Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

June 2011 **Volume III** **Number 2** **ISSN 2061-2079**

PAPERS

CALLS FOR PAPERS

ADDITIONAL

hte IEEE HUNGARY SECTION

Indexing information

Infocommunications Journal is covered by INSPEC and Compendex.

The journal is supported by the National Civil Fund.

Infocommunications Journal

Technically co-sponsored by IEEE Hungary Section

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest, Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu
Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: FOM Media

www.infocommunications.hu

Utilizing the power of Desktop Grid systems by Web 2.0 communities

(Invited Paper)

Attila Csaba Marosi, József Kovács, Péter Kacsuk

Abstract — **During the life of a Web 2.0 community, members are creating, sharing and manipulating media contents. The various media applications used by the individuals sometimes require much more computational capacity than it is available. This paper details how we integrated a grid infrastructure behind a Facebook application where computational resources are also provided by the community members. For this purpose, we integrated the SZTAKI Desktop Grid system together with the gUSE workflow enactor system and provided this support as a web service. In this system processing requirements and computational capacity comes from the same community forming a self-supporting mechanism regarding computation.**

Index Terms **— Web 2.0, desktop grid, workflow, accounting, volunteer computing, portal, community, watermark**

I. INTRODUCTION

In the last decades Grid systems offered large-scale distributed computing and storage services mainly for academic and university research. Researchers have become members of virtual organizations, where they were able to share research equipment, computing and storage facilities in a highly efficient way. In recent years focus is shifting towards to services and knowledge sharing, and extension of the application domain with standardized access methods and to overcome system heterogeneity become crucial point. The European scientific research community in the topic of Grid computing focuses on two main categories: service grids (SG) and desktop grids (DG).

Service grids are typically organized either from managed clusters (e.g. the gLite based EGEE grid) or from supercomputers (e.g. the Unicore based DEISA grid), and provide a 24/7 service for a large number of users who can execute their applications on those grids. The service grid middleware is quite complex and hence only a relatively few managed clusters or supercomputers take the responsibility of providing grid services. As a result, the number of processors in a cluster-based SG is moderate typically in the range of 1.000-100.000. Even the largest SG system, EGEE [1] has collected about 250.000 computers, however resources are divided between about 200 virtual organizations which means that a typical VO user can access only 1250 processors on average. Supercomputer grids can provide much larger number of processors but they are extremely expensive.

Desktop Grids (DGs) are strongly related to Volunteer Computing and are much simpler than SGs related to the software infrastructure they are relying on. Installation and maintenance of DG resources (i.e. machines donating resources to the grid) is extremely simple, requiring no special expertise thus, DGs are able to utilize non-dedicated machines even home computers. This means that, large number of donors can easily contribute to the pool of shared resources as in Volunteer Computing projects. Unlike Service Grids though, Volunteer Computing projects usually serve only a very limited user community (or target applications) who are able to use the resources for computation. The common architecture of Desktop Grids and Volunteer Computing systems typically consists of one or more central servers and (typically large number of) clients that connect to them from time to time. The central servers provide the applications and their input data in form of work units (WUs, i.e. conveniently sized chunks of computation to keep clients busy between contacting the server). Clients can join voluntarily, offering to download and run WUs of an application with a set of input data. When the computation has finished, the client uploads the results to the server where the master part of the application assembles the final output from the partial results returned by clients.

Desktop grids [2] are collecting large number of volunteer desktop machines to exploit their spare cycles. These donor desktops have no SLA requirement, their middleware code is extremely simple and hence the typical number of processors in volunteer desktop grids is in the range of 10.000-1.000.000. However, their drawback is that they can execute only some very limited number of pre-registered applications and users can use only these pre-registered applications on these systems.

The most well-known volunteer desktop grid is SETI@home [3] that collected over 3 Million CPUs worldwide. This Desktop Grid is based on BOINC which is an open-source software platform for computing using volunteered resources. It provides a framework in which public computing projects can be launched. A public BOINC project is aiming at providing workunits belonging to a particular scientific area, for which people can donate the spare computing cycles of their home PCs. Donation is realized by a BOINC client running on the PC in the background and performing workunit download, execution and result upload automatically. BOINC

Manuscript received May 19, 2011.

Attila Csaba Marosi is with the MTA SZTAKI, Computer and Automation Research Institute of the Hungarian Academy of Sciences, P.O.Box 63, 1528 Budapest, Hungary (phone/fax: +3613297864; e-mail: atisu@ sztaki.hu).

Jozsef Kovacs is with the MTA SZTAKI, Computer and Automation Research Institute of the Hungarian Academy of Sciences, P.O.Box 63, 1528 Budapest, Hungary (e-mail: smith@ sztaki.hu).

Peter Kacsuk is with the MTA SZTAKI, Computer and Automation Research Institute of the Hungarian Academy of Sciences, P.O.Box 63, 1528 Budapest, Hungary (e-mail: kacsuk@ sztaki.hu).

is the most popular distributed computing software and can be downloaded at [4].

SZDG [5] (SZTAKI Desktop Grid) is a software package with many extensions compared to BOINC: installation/deployment, application support, security, interfaces. SZDG provides easy-to-deploy software packages, and easy-to-setup project setup facilities. For application support there are several enhancements, like DC-API [6]for developing master/worker applications, or GenWrapper [7] to support the execution of legacy application. On the security side SZDG enhancements cover the utilization of X.509 certificates for application trust [8] while the provided virtualization/sandboxing technology [9] allows the execution of the application on the client resources in a safer way. Regarding the interfaces, different implementations of the DC-API enable the utilization of other distributed computing infrastructures (DCIs) like Condor [10] and a gateway called 3G Bridge [11] has been developed to forward jobs between different computing infrastructures, like gLite, BOINC, XtremWeb and to those supporting BES interface. Finally, the SZDG package also contains improved monitoring facilities.

The Web 2.0 term [12] stands for the second generation internet services, based mainly on communities, where users create the content as joint effort, or they share information. In Web 2.0 services the owner of the server provides only the service framework, the real content is uploaded and maintained, shared or evaluated by the users. Users communicate with each other, and they maintain connections, too. Generally, due to the large number of connections and users, Web 2.0 systems should handle heavy data traffic and complex relations, which need extraordinary large computational power in many cases. Many Web 2.0 applications deal with large files (e.g. video sharing); conversion or processing of such files demands large amount of resources. From Web 2.0 side, numerous application topics with large IT capacity can be solved by Grid technologies, like the already available advisor service system, or the media conversion service, or new envisioned services like digital signature and watermark service. Further community specific applications have also got high chances.

II. THE WEB2GRID PROJECT

The main aim of the Hungarian Web2Grid [13] project is to support Web 2.0 communities in dynamically building the necessary infrastructure that is needed for processing the high-end tasks of their community. The same way as the content is generated as joint effort of the members, Web2Grid technology enables the community to build the required infrastructure as a joint community effort exploiting the volunteer computing resources of the community members. In order to achieve this goal, Web2Grid exploits the BOINC and 3G Bridge technologies and other earlier initiatives focusing on the integration of Desktop Grid and Web 2.0 technologies. On the one hand, the Web 2.0 technologies assist to advertise Desktop Grid goals and attract computational resources for Desktop Grid communities. On the other hand, the developed new Web2Grid platform that includes Desktop Grid technology can offer back-end infrastructure for several compute-intensive Web 2.0 related problems:

- for operational requirements of Web 2.0 portals (i.e. large number of users, connections, file processing and large file conversions)
- for security requirements (watermarking, periodic revision of digital signature, etc.)
- for systems requiring large IT capacity
- and for tasks defined by community member end-users.

The combination of the Web 2.0 and Desktop Grid technologies provides advantages for both sides (as depicted on Figure 1), they can mutually benefit from each other's approach. In such combination Web 2.0 can extend its capabilities further from the community contents towards to shared services with the help of Grid technologies. From Grid point of view (right side of Figure 1), the project relies on the results of the EU FP7 EDGeS [14] infrastructure project which later has been further developed by the EDGI [15] project.

The main aims of these projects were the examination of the above mentioned targets and development of the tools, interfaces, methodologies to establish the integrated services both for private (local Desktop Grid), and public (volunteer Desktop Grid) environments.

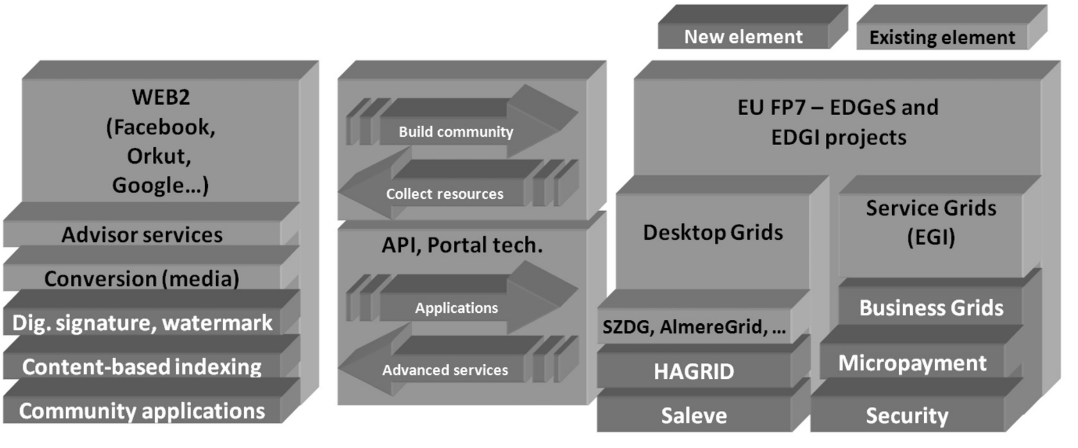


Figure 1 General concept & join efforts in the Web2Grid project

The projects focus(ed) heavily on security and accounting to facilitate the usage of service and commercial Grid systems. In these projects a new bridge was developed that integrates the usual Service Grids like EGI [16], and the volunteer computing Desktop Grids (like the SZTAKI Desktop Grid) that are very close to the main philosophy of the Web 2.0 technology (right side of Figure 1), since grid resources are provided by a community. During the Web2Grid project the aim is to extend these grids with technologies like micropayment and business security. Furthermore, the project also integrates other types of desktop grid like HAGRID [17] and Saleve [18].

This enhanced grid system then forms the computational resource needed for the execution of Web 2.0 community services (middle of Figure 1) while Web 2.0 community environments helps the public desktop grid system to collect new resources for the applications. Web 2.0 applications also gain advantages from advanced services like workflow execution provided by the grid side of the overall system.

On the left side of Figure 1, the web2 community using various applications (which may rely on grid services) and special Web 2.0 applications is being developed during the project to show how these applications can utilize the power of the grid or desktop grid systems. To demonstrate the developed system solution, the Web2Grid project consortium created Web 2.0 based demo applications from different fields, and provided easy-to-use interfaces and recommendations to foster the potential usage of the technologies.

III. SYSTEM ARCHITECTURE

One of the main goals of the Web2Grid project is to solve computation intensive tasks of Web 2.0 communities with the help of volunteer computing by harnessing the power of open (global) desktop grids. BOINC is the most popular volunteer computing system, thus the project heavily relies on the BOINC architecture. The connection between Web 2.0 and Desktop Grid is realized with well defined, low-level application development interfaces (API), and high-level, graphical application development solutions developed by the project partners. During the definition of the interfaces it was considered that Web 2.0 applications can come from various fields, and also Web 2.0 service providers, external consultants (e.g.: like security domain), or community users can develop applications (e.g.: Facebook has an API). These justify needs for the development of multi-level interfaces, and the project can benefit from numerous EU FP6 and EU FP7 project achievements. During the development of the interfaces, partners are focusing on protocols that are needed for commercial services that realize security and accounting functionalities. Figure 2 depicts the architecture of the platform developed by the project.

The architecture presented in Figure 2 contains four main blocks. First the platform may contain arbitrary number of „Web 2.0 applications" blocks. These are represented by web user interfaces of Web 2.0 community sites where users can use the services provided by the applications (e.g.: submit images for watermarking). Second block is the "Payment system" that is used for charging for the Web 2.0 services. Currently the Abaqoos micro-payment system is supported but the use of the Payment system is optional and always decided by the developers and service providers of a given application/service. The third block is the "Grid" that executes the compute-intensive tasks created by the Web 2.0 services: each service is represented as a workflow and computed using volunteer resources. Finally, the "Accounting system" amongst others coordinates the previous three blocks and accounts for the users, work done, and volunteer resources contributed by donors. In the following sub-sections we detail the main components of the platform.

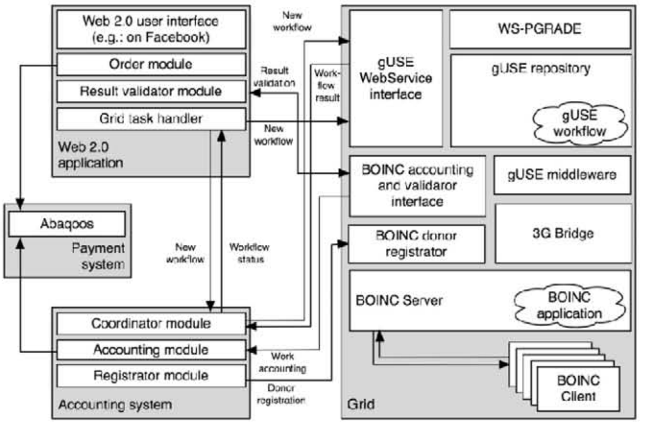


Figure 2 Architecture of the Web2Grid platform

A. The Web 2.0 application

A Web 2.0 application (depicted on Figure 2) consists of several components. First there is a user interface on one or more Web 2.0 social networking platforms for the application. The goal of the interface is to take "orders" from users and deliver the results for them. E.g. in case of a watermark embedding service the interface is used for selecting and submitting a batch of images to the service, and viewing the results later. A submission on the Web 2.0 interface is referred to as "order", since the user makes an order in the application (e.g.: to watermark a batch of images) and will get the results (e.g.: watermark embedded images) after the order is filled. It is the task of the Order module to track the submissions and to interface with the optional Payment system. If a Web 2.0 application requires payment for its services, the user is redirected to the optional Payment gateway. The order can be filled (e.g.: watermarks embedded) only after the payment transaction has finished. The Grid task handler module interfaces with the Grid and with the Accounting system: the Accounting system provides a unique identifier for the order, which is used throughout the system. The Result validator module is responsible for checking the results returned by the volunteer resources of the Grid. Usually the same task is replicated to different volunteer resources. The returned results are compared and if the majority of the results match they are accepted. We describe this mechanism in more detail in subsection E. Each Web 2.0 application has two additional components but they are not part of the "Web 2.0 application" block (on Figure 2), rather they are deployed directly in the Grid. First there is a gUSE workflow representing the Web 2.0 application, second there are one or more BOINC applications that are invoked from the workflow. We detail these components in subsection D.

B. The Accounting system

Although BOINC provides a credit system, it is not suited for the Web2Grid platform. BOINC uses a mechanism where the clients running on the computers of the donors monitor the CPU time used for computing a given task. On completion the clients report the computed results and the credits requested for it (based on the used CPU time) to the BOINC project. Any project may implement its own credit giving policy but usually the median requested credit is granted to each donor who uploaded a valid result for a given replicated task. In the case of Web2Grid where the virtual credit can be later exchanged for real money (which will be donated for charity) and where different desktop grid middleware can be used together, clearly this is not a solution, since donors might be able to manipulate the data reported by their clients. The project developed an own Accounting system which takes over some tasks from BOINC but its role is far more complex in the Web2Grid platform. The Accounting system is the key component since it is responsible (amongst many other things) for integrating the main components. The Web2Grid platform motivates donors to join and donate resources by giving virtual credits for their donated CPU time. These credits can later be exchanged for real money and donated for charity. Figure 2 shows the three major components of the Accounting system. The Coordinator module maps Desktop Grid tasks to one of the orders submitted on the Web 2.0 interface(s) of a given application (one such Web 2.0 orders may result in many Desktop Grid tasks). These tasks need to be accountable, meaning that they must be assigned to a user who sent the order on the Web 2.0 interface and to the donors who should get virtual credit for computing the tasks belonging to the orders. It is achieved by assigning a unique id for each order. This id is used by the Grid system for identification, since a single order may be split to many Desktop Grid tasks. Finally, the Registrator module registers the users signed up on the Web 2.0 interface to be donors in the Desktop Grid server and synchronizes the user database of the two systems.

C. User-donor registration and the Payment system

The Web2Grid platform recruits users and donors from the Web 2.0 social networking platforms (e.g.: Facebook or Orkut). Resource donors (or simply donors) are persons who donate their idle resource capacities (usually CPU time) to the platform, while users are the entities who want to use the services provided by the platform (e.g.: want to embed watermarks in the pictures of their albums uploaded to Facebook). Registration for both donors and users is done through web interfaces on Web 2.0 social networking platforms like Facebook or Orkut. Users can sign up for the specific Web 2.0 application or service, while donors can sign up on a central page rather than signing up directly at the web page provided by the Desktop Grid. Since the Web2Grid platform can utilize different Desktop Grid middleware, some kind of centralized administration of donor data is required. It is the task of the Registrator module of the Accounting system to maintain and synchronize donor registration data between the different Desktop Grid middleware and the Web2Grid platform.

The Payment system depicted on Figure 2 is an optional component. The developer and service provider of each Web 2.0 application/service decides whether she/he wants to charge for the service. Currently the Abaqoos micro-payment system by E-Group is supported. The payment service, where all payment transactions are authorized, is not part of the Web2Grid platform; rather it is an external service. Any registered user of the platform who wants to use a non-free Web 2.0 service needs to register at his/her first transaction at the Payment system too.

D. The Grid system

The Grid system executes the computation intensive tasks created by the Web 2.0 services. It has three main components: a) the gUSE workflow enactor; b) the BOINC Desktop Grid middleware and c) the 3G Bridge service connecting the previous two components. Each Web 2.0 application is represented by a workflow that consists of several nodes.

Figure 3 shows a sample workflow for a watermark embedding service that we detail in Section IV. The gUSE workflow engine stores and executes the workflows and each node in the workflow represents one or more computation task for the (Desktop) Grid. These may be executed on arbitrary supported distributed computing infrastructures (DCI), in this case on the supported Desktop Grid middleware.

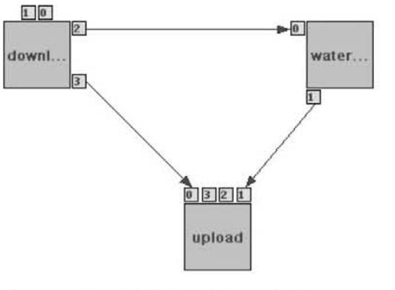


Figure 3 A sample Web2Grid workflow: watermark embedding into user uploaded images

An instance of the workflow is started for each order arriving from the Web 2.0 application. Each instance contains the unique id assigned by the Coordinator module to the order representing the workflow. Tasks belonging to these workflow instances are then executed by the Desktop Grid. The 3G Bridge is used to interface the different components and Grid middleware. Figure 4 depicts the main components of 3G Bridge. The SOAP web service interface is used to submit jobs to the 3G Bridge. Besides gUSE, different distributed computing infrastructures (e.g.: EGEE gLite, BOINC, Condor, etc) can submit jobs to various DCIs connected by the bridge. 3G Bridge stores all incoming jobs in an internal database and the Queue Manager component assigns them to a particular queue. The role of the queuing system is twofold; first, each queue is paired with an output plug-in on the Grid Handler Interface; second, multiple jobs within the same queue can be bundled (batched) together and sent to the destination computing infrastructure as a single job. This is usually used for tasks where the overhead of communication compared to the computation time required would be large. This plug-in system allows to easily connect different destination computing infrastructures (desktop grids in our case), since only an output plug-in for the destination needs to be added to 3G Bridge. In the Web2Grid project the gUSE submitter is used to send tasks to 3G Bridge and a BOINC destination plug-in is used to submit those jobs to a volunteer BOINC grid. 3G Bridge also provides an interface to query the status of jobs and retrieve their outputs once they are finished. This means that 3G Bridge hides the specific features of the underlying computing infrastructure from the submitter component.

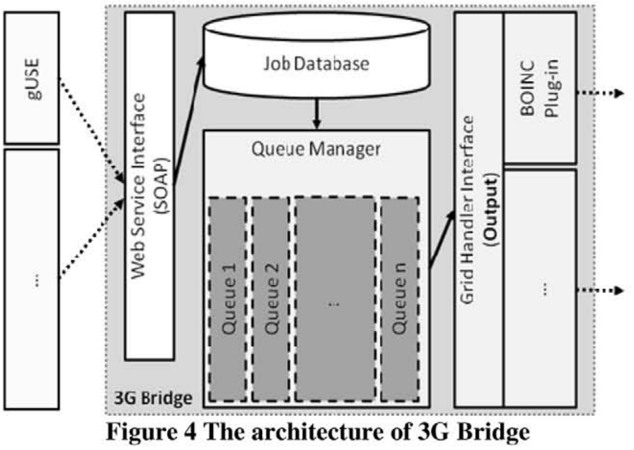


Figure 4 The architecture of 3G Bridge

Although 3G Bridge allows for connecting arbitrary Desktop Grid middleware, the Web2Grid platform requires three additional services from any connected middleware. These services are used by the Accounting system. First an interface for administering donors is required. This interface is used to keep the donor account synchronized in the system. The donors never register directly at the middleware, rather at a central interface operated by the Web2Grid platform. The second interface is provided for accounting information. The interface publishes data at a predefined interval (usually hourly) for the tasks executed and finished by the Desktop Grid middleware in the last period. The most important data is the unique id of the workflow instance the task belongs to; the computation time required; and the id of the donor who computed the task. This data is used by the Accounting system for giving virtual credits for the donors. Finally, a service is required for each application hosted by the Desktop Grid which is able to check the correctness or validity of the task computed on volunteer ("non-trusted") resources. The next subsection details this concept.

E. Remote validation of tasks

One of the challenges of using volunteer resources in desktop grids is that the administrator has no influence over the connected volunteer resources. These resources might be returning faulty results, either because they do not operate correctly (e.g.: they are overclocked) or malicious users are trying to get credits without performing any actual work. In any case the correctness (or validity) of the returned results must be checked. The checking procedure is referred as validation. Validation of the results can be done in two different ways. First, only the format of the returned output files is checked, e.g.: if an image is expected then a check is run to ensure that indeed an image was returned with the correct size, color depth etc. The second method is to use redundancy: multiple instances of the same work unit are sent to different resources and the returned results are compared. If a predefined number of the results of the same task match then they are accepted and declared as valid. Matching results is an application specific task, since it depends on the application what differences are acceptable (e.g.: when comparing text files results from different operating systems might use different characters for end of line markings, or when returning floating-point numbers, some minimal difference might be acceptable) and what not. The Web2Grid platform uses the second method, thus every Web 2.0 application needs to have a result validator module developed (as shown on Figure 2). BOINC provides a framework for developing result validator modules, it is detailed on Figure 5. For each application deployed at the BOINC server there is a running validator module ("Validator daemon"). The task of such a daemon is to perform the validation procedure described above. To be able to perform the validation it accesses the output files belonging to the results from the data server of BOINC and retrieves the required metadata from the BOINC database.

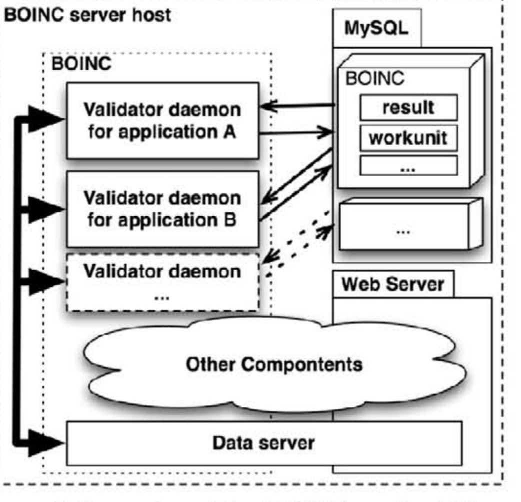


Figure 5 Operation of the BOINC result validator framework

The problem is with this approach is that the validator component is running on the Desktop Grid server, but in reality it should be part of the Web 2.0 application. A component, that was developed by a third party (the application developer) is running on the BOINC server (see Figure 5) which can be considered as a security risk since the validator daemon a) has to be validated itself that it contains no bugs or harmful code; b) has full access to the Data server and is able to read output files of results belonging to other applications and c) is running with the same privileges as the other server components.

One of the goals of the Web2Grid project is to make the Grid technologies available for as many Web 2.0 applications as possible, but this cannot compromise security. The first step to avoid this is to remove or separate all third party components from the Desktop Grid servers in a way that guarantees security but still provides a transparent solution for the application and validator developers. As a consequence, the monolithic single component architecture of the validator framework was replaced by a server-client architecture (detailed on Figure 6). This way the Desktop Grid server only runs the client part, which contains no application specific elements; all application specific code resides in the server part. This guarantees that the BOINC server only runs trusted code. The application developers need only to develop a server component, by using a framework provided by Web2Grid, and only a couple of application specific methods need to be implemented. The resulting server component may run on any host, it is always contacted by the client component running on the Desktop Grid server. The validation procedure is as follows.

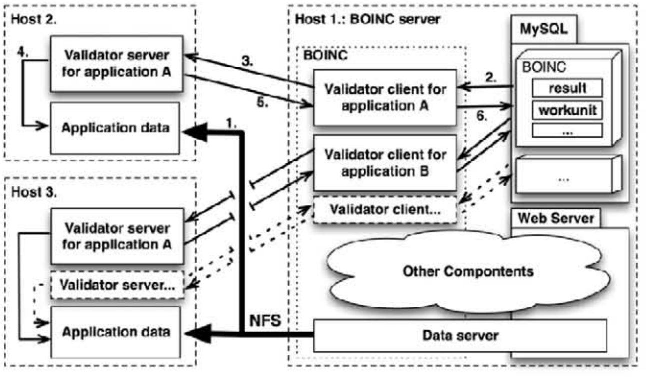


Figure 6 Operation of the improved BOINC validator framework

The first step is (Figure 6, Step 1) to grant read-only access to the Desktop Grid Data server for the Validator server. This can be achieved in many ways, for example by using NFS (Network File System) or by using remote access for the files via Parrot. From data security perspective it is important that the Validator server has read-only access to the input and output files belonging to the task it is validating. The Validator clients continuously poll the database of BOINC for tasks needing validation (Figure 6, Step 2). If a task is found its metadata is forwarded to the appropriate Validator server (Step 3). Since there can be quite a large number of tasks to be forwarded, a permanent connection and batched (streaming) communication is required between client and server or else the overhead of establishing and terminating connections would be enormous. The Validator server after receiving one or more tasks decides based on the data (and the input-output files) whether the task can be validated (if not it will send a “validation deferred” message to the client) and validates the results (Step 4). The server only performs data updates in memory (has no access to the database of the Desktop Grid server), all updated data is returned to the client (Step 5), which in turn updates the database (Step 6).

IV. USAGE SCENARIO: AN IMAGE WATERMARKING SERVICE

The goal of the watermarking service is to demonstrate the capabilities of the Web2Grid platform both for users and developers. The users can submit images for watermarking from their profile images on Facebook. The selected pictures are sent to the Grid to find the optimal parameter sets for embedding invisible watermarks using steganography technology into the selected pictures. After embedding the watermarks some attacks are performed in order to check the robustness of the embedded invisible watermark: e.g. after zooming, cropping, color depth reduction or JPEG compression of original pictures the watermarks can be still retrieved. In the watermarking application the user can upload images and submit them for processing.

Figure 7 Screenshot of the watermarking application

While the user is waiting for the watermarked images to be returned, the platform (architecture shown in Figure 2) per-forms the following: an order is created in the system and a workflow instance is started using the gUSE workflow engine in the Grid. The watermarking application workflow (see
Figure 7) consists of three nodes. The first node downloads the images from the application to gUSE; the second node submits the images to 3G Bridge; and the third node uploads the results. In 3G Bridge there is a queue defined for the watermarking application, this is where all images are sent, and 3G Bridge will submit those images as tasks to BOINC. 3G Bridge will periodically check the status of the submitted jobs in BOINC until they finish.

The application - doing the actual watermarking procedure and the simulated attacks - is deployed on the BOINC server. BOINC uses the volunteer resources of the donors: these resources run a BOINC client application that downloads the watermarking BOINC application and some images as input files from the BOINC server. The client starts the application that performs the watermarking and checking procedure. Once finished the client uploads the resulting watermarked images to the BOINC server. Once enough results are available (the Web2Grid platform uses redundancy to ensure the correctness of the returned results) the validation procedure starts to determine the correctness of the results. After the validation procedure finished and a correct result is chosen, the task is considered to be finished by BOINC, 3G Bridge will notice this and update the status of the given job in its queue. Once all jobs belonging to the workflow instance are finished the workflow terminates. The DocMark Web 2.0 application polls periodically the status of the workflows belonging to one of its orders, and if one finishes, then it will upload the resulting images from which the user can select the best quality.

V.RELATED RESEARCH

A community driven public service where the service is utilized/used and supported at the same time by the same community is the renderingfarm.fi [19] BOINC project where the members of the community can upload blender files for processing. The real processing is also distributed among the members attached to the BOINC server. As an application, this is a good example, how a community can use its own computational resources for calculating its own requests, however this project is currently not very strong in Web 2.0 aspects since they have not integrated their project into an existing Web 2.0 environment like Facebook.

An experimental project on linking Facebook and BOINC based desktop grids can be observable in one of Intel's initiative [20] called "Progress Thru Processors" in 2009. The aim of this initiative was to promote 3 different BOINC-based desktop grid projects through the Facebook communication channels. It was done in collaboration with GridRepublic which is one of the biggest BOINC user community building project. Similarity with our system is that Facebook was used as a frontend for the BOINC projects, however, contrary to our solution, they did not enable the job submission mechanism for the Facebook community members which is one of the main goal in our solution. We can say that in this particular case they relied primarily on the power of community building feature of Facebook.

The paper [21] presents an experiment on how to implement a Grid-based High Performance Computing solution for a rendering farm utilizing existing institutional desktop resources by scavenging the idle cycles. To achieve this goal, the open source Condor High Throughput Computing software was selected and implemented as a desktop Grid computing solution. The solution mainly focuses on building up a rendering farm based on condor. This service then could be accessed by institutional people, however public availability is not granted. This solution could be a good candidate for a web2 application, however the infrastructure is based on Condor which is not really suitable for collecting computational power of PCs owned by people at home.

A good example how a community can provide the needed computational capacity to operate its own services is the CancerGrid Computing System [22] developed by the EU FP6 CancerGrid project. This system aimed to provide a web-based portal combined with a desktop grid infrastructure to perform various chemical related calculations. CancerGrid consists of the gUSE web-based portal, an integrated molecule database browser and a BOINC based desktop grid underneath. From application point of view, several chemical algorithms have been deployed on the BOINC server and these algorithms were combined into workflows handled by the gUSE portal. In this infrastructure the community was formed by chemists, biologists, physicians and the computational resources were collected from the institutes these users came from. In this solution, the community was a private one, i.e., new members could join only after authentication. The reason behind it is the sensitivity of the data the system was handling, i.e., no volunteers were allowed to see the results of the computations.

The most interesting example which is closely related to the proposed solution in this paper is a high performance online service for genetic linkage analysis, called Superlink-online [23]. This system enables anyone with Internet access to submit genetic data and analyze it easily and quickly. The running environment is provided by Superlink@Technion - GridBot - Distributed Computing System which allows the usage of idle cycles of tens of thousands of home desktop PCs as well as high performance grids and cloud all over the world. This infrastructure receives jobs from the public and processes them by its on institutional computing resources as well as by a volunteer BOINC project. The two different type of grid is conducted by a high-level scheduler and using redundancy at job submission. However, this project was intended to support a single application and a single community. Our solution provides a framework for any web2 community to create and run their own applications.

VI. CONCLUSION

The Web 2.0 term is used as name for the collection of second generation services, which primarily based on communities. This means that members of the communities create the content together or share information. In such systems the owner of the server only provides the framework, the content is uploaded, created or shared by the community. Users communicate and create links among each other. Usually, a huge relational network is created by the members which causes heavy load on the server because of the huge data and relations to be handled. Some of the Web 2.0 applications themselves require relatively big computational capacity like video sharing, processing and conversion. So, in Web 2.0 system the need for computational grid system is a natural requirement due to the numerous tasks related to media conversion and advice systems which are popular nowadays. At the same time we expect new requirements to be raised like digital signature and watermarking.

This paper showed a solution how Facebook applications can utilize an underlying grid system with its computational power supported by the individuals of the community. In our system we have integrated a demo Facebook application for watermarking, an accounting system handling micro-payment if needed and the grid system. The underlying grid system is represented by a web service behind which there is a workflow enactor system called gUSE, a bridge and a volunteer desktop grid system based on BOINC.

The overall infrastructure is operated by the consortium to demonstrate the solution and to collect experimental information. The demonstration system was compiled together mainly by reutilizing already existing software components developed by European infrastructure projects.

Acknowledgment

The research has been supported by the Hungarian National Office for Research and Technology (NKTH) under grant No TECH_08-A2/2-2008-0097 (WEB2GRID). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 261556 (EDGI).

References

[1] M. Lamanna and E. Laure (editors), The EGEE User Forum – Experiences in using Grid Infrastructures. Special issue in Journal of Grid Computing, **6**(1) (2008), 1-123.

[2] D. Anderson. BOINC: A System for Public-Resource Computing and Storage, In Proceedings of the 5th IEEE/ACM International GRID Workshop, Pittsburgh, USA, 2004.

[3] D. Anderson, J. Cobb, E. Korpela, M. Lebofsky, and D. Werthimer, Seti@home: an experiment in public-resource computing. Communication ACM, **45**(11) (2002), 56–61.

[4] BOINC software homepage: http://boinc.berkeley.edu

[5] Peter Kacsuk, Jozsef Kovacs, Zoltan Farkas, Attila Csaba Marosi, Gabor Gombas, Zoltan Balaton, SZTAKI desktop grid (SZDG): a flexible and scalable desktop grid system, J. Grid Comput. 7 (4) (2009) 439–461.

[6] Gombas, G., Kiss, T., Terstyanszky, G.: DC-API: Bringing distributed applications to desktop grids. In: Proc. of CoreGrid Integration Workshop, 2006

[7] Attila Csaba Marosi, Zoltan Balaton, Peter Kacsuk: "GenWrapper: A Generic Wrapper for Running Legacy Applications on Desktop Grids", Proceedings of the 23rd IEEE International Parallel and Distributed Processing Symposium, pp. 1-6, 2009

[8] Attila Marosi, Gabor Gombas, Zoltan Balaton: "Secure application deployment in the Hierarchical Local Desktop Grid", Distributed and Parallel Systems - Cluster and Grid Computing (Proc. of the 6th Austrian-Hungarian Workshop on Distributed and Parallel Systems (DAPSYS)), P. Kacsuk, T. Fahringer, Zs. Nemeth (eds.), Springer, 2007, pp.:145-154

[9] Marosi, A.C.; Kacsuk, P.; Fedak, G.; Lodygensky, O.; , "Sandboxing for Desktop Grids Using Virtualization," Parallel, Distributed and Network-Based Processing (PDP), 2010 18th Euromicro International Conference on , vol., no., pp.559-566, 17-19 Feb. 2010 doi: 10.1109/PDP.2010.90

[10] Douglas Thain, Todd Tannenbaum, Miron Livny: Distributed computing in practice: the Condor experience. Concurrency - Practice and Experience 17(2-4): 323-356 (2005).

[11] Zoltán Farkas, Péter Kacsuk, Zoltán Balaton, Gábor Gombás: "Interoperability of BOINC and EGEE", Future Generation Computer Systems, Volume 26, Issue 8, October 2010, pp.:1092-1103

[12] T. O'Reilly. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. 2005

[13] The Web2grid project homepage: http://web2grid.econet.hu

[14] The EDGeS project homepage: http://www.edges-grid.eu

[15] The EDGI project homepage: http://www.edgi-project.eu

[16] The EGI organization homepage: http://www.egi.eu

[17] The Hungarian Advanced Grid: http://hagrid.econet.hu

[18] Péter Dóbé, Richárd Kápolnai, and Imre Szeberényi: "Saleve: toolkit for developing parallel grid applications", Infocommunications Journal, LXIII(1):60-64, January 2008

[19] Public rendering site: http://renderfarm.fi

[20] Intel's Progress Thru Processors: http://www.facebook.com/progressthruprocessors

[21] Patoli, Zeeshan; Gkion, Michael; Al-Barakati, Abdullah; Zhang, Wei; Newbury, Paul; White, Martin: "How to Build an Open Source Render Farm Based on Desktop Grid Computing", Book chapter, Title: Wireless Networks, Information Processing and Systems, Series: Communications in Computer and Information Science, 2009, Springer Berlin Heidelberg, ISBN: 978-3-540-89853-5, pp.: 268-278, vol. 20

[22] Jozsef Kovacs, Peter Kacsuk, Andre Lomaka, Using a private desktop grid system for accelerating drug discovery, Journal of Future Generation Computer Systems, Volume 27, Issue 6, June 2011, Pages 657-666, ISSN 0167-739X, DOI: 10.1016/j.future.2010.12.008.

[23] Mark Silberstein: "Building an online computing service over volunteer grids", Invited paper in proceedings of the 5th Workshop on Desktop Grids and Volunteer Computing Systems (PCGrid 2011), Anchorage (Alaska) USA, May 20, 2011

Attila Csaba Marosi is a member of the Laboratory of Parallel and Distributed Systems at the Computer and Automation Research Institute of the Hungarian Academy of Sciences since 2001. He received his M.Sc. from the Budapest University of Technology and Economics in 2006. He started PhD studies at the Budapest University of Technology and Economics in 2009. His research interest includes desktop grid and cloud computing. He participated in many national (HAGRID, Web2Grid) and international (EGEE, CoreGrid, EDGeS, EDGI) research and development projects. He is the coauthor of more than 20 scientific papers in journals and conference papers related to grid computing.

József Kovács was born in 1975 in Budapest, Hungary. He is a Senior Research Fellow at the Laboratory of Parallel and Distributed Systems (LPDS) at the Computer and Automation Research Institute of the Hungarian Academy of Sciences. He got his B.Sc. (1997), M.Sc. (2001) and Ph.D. (2008). In the meanwhile he was continuously involved in numerous national, international and European (Esprit, FP5, FP6, FP7) research projects. From 2002 he was doing research in the field of parallel checkpointing techniques in grids and later in in the field of DesktopGrid computing. From 2006 he is the leader of the DesktopGrid team of LPDS. He is the author and co-author of more than 40 scientific papers.

Peter Kacsuk is the Head of the Laboratory of Parallel and Distributed Systems in MTA SZTAKI Computer and Automation Research Institute of the Hungarian Academy of Sciences. He received his M.Sc. and university doctorate degrees from the Technical University of Budapest in 1976 and 1984, respectively. He received the Candidate degree (equivalent to Ph.D.) from the Hungarian Academy in 1989. He habilitated at the University of Vienna in 1997. He received his professor title from the President of Hungary in 1999 and the Doctor of Academy degree (D.Sc.) from the Hungarian Academy of Sciences in 2001. He has been a part-time full professor at the Cavendish School of Computer Science of the University of Westminster and the Eötvös Lóránd University of Science Budapest since 2001. He has published two books, two lecture notes and more than 200 scientific papers on parallel computer architectures, parallel software engineering and Grid computing. He is co-editor-in-chief of the Journal of Grid Computing published by Springer.

Improved On Demand Clustering on Scale-Free Topologies

Borbála K. Benkő, Márton Legény

Abstract—**We define and analyze two directions to improve On Demand Clustering (ODC) in self-organizing networks. The first direction, Spyglass, uses excessive lookup in the match maker's two-hop neighborhood. The second direction, Shuffling ODC, applies link reconfiguration also in case of unsuccessful search, causing minor modifications in the adjacent topology without immediate gain. Both directions were evaluated with simulation using scale-free graphs as network abstraction. Spyglass results in significantly faster reaction times and larger clusters than the basic ODC, but with a significantly higher communication overhead. Shuffling ODC brings similar advantages, significantly larger clusters and less separated nodes, but without the increment in the message count. Shuffling ODC, due its emergent self-shuffling effect, is able to reach more nodes than the other two, without breaking the rule of locality. We found 20-30% improvement in job processing and 35-50% gain in job sharing with our algorithms, compared to the baseline ODC.**

Index Terms— **Ad-hoc networks, Clustering, Load-balancing, Self-organization**

I. Introduction and Background

Self-organization in networks is gaining significant interest today in various fields from 3GPP LTE [1] to pervasive computing [2], [3] and other mobile ad-hoc network scenarios. Elements in these networks need to share resources, communicate, perform load balancing or provide services for each other without the help of a central management entity. While human network managers possess information about the state of the whole network, in case of self-organization, nodes need to rely on *locally available information,* even if it is incomplete and non-objective, when making their autonomous decisions.

Self-organization often uses simple algorithms with *emergent* properties, i.e. the multitude of executions result in a complex, "intelligent" high-level behavior which is of a different quality than the simple building blocks themselves. Emergent algorithms are often inspired by biological, chemical or physical phenomena like the behavior of swarms, insect colonies, the human brain or immune system. These paradigms were used in numerous ways to solve problems in computer networks. A thoroughly researched application field is load balancing [4], [5], [6], [7]. Self-organization in overlay networks, without the need for a supervising entity, is also widely used in data grids, peer-to-peer systems and other types of distributed applications [8], [9], [10]. It is notable that due to the low processing power these algorithms demand, self-organization can even be used in wireless sensor networks [11], [12].

Clustering, in a self-organizing network, means that entities of the network search for other entities that meet a certain criterion (similarity in case of normal clustering and complementariness in inverse clustering) and establish connections with them in form of overlay network links. [13] [14] The resulting overlay network can be used for various purposes. *Load balancing* is one of these use cases; here, entities share their local workload with members of the cluster. While the efficient creation of clusters is a prerequisite for efficient load balancing, it is not a sufficient condition: the load balancing algorithm must also answer the questions when and with which cluster member to share the load, and what information to use (collect, store and update) when making these decisions.

Our work focuses on biologically inspired, fully distributed self-organization algorithms for large overlay networks, with an emphasis on clustering and load balancing. We considered the problem of on-demand clustering, where the goal is to expand the cluster when demand for that occurs, for example the cluster has more workload than it can process. The demand-driven problem statement has two main differences from the general case: (i) the amount of clustering is proportional to the demand, hence there's no driving force to expand clusters to their theoretical maximum size or to assign each and every node to a cluster, and (ii) the speed of cluster formation is more crucial, because the demand depicts an immediate request that cannot wait.

The novelty in the paper is the improved clustering ability which we demonstrate with a load balancing scenario. We extend a well-known baseline clustering algorithm in two directions (Spyglass and Shuffling ODC), and analyze how the modifications affect the results. Shuffling ODC is introduced in this paper. Early versions of Spyglass have been discussed in [15] and [16], however, always evaluated with random topologies. Now we evaluate and compare the algorithms using scale-free topologies, which provide a much more realistic abstraction of real networks.

The rest of the paper is organized as follows. Section II describes the three clustering algorithms: the baseline On Demand Clustering, and the two extensions Spyglass and Shuffling ODC. Section III defines the load balancing mechanism

Manuscript received March 28, 2011.
Borbála Katalin Benkő is with the Budapest University of Technology and Economics, Hungary, H1117 Budapest, Magyar tudosok krt 2. (phone: 36-1-4633219; fax: 36-1-4633263; e-mail: bbenko@hit.bme.hu).

Márton Legény was with the Budapest University of Technology and Economics, Hungary, H1117 Budapest, Magyar Tudosok krt 2. (e-mail: legeny@hit.bme.hu). Now he works as a Software Engineer at Nokia Siemens Networks Hungary.

where the clustering is used. In Section IV and V we investigate the algorithms along different network abstractions through simulation and compare the results. Finally, Section VI concludes the paper.

II. Clustering Algorithms

While numerous clustering algorithms are known today, we chose a specific algorithm family, known as On Demand Clustering. This section, after drafting how the base algorithm works, describes our two extensions to it, namely Spyglass and Shuffling ODC.

A. On Demand Clustering

On Demand Clustering, or ODC, is an emergent self-organization algorithm with beneficial properties on the node degree, developed at BT Labs [17], [18], [19]. The algorithm can be summarized as follows.

1. The clustering process is initiated on demand, i.e. when a node is in need of expanding its cluster.
2. The node where the demand for clustering raises, called initiator, selects one of its neighbors to serve as match maker.
3. The match maker looks for a matching node, one that meets the initiator's clustering criterion, among its own neighbors. Returns the match if found.
4. When the match is not already a neighbor for the initiator, a process called rewiring begins. The initiator and the match establish a new link, while, in order to keep the total number of links under control, the match maker removes its own link to the match.

It has been shown that ODC results in an emergent self-organization behavior, clusters are formed and expanded in accordance with the local demand. However, ODC may not perform equally well in all cases. While the strict locality principle (when searching for a match) guarantees that the communication overhead remains low, it sometimes prevents the fast formation of clusters. This is especially the case in sparse or type-wise highly diverse networks, where often no suitable match is present in the neighborhood of most match makers. The locality principle here causes not large enough clusters or not fast enough clustering.

ODC uses message based communication; each request, answer, link creation and link removal step costs a message.

B. Spyglass

The motivation behind Spyglass was to overcome the strict locality principle of ODC by extending the range of the lookup.

In Spyglass [15], [16] the match maker is able to look one hop farther than in the original ODC, i.e. at its two-hop neighborhood. Spyglass differs from ODC in the last two steps: in the behavior of the match maker and in the details of rewiring.

3. The match maker looks for a matching node, one that meets the initiator's clustering criterion, among its own neighbors. When no match is found there, the match maker continues with checking its two-hop neighbors (the neighbors of its neighbors) for a match. Returns the match if found.
4. When the match is not already a neighbor for the initiator, the rewiring begins. The initiator and the match establish a link, while, in order to keep the total number of links under control, the original connection node to the match removes its own link towards it. The connection node is the match maker if the match is originally its own direct neighbor. Otherwise the connection node is the node between the match maker and the match.

The two-hop search horizon, clearly, promises more lookup success. However, looking at two-hop neighbors is an operation of exponential cost, so some kind of optimization is inevitable if we want to avoid the exponential communication overhead. Hence, Spyglass nodes are equipped with a local data structure called Neighbor Cache (NC) storing information about nodes in the vicinity. NC needs to be established and updated with care; while both the build-up and refresh steps are expensive, having outdated content is even worse as it hinders the self-organization process.

In [20] we discussed various caching strategies that provide good tradeoff between communication overhead and keeping the cached contents up-to-date.

The biggest advantage of Spyglass is its reaction speed to requests. The ability of returning a match fast and with a high probability is a very beneficial property, because clustering happens on demand, so a request to the match maker means that the initiator really needs the match as soon as possible. If the expansion of the cluster is too slow the initiator may reach critical levels of overload.

Spyglass, just like ODC, uses message based communication. It introduces new lookup messages for detecting the match maker's two-hop neighbors.

C. Shuffling ODC

The motivation behind Shuffling ODC was to overcome the strict locality principle of ODC by allowing small changes in the topology without direct gain (i.e. a useful overlay link).

Shuffling ODC extends the original ODC algorithm with two new ideas: (i) preferring non cluster member nodes as match makers, and (ii) "mingling with the vicinity", i.e. introducing minor topological changes, when no match is found. These modifications are achieved with minor changes in steps 2-4 of ODC.

2. The node where the demand for clustering raises, called initiator, selects one of its non-cluster neighbors to serve as match maker. When there are no non-cluster neighbors available, a cluster neighbor is selected as match maker.
3. The match maker looks for a matching node, one that meets the initiator's clustering criterion, among its own neighbors. If found, returns the match. If not found, returns a random node from its neighbor list.
4. When the returned node is not already a neighbor for the initiator, a rewiring process begins. The initiator and the match establish a new link, while, in order to keep the total number of links under control, the match maker removes its own link to the match.

Shuffling ODC tries to overcome the topologic burden caused by ODC's positive-feedback-only rewiring. We found that in some cases, in ODC, the set of available match makers for a node is not dynamic enough: only the node's original neigh-

bors and newly found matches can play this role. After a while, this set easily runs out of matching but not-already-cluster-member neighbors, i.e. the initiator is unable to expand its cluster via either of the match makers. Shuffling ODC tries to overcome that by introducing new potential match makers in case of a no-match, and, even better, doing this at nearly no additional cost.
The preference of non-cluster (non matching color) nodes as match makers comes from the observation that often, direct cluster members of the initiator also get overloaded thus will initiate clustering. Hence, the potential matches within a cluster member's neighbors are there with a reason (they are already used for load sharing by the neighbor). On the other hand, a non cluster member node could potentially bring true new matches, ones not already participating in the load sharing chain.
Shuffling ODC uses the same communication messages as ODC.

III. The Load Balancing Task

The load balancing task is a common application area of clustering. We use a model where a load balancing problem generates the demand for the clustering.
The model is as follows (see Figure 1).

- The overlay network consists of colored nodes and links connecting them.
- Nodes are able to process matching-color jobs.
- Links are not colored.
- Jobs enter the overlay network via colored workload generators, each statically attached to a matching-color node.
- Workload generators generate jobs and put them on the queue of the attached node. The expected value of the generation rate is constant.
- Nodes consume jobs from their local queue.
- When a node feels to be overloaded, i.e. its local queue length exceeds a limit, it shares the local workload with its cluster neighbors, i.e. matching-color neighbors. Load sharing means to transfer a job from the local queue to each neighbor via the overlay link. The sharing decision may also be bound to specific conditions on the remote node (e.g. the acceptor is not overloaded) and the capacity of the link.
- When a node feels overloaded and cannot find enough appropriate neighbors to share the load with, a demand for clustering occurs.

Hence, clustering is aimed to reorganize the overlay topology and to create a new link to a suitable node on demand.
Note that the load balancing task introduces a certain amount of new requirements towards the clustering algorithm. Normally, the resulting cluster size is the most important performance metric of a clustering algorithm. However, when clustering serves load balancing purposes, the need for cluster expansion is limited. We do not need to generate the largest possible clusters, instead, just large enough clusters for the local excess workload (the creation of clusters larger than that would not bring further advantages but would cost communication messages). On the other hand, the clustering speed becomes vital for load balancing: the initiator needs the match urgently. Given that the job generation rate is constant, every unsuccessful search for a match just worsens the initiator's situation.

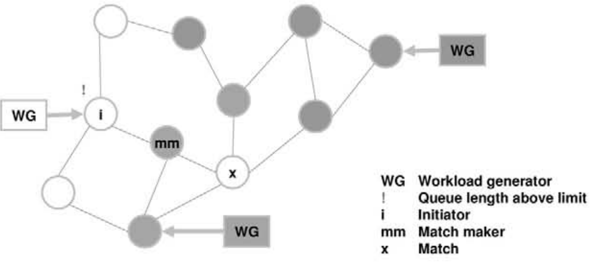


Figure 1: Graphical representation of the model: colored nodes and colored workload generators. If the local queu length is exceeds a limit, the node shares its workload with cluster neighbors, and when that is not enough, initiates a self-organization step too (selects a match maker and asks for a match).

IV. Evaluation Setting

A simulation environment and a set of evaluation scenarios were created in order to objectively evaluate the clustering algorithms through a load balancing scenario.

A. Network Abstraction

For the evaluation we used the scale-free graph model as the abstraction of the network. While the random graph model is also widely used for evaluating networked algorithms, the scale-free model matches many real-life problems more precisely [20]. In both models, the total number of links between the nodes is fixed. However, while the node degree distribution in case of a random graph is a Poisson distribution, for scale-free graphs it follows the power law, i.e. there are a few nodes in the network with very high node degrees (similar to the presence of hubs and routers on the Internet).
Our scale-free topologies were generated with Kumar's copy method [21], which is an iterative model. In each generation step, one new node is added to the network. The new node gets connected to some of the existing nodes, with a probability proportional to the current node degree of each existing node. Separated (zero-degree) nodes, if any, at the end of the generation phase were connected to randomly selected other nodes in order to guarantee that all nodes have at least one neighbors.
We used networks of 1k nodes with ~2k links, and 10k nodes with ~20k links for the evaluation. The exact numbers are:

- Basic scale-free graph: 1000 nodes, 2029 links.
- Large scale free graph: 10,000 nodes, 20,028 links

B. Evaluation Environment

The simulation framework was written in Java 6 SE, and the experiments took place on a desktop PC with 2GHz dual core processor and 2 GByte RAM.

C. Evaluation Scenario

Measurements were conducted on the above showcase networks. Workload generators were attached to 30% of the nodes. Nodes belonged to 10 different classes (colors), and load balancing swung into action once the local queue length exceeded the static limit of 5 unprocessed jobs. The choice of network size and workload generation density was motivated by our previous work on random networks [15], where this problem size was found to be convenient for demonstrating the scalability of the emergent algorithms as well as for pointing out the differences between directions.

The simulation was executed for 500 rounds. The generation rate of workload generators changed between 1 and 10 jobs per round (selected randomly at startup, and constant during the life time of the generator), and the duration of the generation lasted in some cases for 200 rounds while in others for 500 rounds. When the generation is ceased before the end of the experiment, the shape of the tail curve gets visible too.

We simulated the same scenario, excess workload triggering self-organization, in all cases. Network dynamics, i.e. the appearance of disappearance of nodes, was not considered during simulation.

D. Evaluation Criteria

The following metrics were applied for evaluation:

- Message count depicts the amount of ***communication overhead*** generated by the algorithm variant. Small message counts are preferred over high message counts.
- ***Number of clustered nodes*** during simulation and at the end. Intuitively, the larger this number is the better the self-organization performs. However, the demand for clusters is constrained by the amount of workload, so the cluster size cannot grow indefinitely.
- The ***unprocessed job curve*** drafts the processing dynamics of the system by plotting the number of injected but unprocessed jobs versus time. The unprocessed job curve displays two very important properties of the clustering algorithm: (i) the *reaction time* is manifested in the shape and in the peak of the curve, smaller peaks and flatter and more rapidly decreasing curves are preferred, and (ii) the area under the unprocessed job curve depicts the total waiting time, which is the smaller the better.
- The ***number of overloaded nodes*** depicts the dynamics of the demand for clustering. This metrics can be used for two purposes: (i) to understand the characteristics of the demand that triggers clustering; and (ii) to understand how the demand is silenced or excited by the clustering and load balancing algorithm. Note that a node without an attached workload generator may also get overloaded (e.g. receives a new job per round from more than one sources), hence trigger additional clustering. We call this phenomenon *secondary overloading*.
- We also use a ***visual evaluation*** of the self-organized network, using on a spring based layout engine that draws a 2D approximation of the topology, considering each link as a spring and each non-cluster node as a magnet with the same pole. Visually evaluated properties of the network include balance between colors, distribution of non-clustered nodes, etc.

V. Evaluation

The evaluation section first discusses the results measured on the Basic Scale-Free Network, then, elaborates on the differences and similarities observed on the Large Scale-Free Network. On both networks, we used two settings: (i) an excitation setting where job generation lasted for 500 rounds (workload generators generate more work than what the network is able to process, so this setting excites self-organization), and (ii) a run-up tail-off setting, where job generation stops after 200 rounds, giving time for the network to consume some of the excess workload.

In the second part of the section we investigate how the change in different parameters affect the results. We analyze what happens if we take the physical topology into account. Afterwards, we provide and example how the system changes if the task becomes easier or harder, i.e. the number of colors is decreased or increased, and also discuss the effect of the network abstraction.

A. Basic Scale-Free Network

1) Excitation Shapes

The first experiment tackles with the case when, with generation throughout all 500 rounds of the experiment, workload generators flood their attached node with jobs (for generation rates ≥ 1, the attached nodes must use clustering and load balancing in order to keep track with the load). The three algorithms show different behaviors.

Figure 2 shows the number of clustered nodes per round. With each algorithm, the curve starts with a rapid rise; the more rapid the rise the better the algorithm's reaction time. ODC is the slowest from the three, and after the increment phase, it soon reaches a static limit (588 clustered nodes), where match makers do not have matching neighbors any more, hence no further match can be returned. Spyglass reacts to the clustering request very rapidly, in the first few rounds it is by far speedier than the other two algorithms. However, after reaching a peak at 740 clustered nodes, the curve starts to decrease. The explanation for that is the following. When the match maker returns a new indirect match (there is a connecting node between the match maker and the match), then the match maker and the connecting node need to remove their link. This, at the first sight, may not decrease the number of clustered nodes. However, if the initiator and the match are of color A, and the match maker and the connection are of color B, then, this step disconnects the cluster members of color B. And, if the match was already clustered before the step (in the cluster of some other node than the initiator), and the match maker and the connection node are one another's only cluster center, then, all together, this step decreases the number of clustered nodes by disconnecting one of the color B nodes from their group. In Shuffling ODC, the same decrement effect can be observed, but, due to the small local rewirings without gain, the algorithm manages to turn back this trend.

The number of processed jobs per round in Figure 3 shows exactly the same trend. For each algorithm the job processing curves are slightly below the clustered nodes curve, showing that in each round there are a few clustered nodes without load. This is mostly due to the natural clusters coming from the initial topology: some matching-color nodes are neighbors, i.e. the count as clustered nodes, but never receive a job in their life.

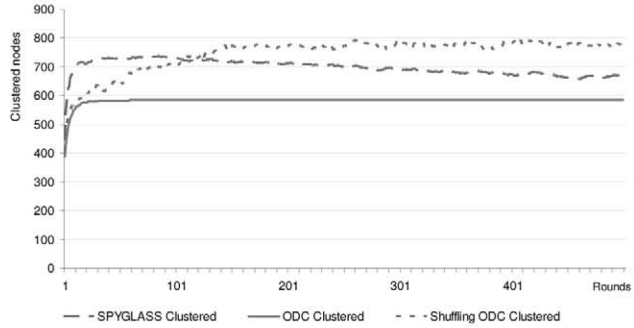


Figure 2: Number of clustered nodes (Basic Scale-Free network, generation time 500 rounds)

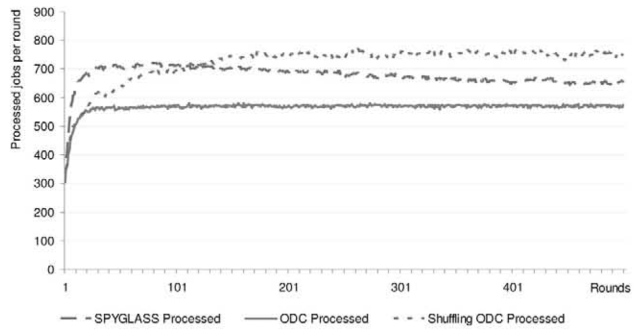


Figure 3: Number of processed jobs per round (Basic Scale-Free network, generation time 500 rounds)

Figure 4 depicts the number of shared jobs per round. The curves, again, follow the trend already seen above, however, with smaller numbers (the work done by the nodes with generators is not shown here.)

Figure 5 visualizes the nature of the overload. The set of darker curves shows the number of overloaded nodes that have workload generators attached (primary overload) per round, while lighter shades display the number of overloaded nodes without workload generators (secondary overload). Primary overload curves follow the same shape, in Shuffling ODC a slight constant rise can be observed, which is due to the fall in the number of clustered nodes (nodes that could bear the load with the help of their cluster get overloaded after losing a cluster member). Secondary overload curves have an initial peak, but with time, when these nodes also find partners for workload sharing, their number decreases rapidly.

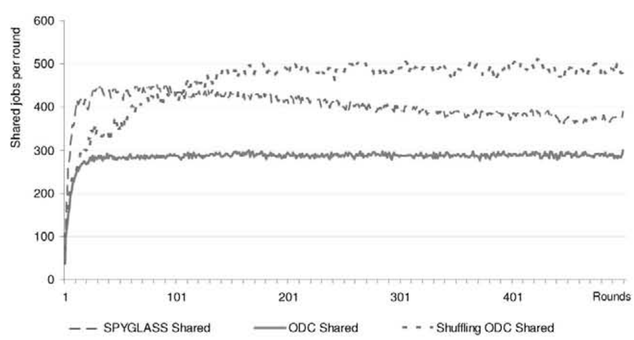


Figure 4: Number of shared jobs per round (Basic Scale-Free network, generation time 500 rounds)

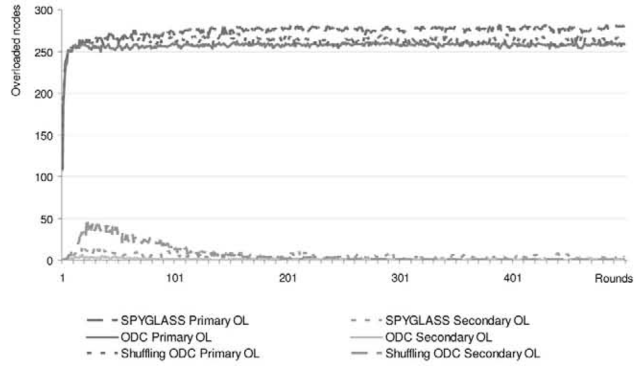


Figure 5: Number of overloaded nodes, with primary load (direct from generators) and secondary load (from other nodes) (Basic Scale-Free network, generation time 500 rounds)

The unprocessed job curve in Figure 6 confirms the previous result. ODC is the weakest in bearing with the overload. Spyglass, after initially being the strongest of the three, falls back to the second place due to the breakup of clusters. Shuffling ODC is similar to Spyglass in the overall performance, but without the performance fallback.

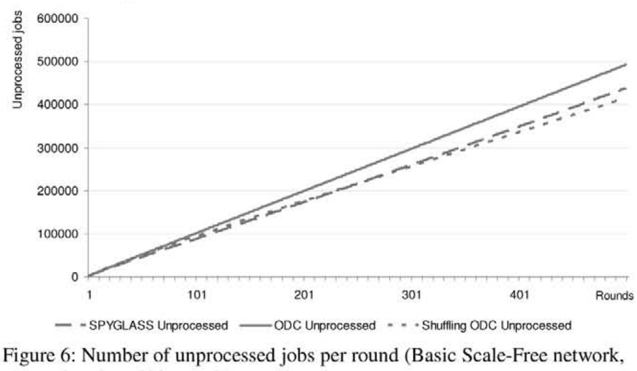


Figure 6: Number of unprocessed jobs per round (Basic Scale-Free network, generation time 500 rounds)

Figure 7 displays the messaging overhead of each algorithm. Clearly, Spyglass with its excessive lookup consumes one magnitude more messages than the other two. ODC is the most economic in this sense. Shuffling ODC produces a linear overhead compared to ODC, with its rewirings without direct gain.

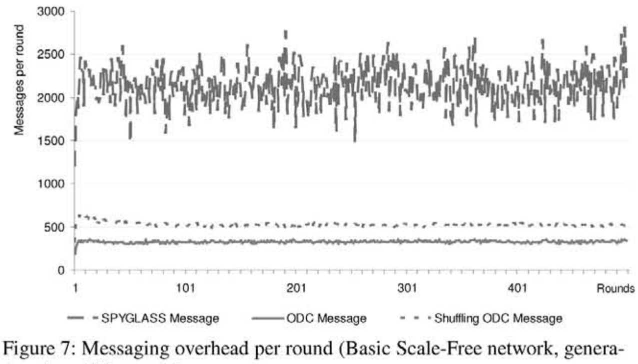


Figure 7: Messaging overhead per round (Basic Scale-Free network, generation time 500 rounds)

The resulting topologies were also evaluated with the help of a layout engine that produces a contraction and distraction based 2D layout for the overlay network. Without elaborating on the details, large circles display the nodes with workload generators (the diameter is proportional to the count of cluster neighbors), small empty circles are clustered nodes, and small squares are non-clustered nodes. Workload generators, when they are each other's neighbors, are connected by a visual link in the figure. (The layout algorithm concentrates on the non-generator-attached neighborhood of generator nodes, and sometimes fails to place workload generators near enough, even if they're connected. The link helps detecting these rendering failures.)

Figure 8, Figure 9, and Figure 10 shows the layout produced by ODC, Spyglass and Shuffling ODC, accordingly. ODC produces the less connected layout: a large number of solitary, un-clustered nodes and only small clusters can be observed. Clusters, if any, tend to stay of strictly local nature. Spyglass and Shuffling ODC produce approximately the same level of organization, a level that is significantly higher than ODC's. The difference in the number of solitary nodes can also be observed: in Shuffling ODC there are clearly less standalone squares. The homogeneity of all algorithms is good: cluster sizes and the distribution of non-clustered nodes are balanced.

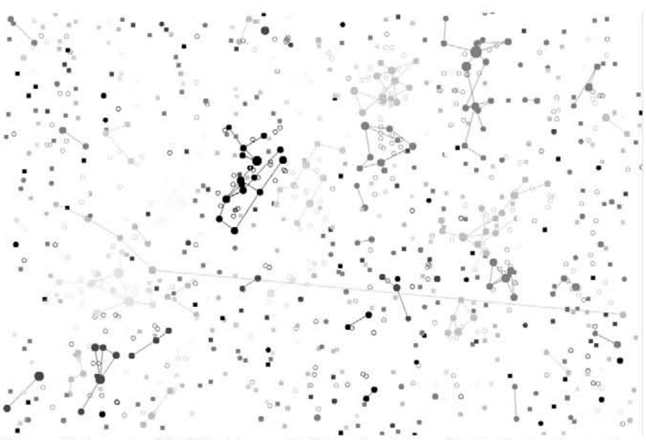


Figure 8: Layout of ODC at the end of the simulation (Basic Scale-Free network, generation time 500 rounds)

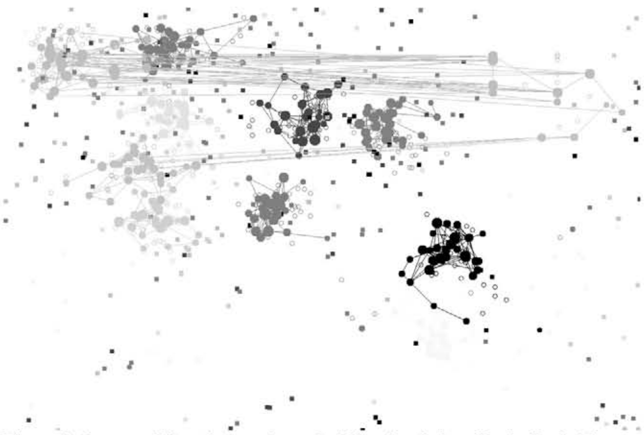


Figure 9: Layout of Spyglass at the end of the simulation (Basic Scale-Free network, generation time 500 rounds)

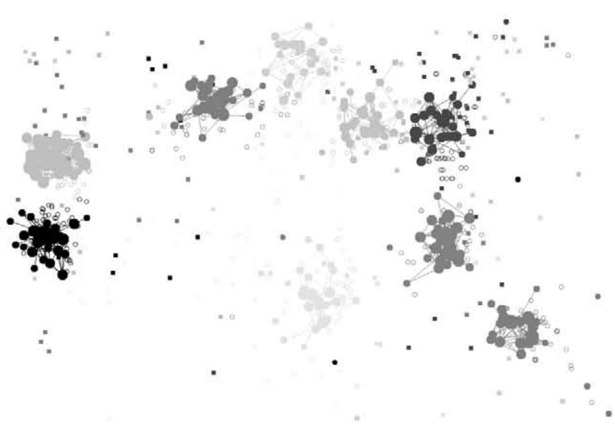


Figure 10: Layout of Shuffling ODC at the end of the simulation (Basic Scale-Free network, generation time 500 rounds)

2) Run-up and Fall-off

The second experiment examines the behavior of the algorithms when the generation of load stops after 200 rounds. We concentrate on the differences here.

In the shape of the clustered node curve in Figure 11, it can be observed that as the demand for clusters gets weaker, in general, both Shuffling ODC and Spyglass tend to destruct some of the existing groups, and not rebuild them. This is an interesting emergent property: clusters may get weakened when no more demand for them is present (but demand for some other cluster colors is available).

The processed job curve in Figure 12 emphasizes the practical difference between the processing speed of the algorithms. The rapid drop after 200 rounds illustrates the presence of very successful workload generator-attached nodes: they process and/or share all incoming jobs immediately, so, as the workload stops arriving, they finish their operation within a few rounds. The tail shapes of the curves show that ODC needs more time to finish the task than the two optimized algorithms.

The number of shared jobs in Figure 13 shows that after step 200, other clusters find and incorporate the former members of the successful clusters (that run out of jobs as the generation is ceased).

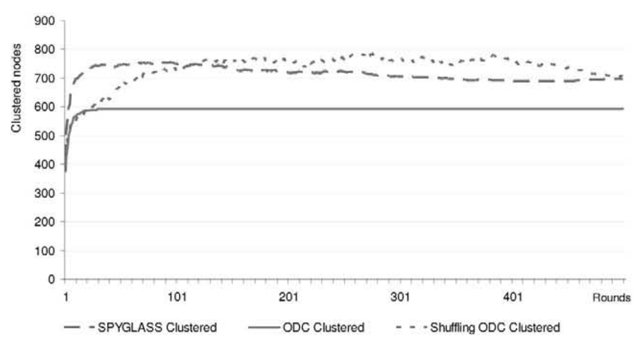


Figure 11: Number of clustered nodes (Basic Scale-Free network, generation time 200 rounds)

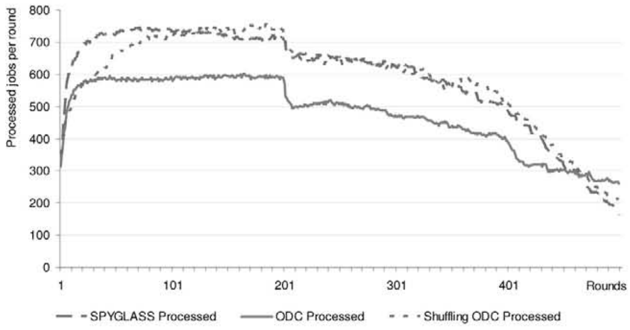


Figure 12: Number of processed jobs per round (Basic Scale-Free network, generation time 200 rounds)

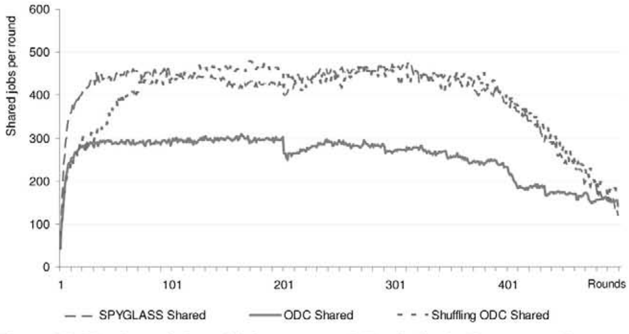


Figure 13: Number of shared jobs per round (Basic Scale-Free network, generation time 200 rounds)

Figure 14 shows the number of overloaded nodes, and Figure 15 the number of unprocessed jobs. Note that Spyglass produces less overloaded nodes, but with less clustered nodes and the same unprocessed job curve, suggesting that the overloaded nodes in Spyglass have higher amounts of overload than on Shuffling ODC. The original ODC, while following the same overload shape, leaves a significantly higher amount of unprocessed jobs. The two improved algorithms have only a few unprocessed jobs left at the end of the 500th round, compared to the total load, while ODC has significantly more left, and shows much slower processing characteristics.
The messaging overhead, shown in Figure 16, suggests that the communication follows the demand, and when the demand for clustering decreases, the message count does, too.

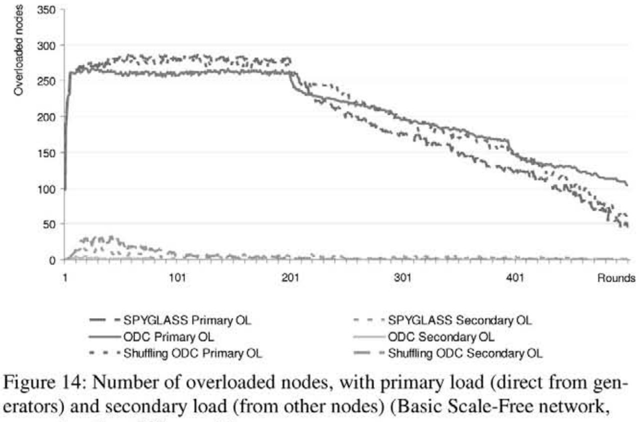


Figure 14: Number of overloaded nodes, with primary load (direct from generators) and secondary load (from other nodes) (Basic Scale-Free network, generation time 200 rounds)

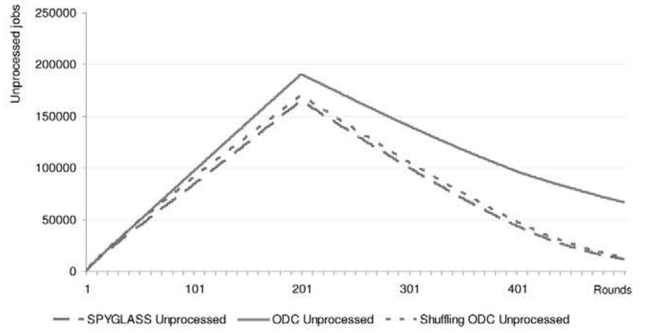


Figure 15: Number of unprocessed jobs per round (Basic Scale-Free network, generation time 200 rounds)

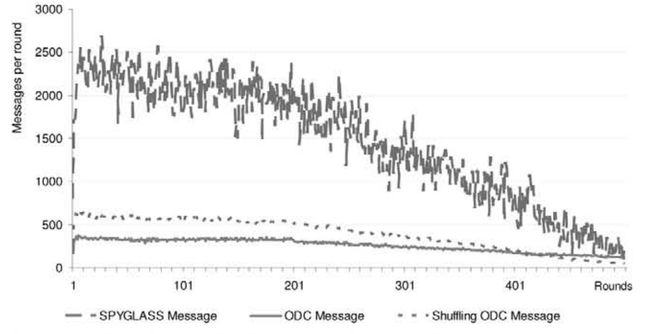


Figure 16: Messaging overhead per round (Basic Scale-Free network, generation time 200 rounds)

B. Algorithm Scaling

The scalability of the algorithm was demonstrated with experiments on the Large Scale-Free Network.

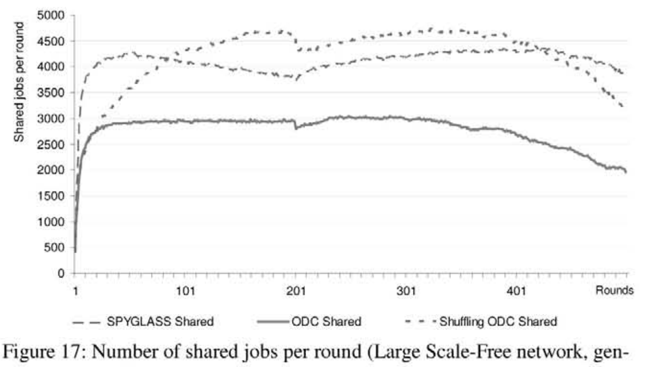


Figure 17: Number of shared jobs per round (Large Scale-Free network, generation time 200 rounds)

Experiments resulted in the same curve shapes as observed on the small network. No significant difference was detected. At some points the details of the mechanisms are clearer on the large network. As two examples, let us have a look at Figure 17 and Figure 18, showing the shared and processed job curves for the 200 rounds long workload generation scenario. The change after ceasing job injection can be clearly seen: after a short reaction period, the amount of job sharing reaches the original level. Shuffling ODC here slightly outperforms Spyglass, and both improved algorithms outperform ODC by at least 20-30% in the number of processed jobs, and by 35-50% in means of workload sharing. The relationship between the communication overheads of the algorithms does not change (see Figure 19).

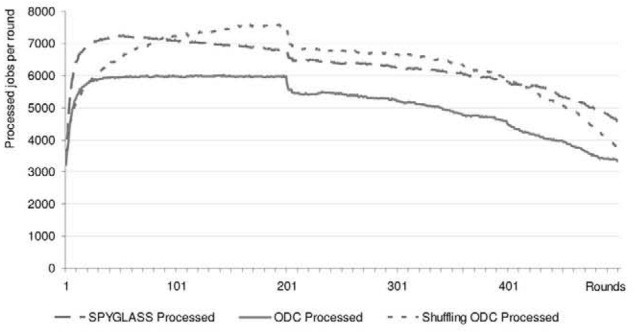


Figure 18: Number of processed jobs per round (Large Scale-Free network, generation time 200 rounds)

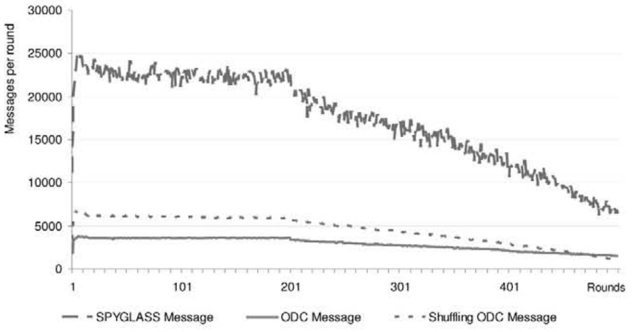


Figure 19: Messaging overhead per round (Large Scale-Free network, generation time 200 rounds)

C. *Effect of the Physical Topology*

The next set of experiments tackles with the question how Spyglass's and Shuffling ODC's somewhat weakened locality principle affect the real efficiency of workload sharing. One may intuitively assume that these algorithms cause workload sharing between physically more distant nodes than ODC does, and this difference may influence their real-life performance. Another question here is if the use of more realistic, distance-dependent sharing overheads affect the curve shapes discussed previously. We executed the next set of simulations with a setting where the transfer time of the jobs was visible and proportional to the physical distance between the nodes.

Figure 20 shows the average workload sharing distance, i.e. the average physical distance to the receiver node, throughout the simulation. Clearly, the average distance in Spyglass and Shuffling ODC is higher than in ODC, i.e. sharing happens between physically more distant nodes, but the difference never exceeds the +10% limit.

Table 1 summarizes the difference measured in sharing distance values and average cluster sizes of the two algorithms compared to the values produced by the basic ODC. The increment in the sharing distance is significantly lower (+7 and +8%) than the increment in the cluster size (+19 and +18%). In other words, the algorithms do not violate the locality principle by spreading unnecessarily over the network. The workload remains physically more concentrated than the expansion of the cluster would suggest.

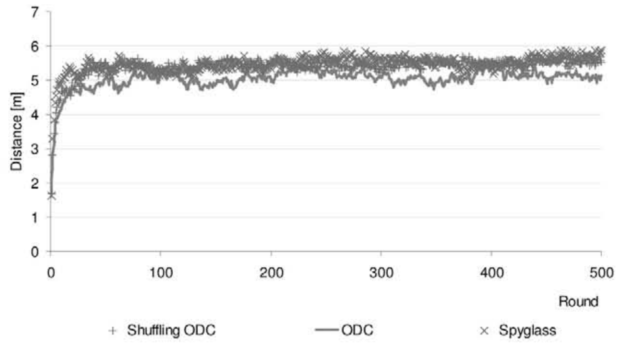


Figure 20: Average workload sharing share distance per round with physical distance dependent sharing overhead. (Basic Scale-Free network, generation time 200 rounds)

	Shuffling ODC	Spyglass
Avg. share distance	107% of ODC	108% of ODC
Avg. cluster size	119% of ODC	118% of ODC

Table 1: Average sharing distances and cluster sizes with significant, distance proportional workload sharing delay, relatively to ODC

The simulations with visible, distance-proportional sharing overheads showed no significant change in the trends. The higher overhead introduced some delay, but otherwise, lines remained the same. Figure 21 and Figure 22 show two of the charts. The sharing overhead, i.e. the time the job transfer takes, brings a certain amount of delay into the system. The receiver node receives the job later, senses the overload later, and initiates self-organization later.

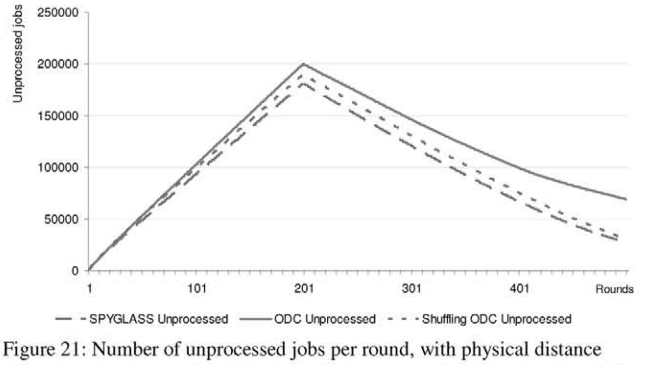


Figure 21: Number of unprocessed jobs per round, with physical distance dependent sharing overhead (Basic Scale-Free network, generation time 200 rounds)

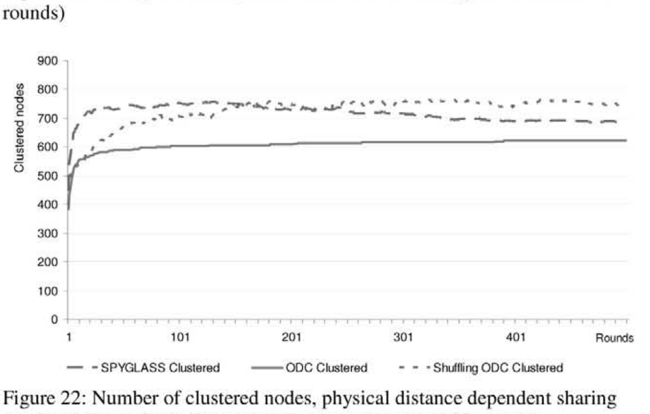


Figure 22: Number of clustered nodes, physical distance dependent sharing overhead (Basic Scale-Free network, generation time 200 rounds)

Simulations were, additionally, also repeated with a sharing delay model that is proportional to the square of the physical distance between the nodes (as often the case in wireless networks is). The trends were also confirmed here. Table 2 summarizes the increments observed. The sharing increment is still significantly smaller than the increment in the cluster size, hence, the locality principle is not violated. In means of locality preservation, Shuffling ODC outperformed Spyglass: it produced both larger clusters and shorter sharing distances. The explanation is that Spyglass tends to find more remote matches initially, hence it, at once, brings a delay into the system. Not only the actual job sharing but also its further effects are delayed, e.g. the secondary overloads and the triggered further clustering steps. On the other hand, depending on the situation, transferring a job to a more distant node may also have its benefits, for example, it may, so to day, 'guarantee' that the receiver is not already overloaded.

	Shuffling ODC	**Spyglass**
Avg. share distance	108% of ODC	109.5% of ODC
Avg. cluster size	118.5% of ODC	116% of ODC

Table 2: Average sharing distances and cluster sizes with significant, distance square proportional workload sharing delay, relatively to ODC

D. *Effect of the Problem Complexity*

The next set of experiments investigate how the complexity, i.e. hardness, of the problem influences the performance of the algorithms. We consider the problem to be more complex when it is harder to find a match, for example more colors are present or the system is less interconnected. We repeated the previous experiments by increasing the number of colors from 10 to 50.

Figure 23 shows the processed job curve of the 200-rounds workload generation setting on the basic scale-free network with 50 different colors. Note that the average rate of workload generation was not changed, so, the same amount of workload now needs to be processed by 5 times less nodes. Spyglass, here, outperforms the other two in speed, it sometimes processed even 50% more jobs than ODC. Shuffling ODC, even if not in such a high amount as Spyglass, clearly outperforms ODC, too. Figure 24 shows the cost of the better performance. The messaging overhead generated by Spyglass is immense, it is by two magnitudes over ODC. Shuffling ODC produces only a linear communication overhead.

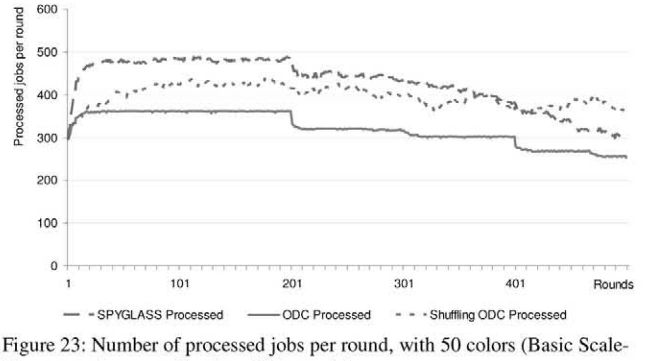


Figure 23: Number of processed jobs per round, with 50 colors (Basic Scale-Free network, generation time 200 rounds)

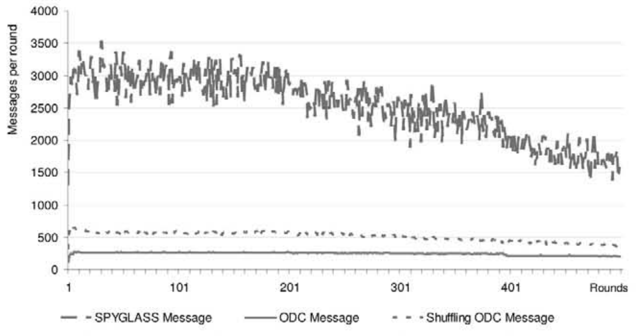


Figure 24: Messaging overhead per round, with 50 colors (Basic Scale-Free network, generation time 200 rounds)

E. *Effect of Network Abstraction*

The last set of experiments investigate how the network abstraction influences the results. We repeated the experiments with the same node and link counts, but with random networks. Results confirmed the general trends, however, sometimes with an even less pronounced performance difference between Spyglass and Shuffling ODC. Figure 25 shows an example, the unprocessed job curve which is basically the same as what we saw with the scale-free setting.

With random graph based network abstraction, the differences in the nature of cluster formation are more visible (the degree distribution of the nodes is homogeneous, i.e. random picks work better than in the scale-free case). Figure 26 visualizes the number of clustered nodes along time. The fallback effect, discussed earlier, is less visible here.

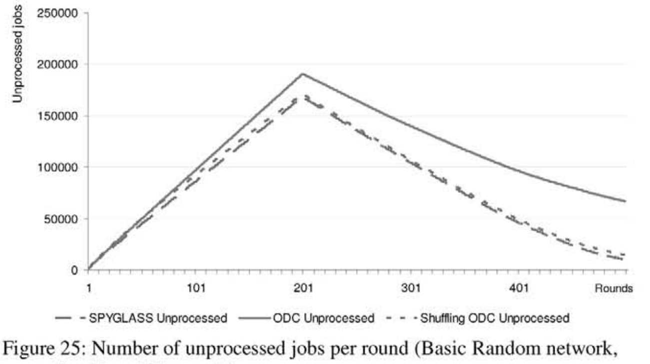


Figure 25: Number of unprocessed jobs per round (Basic Random network, generation time 200 rounds)

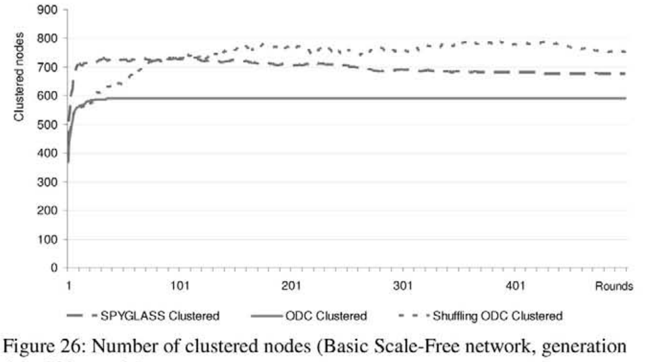


Figure 26: Number of clustered nodes (Basic Scale-Free network, generation time 200 rounds)

VI. CONCLUSIONS AND FUTURE DIRECTIONS

We discussed two directions to improve the On Demand Clustering for self-organizing networks. One of the directions, Spyglass, with excessive search in a wider area, produces larger clusters rapidly, but with a significant messaging overhead. The other direction, Shuffling ODC, with rewiring also in no-match cases, also results in large clusters, more clustered nodes, faster and more reliable job processing, but without an excessive increment in the number of messages.

Experiments confirmed that our two algorithms, Spyglass and Shuffling ODC, provide improved clustering and load-balancing capabilities compared to the baseline ODC algorithm. We observed improvement in the number of clustered nodes, clustering speed, the shape of the job processing and unprocessed job curves, and in primary and secondary overload characteristics. Visual evaluation confirmed the balanced operation of the algorithms. Scalability was also demonstrated.

Spyglass and ODC, while similar in some results, have different strong and weak points. The biggest advantage of Spyglass is its fast reaction time: clusters are extensively and very successfully formed from the arrival of the first trigger. The massive messaging overhead measured during the experiments may be reduced with caching, but not indefinitely. Shuffling ODC on the long term provides the same advantages as Spyglass, without the massive messaging overhead, but with slightly slower initial reaction time. The unique advantage of Shuffling ODC is its emergent self-shuffling effect which enables overloaded initiator nodes to overcome the boundaries of using only those matches that were nearby already in the initial topology.

Future directions include: investigating the effects of network density, experimentation with underlying physical topology and routing, and design of new hybrids from the ideas behind Spyglass and Shuffling ODC.

VII. ACKNOWLEDGEMENT

The support of the Hungarian Government through the TÁMOP-4.2.1/B-09/1/KMR-2010-0002 project at the Budapest University of Technology and Economics is acknowledged.

VIII. REFERENCES

[1] E. Dahlman, H. Ekström, A. Furuskär, Y. Jading, J. Karlsson, M. Lundevall, and S. Parkvall. 2006. "The 3G Long-Term Evolution - Radio Interface Concepts and Performance Evaluation," In *IEEE Vehicular Technology Conference (VTC) 2006 Spring*, Melbourne, Australia

[2] D. J. Cook, J. C. Augusto, V. R. Jakkula. 2009. "Ambient Intelligence: Technologies, Applications, and Opportunities." in *Journal of Pervasive and Mobile Computing*, 5(4): 277-298.

[3] J. C. Augusto. 2007. "Ambient Intelligence: the Confluence of Ubiquitous / Pervasive Computing and Artificial Intelligence." in *Intelligent Computing Everywhere*, Springer Verlag, pp. 213-234.

[4] G. Di Caro, F. Ducatelle, L. M. Gambardella. 2005. "Swarm Intelligence For Routing In Mobile Ad Hoc Networks." in *Proceedings of the 2005 IEEE Swarm Intelligence Symposium (SIS)*, pp. 76-83.

[5] A. Montresor, H. Meling, Ö. Babaoglu. 2002. "Messor: Load-Balancing through a Swarm of Autonomous Agents." in *Proceedings of 1st Workshop on Agents and Peer-to-Peer Computing (AP2PC 2002)*, Springer Verlag, pp. 125-137.

[6] A. Montresor, H. Meling, Ö. Babaoglu. 2003. "Toward Self-Organizing, Self-Repairing and Resilient Distributed Systems." in *Number 2584 in Lecture Notes in Computer Science*, Springer Verlag, pp. 119-124.

[7] IST CASCADAS Deliverable D3.2: "Report on rule-based modules for unit differentiation using cross-inhibition and/or resource competition." IST CASCADAS Project, 2007.

[8] S. Jain, R. Mahajan, B. Niswonger. 2000. "Self-Organizing Overlays." Technical report, University of Washington, DC, USA.

[9] S. Apel, K. Böhm. 2005. "Self-Organization in Overlay Networks." in *Proceedings of CAISE'05 Workshops (Workshop on Adaptive and Self-Managing Enterprise Applications)*, Vol. 2, pp. 139-153.

[10] I. Clarke, O. Sandberg, B. Wiley, T. W. Hong. 2000. "Freenet: A Distributed Anonymous Information Storage and Retrieval System." in *Proceedings of Workshop on Design Issues in Autonymity and Unobservability, Berkeley*, CA, USA.

[11] P. Gburzynski, B. Kaminska, W. Olesinski. 2007. "A Tiny and Efficient Wireless Ad-hoc Protocol for Low-cost Sensor Networks." in *Proceedings of the conference on Design, Automation and Test in Europe (DATE'07)*, pp. 1557-1562.

[12] R. Krishnan. 2004. "Efficient Self-Organization of Large Wireless Sensor Networks". Ph.D. Dissertation, Boston University, College of Engineering.

[13] A. A. Abbasi, M. Younis. 2007. "A Survey on Clustering Algorithms for Wireless Sensor Networks". Elsevier Computer Communications, vol. 30, pp. 2826–2841

[14] K. Farkas, F. Maurer, L. Ruf, B. Plattner. 2006. "Dominating Set Based Support for Distributed Services in Mobile Ad Hoc Networks". In Proceedings of the 10th IEEE/IFIP Network Operations and Management Symposium (NOMS 2006); Vancouver, Canada; April 2006

[15] M. Legeny. 2010. "Distributed and hybrid algorithms in Self-Organizing Networks" (In Hungarian). Master's thesis, Department of Telecommunications, Budapest University of Technology and Economics (unpublished).

[16] M. Legeny, B. K. Benko. 2010. "Design of Novel Self-Organization Algorithms through Simulation." in *Proceedings of EUROSIS ISC'2010*, pp. 10-15.

[17] F. Saffre, R. Tateson, J. Halloy, M. Shackleton, J.L. Deneubourg. 2008. "Aggregation Dynamics in Overlay Networks and Their Implications for Self-Organized Distributed Applications". in *The Computer Journal*.

[18] E. Di Nitto, D. J. Dubois, R. Mirandola, F. Saffre, R. Tateson. 2008. "Self-Aggregation Techniques for Load Balancing in Distributed Systems." in *Proceedings of SASO 2008*, pp 489-490.

[19] IST CASCADAS Deliverable D3.1: “Aggregation Algorithms, Overlay Dynamics and Implications for Self-Organised Distributed Systems.” IST CASCADAS Project, 2006.
[20] B. K. Benko, M. Legeny. 2011. “Topology Cache aided Self-Organization for Ad-hoc Mobile Networks” in Cyber Journals - Journal of Selected Areas in Telecommunications. 2011 (3), pp. 56-63
[21] A. L. Barabási, R. Albert. 1999. “Emergence of scaling in random networks”. Science 286 (5439): 509–512. doi:10.1126/science.286.5439.509
[22] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, E. Upfal. 2000. “Stochastic models for the web graph”. Proceedings of the *41st Annual Symposium on Foundations of Computer Science (FOCS)*. Redondo Beach, CA: IEEE CS Press. pp. 57–65.

CALLS FOR PAPERS

Call for Papers

Prospective authors are invited to submit original research papers for publication in the upcoming issues of our Infocommunications Journal.

Topics of interests include the following areas:

- Data and network security
- Digital broadcasting
- Infocommunication services
- Internet technologies and applications
- Media informatics
- Multimedia systems
- Optical communications
- Society-related issues
- Space communications
- Telecommunication software
- Telecom. economy and regulation
- Testbeds and research infrastructures
- Wireless and mobile communications

Theoretical and experimentation research results achieved within the framework of European ICT projects are particularly welcome. From time to time we publish special issuesand feature topics so please follow the announcements. Propo-sals for new special issues and feature topics are welcome.

Our journal is currently published quarterly and the editors try to keep the review and decision process as short as possible to ensure a timely publication of the paper , if accepted. As for manuscript preparation and submission, please follow the guidelines published on our website
http://www.infocommunications.hu/for_our_authors.

Authors are requested to send their manuscripts via electronic mail (preferably) or on a CD by regular mail to the Editor -in-Chief:

Csaba A. Szabo
Dept. of Telecommunications,
Budapest University of Technology and Economics
2 Magyar Tudosok krt. Budapest 1117 Hungary
E-mail: szabo@hit.bme.hu

Call for Proposals of Special Issues

Infocommunications Journal welcomes proposals for Special Issues – collections of papers dedicated to a particular topic of interest for the readers of the journal.

A Special Issue can be based on a recent high quality workshop or conference or papers can be collected from open call. Invited papers can be part of the special issue as well.

A Special Issue can fill in a whole issue, in which case the number of papers is expected to be 8 to 10, or it can be a Mini – Special Issue. In the latter case, at least 3, preferably 4 papers are required.

Proposals for special issues should include:

– contact information (name, e-mail, title, affiliation and address);
– resume(s) of the proposer(s), with a representative list of recent publications and related experience (Editorial Board memberships, Guest Editorships, or roles in relevant conferences’ program committees);
– the proposed title for the special issue;
– the way the special issue will be compiled (contributions solicited from a technical event, or to be collected from call for this special issue;
– intent to include invited papers should be also indicated, if possible with the names of professionals who are planned to be invited;
– scope and motivation and description of the special issue;
– guest editors (if different from the proposers) with detailed contact information and resumes.

Proposals should be sent to the Editor-in-Chief:

Csaba A. Szabo
Dept. of Telecommunications,
Budapest University of Technology and Economics
2 Magyar Tudosok krt. Budapest 1117 Hungary
E-mail: szabo@hit.bme.hu

Meta-level Performance Management of Simulation of Organizational Information Systems: The Problem Context State Approach

László Muka, Gábor Lencse

***Abstract*—Simulation has become a frequently used tool for the analysis of ICT and BP systems and for fitting the features of these systems with each other and with the goals of the enterprise. For example, the change management of ERP (Enterprise Resource Planning) systems is a significant generator of the need for the common analysis of ICT/BP systems and the use of simulation may play crucial role in their analysis. The paper formulates the problem context state approach to the meta-level performance management of simulation in the form of efficiency management principles. The formulation is based on the investigation of the features of the dynamic behavior of problem contexts – using the 4-state and 2-state models of problem context types – for the common modeling and simulation of organizational ICT/BP systems. The process of the occurrence and elimination of the methodological gap is explained too.**

***Index Terms*—efficiency of simulation, problem context state model, efficiency principles, ICT and BP systems, efficiency management**

I. INTRODUCTION

SIMULATION has been accepted as an appropriate tool for the analysis of ICT and BP systems and for fitting the features of these systems with each other and with the goals of the business.

Examining in an organizational environment, the *simulation process* is a *participative* and *collaborative* process with many *participants* [10]. Sierhuis and Selvin define the *simulation process* as a *holon*[1] in terms of Soft Systems Methodology (SSM, [2]) [11]. As the system approach to the simulation, the *simulation methodology* may be defined as a *structured set of methods* applied by a *HAS* (Human Activity System, [2]) performing the process of simulation. In an organizational environment, the process of simulation may also be treated as a *project process* with predefined goals aimed to be reached within time and cost limits with prescribed quality requirements. The *phase* of the simulation process is determined by the method of the simulation methodology being executed. *Discrete-Event Simulation* (DES) is a frequently used method for the analysis of the ICT and BP systems [5].

Simulation projects aimed at supporting the analysis and design of the *dynamic behavior* of organizational ICT (Information and Communication Technology) systems and BP (Business Process) systems are usually separate projects but these systems may have significant influence on each other. (ICT and BP systems of an organization may also be referred to as an *Organizational Information System* or OIS). Thus, the common analysis of these systems may have *significant benefits* – this is why there is an increasing need for the *common modeling and simulation* occurs. In the common analysis, we need models of ICT and BP systems that can *interact* with each other just as these systems interact with each other in the real world.

Depending on the task, *distinct* or *integrated* ICT and BP models may be used for the common analysis.

- Examples of tasks for which distinct models are appropriate: Support for the BSM (Business Service Management) method [12] in defining the performance and availability of ICT and the features of BPs operating the ICT systems (for example, in an ERP (Enterprise Resource Planning) system) according to the service-level requirements determined by the current and future BPs. Another example is to analyze the dynamic relationship between ICT and BP performance of system functions by using simulation in order to help BP and ICT designers and analysts [9].
- Examples for tasks with integrated ICT and BP model: It may be beneficial to integrate the model of the BP system into the model of the ICT system if, for example, the BP system is an intensive traffic source for the ICT system (for example, customer service offices in the ICT infrastructure). The BP model may integrate ICT model, for example, in the task of the optimization of the proportion of automatic (produced by some answering software) and operator performed activities of a help desk system.

In the common simulation analysis of ICT and connected

Manuscript received March 29, 2011.
László Muka is with the Elassys Consulting Limited, Budapest, Hungary (e-mail: muka.laszlo@elassys.hu).
Gábor Lencse is with the Department of Telecommunications, Széchenyi István University, Győr, Hungary (e-mail: lencse@sze.hu).
[1] An SSM-holon is a whole - with emergent properties - from some point of view of abstraction.

BP systems, we may easily be faced with the case of large and complex systems where the necessary computing capacity may reach or even exceed the reasonably available. The increase of the *efficiency of simulation* may be an answer to this problem.

The efficiency of simulation is influenced by many factors including *methodological fact*ors too (for example, the occurrence of *unstructured problems* and the problem of *efficient applicability of methods*).

The aim of this paper is to address the problem of the increase of the efficiency of simulation on *meta-level*, by developing *the principles of managing the efficiency of simulation* on the level of methods and problem contexts.

The *new results* in the paper can be summarized as follows. On the basis of 4-state and 2-state models of problem context types, the *features of the dynamic behavior of problem contexts* are investigated for the case of common modeling and simulation of organizational ICT/BP systems (OIS). Using the set of efficiency principles – referring to the re-definition of Checkland's systems performance criteria and to the criterion of gap-efficiency with an explanation of the process of the occurrence and elimination of the methodological gap – *efficiency management principles* are formulated serving for the managing of the efficiency of simulation of OIS on meta-level.

The new approach introduced in the paper has significant advantages comparing with other approaches. The classic simulation methodologies (for example, those described in [15]) are efficient only for the hard-systems type of problem contexts. Other context based systems approaches ([13], [14]) have static approach and do not take into account the occurrence of context-type changes in the execution of a simulation task. Furthermore, they do not use an explicit and general approach to efficiency such as the rules for the management of efficiency formulated in this paper.

The paper is organized in the following way. Section 2 describes the problem context state models and analyses the simulation of ICT and BP systems from the point of view of problem contexts. Section 3 introduces the efficiency criteria and defines the efficiency principles. Section 4 formulates the principles for the managing of efficiency of simulation on meta-level. Section 5 examines the work of the efficiency management principles. Section 6 refers to the current and potential applications. Section 7 summarizes the work.

II. The Problem Context State Models

In this section, the meta-level analysis of efficiency of simulation, the problem context state models[2] – the environment for the functioning of the simulation process – will be defined and explored, then the features of the process of simulation will be analyzed using the defined models.

A. *The Jackson-Keys Classification of Problem Contexts*

Jackson and Keys [4] defines the classes of problem contexts according to two dimensions: the *simple-complex* (or simple-systemic) dimension describes the *system feature* and the *unitary-pluralist*[3] dimension characterizes the *actors* (decision makers) of the problem context. According to this classification, the *types of problem contexts* may be: *simple-unitary, simple-pluralist, complex-unitary and complex-pluralist*.

B. *The 4-state model of the dynamic problem contexts*

The *4-state-type model of the problem contexts* shown in Figure 1 is created in the way of utilizing the *subset* of classes of Jackson and Keys in a different way, using them in a dynamic manner.

The extension-restriction relationship of the *generality* of the problem context types is demonstrated with dashed lines in Figure 1: the simple-unitary context type is a special case of the complex-unitary and of the simple-pluralist context types and all these three context types are the special cases of the complex-pluralist context type. The relationship in the level of determination of problem contexts is also shown: the lighter the shade of color a problem context in Figure 1 has the more determined the context is.

In Figure 1, straight lines with arrows on both ends show the transitions between different types of problem contexts which are reached by changes of dimensions (A-transitions), the straight lines with single arrow show the possibility of the occurrence of a new context (B-transitions), curved lines with arrows demonstrate the transitions between similar contexts (inside of a problem context type) (C-transitions).

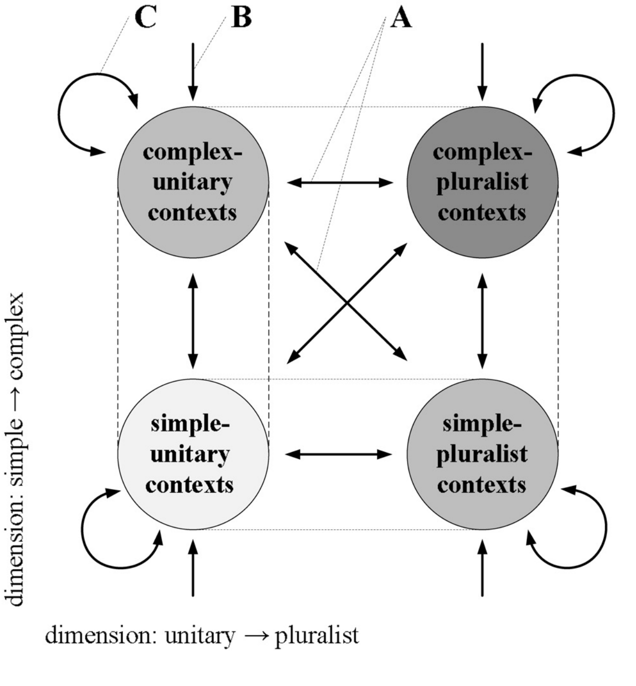


Fig. 1. Problem contexts and transitions

[2] The reader is referred to [17] too, regarding context state models in. The present paper analyses their operation in the way necessary for the formulation of the efficiency management principles.

[3] The "coercive" category of the actors has no significance in our analysis.

C. The 4-state model of the dynamic simulation problem context

The *dynamic simulation problem context (DSPC)* – which is the problem context environment of the functioning of the simulation process [7, 8] – *may be defined as the sequence of problem contexts and transitions of contexts* that occur in the process of simulation.

The following *propositions* about the features of DSPC are examined:

- The dynamic simulation problem context may contain all the problem context types.
- In the dynamic simulation problem context any type of problem context may occur in any phase of the simulation process.
- The occurrence of the problem contexts may happen independently from the process of simulation too.

The above propositions are examined for the case of the common modeling and simulation of ICT and BP systems.

Defining the simple-complex dimension

Vemuri (Vemuri, in [4]) defines the system feature of complex systems: *partially observable, subject to behavioral influences*[4], *probabilistic and evolving*. According to Jackson and Keys, [4] complex systems – in addition to Vemuri's criteria – have a *large number of elements* (that are highly interrelated) and the *evolving* feature may be replaced by the features that complex systems are *open* and they *have purposeful parts (subsystems)* as well as the selection of *boundaries* of a system may have influence on its complexity.

Taking into account the previous points, the following features to characterize complex systems are defined: *(1) partial observability,* (*2) wide boundaries and high resolution*, both in *structures* and in *time* (which results in large number of elements, relations and events), *(3) openness* and *(4) purposeful subsystems (with behavioral attributes)*. (The *probabilistic feature* is also taken into account as it is explained later.) Thus, the *system features* of the *simple-complex dimension* of the *simulation problem contexts* for the *modeling and simulation of ICT and BP systems* are defined as follows:

1. Systems of interest are often only *partially observable*: this may be caused by data availability problems (for example: data are not collected or cannot be collected because of technical reasons, cost, time and resource limits; collected and available data are enough only for partial description of the system; data sources may be located in other systems and are not available for the modeling purposes, etc.).
2. The wide *boundaries* (including both structural and time limits) of the models of systems of interest and their high *resolution* too (including both structural and time boundaries and resolution) may make the problem complex: the wider the boundary is set the more complex the system may become and the same is true for the resolution.
3. The complexity is increased by taking into account the influences among systems (subsystems) – including, of course, the influences between ICT and related BP systems. Interacting systems are *open* to influences between each other. The more detailed the model of *interactions* is the more complex the system may become.
4. BP systems may have *active*, *purposeful parts*: their *behavior* cannot be predicted exactly (for example people in the system may act in opposition to simulation project goals).

A *simulation problem context* is *simple* if the systems of interest are observable, the boundaries and the resolution of modeling of the systems are set at a necessary but low level, the influences among the systems (subsystems) of interest are limited in the model (systems are reasonably closed) and the purposeful parts of processes are passive. Any of the above listed conditions may make the *simulation problem context complex*: if the systems of interest are not observable (partially observable), the boundaries and the resolution of modeling of the systems are set at a too wide/high level for simulation, the influences among the systems (subsystems) of interest are not limited enough in the model (systems are open) and the purposeful parts of processes are active.

Remarks:

- The *probabilistic* feature of the behavior of the analyzed systems is the basic object of the simulation investigation.
- The behavioral influences of systems of interests are taken into account in the examination of active, purposeful features of the BP systems.
- In determination of system features the *emergent properties*[5] has to be taken into account. (For example, on the one hand, the boundary for modeling should be set wide enough and the resolution of models high enough to examine the emergent properties and to get the necessary answer and on the other hand, the boundary should be narrow enough and the resolution low enough to be able to simulate the system.)

Defining the unitary-pluralist dimension

The decision makers of the simulation problem contexts in an organization environment are determined by the simulation project. The problem context is *unitary* if the *set of decision makers* have a *common set of goals* (agree) and *pluralist* (disagree) if they do not. *Problem solvers* (as participants of the problem context) may also become decision makers in the simulation process.

D. Defining the 2-state model of DSPC

In the following, according to the 4-state model of DSPC, the *2-state-type of model (or hard-soft model)* of DSPC will be defined (Figure 2).

[4] Political, cultural, ethical and other similar type of influences should be taken into account in the analysis of these systems.

[5] The *emergent property* may be, for example, an analysed functional capability of the system of interest. This capability may disappear or occur in correlation with the selected system features of a problem context.

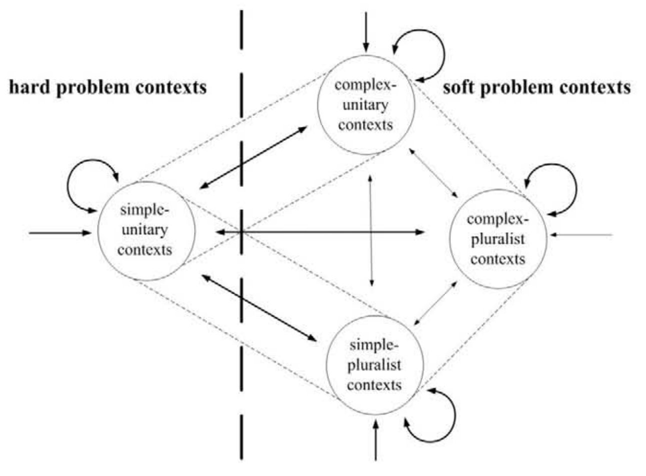


Fig. 2. The hard-soft model of problem contexts

Problem contexts could also be classified from the point of view of appropriate approaches. *Hard-systems approaches*[6] are suitable for looking for solutions to well-defined problem situations, starting from clearly defined objectives. Thus, *simple-unitary problem contexts* with well defined system features and with a common set of goals of decision makers are *hard problem contexts*. *Soft-systems approaches* are to cope with ill-defined, unstructured problem situations, in which objectives are themselves problematical. The *complex-pluralist problem contexts* with undefined system feature and with pluralist set of decision makers are *soft problem contexts*. The *complex-unitary* problem context may have an active purposeful part and thus, it will also require a soft-systems method to deal with the situation therefore this problem context can be classified as a *soft problem context*. The *simple-pluralist* problem context requires a soft approach to deal with the pluralist set of decision makers that is it is a *soft problem context* too.

(Remark: Soft-systems approaches may be appropriate both hard- and soft problem contexts but hard-systems approaches are suitable only for hard problem contexts.)

In the following, the 2-state and the 4-state models of DSPC are applied in the argumentations in a mixed way.

E. *Analyzing the DSPC transitions*

Now, in order to reveal the features of DSPC, the transitions of types of problem contexts will be examined. (The transitions are investigated as they are shown in Figure 1.)

A-transitions:

A change of the simulation process phase may generate transition: for example, in the simulation process, after the phase of the analysis of results a need occurs to change the resolution of the simulation model. (The simulation process phases are described, for example, in [8].) The system feature of the problem context may remain simple or change for complex and there can be agreement or disagreement about the measure of the resolution: these are transitions from simple-unitary to complex-unitary, simple-pluralist and complex-pluralist contexts. When entering a new phase, a new problem situation is created thus *any type* of problem contexts may be identified. In general, *any phase* of the simulation process may lead to a *pluralist* problem context: different opinions may occur concerning the goal setting for the phase and concerning the further use of the results of the phase.

Transition may also be generated by *transformation decision*: for example after entering a simulation process phase it is found that the problem context is one of the complex-unitary, simple-pluralist or complex-pluralist contexts. There should be made transformation decision for the transition into the simple-unitary context, because the simulation methodology is appropriate only for the simple-unitary context [3, 4]. This may be done by the way of agreed changes of system features (if it is necessary) and by the way of finding the consensus about the set of goals. The system features for modeling may be changed into the *simple* direction by decisions (and actions taken according to the decisions) about data availability (1), by decisions about setting up the boundaries and the resolution of systems of interest (2) and decisions about the modeling of interactions among systems (3). The passivity of the purposeful part of the system (4) may be reached typically by using some consensus building method.

In the decision process, transitions between complex-unitary, simple-pluralist and complex-pluralist contexts may also occur: for example, the purposeful part of the system feature may change between active and passive (all other system features show a simple system) and the set of the opinion of decision makers may change between agree and disagree.

B-transitions:

These transitions are of "*insertion*" type: a new problem context may be generated *independently* from the earlier problem contexts. For example, because of the influence of the changes in the organization (in the wider environment of the simulation process) new requirements may occur concerning the system feature of the problem context and a new set of the decision makers opinion may occur too. The starting problem contexts of the simulation process may also be of *any type* and the initial problem analysis (*structuring)* may lead to a pluralist set of opinions about the goals even if there was an agreement about the initiation of the simulation project.

C-transitions:

These transitions show the change of the problem context features without changing the problem context type. If the system features show a complex (simple) system it remains complex (simple) after this type of transition only with other set of problem context features and if the set of decision makers is unitary it remains unitary or if the set of decision makers is pluralist it remains pluralist only with other set of disagreeing decision makers. These transitions may occur, for example, in the decisions process but this type of transitions

[6] A more detailed description of hard- and soft-systems approaches can be read, for example, in [1].

may take place between simple-unitary contexts in the execution of the phases of simulation too.

Remarks to problem context transitions:

- The unitary-pluralist dimension of the problem contexts may also be changed if there is a *change in the set of actors* (for example, there are two collaborating teams of actors in the process of simulation).
- The interaction of actors with purposeful parts of the systems of interest may also change the unitary-pluralist dimension of a problem context.
- It is important to notice, that the *worldview* (Weltanschauung) of actors may also influence the simulation problem context through the decisions made about the features of the simple-complex dimension [4].
- A pluralist set of opinions may also occur about the different ways of *implementation* of results (for example: who is responsible for what during the implementation).

III. INTRODUCING EFFICIENCY CRITERIA

In the following, the rather static approach of the Jackson-Keys method [4] will be *extended*: for the DSPC, efficiency criteria are introduced and developed and the efficiency principles are defined.

A. Efficiency in the Jackson-Keys method

According to Jackson and Keys, a method is appropriate for a problem context if it is selected to be the same type as the explored type of the problem context. Methods, which are suitable for complex-pluralist problem contexts, are potentially able to address problems in all other problem contexts but using methods for complex-pluralist problem contexts in other problem contexts may lead to *inefficiency.*

B. Checkland's systems performance criteria

If the system approach of simulation is used – the holon and HAS concepts of the simulation process –, it seems to be fruitful to apply the approach to *efficiency of activity systems* [2]. According to Checkland, the problem of *efficiency* is addressed together with the examination of questions of *efficacy* and *effectiveness*. According to Checkland, there is a *hierarchy-like relationship between* the three criteria efficiency, efficacy and effectiveness: the question of the adequacy for the longer term and for the wider environment is checked by the *effectiveness (criterion E3)*, the *efficacy (criterion E2)* investigates the question whether the solution will be suitable and work in all circumstances and the *efficiency (criterion E1)* examines the traditional question of efficiency (the question of direct efficiency) which can be measured by the proportion of the required outputs and the resources used to produce the outputs.

C. Defining the efficiency principles

Now, the *principles of efficiency* will be formulated applying Checkland's *criteria* for the process of simulation and introducing a new criterion, the criterion of *gap-efficiency*. Efficiency principles are defined for the relationship of fitting of methodology, method and problem context.

The *extension-restriction relationship* of the scopes of method types is the same as the extension-restriction relationship of the *generality* of the problem context types.

(E1) The *principle of methodological efficiency* can be defined as follows: for a methodology to be efficient the best fit with a specific problem context should be found. It means that the type of the selected method should be the same as the type of the problem context. Furthermore, it also means that the best fit with a specific problem context should be found within the set of methods of the same type of the methodology – if there are more methods of the same type in the methodology.

(E2) The *principle of hardening up and softening up (or the principle of methodological efficacy)* is the principle of dealing with a problem context that does not fit into one type of problem contexts in the sense that it has some aspects belonging to other problem context type. In other words, an aspect of a problem context, which has been revealed by applying the selected (according to the condition E1) method of the methodology, defines a problem context with a type different from the original one. In this case, in order to find the exact fit and to avoid inefficiency, the methodology should be *hardened up* or *softened up* by involving a method which is efficient for that different-type problem context.

(Eg) The *principle of the elimination of the methodological gap (or the principle of gap-efficiency*) is the principle for dealing with the problem of inefficiency that may be caused by soft-hard *problem context transitions*. These transitions are necessary and crucial because the traditional *simulation methodology,* as it was mentioned before*, is a method appropriate only for simple-unitary context* [3, 4]. A *methodological gap* ([7, 8]) may occur in the execution of the process of simulation if a soft-systems method and a hard-systems method of the methodology is applied for two sequencing problem contexts: the set of hard-level information for further processing by some hard-systems method is produced from the set of soft-level information by the way of using some soft-systems method and executing *ad-hoc, occasional condensing*[7] (Figure 3). For example, in the process of simulation, executed according to the framework for collaborative modeling and simulation [11], a methodological gap may occur when the team of hard modelers builds the simulation model using modeling data got from the team of soft modelers. The methodological gap may lead to *inefficiency* because of the fact that not the necessary condensing has been carried out which results in that not the required simulation model will be built.

The methodological gap may be *eliminated* by a *methodology constructed to connect the soft and hard levels* [6-8]. In order to tell whether a methodological gap has occurred or not, a new criterion the *criterion of gap-efficiency (Eg)* is introduced. The principle based on this criterion is *the*

[7] Checkland defined occasional condensing as the relationship between the soft-systems thinking and the hard systems thinking [1].

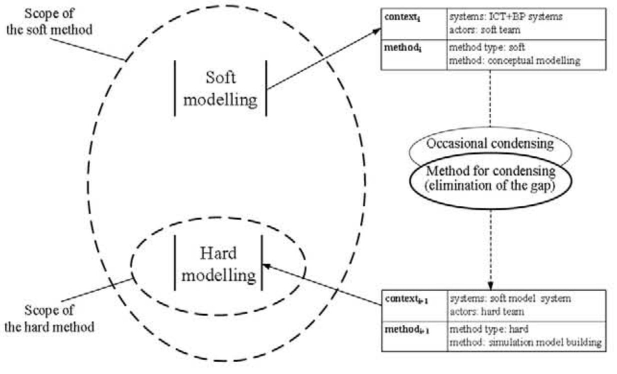


Fig. 3. The occurrence and elimination of a methodological gap

principle of elimination of the methodological gap (or principle of gap-efficiency).

(E3) The *principle of methodological effectiveness* expresses the efficiency requirement for the whole process of simulation resulting in the reduction of the number of problem contexts to deal with, reduction of the number of methodological cycles (number of iterations) in the process.

E1, E2 and *Eg* refer to the *step-by-step efficiency* (efficiency in problem context states and transitions) of application of methods while *E3* stands for *the efficiency on long-range* (to have less states and transitions in the process of application of the methodology).

IV. FORMULATING EFFICIENCY MANAGEMENT PRINCIPLES

Now, the efficiency management principles below will be formulated taking into account the requirements set by to the DSPC and by the efficiency principles described above.

Points a – c are the principles for collecting an efficient set of methods. For finding the "best fit", only the methods of the set of methods can be taken into account. For the methodology to be efficacious, the set of problem contexts should contain an efficacious method for any problem context of the DSPC. *Points d – i* refer to the structural and application (operation) type efficiency management principles.

a As a *soft-systems method*, we need a method appropriate for the complex-pluralist problem context. In theory, it may be used to any problem context but the *function of application* of the soft-systems method *is different in* the different phases of the simulation process (for example: scanning the relevant set of systems, scanning for simulation scenarios, etc.).

b The set of *hard-systems methods* should contain the methods of the traditional *simulation methodology* – according to the requirement of the methodological efficiency it should be a set of methods for the typical hard problem contexts in the simulation process that has been identified – and further methods required by the principle of the methodological efficacy and effectiveness (for example methods supporting goal setting, or methods supporting fast modeling).

c For the *elimination of the methodological gaps*, the set of methods is proposed to contain a *methodology connecting the soft-systems and hard-systems levels*. This is a methodology consisting of a soft-systems method and hard-systems methods defined on the basis of the identified soft- and hard-systems level contexts and the constraints for condensing (hard contexts may be for example contexts relating to the tasks of the analysis of time relations in ICT and BP systems).

d It should be taken into account that the traditional simulation methodology – which is based on a set of hard-systems methods – is a hard-systems approach. Hard-systems methods are appropriate only for hard problem contexts, therefore soft problem contexts of DSPC should be transformed into hard problem contexts.

e For managing efficiency, it is proposed to help to realize the change of problem contexts: hard-systems methods of the set of methods cannot see beyond the hard problem context, thus the need for insertion for a new problem context (generated by the observed systems) can be realized only by a soft-systems method.

f The hardening up and softening up of the methodology is proposed to be supported in any phase of the simulation process: inefficacy (and inefficiency) may occur in any phase of the simulation process when applying a method to a problem context which does not fit exactly into one problem context type. For managing similar situations, it is proposed to have soft-hard (hard-soft) method pairs for such problem contexts.

g It is proposed to support the method-selection decisions inside a context type: according to the principle of the methodological efficiency, it is necessary to find the best fit of a problem context and the method inside the given type too.

h To be efficient, the whole set of methods should be taken into account in the method-selection decisions: to reach the best fit of a problem context and a method, it may be necessary to choose a method other than the next method in the process of simulation.

i It is proposed to support to take into account – with an appropriate use of the necessary methods – the principle of the methodological effectiveness. It is necessary to find the best fit for a method and for the whole sequence of methods taking into account a wider systems environment and longer time frame generated by the observed systems and by the process of simulation itself.

V. EFFICIENCY MANAGEMENT EXAMPLE

Let us examine the work of the efficiency management principles on the example of a simulation task execution.

Let us use the *4-state model* of problem contexts and let X denote the set of problem contexts that have been identified in the process of simulation. In the 4-state model, the *types of problem contexts* are *simple-unitary (su), simple-pluralist (sp), complex-unitary (cu) and complex-pluralist (cp)*. Let M denote the set of methods which contains the methods of a *classic simulation methodology [8] (SM1(Goal-definition), SM2(Data-gathering), SM3(Modelling), SM4(Simulation), SM5(Evaluation), SM6(Implementation-support)), the SSM (Soft Systems Methodology [2]), and the MCMM (Modified Conceptual Modelling Methodology [8])*. Thus, the method set is $M = \{SM1; SM2; SM3; SM4; SM5; SM6; SSM; MCMM\} = \{1; \ldots 6; 7; 8\}$.

The efficiency management principles listed in the previous section will be referred as $p_a, p_b, \ldots, p_i$.

The transition between two subsequent problem context states of the process of the simulation task execution may be described by the expression:

$$x_{k,(type_x)}\left(m_{(type_m),(method)}\right) \xrightarrow[\ ]{\overset{(insertion)}{(transition\ type)}} x_{k+1,(type_x)}\left(m_{(type_m),method}\right),$$

where $x_{k,(type_x)}$ is k-th probem context in the problem context sequence ($k = 1,2,3, \ldots, |X|$, $\boldsymbol{type_x = \{su; cu; sp; cp\} = \{1; 2; 3; 4\}}$) and $\boldsymbol{m_{(type_m),(method)}}$ is the method assigned to the problem context ($\boldsymbol{type_m = \{su(SM1, \ldots, SM6, MCMM); cu; sp; cp(SSM, MCMM\} = \{1; 2; 3; 4\}}$, $\boldsymbol{method = method\ identifier\ in\ the\ set\ M = \{1; 2; \ldots 6; 7; 8\}}$) and the variables of the transition operator are: $\boldsymbol{transition\ type = \{A; C\}\ and\ insertion = \{B\}}$.

Now, let us examine the problem context sequence

$$x_{1,4} \overset{A}{\rightarrow} x_{2,1} \overset{C}{\rightarrow} x_{3,1} \overset{A}{\rightarrow} x_{4,4} \overset{A}{\rightarrow} x_{5,1} \overset{C}{\rightarrow} x_{6,1} \overset{C}{\rightarrow} x_{7,1} \overset{B}{\overset{A}{\rightarrow}}$$

$$\overset{B}{\overset{A}{\rightarrow}} x_{8,4} \overset{A}{\rightarrow} x_{9,1} \overset{C}{\rightarrow} x_{10,1} \overset{B}{\overset{A}{\rightarrow}} x_{11,4} \overset{C}{\rightarrow} x_{12,4} \overset{A}{\rightarrow} x_{13,1} \overset{C}{\rightarrow}$$

$$\overset{C}{\rightarrow} x_{14,1} \overset{C}{\rightarrow} x_{15,1} \overset{C}{\rightarrow} x_{16,1} \overset{()}{\rightarrow} \ldots \overset{()}{\rightarrow} x_{|X|-1,1} \overset{A}{\rightarrow} x_{|X|,4}$$

which shows a part of the process of a simulation task execution. (Let us denote this example sequence by DSPC(E)).

Showing also the methods assigned to the contexts, the DSPC(E) has the form:

$$x_{1,4}(m_{4,7}) \overset{A}{\rightarrow} x_{2,1}(m_{1,1}) \overset{C}{\rightarrow} x_{3,1}(m_{1,2}) \overset{A}{\rightarrow} x_{4,3}(m_{4,7}) \overset{A}{\rightarrow}$$

$$\overset{A}{\rightarrow} x_{5,1}(m_{1,2}) \overset{C}{\rightarrow} x_{6,1}(m_{1,3}) \overset{C}{\rightarrow} \overset{C}{\rightarrow} x_{7,1}(m_{1,4}) \overset{B}{\overset{A}{\rightarrow}} x_{8,4}\ (m_{4,7}) \overset{A}{\rightarrow}$$

$$\overset{A}{\rightarrow} x_{9,1}\ (m_{1,2}) \overset{C}{\rightarrow} x_{10,1}(m_{1,3}) \overset{B}{\overset{A}{\rightarrow}} x_{11,4}(m_{4,7}) \overset{C}{\rightarrow} \mathrm{x}_{12,4}(\mathrm{m}_{4,8}) \overset{A}{\rightarrow}$$

$$\overset{A}{\rightarrow} x_{13,1}(m_{4,8}) \overset{A}{\rightarrow} x_{14,1}(m_{1,4}) \overset{C}{\rightarrow} x_{15,1}(m_{1,5}) \overset{C}{\rightarrow} x_{16,1}(m_{1,6}) \overset{()}{\rightarrow} \ldots$$

$$\overset{()}{\rightarrow} x_{|X|-1,1}\ (m_{1,5}) \overset{A}{\rightarrow} x_{|X|,4}(m_{4,7}).$$

The goal is to introduce the work of the approach through the examination of *typical sequence patterns* of DSPC(E).

The sequence fragment $x_{1,4} \overset{A}{\rightarrow} x_{2,1}$ shows the transformation of the starting *cp* problem contest to a *su* problem context (p_d) which is a problem structuring pattern. Using the *su*-type method $(m_{1,1})$ for the transformation would be inefficacious and thus inefficient.

The sequence fragment $x_{3,1} \overset{A}{\rightarrow} x_{4,3} \overset{A}{\rightarrow} x_{5,1}$ is a softening up pattern: for $x_{3,1}$, it is necessary to involve a soft method (p_f). There is no M_3-type method in the set of methods (lack of soft method), thus SSM used $x_{4,3}$ which is efficacious for the case. (Using the *su*-type method $(m_{1,2})$ for $x_{4,3}$ would be inefficacious.)

In the DSPC(E), the $\overset{C}{\rightarrow} x_{7,1} \overset{B}{\overset{A}{\rightarrow}} x_{8,4}$ and the $x_{10,1} \overset{B}{\overset{A}{\rightarrow}} x_{11,4}$ *insertion* transitions occur. Without taking into account *insertions* (p_e does not function in the methodology), the examined steps of DSPC(E) may have, for example, the following forms

$$x_{7,1}\ (m_{1,4}) \overset{B}{\overset{C}{\rightarrow}} x_{8,1}(m_{1,5}) \text{ and } x_{10,1}(m_{1,3}) \overset{B}{\overset{C}{\rightarrow}} x_{11,4}(m_{1,4}).$$

The execution of the next phase of the simulation (the use of methods $m_{1,5}$ and $m_{1,4}$ for the problem context $x_{8,4}$ and $x_{11,4}$) is inefficacious, or from other point of view, the processing of the contexts $x_{8,1}$ and $x_{11,4}$ are not the contexts to process.

The $x_{12,4} \overset{A}{\rightarrow} x_{13,1} \overset{A}{\rightarrow} x_{14,1}$ sequence of DSPC(E) shows the pattern of elimination of the methodological gap (p_c). In the examined sequence, the use $m_{4,8}$ is used for the sequence $x_{12,4} \overset{A}{\rightarrow} x_{13,1}$ before the use of $m_{1,4}$ for $x_{14,1}$ (before the simulation phase). The use of $m_{1,4}$ would be inefficacious for $x_{12,4} \overset{A}{\rightarrow} x_{13,1}$ and $m_{4,7}$ would also be inefficient for the context $x_{13,1}$.

The closing sequence fragment of DSPC(E) $x_{|X|-1,1} \overset{A}{\rightarrow} x_{|X|,4}(m_{4,7})$ is the reverse case of the starting sequence $x_{1,4} \overset{A}{\rightarrow} x_{2,1}$: the use of $m_{1,5}$ instead of $m_{4,7}$ would be inefficacious for the $x_{|X|,4}$ context.

The set of methods may be described as $M = M_{su} \cup M_{cu} \cup M_{sp} \cup M_{cp}$ and in the examined case, $M_{sp} = \emptyset$ and $M_{cu} = \emptyset$. To improve efficiency, for example, the user methods may be used in the design of set of methods (p_a, p_b).

Remark: The example does not give an exhaustive analysis: there are no example patterns for the principles p_g, p_h and p_i.

VI. APPLICABILITY OF THE RESULTS

The results of the paper have already been successfully used in several large projects such as, for example: improving the performance of BP of a large telecommunication service company by integration of ICT services [16], modelling and simulation of a large CRM system (together with the problem context retrieval approach [17]), simulation project of BCP-DRP (Business Continuity Plan – Disaster Recovery Plan) in order to support planning at a large power service company.

The approach proposed by the paper is general: it is generally applicable for the meta-modelling of any ICT/BP because no domain-specific restrictions were used. Results may also be applied for other systems with components *technical-subsystem/human-subsystem*: for example, *traffic-subsystem/service-process-subsystem*, *environment-protection-subsystem/human-supervision-subsystem*. For these cases, the system features of the 4-state and 2-state models and the question of the necessary methods should be revised.

VII. CONCLUSION

In this paper, a new method for the management and increase of the efficiency of modelling and simulation of Organisational Information Systems has been formulated.

Using the 4-state and the 2-state models of problem contexts, the features of the dynamic simulation problem context for the case of modelling and simulation of organisational ICT and BP systems were investigated.

The method of Jackson and Keys (the approach for appropriate fitting of problem contexts and methods) was extended: Taking into account the features of the dynamic simulation problem context (the 2-state and 4-state models of problem context types) and the efficiency principles – the Checkland's systems performance criteria *re-defined* simulation efficiency measures and the criterion of *gap-efficiency* with the concept of explaining the occurrence and elimination of the methodological gap in the process of simulation –, the *principles of the managing the efficiency of simulation* were formulated. These principles of the problem context state approach are able to deal with both aspects of the efficiency of simulation: with the step-by-step efficiency of fitting of states and state-transitions with methods and with the long-range requirement of efficiency according to the amount of states and transitions.

The newly formulated set of principles *of the managing the efficiency of simulation* may also be taken as general requirements (including requirements for the set of methods, for the structure and operations) for building and implementation of a simulation meta-methodology.

The work of the new approach is illustrated by the analysis of an example of a simulation task. The applicability of the approach is shortly overviewed for the common analysis of systems with different cooperating subsystem components.

REFERENCES

[1] P. Checkland, "From Optimizing to Learning: A Development of Systems Thinking", *J. Opl. Res. Soc.* Vol. 36, No. 9, 1985. pp. 757-767.

[2] P. Checkland, "Soft Systems Methodology", in *Rational Analysis for a Problematic World*, Edited by J. Rosenhead, John Wiley & Sons, 1989.

[3] M. C. Jackson, *Systems Methodology for the Management Sciences*, Plenum Press, New York, 1991.

[4] M. C. Jackson, P. Keys, "Towards a System of Systems Methodologies", *J. Opl. Res. Soc.* Vol. 35, No. 6, 1984.

[5] R. Jain, *The Art of Computer Systems Performance Analysis*, John Wiley & Sons, New York, 1991.

[6] L. Muka, G. Lencse, "Developing a Meta-Methodology Supporting the Application of Parallel Simulation", in *Proc. 2006 European Simulation and Modelling Conference,* Toulouse, 2006, pp. 117-121.

[7] L. Muka, G. Lencse, "Hard and Soft Approaches in a Simulation Meta-methodology", in. *Proc. 5th Industrial Simulation Conference,* Delft, 2007, pp. 17-22, 2007.

[8] L. Muka, G. Lencse, "Developing a meta-methodology for efficient simulation of infocommunication systems and related processes", *Infocommunications Journal*, Vol. LXIII, No. 7, pp. 9-14, 2008.

[9] P. Ray, A. Serrano, "Collaborative Systems and Business Process Design Using Simulation", *Proc. 37th Hawaii International Conference on Systems Sciences*, 2004.

[10] M. Sierhuis, W.J. Clancey, "Modeling and Simulating Work Practice: A Method for Work System Design", *IEEE Intelligent Systems*, September/October, Vol. 17, No. 5, pp. 32-41. 2002.

[11] M. Sierhuis, A. M. Selvin, "Towards a Framework for Collaborative Modeling and Simulation", *Workshop on Strategies for Collaborative Modeling & Simulation*, CSCW '96, Boston, MA

[12] M. Warmerdam, P. Bredveld, "A Holistic Approach to Delivering the Value of IT: Business Service Management", *IDC White Paper*, Framingham USA, www.idc.com, 2003.

[13] T. Nakamura, K. Kijima, *System of system failures: Meta methodology for IT engineering safety*, John Wiley & Sons, 2008.

[14] M. C. Jackson, *System Thinking - Creative Holism for Managers*, John Wiley & Sons, 2003.

[15] R. J. Paul, V. Hlupic, G. Giaglis, "Simulation Modelling of Business Processes", Accepted for UKAI'98 – *UK Academy of Information Systems Conference*, Lincoln, UK, 1998.

[16] L. Muka, G. Muka, "BPR projekt támogatása komplex szinulációs modellel", *Minőség és Megbízhatóság,* 2. szám, 88-93 oldal, 2009.

[17] L. Muka, B. K. Benko, "Meta-level performance management of simulation: The problem context retrieval approach" *Periodica Polytechnica*, to be published.

László Muka graduated in electrical engineering at the Technical University of Lvov in 1976. He got his special engineering degree in digital electronics at

the Technical University of Budapest in 1981, and became a university level doctor in architectures of CAD systems in 1987. Dr Muka finished an MBA at Brunel University of London in 1996. Since 1996 he has been working in the area of simulation modeling of telecommunication systems, including human subsystems. He is a regular invited lecturer in the topics of application of computer simulation for performance analysis of telecommunication systems at the Multidisciplinary Doctoral School of Engineering, Modelling and Development of Infrastructural Systems at the Széchenyi István University of Győr.

Gábor Lencse received his M.Sc. in electrical engineering and computer systems at the Technical University of Budapest in 1994 and his Ph.D. in 2001. The area of his research is (parallel) discrete-event simulation methodology. He is interested in the acceleration of the simulation of info-communication systems. Since 1997, he has been working for the Széchenyi István University in Győr. He teaches computer networks and networking protocols. Now, he is an Associate Professor. He is a founding member of the Multidisciplinary Doctoral School of Engineering, Modelling and Development of Infrastructural Systems at the Széchenyi István University. He does R&D in the field of the simulation of communication systems for the Elassys Consulting Ltd. since 1998. Dr Lencse has been working part time at the Budapest University of Technology and Economics (the former Technical University of Budapest) since 2005. There he teaches computer architectures.

Analysis of a new Markov model for packet loss characterization in IPTV Solutions

T. Jursonovics, S. Imre

***Abstract*—Several methods have been developed for packet loss characterization in IP networks, typically introducing Markov models. On the one hand these techniques are performing well in the description of real time communication but on the other hand they are not able to perfectly describe the content delivery of an IPTV Solution extended with application level packet retransmission. This specific problem has been introduced and addressed by this paper with the aim of creating a simple but more accurate traffic model on the UNIT-17 interface for IPTV systems. Analysis of the implemented new Markov chain has been conducted to express its major parameters in a closed mathematical form. The results show that the proposed model is reducible to the Gilbert Model and its performance is validated and compared in different network situations with legacy models.**

***Index Terms*—IPTV, Markov process, OIPF, Packet loss, Retransmission**

I. PREFACE

This article discusses several IPTV related technologies and features which are usually part of today's IPTV Service Platforms but in real life described and referenced on a different way vendor by vendor. These diverse terms may confuse the reader, therefore we have decided to implement and consistently use the terminology of the Open IPTV Forum (OIPF) release 2 specifications [1]. We believe that the adaptation of the OIPF terms to any specific IPTV Solution is straightforward therefore all our results can be easily interpreted on any IPTV environment.

The scope of this paper includes the mathematical description and analysis of a proposed Markov model but the detailed explanation of any IPTV specific feature or system architecture are beyond the limits of this article. In case of any uncertainty we ask the reader to visit the OIPF specifications [1] for further clarification.

II. INTRODUCTION

Customers expect high quality and reliable IPTV services therefore Service Platform Providers must introduce enablers like delivery over managed networks [10], QoS or Application Level Retransmission (RET) [6] to assure the required level of quality. Against all attempt, quality deterioration still occurs, therefore Service Providers must implement network characterization models and quality measurement features to adjust and monitor the performance of their services.

In the upcoming section we are going to describe the technical features of an IPTV Solution (Section III) with the aim of providing the necessary technical overview to understand our further modeling efforts. Section IV gives a short summary about modeling requirements highlighting the role of the packet loss characterization. The main work and most of our effort to create a better model for packet losses in the IPTV Solutions will be introduced and analyzed by Section V. Finally Section VI concludes the paper with our results.

Tamás Jursonovics was with the Budapest University of Technology and Economics, Budapest, Hungary. He is now with the Service Development Department of Magyar Telekom Plc, Budapest, Hungary (e-mail: jursonovicst@gmail.com).

Sándor Imre Author is with the Department of Telecommunications, Budapest University of Technology, Budapest, Hungary (e-mail: imre@hit.bme.hu).

III. CONTENT DELIVERY IN IPTV

Scheduled Content (also known as Linear TV) and Content on Demand (CoD, aka. Video on Demand) are the general – and probably most challenging – services of any IPTV Solution. The user experience is mainly determined by the role of Content Production, Content Delivery and Content Reconstitution. Although none of these roles are less important that the others, our work concentrates on the Content Delivery and discusses only this aspect from the network point of view.

A. Provider(s) Network

The architecture of the Content Delivery has two main segments [2]. The Provider(s) Network distributes the content to the Residential Network. Figure 1 shows the relevant part of the high level architecture of OIPF.

Multicast Content Delivery Function

This entity is responsible for the delivery of content and generic data to the OITF using multicast streams and the multicast data channel respectively. This is the so-called head end, the source of the multicast data channel.

The Content Delivery Network

This is a fundamental functionality in an IPTV CoD solution. For CoD, it allows the optimization of the network

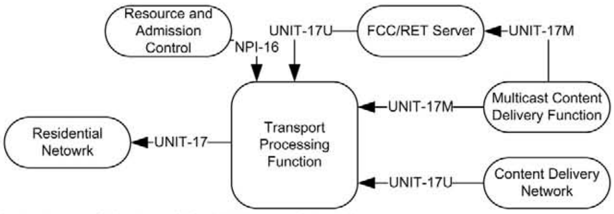


Fig. 1. Part of the Provider(s) Network Architecture

use through a distribution of the media servers in the physical network, and the optimization of the storage resources through a popularity-based distribution of the content on the media servers. This results in having popular content massively distributed on media servers at the edge of the network (as close as possible to the customer) while less popular content is distributed on a reduced number of media servers. For scheduled content, it enables the support of enhanced services like Personal Channel (PCh) and Network Personal Recording (nPVR).

Fast Channel Change/Retransmission Server

The functional entity that delivers ancillary data for multicast streams when triggered by the OITF, in the context of FCC/RET service.

Transport Processing Function

This functional entity includes the functions needed to support real-time multicast and unicast streams, optimizing network usage in the physical network and enforcing related traffic policies coming from Resource and Admission Control.

UNIT-17

Content stream including content, content encryption (for protected services) and content encoding. This reference point can be used for both multicast and unicast (UNIT-17M and

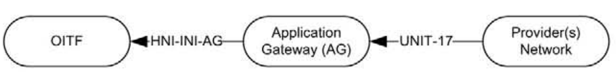


Fig. 2. Part of the Residential Network Architecture

UNIT-17U, respectively) and uses IP over UDP over RTP and HTTP (unicast only) protocols. It includes the FCC/RET RTP packets issued by the FCC/RET server.

B. Residential Network

The Residential Network (Figure 2) delivers the content within the household to the OITF.

Open IPTV Terminal Functional Entity

The OITF includes the functionality required to access IPTV service for both the unmanaged and the managed network models through the HNI-INI interfaces and renders the content on a display device. Typically this entity is implemented in an STB.

Application Gateway Functional Entity

The Application Gateway is an optional gateway function that incorporates a procedural language based application execution environment where applications can be remotely downloaded for execution.

C. QoS Framework

Network Providers usually realize a QoS Framework in the Managed Network model to assure service quality by deploying policy based transport control in the Transport Processing Function [11]. Although this provides reliable transport lines and avoids from router packet drops in the metro area of the Access Network, packet loss still occurs on the last mile, especially with xDSL and EuroDocsis technologies. Network Providers are pushed by customers to increase the access bandwidth which forces the xDSL technology to its limit mostly over long and old twisted copper pairs. Cable networks use shared media which is also effected by noise.

To overcome these difficulties, OIPF Release 2 specification introduces Application Level Retransmission (RET). This ARQ feature corrects the lost packets on the fly but requires several QoS attributes from the Transport Processing Function like bandwidth allocation or delay and jitter protection.

IV. Challenges of Traffic Modeling in IPTV

This section introduces the motivation behind our effort to create a new traffic model for IPTV Solutions. The UNIT-17 interface specifies RTP/UDP/IP protocols for the real-time content delivery of the IPTV service which flow consists of individual RTP packets carrying MPEG frames [14]. Several models [3]-[8] have been developed for quality prediction and packet loss characterization in IP networks, like Gilbert-Elliot, Extended Gilbert or HMM but they are not considering the selective effect of the Application Level Retransmission.

RET is widely deployed in the state of the art IPTV systems. In the first place this feature generally decreases the performance of the current models with the implementation of a selective repeat packet loss correction. On the top of that legacy models are not able to describe the required additional bandwidth for the error correction packets transmitted over the same communication channel which is essential for the network characterization to provide a reliable IPTV service.

The steady-state packet loss rate (P_L) is calculated according to the number of the received/lost packets but legacy models predict higher probability because RET selective retransmits lost packets.

The average loss burst length (B_L) is essential to plan the GOP size for an MPEG transmission to avoid from the loss of an I-Frame [12]. Legacy models predict higher B_L or they are not able to capture the real distribution of loss bursts because RET usually divides or minimizes lost runs with error correction.

For the optimal network resource allocation Resource Admission Control needs accurate bandwidth planning but

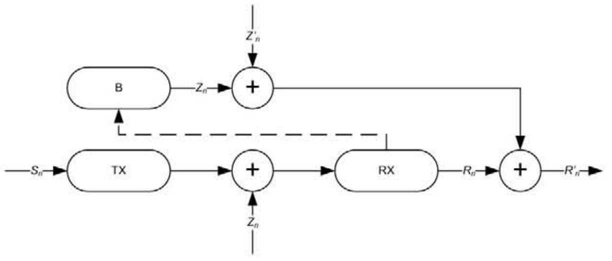


Fig. 3. Discrete communication system model

legacy models are not considering the required bandwidth for RET packets.

RET dramatically influences the characteristic of packet transmission, this change has to be taken into account. Therefore we developed and introduced a new simple communication model based on a three-state Markov chain to address this specific question and to provide better characterization results.

V. Three-state Markov model

Consider the discrete communication system for the mathematical model for the packet transmission over UNIT-17 interface of the IPTV architecture in Figure 3. Let S_n={0, for all n} represents the sent RTP packets at the time n, Z_n={0;1} a random noise in the communication channel, TX the Transmitter and RX the Receiver device. We can express the received packets with R_n={0, for successful packet transmission; 1, for packet loss} as the function of the channel noise (1) because S_n does not carry any real information (we assume that all packet were sent out properly):

$$R_n = S_n + Z_n = Z_n \,. \tag{1}$$

The IPTV RET feature extends the system with the element of B, the retransmission buffer. We assume, that all lost packets will be requested only once for retransmission, therefore the retransmission signal equals with Z_n and travels on the same channel. This means that the retransmission signal will also be effected by channel noise (Z'_n). *Expressing* the final, corrected signal:

$$R'_n = R_n + (Z_n + Z'_n) = Z_n + Z_n + Z'_n = Z'_n \,. \tag{2}$$

This shows that the packet transmission (R'_n) can be described by a discrete time random process.

A. Markov model

For the packet loss characterization in the communication system above, introduce X_n, a discrete time random process to describe the transmission state of the n-th RTP packet: X_n=0, if the n-th packet is received correctly; X_n=1, if the n-th packet is lost and has not been retransmitted and finally, X_n=2, if the n-th packet is successfully retransmitted after loss.

We showed that the probability of the packet loss and the retransmission are represented by stochastic processes which do not depend on any previous state (2), so X_n possesses the Markov property (3), therefore X_n can be described with a Markov model.

$$\begin{aligned} &P(X_n = x_n \mid X_{n-1} = x_{n-1} \ldots X_0 = x_0) = \\ &= P(X_n = x_n \mid X_{n-1} = x_{n-1}) \end{aligned} \tag{3}$$

where $x_0, x_1, \ldots x_n \in \varepsilon$ for every $n \geq 0$. Let us construct a Markov chain (Figure 4) with the $p_{ij}[n]$ transition probabilities from the state i to j at the time n:

$$p_{ij}[n] = P(X_{n+1} = j \mid X_n = i) \,. \tag{4}$$

We also assume that the main characteristic of the communication channel will not change in time, therefore transition probabilities do not depend on time and the Markov chain is time-homogeneous:

$$p_{ij}[n] = p_{ij} \,. \tag{5}$$

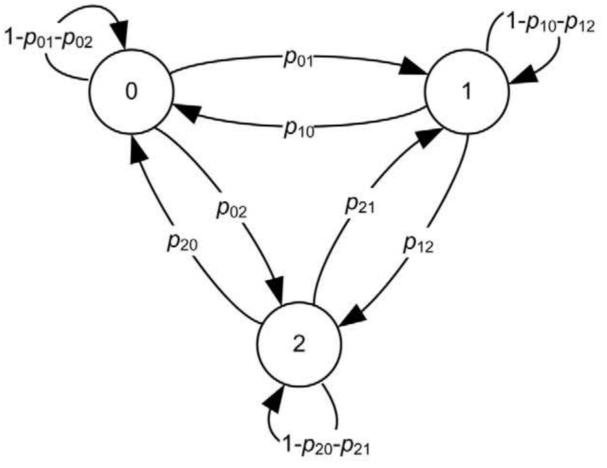


Fig. 4. State Transition Diagram

B. Analysis

We start our analysis from describing the system equations of the Markov chain:

$$y_0[n+1] = (1 - p_{01} - p_{02}) y_0[n] + p_{10} y_1[n] + p_{20} y_2[n] \tag{6}$$

$$y_1[n+1] = (1 - p_{10} - p_{12}) y_1[n] + p_{21} y_2[n] + p_{01} y_0[n] \tag{7}$$

$$y_2[n+1] = (1 - p_{20} - p_{21}) y_2[n] + p_{02} y_0[n] + p_{12} y_1[n] \tag{8}$$

$$p_{00} + p_{01} + p_{02} = 1 \tag{9}$$

$$p_{11} + p_{10} + p_{12} = 1 \tag{10}$$

$$p_{22} + p_{20} + p_{21} = 1 \tag{11}$$

where p_{ij} are the state transition probabilities, $y_0[n]$ is the probability at the time n of successfully packet transmission, $y_1[n]$ of loss and $y_2[n]$ of retransmission for $n \geq 0$. Taking the z-transform from both sides of the system equations (6, 7, 8) we move to the complex frequency domain where $Y_0(z)$, $Y_1(z)$, $Y_2(z)$ are the z-transforms of $y_0[n]$, $y_1[n]$ and $y_2[n]$:

$$z(Y_0(z)-y_0[0])=(1-p_{01}-p_{02})Y(z)+p_{10}Y_1(z)+ \\ +p_{20}Y_2(z) \tag{12}$$

$$z(Y_1(z)-y_1[0])=(1-p_{10}-p_{12})Y_1(z)+p_{21}Y_2(z)+ \\ +p_{01}Y_0(z) \tag{13}$$

$$z(Y_2(z)-y_2[0])=(1-p_{20}-p_{21})Y_2(z)+p_{02}Y_0(z)+ \\ +p_{12}Y_1(z) \tag{14}$$

From the packet characterization point of view, one key attribute is the steady-state packet loss probability, therefore let us express $Y_1(z)$ by solving the equations above (12, 13, 14):

$$\begin{aligned} Y_1(z)= {} & z(2y_1[0]z+y_1[0]-p_{21}y_2[0]-y_1[0]p_{20}- \\ & -y_1[0]p_{21}+p_2y_0[0]p_{21}+y_2[0]zp_{21}+p_{01}p_{20}y_2[0]+ \\ & +y_2[0]p_{01}p_{21}+y_2[0]p_{02}p_{21}-p_{02}y_1[0]+zy_1[0]p_{20}+ \\ & +zy_1[0]p_{21}+p_{01}y_1[0]z+p_{01}y_1[0]p_{20}+p_{01}y_1[0]p_{21}+ \\ & +p_{01}y_0[0]z+p_{01}y_0[0]p_{20}+p_{01}y_0[0]p_{21}+p_{02}y_1[0]z+ \\ & +p_{02}y_1[0]p_{21}+y_1[0]z^2-p_1y_1[0]-p_1y_0[0])/ \\ & (-1+3z+p_{01}+p_{02}+p_{10}+p_{20}+p_{12}+p_{21}-2zp_{01}- \\ & -3z^2+z^3+zp_{20}p_{12}+zp_{20}p_{10}+zp_{21}p_{10}+zp_{02}p_{10}+ \\ & +zp_{01}p_{20}+zp_{01}p_{21}+zp_{02}p_{21}+p_{12}p_{01}z+p_{12}p_{02}z- \\ & -2zp_{02}-2p_{10}z-p_{10}p_{02}-2p_{12}z-p_{12}p_{01}-p_{12}p_{02}- \\ & -p_{21}p_{10}-2p_{20}z-p_{20}p_{10}-p_{20}p_{12}+z^2p_{21}-2zp_{21}+ \\ & +z^2p_{01}+z^2p_{02}+p_{12}z^2-p_{01}p_{20}-p_{01}p_{21}-p_{02}p_{21}+ \\ & +z^2p_{10}+z^2p_{20}) \end{aligned} \tag{15}$$

Let us observe that $Y_1(z)$ can be written as a fraction of two polynomials in the following form:

$$Y_1(z)=\frac{B(z)}{A(z)}=\frac{\sum_{i=0}^{M}b_iz^{-i}}{\sum_{i=0}^{N}a_iz^{-i}}. \tag{16}$$

By applying partial fraction expansion, we obtain:

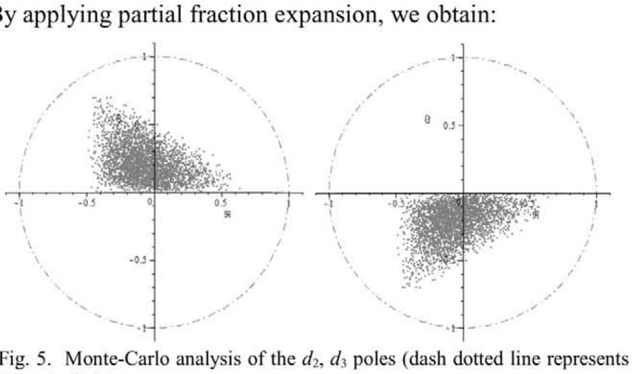


Fig. 5. Monte-Carlo analysis of the d_2, d_3 poles (dash dotted line represents the unit circle)

$$\begin{aligned} & Y_1(z)=\sum_{r=0}^{M-N}B_rz^{-r}+\sum_{i=1;i\neq n}^{N}\frac{A_i}{1-d_iz^{-1}}+\sum_{j=1}^{s}\frac{C_j}{\left(1-d_nz^{-1}\right)^j}, \\ & A_i=\left(1-d_iz^{-1}\right)\cdot Y_1(z)\Big|_{z=d_i} \\ & C_j=\frac{1}{(s-j)!(-d_i)^{s-j}}\left\{\frac{d^{s-j}}{dw^{s-j}}\left[(1-d_iw)^sX\left(w^{-1}\right)\right]\right\}_{w=d_i^{-1}} \end{aligned}, \tag{17}$$

where B_r can be calculated with a long division, and d_i are the poles of $A(z)$. d_i can be calculated from (15, 16):

$$d_1=1 \tag{18}$$

$$\begin{aligned} d_2= {} & 1-\frac{1}{2}(p_{01}+p_{02}+p_{10}+p_{12}+p_{20}+p_{21})+ \\ & +\frac{1}{2}(p_{01}^2+p_{02}^2+p_{10}^2+p_{12}^2+p_{20}^2+p_{21}^2+2p_{21}p_{20}+ \\ & +2p_{21}p_{12}+2p_{12}p_{10}+2p_{20}p_{02}+2p_{01}p_{02}+2p_{01}p_{10}- \\ & -2p_{10}p_{02}-2p_{12}p_{01}-2p_{12}p_{02}-2p_{21}p_{10}-2p_{20}p_{10}- \\ & -2p_{20}p_{12}-2p_{01}p_{20}-2p_{01}p_{21}-2p_{02}p_{21})^{\frac{1}{2}} \end{aligned} \tag{19}$$

$$\begin{aligned} d_3= {} & 1-\frac{1}{2}(p_{01}+p_{02}+p_{10}+p_{12}+p_{20}+p_{21})- \\ & -\frac{1}{2}(p_{01}^2+p_{02}^2+p_{10}^2+p_{12}^2+p_{20}^2+p_{21}^2+2p_{21}p_{20}+ \\ & +2p_{21}p_{12}+2p_{12}p_{10}+2p_{20}p_{02}+2p_{01}p_{02}+2p_{01}p_{10}- \\ & -2p_{10}p_{02}-2p_{12}p_{01}-2p_{12}p_{02}-2p_{21}p_{10}-2p_{20}p_{10}- \\ & -2p_{20}p_{12}-2p_{01}p_{20}-2p_{01}p_{21}-2p_{02}p_{21})^{\frac{1}{2}} \end{aligned} \tag{20}$$

We note that d_1 is on, and d_2, d_3 are within the unit circle showed by a Monte-Carlo simulation (Figure 5) performed with 100 k random samples with $0\leq p_{ij}\leq 1$, where $0\leq p_{01}+p_{02}\leq 1$, $0\leq p_{10}+p_{12}\leq 1$ and $0\leq p_{20}+p_{21}\leq 1$ given by (9, 10, 11):

$$\begin{aligned} & |d_2|<1 \\ & |d_3|<1 \end{aligned} \tag{21}$$

While in our model M=3, N=3 and all the poles of $A(z)$ are first order poles, the third term of (17) does not exists, $Y_1(z)$ can be reduced to the form of:

$$\begin{aligned} & Y_1(z)=\sum_{r=0}^{M-N}B_rz^{-r}+\sum_{i=1}^{3}\frac{A_i}{1-d_iz^{-1}} \\ & A_i=\left(1-d_iz^{-1}\right)\cdot Y_1(z)\Big|_{z=d_i} \end{aligned} \tag{22}$$

Each term can be inverse transformed by inspection, where $u[k]$ is the unit step function and $\delta[k]$ is the Dirac-delta function:

$$y_1[k] = B_r\delta[k] + \sum_{i=1}^{3} u[k]A_i d_i^k = \qquad (23)$$
$$B_r\delta[k] + u[k]A_1 d_1^k + u[k]A_2 d_2^k + u[k]A_3 d_3^k$$

C. *Steady-state analysis of packet loss*

The steady-state packet loss rate P_L of $y_1[k]$ can be derived by:

$$P_L = \lim_{k\to\infty} y_1[k] = \qquad (24)$$
$$= \lim_{k\to\infty}\left(u[k]A_1 d_1^{\,k} + u[k]A_2 d_2^{\,k} + u[k]A_3 d_3^{\,k}\right)$$

Let us observe that the second and third terms are converging to zero (21), so P_L can be expressed in a closed form:

$$P_L = \lim_{k\to\infty} u[k]A_1 d_1^{\,k} =$$
$$= (p_{02}p_{21} + p_{01}p_{20} + p_{01}p_{21}) / \qquad (25)$$
$$(p_{10}p_{02} + p_{12}p_{01} + p_{12}p_{02} + p_{21}p_{10} + p_{20}p_{10} +$$
$$+ p_{20}p_{12} + p_{01}p_{20} + p_{01}p_{21} + p_{02}p_{21})$$

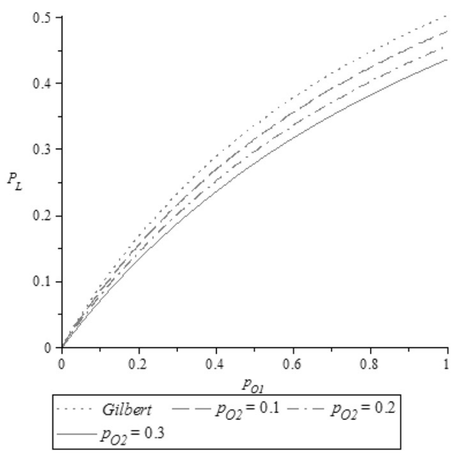


Fig. 6. Performance – Retransmission

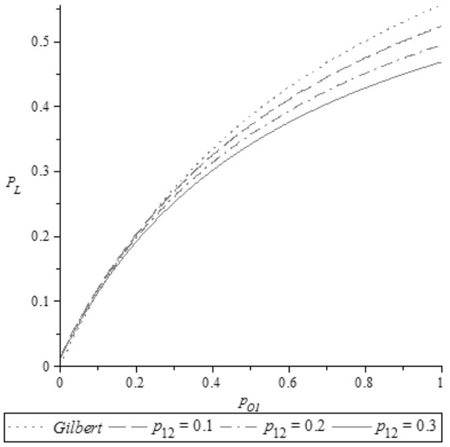


Fig. 7. Performance – Error correction on packet loss burst

D. *Model performance*

We validate the performance of the proposed model by calculating packet loss rate $P_L = f(p_{01})$ on different types of channels represented by different model parameters (p_{ij}). To highlight the strength of our model, we also indicated the loss rate of the Gilbert model.

In Figure 6 we assume a simple loss channel (p_{10}=0.99, p_{20}=0.99, p_{12}=0, p_{21}=0) where the correction effect of the retransmission feature lovers the probability of the packet loss.

Figure 7 describes a loss burst channel (p_{10}=0.8, p_{02}=0.01, p_{20}=0.01, p_{21}=0.9). The retransmission feature takes effect only at high loss rates (p_{01}>0.3) and improves the success of the packet transmission.

E. *Reduction to the Gilbert model*

If we reduce the proposed model to the well known Gilbert model by defining the state probabilities to newer entries to the state "2" (p_{02}=0, p_{20}=1, p_{12}=0 and p_{21}=0) then the packet loss rate from (25):

$$P_L\Big|_{p_{02}=0,\ p_{20}=1,\ p_{12}=0,\ p_{21}=1} = \frac{2p_{01}}{2p_{10} + 2p_{01}} \qquad (26)$$

which is exactly the loss rate of the Gilbert model.

VI. CONCLUSION

New model was created and proposed for packet loss attribute specification in IPTV Solutions where we showed that the existing models are not considering the effect of the retransmission feature. We designed a Markov chain representation and delivered its analytical analysis. As a result of our work we described the packet loss rate – a key performance indicator –in a closed form and we compared and evaluated the performance of our model which proves our assumptions.

VII. FURTHER WORK

We have decided to give the primary focus in this paper on the foundation and the analytical analysis of our proposed model, however we understood that the application is equally important therefore we present our further plans below.

Firstly we are going to evaluate the proposed model with real-time IPTV traffic measurement. The testbed environment was successfully set up but the data processing has not been finished till the preparation of this paper. We are mainly interested in evaluation of our theoretical results in real life IPTV Solutions.

Secondly we propose to use our model and results for accurate bandwidth estimation of the Resource Admission Control on the Unit-17 interface and for bandwidth profile planning on different access technologies. As we emphasized in Section IV our work considers the requirements and effects of the Application Level Retransmission in IPTV Solutions which requires a proper bandwidth management for the delivery of the additional correction data over the same communication channel.

Finally in our previous [13] work we proposed an application area for loss characterization in online charging solutions for multimedia streams over mobile networks. We are interested in evaluating our models in the mobile environment as well.

REFERENCES

[1] Open IPTV Forum – Release 2 Specifications (http://www.oipf.tv/specifications.html)
[2] OIPF – Functional Architecture [V2.1] – [2011-03-15], (http://www.oipf.tv/specifications.html)
[3] E.N. Gilbert, “Capacity of A Burst-Noise Channel,” Bell Syst. Tech J., vol. 39, pp. 1253-1265, September 1960.
[4] E. O. Elliott, “Estimates of error rates for codes on burst-noise channels,” Bell Syst. Tech. J., vol. 42, pp. 1977–1997, 1963.
[5] H. Sanneck, G. Carle, and R. Koodli, “A framework model for packet loss metrics based on loss run length,” in Proceedings of SPIE/ACM SIGMM Multimedia Computing and Networking Conference, January 2000.
[6] K. Salamatian and S. Vaton –“Hidden markov modeling for network communication channels,” in Proc. of SIGMETRICS, 2001
[7] Zs. Butyka, T. Jursonovics, S. Imre – “New fair QoS-based charging solution for mobile multimedia streams”, Int. J. Virtual Technology and Multimedia, Vol. 1, No. 1, 2008.
[8] Fengdan Wan, Lin Cai, and T. Aaron Gulliver, “A Simple, Two-Level Markovian Traffic Model for IPTV Video Sources”, GLOBECOM 2008
[9] ITU-T H.701 (03/2009) – “Content delivery error recovery for IPTV services”, 2009.
[10] ITU-T Y.1541 –“Network performance objectives for IP-based services”, 2006.
[11] ITU-T H.360 – “An architecture for end-to-end QoS control and signaling”, 2004.
[12] T. Jursonovics, Zs. Butyka, S. Imre – “Examination and New Charging Solution for Multimedia Streams over Mobile Network”, MoMM2005, pp. 311--320. (ISBN 3-85403-195-5), 2005.
[13] T. Jursonovics, Zs. Butyka, S. Imre – “Multimedia Transmission over Mobile Networks”, SoftCOM2004, pp. 453--457. (ISBN 9-53611-469-0), 2004
[14] ETSI TR 102 033 – “Digital Video Broadcasting (DVB); Architectural framework for the delivery of DVB-services over IP-based networks”, V1.1.1 2002
[15] S. Meyn, R. L. Tweedie, *Markov Chains and Stochastic Stability*, London: Springer-Verlag, 1993

Healthcare Service Based on ICT: the eHealth8 Telemedicine System

(Design Studies Paper)

László Daragó[1], Csaba Engi[2], György Ferenczi[3], István Pesti[4], Dezső Vass[5]
[1]Semmelweis University, [2]Answare Ltd., [3]Thormed Ltd., [4]Humansoft Ltd., [5]Bay Zoltán Foundation

Abstract **- Telemedicine is a tool of healthcare (and education) based on telecommunication and information technology (IT), where the scene of patient and care (diagnosis/therapy/monitoring) are separated in time and/or place. This article discusses the concepts and design of a telemedicine system developed within the framework of the project implemented by the *eHealth8* consortium, first of all with focus on its architecture and communication solutions.**

I. Introduction

The idea of performing medical examinations and evaluations by means of telecommunication networks is not new. Shortly after the invention of telephone, experiments have already been started to transmit heart and lung sounds but transmission capabilities have always arisen an issue.
Some recent applications are as follows.

International Virtual e-Hospital, Kosovo

Four telemedicine centers were established during a two-year working period. Telecommunication infrastructure was set up that enabled major medical institutions and medical universities to establish wideband telecommunication connections with each other and with the outside world. Telecommunication connections and Internet are being used both in active medical care and education usually by means of ordinary Internet technologies (e-mail, web, wiki, Skype, Groove etc.). The system includes applicable solutions about the usage of open source Internet standards. Open connection of general purpose elements (e. g. Skype), and the realized solutions suitable for operation in heterogeneous manufacturer environment were useful for the work of the Consortium.

Emergency Telemedicine Applications

During the evolutional process it can be observed that only the minimal amount of the necessary data collecting and communication devices are deployed on site, then analyses and professional intelligence services are provided from remote locations by means of data communications. This way professional knowledge of the centers can be utilized everywhere.
It is very useful that the technology of applying unified decisions made from a geographically remote location. Especially, special, differentiated devices have been set at patients' homes that report on the same communication channel, in the same form to decision makers.

Hungarian Atomic Energy Authority (HAEA) – Nuclear Emergency Response Case Management Support

HAEA is an administrative authority controlled by the Hungarian Government, commissioned with independent tasks and empowered with official rights. Its duty is the official supervision of safe utilization of nuclear energy; especially the safety control of nuclear plants and materials, and the prevention of nuclear proliferation; in addition catalogueing documents in a consistent system. The workflow based on documents and detailed documentation are worth to mention in this study. Also the methodology of tracking and automatic developing of frequently changing processes.

The Role of Service Oriented Architecture in Telemedicine Healthcare System / B&M

Due to the great diversity of devices and manufacturers, the most significant challenge in telemedicine projects is their interoperability. It is difficult to develop a system with properly flexible data exchange among different systems without restricting the groups of data processed. Numerous solutions exist but all of them aim the same from software development point of view. The DICOM system of B&M Company is a typical representative of this tendency. The utilization of SOA architecture, the outsourced expert pool, marketable, commercial off-the-shelf components and ASP function capability also should be noted from this project.

DISPEC TeleTriage and Dispatch System

The Bucharest Ambulance Service (SAMB) functions as an independent organization, and serves 2.5 million people. Beyond emergency care service its tasks are to provide home care in difficult cases, event control, organ, blood and medicine transportation, moreover issue of official certificates. The experiences of practical establishment of return calculations and utilization of pieces of experience accumulated in databases (knowledge base) could be used for the eHealth8 consortia, as well, as the functioning of the expert dispatcher team.

Electronic Immunization Registry and Tracking System in Bangladesh

Even nowadays, the state administration – as well as healthcare authorities - in this country are informed only about 74% of childbirths. Unfortunately vaccination of newborn babies and small children (most of them should be injected in 3-4, comparing to each other 3-4-9-month intervals) has considerable difficulties in getting to families that meant 40% of dispersion at each phase. As a result, some well-known and preventable diseases can still cause serious casualties. The technique of multiple utilization of the same data and the focus on everyday needs of patients concerned are valuable for our project too.

The eHealth8 project presented in this article aims at developing and implementing protocols of telemedicine based on the utilization of professional guidelines of traditional medicine on telemedicine, with focus on the support of patient-physician relationship. The fundamental requirements of the newly developed system intend to eliminate the deficiencies of current telemedicine systems mentioned above, at the same time expectations that are beneficial in domestic medical care environment are drafted.
After presenting the objectives and participants of the project, the article describes the architecture of the system developed, then goes into detail about the patient side interface, roles of components, their communication, and finally the business model of the system is outlined by a couple of sentences.

II. THE eHEALTH8 PROJECT

In spite of the fact that utilization of telemedicine – in addition to its necessary application in certain cases –, can result in more efficient operation of health care systems and in decrease of the employment of expensive and scant resources, it is less widespread than expected. The fact that technologies for telemedicine applications are available and there are numerous successfully implemented experimental pilot projects – though in many cases only in order to separately check and analyze the medical effectiveness of new technologies – urge the spread of telemedicine. However, there are several factors that present hurdles in the way of wide penetration of telemedicine:

- lack or poor quality of economic models, case studies, analyses that supplement available medical evidences,
- lack of structured processes of medical evidences available in certain fields and their integration into medical professional directives, protocols that are the basis of medical professionals' approval and availability of financial resources as well,
- above all the existing technology-based attitude instead of a complex approach considering other aspects of telemedicine introduction and successful long-term utilization, operation.

Successful utilization of telemedicine systems in the long run that brings real benefit for all participants is possible only if these impeding factors are eliminated. The **eHEALTH8** consortium of medical, IT, economic and measuring technique experts was established in 2008, in order to design a telemedicine system that is remunerative for patients, society and care system and develop its infocommunication and economic system.
Members of the consortium are, in alphabetical order, Answare Ltd. Bay Zoltán Foundation for Applied Research, Humansoft Ltd., Semmelweis University, Thormed Ltd.

The members of the consortium have prepared a project proposal to implement the tasks mentioned above, that was granted by the National Office for Research and Technology (NKTH) 'Lifestyle supported by info-communication devices' (AAL – Ambient Assisted Living) subprogram. The objective of the project was to establish the IT and medical basis and construct the prototype of a telemedicine service system that does not rely on presently available, separately operating telemedicine sub-processes but focuses on rearranging medical protocols on the basis of domestic and international experiences. By extending towards telemedicine the project makes efforts to apply protocol-based approach that fits in new, already existing medical therapeutic directives during the design of IT support of telemedicine processes and the development of services.
In order to increase the growth potential in efficiency as much as possible, protocol-based approach is supplemented by cost efficiency and feasibility study. The objective of this activity is to highlight application areas and processes, which are likely to result in the most considerable efficiency growth rate and/or cost reduction for the participants of healthcare service and/or financing system while achieving the same level of individual health benefits or higher.

The **eHEALTH8** project emphasizes the support of doctor-patient relationship within telemedicine. Fundamental requirements for the presently developed system are aimed at eliminating the deficiencies of current telemedicine systems mentioned above; moreover expectations that are beneficial in the domestic medical care environment are drafted:

- the system of devices and processes developed need to be based on medical protocols approved by medical profession;
- the system of devices and processes developed should bring the highest level health improvement for the patient;
- it should be simply extendable with measuring processes, examination and nursing protocols without any changes in the central base system but supported by its functions;

- it should ensure mobility and as extended independence for patients as possible during different medical treatments;
- it should have the capability to connect to the inter-institutional information systems established in HEFOP 4.4;
- it should be applicable to preventive (status and lifestyle assessment) and aftercare purposes.

III. THE ARCHITECTURE AND COMMUNICATION SCHEME OF TELEMEDICINE SYSTEM DEVELOPED BY eHEALTH8

The architecture (Figure 1) includes the main participants and the directions of service flow. The protocol service provider uses professional support of the protocol expert to interpret, assess the application and maintain the protocol itself. The telemedicine service provider also resorts to medical, health care supplier, technological and logistic service background. The figure demonstrates that patients can use service at least three ways: (i) within the framework of a health care service, (ii) as a fitness, wellness, sport, leisure activity or (iii) as an independent health surveillance or management service.

BASIC REQUIREMENTS FOR PATIENT SIDE INTERFACE

Telemedicine methods do not simply mean the utilization of traditional methods the way healthcare assistance is performed by patients or their surroundings. Complying with the rules regarding to the measurements requires self-disciplined, conscious behavior of the patient. In case the rules are disregarded but the measurement is performed, measured results will not be real values and the therapeutic decisions based on measurement will not bring the expected result.

Elderly people need a special unit which enables simple use, possibly with the simplest and most robust design. Due to weakened visual capabilities, large, in some cases extra large display is needed with an audio unit. Any data recording can be problematic due to trembling hands. Usually patients are afraid of busting up the devices they handle, so persistent, motivating education and training is a priority for them. Regular feedback and knowledge update are necessary during personal visits. In their case devices must contain instant emergency call function, too. Traditional medical devices were not designed in a way to enable the patients to use them on their own. The patient is a participant of measurement and intervention but the operator of the device is always a physician or medical professional. Measurement results are analyzed by a professional who decides if the result is within the expected range, i.e. whether it can be valid at all or the result of the measurement is false, loaded with disturbance, or the behavior of the patient could have an effect on it. During telemedicine processes it is not always possible to validate the way that focuses on selecting the right method and the autonomic validating and pre-evaluative functions of the examining system.

Examinations must start with identification of the patient in every case, so (theoretically!) it is impossible that another patient's measurement data are recorded in someone's documentation. Identification is of special importance during a telemedicine application when patients are not present at the site of healthcare provider but they are e.g. at their homes, where anyone who can have an access (e. g. family members) can use the device, even can substitute the patient. Another essential problem is how it can be stated with responsibility that the results of a measurement performed at a patient's home are acceptable, were created under the right conditions that are determined by measurement protocols, thus the medical decisions can be based on them. Fundamental precondition of telemedicine utilization is to solve the identification of the patient with a safety level corresponding to the cost of measurement or to the extent of possible damages caused by deliberate or accidental misidentification.

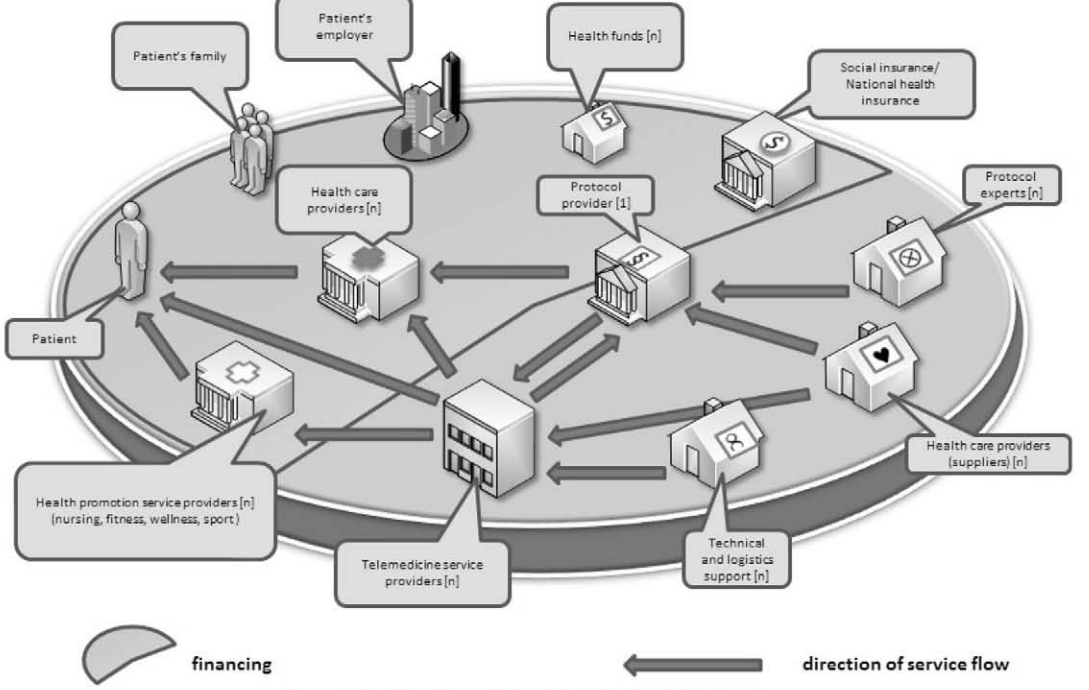


Figure 1: Service model of the eHealth8 telemedicine system

Currently available technologies– based on ownership, special and biometric information – make identification possible in various ways:

- ownership
  - RFID identification
  - bar code
- biometry
  - fingerprint identification
  - retina identification
  - based on typical characteristics of the individual of the examined parameter (e.g. exhaled breath)
- special knowledge
  - password (PIN code)

The utilization of these methods and combinations make different level of reliability in identification possible.

At present most of the currently available devices do not have any identification facilities so the patient is responsible for the origin of the measured values. The objective of the utilization of these devices is not to supplement the tele-monitoring system as an interface on the patient side but support patients' self-check. From technological point of view the correct solution could be fingerprint identification or an RFID chip implanted under the skin or worn on a bracelet used together with a reader built into the device. Taking into consideration the characteristic of the currently supported applications, development costs and applicability, at present fingerprint identification is practical to be accomplished.

To sum up, the main requirements for patient side devices are as follows:

- safety and reliability:
  - operation:
    - do not require installation performed by experts as far as possible;
    - the device should give information automatically on its technical condition to the centre (self-test);
    - its basic maintenance (recharge or replacement of battery) should be simply performed even by the user;
    - it should be possible to remotely update the device's firmware
  - data transmission:
    - expected reliability of data transmission and patient identification should be adjusted to sensitivity and risks of the telemedicine process;
    - in case of using public data transmission networks, data transmission has to be protected from accessing by a third party;
    - data consistency has to be ensured from measuring device to the centre;
- ergonomics – surfaces, utility procedures:
  - measuring unit should be suitable for independent measurements performed by the patient;
  - it should be easy-to-use, mounted by as few operation buttons as possible;
  - patients should get constantly available information on the specified method of measurement (automatic audio, video comments – downloadable or displayable on the device – chart of a few pictographs etc.).
- bi-directional communications that should work not only in the patient -> telemedicine centre direction but also in the opposite direction, including the communication from the physician towards the patient through the centre.

IV. AN OVERVIEW OF THE ARCHITECTURE (SERVER, CLIENT-SIDE CENTRE, END POINT DEVICES / SENSORS), ROLE OF COMPONENTS, COMMUNICATION AMONG THEM

Vast majority of telemedicine centres implemented worldwide including those in pilot stage are self-developed (or are derived from a small range of suppliers). The advantage of this solution is that the data communication protocols of devices are known, further development of device's functions can easily be realized thus this model can be considered as optimal during the development phase. In case of business model of extended range that covers more groups of chronic diseases, devices have to be selected from a more extensive range of manufacturers; due to the lack of mass production the costs of self-developed devices are usually high, their support is usually weaker than those of on the market. For these reasons it seems to be practical to develop an architecture, in which self-developed measuring devices and constantly wearable sensors take place but devices of a third manufacturer are also applicable in the system. In this latter case the project aimed at implementation of devices that accept the recommendations of **Continua Health Alliance** for standardized communication connections. Measuring instruments (end point devices) are always in contact with the Homecare Central Unit (HCU). HCU processes and transmits the received data to the telemedicine centre. Also the HCU provides communication connection between the patient and the medical staff.

a) End point devices and their communication

The main measuring devices used for telemedicine processes developed in the project are the followings: body weight scale, blood-pressure monitor, oxygen saturation meter, activity sensor, fall sensor, tremorometer, ECG and spirometer.

Different potential solutions were examined during the design phase of devices suitable also for home applications; among others MMS-based and then SMS-based data transmission that are able to establish teledoctor communication cheaply and simply by means of Bluetooth mobile phones. Efficient database solution was developed for data storage that provides sufficient storage capacity at reasonable costs for storing measured data by means of cheap SPI flash memories, while adjusting to the properties of those (e.g. equal distribution of sector employment).

An example for the use of the short message based telecommunication method is the SpiroTube Mobile Edition home care spirometry system with the ThorSoft Mobile Edition interface software. The secret of success with short messages in telemedicine is efficient encoding and data compression. Spirometry results can be stored in a relatively small data space, the most frequent measurement results, that the FVC curve can be stored in one simple SMS with a minimal data loss.

The SpiroTube Mobile Edition device can also send the raw respiratory flow-volume signal in real time via Bluetooth to virtually any base station, including the OKE system developed by the Zoltan Bay Foundation. The OKE system can run the spirometry data processing software, based on an extracted library that facilitates communication with the terminal device.

During the system design phase efforts were made by the technical experts of the project to achieve high level reliability and low cost of operation. There are two ways to realize these objectives. On the one hand we should avoid the use of devices that can only be installed by professionals (movement sensors fitted on walls etc.); on the other hand in case of located devices every effort has to be made to provide remote management. This can be achieved for devices developed by the consortium (HCU and end point devices) but unfortunately not for products on the market.

b) Homecare Central Unit (HCU)

Special requirements related to certain age groups have already been outlined in the description of requirements for devices section. In order to fulfill the demands of all three groups and to decrease costs by the exploitation of already available devices (computers, mobile phones) of user groups, HCU has been developed in three different ways.

Besides the installed device with full functionality, a web-based, virtual HCU has also been developed that can be accessed through a web browser with functionality decreased to a certain extent but available for considerable parts of the juvenile and middle-aged groups. Data provided by the measuring units are received by the virtual HCU via Bluetooth connection of the mobile phone or the computer.

The soul of the located HCU is an iGEP V2 embedded controller, to which end point devices are able to connect via USB, Bluetooth and ZigBee. The system is able to communicate with the telemedicine centre by both Ethernet network and GSM modem. The base unit consists of three high-intensity LED-s and three press buttons but it is able to provide appropriate sized image for dim-sighted patients by a pico-projector connected to its VGA port. In case of projected image, navigation can be performed by simple movements of hands ignoring positioning difficulties that occur due to trembling hands. Patient identification is performed by the device on the basis of PIN code and/or fingerprint identification.

Main characteristics and functions of the unit are as follows:

- communication with the telemedicine centre (Oracle BPM server (SOAP));
- communication with end point devices (ZigBee, Bluetooth, Continua);
- patient identification;
- measurements scheduled on the basis of measurement plan determined in the protocol by the telemedicine centre and customized by the physician and initiated by the patient
    - launch;
    - receive and pre-process data;
    - transmit data to the telemedicine centre;
- providing audio and video connection;
- running and displaying applications;
    - filling in forms;
    - tests.

c) Telemedicine Centre

Utilization of telemedicine processes in healthcare can contribute to their more efficient accomplishment in total: utilization of expensive medical resources can be decreased; meanwhile beyond preserving efficiency level of treatments value creating ability of active population can be increased by decreasing the amount of wasted working hours; moreover quality features of care experienced by the patients can be improved. In order to realize the positive effects mentioned above, we must have **process and device systems** that fit to new methods and fulfill the requirements (security, reliability, ergonomics etc.) mentioned in the previous chapters at system level.

One of the central elements of the eHealth8 project was the development of the processes (telemedicine protocols) that determine the methods and rules of telemedicine applications within the given professional field for particular groups of diseases. Beyond the fact that accurate application descriptions created this way contribute to more carefully considered and more efficient utilization of resources; they play key role in telemedicine because accomplishment of virtual patient-physician relationships separated in time and place is unimaginable without prior, detailed determination of its rules. Benefits expected from the utilization of telemedicine processes can only be realized if processes are controllable in accordance with the previously determined rules. So system level requirements concerning to the central system described in the previous chapters are supplemented by new ones. The most important new requirements are:

- support standardized representation of telemedicine protocols that enable their unambiguous IT interpretation and processing;
- support controllable and safe performance (run) of telemedicine processes;
- medical professional regulations and realization in running environment have to be accessible at user level too (not only for the software developer);
- ability to get complied with the protocol rules and give quality assurance from the centre as well;
- concerning a given protocol, the ability to easily change the type of measuring device, integrate new ones;
- since telemedicine service is a part of the care process; the ability for appropriate data transfer with in- and outpatient management systems;
- due to data exchange the obligation of long term storage of data (legal state) that is not allowed to be performed in the centre;
- endpoint devices and HCU are connected to the centre via Internet, high level of safety and reliability without the need of any technical assistance are fundamental.

Standardized representation and IT interpretation focus on knowledge share and portability; partly in case of separately developed systems, partly among special medical and IT fields.

Endpoint devices do not connect directly to the centre; they receive data solely through the HCU. Consequently HCU accomplishes direct, individual physical attachment of different endpoint devices to the system. The telemedicine processes running in the centre control the endpoint devices through an abstraction layer provided by the HCU; so they have to be made independent from particular types of devices. This architecture provides simple integration of new devices and device independency of protocols. HCU-s communicate with the centre via Internet (mobile or wired), by SOAP calls, standardized XML messages.

Even if many components of the system (centre, CHU or the device itself) are able to operate autonomous operation and to make decisions on different levels during measurements, where are data evaluation and making decisions practical to perform? Pre-processing of data is performed on device and CHU level. The CHU downloads technical and physiological threshold values typical to the particular measurement. Evaluation for these values is locally performed as values beyond threshold values can refer to false measurement or device failure, so in any case repetition of measurements is necessary in order to decrease and validate the number of false values. However, the evaluation regarding every process runs in the centre that is important regarding to quality insurance, as the entire sequence can be tracked and audited at the same place.

V: EFFICIENCY AND BUSINESS MODEL

The project is currently considering the possibilities of the integration into the present healthcare financial system.

In accordance with its original objective, the project intends to support methods that are the most economical for the patient, the employer, the national economy, the society and the care system in total and/or result in improvement of quality of life for patients. This does not exclude the possibility that its particular cost elements are more expensive but it is a more economical solution in total.

According to the economic model, the telemedicine service provider is a participant of the care system. Basically, the new methods developed on the basis of telemedicine for the community by the care system create new values; therefore the expenditures of value creation can be covered by them. In this respect the telemedicine service provider is an element of the system. Theoretically the equivalent of benefits (profit) created in the system can be transferred to value producers different way, as there exist relationships among participants at different levels and of different nature.

VI. SUMMARY

During project implementation the consortium examined thirty different medical professional guidelines; from these guides three telemedicine protocols were developed by the cooperation of medical and technical experts. On the basis of economical considerations, three actual realizations were performed in this project (asthma diagnostics and therapy monitoring, stroke homecare and rehabilitation). Realization of asthma diagnostics by means of telemedicine devices makes possible to identify the disease without getting the patient travel to or stay in county centre for a longer period. According to the medical guide, inhalation and exhalation of the patient should be monitored for at least ten days, by this method patients can get over this process at their home, living their normal way of life instead of staying in hospital. Asthma therapy monitoring serves the patient's personal therapy. Stroke rehabilitation realizes communication and co-operation among specialists, rehabilitation experts and patients allowing drastic cost reduction and increase of quality of life and sense of comfort of the patient.

The final objective is to have these processes accepted by healthcare representatives and to achieve that these processes are integrated into the national finance system that would make possible to widen service availability for those in need.

REFERENCES

[1] Telemedicine in the NHS for the millennium and beyond; Simon Wallace, Jeremy Wyatt, Paul Taylor; Postgrad. Med. J. 1998, 74:721–728 © The Fellowship of Postgraduate Medicine, 1998

[2] Csaba A. Szabó, András Jávor: Telemedicine Applications, Hiradástechnika, Vol. LXIII, 2008/2, pp. 9-14 (in Hungarian).

[3] László Daragó, András Jávor: How to make an eProtokoll?, Egészségügyi gazdasági szemle, 48/3, June 2010, pp. 32-3 (in Hungarian).

[4] László Daragó, Csaba Engi, István Pesti, Dezső Vass˙ Telemedicine: ICT-based health care service, Part I. System concept and architecture. Híradástechnika, Vol. LXV, No 4, 2010, pp. 33-37 (in Hungarian).

[5] László Daragó, András Jávor: How to make eProtocol?
Proceedings of The International eHealth, Telemedicine and eHealth ICT Forum for Education, Networking and Business, April 14-16, 2010 Luxembourg, pp. 203-206.

[6] Technology and Informatics - Current Principles and Practices of Telemedicine and e-Health, 2008, IOS Press, Amsterdam, ISBN 978-1-58603-806-9.

[7] eHealth – Regional case studies, http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=2892, uploaded: 6 October 2006, accessed: 8 May 2011.

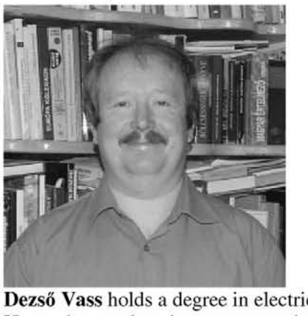
Authors:

László Daragó graduated from Kossuth Lajos University of Sciences in 1984 as a physicist, received his PhD in 2006 at Mathematics and IT PhD Department of the University of Debrecen. He is Associate Professor of Health Informatics Institute of Semmelweis University. His main areas of research are telemedicine, artificial intelligence, designing and planning of information systems, pixelisation, DRG.

Csaba Engi graduated from the Technical University of Budapest in 1995 as electrical engineer. He received his MBA degree from Buckighamshire New University (BUCKS) in 2008. He has been with Answare Ltd. since 1998. During his career he has obtained experience in integration of infocommunication infrastructure and systems, planning of integration activity, and carrying out projects in many areas: electronic message management, infrastructure management, IT safety, medical IT. Presently he is coordinating the tendering and competition activity of Answare Ltd.

György Ferenczi graduated from the Technical University of Budapest in 1999 as an electrical engineer. He performed post-graduate studies at the State University of New York, at the Tampere University of Technology and later at the Helsinki University of Technology. Currently he is one of the directors of a medical technology company.

István Pesti graduated from the Technical University of Budapest in 1991, participated in MBA education between 1992 and 2002, then studied at MIT in Boston in 2006-2007. Presently he is the managing director of P-Invent Ltd. His professional experience includes planning, coordinating, managing, financing different projects.



Dezső Vass holds a degree in electrical and medical engineering. He used to work as instrument engineer of Semmelweis Hospital and Clinic in Miskolc. He is currently with the Bay Zoltán Foundation for Applied Research, Institute for Logistics and Production Systems, where he leads the IT department. The priority of his research is telemedicine, more closely teleradiology and development of telemonitoring systems.

Equalization of Multicarrier Cognitive Radio Transmission Over Multipath Channel with Large Delay Spreads

(Short Paper)

Zsolt Kollár and Péter Horváth

***Abstract*—This paper investigates the channel equalization problem in opportunistic white space communication systems, with an emphasis on transmission over channels with high delay spread. We consider the equalization problems of Orthogonal Frequency Division Multiplexing (OFDM) and three alternative multicarrier modulation schemes, namely DFT-spread OFDM (DFTS-OFDM), Constant Envelope OFDM (CE-OFDM) and Filter Bank Multicarrier (FBMC). After a brief description of these schemes, we show their performance assuming frequency domain per subcarrier minimum mean square error (MMSE) channel equalization. The simulations are performed using the channel parameters specified in the IEEE 802.22 standard. We show that bit error rate (BER) performance of FBMC and DFTS-OFDM are comparable to OFDM in frequency selective channels which make them a very strong candidate for cognitive applications.**

***Index Terms*—Channel equalization, Cognitive Radio, multipath propagation, OFDM, modulation schemes.**

I. INTRODUCTION

As some licensed bands become crowded and white spaces [1] are left behind by the ceased analog TV broadcast, the need for better utilization of the existing, under-utilized spectrum becomes evident. Cognitive radio-based (CR) opportunistic exploitation of spatial and temporal white spaces is considered as a feasible approach to improve the spectral efficiency and to introduce new services into legacy bands.

Due to strict regulations and high requirements on analog components, the characteristics for the opportunistic transceiver are crucial.

Today, OFDM is considered as the most widespread multicarrier modulation scheme for wireless data transmission. It is used in many broadcasting and communication systems such as DVB, DAB and certain types of IEEE 802.11 WLAN. Low-complexity demodulation and modulation can be performed by the Fast Fourier Transform (FFT) and the Inverse FFT (IFFT) operations, respectively. With the use of the cyclic prefix (CP), the channel equalization can be implemented effectively in the frequency domain. Nevertheless, this scheme exhibits some drawbacks. Due to the large dynamic range of the transmitted signal, OFDM is highly sensitive to nonlinear distortions caused by e.g. power amplifiers. This nonlinearity might degrade the system's performance because of the induced in-band and out-of-band radiation. OFDM system performance can strongly depend on the frequency mismatch of the transmitter and receiver local oscillators, requiring very robust synchronization techniques. Moreover, without additional filtering of the transmitted signal, the out-of-band radiation is considered moderate, which is a major drawback in the opportunistic context, where significant adjacent-channel leakage might lead to interference to incumbent users.

There are many issues regarding CR-scenarios which cannot be met by OFDM. This is why other multicarrier schemes have become a point of interest. In this paper we focus on three alternatives:

• DFT-spread OFDM (DFTS-OFDM) [2],
• Constant Envelope OFDM (CE-OFDM) [3],
• Filter Bank Multicarrier (FBMC) [4], [5].

Each of these schemes has advantages over OFDM in some aspects. [6] gives an overview on their sensitivity to residual synchronization errors and non-linear impairments. In this paper we intend to compare these schemes in multipath channel propagation scenarios. Also the corresponding equalization algorithms will be compared and evaluated.

The paper is organized as follows. In Section II, each of these modulation schemes are explained in details. The main signal processing blocks are explained separately for the transmitted and for the receiver. In Section IV the channel equalization methods are explained. In Section V the simulation results of the four systems are compared via bit error rates over various multipath channels. Also the applied IEEE 802.22 [7] channel parameters are presented together with the system parameters which are used for the simulations. Finally the conclusions are drawn from the results of the simulation.

Manuscript received May 15, 2011. The research leading to these results was derived from the European Communitys Seventh Framework Programme (FP7) under Grant Agreement number 248454 (QoSMOS).

Zsolt Kollár is with the Budapest University of Technology and Economics, Department of Measurement and Information Systems (e-mail: kollarzs@mit.bme.hu).

Péter Horváth with the Budapest University of Technology and Economics, Department of Broadband Infocommunications and Electromagnetic Theory (e-mail: hp@mht.bme.hu).

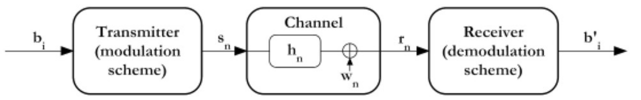


Fig. 1. Baseband digital model of the transceiver.

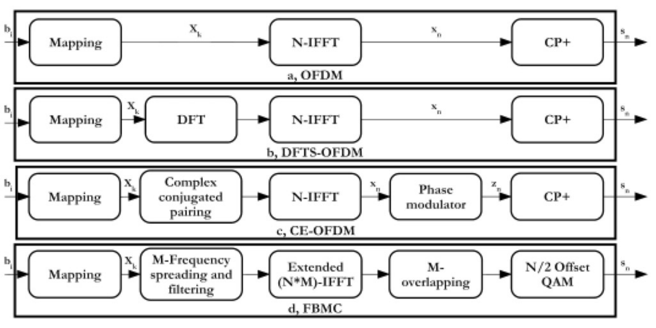


Fig. 2. Block diagram of the transmitters of the four modulation schemes: OFDM (a), DFTS-OFDM (b), CE-OFDM (c) and FBMC (d).

II. DISCRETE BASEBAND TRANSCEIVER MODEL

Assuming ideal analog components – neglecting their impairments – and perfect knowledge of the multipath channel characteristics, the analog transmission chain can be modeled in the digital baseband. In this section the applied discrete baseband model for the transceiver chain is explained. The block diagram of the model can be seen in Fig. 1.

First the transmitted bit stream b_i is converted to a discrete transmission signal s_n according to the applied modulation scheme. The multipath channel is modeled as a discrete FIR filter h_n with the length of *L*. After filtering the transmitted signal the Additive White Gaussian Noise (AWGN) w_n is added. Finally, the received signal r_n is demodulated with the demodulation scheme corresponding to the transmitter's modulation forming the received bitstream $b_i^{'}$. The received signal r_n can be written as

$$r_n = s_n * h_n + w_n \,. \tag{1}$$

III. MULTICARRIER MODULATION SCHEMES

In this section we give a detailed explanation of the four investigated modulation schemes. Although OFDM has a well-known architecture, we give a longer explanation, because the other schemes are based on it. For each modulation we first summarize the procedure of the signal generation in the transmitter then the demodulation of the received signal in the receiver. The block diagram of the four transmitters is depicted in Fig. 2. The signal processing blocks of the receiver can be seen in Fig. 3. One can see that the signal processing differs only slightly from that of the OFDM.

A. OFDM

Transmitter

The basic idea of OFDM is to modulate *N* complex carriers. This can be performed effectively using the IFFT. First the bitstream b_i is transformed to complex Quadrature Amplitude Modulation (QAM) constellation symbols X_k selected from the constellation alphabet *C*. The selected symbols are used to modulate the subcarriers via N-IFFT to form samples of a time domain OFDM symbol from the transmit signal s_n as

$$x_n = \sum_{k=0}^{N-1} X_k e^{j\frac{2\pi}{N}nk}, \qquad n \in 0 \ldots N-1 \,. \tag{2}$$

In real applications, some subcarriers are left unmodulated or carry reference signals. Before transmission each OFDM symbol is extended with a CP which helps to reduce multipath propagation effects so the compensation can be performed directly in the frequency domain. The CP is usually a repeated part of *P* samples taken from the end of the time domain symbol and inserted to the beginning resulting in a symbol length of *N* +*P* samples. The transmitted signal s_n is formed from these extended OFDM symbols.

Receiver

The received signal r_n as explained in Section II can be expressed according to equation (1). As long as the CP is longer than the channel delay spread, for one OFDM symbol after removing the CP, the following frequency domain equation is valid:

$$Y_k = X_k H_k + W_k, \qquad k \in 0 \ldots N-1 \,. \tag{3}$$

Here Y_k, X_k, H_k and W_k are N-FFT of the signals y_n, x_n, h_n, and w_n respectively.

B. DFTS-OFDM

Transmitter

The transmitter of the DFTS-OFDM differs only by an additional block compared to OFDM. An additional DFT is applied to the complex modulation symbols. It can be considered as a spreading of the information through the available subcarriers. In strict sense, this can be considered as a single carrier modulation scheme. This extra spreading block will lead to many advantages such as lower peak-to-average power ratio (PAPR), which is desired when nonlinearity is present in the system. It can also be considered as a diversity leading to a gain over frequency selective channels over OFDM systems without interleaving and coding – which is of little concern in modern wireless systems.

Receiver

The receiver block are almost the same as for OFDM, the difference is only the extra IDFT operation which despreads the information after channel equalization. The remaining blocks of the receiver are the same.

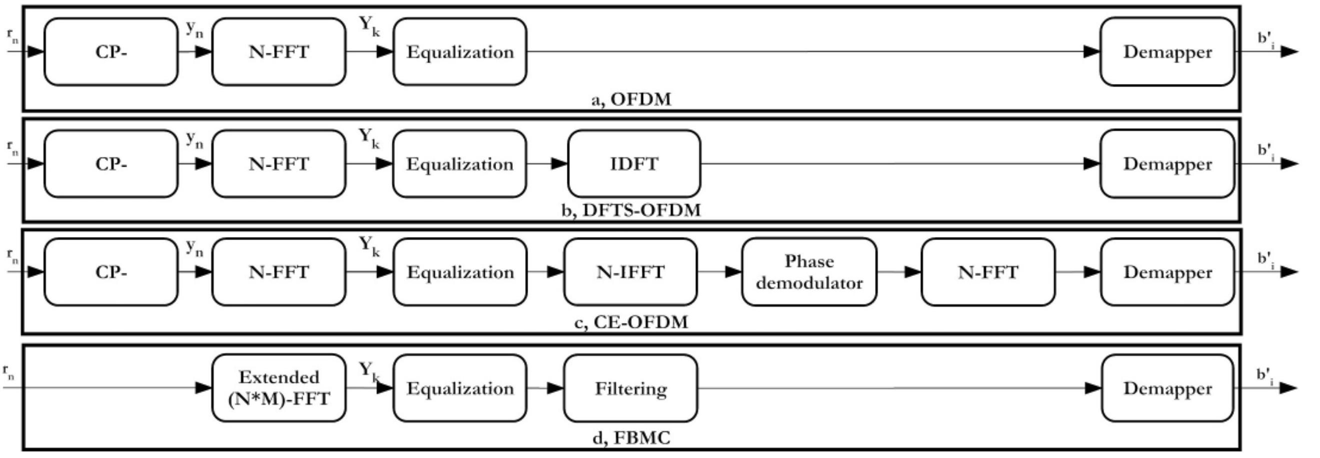


Fig. 3. Block diagram of the receivers of the four modulation schemes: OFDM (a), DFTS-OFDM (b), CE-OFDM (c) and FBMC (d).

C. CE-OFDM

Transmitter

The main aim of CE-OFDM is to achieve a constant envelope of the transmitted signal. The first step is to modulate the subcarriers in a way to achieve a real valued signal, which can be done by modulating the subcarriers by complex conjugated pairs $X_k = X^*_{N-k}$. This way the output of the IFFT will be a real valued signal x_n. This signal will be used as the input of a phase modulator

$$z_n = e^{j2\pi m x_n}, \quad n \in 0 \ldots N-1. \tag{4}$$

where m is the modulation index. The transmitted signal s_n is then generated from z_n with the additional CP. One of the advantages of CE-OFDM is that it has a constant envelope, on the other hand the data rate is halved, due to the complex conjugated pairing. Also the spectral behavior of the transmitted signal is highly dependent on the modulation index *m* and a large DC component will appear which is difficult to handle in the analog chain.

Receiver

The receiver for CE-OFDM is more complex compared to OFDM. First, frequency domain equalization is performed, similar to OFDM. Then, the equalized signal is transformed back to time domain where the phase demodulation can be performed. After the phase demodulation, the N-FFT is applied to retrieve the complex modulation values.

D. FBMC

Transmitter

FBMC systems are derived from the orthogonal lapped transforms [8] and filter bank theory [9]. Similar to OFDM the bits are first mapped to complex constellation symbols X_k, then the frequency domain data is spread over *M* subcarriers forming a subband, then it is filtered by a prototype filter of the k^{th} subband $F_k(z)$ which is designed so that it fulfills the Nyquist criterion. The real parts of X_k are modulated by a cosine filter bank $F_k^c(z)$, where only the even-index sub-bands are used and the imaginary parts are modulated by a sine filter bank $F_k^s(z)$ where only the odd-index sub-bands are modulated. In order not to lose data rate an offset of half of the original symbol duration *N*=2 is applied – similarly to offset quadrature amplitude modulation technique. In this case the imaginary parts of X_k are modulated with the cosine filter bank and the real parts are modulated with the sine filter bank respectively. Due to the design, the inter-symbol-interference (ISI) between the filters is negligible.

In FBMC applications these filter bank structures are implemented in a computationally efficient manner using an N-IFFT and a polyphase network [9]. The filter bank output provides a symbol length of *N*M* samples. For example if *M* = 4 then 4 consecutive FBMC symbols overlap. The resulting transmitted signal is the sum of overlapping FBMC symbols generated by the filter banks.

It can be observed that no CP is used, which will lead to ISI in the presence of frequency-selective multipath propagation. On the other hand the FBMC symbol length is *M* times larger which decrease the effect of the ISI in comparison with CP-OFDM.

Receiver

The FBMC receiver is rather simple. Each FBMC symbol is turned to the frequency domain using an extended FFT for *N*M* samples. The equalization is done in the frequency domain. After equalizing all subcarriers in the subband separately each subband is filtered with the corresponding analysis filterbank to get an estimate of the transmitted constellation value X_k which can be fed to the demapper.

IV. CHANNEL EQUALIZATION

Zero forcing (ZF) is known to be the simplest method for channel equalization in the frequency domain. We simply assume that the received noise is zero in equation (3) for the received frequency domain OFDM symbol, so the transmitted complex constellation value on the k^{th} subcarrier can be simply calculated as $\hat{X}_k^{ZF} = Y_k / H_k$. The MMSE technique gives a better result if we take the noise also into account. The problem of ZF occurs when H_k is small, the noise values will be also amplified. The equalization coefficient for the k^{th} subcarrier is calculated according to [10] as

$$\frac{1}{H_k^{MMSE}} = \frac{H_k^*}{\left(\left|H_k\right|^2 + \frac{N_0}{E_s}\right)}, \quad (5)$$

where N_0 is the noise power and E_s is the signal power. It can be seen that with small N_0 / E_s compared to $\left|H_k\right|^2$ values the MMSE solution is equal to the ZF. In case of an FMBC system equation (5) has to be modified according to the ISI if we use a per subcarrier equalization scheme similar to [11] as

$$\frac{1}{\hat{H}_k^{MMSE}} = \frac{H_k^*}{\left(\left|H_k\right|^2 + \frac{N_0 + I}{E_s - I}\right)}, . \quad (6)$$

where I is the power of the ISI, for which we present the following equation:

$$I = E_s \sum_{n=0}^{L-1} \frac{n}{NM} \left|h_n\right|^2 . \quad (7)$$

So we can write for the MMSE estimate of the k^{th} subcarrier

$$\hat{X}_k^{MMSE} = \frac{Y_k}{\hat{H}_k^{MMSE}} . \quad (8)$$

For OFDM systems this estimate can be directly fed to the demapper, for the other schemes some extra operations have to be performed before demapping, which were explained previously in Section 3.

A detailed analysis regarding equalization techniques for FBMC can be found in [12], where also a novel iterative decision feedback equalization technique is presented. The main idea is to iteratively minimize the ISI achieving a better BER performance. In our simulation we will present also the results of this technique for FBMC systems.

V. Simulation Results

To compare the previously described equalization scheme for the modulation schemes, simulations were performed. The simulation parameters used for the various systems are summarized in Table I. The parameters of the IEEE 802.22 channel profiles B and C can be found in Table II. [7]. In order to obtain comparable bit error rates for the multicarrier systems we define the SNR normalized to one bit energy.

$$SNR_{dB} = 10 \log_{10}\left(\frac{E_s}{N_0}\right) = 10 \log_{10}\left(\frac{E_b N_c D}{(N+P) N_0}\right), \quad (9)$$

TABLE I
Simulation parameters for FBMC, OFDM, DFTS-OFDM and CE-OFDM system

Parameter	FBMC	OFDM	DFTS-OFDM	CE-OFDM
Bandwidth	8 MHz			
CP	0	1/16 & 1/4		
N	2048			
M	4	1		
Modulation *(D)*	16-QAM (4)			
Modulated subcarriers/subbands	1536			

TABLE II
Channel paramters for IEEE 802.22 B and C channel profiles
(excess delay and relative amplitude)

Profile B	**Path 1**	**Path 2**	**Path 3**	**Path 4**	**Path 5**	**Path 6**
Excess delay	-3 µs	0 µs	2 µs	4 µs	7 µs	11 µs
Relative amplitude	-6 dB	0 dB	-7 dB	-22 dB	-16 dB	-20 dB
Profile C	**Path 1**	**Path 2**	**Path 3**	**Path 4**	**Path 5**	**Path 6**
Excess delay	-2 µs	0 µs	5 µs	16 µs	24 µs	33 µs
Relative amplitude	-9 dB	0 dB	-19 dB	-14 dB	-24 dB	-16 dB

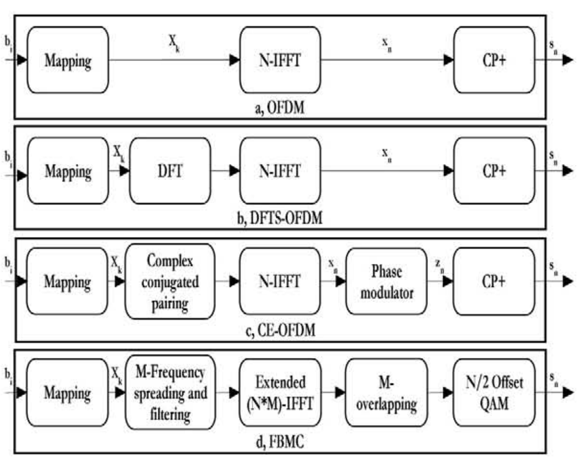


Fig. 4. BER performance of the four systems over a channel with no multipath propagation and over Channel B.

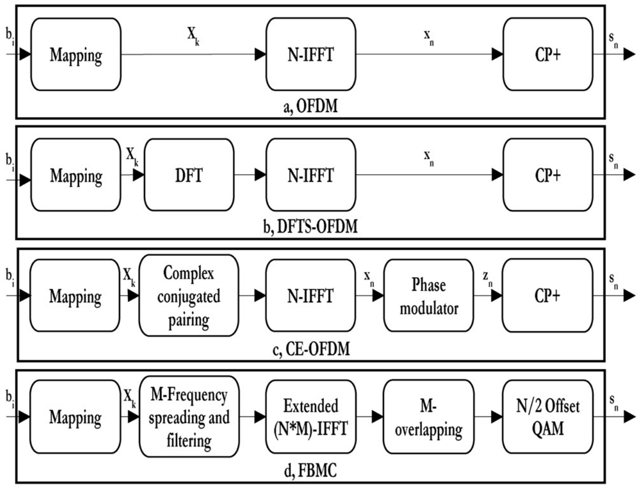


Fig. 5. BER performance of the four systems over a channel with no multipath propagation and over Channel B.

where N is the number of the subcarriers/subbands available and N_c is the number of subcarriers/subbands used. E_b is the bit energy and *P* is the length of the CP and *D* is the number of bits transmitted by one subcarrier/subband.

In the first simulation we compared the BER results of the four schemes for channel profile B together with the case when no multipath propagation is present. The resulting BER curves can be seen in Fig. 4. In multipath-free environment FBMC outperforms all other techniques because it does not apply any CP, resulting in higher data rate. The results for DFTS-OFDM and conventional OFDM are the same. CE-OFDMs performance is the worst, due to the fact that only half of the available bandwidth is used for data transmission. On the other hand if Channel B is introduced, DFTS-OFDM has the best performance due to the frequency spreading. FBMC still slightly outperforms OFDM, CE-OFDM still has the worst BER values under 20 dB SNR.

The simulation results of the second scenario can be seen in Fig. 5. Here the BER curves for channel C are evaluated for a CP which is shorter than the largest excess delay and also with a CP which is longer. As FBMC does not use CP, the results for the third iteration of a novel iterative compensation technique, presented in [12], are shown together with the per subcarrier MMSE equalization technique. We can conclude that until the CP is sufficiently long DFTS-OFDM has the best performance. The BER values of the other three modulations show a similar behavior as shown in Fig. 4. Again, CE-OFDM can only compete above an SNR value of 20 dB. If the CP is shorter than the maximum excess delay, ISI is introduced. It leads to an error floor for all four modulations. With the previously mentioned iterative technique, BER results of FBMC can be further improved. In this scenario OFDM and DFTS-OFDM perform the worst, both reaching a bit error floor of $2*10^{-3}$ in comparison with CE-OFDM which reaches $5*10^{-4}$.

VI. CONCLUSION

In this paper we have presented four modulation schemes which can be promising candidates for future cognitive radio applications. We have described their main signal processing blocks and we have investigated a simple frequency domain MMSE equalization technique. The systems' performances were compared via simulations for various channels defined by IEEE 802.22. The results show that both DFTS-OFDM and FBMC are very strong candidates for cognitive radio physical layer. CE-OFDM is only worth consideration if dominant nonlinearities are present in the transceiver chain.

ACKNOWLEDGMENT

The research leading to these results was derived from the European Communitys Seventh Framework Programme (FP7) under Grant Agreement number 248454 (QoSMOS).

REFERENCES

[1] S. J. Shellhammer, A. K. Sadek, and W. Zhang, "Technical challenges for cognitive radio in the TV white space spectrum," *Information Theory and Applications Workshop*, pp. 323–333, Feb. 2009.G. O.

[2] D. Galda and H. Rohling, "A low complexity transmitter structure for OFDM-FDMA uplink systems," in *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, vol. 4, 2002, pp. 1737–1741.

[3] S.C. Thompson, A.U. Ahmed, J.G. Proakis, J.R. Zeidler, and M.J. Geile, "Constant envelope OFDM," *Transactions on Communications*, vol. 56, no. 8, pp. 1300–1312, Aug. 2008.

[4] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," vol. 50, no. 5, pp. 1170–1183, 2002.

[5] F. Schaich, "Filterbank based multi carrier transmission (FBMC) – evolving OFDM: FBMC in the context of WiMAX," in *Wireless Conference (EW), 2010 European*, 2010, pp. 1051–1058.

[6] Zs. Kollár and P. Horváth, "Modulation schemes for cognitive radio in white spaces," *Radioengineering*, vol. 19, no. 4, pp. 511–517, Dec. 2010.

[7] IEEE 802.22-05/0055r7, "WRAN channel modeling," E. Sofer and G. Chouinard, Aug. 2005.

[8] H. S. Malvar, "Extended lapped transforms: Properties, applications, and fast algorithms," *IEEE Trans. on signal processing*, vol. 40, no. 11, pp. 2703–2714, Nov. 1992.

[9] P. P. Vaidyanathan, *Multirate systems and filter banks*. Prentice-Hall Inc., 1993.

[10] J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," vol. 2, pp. 815–819, Jul. 1995.

[11] T. Ihalainen, T. H. Stitz, M. Rinne, and M. Renfors, "Channel equalization in filter bank based multicarrier modulation for wireless communications," *EURASIP Journal Appl. Signal Process.*, 2007.

[12] Zs. Kollár, G. Péceli, and P. Horváth, "Iterative decision feedback equalization for FBMC systems," in *Proceedings of IEEE First International Conference on Advances in Cognitive Radio, COCORA 2011*, Apr. 2011.

Authors:

Zsolt Kollár was born in Budapest, Hungary, 1983. He received his MSc. in electrical engineering from the Budapest University of Technology and Economics in 2008. He is currently working toward his Ph.D. degree.

He has been a member of the Rohde & Schwarz Reference Laboratory since 2007. His current research and study interests lie in the area of digital signal processing in wireless communication, especially regarding channel equalization and synchronization.

Péter Horváth was born in Budapest, Hungary, 1978. He received his MSc. in electric engineering from the Budapest University of Technology and Economics in 2002. He received his Ph.D. degree in 2011 at the Budapest University of Technology and Economics.

He is currently an assistant lecturer at the Department of Broadband Infocommunications and Electromagnetic Theory. His research activities include MIMO systems, channel modeling and aspects of the physical layer for cognitive radio.

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on
http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2,
"Template and Instructions on How to Create Your Paper".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *index terms (Keywords)*. For the final version of accepted papers, please send the *short cvs* and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984.

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

Authors are requested to send their manuscripts via electronic mail or on an electronic medium such as a CD by mail to the Editor-in-Chief:

Csaba A. Szabo
Dept. of Telecommunications
Budapest University of Technology and Economics
2 Magyar Tudosok krt.
Budapest 1117 Hungary
szabo@hit.bme.hu

2ND INTERNATIONAL CONFERENCE ON COGNITIVE INFOCOMMUNICATIONS, COGINFOCOM 2011

July 7-9, 2011 Budapest, Hungary

Web: http://www.hte.hu/coginfocom2011

ORGANIZERS

ETO COM · HUNOROB · BME Department of Telecommunications and Media Informatics (BME-TMIT) · MTA SZTAKI Computer and Automation Research Institute, Hungarian Academy of Sciences · hte

General Chair

Péter Baranyi, *3D Internet Communication and Control Lab. · The Consortium of MTA SZTAKI and BME, Hungary*

Honorary Chairs

Hideki Hashimoto, *iSpace Research · The University of Tokyo, Japan*

Csaba Pléh, *Cognitive Science · BME, Hungary*

Gyula Sallai, *Infocommunication Strategy and Regulation · BME, Hungary*

Liaison Chair

Imre J. Rudas, *IEEE Hungarian Section Chair · Óbuda University, Hungary*

Area Chairs *(Preliminary list)*

Martin Buss, *Cognitive Technical Systems – Autonomous and Cognitive Robotics · Technical Univ. of Munchen, Germany*

Nick Campbell, *Multimodal Speech Communication Technology · Trinity College Dublin, Ireland*

Kerstin Dautenhahn, *Human Robot Interaction · University of Hertfordshire, UK*

Tom Gedeon, *Australian National University, Australia*

Károly Hercegfi, *Cognitive Ergonomics, Human Factors of Infocommunication · BME, Hungary*

Kristiina Jokinen, *Intelligent Interfaces and Language Technology · University of Helsinki, Finland*

Péter Korondi, *Cognitive Features in Mechatronics · BME, Hungary*

Ilona Kovács, *Cognitive Science · BME, Hungary*

Joo-Ho Lee, *Ritsumeikan University, Kyoto, Japan*

Gábor Magyar, *Cognitive Content Management · BME, Hungary*

Ádám Miklósi, *Ethology, (Etho-communication) · ELTE, Hungary*

Géza Németh, *Multimodal Speech Interaction & Applications · BME, Hungary*

Mihoko Niitsuma, *Spatial Memory in iSpace · Chuo University, Tokyo, Japan*

Primoz Podrzaj, *Automated Processing · University of Ljubljana, Slovenia*

Claudiu Pozna, *Cognition and Robotics · Széchenyi István Univ., Hungary, University Transilvania of Brasov, Romania*

Péter Sincak, *Computational Intelligence · Technical University of Kosice, Slovakia*

Trygve Thomessen, *Production Engineering · PPM AS, Norway*

Klára Vicsi, *Verbal and Non-verbal Communication · BME, Hungary*

Local Organizing Committee

Gábor Magyar, *BME, Hungary*
Péter Nagy, *HTE, Hungary*
Géza Németh, *BME, Hungary*
Anna Szemereki, *BME, Hungary*
Mária Tézsla, *HTE, Hungary*

The 2nd International Conference on Cognitive Infocommunications (CogInfoCom2011) will be held in Budapest, Hungary, at the premises of the Budapest University of Technology and Economics, in July 7-9, 2011.

Location: **Building of Informatics, Budapest University of Technology and Economics**
Magyar tudósok krt. 2., H-1117 Budapest, Hungary

Cognitive infocommunications (CogInfoCom) investigates the link between the research areas of infocommunications and cognitive sciences, as well as the various engineering applications which have emerged as the synergic combination of these sciences.

The primary goal of CogInfoCom is to provide a systematic view of how cognitive processes can co-evolve with infocommunications devices so that the capabilities of the human brain may not only be extended through these devices, irrespective of geographical distance, but may also interact with the capabilities of any artificially cognitive system. This merging and extension of cognitive capabilities is targeted towards engineering applications in which artificial and/or natural cognitive systems are enabled to work together more effectively.

Patrons

New Hungary Development Plan · NFÜ · eea grants Iceland Liechtenstein Norway

Technical Co-sponsors

ISPace · IEEE Hungary Section

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its more than 1300 individual members, the Scientific Association for Infocommunications (in Hungarian: Hírközlési és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society. HTE is corporate member of International Telecommunications Society (ITS).

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **DR. GÁBOR HUSZTY** • *ghuszty@entel.hu*

Secretary-General: **DR. ISTVÁN BARTOLITS** • *bartolits@nmhh.hu*

Managing Director, Deputy Secretary-General: **PÉTER NAGY** • *nagy.peter@hte.hu*

International Affairs: **ROLLAND VIDA, PhD** • *vida@tmit.bme.hu*

Addresses

Office: H-1055 Budapest, V. Kossuth Lajos square 6-8, Room: 422.

Mail Address: 1372 Budapest, Pf. 451., Hungary

Phone: +36 1 353 1027, Fax: +36 1 353 0451

E-mail: *info@hte.hu*, Web: *www.hte.hu*