse is the result of a change in the products' storage (reloading), which is entirely a loss, and movement loss because it does not add value to the product. This is mainly due to scarcity of capacity with the reasons previously mentioned.

Since we are talking about commercial activities and meeting the needs of the points of sales, over-processing equals to over-ordering. This loss source also has a close relationship with overstocking which can come from the supplier's order composition. Certain suppliers from the Middle East are unreliable in managing items the company did not order at all. Reducing this loss is possible with a medium-term strategy, which is not the target of our investigation.

The table below shows the potentials available in the development areas of the Lean approach discussed earlier. We made some comparisons and projected the following expected results as in Table 2.

Table 2

An overview of the results (Source: own edition based on the sales data of the investigated company)

	Benchmark	Before Lean concept		By Lean concept tools		Results	
		Time required	No of workers	Time required	No of workers	Time	Workers
Receiving (Reloading)	1 shipment (4,000 items)	480'	5	260'	3	**54%**	**40%**
Picking	10,000 items	420'	9	315'	5	**75%**	**55%**
Loading and checking	10,000 items	400'	2	200'	4	**50%**	**200%**
Total input		1,300'	16	775'	12	**60%**	**75%**

5 AnyLogic® - Simulation before Introducing New Processes

Modeling physical processes provides the user with benefits, which can only be learned from its own faults and losses if a company does not employ experts or modeling software. This system can model any phenomenon such as the spread of a virus and its consequences, or even logistics processes. In other words, we can

develop processes, design new systems, or even plan tool development (for example buying a trolley); simulating operation and results can be performed without risk.

This solution as a decision support system is useful in a company's life. A logistics provider company which operates a cross-docking warehouse is able to model the operation for resource planning meeting the needs and consequently optimizing its costs.

Modeling step-by-step [13]:

- exploring the real situation (realisation)
- determining the abstraction level (modeling perspective and its details)
- selecting the modeling language (formalism)

One of the options for modeling is Excel, which is one of the simplest versions of simulation. We call it an analytical model, which is formula based so the area of its application is very limited. Static dependencies are expressed by mathematical formul as though they are not suitable for modeling dynamic systems.

Simulation models are of a higher level than analytical models. Their development is much simpler than that of analytical systems. Simulation is almost the same as reality. Values are measured, entities are tracked and statistical analyses can be performed based on data measured. We can observe how the system behaves, and we can find 'playing', 'monitoring' and 'controlling' functions. As visualisation, simulation is a lot more than a presentation of numbers. Visual programming uses graphical elements to represent and build-up algorithms while focusing on the underlying logic of the application under development [14].

Three types of simulation systems:

- system dynamics simulation - used at macro level, high level abstraction
- discrete event (or event-driven) modeling – medium level abstraction
- agent-based modeling – from low to high abstraction levels, for simulating business-to-business competitions (mainly used at high level of realism and low level of abstraction as it really is)

There is also a fourth level in advanced simulation environments which is a free combination of these three types without restrictions termed as dynamic simulation.

To analyse a difficult process it is best to apply a complex method, which can simulate the sub-processes concerning the level of the simulation [15].

Two more important and longer processes out of three in the study are worth simulating by agent-based modeling technique. In case of initial simpler models it does not matter either we use entities or agents but later in the development we can integrate more complex behavioral patterns and more realistic decision points into the simulations with agents. Therefore, we recommend the more advanced agent based modeling.

Throughout our analysis presented here, we examined the process of receiving and picking of goods in detail. By analyzing the theoretical bases mentioned above the efficient operation became visible and expected concerning the processes' resource need (workers) and time, we wanted to test this 'development based on Lean concept' as well.

Although theoretical calculations showed positive results the question remains the same, namely whether the improvement can be predicted if uncertainties are taken into account (in a more realistic way).

In the modeled processes through formalization, the carpet varieties (size/type combinations, usually on a trolley) from which shipmentsare made were considered as agent objects. (Based on the distributions above the type, size, and specific quantity by type/size are these agents' own random characteristics.) These agents are picked up by another agent object (Pickup object), the truck entity whose role is virtual transport receiving according to the seasonal needs. The arrival of the truck agents was matched to the measured distributions, however, we must notice that in reality the distributions are determined by the consumption (demand) statistics of the consumers which we did not consider necessary to model in this simple case. Unloading takes place in the second framed area of the model (arrival area), this was the actual purpose of our simulation model to examine the cumulative effect of randomness on the whole receiving process. We did not use animation graphics in this model, only the changes in quantitative and temporal transitions were observed. In our case, the former (the change of volume) is the goal of the simulation and the latter (visualisation) is intensifying the convincing power of simulation, leading to a further realization.

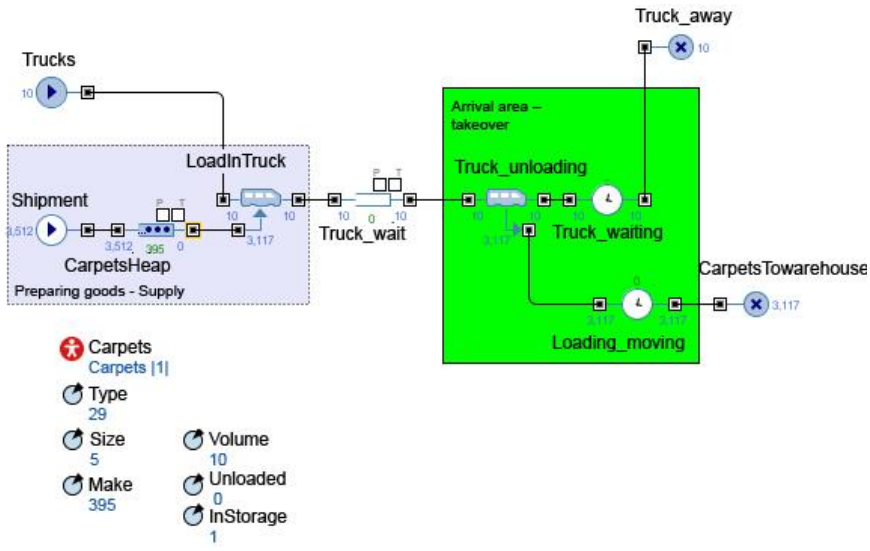


Figure 3

The flowchart of the simulation model(Source: AnyLogic simulation running chart)

As a second step the process of picking can also be modeled, for example we can examine the efficiency of commission tasks for 1 or 2 workers concerning different amounts to pick and consequently how this affects the time requirement of the whole process.

Our study says that modeling reveals what we are not able to analyse(simulate) in our everyday processes at all, or with more risks and more costs, or only in a speculative way, is the effect and relationship on the development analysed due to seasonal variations. It means eventually if we perform a task improved with the help of Lean management tools at any time of the year, this will prove or deny our assumptions depending on the characteristics of that period. Seeing this dilemma we realised that it would be worth analysing the processes with an effective simulation tool while placing our empirical knowledge of the process into a wider context with different loads.

Based on experience, AnyLogic® simulation is able to simulate any or all of the periods during a short period of modeling, this way we can analyse these experiments over various periods of time and we can test our hypotheses in a realistic and spectacular way without serious decisions and costly investments.

Conclusions

A simulation system makes it possible to carry out process development without involving expensive practical experiments. It is also an excellent tool to validate the correctness of our existing processes and concepts, taking into account changing circumstances and fluctuations due to unpredictable demand. It is not

only an advantage but also an important statement that with this method we can get much better understanding and deeper knowledge about our operation linked to our physical processes, both in running the model and in answering the questions arising inevitably in developing the model. In running our model we stated that our quantitative ideas anticipate realistic improvements, hidden bottlenecks or congestion do not block the achievement of estimated results, and seasonal fluctuations are manageable.

Furthermore the material flow processes can be simulated both inside and outside the company (in the supply chain) provided we know the suppliers or buyers' capabilities, flexibility and unpredictability. It is good to know what our supply chain can do in the ongoing competition for customers and market positions. We found that in modeling economic processes the main strength of the simulation model isnot primarily in the specific answer or in the evidence given to the question but mainly, as in our experience, through modeling we met more new questions than we thought earlier. For example, what the role of product distribution in service is, what the real cost of availability (in terms of investing assets and storage or work organization) is or what degree of volatility and intensity of customer demand causes congestion in orders.

As a conclusion, we are capable of defining changes without simulation, and we may possibly have (good or bad) expectations, however, with simulation (modeling) we can also get the probability of the expected events and discover completely new perspectives to the problem.

References

[1] R. A. Thietart, R. Vivas: Strategic intelligence activity: The management of the sales force as a source of strategic information: Strategic Management Journal, Vol. 2, Issue1, January 1981, Pages 15-25 https://doi.org/10.1002/smj.4250020103

[2] M. J. Tippins, R. S. Sohi: IT competency and firm performance: is organizational learning a missing link?: Strategic Management Journal, Vol. 24, Issue8, August 2003, Pages 745-761https://doi.org/10.1002/smj.337

[3] M. Stojanović, D. Regodić: The Significance of the Integrated Multicriteria ABC-XYZ Method for the Inventory Management Process, Journal: Acta Polytechnica HungaricaVol. 14, No. 5, 2017 http://uni-obuda.hu/journal/

[4] Sarkis. J.;Gonzalez-Torre. P. & Adenso-Dian. B.: Stakeholder pressure and the adoption of environmental practices: The mediating effect of training. Journal of Operations Management.28(2). 163-176, ISSN 0272-6963

[5] F. Azemi, G. Šimunović, R. Lujić, D. Tokody, Z. Rajnai: The Useof Advanced Manufacturing Technologyto ReduceProduct Cost: Acta Polytechnica Hungarica Vol. 16, No. 7, 2019 https://www.uni-obuda.hu/journal/Azemi_Simunovic_Lujic_Tokody_Rajnai_94.pdf

[6] KSH: STADAT 6.3.2.1.1. Tite in engish:*Number of registred enterprices* - A regisztrált vállalkozások száma (2013–2017) http://www.ksh.hu/docs/hun/xstadat/xstadat_evkozi/e_qvd017c.html, 2018, Downloaded 02.09.2018

[7] Kozma, T., Pónusz, M. (2016): Tite in engish: *Supply Chain Management Theory and Practice* - Basics – Ellátásilánc-menedzsment elmélete és gyakorlata – alapok, Gyöngyös, Károly Róbert Kutató – Oktató Közhasznú Nonprofit Kft.

[8] Gy. Czifra,P. Szabó, M. MĺkvaJ. Vaňová.:Lean Principles Application in the Automotive Industry, Acta Polytechnica Hungarica, Vol. 16, No. 5, 2019, https://www.uni-obuda.hu/journal/Czifra_Szabo_Mlkva_Vanova_92.pdf

[9] Womack, J. P., Jones, D. T.: Lean Thinking Banish Waste and Create Wealth in Your Corporation. Simon & Shuster, New York, 1996, p. 400

[10] Gyenge, B., Szilágyi, H., Kozma, T.: Tite in engish: *Usage of Lean management aproach in case of a service company* - Lean menedzsment alkalmazása szolgáltatóvállalat esetében. Budapest, Vezetéstudomány, (4), 2015, p. 46

[11] Blanchard, D.: Supply Chain Management – Best practices, Second edition, Hoboken, New Jersey, John Wiley & Sons, Inc., 2010, p. 215

[12] Demeter, K.: Termelés, szolgáltatás, logisztika. Tite in engish:*Process of Value Creation* - Az értékteremtés folyamatai: Budapest, Wolters Kluwer Publishing, 2014

[13] Grigoryev, I.: AnyLogic7 in Three Days. Createspace Independent Pub., 2016

[14] Gábor, Cs.: Placing Event-Action-based Visual Programming in the Process of Computer Science Education, Journal: Acta Polytechnica Hungarica, Vol. 16, No. 2, 2019, http://uni-obuda.hu/journal/

[15] Borshchev, A.: The Big Book of Simulation Modeling. Multimethod modelingwith AnyLogic 6. AnyLogic North America, 2016

On an Alternative to Four Notable Distribution Functions with Applications in Engineering and the Business Sciences

József Dombi[1], **Tamás Jónás**[2]

[1]Department of Computer Algorithms and Artificial Intelligence
University of Szeged
Árpád tér 2, H-6720 Szeged, Hungary
dombi@inf.u-szeged.hu

[2]Institute of Business Economics
Eötvös Loránd University
Egyetem tér 1-3, H-1053 Budapest, Hungary
jonas@gti.elte.hu

Abstract: A new parametric probability distribution function is introduced and its connections with some well-known distribution functions are discussed. Due to its flexibility, we call the novel distribution function the pliant distribution function. We show that the asymptotic pliant probability distribution function can coincide with the Weibull-, exponential and logistic probability distribution functions. Furthermore, we demonstrate that with appropriate parameter settings, the novel distribution gives a simple and accurate approximation to the standard normal probability distribution. Next, we show that the pliant probability distribution function, as an alternative to the Weibull- and exponential distribution functions, can be used to model constant, monotonic and bathtub-shaped hazard functions in reliability theory. We also point out that a function transformed from the new probability distribution function can be applied in the so-called kappa regression analysis method, which may be viewed as an alternative to logistic regression.

Keywords: pliant distribution; Weibull distribution; logistic distribution; normal distribution; approximations

1 Introduction

In this article, a new four-parameter probability distribution function is introduced and some of its applications are discussed. As the novel distribution function is so flexible that it may be viewed as an alternative to familiar distribution functions, we call it the pliant distribution function. The cumulative distribution function of

the novel probability distribution is based on the so-called omega function. We will demonstrate that both the omega and the exponential function $f(x) = \exp(\alpha x^{\beta})$ $(x, \alpha, \beta \in \mathbb{R}, \beta > 0)$ may be deduced from a common differential equation that we call the generalized exponential differential equation. Furthermore, we will show that the omega function, which has the α, β and d parameters $(\alpha, \beta, d \in \mathbb{R}, \beta, d > 0)$, is asymptotically identical with the exponential function $f(x) = \exp(\alpha x^{\beta})$. Exploiting this result, we will discuss in detail how the pliant probability distribution function can be utilized to approximate some remarkable probability distribution functions, whose formulas include exponential terms. Namely, we will show that the asymptotic pliant distribution function can coincide with the Weibull-, exponential and logistic probability distribution functions. Moreover, we will demonstrate that with appropriate parameter settings, the pliant probability distribution function can approximate the standard normal probability distribution function quite well, while the approximating formula is very simple and contains only one parameter. Here, we will also highlight the interesting fact that this approximation formula may be viewed as a special case of a modifier operator in the continuous-valued logic. The flexibility of the novel probability distribution function lays the foundations for its applications in different areas of science and in a wide range of modeling problems. We will point out that the pliant probability distribution function, as an alternative to the Weibull- and exponential distribution functions, can be used to model constant, monotonic and bathtub-shaped hazard functions in reliability theory. We will also demonstrate that a function transformed from the new probability distribution function can be utilized in the so-called kappa regression analysis method, which may be viewed as an alternative to logistic regression.

The remaining part of the paper is organized as follows. In Section 2, the omega function and the pliant probability distribution function are introduced and some important connections between the omega and the exponential functions are discussed. In Section 3, we demonstrate how the new probability distribution function can be utilized to approximate the Weibull, exponential, logistic and standard normal distribution functions and highlight some practical applications of our theoretical findings. Lastly, we will draw some key conclusions about the new probability distribution and make some suggestions for future research.

2 The Pliant Probability Distribution Function

Now, we will introduce a new four-parameter probability distribution function which we call the pliant probability distribution function. This novel function has the parameters α, β, γ and d, where $\alpha > 0$, $d > 0$, $\gamma \in \{-1, 1\}$ and $\beta \in \mathbf{B}_{\gamma}$. First of all, we will define the domain $\mathbf{B}_{\gamma}$ of parameter β.

Definition 1. *The set* $\mathbf{B}_{\gamma}$ *is given by*

$$\mathbf{B}_{\gamma} = \{b^{\frac{1}{2}(\gamma+1)} : b \in \mathbb{R}^{+}, \gamma \in \{-1, 1\}\}. \tag{1}$$

Notice that if $\gamma = 1$, then $\mathbf{B}_{\gamma} = \mathbb{R}^{+}$, and if $\gamma = -1$, then $\mathbf{B}_{\gamma} = \{1\}$. From here on, a probability distribution function is always a cumulative distribution function (CDF).

The pliant probability distribution function is founded on an auxiliary function that we call the omega function, the appropriate linear transformation of which is the generator function of certain unary operators in continuous-valued logic [7]. Firstly, we will introduce the omega function.

Definition 2. *The omega function* $\omega_d^{(\alpha,\beta)}(x)$ *is given by*

$$\omega_d^{(\alpha,\beta)}(x) = \left(\frac{d^\beta + x^\beta}{d^\beta - x^\beta}\right)^{\frac{\alpha d^\beta}{2}}, \tag{2}$$

where $\alpha, d \in \mathbb{R}$, $d > 0$, $\beta \in \mathbf{B}_\gamma$, $x \in \left(\frac{d}{2}(\gamma-1), d\right)$, $\gamma \in \{-1, 1\}$.

Later, we will explain why this formula is so useful. Making use of this function, we can define the pliant probability distribution function.

Definition 3. *The pliant probability distribution function* $F_P(x;\alpha,\beta,\gamma,d)$ *is given by*

$$F_P(x;\alpha,\beta,\gamma,d) = \begin{cases} 0, & \text{if } x \le \frac{d}{2}(\gamma-1) \\ \left(1 - \gamma\omega_d^{(-\alpha,\beta)}(x)\right)^\gamma, & \text{if } x \in \left(\frac{d}{2}(\gamma-1), d\right) \\ 1, & \text{if } x \ge d, \end{cases} \tag{3}$$

where $\alpha, d \in \mathbb{R}$, $\alpha > 0$, $d > 0$, $\beta \in \mathbf{B}_\gamma$, $\gamma \in \{-1, 1\}$.

In order to demonstrate that the function $F_P(x;\alpha,\beta,\gamma,d)$ is in fact a probability distribution function of a continuous random variable, we will describe the main properties of the omega function.

2.1 Main Properties of the Omega Function

Here, we state the most important properties of the omega function, namely domain, differentiability, monotonity, limits and convexity. Note that we demonstrated these properties in [11].

Domain We will utilize the omega function either with the domain $x \in (0, d)$, or with the domain $x \in (-d, d)$. Note that the domain $\mathbf{B}_\gamma$ of parameter β is connected with the domain of x; that is, since $x \in \left(\frac{d}{2}(\gamma-1), d\right)$, $\beta \in \mathbf{B}_\gamma$ and $\gamma \in \{-1, 1\}$, one of the following two cases holds:

- if $\gamma = 1$, then $x \in (0, d)$ and $\beta > 0$
- if $\gamma = -1$, then $x \in (-d, d)$ and $\beta = 1$.

Notice that we allow the parameter β to have the value of 1 when $x \in (0, d)$, but β just has the value of 1 when $x \in (-d, d)$.

Differentiability $\omega_d^{(\alpha,\beta)}(x)$ is differentiable in the interval $\left(\frac{d}{2}(\gamma-1), d\right)$.

Monotonicity If $\alpha > 0$, then $\omega_d^{(\alpha,\beta)}(x)$ is strictly monotonously increasing; if $\alpha < 0$, then $\omega_d^{(\alpha,\beta)}(x)$ is strictly monotonously decreasing; if $\alpha = 0$, then $\omega_d^{(\alpha,\beta)}(x)$ has a constant value of 1 in the interval $\left(\frac{d}{2}(\gamma-1), d\right)$.

Limits

- If $x \in (0,d)$ and $\beta > 0$, then

$$\lim_{x\to 0^+} \omega_d^{(\alpha,\beta)}(x) = 1, \tag{4}$$

$$\lim_{x\to d^-} \omega_d^{(\alpha,\beta)}(x) = \begin{cases} \infty, & \text{if } \alpha > 0 \\ 0, & \text{if } \alpha < 0. \end{cases} \tag{5}$$

- If $x \in (-d,d)$ and $\beta = 1$, then

$$\lim_{x\to -d^+} \omega_d^{(\alpha,\beta)}(x) = \begin{cases} 0, & \text{if } \alpha > 0 \\ \infty, & \text{if } \alpha < 0, \end{cases} \tag{6}$$

$$\lim_{x\to d^-} \omega_d^{(\alpha,\beta)}(x) = \begin{cases} \infty, & \text{if } \alpha > 0 \\ 0, & \text{if } \alpha < 0. \end{cases} \tag{7}$$

Convexity It can be shown that if $x \in (0,d)$ and $\beta > 0$, then the convexity of the function $\omega_d^{(\alpha,\beta)}(x)$ in the interval $(0,d)$ is as follows:

- If $d^{2\beta} < \frac{4(\beta^2-1)}{\alpha^2\beta^2}, \alpha \neq 0$, then $\omega_d^{(\alpha,\beta)}(x)$ is convex when $\alpha > 0$ and $\omega_d^{(\alpha,\beta)}(x)$ is concave when $\alpha < 0$.
- If $d^{2\beta} \geq \frac{4(\beta^2-1)}{\alpha^2\beta^2}, \alpha \neq 0$, then we can distinguish the following cases:
  - if $\alpha > 0$ and $0 < \beta < 1$, then $\omega_d^{(\alpha,\beta)}(x)$ changes its shape from concave to convex at x_r
  - if $\alpha > 0$ and $\beta \geq 1$, then $\omega_d^{(\alpha,\beta)}(x)$ is convex
  - if $\alpha < 0$, $0 < \beta \leq 1$ and $x_r < d$, then $\omega_d^{(\alpha,\beta)}(x)$ changes its shape from convex to concave at x_r
  - if $\alpha < 0$, $0 < \beta \leq 1$ and $x_r \geq d$, then $\omega_d^{(\alpha,\beta)}(x)$ is convex
  - if $\alpha < 0$, $\beta > 1$ and $x_r < d$, then $\omega_d^{(\alpha,\beta)}(x)$ changes its shape from concave to convex at x_l and from convex to concave at x_r
  - if $\alpha < 0$, $\beta > 1$ and $x_r \geq d$, then $\omega_d^{(\alpha,\beta)}(x)$ changes its shape from concave to convex at x_l,

where

$$x_l = \left(\frac{-\alpha\beta d^{2\beta} - \sqrt{\alpha^2\beta^2 d^{4\beta} - 4(\beta^2-1)d^{2\beta}}}{2(\beta+1)} \right)^{1/\beta} \tag{8}$$

$$x_r = \left(\frac{-\alpha\beta d^{2\beta} + \sqrt{\alpha^2\beta^2 d^{4\beta} - 4(\beta^2-1)d^{2\beta}}}{2(\beta+1)} \right)^{1/\beta}. \tag{9}$$

It can be also shown that if $x \in (-d,d)$ and $\beta = 1$, then the shape of the function $\omega_d^{(\alpha,\beta)}(x)$ in the interval $(-d,d)$ is as follows:

- if $|\alpha|d \geq 2$, then $\omega_d^{(\alpha,\beta)}(x)$ is convex
- if $\alpha > 0$ and $\alpha d < 2$, then $\omega_d^{(\alpha,\beta)}(x)$ changes from concave to convex at $-\alpha d^2/2$
- if $\alpha < 0$ and $\alpha d > -2$, then $\omega_d^{(\alpha,\beta)}(x)$ changes from convex to concave at $-\alpha d^2/2$.

Based on the above-mentioned properties of the omega function, the basic semantics of the parameters α, β and d can be summarized as follows. The parameter d determines the domain of the omega function (either $x \in (-d,d)$, or $x \in (0,d)$), the parameter α influences its monotonicity and steepness, while the parameter β also affects the steepness of $\omega_d^{(\alpha,\beta)}(x)$, ($\alpha, d \in \mathbb{R}$, $d > 0$, $\beta > 0$). Figure 1 shows plots of some omega functions.

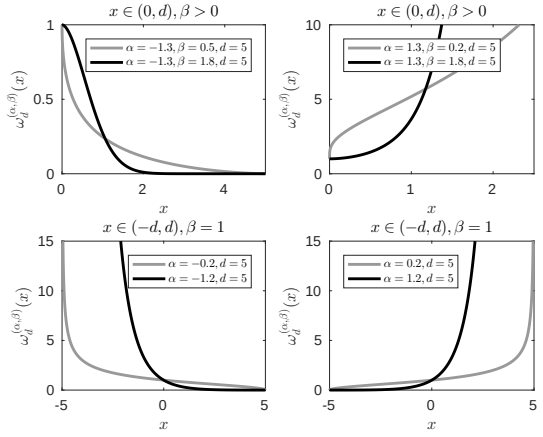


Figure 1
Plots of some omega functions

2.1.1 The Generalized Exponential Differential Equation

Next, we will introduce the generalized exponential differential equation and show how it is related to the exponential function $f(x) = \exp(\alpha x^{\beta})$, $(\alpha, \beta \in \mathbb{R}, \beta > 0)$ and to the omega function.

Definition 4. *We define the generalized exponential differential equation as*

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \alpha\beta x^{\beta-1}\left(\frac{d^{2\beta}}{d^{2\beta}-x^{2\beta}}\right)^{\varepsilon} f(x), \tag{10}$$

where $\alpha, d \in \mathbb{R}$, $d > 0$, $\beta \in \mathbf{B}_{\gamma}$, $x \in \left(\frac{d}{2}(\gamma-1), d\right)$, $\gamma \in \{-1, 1\}$, $\varepsilon \in \{0, 1\}$, $f(x) > 0$.

Lemma 1 (Lemma 2 in [11])**.** *The solutions of the generalized exponential differential equation are:*

$$f(x) = \begin{cases} C\exp(\alpha x^{\beta}), & \text{if } \varepsilon = 0 \\ C\left(\frac{d^{\beta}+x^{\beta}}{d^{\beta}-x^{\beta}}\right)^{\frac{ad^{\beta}}{2}}, & \text{if } \varepsilon = 1, \end{cases} \tag{11}$$

where $C \in \mathbb{R}$ *and* $C > 0$.

Proof. See the proof of Lemma 2 in [11]. □

2.1.2 Connections Between the Exponential and Omega Functions

Lemma 1 suggests that there is a key connection between the exponential function $f(x) = \exp(\alpha x^{\beta})$ and the omega function. Namely, the solution of the generalized exponential differential equation with $\varepsilon = 0$ and $C = 1$ is simply the exponential function $f(x) = \exp(\alpha x^{\beta})$, while the solution of (10) with $\varepsilon = 1, C = 1$ is the omega function. Furthermore, if d is much greater than x, then

$$\frac{d^{2\beta}}{d^{2\beta}-x^{2\beta}} \approx 1. \tag{12}$$

In this case the generalized exponential differential equation for $\varepsilon = 1$ becomes the following approximate equation:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} \approx \alpha\beta x^{\beta-1} f(x), \tag{13}$$

which is nearly the generalized exponential differential equation with $\varepsilon = 0$, the solution of which is the exponential function $f(x) = \exp(\alpha x^{\beta})$. The following proposition provides the theoretical basis for this result (see Theorem 1 in [11]).

Proposition 1 (Theorem 1 in [11])**.** *For any* $x \in \left(\frac{d}{2}(\gamma-1), d\right)$,

$$\lim_{d\to\infty} \omega_d^{(\alpha,\beta)}(x) = \exp(\alpha x^{\beta}), \tag{14}$$

where $\alpha, d \in \mathbb{R}$, $d > 0$, $\beta \in \mathbf{B}_{\gamma}$, $\gamma \in \{-1, 1\}$.

Proof. See the proof of Theorem 1 in [11]. □

Based on Proposition 1, we can state that the asymptotic omega function is just the exponential function $f(x) = \exp(\alpha x^{\beta})$. Actually, if $x \ll d$, then $\omega_d^{(\alpha,\beta)}(x) \approx \exp(\alpha x^{\beta})$; that is, if d is sufficiently large, then the omega function suitably approximates the exponential function $f(x) = \exp(\alpha x^{\beta})$.

3 Approximations and their Applications

Here, we will demonstrate how the pliant probability distribution function $F_P(x;\alpha,\beta,\gamma,d)$ can be utilized for approximating the Weibull, exponential, logistic and standard normal probability distribution functions. First, we will show that the function $F_P(x;\alpha,\beta,\gamma,d)$ is in fact a probability distribution function.

Lemma 2. *The function $F_P(x;\alpha,\beta,\gamma,d)$ given by Definition 3 is a probability distribution function.*

Proof. The function $F_P(x;\alpha,\beta,\gamma,d)$ is a probability distribution function if it is:

1. monotonously increasing
2. left continuous,
3. and $\lim_{x\to-\infty} F_P(x;\alpha,\beta,\gamma,d) = 0$ and $\lim_{x\to+\infty} F_P(x;\alpha,\beta,\gamma,d) = 1$.

These properties of the function $F_P(x;\alpha,\beta,\gamma,d)$ readily follow from properties of the omega function $\omega_d^{(\alpha,\beta)}(x)$. □

We have demonstrated that the omega function in the asymptotic limit is just the exponential function $f(x) = \exp(\alpha x^{\beta})$, where $\beta > 0$. Exploiting this result and the fact the function $F_P(x;\alpha,\beta,\gamma,d)$ is a probability distribution function, we will show how the pliant probability distribution function can be used to approximate some well-known probability distribution functions that include exponential terms. From now on, we will use the notation $\xi \sim D_p(\alpha,\beta,\gamma,d)$ to indicate that the random variable ξ has a pliant probability distribution with the parameters α,β,γ,d. That is, if $\xi \sim D_p(\alpha,\beta,\gamma,d)$, then $P(\xi < x) = F_P(x;\alpha,\beta,\gamma,d)$ for any $x \in \mathbb{R}$.

3.1 Approximation to the Weibull probability Distribution Function

The 2-parameter distribution function $F_W(x;\beta,\lambda)$ of the random variable which has a Weibull probability distribution is usually given by the formula:

$$F_W(x;\beta,\lambda) = \begin{cases} 0, & \text{if } x \le 0 \\ 1-\exp(-(x/\lambda)^{\beta}) & \text{if } x > 0, \end{cases} \tag{15}$$

where $\beta, \lambda \in \mathbb{R}$ and $\beta, \lambda > 0$ are the shape and scale parameters of the distribution, respectively [24, 33]. Making the substitution $\alpha = \lambda^{-\beta}$, (15) may be written in the form

$$F_W(x;\alpha,\beta) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1 - \exp(-\alpha x^{\beta}), & \text{if } x > 0, \end{cases} \tag{16}$$

where $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta > 0$. Hereafter, we will use this alternative definition of the 2-parameter distribution function of the random variable that has a Weibull probability distribution. Furthermore, we will use the notation $\eta \sim W(\alpha,\beta)$ to indicate that the random variable η has a Weibull probability distribution with the parameters $\alpha, \beta > 0$; that is, if $\eta \sim W(\alpha,\beta)$, then $P(\eta < x) = F_W(x;\alpha,\beta)$ for any $x \in \mathbb{R}$.

Proposition 2. *If $\xi \sim D_p(\alpha,\beta,\gamma,d)$, $\eta \sim W(\alpha,\beta)$ and $\gamma = 1$, then for any $x \in \mathbb{R}$,*

$$\lim_{d\to\infty} P(\xi < x) = P(\eta < x), \tag{17}$$

where $\alpha, d \in \mathbb{R}$, $\alpha, d > 0$, $\beta \in \mathbf{B}_\gamma$.

Proof. Utilizing the definitions of $F_P(x;\alpha,\beta,\gamma,d)$ and $\mathbf{B}_\gamma$, if $\gamma = 1$, then the pliant probability distribution function $F_P(x;\alpha,\beta,\gamma,d)$ may be written as

$$F_P(x;\alpha,\beta,\gamma,d) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1 - \omega_d^{(-\alpha,\beta)}(x), & \text{if } x \in (0,d) \\ 1, & \text{if } x \geq d. \end{cases} \tag{18}$$

Let $x \in \mathbb{R}$ be fixed. We will now distinguish the following two cases.

1. If $x \leq 0$, then $F_P(x;\alpha,\beta,\gamma,d) = F_W(x;\beta,\lambda)$ holds by definition.
2. If $x \in (0,d)$, $d > 0$, then

   $$F_P(x;\alpha,\beta,\gamma,d) = 1 - \omega_d^{(-\alpha,\beta)}(x). \tag{19}$$

   Next, following Proposition 1, if $d \to \infty$, then

   $$F_P(x;\alpha,\beta,\gamma,d) = 1 - \omega_d^{(-\alpha,\beta)}(x) \to 1 - \exp(-\alpha x^{\beta}) = F_W(x;\alpha,\beta). \tag{20}$$

   That is,

   $$\lim_{d\to\infty} P(\xi < x) = P(\eta < x). \tag{21}$$

□

Some example plots of Weibull probability distribution functions and their approximations by pliant probability distribution functions are shown in Figure 2. In each subplot of Figure 2, the left hand side scale is associated with functions $F_W(x;\alpha,\beta)$ and $F_P(x;\alpha,\beta,\gamma,d)$, while the right hand side scale is related to the difference function $F_W(x;\alpha,\beta) - F_P(x;\alpha,\beta,\gamma,d)$. We can see that, in line with Proposition 2, the goodness of approximation improves with the increasing value of the parameter d.

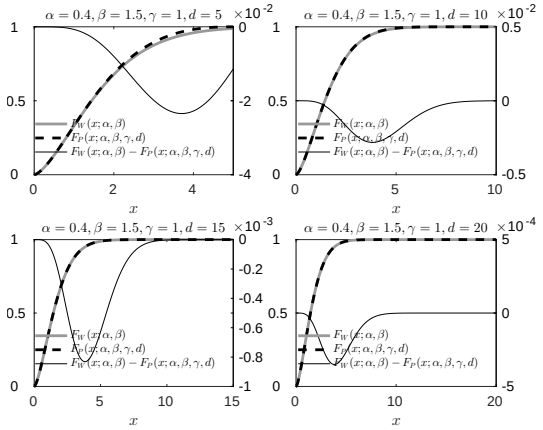


Figure 2
Examples of Weibull- and pliant probability distribution function plots

3.1.1 Applications in Reliability Theory

Let the continuous random variable τ be the time-to-first-failure of a component or system. In reliability theory, the failure rate function $h(t)$ for τ is given by

$$h(t) = \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t R(t)} = \frac{f(t)}{R(t)}, \tag{22}$$

where $f(t)$ is the probability density function of τ. The hazard function $h(t)$ is also called the failure rate function. In practice, the quantity $h(t)\Delta t$ represents the conditional probability that a component or a system will fail in the time interval $(t, t+\Delta t]$, given that it has survived up to time t, $(t, \Delta t > 0)$.

A typical hazard function curve of a component or a system is bathtub-shaped; that is, it can be divided into three distinct phases called the infant mortality period, useful life, and wear-out period. It is typical that the probability distribution of τ is different in the three characteristic phases of the bathtub-shaped hazard function. If τ has a Weibull probability distribution with the parameters $\alpha, \beta > 0$, then using (22), the hazard function $h_W(t;\alpha,\beta)$ of τ (which we call the Weibull hazard function) is

$$h_W(t;\alpha,\beta) = \frac{f_W(t;\alpha,\beta)}{1 - F_W(t;\alpha,\beta)} = \frac{\alpha\beta t^{\beta-1}\exp(-\alpha t^{\beta})}{\exp(-\alpha t^{\beta})} = \alpha\beta t^{\beta-1}, \tag{23}$$

where $f_W(t;\alpha,\beta)$ is the probability density function of τ. Equation (23) suggests an important property of the hazard function $h(t;\alpha,\beta)$. Namely,

- if $0 < \beta < 1$, then $h_W(t;\alpha,\beta)$ is decreasing with respect to time
- if $\beta = 1$, then $h_W(t;\alpha,\beta)$ is constant and has a value of α
- if $\beta > 1$, then $h_W(t;\alpha,\beta)$ is increasing with respect to time.

That is, if $\tau \sim W(\alpha,\beta)$, then the failure rate function $h_W(t;\alpha,\beta)$, with appropriate values of its parameters, can characterize each of the three phases of a bathtub-shaped failure rate curve quite well. On the one hand, this flexibility of the Weibull probability distribution makes it suitable for modeling the probability distribution of time-to-first-failure random variable in a wide range of reliability analyses [23]. On the other hand, the Weibull hazard function $h_W(t;\alpha,\beta)$ is either monotonic or constant; that is, its curve cannot be bathtub-shaped. However, lifetime data of a component or system typically require non-monotonic shapes like the bathtub shape. Many modifications have been suggested for the Weibull probability distribution in order to have non-monotonic shapes. See, for example, the publications [17], [25], [22], [34], [37], [14], [13], [18], [27], [28], [4], [6]. A comprehensive review of the known modifications to the Weibull probability distribution can be found in a quite recent article [1]. Here, we will demonstrate that the hazard function of the pliant probability distribution can be used to model both monotonic and bathtub-shaped hazard rate curves.

Now let us assume that τ has a pliant probability distribution with the parameters $\alpha,d>0,\gamma=1$. In this case, $\beta>0$ and the hazard function of τ, which we will call the pliant hazard function, is

$$h_P(t;\alpha,\beta,d)=\frac{f_P(t;\alpha,\beta,d)}{1-F_P(t;\alpha,\beta,d)}=\frac{\alpha\beta t^{\beta-1}\frac{d^{2\beta}}{d^{2\beta}-t^{2\beta}}\omega_d^{(-\alpha,\beta)}(t)}{\omega_d^{(-\alpha,\beta)}(t)}= \\ =\alpha\beta t^{\beta-1}\frac{d^{2\beta}}{d^{2\beta}-t^{2\beta}}, \tag{24}$$

if $0<t<d$, where $f_P(t;\alpha,\beta,d)$ is the probability density function of τ. Utilizing (23) and (24), the pliant hazard function $h_P(t;\alpha,\beta,d)$ may be written as

$$h_P(t;\alpha,\beta,d)=h_W(\alpha,\beta)g(t;\beta,d), \tag{25}$$

where

$$g(t;\beta,d)=\frac{d^{2\beta}}{d^{2\beta}-t^{2\beta}}, \tag{26}$$

and $\alpha,\beta,d>0$, $t\in(0,d)$. That is, the pliant hazard function may be interpreted as the Weibull hazard function multiplied by the corrector function $g(t;\beta,d)$. The pliant hazard function $h_P(t;\alpha,\beta,d)$ has some key properties that make it suitable for modeling bathtub-shaped failure rate curves. The following lemma allows us to utilize the pliant hazard function as an alternative to the Weibull hazard function.

Lemma 3. *For any $t\in(0,d)$, if $d\to\infty$, then $h_P(t;\alpha,\beta,d)\to h_W(t;\alpha,\beta)$, where $\alpha,\beta,d>0$.*

Proof. If $t\in(0,d)$ is fixed and $d\to\infty$, then $g(t;\beta,d)\to 1$ and so

$$h_P(t;\alpha,\beta,d)=h_W(\alpha,\beta)g(t;\beta,d)\to h_W(\alpha,\beta). \tag{27}$$

□

The practical implication of this result is as follows. Since the Weibull hazard function can be utilized as a model for each phase of a bathtub-shaped failure rate curve and $h_P(t;\alpha,\beta,d) \approx h_W(\alpha,\beta)$, if t is small compared to d, the pliant hazard function can also model each phase of the same bathtub-shaped failure rate curve, if d is sufficiently large. That is, the pliant hazard function, as an alternative to the Weibull hazard function, can be utilized as a phase-by-phase model of a bathtub-shaped failure rate curve.

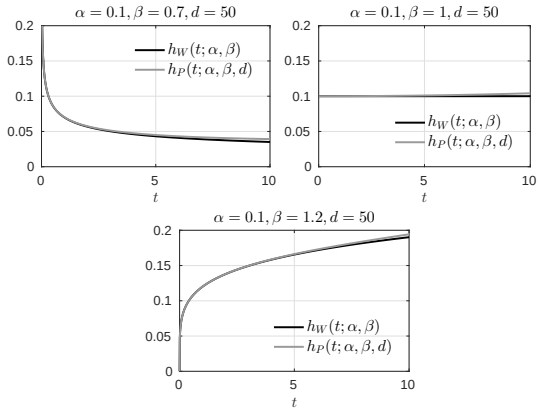


Figure 3
Plots of Weibull- and pliant hazard functions (Figure 5 in [11])

Figure 3 shows how the Weibull- and pliant hazard function curves can model each characteristic phase of a failure rate curve. The plots in Figure 3 demonstrate the results of the previous lemma; that is, if $t \ll d$, then the pliant hazard function approximates the Weibull hazard function quite well.

The next lemma shows how that the pliant hazard function $h_P(t;\alpha,\beta,d)$ can be utilized as a model for all the three phases of a bathtub-shaped failure rate curve.

Lemma 4 (Lemma 5 in [11]). *If $0 < \beta < 1$, then $h_P(t;\alpha,\beta,d)$ is strictly convex in the interval $(0,d)$ and $h_P(t;\alpha,\beta,d)$ has its minimum at*

$$t_0 = d\left(\frac{1-\beta}{1+\beta}\right)^{\frac{1}{2\beta}}. \tag{28}$$

Proof. The lemma follows from the elementary properties of the pliant hazard function $h_P(t;\alpha,\beta,d)$ by using its first and second derivatives. □

Figure 4 shows some plots of the pliant hazard function $h_P(t;\alpha,\beta,d)$ with $0 < \beta < 1$.

Based on the above properties of the pliant hazard function, we may conclude that the pliant probability distribution with $\gamma = 1$ can be employed to describe the probability distribution of the time-to-first-failure random variable in each characteristic phase of a bathtub-shaped failure rate curve. Moreover, we have two possibilities for

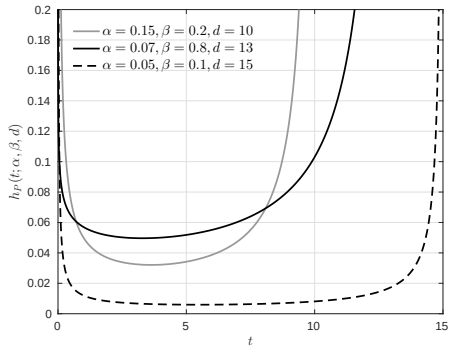


Figure 4
Examples of bathtub-shaped pliant hazard function plots (Figure 6 in [11])

modeling a bathtub-shaped failure rate curve. Namely, either we piecewise describe each phase by a pliant hazard function, or we apply one pliant hazard function that models the entire failure rate curve. It is worth noting here that the Weibull distribution is widely applied to model reliability of software systems (see, e.g. [19, 29]), therefore, our pliant probability distribution, as an alternative to the Weibull probability distribution, can also be applied in this area.

3.2 Approximation to the Exponential probability Distribution Function

If the random variable η has an exponential probability distribution with the parameter $\alpha > 0$, then the probability distribution function $F_{exp}(x;\alpha)$ of η is given by

$$F_{exp}(x;\alpha) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1 - \exp(-\alpha x), & \text{if } x > 0 \end{cases} \tag{29}$$

[24]. We will use the notation $\eta \sim \exp(\alpha)$ to indicate that the random variable η has an exponential probability distribution with the parameter $\alpha > 0$. The probability distribution function $F_{exp}(x;\alpha)$ is a special case of the Weibull probability distribution function $F_W(x;\alpha,\beta)$. Namely, if $\beta = 1$, then $F_{exp}(x;\alpha) = F_W(x;\alpha,\beta)$. Based on this, we can state the following proposition.

Proposition 3. *If $\xi \sim D_p(\alpha,\beta,\gamma,d)$, $\eta \sim \exp(\alpha)$, $\gamma = 1$ and $\beta = 1$, then for any $x \in \mathbb{R}$,*

$$\lim_{d\to\infty} P(\xi < x) = P(\eta < x), \tag{30}$$

where $\alpha, d \in \mathbb{R}$, $\alpha, d > 0$.

Proof. The proposition follows from Proposition 2. □

This result tells us that if $\gamma = 1$, $\beta = 1$, then the asymptotic pliant probability distribution function, for the parameter d, is just the exponential probability distribution function. Note that the pliant probability distribution is a generalization of the epsilon probability distribution that we introduced in [10].

3.2.1 Applications in Reliability Theory

The exponential probability distribution plays a significant role in the theory and practice of reliability management [36, 32]. This distribution also appears frequently in lifetime and reaction time studies. Here, we will discuss how the pliant probability distribution, as an alternative to the exponential distribution, can be applied to model failure rates.

Since the exponential probability distribution may be viewed as a special case of the Weibull probability distribution, namely when $\beta = 1$, utilizing the Weibull hazard function in (23) with $\beta = 1$ gives us the hazard function $h_{exp}(t;\alpha)$ of the exponential distribution with the parameter $\alpha > 0$:

$$h_{exp}(t;\alpha) = h_W(t;\alpha,\beta)_{|\beta=1} = \alpha\beta t^{\beta-1}{}_{|\beta=1} = \alpha. \tag{31}$$

That is, if $\tau \sim \exp(\alpha)$, then the hazard function $h_{exp}(t;\alpha)$ of τ is constant with the value α. Based on Lemma 3, the pliant hazard function tends to the Weibull hazard function, if $d \to \infty$. Applying this results for the special case where $\beta = 1$, we get the following reduced pliant hazard function

$$h_P(t;\alpha) = h_P(t;\alpha,\beta)_{|\beta=1} = \alpha\beta t^{\beta-1}\frac{d^{2\beta}}{d^{2\beta}-t^{2\beta}}\Big|_{\beta=1} = \alpha\frac{d^2}{d^2-t^2}, \tag{32}$$

where $t \in (0,d)$. There are two key properties of the pliant hazard function $h_P(t;\alpha)$ that should be mentioned here. These are:

1. If $t \in (0,d)$ is fixed and $d \to \infty$, then $h_P(t;\alpha) \to \alpha$.
2. $h_P(t;\alpha)$ is monotonously increasing in the interval $(0,d)$.

The practical implications of the above properties of the hazard function $h_P(t;\alpha)$ are as follows. If d is sufficiently large compared to t, then $h_P(t;\alpha)$ is approximately constant with the value α; that is, function $h_P(t;\alpha)$ can be utilized to model the quasi constant second phase of a bathtub-shaped failure rate curve. As $h_P(t;\alpha)$ is monotonously increasing in the interval $(0,d)$, it can also be treated as a model of the increasing third phase of a bathtub-shaped hazard curve (see also [10]).

3.3 Approximation to the Logistic Probability Distribution Function

The 2-parameter probability distribution function $F_L(x;\mu,s)$ of the random variable that has a logistic probability distribution is commonly given by

$$F_L(x;\mu,s) = \frac{1}{1+\exp(-\frac{x-\mu}{s})}, \tag{33}$$

where $\mu, s \in \mathbb{R}$ and $s > 0$ are the location and scale parameters of the distribution, respectively [16, 3].

By applying the $\alpha = 1/s$ substitution and setting the location parameter μ to zero, (33) may be written as

$$F_L(x;\alpha) = \frac{1}{1+\exp(-\alpha x)} \tag{34}$$

where $\alpha \in \mathbb{R}$, $\alpha > 0$. From now on, we will utilize the logistic probability distribution function in the latter form and use the notation $\eta \sim L(\alpha)$ to indicate that the random variable η has a logistic distribution with the parameter $\alpha > 0$; that is, $P(\eta < x) = F_L(x;\alpha)$. The following proposition will demonstrate that the pliant probability distribution function can be used to approximate the logistic distribution function $F_L(x;\alpha)$.

Proposition 4. *If $\xi \sim D_p(\alpha,\beta,\gamma,d)$, $\eta \sim L(\alpha)$, $\gamma = -1$, then for any $x \in \mathbb{R}$,*

$$\lim_{d\to\infty} P(\xi < x) = P(\eta < x), \tag{35}$$

where $\alpha, d \in \mathbb{R}$, $\alpha, d > 0$.

Proof. Since $\gamma = -1$ and $\beta = 1$, the pliant probability distribution function $F_P(x;\alpha,\beta,\gamma,d)$ may be written as

$$F_P(x;\alpha,\beta,\gamma,d) = \begin{cases} 0, & \text{if } x \le 0 \\ \frac{1}{1+\omega_d^{(-\alpha,\beta)}(x)}, & \text{if } x \in (-d,d) \\ 1, & \text{if } x \ge d. \end{cases} \tag{36}$$

Let $x \in \mathbb{R}$ be fixed. Now, utilizing the fact that $\beta = 1$ and exploiting Proposition 1 gives

$$F_P(x;\alpha,\beta,\gamma,d) = \frac{1}{1+\omega_d^{(-\alpha,\beta)}(x)} \xrightarrow{d\to\infty} \frac{1}{1+\exp(-\alpha x)} = F_L(x;\alpha). \tag{37}$$

This result means that

$$\lim_{d\to\infty} P(\xi < x) = P(\eta < x). \tag{38}$$

□

Based on Proposition 4, we can state that the pliant probability distribution function $F_P(x;\alpha,\beta,\gamma,d)$ can be used to approximate the logistic distribution function $F_L(x;\alpha)$. A few sample plots of logistic probability distribution functions and their approximations by pliant probability distribution functions are shown in Figure 5. In each subplot of Figure 5, the left hand side scale is associated with functions $F_L(x;\alpha)$ and $F_P(x;\alpha,\beta,\gamma,d)$, while the right hand side scale is related to the difference function $F_L(x;\alpha) - F_P(x;\alpha,\beta,\gamma,d)$. The plots show how the goodness of approximation improves with increasing values of the parameter d.

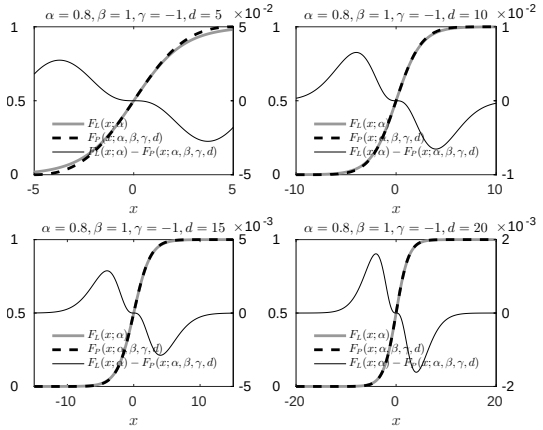


Figure 5
Examples of logistic- and pliant probability distribution function plots

3.3.1 The Kappa Regression Function

The logistic function as a regression function has a wide range of applications in many fields, including economics, business, biology, the medical sciences and engineering. Here, we will demonstrate that a function which can be derived from the pliant probability distribution function may be viewed as an alternative to the logistic regression function. If $\gamma = -1$, then $\beta = 1$ and the formula of the pliant probability distribution function for $x \in (-d, d)$ gives the following $\kappa_d^{(\alpha)}(x)$ function:

$$\kappa_d^{(\alpha)}(x) = \frac{1}{1 + \left(\frac{d+x}{d-x}\right)^{\frac{-\alpha d}{2}}}, \tag{39}$$

where $d > 0$ and $\alpha \in \mathbb{R}$. Notice that here we allow α to take any value. Let Y be a dichotomous random variable, and let 0 and 1 code its possible values. Here, we model the conditional probability $P(Y = k|x)$ as a function of the independent variable x; $x \in (a, b)$, $k \in \{0, 1\}$. In our model, which we call the kappa regression model [9], the odds $O(x)$ are given by

$$O(x) = \frac{P(Y = 1|x)}{1 - P(Y = 1|x)} = C' \left(\frac{x-a}{b-x}\right)^{\lambda}, \tag{40}$$

where $\lambda \in \mathbb{R}$ and $C' > 0$. If C' is written in the form $C' = \exp(-c)$, then from (40):

$$P(Y = 1|x) = \frac{1}{1 + \exp(c)\left(\frac{b-x}{x-a}\right)^{\lambda}} = \kappa(x; \lambda, a, b, c). \tag{41}$$

Notice that the kappa function $\kappa_d^{(\alpha)}(y)$ can be derived from the kappa function $\kappa(x; \lambda, a, b, c)$ by setting $c = 0$, $\lambda = \alpha d/2$ and applying the linear transformation $x = \frac{y+d}{2d}(b-a) + a$ with $d > 0$.

In the logistic regression model, the odds $O(x)$ are modeled by

$$O(x) = \frac{P(Y=1|x)}{1-P(Y=1|x)} = \exp(\beta_1 x + \beta_0), \tag{42}$$

and so

$$P(Y=1|x) = \frac{1}{1+\exp(-\beta_1 x - \beta_0)}, \tag{43}$$

where $\beta_0, \beta_1 \in \mathbb{R}$. The next lemma lays the foundations for using kappa regression as an alternative to logistic regression.

Lemma 5. *If $a < b$, $(a+b)/2 = x_0$ is constant, $\lambda = \beta_1 \frac{(x_0-a)(b-x_0)}{b-a}$ and $a \to -\infty$, $b \to +\infty$ so that $(a+b)/2 = x_0$, then for any $x \in (a,b)$,*

$$C' \left(\frac{x-a}{b-x} \right)^{\lambda} \to \exp(\beta_1 x + \beta_0), \tag{44}$$

where $\beta_0 = \ln C' - \beta_1 x_0$, $(C' > 0)$.

Proof. Here, $a \to -\infty$, $b \to +\infty$ so that $(a+b)/2 = x_0$; that is, $a \to -\infty$ and $b \to +\infty$ may be written as $a = x_0 - \Delta$, $b = x_0 + \Delta$, where $\Delta \to \infty$. Then, if the conditions of the lemma are satisfied,

$$\begin{aligned} \lim_{\substack{a\to-\infty \\ b\to+\infty}} \left(C' \left(\frac{x-a}{b-x} \right)^{\lambda} \right) &= \exp(\beta_1 x_0 + \beta_0) \lim_{\substack{a\to-\infty \\ b\to+\infty}} \left(\frac{x-a}{b-x} \right)^{\beta_1 \frac{(x_0-a)(b-x_0)}{b-a}} = \\ &= \exp(\beta_1 x_0 + \beta_0) \lim_{\Delta\to\infty} \left(\frac{x-x_0+\Delta}{\Delta-(x-x_0)} \right)^{\beta_1 \frac{\Delta}{2}}. \end{aligned} \tag{45}$$

Similar to the proof of Proposition 1, it can be shown that

$$\lim_{\Delta\to\infty} \left(\frac{x-x_0+\Delta}{\Delta-(x-x_0)} \right)^{\beta_1 \frac{\Delta}{2}} = \exp(\beta_1 (x-x_0)), \tag{46}$$

and so the chain in (45) can be continued as

$$\begin{aligned} &\exp(\beta_1 x_0 + \beta_0) \lim_{\Delta\to\infty} \left(\frac{x-x_0+\Delta}{\Delta-(x-x_0)} \right)^{\beta_1 \frac{\Delta}{2}} = \\ &= \exp(\beta_1 x_0 + \beta_0) \exp(\beta_1 (x-x_0)) = \exp(\beta_1 x + \beta_0). \end{aligned} \tag{47}$$

□

Hence, we have demonstrated that logistic regression may be regarded as asymptotic kappa regression. Based on our findings, kappa regression may be treated as a quasi logistic regression, where the explanatory variable is defined over a bounded subset of the real numbers. This property of kappa regression is advantageous in situations where the explanatory variable is defined over a bounded subset of real numbers, the empirical conditional probabilities exhibit an asymmetric trend and there are empirical conditional probabilities very close to zero or one at the terminal locations of the domain of the explanatory variable [9].

3.4 Approximation to the Standard Normal Probability Distribution Function

The probability distribution function $\Phi(x)$ of the standard normal random variable is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t. \tag{48}$$

We will use the common notation $\eta \sim N(0,1)$ to indicate that the random variable η has the standard normal probability distribution. The fact that the probability distribution function $\Phi(x)$ cannot be expressed in a closed form and the practical needs for computing its values provided the motivation for researchers and practitioners to approximate the standard normal probability distribution function. These research efforts resulted in an extremely wide range of approximations with various applications. Most of the approximations belong to the group of ad-hoc methods which typically utilize an a priori selected parametric function and apply various mathematical techniques to estimate the parameters in order minimize the approximation error. [21], [30] and [35] gave comprehensive overviews of the approximation formulas in their reviews. In general, we can say that the accuracy of approximations increases with the complexity of formulas and with the number of parameters they have.

Among the many ad-hoc approximations available, the 1-parameter logistic function, which has the same form as the logistic probability distribution function $F_L(x;\alpha)$, can also be applied to approximate the standard normal probability distribution function (see, for example, the books of [31], [3], and [15], and the paper of [5]). Here, we will utilize Tocher's approximation and the pliant probability distribution function to obtain a novel approximation (which has only one parameter and a very simple formula) to the standard normal probability distribution function. Tocher's approximation, which we denote by $\Phi_T(x)$, utilizes the logistic probability distribution function $F_L(x;\alpha)$ with $\alpha = 2\sqrt{2/\pi}$. That is,

$$\Phi_T(x) = \frac{1}{1+\exp\left(-2\sqrt{2/\pi}x\right)}. \tag{49}$$

It is worth mentioning here that the probability density function of Tocher's approximation can be derived from sigmoid fuzzy membership functions by using operators of continuous-valued logic [8]. Note that setting α to $2\sqrt{2/\pi}$ ensures that function $\Phi_T(x)$ is identical with function $\Phi(x)$ to first order at $x = 0$. Furthermore, based on Proposition 4, if $\alpha = 2\sqrt{2/\pi}$, $\gamma = -1$ and $\beta = 1$, then for any $x \in \mathbb{R}$,

$$\lim_{d\to\infty} F_P(x;\alpha,\beta,\gamma,d) = \Phi_T(x). \tag{50}$$

These results tell us that the following function, which we call the quasi logistic

probability distribution function

$$\Phi_{\kappa,d}(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ \frac{1}{1+\left(\frac{d+x}{d-x}\right)^{-\sqrt{2/\pi}d}}, & \text{if } x \in (-d,d) \\ 1, & \text{if } x \geq d \end{cases} \tag{51}$$

can be applied to approximate the standard normal probability distribution function $\Phi(x)$, where $d > 0$. Notice that $\Phi_{\kappa,d}(x) = F_P(x;\alpha,\beta,\gamma,d)$ with the parameter values $\alpha = 2\sqrt{2/\pi}$, $\gamma = -1$ and $\beta = 1$, so the function $\Phi_{\kappa,d}(x)$ is in fact a probability distribution function.

In fuzzy logic, the linguistic modifiers like "very", "more or less", "somewhat", "rather" and "quite" over fuzzy sets that have strictly monotonously increasing or decreasing membership functions, can be modeled by the following unary operator, which is also known as the kappa function [7].

Definition 5. *The kappa modifier operator (kappa function) is given by*

$$\kappa^{(\lambda)}_{\nu,\nu_0}(x) = \frac{1}{1+\frac{1-\nu_0}{\nu_0}\left(\frac{\nu}{1-\nu}\frac{1-x}{x}\right)^{\lambda}}, \tag{52}$$

where $\nu, \nu_0 \in (0,1)$, $\lambda \in \mathbb{R}$, *and x is a continuous-valued logic variable.*

Notice that if $\nu = \nu_0 = 0.5$, then the function $\Phi_{\kappa,d}(x)$ for $x \in (-d,d)$ can be derived from the kappa function $\kappa^{(\lambda)}_{\nu,\nu_0}(x)$ in (52) by setting $\lambda = \sqrt{2/\pi}d$ and applying the $x' = (x+d)/(2d)$ linear transformation ($d > 0$). The κ index in the notation $\Phi_{\kappa,d}(x)$ indicates that this function is connected with the kappa modifier operator.

It can be shown numerically that $|\Phi(x) - \Phi_{\kappa,d}(x)|$ is approximately minimal, if $d = 3.1152$. In this case, the maximum absolute approximation error is $2.15 \cdot 10^{-3}$. Considering that 3.1152 is close to π, using $d = \pi$ instead of $d = 3.1152$ does not worsen significantly the approximation accuracy. If $d = \pi$, then the maximum absolute approximation error is $2.3570 \cdot 10^{-3}$. Thus, we propose the use of function $\Phi_{\kappa,\pi}(x)$ for approximation as this function has a very simple formula and its maximum absolute approximation error is just slightly greater than that of function $\Phi_{\kappa,d}(x)$ with $d = 3.1152$ [8]:

$$\Phi_{\kappa,\pi}(x) = \Phi_{\kappa,d}(x)\big|_{d=\pi} = \begin{cases} 0, & \text{if } x \leq -\pi \\ \frac{1}{1+\left(\frac{\pi-x}{\pi+x}\right)^{\sqrt{2\pi}}}, & \text{if } x \in (-\pi,+\pi) \\ 1, & \text{if } x \geq +\pi. \end{cases} \tag{53}$$

Recall that $\Phi_{\kappa,\pi}(x) = F_P(x;\alpha,\beta,\gamma,d)$ with the parameter values $\alpha = 2\sqrt{2/\pi}$, $\beta = 1$, $\gamma = -1$ and $d = \pi$; that is, $\Phi_{\kappa,\pi}(x)$ is a special case of the pliant probability distribution function. It should be added that there are just a few known approximations with a single constant parameter in this accuracy range (e.g. [26], [20], [2], [12]), and all these approximations include exponential terms, while $\Phi_{\kappa,\pi}(x)$ does not contain any. That is, to the best of our knowledge, in this accuracy range, there

is no other known approximation that has such a simple formula as $\Phi_{\kappa,\pi}(x)$. The known approximations with a better accuracy have more complex formulas, while the ones with similar complex formulas do not have a higher accuracy.

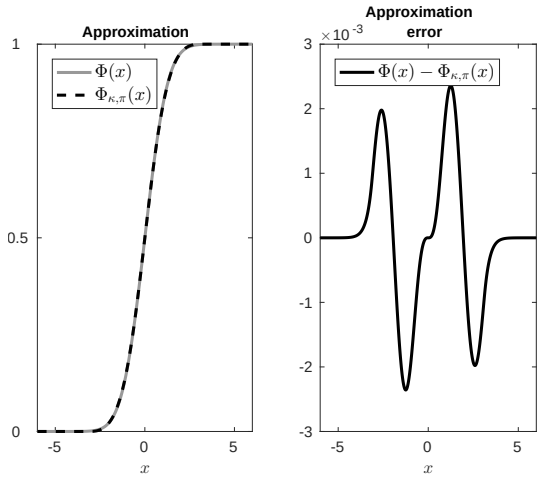


Figure 6
Approximation to the standard normal distribution function by a pliant (quasi logistic) distribution function

Figure 6 shows that the pliant probability distribution function fits quite well to the standard normal probability distribution function. Based on the above results, we can state the following proposition.

Proposition 5. *If* $\xi \sim D_p(\alpha,\beta,\gamma,d)$, $\eta \sim N(0,1)$, $\alpha = 2\sqrt{2/\pi}$, $\gamma = -1$ *and* $d = \pi$ *then for any* $x \in \mathbb{R}$,

$$\max_{x\in\mathbb{R}} |P(\eta < x) - P(\xi < x)| < 2.36 \cdot 10^{-3}. \tag{54}$$

Proof. See the previous results of this section. □

Conclusions

Table 1 summarizes how the pliant probability distribution function (pliant CDF)

$$F_P(x;\alpha,\beta,\gamma,d) = \begin{cases} 0, & \text{if } x \le \frac{d}{2}(\gamma-1) \\ \left(1-\gamma\omega_d^{(-\alpha,\beta)}(x)\right)^{\gamma}, & \text{if } x \in \left(\frac{d}{2}(\gamma-1), d\right) \\ 1, & \text{if } x \ge d, \end{cases} \tag{55}$$

where $\alpha, d \in \mathbb{R}$, $\alpha > 0$, $d > 0$, $\beta \in \mathbf{B}_\gamma$, $\gamma \in \{-1,1\}$, can be applied to approximate some well-known probability distribution functions. Based on the theoretical results and findings of our study, we may conclude that the pliant probability distribution

Table 1
Summary of approximations by the pliant CDF

Parameters and domain of $F_P(x) = F_P(x;\alpha,\beta,\gamma,d)$	Approximated CDF	Approximation error
$\alpha > 0, \beta > 0$ $\gamma = 1, d > 0$ $x \in (0,d)$	$F_W(x;\alpha,\beta) =$ $= \begin{cases} 0, & x \le 0 \\ 1 - \exp(-\alpha x^{\beta}), & x > 0 \end{cases}$	$\lim_{d\to\infty} \lvert F_W(x;\alpha,\beta) - F_P(x)\rvert = 0$
$\alpha > 0, \beta = 1$ $\gamma = 1, d > 0$ $x \in (0,d)$	$F_{exp}(x;\alpha) =$ $= \begin{cases} 0, & x \le 0 \\ 1 - \exp(-\alpha x), & x > 0 \end{cases}$	$\lim_{d\to\infty} \lvert F_{exp}(x;\alpha) - F_P(x)\rvert = 0$
$\alpha > 0, \beta = 1$ $\gamma = -1, d > 0$ $x \in (-d,d)$	$F_L(x;\alpha) =$ $= \frac{1}{1+\exp(-\alpha x)}$	$\lim_{d\to\infty} \lvert F_L(x;\alpha) - F_P(x)\rvert = 0$
$\alpha = 2\sqrt{2/\pi}, \beta = 1$ $\gamma = -1, d > 0$ $x \in (-d,d)$	$\Phi(x) =$ $= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t$	$\max_{x\in\mathbb{R}} \lvert \Phi(x) - F_P(x)\rvert < 2.36 \cdot 10^{-3}$

function may be viewed as an alternative to some key probability distribution functions including the Weibull-, exponential, logistic and standard normal probability distribution functions. We showed that our results can be utilized in a wide range of fields that include engineering, economics and the social sciences.

References

[1] S. J. Almalki and S. Nadarajah. Modifications of the Weibull distribution: A review. *Reliability Engineering & System Safety*, 124:32 – 55, 2014.

[2] K. M. Aludaat and M. T. Alodat. A note on approximating the normal distribution function. *Applied Mathematical Sciences*, 2(9):425–429, 2008.

[3] N. Balakrishnan. *Handbook of the Logistic Distribution*. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 2013.

[4] M. Bebbington, C. Lai, and R. Zitikis. A flexible Weibull extension. *Reliability Engineering & System Safety*, 92(6):719 – 726, 2007.

[5] S. Bowling, M. Khasawneh, S. Kaewkuekool, and B. Cho. A logistic approximation to the cumulative normal distribution. *Journal of Industrial Engineering and Management*, 2(1):114–127, 2009.

[6] G. M. Cordeiro, E. M. Ortega, and S. Nadarajah. The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 347(8):1399 – 1429, 2010.

[7] J. Dombi. On a certain type of unary operators. In *2012 IEEE International Conference on Fuzzy Systems*, pages 1–7, June 2012.

[8] J. Dombi and T. Jónás. Approximations to the normal probability distribution function using operators of continuous-valued logic. *Acta Cybernetica*, 23(3):829–852, 2018.

[9] J. Dombi and T. Jónás. Kappa regression: an alternative to logistic regression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, accepted paper, in press, 2020.

[10] J. Dombi, T. Jónás, and Z. E. Tóth. The epsilon probability distribution and its application in reliability theory. *Acta Polytechnica Hungarica*, 15(1), 2018.

[11] J. Dombi, T. Jónás, Z. E. Tóth, and G. Árva. The omega probability distribution and its applications in reliability theory. *Quality and Reliability Engineering International*, 35(2):600–626, 2019.

[12] O. Eidous and S. Al-Salman. One-term approximation for normal distribution function. *Mathematics and Statistics*, 4(1):15–18, 2016.

[13] M. Ghitany, D. Al-Mutairi, N. Balakrishnan, and L. Al-Enezi. Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis*, 64:20 – 33, 2013.

[14] M. E. Ghitany, E. K. Al-Hussaini, and R. A. Al-Jarallah. Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, 32(10):1025–1034, 2005.

[15] F. S. Hillier and G. J. Lieberman. *Introduction to Operations Research, 7th Ed.* McGraw-Hill, New York, USA, 2001.

[16] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*. Number 2 in Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley & Sons, 1995.

[17] J. Kies, N. R. L. (U.S.), and N. P. S. (U.S.). *The Strength of Glass*. NRL report. Naval Research Laboratory, 1958.

[18] C. D. Lai, M. Xie, and D. N. P. Murthy. A modified Weibull distribution. *IEEE Transactions on Reliability*, 52(1):33–37, 2003.

[19] Q. Li and H. Pham. A generalized software reliability growth model with consideration of the uncertainty of operating environments. *IEEE Access*, 7:84253–84267, 2019.

[20] J.-T. Lin. A simpler logistic approximation to the normal tail probability and its inverse. *Applied Statistics*, 39:255–257, 1990.

[21] I. Matic, R. Radoicic, and D. Stefanica. A sharp Pólya-based approximation to the normal cdf. *SSRN*, 2016. http://dx.doi.org/10.2139/ssrn.2842681.

[22] G. S. Mudholkar, D. K. Srivastava, and G. D. Kollia. A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91(436):1575–1583, 1996.

[23] D. Murthy, M. Xie, and R. Jiang. *Weibull Models*. Wiley Series in Probability and Statistics. Wiley, 2004.

[24] A. Papoulis and S. U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Higher Education, 4th edition, 2002.

[25] K. K. Phani. A new modified Weibull distribution function. *Journal of the American Ceramic Society*, 70(8):C–182–C–184, 1987.

[26] G. Pólya. Remarks on computing the probability integral in one and two dimensions. In *Proceedings of the 1st Berkeley Symposium on Mathematical Statistics and Probability*, pages 63–78, Berkeley, Calif., 1949. University of California Press.

[27] A. M. Sarhan and J. Apaloo. Exponentiated modified Weibull extension distribution. *Reliability Engineering & System Safety*, 112:137 – 144, 2013.

[28] G. O. Silva, E. M. M. Ortega, and G. M. Cordeiro. The beta modified Weibull distribution. *Lifetime Data Analysis*, 16(3):409–430, Jul 2010.

[29] K. Y. Song, I. H. Chang, and H. Pham. A software reliability model with a weibull fault detection rate function subject to operating environments. *Applied Sciences*, 7(10), 2017.

[30] A. Soranzo and E. Epure. Very simply explicitly invertible approximations of normal cumulative and normal quantile function. *Applied Mathematical Sciences*, 8(87):4323–4341, 2014.

[31] K. D. Tocher. *The Art of Simulation*. English University Press, London, 1963.

[32] L. Wang and Y. Shi. Reliability analysis of a class of exponential distribution under record values. *Journal of Computational and Applied Mathematics*, 239:367 – 379, 2013.

[33] W. Weibull. A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18:293–297, 1951.

[34] M. Xie and C. Lai. Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 52(1):87 – 93, 1996.

[35] R. Yerukala and N. K. Boiroju. Approximating to the cumulative normal function and its inverse. *International Journal of Scientific & Engineering Research*, 6(4):515–518, 2015.

[36] T. Yuge, M. Maruyama, and S. Yanagi. Reliability of a k-out-of-n system with common-cause failures using multivariate exponential distribution. *Procedia Computer Science*, 96:968 – 976, 2016.

[37] T. Zhang and M. Xie. Failure data analysis with extended Weibull distribution. *Communications in Statistics - Simulation and Computation*, 36(3):579–592, 2007.