NARMA-L2 Control of a Nonlinear Half-Car Servo-Hydraulic Vehicle Suspension System

Jimoh Pedro, John Ekoru

School of Mechanical, Aeronautical and Industrial Engineering, University of the Witwatersrand, Private Bag 03, WITS2050, Johannesburg, South Africa
e-mail: Jimoh.Pedro@wits.ac.za; John.Ekoru@students.wits.ac.za

Abstract: In this paper, the performance of a nonlinear, 4 degrees-of-freedom, servo-hydraulic half-car active vehicle suspension system is compared with that of a passive vehicle suspension system with similar model parameters. The active vehicle suspension system is controlled by an indirect adaptive Neural Network-based Feedback Linearization controller (NARMA-L2). Hydraulic actuator force tracking is guaranteed by an inner Proportional+Integral+Derivative-based force feedback control loop. The output responses of the vehicles are presented and analyzed in the frequency and time domains, in the presence of model uncertainties in the form of variation in vehicle sprung mass loading. The results show that the NARMA-L2-based active vehicle suspension system performed better than the passive vehicle suspension system within the constraints.

Keywords: Neural Networks; NARMA-L2; PID Control; Nonlinear Vehicle Suspension System; Hydraulic Actuator Dynamics; Model Uncertainty

1 Introduction

Vehicle Suspension Systems (VSS) are highly nonlinear uncertain systems [1]. In VSS design, a tolerable compromise between conflicting vehicle parameters (ride comfort, handling and road holding) must be found, within the limits of suspension travel [1; 2]. VSS design is an active area of research involving the development of Passive Vehicle Suspension Systems (PVSS), Semi-Active Vehicle Suspension Systems (SAVSS) and Active Vehicle Suspension Systems (AVSS) [3; 4]. The rapid progress within this field can be attributed to advances made in optimal control and in computer processing power and to the increasing affordability of actuators and sensors [2; 3]. Compared with PVSS and SAVSS, AVSS are better at addressing the VSS design trade-off; PVSS and SAVSS can only dissipate forces incident on the VSS, whereas AVSS introduce forces into the VSS. However, AVSS are more complex and consume greater amounts of energy [5].

Examples of control techniques succesfully applied to AVSS include: PID control [6], optimal control [2; 7], robust control methods such as H_2 [8], H_∞ [9], H_2/H_∞ [10], linear parameter varying (LPV) [11], nonlinear control techniques like sliding mode control (SMC) [12], backstepping control [13] and feedback linearization (FBL) [14]. Intelligent control techniques such as fuzzy logic control (FLC) [15] and various Neural Network (NN)-based control methods [1; 16; 17] have also been applied.

Neglecting actuator dynamics in the study of AVSS has restricted the amount of experimental validation possible in the past [18; 19]. In research works where actuator dynamics are considered, hydraulic actuators are most often selected because they have a fast response time, high stiffness, a superior power-to-weight ratio, low cost and low heat dissipation during periods of sustained force generation, compared to other actuators [1; 17]. However, hydraulic actuators are highly nonlinear and prone to chattering in AVSS applications [18]. Backpressure caused by tight coupling between motion of the vehicle body and actuator force generation requires the use of hydraulic actuator force feedback [19]. Actuator force feedback improves vehicle ride comfort and road holding by stabilizing the hydraulic actuator, and guaranteed desired forces levels are attained. In the literature, force control is commonly applied utilizing a system of control loops; i.e., actuator force feedback control in the inner loop and sprung mass displacement control in the outer loop [14; 18; 19].

In industry, PID control is the most common control method employed (it is simple in structure, and tuning is straightforward) [20]. PID control has been applied both to benchmark AVSS performance [17; 21] and as the main control method [6]. However, its lack of robustness to parameter variation and the requirement of high loop gains have motivated research with the aim of enhancing PID controller performance for AVSS applications [22; 23; 24].

Highly nonlinear systems can be transformed into linear systems to enable the application of linear control methods by making use of FBL [25]. The implementation of FBL may prove difficult as feedback of all system states is required, and issues may arise regarding the robustness of the linearized system. This is solved by a combination of FBL with intelligent control methods [1]. In Buckner et al. [16], radial basis function (RBF) NNs were trained to estimate the nonlinear suspension damping and spring forces in order to cancel out nonlinear suspension dynamics by using NNFBL and to control a quarter-car AVSS with a linear electro-mechanical force actuator. Offline training of the NNs was performed using input-output data obtained from a quarter-car suspension test-rig. The RBF NN weights were updated online. Pedro and Dahunsi [1] proposed a Multi Layer Perceptron (MLP) NNFBL controller for a nonlinear, electro-hydraulic quarter-car AVSS. MLP NNs were trained to approximate the nonlinear functions using the Levenberg-Marquardt (LM) algorithm. Compared with a PID-controlled AVSS, the NNFBL-controlled AVSS performed better at tracking a square-wave suspension travel reference signal, consuming a lower amount of energy.

The exact values of the VSS model parameters such as sprung mass loading, suspension damping and suspension spring stiffness, etc. are uncertain. Therefore, consideration should be given to include acceptably bounded parametric uncertainty in the design of control systems for VSS [4; 9; 26; 27; 28].

In this paper, the effects of the uncertainties in vehicle sprung mass loading on the performance of AVSS and PVSS are explored. A NNFBL-controlled nonlinear half-car AVSS, with PID hydraulic actuator force feedback control and a nonlinear half-car PVSS with similar model parameters are compared and their performance is analyzed in the frequency and time domain. This paper is organized as follows: the mathematical model of the nonlinear, half-car AVSS with hydraulic actuator dynamics is presented in Section 2, followed by the performance specifications in Section 3. The overall control architecture and design are given in Section 4. In Section 5, simulation results are presented and discussed and the paper is concluded in Section 6.

2 System Model

2.1 Modelling Assumptions

The following simplifying assumptions are made during the mathematical modelling [8]:

1) All joints connecting the suspension system components are considered to be ideal.

2) The vehicle is moving in a straight line in the horizontal direction at a constant velocity. Forces and moments due to cornering, accelerating and braking of the vehicle are neglected.

3) Both the sprung and unsprung masses are assumed to be uniform in mass.

4) Sprung mass loading does not vary with time [4]. However, different sprung mass loadings are applied to test controller's robustness under parameter variations.

5) The vehicle body is taken as rigid.

6) Road surface roughness and irregularities are the sole source of vehicular vibration; road surface elastic deformation and engine-induced vibrations are ignored [2; 4].

7) Non-uniformity of the tyre as well as wheel unbalance effects are disregarded.

8) The sprung mass centre of mass rests along the vehicle body's longitudinal centreline.

9) The sensors and actuators used respond instantly to changes in measured parameters.

10) Heaving and pitching are taken as the main vehicle body movements. Other possible motions such as yawing, rocking and lateral motions of the sprung mass centre of gravity about the nominal travel path are neglected.

11) The simulation is carried out at the point of dynamic equilibrium. Therefore, the weight of the vehicle is ignored.

12) The vehicle body pitch angular displacement, θ, is varied through large angles about the point of equilibrium [29].

13) Tyre damping couples the wheel and the vehicle body motion at the wheel-hop frequency and thus is not ignored [10].

2.2 Physical and Mathematical Modelling

Figure 1 illustrates a half-car AVSS physical model of sprung mass Ms, pitch moment of inertia I_θ and pitch angular displacement θ, front and rear unsprung masses mu_f and mu_r, respectively.

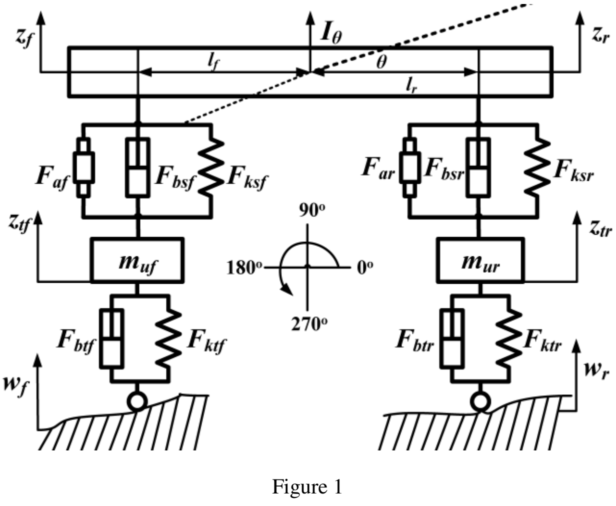


Figure 1
Schematic of a half-car AVSS

z_C, z_{tf} and z_{tr} are the vertical displacements of the sprung mass at the centre of gravity, the front tyre and the rear tyre, respectively. The lengths between the front and rear axles and the vehicle centre of gravity are given by l_f and l_r, respectively. The front and rear suspension travels are expressed as $y_f = z_{tf} - (z_C - l_f \sin\theta)$ and $y_r = z_{tr} - (z_C + l_r \sin\theta)$, respectively.

The application of Newton's 2nd law of motion to the nonlinear half-car AVSS gives the governing equations of motion in state-space form as [7; 11; 29; 30]:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} + \mathbf{p}\mathbf{w} \tag{1}$$

where $\mathbf{x}$ is the state vector, u is the control input vector, $\mathbf{w}$ is the disturbance input vector, $\mathbf{f}(\mathbf{x})$ is the system vector, $\mathbf{g}(\mathbf{x})$ is the control input matrix and $\mathbf{p}$ is the disturbance input matrix.

$$\begin{aligned}\mathbf{x} &= \left[\mathrm{z_c}, \theta, \mathrm{z_{tf}}, \mathrm{z_{tr}}, \dot{\mathrm{z}}_\mathrm{c}, \dot{\theta}, \dot{\mathrm{z}}_\mathrm{tf}, \dot{\mathrm{z}}_\mathrm{tr}, \mathrm{P_{lf}}, \mathrm{P_{lr}}, \mathrm{x_{vf}}, \mathrm{x_{vr}}\right]^\mathrm{T} \\ &= \left[x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\right]^\mathrm{T}\end{aligned} \tag{2}$$

$$\mathbf{u} = \left[\mathrm{u_1}, \mathrm{u_2}\right]^\mathrm{T} = \left[v_f, v_r\right]^\mathrm{T} \tag{3}$$

$$\mathbf{w} = \left[w_f, w_r, \dot{w}_f, \dot{w}_r\right]^\mathrm{T} \tag{4}$$

$$\mathrm{f} = \left[f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{12}\right]^\mathrm{T} \tag{5}$$

The output equation is given as:

$$\mathbf{y} = \begin{bmatrix} y_f \\ y_r \end{bmatrix} = \mathbf{h}(\mathbf{x}) = \begin{bmatrix} x_3 - x_1 + l_f \sin x_2 \\ x_4 - x_1 + l_r \sin x_2 \end{bmatrix} \tag{6}$$

The components of vector $\mathbf{f}(\mathbf{x})$, $\mathbf{g}(\mathbf{x})$ and $\mathbf{p}$ are given in [29].

2.2 Road Input Disturbance Modelling

Equations 18 and 19 express the front and rear wheel input disturbances, w_f and w_r, respectively.

$$w_f = \begin{cases} 0.5a\left(1 - \cos\left(2\pi Vt/\lambda\right)\right) & 1 \le t \le 1 + \lambda/V \\ 0 & otherwise \end{cases} \tag{18}$$

$$w_r = \begin{cases} 0.5a\left(1 - \cos\left(2\pi Vt/\lambda\right)\right) & 1 + t_d \le t \le 1 + t_d + \lambda/V \\ 0 & otherwise \end{cases} \tag{19}$$

where a is the bump amplitude, V is the vehicle forward velocity, λ is the disturbance wavelength, and t is the simulation time. t_d is the time delay between the front and rear wheels written as:

$$t_d = \left(l_f + l_r\right)/V \tag{20}$$

The bump profile is illustrated in Fig. 2. The half-car, hydraulic actuator and road input disturbance model parameters are given in [11; 25; 29].

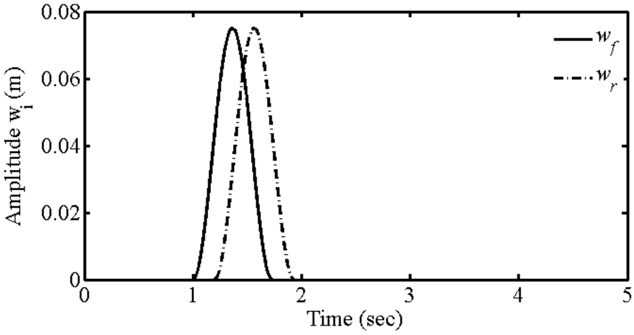


Figure 2
Bump road input disturbance

3 Performance Specifications

The performance specifications utilized in this work are:

1) The closed loops should be nominally stable and the controller must possess good command following and disturbance rejection.

2) The maximum allowable suspension travel should not exceed the limit given as:

$$|y_i| \le z_{\max} \tag{21}$$

where $i \in (f, r)$. $z_{\max}$ is the maximum suspension travel, equal to $0.08m$.

3) The maximum allowable control voltage, $u_{\max}$, is expressed as:

$$|u_i(t)| \le u_{\max} \tag{22}$$

where $u_{\max}$ is the maximum allowable control voltage equal to $10V$.

4) The maximum allowable controlled force, F_{ai}, is given as:

$$|F_{ai}| \le \pm Ms \times g \tag{23}$$

where g is the acceleration due to gravity, equal to $9.81 m/s^2$.

5) To maintain good road holding, the dynamic tyre load, F_{ti}, should not exceed the static load, F_{ti}^{stat} [4]:

$$F_{ti} \le F_{ti}^{stat} \tag{24}$$

where

$$F_{ti} = kt_i(\dot{z}_{ti} - \dot{w}_i) + bt_i(\dot{z}_{ti} - \dot{w}_i) \tag{25}$$

$$F_{ti}^{stat} = g\left[Ms \times l_i / (l_f + l_r) + mu_i\right] \tag{26}$$

6) The RMS values of the performance parameters will be used to enable detailed performance comparison of the AVSS with the PVSS. For n simulation samples:

$$\Theta_{RMS} = \sqrt{n^{-1}\sum{}_0^n(\Theta)^2} \tag{27}$$

where

$$\Theta = \left[y_i, F_{ti}, \ddot{z}_c, \ddot{\theta}, u_i, F_{ai}\right]^T \text{`} \tag{28}$$

7) Ride Comfort: The evaluation of vehicle ride comfort is based on the ISO 2631-1 frequency weighted RMS acceleration [31]. W_k, the ISO 2631-1 frequency weighting for acceleration input at the feet, is selected since theVSS models do not include vehicle seats. A fifth order approximation of W_k is expressed as [32]:

$$\begin{aligned} W_k(s) = &\left(87.72s^4 + 1138s^3 + 11336s^2 + 5453s + 5509\right) \div (s^5 + 92.6854s^4 \\ &+ 2549.83s^3 + 25969s^2 + 81057s + 79783) \end{aligned} \tag{29}$$

The weighted RMS acceleration, a_{wi}^{RMS}, for n samples is given by:

$$a_{wi}^{RMS} = \sqrt{n^{-1}\sum{}_0^n\left(k_{axis}W_k(\ddot{z}_c)\right)^2} \tag{30}$$

where the axis multiplication factor $k_{axis} = 0.40$ for vertical sprung mass acceleration along the z axis. A vibration induced discomfort scale for various values of a_{wi}^{RMS} is given in [31].

4 Controller Design

4.1 Control Architecture

The AVSS control configuration shown in Fig. 3 consists of two control loops: an outer NNFBL suspension travel feedback control loop and an inner PID hydraulic actuator force feedback control loop. Suspension travel is chosen as the controlled output variable on the outer loop because it is easily measured with displacement transducers [15].

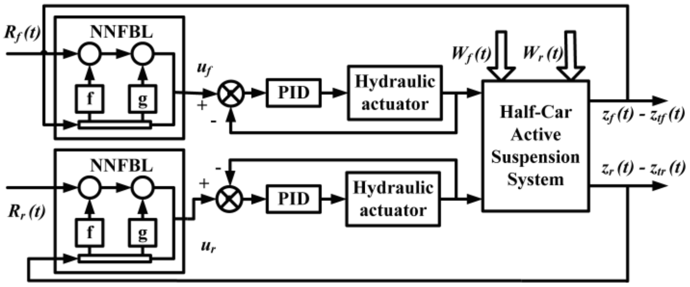


Figure 3
Control architecture

4.2 PID Force Control Loop Design

The PID force control input, u_i to the half-car AVSS actuators is given in Equation 31:

$$u_i = K_{Pi} e_i(t) + K_{Ii} \int e_i(t) dt + K_{Di}\, d(e_i(t))/dt \qquad (31)$$

$$e_i(t) = F_{airef}(t) - F_{ai}(t) \qquad (32)$$

where K_P is the proportional gain, K_I is the integral gain, K_D is the derivative gain, and e_i is the error between the desired actuator force reference signal, F_{airef}, and the actual actuator force, F_{ai}. The PID controller gains obtained by the Ziegler-Nichols tuning method are $K_{Pf} = K_{\mathrm{Pr}} = 0.0010$, $K_{If} = K_{Ir} = 0.0145$ and $K_{Df} = K_{Dr} = 0.0003$.

4.3 NNFBL Suspension Travel Control Loop Design

FBL involves the cancellation of nonlinear system dynamics, thereby enabling the application of linear control techniques to highly nonlinear systems [1; 24]. A system expressed in state-space form can be linearized with an input $\mathbf{u}$ [33]:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} \qquad (33)$$

$$\mathbf{u} = \mathbf{g}(\mathbf{x})^{-1}\left[-\mathbf{f}(\mathbf{x}) - \mathbf{K}^{\mathbf{T}}\mathbf{x} + \mathbf{R}\right] \qquad (34)$$

where $\mathbf{K}$ is a matrix of controller gains and $\mathbf{R}$ is the reference input. Substituting equation (34) into (33) cancels out the nonlinear dynamics of the system, $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$. This gives a linear system able to track reference trajectory using a linear controller expressed as [33]:

$$\dot{\mathbf{x}} = -\mathbf{K}^{\mathbf{T}}\mathbf{x} + \mathbf{R} \qquad (35)$$

Using NNs to approximate $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$, equation (34) becomes [43]:

$$\mathbf{u} = \mathbf{g_{NN}}(\mathbf{x})^{-1}\left[-\mathbf{f_{NN}}(\mathbf{x}) - \mathbf{k^T x} + \mathbf{R}\right] \tag{36}$$

where $\mathbf{f_{NN}}(\mathbf{x})$ and $\mathbf{g_{NN}}(\mathbf{x})$ are the NN approximations of $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$, respectively. The system is required to track a reference model [33]:

$$\dot{\mathbf{x}}_\mathbf{m} = -\mathbf{K^T x_m} + \mathbf{R} \tag{37}$$

Substituting equation (36) into equation (33) gives [33]:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + (\mathbf{g}(\mathbf{x})/\,\mathbf{g_{NN}}(\mathbf{x}))\left[-\mathbf{f_{NN}}(\mathbf{x}) - \mathbf{k^T x} + \mathbf{R}\right] \tag{38}$$

The controller error $\mathbf{e}$ is calculated as [33]:

$$\mathbf{e} = x - x_m \tag{39}$$

The error dynamics given by [33]:

$$\dot{\mathbf{e}} = -\mathbf{K^T e} + [\mathbf{f}(\mathbf{x}) - \mathbf{f_{NN}}(\mathbf{x})] + [\mathbf{g}(\mathbf{x}) - \mathbf{g_{NN}}(\mathbf{x})]\mathbf{u} \tag{40}$$

The MATLAB/Simulink® NARMA-L2 control toolbox was used to perform NNFBL on the outer control loop. Typically, there are two steps involved in indirect adaptive control using NNs: system identification and control design. Since this tool is a SISO tool, two separate NARMA-L2 controllers were used (the half-car AVSS is a MIMO system). Since the NARMA-L2 controller is a SISO tool, the AVSS model was modified to allow for collection of the Input/Output (I/O) data required for system identification; by grounding and terminating part of the AVSS inputs and outputs, the AVSS appears as SISO to the controller (see Fig. 4). The system identification was then carried out in four steps: (a) Experimentation, (b) Model Structure Selection, (c) Model Estimation, and (d) Model Validation.

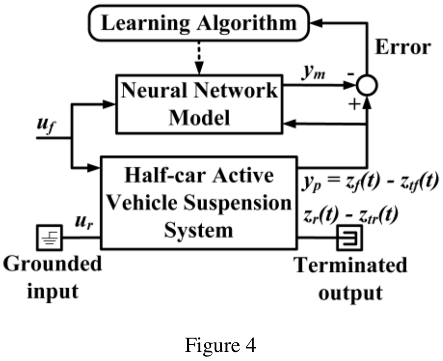


Figure 4

System identification of the half-car AVSS for NNFBL control

4.3.1 Experimentation

Two random non-saturating input signals covering the entire operational range of the half-car AVSS were used to generate the I/O data. The data, Z_i^n, was collected in the form presented in Equation 41 in which $u_i(k)$ and $y_i(k)$ are the input and output to the system, respectively, k is the number of the sampling instant, and n is the total number of samples taken [34; 35] (see Figure 4).

$$Z_i^n = f\{[u_i(k), y_i(k)]; k = 1, \ldots, n\} \tag{41}$$

4.3.2 Model Structure Selection

The NARMA models given in Equation 42, in which *d* is the system delay, na is the number of past inputs, nb is the number of past outputs and *F* is a nonlinear function, were trained to present the half-car AVSS model's forward dynamics. The nonlinear function F is approximated by the NN during identification [34; 35].

$$\begin{aligned} y_i(k+d) = F[&y_i(k), y_i(k-1), \ldots y_i(k-na+1), u_i(k), u_i(k-1), \ldots, \\ &u_i(k-nb+1)] \end{aligned} \tag{42}$$

FBL involves cancellation of the nonlinear system dynamics. The NARMA-L2 controller achieves this by training two MLP NN to approximate the nonlinear functions $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$. The MLP NN models have two layers. The first layer (hidden layer) contains neurons with tangent-sigmoid activation function [34; 35]:

$$\tan sig(n) = (2)/\left(1 + e^{-2(n)}\right) - 1 = \left(1 - e^{-2(n)}\right)/\left(1 + e^{-2(n)}\right) \tag{43}$$

The second layer (output layer) contains linearly activated neurons. Both the hidden and output layers contain a bias. The companion form is expressed in Equation 44. This form enables tracking of reference signal $y_{ri}(k+d)$. However, because $y_i(k)$ is required to calculate $u_i(k)$ and both occur at the same sampling instant, the resulting controller (see Eq. 37) is not practically feasible [34; 35].

$$\begin{aligned} \hat{y}_i(k+d) = f[&y_i(k), y_i(k-1), \ldots y_i(k-n+1), u_i(k), u_i(k-1), \ldots, \\ &u_i(k-na+1)] + g[y_i(k), y_i(k-1), \ldots, y_i(k-n+1), \\ &u(k-1), \ldots, u_i(k-nb+1)] \cdot u_i(k) \end{aligned} \tag{44}$$

$$\begin{aligned} u_i(k) = [&\hat{y}_{ri}(k+d) - f[y_i(k), y_i(k-1), \ldots y_i(k-n+1), u_i(k), \\ &u_i(k-1), \ldots, u_i(k-na+1)] \times [g[y_i(k), y_i(k-1), \ldots, \\ &y_i(k-n+1), u(k-1), \ldots, u_i(k-nb+1)]]^{-1} \end{aligned} \tag{45}$$

By setting the plant delay $d \geq 2$ with a model order $n = na = nb = 2$, a NARMA model and therefore a practical NARMA-L2 controller can be developed as given in Equations 46 and 47, respectively.

Fig. 5 illustrates the NARMA-L2 controller implementation [34; 35].

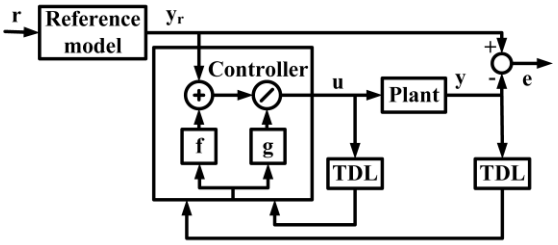


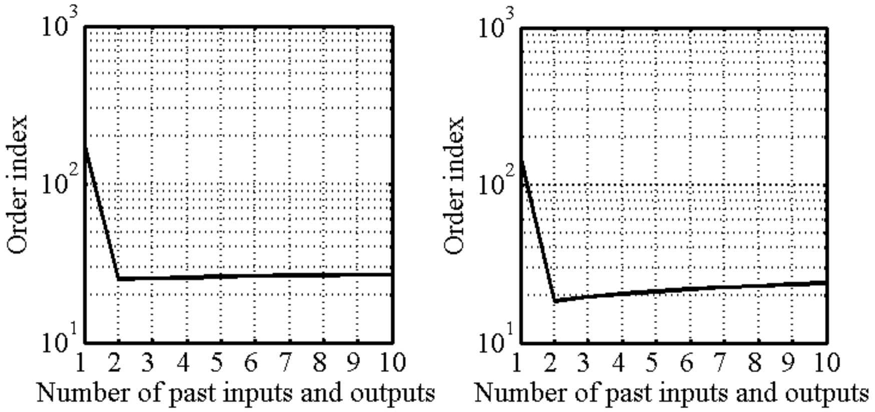
Figure 5

NARMA-L2 Controller implementation

$$\begin{aligned} y_i(k+d) = {} & f\left[y_i(k), y_i(k-1), \ldots y_i(k-n+1), u_i(k), u_i(k-1), \ldots, \right. \\ & \left. u_i(k-n+1)\right] + g\left[y_i(k), y_i(k-1), \ldots, y_i(k-n+1), \right. \\ & \left. u(k-1), \ldots, u_i(k-n+1)\right] \cdot u_i(k+1) \end{aligned} \tag{46}$$

$$\begin{aligned} u_i(k) = {} & \left[y_{ri}(k+d) - f\left[y_i(k), \ldots y_i(k-n+1), u_i(k), u_i(k-1), \ldots, \right.\right. \\ & \left.\left. u_i(k-n+1)\right]\right] \times \left[g\left[y_i(k), y_i(k-1), \ldots, y_i(k-n+1), \right.\right. \\ & \left.\left. u(k-1), \ldots, u_i(k-n+1)\right]\right]^{-1} \end{aligned} \tag{47}$$

The plant model order is determined by evaluating the Lipschitz quotients of the I/O data, plotting the model order index against the lag space, i.e., the number of past inputs and outputs [17; 23]. Figs. (a) and (b) show that the slopes of both graphs increase at the "knee point" where the number of past inputs and outputs is greater than or equal to two. Therefore, setting a model order greater than two could cause overfitting of data when training the NNs [17; 23].



(a) Front suspension

(b) Rear suspension

Figure 6

Two dimensional plot of the Order of Index versus lag space

4.3.3 Model Estimation

Model estimation involves using the I/O data collected to train an NN to approximate the desired function. The NNs were trained using backpropagation method, by minimizing the Mean Square Error (MSE) [35]. For an AVSS plant output $\mathbf{y_i}$ and NN model output $\mathbf{y_{mi}}$, the NN model MSE is calculated as follows:

$$MSE = n^{-1}\sum_{0}^{n}(e_i)^2 = n^{-1}\sum_{0}^{n}(\mathbf{y_i} - \mathbf{y_{mi}})^2 \tag{48}$$

where e_i is the error between the AVSS plant output and the NN model output and n is the number of training samples. The Levenberg-Marquardt (LM) algorithm was used for training the NNs, producing MSEs of 1.6736×10^{-7} after 13 training epochs and 3.8147×10^{-7} after 28 training epochs for the front and rear suspension NNs, respectively. The NNFBL parameters for both the Front and Rear Suspension are listed in Table 2 for a maximum number of 1000 training epochs. The LM algorithm produced the best training results compared to the other training algorithms given in Table 3 for the rear suspension.

Table 2

NNFBL parameters for front and rear suspension

Parameters	Value
Number of hidden layer neurons	5
Number of delayed plant inputs	2
Number of delayed plant outputs	2
Sampling interval T_s	$0.001\sec$
Normalize training data	No
Plant input range	$[-5000N:5000N]$
Plant output range	$[-0.08m:0.08m]$
Maximum interval value	0.5sec
Minimum interval value	$0.1\sec$
Number of training samples	10000

4.3.3.1 Levenberg-Marquardt (LM)

The LM algorithm is an algorithm that enables rapid training of NNs by calculating an approximation of the Hessian matrix, H, rather than the actual Hessian matrix [35]. For performance functions expressed in the form of a sum of squares, such as the MSE, the Hessian matrix can be approximated as [35]:

$$H = J^T J \tag{49}$$

where J is the Jacobian matrix containing the first derivatives of the NN errors with respect to the NN weights and biases. The gradient, G, is given by [35]:

$$G = J^T e \tag{50}$$

where e is a vector of NN errors. At sampling instant k, the vector of the NN weights and biases, x_k, is updated using the approximation of the Hessian matrix [35]:

$$x_{k+1} = x_k - \left(J^T e\right)/\left(J^T J + \mu I\right) = x_k - G/\left(H + \mu I\right) \tag{51}$$

where I is the identity matrix and μ is a scalar parameter that is adjusted at each sampling instant to ensure that the performance function is minimized [45].

Table 3

Performance of NNFBL training algorithms for rear suspension

Full name of algorithm	Rear suspension	
	No. of epochs	MSE
Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton Backpropagation	24	5.0564×10^{-3}
Bayesian Regulation	32	1.1742×10^{-3}
Powell-Beale conjugate gradient backpropagation	9	7.5448×10^{-4}
Fletcher-Powell conjugate gradient backpropagation	10	0.25491
Polak-Ribiereconjugate gradient backpropagation	4	0.054292
Gradient descent backpropagation	1000	7.24×10^{-3}
Gradient descent with momentum backpropagation	6	2.16
Gradient descent with adaptive learning backpropagation	33	3.3688
Gradient with momentum and adaptive learning backpropagation	48	0.24551
Levenberg-Marquardt backpropagation	28	3.8147×10^{-7}
One step secant backpropagation	27	9.6757×10^{-4}
Resilient backpropagation	39	8.6774×10^{-4}
Scaled conjugate gradient backpropagation	29	0.039014

4.3.4 Model Validation

50% of the I/O data is used for NN training, 25% for validation and the remaining 25% for testing of the NN models.

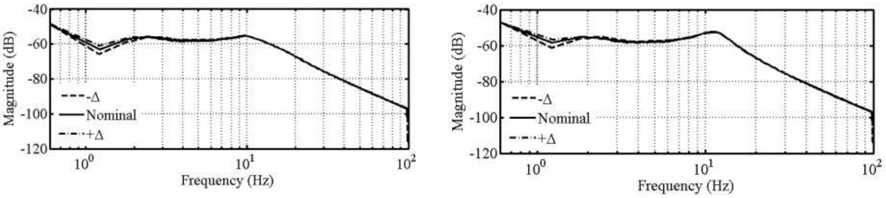
5 Simulation Results and Discussion

The VSS models were built in the MATLAB/ Simulink® environment. The fixed step solver ODE-3 (Bogacki-Shampine) was utilized, with the sampling time $T_s = 0.0001\sec$. T_s is smaller than the fastest half-car AVSS model dynamics, enabling observation of all model dynamics [1]. Robustness to parameter variation was tested by varying Ms and I_θ by $\pm 20\%$ about their nominal values [9; 36].

5.1 Frequency Domain Results

A frequency sweep method similar to that used by [37] was applied here, by use of a chirp road input disturbance signal of amplitude $\pm 15mm$ [38]. Its frequency was set to vary between $0-100Hz$ over a $100\sec$ period, in order to expose the vehicle to a wide range of input disturbance frequencies. Spectral analysis of the AVSS and PVSS model outputs was performed utilizing the MATLAB Welch algorithm/spectral estimator [39]. The "Hamming" window setting with a segment length of $n/100$ (n is the total number of samples) and a percentage overlap of 2^{14} produced optimal frequency response plots. Due to space contrainsts, only the results for the rear suspension travel will be shown.

Figs. 7(a) and 7(b) show the AVSS and PVSS rear suspension travels Power Spectral Density (PSD) responses. The magnitude of the AVSS response at the "knee" between $1Hz$ and the body-hop frequency $(\approx 2Hz)$ is smaller than that of the PVSS. In the frequency range below $4Hz$ and between $4-8Hz$, both the AVSS and PVSS perform similarly. At the wheel-hop frequency $(\approx 12Hz)$, the magnitude of the AVSS response is smaller than that of the PVSS. At high frequencies $(> 20Hz)$, the AVSS and PVSS behave similarly.



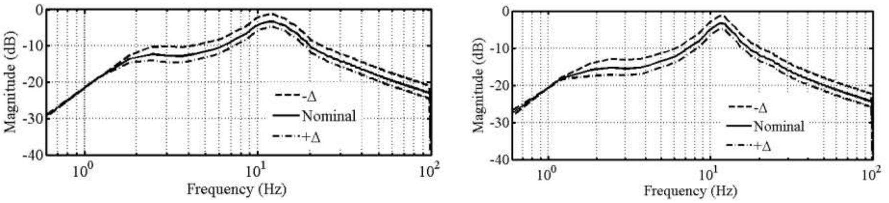
(a) AVSS

(b) PVSS

Figure 7

Rear suspension travel PSD

Figs. 8(a) and 8(b) show the AVSS and PVSS sprung mass acceleration PSD response, respectively. At low frequencies (below $4Hz$), the AVSS is less sensitive to variation in Ms and I_θ than the PVSS. The response of the AVSS about wheel-hop frequency $(\approx 12Hz)$ appears more rounded than that of the PVSS. Above $20Hz$, both the AVSS and PVSS behave similarly.



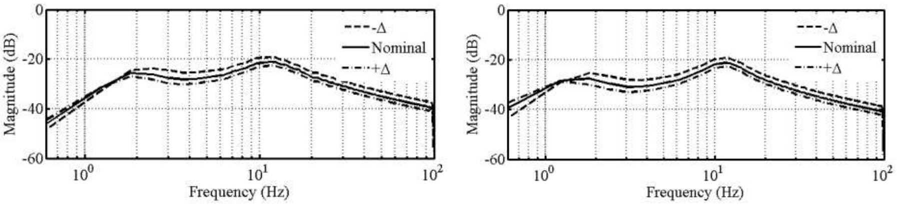
(a) AVSS

(b) PVSS

Figure 8

Sprung mass acceleration PSD

Figs. 9(a) and 9(b) give the AVSS and PVSS pitch angular acceleration PSD responses, respectively. At frequencies less than $4Hz$, the AVSS is less sensitive than the PVSS to variations in Ms and I_{θ} , although the magnitude of the AVSS response is greater than that of the PVSS within this frequency range. In the $4-8Hz$ frequency range, around the wheel-hop frequency and above $20Hz$, the AVSS and PVSS behave similarly.



(a) AVSS

(b) PVSS

Figure 9

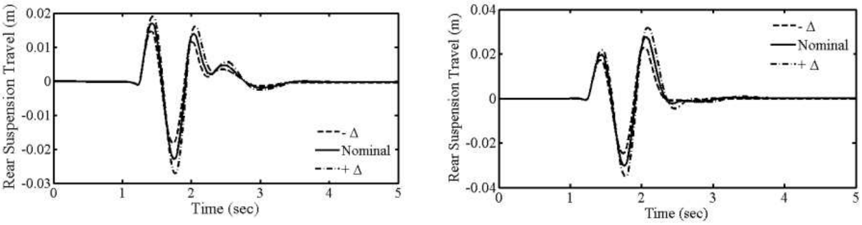
Pitch angular acceleration PSD

5.2 Time Domain Results

Due to space contrainsts, only the results for the rear suspension will be shown. Table 4 contains the RMS values obtained as the VSS models traversed the road input disturbance given in Section 2.2 for suspension travel regulation $\left(R_i(t)=0\right)$, at a constant forward velocity. Although the AVSS performance is significantly better than that of the PVSS, the RMS results for both the front and rear suspension (see Table 4) indicate that an increase in vehicle sprung mass causes a reduction in AVSS suspension travel performance. The minimum and maximum peak front and rear suspension travel values obtained by the AVSS are also lower than those of the PVSS. An increase in vehicle mass causes the suspension travel workspace to reduce. Fig. 10 suggests that the AVSS is less likely to hit the suspension travel limits $\left(\pm 0.08m\right)$ than the PVSS as the mass of the vehicle increases.

Table 4
RMS values for $\Delta = \pm 20\%\{Ms, I_\theta\}$ about their nominal values

Parameters	$-\Delta$			Nominal			$+\Delta$		
	PVSS	AVSS	% Reduction by AVSS	PVSS	AVSS	% Reduction by AVSS	PVSS	AVSS	% Reduction by AVSS
Rear Suspension Travel (m)	0.0069	0.0047	31.88	0.0085	0.0059	30.59	0.0099	0.0069	30.3
Rear Dynamic Tyre Force (N)	222.03	228.15	-3.69	262.29	268.35	-2.31	302.09	305.28	-1.06
Sprung Mass Acceleration (m/s^2)	1.0753	0.8285	23.2	1.1293	0.8658	23.33	1.1395	0.8673	23.89
Pitch Angular Acceleration (rad/s^2)	0.6453	0.6166	4.45	0.5849	0.5559	4.96	0.5231	0.4944	5.49
Rear Actuator Voltage (V)	–	0.1585	–	–	0.1845	–	–	0.2073	–
Rear Actuator Force (N)	–	77.934	–	–	95.713	–	–	112.3	–

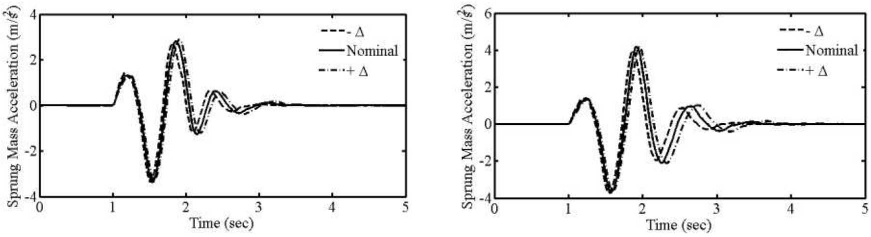


(a) AVSS

(b) PVSS

Figure 10
Rear suspension travel

The AVSS and PVSS sprung mass acceleration are shown in Fig. 11. Although both the AVSS and PVSS maintained a “Not Uncomfortable” level of discomfort throughout, Table 4 shows that the AVSS reduced the ISO 2631-1 weighted RMS acceleration better than the PVSS. With regards to pitch angular acceleration, the RMS values in Table 4 indicate that AVSS performance deteriorated with an increase in vehicle mass. The AVSS peak pitch angular acceleration values tend to be higher when compared with the corresponding PVSS values. However, Fig. 12 shows the AVSS dampens oscillation approximately $0.5\,\text{sec}$ faster than the PVSS.



(a) AVSS

(b) PVSS

Figure 11

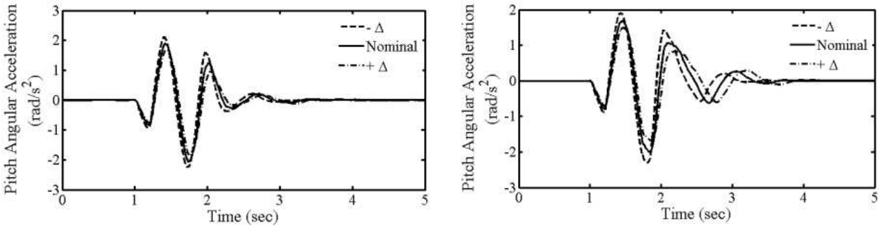
Sprung mass acceleration

Table 4

Weighted RMS acceleration and discomfort levels

	PVSS		AVSS		
	Weighted RMS Acceleration $a_{wi}^{RMS}\ (m/s^2)$	ISO 2631-1 Level of Discomfort	Weighted RMS Acceleration $a_{wi}^{RMS}\ (m/s^2)$	ISO 2631-1 Level of Discomfort	% Reduction by AVSS
$-\Delta$	0.2117	N.U.	0.1663	N.U.	21.45
Nominal	0.2205	N.U.	0.1714	N.U.	22.27
$-\Delta$	0.2217	N.U.	0.1703	N.U.	23.18

where N.U. = Not Uncomfortable



(a) AVSS

(b) PVSS

Figure 12

Pitch angular acceleration

5.3 Effect of the Inner Force Control Loop

To investigate the effect of the inner force control loop on AVSS performance, an AVSS with both inner PID force control loop and outer NARMA-L2 suspension travel control loop is compared with a second AVSS without the inner force control loop. The sprung mass acceleration time domain responses of both AVSS are plotted against each other for nominal Ms and I_θ. As shown in Fig. 13, the AVSS with force control is able to dampen the sprung mass oscillation due to the road input disturbance faster than the one without force control, by stabilizing the hydraulic actuator.

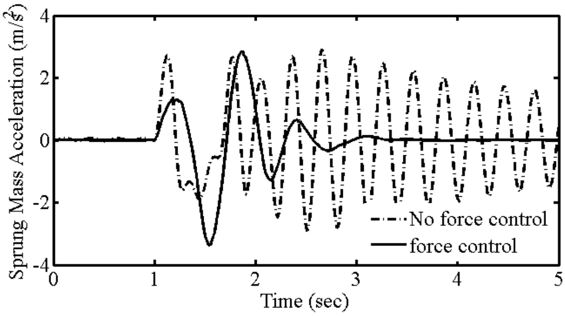


Figure 13

AVSS sprung mass acceleration with and without force control

Conclusion

In this work, a NNFBL controlled nonlinear AVSS with PID based hydraulic actuator force feedback has been presented to improve suspension system model of the AVSS. Robustness to parameter variation was tested by varying the performance. Two-layer tansig activated MLP NNs were used to identify the NN vehicle sprung mass loading and the moment of inertia by $\pm 20\%$, comparing the AVSS and PVSS performance in the frequency and time domains. The frequency domain results showed that the AVSS possesses greater robustness to uncertainties than the PVSS, particularly at frequencies below $4Hz$ and around the wheel-hop frequency. In the time domain, the overall AVSS performance in terms of reducing RMS parameters was better than that of the PVSS. The AVSS suspension travel was significantly lower than the PVSS suspension travel, both remaining within the specified limits throughout. The AVSS and PVSS did not exceed the specified front and rear dynamic tyre force limits. The AVSS maintained the actuator force well below the specified limits.

References

[1] Pedro, J. and Dahunsi, O. (2011). Neural Network-based Feedback Linearization Control of a Servo-Hydraulic Vehicle Suspension System, International Journal of Applied Mathematics and Computer Science 21(1): 137-147

[2] Hrovat, D. (1997). Survey of Advanced Suspension Developments and Related Optimal Control Applications, Automatica 33(10): 1781-1817

[3] Guglielmino, E. and Edge, K. A. (2004). A Controlled Friction Damper for Vehicle Applications, Control Engineering Practice 12(4): 431-443

[4] Gao, H., Lam, J. and Wang, C. (2006). Multi-Objective Control of Vehicle Active Suspension Systems via Load-Dependent Controllers, Journal of Sound a nd Vibration 290(3-5): 654-675

[5] Fischer, D. and Isermann, R. (2004). Mechatronic Semi-Active and Active Vehicle Suspensions, Control Engineering Practice 12(11): 1353-1367

[6] Guclu, R. (2003). Active Control of Seat Vibrations of a Vehicle Model Using Various Suspension Alternatives, Turkish Journal of Engineering and Environmental Sciences 27(6): 361-373

[7] Hassanzadeh, I., Alizadeh, G., Shirjoposht, N. P. and Hashemzadeh, F. (2010). A New Optimal Nonlinear Approach to Half Car ACTIVE Suspension, IACSIT International Journal of Engineering and Technology 2(1): 78-84

[8] Pedro, J. O. (2007). H_2 - *LQG/LTR* Controller Design for Active Suspension Systems, R and D Journal of the South African Institution of Mechanical Engineering 23(2): 32-41

[9] Chen, H., Liu, Z. Y. and Sun, P. Y. (2005). Application of Constrained H_∞ Control to Active Suspension Systems on Half-Car Models, Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME 127(3): 345-354

[10] Akcay, H. and Turkay, S. (2009). Influence of Tire Damping on Mixed H_2/H_∞ Synthesis of Half-Car Active Suspensions, Journal of Sound and Vibration 322(1-2): 15-28

[11] Szaszi, I., Gáspár, P. and Bokor, J. (2002). Nonlinear Active Suspension Modelling Using Linear Parameter Varying Approach, Proceedings of the 10th Mediterranean Conference on Control and Automation (MED2002), Lisbon, Portugal, pp. 1-10

[12] Yoshimura, T., Kume, A., Kurimoto, M. and Hino, J. (2001). Construction of an Active Suspension System of a Quarter-Car Model Using the Concept of Sliding Mode Control, Journal of Sound and Vibration 239(2): 187-199

[13] Yagiz, N. and Hacioglu, Y. (2008). Backstepping Control of a Vehicle with Active Suspensions, Control Engineering Practice 16(12): 1457-1467

[14] Fateh, M. M. and Alavi, S. S. (2009). Impedance Control of an Active Suspension System, Mechatronics 19(1): 134-140

[15] Du, H. and Zhang, N. (2009a). Static Output Feedback Control for Electrohydraulic Active Suspensions via T-S Fuzzy Model Approach, Journal of Dynamic Systems, Measurement and Control: Transactions of ASME 131(5): 1-11

[16] Buckner, G. D., Schuetze, K. T. and Beno, J. H. (2000). Active Vehicle Suspension Control Using Intelligent Feedback Linearization, Proceedings of the American Control Conference, Vol. 6, Chicago, IL, USA, pp. 4014-4018

[17] Dahunsi, O. A. and Pedro, J. O. (2010). Neural Network-based Identification and Approximate Predictive Control of a Servo-Hydraulic Vehicle Suspension System, Engineering Letters 18(4): 357-368

[18] Chantranuwathana, S. and Peng, H. (2004). Adaptive Robust Force Control for Vehicle Active Suspension, International Journal of Adaptive Control and Signal Processing 18(2): 83-102

[19] Sam, Y. M. and Hudha, K. (2006). Modelling and Force Tracking of Hydraulic Actuator for an Active Suspension System, Proceedings of the IEEE conference on Industrial Electronics and Applications, (ICIEA 2006), Singapore, Singapore, pp. 1-6

[20] Astrom, K. J. and Hagglund, T. (2001). The future of PID control, Control Engineering Practice 9(11): 1163-1175

[21] Eski, I. and Yildirim, S. (2009). Vibration Control of Vehicle Active Suspension System Using a New Robust Neural Network Control System, Simulation Modelling Practice and Theory 17(5): 778-793

[22] Feng, J. Z., Li, J. and Yu, F. (2003). GA-based PID and Fuzzy Logic Control for Active Vehicle Suspension System, International Journal of Automotive Technology 4(4): 181-191

[23] Dahunsi, O. A., Pedro, J. O. and Nyandoro, O. T. (2010). System Identification and Neural Network-based PID Control of Servo-Hydraulic Vehicle Suspension System, SAIEE Africa Research Journal, Research Journal of the South African Institute of Electrical Engineers 101(3): 93-105

[24] Precup, R.-E., Preit, S., Petriu, E. M., Tar, J. K., Tomescu, M. L. and Pozna, C. (2009). Generic Two-Degree-of-Freedom Linear and Fuzzy Controllers for Integral Processes, Journal of the Franklin Institute 346(10) 980-1003

[25] Deng, J., Becerra, V. and Stobart, R. (2009). Input Constraints Handling in an MPC/Feedback Linearization Scheme, International Journal of Applied Mathematics and Computer Science 19(2): 219-232

[26] Du, H. and Zhang, N. (2009b). Fuzzy Control for Nonlinear Uncertain Electrohydraulic Active Suspensions with Input Constraint, IEEE Transactions on Fuzzy Systems 17(2): 343-356

[27] Gáspár, P., Szabó, Z. and Bokor, J. (2012). LPV Design of Fault-Tolerant Control for Road Vehicles, International Journal of Applied Mathematics and Computer Science 22(1): 173-182

[28] Rodić, A. and Mester, Gy. (2011). The Modeling and Simulation of an Autonomous Quad-Rotor Microcopter in a Virtual Outdoor Scenario, Acta Polytechnica Hungarica 8(4): 107-122

[29] Huang, C., Lin, J. and Chen, C. (2010). Road Adaptive Algorithm Design of Half-Car Active Suspension System, Expert Systems with Applications 37(6): 4392-4402

[30] Ekoru, J. E. D., and Pedro, J. O. (2012). Intelligent Feedback Linearization-based Control of Half-Car Active Suspension Systems, Proceedings of the fifth IASTED Africa International Conference on Modelling and Simulation (AfricaMS 2012), Gaborone, Botswana, pp. 161-168

[31] International Organization for Standardization (1997). Mechanical Vibration and Shock - Evaluation of Human Exposure to Whole Body Vibration (Part 1: General Requirements), (ISO 2631-1:1997) second edition, International Standards, Geneva, Switzerland

[32] Zuo, L. and Nayfeh, S. (2003). Low Order Continuous-Time filters for Approximation of the ISO 2631-1 Human Vibration Sensitivity Weightings, Journal of Sound and Vibration 265(2): 459-465

[33] Hagan, M. T., Demuth, H., B. (1999). Neural Networks for Control, Proceedings of the 1999 American Control Conference, Vol. 3 pp. 1642-1656, San Diego, CA

[34] Pedro, J. O., Dahunsi, O. A. and Nyandoro, O. T. C. (2012). Direct Adaptive Neural Control of Antilock Braking Systems Incorporated with Passive Suspension Dynamics, Journal of Mechanical Science and Technology 26(12): 4115-4130

[35] Beale, M. H., Hagan, M. T. and Demuth, H. B. (2010). Neural Network Toolbox™ 7 User's Guide, The MathWorks, Inc., Natick, MA

[36] Du, H., Zhang, N. and Lam, J. (2008). Parameter-Dependent Input-delayed Control of Uncertain Vehicle Suspensions, Journal of Sound and Vibration 317(3-5): 537-556

[37] Savaresi, S. M., Poussot-Vassal, C., Spelta, C., Sename, O. and Dugard, L. (2010). Semi-Active Suspension Control Design for Vehicles, Butterworth-Heinemann, Boston

[38] Sammier, D., Sename, O. and Dugard, L. (2003). Skyhook and H_∞ Control of Semi-Active Suspensions: Some Practical Aspects, Vehicle System Dynamics 39(4): 279-308

[39] The MathWorks Inc. (2012). Signal Processing Toolbox for use with Matlab® User's Guide version 6.17, Natick, MA

Examination of Complex Optimization Objective Functions of Parameters of Multi-Step Wire Drawing Technology

Sándor Kovács, Valéria Mertinger

Institute of Materials Sciences, University of Miskolc
3515 Miskolc-Egyetemváros, Hungary
femkovac@uni-miskolc.hu, femvali@uni-miskolc.hu

Abstract: In this paper, the algorithmic realization of complex optimization objective functions of parameters of multi-step wire cold drawing technology is described. As a result of utilizing two types of optimization, either identical optimum cone-angles can be obtained in each pass or the (variable) optimum cone angles having an optional size in each pass can be obtained. The calculation time of objective functions was investigated, as it is one of the most important factors of non-linear optimization. The optimum technological parameters obtained by using the two types of objective functions were compared, and the cases in which it is worth using the complex optimization having a shorter calculation time were defined.

Keywords: analytical explicit model; drawing force; drawing temperature; multi-step wire drawing; complex optimization; technology planning

1 Introduction

The main aspects of planning industrial technology can be divided into three main groups. In the first group, the quality of the product is adjusted in accordance with the requirements prescribed by the buyer, and any damage and defects arising are eliminated or decreased to a minimum extent. The objective functions minimizing the specific costs belong to the second group; here the functions minimizing the specific power consumption necessary for a given drawing operation are extremely important. Productivity is maximized by the third, very important group of objective functions, so the highest possible hourly output can be realized by using this group of objective functions.

The technological planning method of multi-step wire drawing was investigated taking into consideration the aforementioned aspects. Wire drawing is one of the most widely used processes to produce wires, strands, ropes, or welding rods. The technological parameters of multi-step wire drawing can be described by analytical methods as well as by finite element methods. In the course of planning the technology, an optimization is realized so that the quantities belonging to the given aspects of planning are maximized or minimized. The non-linear optimization method must be used for multi-step drawing on the basis of the types of functions describing the different parameters.

In addition to the theoretical solution of the tasks, nowadays the presence of the advantageous properties of resolving algorithms and software is a very important aspect as well (e.g. the shortest possible time necessary for the calculations and the lowest possible memory capacity, the increase in size limits, and programs that can be handled and changed easily). Therefore, it can be concluded that in the case of non-linear optimization, the computer implementation of algorithms as well as experimentation are very significant factors, in addition to the mathematical examination of the tasks and resolving algorithms.

Both drawing force and forming energy have been used to optimize wire drawing in [1-3]. Many authors have used different ductile damage or fracture criteria to investigate central burst or other defects and tried to avoid or minimize damage [4-6]. Other studies [7-9] investigate microscopic criteria, based on characteristics of voids and defects, e.g. Kuboki et al. [10] used the void index, based on Oyane's criterion, to evaluate the void fraction during multipass wire drawing.

Automatic optimization processes for wire drawing have been described in [11-15]. The authors have focused on the minimization of the drawing force or/and the forming energy, on the minimization of damage, on the maximization of the wire reduction per pass, on the heterogeneity of the strains or stresses, and on the shape of the inner surface: and they have tended to use optimization algorithms directly coupled with the FEM calculation, each iteration corresponding to one (or several) FEM evaluation(s).

Although these improved optimization processes describe well the real optimums and explain the universal use of angles in the range [4°-8°], these algorithms can be time consuming because they are coupled with FEM evaluations. For instance, the optimal solution of Roy et al. [12] was found after about 100 FEM simulations, in 10 CPU hours. The second disadvantage concerns multi-step wire drawing: these algorithms do not calculate with back tension drawing force in any of the passes. Finally, in most of the articles, the materials are assumed to obey flow rules with no strain hardening.

In [16] a complex model was chosen from among the published models described by explicit analytical formulas on the basis of measurement data. The best approximation of the measured data is given by this complex model, as well as the most important parameters of planning the technology (the wire drawing force, the maximum drawing stress arising in the wire, and the temperature of the wire) are included in this model. The chosen complex model takes into consideration the back tension forces and the flow rules with strain hardening of the materials.

Our earlier research found that a value of exactness similar to the finite element method (FEM) can be obtained by using the analytical method for wire drawing [17].

In order to perform fast calculation of the optimization, it is necessary to choose a model to describe the technological parameters, which requires an exact and a short-time calculation. This requirement is met by the complex model described in [16], as it is very exact, and moreover the time necessary for calculating the analytical models is much shorter than that needed for FEM.

In [17], an optimization objective function is defined which calculates the number of passes by taking into consideration most of the aspects of planning, the geometry of dies and the extent of deformation to be realized up to an intermediate heat treatment (annealing).

Complex optimization differs from standard optimization in that the domain of variability of optimum values is not determined by the equations but rather is determined by another optimization objective function. Owing to this complexity, the time of calculation will become a very important factor, in addition to the exactness of the optimization process. In this paper, the complex optimization objective function described in [17] and a version of it further extended by us are compared by estimating the difference between the optimum values and the length of time necessary for their calculation.

Nomenclature

A_1, A_2	entry and exit wire area of cross section in a pass (mm^2)
b	penetration depth of heat due to friction (m)
c	wire specific heat capacity (Jkg^{-1}K^{-1})
D_0, D_1, D_2	initial diameter, entry and exit wire diameter in a pass (mm)
F, F_{back}	drawing force and back tension drawing force in a pass (N)
k_f	deformation strength (Nmm^{-2})
$k_{f1}, k_{f2}, k_{fk}=(k_{f1}+k_{f2})/2$	entry and exit wire and mean deformation strength in a pass (Nmm^{-2})
$k_{k,back}$	deformation resistance in the case of back drawing force in a pass (Nmm^{-2})
R_M	maximum tensile stress of the exit cross-section of the wire (Nmm^{-2})
t, t_{def}	time and deformation time, i.e. the time spent in the die by the wire (s)
v_1, v_2, v_{mean}, v	entry and exit and the mean wire axial velocity in a pass (ms^{-1}), velocity of drawing (ms^{-1})
v_{diff}	difference between the circumferential velocity of the capstan and the velocity of the coiled wire on the capstan (ms^{-1})
W	specific deformation work (Nmm^{-2})
α	die semi-cone-angle (rad)
$\Delta A=A_1-A_2$	difference of the entry and the exit wire area of cross section (mm^2)

ΔT increase in temperature within a pass (K)

λ thermal conductivity ($Wm^{-1}K^{-1}$)

μ Coulomb friction coefficient (-)

ν_{def} distribution coefficient of the heat of the volumetric deformation, i.e. the part of the heat developed owing to the deformation remaining in the wire (-)

$\nu_{friction}$ distribution coefficient of the heat of the friction, i.e. the part of the friction heat remaining in the wire (-)

ρ density of the wire (kgm^{-3})

$\sigma_{back}=2*F_{back}/(A_2+A_1)$, σ_{max} back tension drawing stress and maximum value of the drawing stress distribution of the exit cross-section of wire (Nmm^{-2})

$\varphi=\ln(A_1/A_2)$ logarithmic plastic stain (-)

ξ, ζ average relative drawing stress; maximum relative drawing stress (-)

$\eta_{drive\_efficiency}$ coefficient of efficiency of the drive of the capstan (-)

i, s sequence number of drawing pass; sequence number of drawing sequence

N_{seq} number of drawing sequences

$N_{pass,s}$ number of drawing passes in the s-th drawing sequence

x,y, N data calculated by identical-angle complex optimization; data calculated by variable-angle complex optimization; number of data

2 Determination of the Complex Optimization Objective Functions

First of all, it is necessary to define the model describing exactly the parameters of multi-step wire-drawing technology in order to calculate the complex optimizing method. This model, described in [16], yields exactness identical with the exactness of FEM, but a much shorter time is necessary for its calculation. The model is as follows:

Wire drawing force:

$$F=k_{k,back}\cdot\Delta A\cdot\left(1+\frac{\mu}{\alpha}\right)+0{,}77\cdot A_2 k_{fk}\cdot\alpha+F_{back} \tag{1}$$

$$k_{k,back}=\frac{k_{fk}(1-0{,}385\alpha)-\sigma_{back}}{1+\dfrac{\Delta A}{2A_2}\left(1+\dfrac{\mu}{\alpha}\right)} \tag{2}$$

Maximum stress arising in the wire:

if φ≥0.3:

$$\sigma_{max} = k_{k,back}\left(\frac{\Delta A}{A_2}\right)\left(1+\frac{\mu}{\alpha}\right) + k_{fk}\alpha\left(1{,}27 + \frac{A_2}{\Delta A\left(1+\frac{\mu}{\alpha}\right)}\right) + \sigma_{back}, \tag{3}$$

if φ<0.3:

$$\sigma_{max} = k_{k,back}\left(\frac{\Delta A}{A_2}\right)\left(1+\frac{\mu}{\alpha}\right) + k_{fk}\alpha\left(1{,}27 + \frac{A_2}{\Delta A\left(1+\frac{\mu}{\alpha}\right)}\right) \cdot \left(\frac{\varphi}{0{,}3}\right)^{1{,}5} + \sigma_{back} \tag{4}$$

Temperature of wire:

$$\Delta T = k_{fk} v_{def} \frac{\varphi + \alpha}{\rho c} + \frac{1 - \left(1 - \frac{2b}{D_2}\right)^2}{3} 1{,}22 v_{friction} \mu k_{fk} v_{mean} \left(\frac{t_{def}}{\lambda c \rho}\right)^{\frac{1}{2}} \tag{5}$$

The complex optimization objective function defined in [17] consists basically of 3 optimization objective functions and of a temperature limit relating to the optimum places.

The extent and size of deformation is maximized (i.e., the number of stages is minimized) by the first optimization objective function in order to ensure the suitable high quality of the product, i.e. ruptures, surface failures and other damage cannot be found in the ready-made wire. In order to avoid damage and failures, the average (Eq. (6)) and maximum (Eq. (7)) relative drawing stresses have been introduced, the values of which shall be set between 0.5…0.55.

$$\xi = \frac{F}{A_2 k_{f2}} \tag{6}$$

$$\zeta = \frac{\sigma_{max}}{R_M} \tag{7}$$

The specific power consumption is minimized by the second optimization objective function. The specific deformation work described by Eq. (8) is minimized by the above function in such a way that it selects the suitable value of $\varphi_{annealing}$. The $\varphi_{annealing}$ determines the extent of deformation at which the intermediate heat-treatment (annealing) process shall be performed on the wire.

$$W = \int_0^{\varphi_{annealing}} k_f(\varphi)d\varphi + \int_0^{\varphi-\varphi_{annealing}} k_f(\varphi)d\varphi \tag{8}$$

The specific power consumption is also minimized by the third objective function. All the power consumption described by Eq. (9) is minimized by this objective function by choosing the optimum cone angles of passes.

$$P = \sum_{s=1}^{N_{seq}} \sum_{i=1}^{N_{pass,s}} \frac{F_i v_i + (F_i - F_{back,i}) v_{diff,i}}{\eta_{drive\_efficiency,i}} \tag{9}$$

As far as the average value of wire temperature is concerned, an upper temperature limit of 60-70°C is prescribed for the wet drawing and an upper temperature limit of 250-300°C is prescribed for the dry drawing. This limit gives the upper boundary value when choosing the drawing velocity.

The complex optimizing objective function can be obtained by including the optimizing objective functions of utilizing factors as well as the optimizing objective functions determining the place of heat treatment in the condition system of optimization relating to all the power consumption. Eventually, the complex optimizing method also examines the holding of the temperature limit. An additional condition is also determined in the complex optimizing objective function described in [17], namely, that the optimum values of cone angles shall be identical in each pass.

In this paper, the complex optimizing objective function is extended in such a way that the process is allowed to take an optional value of cone angle in each pass. It is obvious that all the power consumption obtained by using this extended objective function as well as the number of passes are less or equal to the result of the complex objective function having identical cone angles. We next investigate the difference between the two results as well as the difference between the lengths of time necessary for the calculations.

3 The Development of the Complex Optimizing Objective Functions Using Computer Technology

The algorithmization of a calculation method belonging to the two complex optimization objective functions and its software realization are necessary for determining the length of time of calculation and the differences between the complex optimums. First of all, the algorithm of the objective function searching for the identical cone angles in each pass was described from the complex optimizing methods, and software was developed for it. The flowchart of the algorithm of objective function searching for the identical cone angles in each pass is demonstrated in Fig. 1.

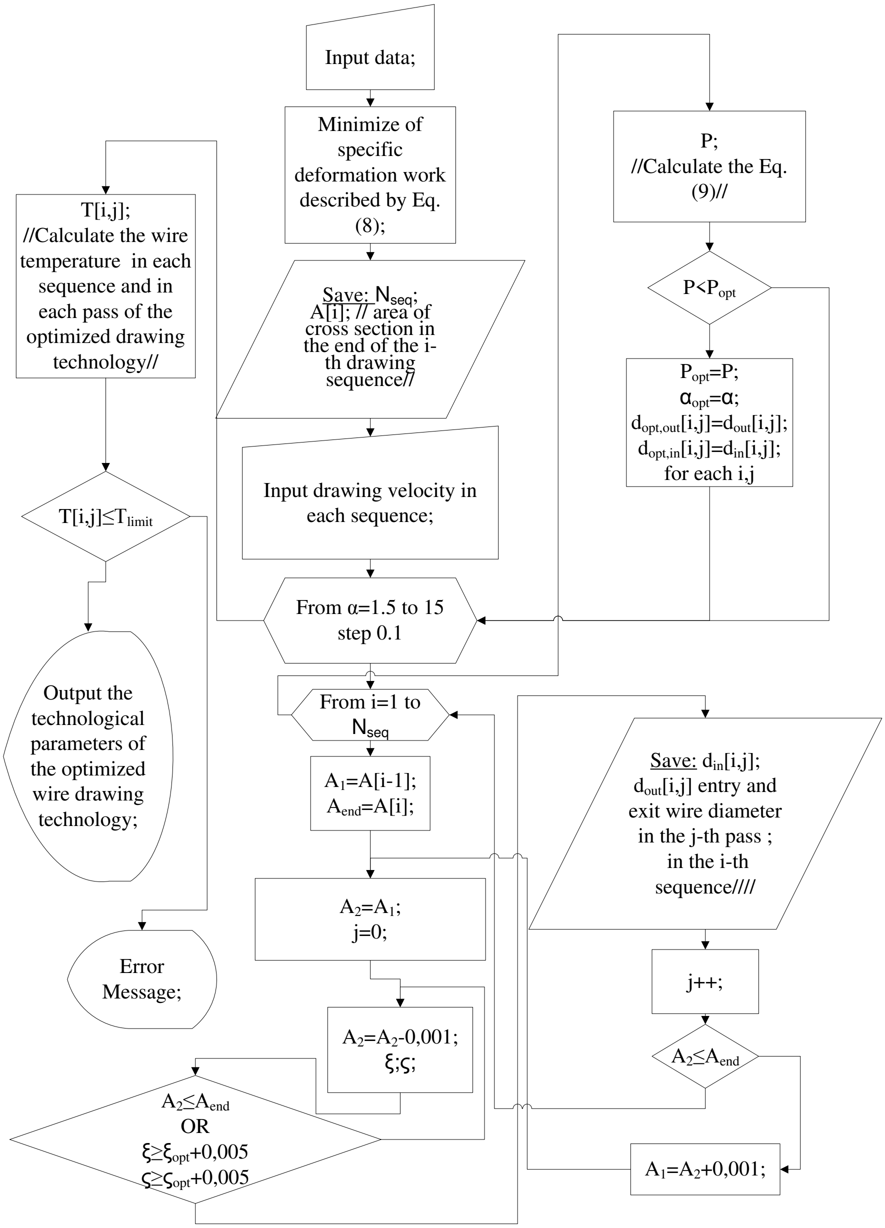

Input data;
Minimize of specific deformation work described by Eq. (8);
Save: N_{seq}; A[i]; // area of cross section in the end of the i-th drawing sequence//
Input drawing velocity in each sequence;
From α=1.5 to 15 step 0.1
From i=1 to N_{seq}
A_1=A[i-1]; A_{end}=A[i];
A_2=A_1; j=0;
A_2=A_2-0,001; ξ;ς;
$A_2 \leq A_{end}$ OR $\xi \geq \xi_{opt}$+0,005 $\varsigma \geq \varsigma_{opt}$+0,005
Save: d_{in}[i,j]; d_{out}[i,j] entry and exit wire diameter in the j-th pass ; in the i-th sequence////
j++;
$A_2 \leq A_{end}$
A_1=A_2+0,001;
P; //Calculate the Eq. (9)//
P<P_{opt}
P_{opt}=P; α_{opt}=α; $d_{opt,out}$[i,j]=d_{out}[i,j]; $d_{opt,in}$[i,j]=d_{in}[i,j]; for each i,j
T[i,j]; //Calculate the wire temperature in each sequence and in each pass of the optimized drawing technology//
T[i,j]≤T_{limit}
Output the technological parameters of the optimized wire drawing technology;
Error Message;

Figure 1

The flowchart of the algorithm of objective function searching for the identical cone angles in each pass

This algorithm has been developed in a software program that treats a group consisting of three different material grades and supposes wet lubrication.

The algorithm requests the suitable material quality and the yield law belonging to it: then the initial diameter of input coarse wire and later the final diameter of the product are given. The algorithm receives the basic parameters of the drawing machine; the most important parameter is whether back drawing force occurs in the individual passes or not.

The pre-determined interval for the cone angle intervals is a set of values between 3^o and 30^o. The values of the cone angles were determined only to an exactness of one decimal because the fact that the cone angles can be developed in the practice was taken into consideration. In this way, the set of optional angles was decreased to a value of 271, thus decreasing the length of time necessary for the calculation of the algorithm. After a value is chosen by the algorithm from a pre-determined interval of cone angles, the algorithm optimizes by using the relative drawing stresses.

It is necessary to use the modified optimization condition in order to decrease the length of time necessary for the calculation. Instead of a definite value, a safety zone is determined for the relative drawing stresses within which their optimum values must fall. In the computer program, the (lower and upper) limits of the safety zone differ from the pre-fixed (0.5….0.55) values by a value of ±0.005. The two relative drawing stresses would still have a small chance of falling at the same time within the safety zone expected by the objective function, and therefore it is reasonable to use an attenuated condition. In accordance with the new optimization condition, only one of the relative drawing stresses is expected to fall within the safety zone, while it is enough if the value of the other factor is lower than the upper limit of the zone. The cone angle has a strong influence on the dominance of one of the relative drawing stresses and whether it falls falling within the safety zone at a time when the value of the other factor is lower than it.

All of the 271 cone angles are examined by the algorithm. At the end of the process, the complex optimization parameters consisting of the saved values of cone angles belonging to the drawing sequences, the values of reduction in each pass, and the values of power consumption are obtained. Finally, the algorithm calculates the wire temperature in each pass and indicates if the upper limit corresponding to the lubrication is exceeded.

The algorithm realizing a complex optimizing objective function allowing different cone angles in each pass is more complicated, and therefore the time necessary for the calculation is longer.

The flowchart of the algorithm realizing a complex optimizing objective function allowing different cone angles in each pass is demonstrated in Figs. 2 and 3.

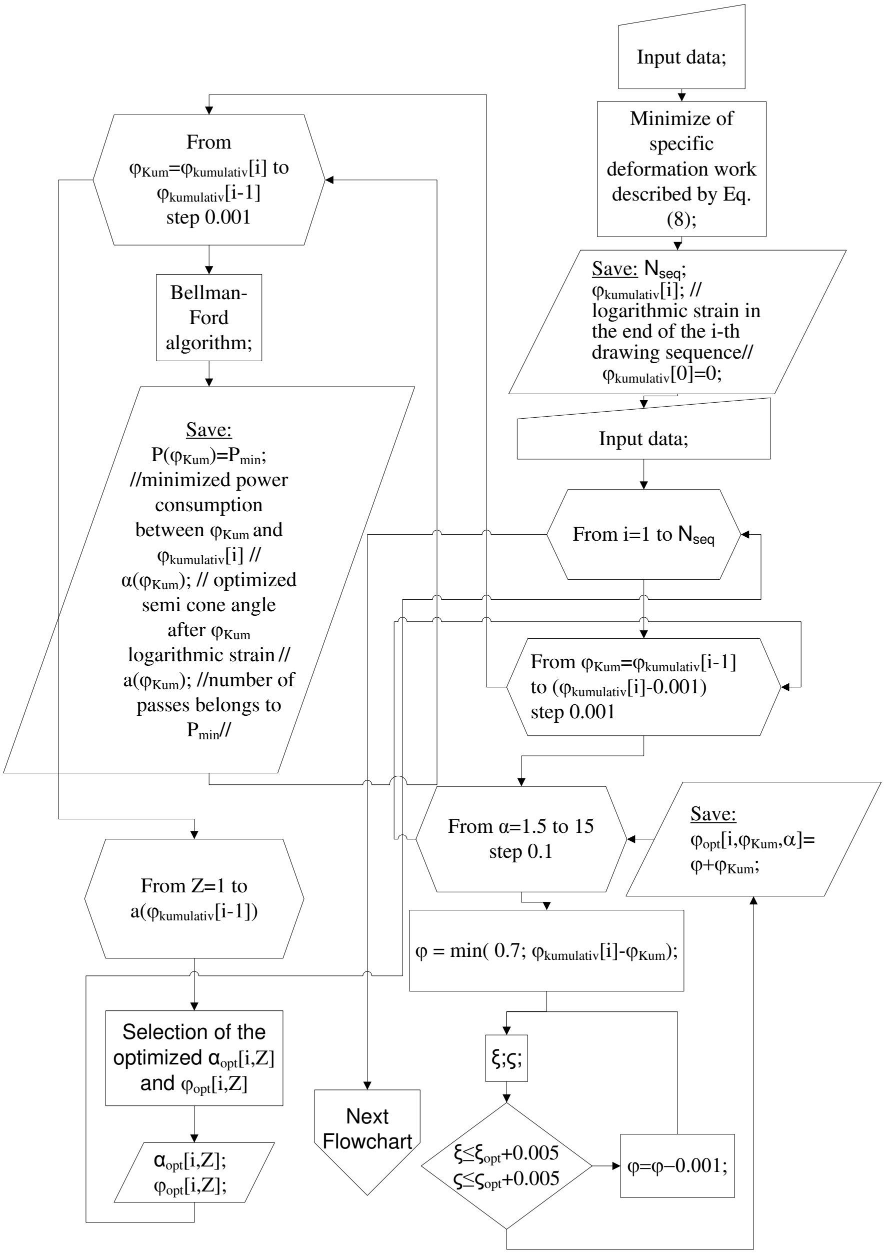

Input data;
Minimize of specific deformation work described by Eq. (8);
Save: N_{seq}; $\varphi_{kumulativ}[i]$; // logarithmic strain in the end of the i-th drawing sequence// $\varphi_{kumulativ}[0]=0$;
Input data;
From i=1 to N_{seq}
From $\varphi_{Kum}=\varphi_{kumulativ}[i-1]$ to $(\varphi_{kumulativ}[i]-0.001)$ step 0.001
From α=1.5 to 15 step 0.1
Save: $\varphi_{opt}[i,\varphi_{Kum},\alpha]=\varphi+\varphi_{Kum}$;
φ = min(0.7; $\varphi_{kumulativ}[i]-\varphi_{Kum}$);
ξ;ς;
$\xi\le\xi_{opt}+0.005$ $\varsigma\le\varsigma_{opt}+0.005$
$\varphi=\varphi-0.001$;
From $\varphi_{Kum}=\varphi_{kumulativ}[i]$ to $\varphi_{kumulativ}[i-1]$ step 0.001
Bellman-Ford algorithm;
Save: $P(\varphi_{Kum})=P_{min}$; //minimized power consumption between φ_{Kum} and $\varphi_{kumulativ}[i]$ // $\alpha(\varphi_{Kum})$; // optimized semi cone angle after φ_{Kum} logarithmic strain // $a(\varphi_{Kum})$; //number of passes belongs to P_{min}//
From Z=1 to $a(\varphi_{kumulativ}[i-1])$
Selection of the optimized $\alpha_{opt}[i,Z]$ and $\varphi_{opt}[i,Z]$
$\alpha_{opt}[i,Z]$; $\varphi_{opt}[i,Z]$;
Next Flowchart

Figure 2

The flowchart of the algorithm realizing a complex optimizing objective function allowing different cone angles in each pass. Part I

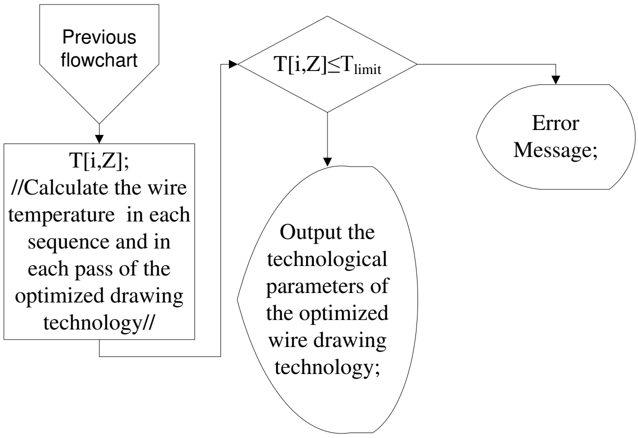

Previous flowchart
T[i,Z];
//Calculate the wire temperature in each sequence and in each pass of the optimized drawing technology//
T[i,Z]≤T_limit
Error Message;
Output the technological parameters of the optimized wire drawing technology;

Figure 3
The flowchart of the algorithm realizing a complex optimizing objective function allowing different cone angles in each pass. Part II

The length of time necessary for the calculation is very long, even if the resources of the algorithm theory are used here. The calculation of the values of angle-reduction pairs arising following all of the possible deformations can increase the calculation time by at least 3 orders of magnitude compared to the calculation time necessary for the complex optimization with identical angle. We need this calculation, as the extent of the previous deformation of wire in the individual passes is not indicated preliminarily during the drawing sequence optimized by the relative drawing stresses. In the next step, it is necessary to carry out a searching action in order to create a drawing sequence from the angle-reduction pairs whose power consumption is the lowest. Only the number of these pairs is larger by at least 2-3 orders of magnitude than the length of time necessary for the calculation of complex optimization having identical angles. The number of “routes” that can be realized on these pairs is of an astronomical order of magnitude; this indicates well how complicated this task is.

The first steps of the algorithm are identical with the algorithm realizing the optimization of identical angles: the request for data and the determination of the number and places of heat treatments.

Following the above procedure, the angle-reduction pairs are calculated. This procedure can be considered as the construction of a map where the algorithm determines the edges coded by the angle-reduction pairs initiating from the point indicated by the deformation occurring up to the given point. So the adjacency created by the edges directing outwards from the given points, as well as the length of route between them (i.e. the value of power consumption), can be obtained.

Further information was collected about the optimum parameters as the algorithm further decreased the length of time of the calculation. The function of the formula describing the explicit average utilization factor was investigated for the values of technological parameters applied in industrial circumstances. On the basis of this investigation, it can be stated that the value of logarithmic deformation carried out in one pass cannot be higher than 0.7 under any circumstances because, in this case, the value of the factor can only be higher than 0.5…0.55. A value of 0.02 can be chosen as the lowest value of deformation, which reduces a rod with a diameter of 100 mm only by 1 mm. As a consequence of the decreasing behaviour of Eqs. 3 and 4, the relative drawing stresses can by no means reach the values of 0.5…0.55 at any of the technological parameters below this deformation range. This means that the range of deformation that can be accepted in one pass is between 0.02…0.7. Owing to the exactness of workability of dies, this range has also been divided into discrete values where the difference between the adjacent reductions is 0.001. With this, the selectable reduction set was decreased to a value of 681. So the algorithm must examine merely a maximum of 681 values if it searches for the optimum value of utilization factor at a given angle-value, so the time necessary for the calculation of constructing the map will be a maximum of 681 times more than the calculation time of optimization.

Afterwards the algorithm searches for the shortest route between the initial and end-points from the short sections of route. The shortest route-set (having the lowest power consumption) is found from the route-set with a number of astronomic orders of magnitude by the clever Bellman-Ford algorithm [18, 19]; the calculation time for this searching procedure does not have a longer order of magnitude than the machine time of making a map. Therefore, it can be concluded that the calculation time of complex optimum with variable angles is longer only by 2 to 3 orders of magnitude, as compared to the version having identical angles; so this algorithm can calculate the objective function even using current computer capacity.

4 Comparison of the Complex Optimization Objective Functions

It can be seen in the previous section that there is a significant difference in calculation time lengths of the algorithms realizing the two complex optimization objective functions. In the case of a multi-step wire drawing, the time necessary for the complex optimization of searching for the identical cone angle is between 5 minutes and 1 hour, while the time necessary for finding the variable angles lasts for days. On the other hand, the complex optimization method with identical angles can never result in an optimum value having lower power consumption than the method with variable angles.

On the basis of the above arguments, it was necessary to compare the two optimum values for the most possible parameter values in order to decide what the extent of error is during the process to search for the identical angle with a short calculation time as compared to the complex optimization with variable angles. The results obtained in the course of running the algorithm with variable angles for ten weeks was compared to the method of identical angles.

The complex optimizations were carried out by using the following three material grades: **Al99.5**; **CuE**; **C10**. The individual optimization processes were started by using different initial wire diameters: the highest value was **20 mm,** while this value was decreased by 0.6 mm down to **0.2 mm** in the course of the further runs. The complete deformation of a multi-step drawing process was equal to the value of approximate deformability of the given material. Though heat treatment is not prescribed in the drawing technology in this case, the two complex optimizing objective functions can well be compared. In addition to changing of the initial wire diameter, the drawing velocity was also changed. The final velocity of drawing was **1, 4** and **7 m/s.** The two complex optimums were calculated for a total of 102 technological adjustments as a function of the velocities, initial wire diameters and material grades. These parameters cover the greatest part of the technological value set applied during wire production, and consequently, a good approximation of the behavior of complex optimums is given for an optional multi-step drawing.

The semi-cone angles belonging to the passes of complex optimums are demonstrated in Fig. 4.

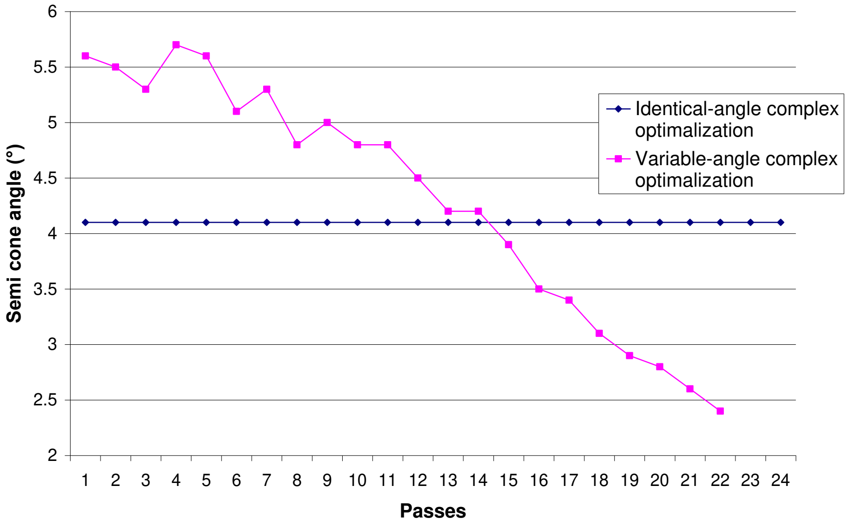


Figure 4

The change of cone angles as a function of the complex optimizing objective functions when the initial wire diameter is 20 mm, the diameter of ready-made wire is 2.58 mm, the final velocity of drawing is 7 m/s and the grade of material is Al99.5

As a result of the complex optimization with variable angles, drawing sequences were obtained consisting of 1 or 2 passes fewer than those in the results of optimization with identical angles (as can be seen in Fig. 4). Independent of the drawing velocity and wire diameter, a decreasing trend can be observed concerning the semi-cone angles if the number of passes increases and the wire diameter decreases when using the version with variable angles, although this change cannot be considered monotonous.

In addition to comparing the number and geometry of dies, the most important task is to compare the value of power consumption of the multi-step wire drawing machine in order to perform the complete deformation.

The absolute error-norm (10) introduced in [16] was used for digitizing the difference between the two objective functions; as input, it substitutes the values for power consumption obtained by the complex optimizations belonging to the identical technological adjustment to the place with identical index.

$$\|,\|_1 = \sum_{j=1}^{N} \frac{|x_j - y_j|}{N * y_j} \tag{10}$$

The value of error norm obtained for the 102 adjustments (samples) is **0.0396.** This value can be considered infinitesimal, so it can be stated that, as far as power consumption is concerned, the difference between the complex optimization with identical angles and the complex optimization with variable angles is not too significant, independent of the cross section, drawing velocity and material grade.

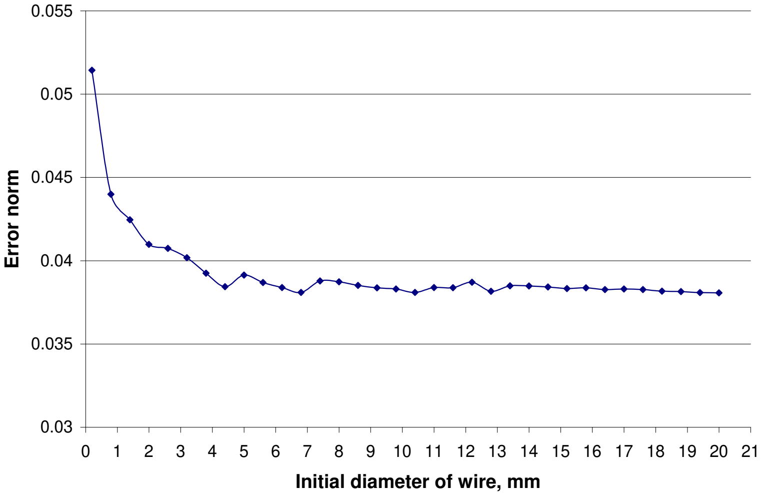


Figure 5

The difference between the power consumption values belonging to the complex optimums expressed by absolute error norms as a function of the initial diameter of wire. The drawing procedure has been carried out up to the limit of deformability for each material grade.

The behaviour of error norms was investigated as a function of the initial wire diameter, the final velocity of drawing and the material grade. The results of these investigations are demonstrated in Figs. 5, 6 and 7. It can be seen in Fig. 5 that the error-norm of power consumption is constantly below 0.04 in the case that the initial diameter is more than 3 to 4 mm. The error norm starts growing in the case of lower values of initial diameters. The angular coefficient of increase becomes significant below 1 mm, though here the value of error norm is still less than 0.045. As was described earlier in the section dealing with the development of the model, the closed analytical relationships installed in some complex optimizing objective functions are not relevant in the case of the finest gauge drawing; therefore, optimization below a wire diameter value of 0.5 mm is not suggested using the optimizing processes described here. However, a small and constant difference can be found between the complex optimizations in the accepted range; therefore, it seems to be reasonable to use an algorithm with identical angles, which has a shorter time of calculation, for the technological planning.

By investigating the difference between the complex optimums as a function of velocity (Fig. 6), it can be stated that the error-norm increases if the final velocity of drawing increases, and the relationship can be approximated by a power function with an exponent with a value of less than 1, in spite of the fact that direct proportionality cannot be observed between them. On the basis of the behavior of the function, it can be predicted that the error norm will not exceed the range of 0.07 to 0.08, even in case of a drawing velocity value of 10 m/s.

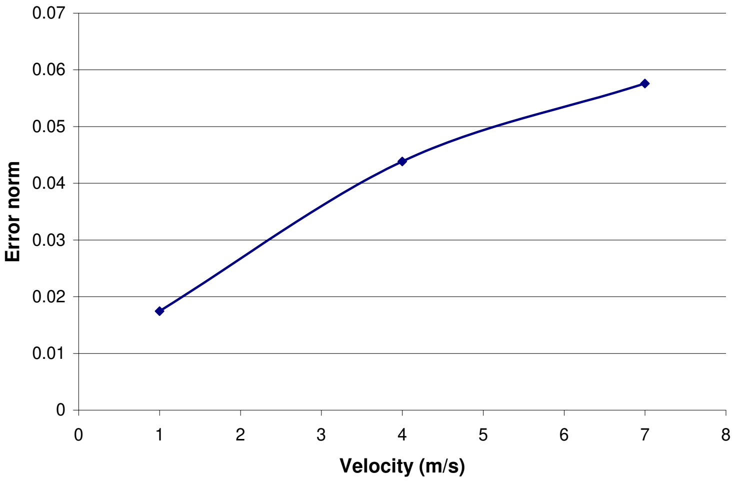


Figure 6

The difference between the power consumption values belonging to the complex optimums expressed by error norms as a function of the final velocity of drawing. The drawing has been carried out up to the limit of deformability in case of each material grade.

By investigating the differences as a function of the three material grades, it can be stated that the n-value of the material influences the value of the error norm. The higher the hardening reaction of the material to the deformation is, the higher the difference is between the power consumption values of complex optimums. However, the effect of material quality is not significant. It can be seen in Fig. 7 that the highest difference between the error norms is only around 0.01 and the error norm is much less than 0.045 in the case of copper.

It can be concluded that the difference between the results of complex optimization with identical angles and complex optimization with variable angles is sensitive to the initial diameter and to the material grade only to a small extent.

However, the drawing velocity has a more significant effect on the error norms belonging to the complex optimums. In the case of coarse drawing machines, medium drawing machines and fine drawing machines, the result of the complex optimization objective function with identical angles is accepted as a drawing-technological line with a minimum error up to a final velocity of 10 m/s, which is the best drawing-technological line from the point of view of material grade, quantity and cost effectiveness. In order to increase the effectiveness, a complex optimization objective function with variable angles can be offered on the basis of the trend of error norms if the final velocity of drawing is higher than 10 m/s.

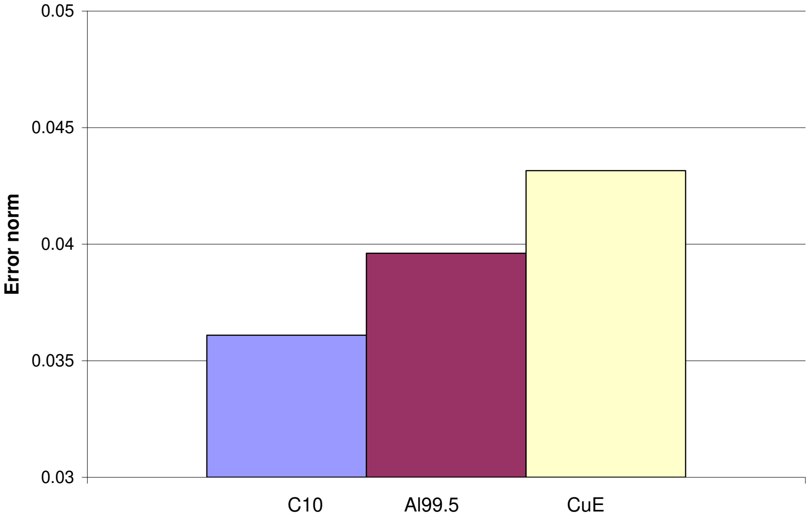


Figure 7

The difference between the power consumption values belonging to the complex optimums expressed by absolute error norms as a function of the material grades. The drawing has been carried out up to the limit of deformability in the case of each material grade.

Conclusions

In the present paper, a complex optimization objective function concerning multi-step wire drawing technology is described and is also interpreted on a complex model. A method has been developed by which the technological planning can be carried out quickly by means of calculations. Therefore, it has become possible to determine the most advantageous optimum technology for the task in a cheaper and faster way.

The objective function focused on in this paper has been extended from a version in which cone angles are identical. In the new objective function, the values of the cone angles are independent of each other in the individual passes. The two versions were compared. The difference between the lengths of time necessary for the calculation of the two objective functions was determined by the algorithmization of the complex optimization processes. It was found that the difference between the length of time necessary for the calculations is of about 2 orders of magnitude, i.e. 100 times higher for the complex optimization with variable angles (e.g. the length of time can be some minutes versus a complete day).

The complex optimization methods were realized using software as well. The differences between the complex optimums were determined by comparing the results of software runs considering the power consumption of the drive and the total number of passes. The absolute error norm of differences is less than 0.04; this means that the two kinds of complex optimization show good conformity.

As a result of the more detailed investigations, it can be concluded that the optimization method with identical angles, due to its significantly shorter calculation time, is suitable for the most effective realization of the planning of an industrial technology – independent of the quality of material – in the case of fine, medium and coarse wires (a diameter of 0.5 to 20 mm) at a final velocity of less than 10 m/s. However, the optimization method with variable angles is suggested for planning the multi-step drawing technology if the velocities are above 10 m/s. Our purpose is to increase the effectiveness of industrial technology; the next step is to test this approach in an industrial setting.

Acknowledgement

The described work was carried out as part of the TÁMOP-4.2.2/B-10/1-2010-00008 project in the framework of the New Hungarian Development Plan. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

[1] J. G. Wistreich: Investigation of the Mechanics of Wire Drawing, Proc. Inst. Mech. Engrs. (London). 169 (1955) pp. 654-665

[2] W. Evans, B. Avitzur: Measurement of Friction in Drawing, Extrusion and Rolling, Trans ASME Series F, J. Lub. Tech. 90 (1968) pp. 72-80

[3] U. S. Dixit, P. M. Dixit: An Analysis of the Steady-State Wire Drawing of Strain-Hardening Materials, J. Mater. Process. Technol. 47 (1995) pp. 201-229

[4] Z. Zimerman, H. Darlington, H. E. Kottkamp: Selection of Operating Parameters to Prevent Central Bursting Defects during Cold Extrusion. In: Hoffmanner AL (ed.) Metal Forming: Interrelation between Theory and Practice. Plenum Press, New York, 1979, p. 47

[5] C. C. Chen, S. I. Oh, S. Kobayashi: Ductile Fracture in Axisymmetric Extrusion and Drawing, Trans ASME Series B, J. Eng. Ind. 101 (1979) pp. 36-44

[6] L. Chevalier: Prediction of Defects in Metal Forming: Application to Wire Drawing, J. Mater. Process. Technol. 32 (1992) pp. 145-153

[7] P. McAllen, P. Phelan: A Method for the Prediction of Ductile Fracture by Central Bursts in Axisymmetric Extrusion, J. Mech. Eng. Sci. 219 (2005) pp. 237-250

[8] P. McAllen, P. Phelan: Numerical Analysis of Axisymmetric Wire Drawing by Means of a Coupled Damage Model, J. Mater. Process. Technol. 183 (2007) pp. 210-218

[9] H. B. Campos, P. R. Cetlin: The Influence of Die Semi-Angle and of the Coefficient of Friction on the Uniform Tensile Elongation of Drawn Copper Bars. J. Mater. Process. Technol. 80-81 (1998) pp. 388-391

[10] T. Kuboki, M. Abe, Y. Neishi, M. Akiyama: Design Method of Die Geometry and Pass Schedule by Void Index in Multi-Pass Drawing, J. Manuf. Sci. Eng. 127 (2005) pp. 173-182, DOI:10.1115/1.1830490

[11] A. Mihelič and B. Štok: Optimization of Single and Multistep Wire Drawing Processes with Respect to Minization of the Forming Energy, Struct. Opt. 12 (1996) pp. 120-126

[12] S. Roy, S. Ghosh, R. Shivpuri: A New Approach to Optimal Design of Multi-Stage Metal Forming Processes with Micro Genetic Algorithms, Int. J. Mach. Tools Manufact. 37 (1997) pp. 29-44

[13] G.Celano, E. Fichera, E. Fratini, F. Micari: The Application of AI Techniques in the Optimal Design of Multi-Pass Cold Drawing Processes, J. Mater. Process. Technol. 113 (2001) pp. 680-685

[14] P. J. M. Van Laarhoven, E. H. L. Aarts: Simulated Annealing: Theory and Applications. Kluwer Academic Publishers, Dordrecht, 1987

[15] T. Massé, L. Fourment, P. Montmitonnet, C. Bobadilla, S. Foissey: The Optimal die Semi-Angle Concept in Wire Drawing, Examined using

Automatic Optimization Techniques, Int. J. Mater. Form. (2012) DOI 10.1007/s12289-012-1092-9

[16] S. Kovács, V. Mertinger, M. Voith: Development of Complex Analytical Model for Optimizing Software of Wire Drawing Technology, Mater. Sci. Forum. 729 (2013) pp. 156-161

[17] S. Kovács, V. Mertinger: Development of a Complex Optimizing Model of Wire Drawing Technology, Mater. Sci. Forum. Manuscript is accepted for publication

[18] R. Bellman: On a Routing Problem, Q. of Appl. Math. 16 (1958) pp. 87-90

[19] L. R. Ford Jr., D. R. Fulkerson: Flows in Networks, Princeton University Press, Princeton, 1962

A Digital Diagnostic System for a Small Turbojet Engine

Rudolf Andoga[*], Ladislav Főző[**], Ladislav Madarász[***], Tomáš Karol'[***]

[*]Technical University of Košice, Faculty of Aeronautics, Department of Avionics, Rampová 7, 04121 Košice Slovakia, e-mail: rudolf.andoga@tuke.sk

[**]Technical University of Košice, Faculty of Aeronautics, Department of Aviation Engineering, Rampová 7, 04121 Košice Slovakia, e-mail: ladislav.fozo@tuke.sk

[***]Technical University of Košice, Faculty of Informatics and Electrical Engineering, Department of Cybernetics and Artificial Intelligence, Letná 9, 04200 Košice Slovakia, e-mail: ladislav.madarasz @tuke.sk, tomas.karol@tuke.sk

Abstract: During the lifecycle of a system not only functionality but also other aspects like safety and reliability are very important. These terms are even more important when connected to aviation in engine and avionic systems control. The reason is simple, a failure must not cause a shutdown of a system during the flight as it would cause a catastrophe. The article deals with the proposal of a progressive diagnostics/backup system using a modified voting method with computational backup models using neural networks. The proposed architecture is expected to be suitable for turbojet engines and was tested on a laboratory object, a small turbojet engine MPM-20 with positive results under operational conditions.

Keywords: digital control; diagnostics; backup; neural network, turbojet engine

1 Introduction – Diagnostics in Aircraft Systems

Progressive approaches in control and advances in the field of artificial intelligence have found their place in diagnostic systems. The application of such systems has shown the potential to increase the reliability and safety of the diagnosed system [1, 2, 3]. The problem of diagnostics is especially important in the field of aircraft avionic systems [1, 2, 3]. Modern aircraft are dependent on the errorless operation of different systems that are usually backed up and have built in diagnostics [1, 2, 3, 4, 5]. With the advent of digital systems, the field of diagnostics offers new possibilities in the implementation of progressive algorithms instead of traditional voting and bulky hardware back-up methods [2, 3]. Modern computers can run back-up controller and sensor models, thus creating highly redundant networks with greatly increased reliability [5, 6, 7, 8]. Moreover, digital/electronic/electric devices and algorithms are more susceptible to errors, either due to interference or algorithmic overload problems; on the other hand miniaturization offers ways to overcome such problems [6, 7]. The article will deal mainly with application of new approaches in diagnostics and back-up systems in the area of turbojet engines. As a test bed, a small turbojet engine MPM-20 that has been transformed into a digitally controlled system will be used. The engine is derived from TS-20/21 engines and is an ideal test bed for scientific purposes [9, 10, 11, 12].

2 Intelligent Engine Control System with Diagnostics

As the engine TS-21 has been adapted to full digital control [9, 10], the safety and diagnostics of critical parameters come into the forefront. The situational control system of the engine had to be expanded with diagnostic and back-up modules that detect faults in the digital control, sensors as well as the back-up systems to transfer the current engine control strategy into a back-up mode [2, 6, 7, 8]. The basic architecture of the designed system is shown in the figure 1 [12].

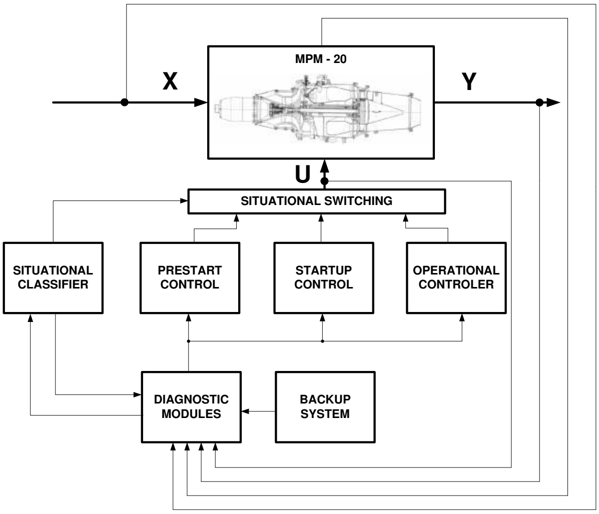


Figure 1
The situational control system with diagnostics and backup modules

The function of the system can be described as follows:

- The measured values (input, state, output and action) are deployed through the diagnostic module where model comparison will show if the measured parameters are erroneous.

- In case of an error in any of the parameters, its value will be replaced by a value of the back-up sensor or an analytic back-up will be used (synthetic value calculated by a model).
- Further these values will enter the situational classifier that will select the actual state of the engine and the appropriate control algorithm.

The whole control process can be broken down into three basic situational frames which are [12]:

- pre-start diagnostics,
- start-up control,
- operational control and control with degraded control modes [2, 12].

Regarding software implementation, the whole situational control and diagnostic system is using Matlab/Simulink and LabView systems. Although both overlap in functionality, LabView is used mainly for data acquisition that is run with National Instruments data acquisition hardware, data visualization, simple calculations for certain situational frames (pre-start control, pre-start diagnostics) and is also used for running the back/up diagnostic module with voting method that will be described later. Matlab/Simulink is used for complex calculations and complex control algorithms and models. Simulink runs all neural networks and dynamic engine models used for diagnostics situational classifier and control, including servo-valve control and operational engine control.

2.1 Pre-start diagnostics

This situational frame is based on the pre-start control of aggregates that are necessary for start and operation of the engine. With its realization a diagnostic expert system has been chosen (Fig. 2) [12].

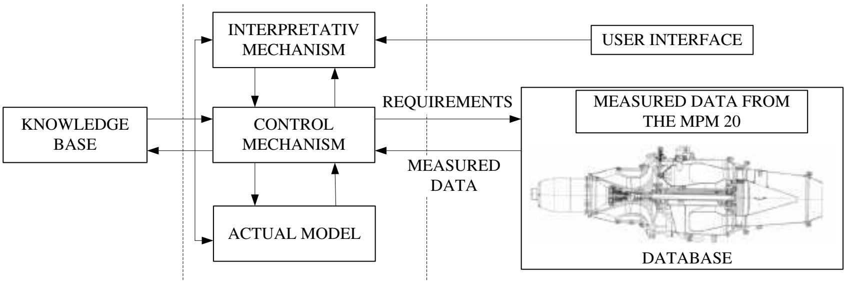


Figure 2
The scheme of the diagnostic expert system for MPM-20

The designed expert system is used to test all observed engine sensors tied to the engine parameters. The system using its knowledge base decides if all the parameters are at their usual levels and decides if the engine can be started.

The whole process can be described as follows:

- Measured data are processed by the expert system and its knowledge base with rules in the form of inference network.
- The results of the process are shown to the user in the graphical interface.

The expert system has been implemented in LabView environment. Its knowledge base in the form of the inference network is shown in the Figure 3.

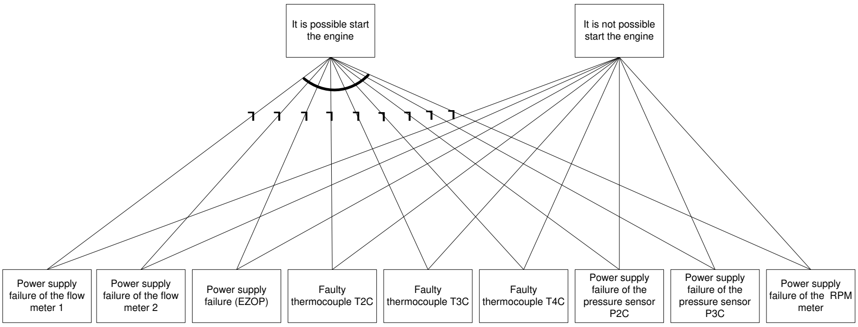

It is possible start the engine
It is not possible start the engine
¬
Power supply failure of the flow meter 1
Power supply failure of the flow meter 2
Power supply failure (EZOP)
Faulty thermocouple T2C
Faulty thermocouple T3C
Faulty thermocouple T4C
Power supply failure of the pressure sensor P2C
Power supply failure of the pressure sensor P3C
Power supply failure of the RPM meter

Figure 3
The inference network of the pre-start diagnostic expert system

2.2 Pre-Start and Start-Up Control

To implement the proposed algorithms, the hydro-mechanical fuel control unit has been replaced by the computer controlled servo valve LUN 6743. This represents the action element for control and it runs in Matlab and LabView environment as described in the beginning of the chapter 2. The flowchart is depicted in the Fig. 4.

Using a digitally-controlled actuator to meter fuel flow into the engine together with digitally-controlled auxiliary systems allows us to build a flexible control and diagnostic system utilizing the concept of a wireless sensor network in the future [14]. All elements can be controlled and turned on or off independently, tied only by algorithmic software links. This allows us to build a complex control system utilizing different control strategies according to situational control methodology. However, the utilization of complete digital control system puts higher demands on the functionality of all elements, because a failed sensor can cause catastrophic failure during digital engine startup that would not appear with hydromechanical control, where the lack of pressure will not allow any actuation. Now the system is dependent on values of speed, temperature and pressure that are presented in a digital form. The start-up of the engine is composed of two phases:

- Phase before ignition (control of auxiliary units – see Figure 4)
- Phase after ignition (situational frame startup – use of fuzzy situational controller) [9]

A part of the start-up control is also formed by implementation of micro-situational start-up fuzzy controller. This controller handles the engine and controls fuel flow according to the speed, temperature and temperature derivation using fuzzy inference system rule base, while also handling atypical situations during start-up after fuel ignition, such as excess temperatures, flameouts, stalls. The start-up controller is also tasked with decreasing the exhaust temperature peaks and securing smooth acceleration towards a normal idle operational state [9].

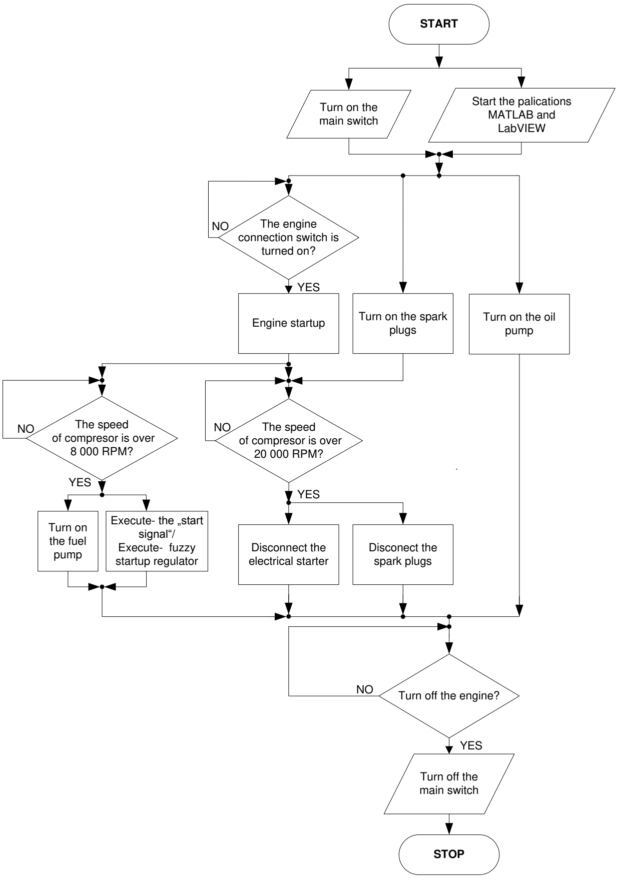

START
Turn on the main switch
Start the palications MATLAB and LabVIEW
NO
The engine connection switch is turned on?
YES
Engine startup
Turn on the spark plugs
Turn on the oil pump
NO
The speed of compresor is over 8 000 RPM?
YES
NO
The speed of compresor is over 20 000 RPM?
YES
Turn on the fuel pump
Execute- the „start signal"/ Execute- fuzzy startup regulator
Disconnect the electrical starter
Disconect the spark plugs
NO
Turn off the engine?
YES
Turn off the main switch
STOP

Figure 4
The inference network of the pre-start diagnostic expert system

2.3 Operational Control

During the operational control of the engine we use a double looped control circuit with inner loop controlling the fuel metering valve and the outer loop controlling the speed or other parameters of the engine using the methodology of situational control [11]. This means that the engine is controlled with different algorithms in different situational frames that are classified by the situational classifier. Normal operational is decomposed into different situational frames, such as stable operation, acceleration, deceleration, degraded modes, etc. [9]

3 Diagnostic and Backup Module

The core of the diagnostics and backup system is represented in the block of diagnostics/backup in Figure 1. The present control of the engine is done via fuel supply parameter and, as such, the engine represents a complex system with a single degree of freedom. The controlled parameter that defines the thrust of the engine is its speed. The speed of the engine determines the mass air flow through it, thus determining its thrust. As a testing parameter for the implementation of the presented diagnostic/backup approach, the parameter of speed has therefore been chosen. Contrary to other diagnostic systems, the presented system acts as a diagnostic but also backup system at the same time [13, 14, 15]. The primary way of obtaining of the engine's speed is the **optical sensor** [10, 11], while the other ways are synthetic model values of:

- **successive integration dynamic model**,
- **a neural network**.

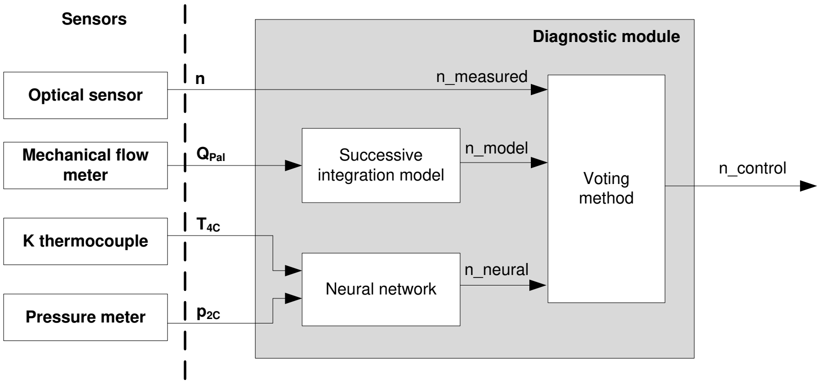


Figure 5
The structure of the diagnostic module

The structure of the diagnostic/backup module is shown in Figure 5. It shows that every input into the selection block has an independent sensor suite. This eliminates their mutual influence and increases reliability. The optical sensor can produce two types of erroneous outputs:

- **A random value –** caused by an electro-magnetic environmental conditions,

- **Sensor failure** – caused by a loss of power, loss of communication channel, or loss of reflex area on the compressor blade.

The designed structural scheme for the realization of the diagnostic/backup module utilizing the concepts of majority methods is shown in the Figure 6.

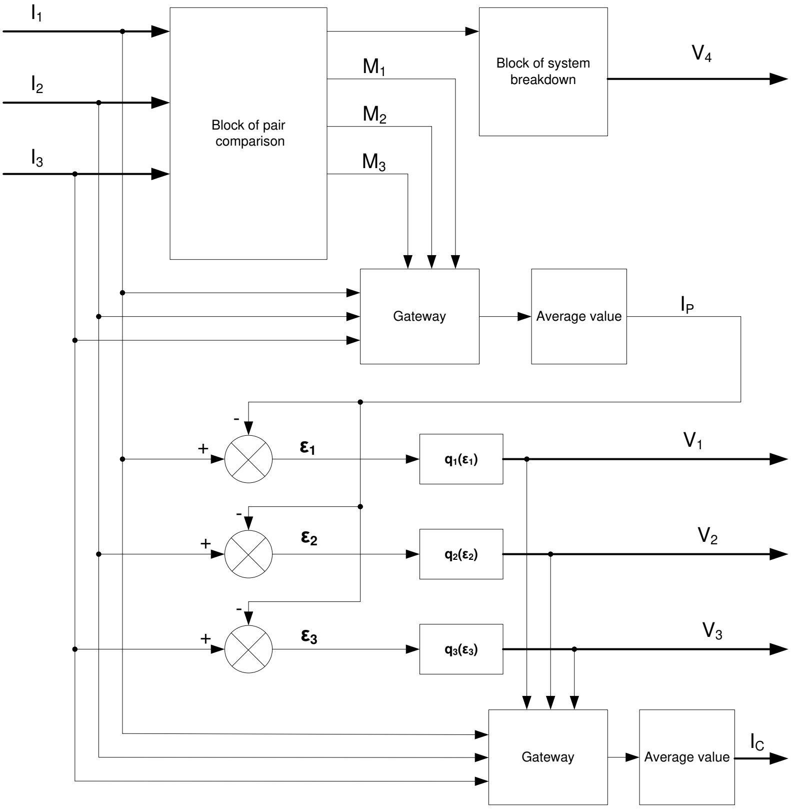


Figure 6

The scheme of the diagnostic/backup system for engine speed measurement

The operation of the diagnostic/backup module can be described as follows:

- The input of the system consists of three parameters representing the speed of engine's compressor *n* (I_1- optical sensor, I_2- successive integration model, I_3- neural network). These signals are transferred through a block of pair comparison, where the maximal errors between the compared pair are defined in Table 1. The allowed error is computed as the sum of the allowed errors for each pair. The optical sensor operates with a maximal error of 200 rpm. The models have defined errors as the maximal absolute error compared to real world testing data. The successive integration model has the tolerated error of 2510 rpm and a neural network has the maximal absolute error of 696 rpm.

Table 1

Pair comparison of individual inputs

Pair comparison of ***i-th*** **and** ***j-th input*****:**	**Allowed error margin** **($\varepsilon_{i,pov}$+ $\varepsilon_{j,pov}$)[rpm]:**
I_1 and I_2	2710
I_1 and I_3	896
I_2 and I_3	3006

- After pair comparison, the inputs in the defined tolerances are set and gating elements values M_1, M_2 and M_3 are set. These values {0,1} directly influence the inputs into the average value I_P.
- If all the gates are set to the value of *0*, the activation of the block – "Total Failure" is executed. The action element is then put into failsafe position and total failure is also signalized.
- I_P at last enters substraction elements where the error residuum ε_i is generated and transferred through the function of q_i. This function sets if the given input is without error or not. State is represented by signalization V_i. The allowed residuum are set as follows:
  - optical sensor

$$q_1 = \begin{Bmatrix} 1, for : \varepsilon_i \leq |2000| \\ 0, for : \varepsilon_i > |2000| \end{Bmatrix} \qquad (1)$$

  - artificial neural network

$$q_2 = \begin{Bmatrix} 1, for : \varepsilon_i \leq |1500| \\ 0, for : \varepsilon_i > |1500| \end{Bmatrix} \qquad (2)$$

  - successive integration model

$$q_3 = \begin{Bmatrix} 1, for : \varepsilon_i \leq |800| \\ 0, for : \varepsilon_i > |800| \end{Bmatrix} \qquad (3)$$

- The values V_i gate the input signals I_i and the arithmetic value I_C is then computed and enters the control circuit.

3 Practical Evaluation of the Proposed Model

The proposed and implemented diagnostic/backup system has been experimentally tested on an MPM-20 engine within its operational speed range. This means that all tests presented in the following figures contain real-world measured data where the diagnostic system is operating in real time and uses data obtained from the data acquisition system. The data acquisition system consists of sensors, National Instruments data acquisition hardware (cDAQ 9172, NI 9263, NI 9205, NI 9472, NI 9423, NI 9213) and National Instruments LabView software used for data acquisition. The system runs at a sampling rate of 100 Hz for analogue channels, at 10 Hz for speed measurements and at 100 Hz for computations. Everything is down sampled to the rate of 10 Hz in result, which is sufficient for a turbojet engine with the time constant of 2 seconds [10, 11].

The faultless operation of the system is shown on the graph presented in Figure 7. The engine was run for 70 seconds and the graph shows the course of its speed obtained by different means (optical sensor, polynomial model, neural network) during accelerations and decelerations and the output from the diagnostic/backup system. During the following experiments, failures of sensors were simulated as we cannot physically damage a sensor or computational model (control system input in general) so a special algorithm in LabView was prepared that triggered preprogrammed sensor failures by adding a preset value to its output or completely zeroing its output, thus generating an error signal.

The following real-world tests were executed:

- simulation of the random added values of individual inputs,
- simulation of failures of individual inputs,
- total failure of the sensors.

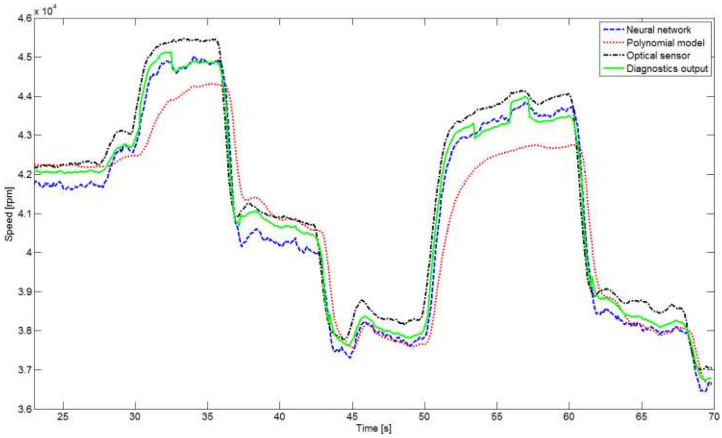


Figure 7

Outputs during faultless operation

Testing with Generated Random Faults

Individual outputs were burdened by added speed of 15000 rpm. The response of the system can be seen in Figure 8. The output in the graph shows that the generated errors of individual outputs did not translate into the final output of the diagnostic system. The figure also shows that some real failures occurred on the optical sensor (at 30 and 45 seconds). This test demonstrates that the proposed diagnostic/backup system can handle such failures and won't translate them further into the control system where the speed of the engine is used.

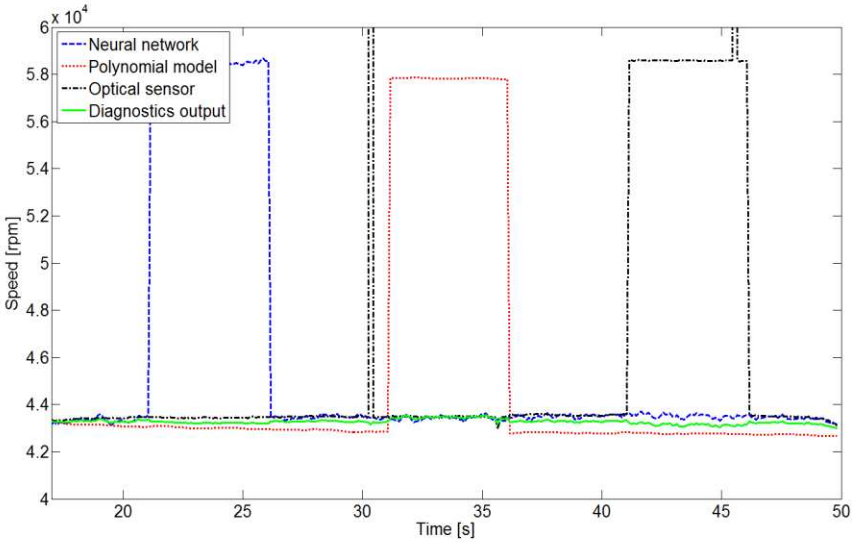


Figure 8
Measured data with applied random errors

Simulation of Failures of the Outputs

In this case during operation of the engine, failures of individual outputs were generated by setting them to zero. The resulting output is shown in Figure 9. The test again shows that output of the diagnostic/back-up module is not influenced by any of the failures. The green line represents the signal, which is sent into the control system and is not influenced by any of the failures.

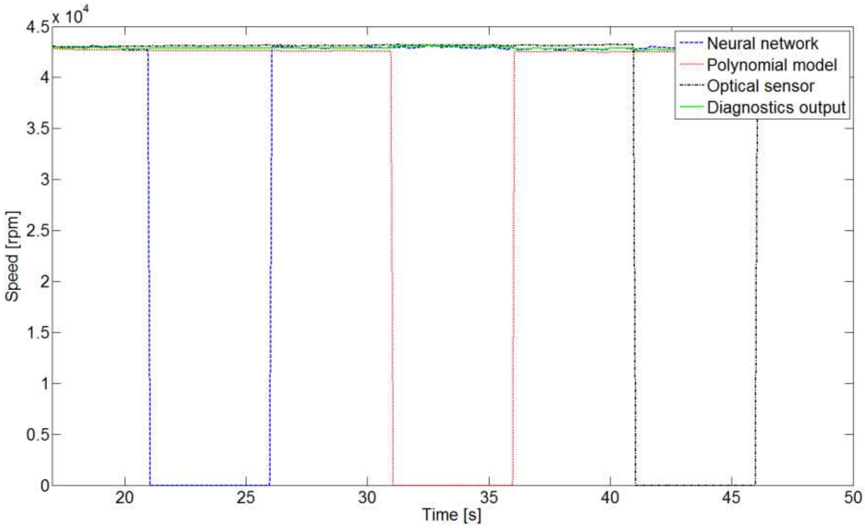


Figure 9

Measured data with failures of individual inputs

Total System Failure

The last experimentally evaluated part was the testing of the total failure mode. In this case, the model value of the neural network was increased by 25 000 rpm and the value of 15 000 rpm was added to the successive iteration model (Fig. 10).

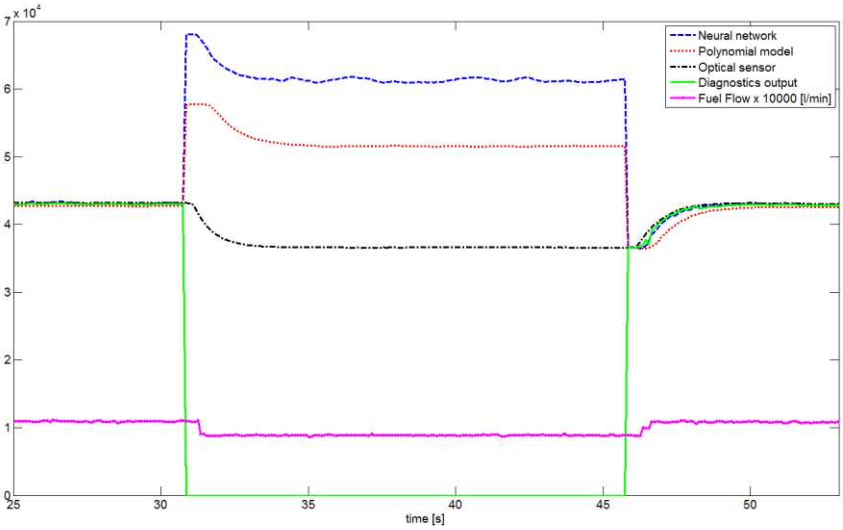


Figure 10

Response of the system to total failure with degraded control mode

Under such conditions, the voting method cannot decide which of the outputs is relevant, because every one of them exceeds the maximal value for pair comparison. The output of the diagnostic module is then set to zero with a signal of total failure mode in the graphic user interface and the activation of a failsafe regime of control. In the failsafe regime of control, the fuel flow supply into the engine is preset to 0.9 l/min. this is shown in the figure 10. If the errors are gone, the fuel flow supply is set back into normal operational mode and the control system transitions into normal control laws for the stable operating regime.

The occurrence of random failures was precisely simulated in the LabView environment, and it is possible to compare the outputs of the diagnostic/backup system with true sensor value of the speed.

During the test with random errors, the diagnostic system operated with a mean average error of MAE=210.116 rpm and maximal absolute average error MAAE=539.40 rpm. The representative graphical depiction of these courses is shown in Fig. 11.

In the experiment with individual input failures the errors were as follows MAE= 241.756 rpm a MAAE=577.2548 rpm. The course of the error is shown in Fig. 11.

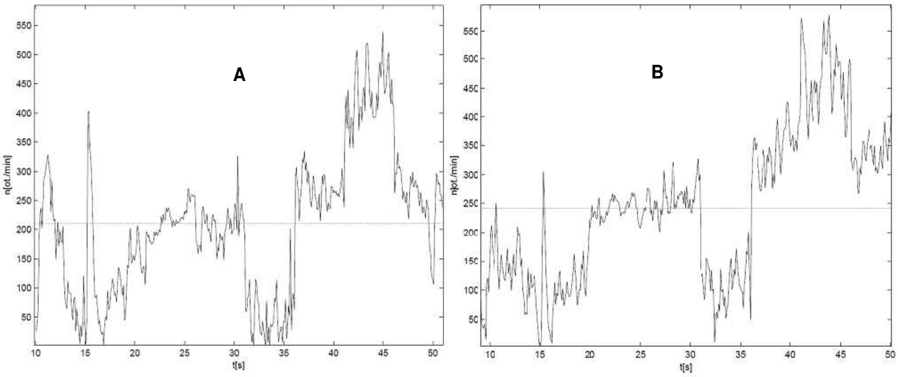


Figure 11

Difference of values from optical sensor and diagnostic system in the experiment with random values (A) and the experiment with input failures (B)

The presented figures show that the highest error occurred with the failure of the optical sensor (time between 42 and 47 seconds), as it operates with the highest precision and was taken as the reference in the creation of models. However, this error is on the level of 500 – 600 rpm (5% of the operating range) and is acceptable for the diagnostic system.

The real-world tests show that the designed and realized diagnostic/backup system operates correctly according to demands put on it. The method is relatively simple for implementation and its main advantage is combination of backup and diagnostic properties in a single unit contrary to other approaches in the field of turbojet engines [6, 7, 13].

Conclusions

The presented diagnostic/backup system utilizes the voting majority method, which is, however, expanded with the methodology of quorum and model based parameter evaluation. The main advantage of such a system is that it combines diagnostics and at the same time operates as a backup system as it uses all available data to synthesize output value. Applying neural networks and successive iteration computational models into the system allows us to increase redundancy of the system without adding any back-up sensor. This redundancy can be increased further by utilizing other measured inputs. The system is designed in a modular way and can be implemented to distributed computational architectures so each module can run on different hardware if needed. The other advantage is its incorporation into situational control systems, which allow controlling the engine under all situations including atypical ones.

Acknowledgement

The work presented in this paper was supported by VEGA, Grant Agency of Ministry of Education and Academy of Science of Slovak Republic under Grant No. 1/0298/12 – "*Digital control of complex systems with two degrees of freedom*" and grant No. 1/1117/11 – *"Integration of automatic flight control algorithms with control algorithms of aircraft turbocompressor engines"*. The work presented in this paper was also supported by KEGA under Grant No. 018TUKE-4/2012 – "*Progressive methods of education in the area of control and modeling of complex systems object oriented on aircraft turbo-compressor engines.*

References

[1] KELEMEN, M., LAZAR, T., KLECUN, R.: Ergatické systémy a a bezpečnosť v letectve (Ergatic systems and aviation safety). AOS gen. M.R. Štefánika, Liptovský Mikuláš, 2009, 316 pp, ISBN 978-80-8040-383-6

[2] LINKE-DIESINGER, Andreas: Systems od Commercial Turbofan Engines: An Introducion to Systems Functions. Hamburg, 2008. 239 pp. ISBN 078-3-540-73619-6

[3] ROLLS-ROYCE: The Jet Engine, Fifth edition, Great Britain 1996, 292 pp., ISBN 0902121235

[4] BOYCE P. MEHERWAN: Gas Turbine Engineering Handbook, Third Edition, Elsevier 2006, 935 pp., ISBN 0-88415-732-6

[5] WEIZHONG, Y., LI., C., J., GOEBEL, K.., F.: Multiple Classifier System for Aircraft Engine Fault Diagnosis, Proceedings of the 60th Meeting of the Society For Machinery Failure Prevention Technology MFPT), pp. 291-300, 6167, pp. 271-279, 2006

[6] ARMSTRONG J. B., SIMON D. L.: Implementation of an Integrated On-Board Aircraft Diagnostic System, NASA/TM—2012-217279, AIAA–2011–5859, 2012

[7] JAW C. Link, MATTINGLY D. Jack: Aircraft Engine Controls Design, System Analysis, And Health Monitoring, American Institute of Aeronautics and Astronautics, 2009, pp. 361, ISBN 978-1-60086-705-7

[8] KULIKOV Gennady G., THOMPSON A. Thompson: Dynamic Modelling of Gas Turbines Identification, Simulation, Condition Monitoring and Optimal Control, Springer 2004, 337 pp., ISBN 1852337842

[9] LAZAR, Tobiáš – MADARÁSZ, Ladislav (Eds.): Inovatívne výstupy z transformovaného experimentálneho pracoviska s malým prúdovým motorom (Inovative Outputs from the Transformed Experimental Laboratory with a Small Turbojet Engine). elfa, s.r.o. Košice, 348 pp. ISBN 978-80-8086-170-4 (2011)

[10] MADARÁSZ, Ladislav et al.: Situational Control Modeling and Diagnostics of Large Scale Systems. In: Towards Intelligent Engineering and Information Technology. Studies in Computational Intelligence 243. ISBN 978-3-642–03736-8, ISSN 1860-969X, 2009, Springer – Verlag Berlin Heidelbeg, pp. 153-164

[11] MADARÁSZ, Ladislav – ANDOGA, Rudolf – FŐZŐ, Ladislav: Intelligent Technologies in Modeling and Control of Turbojet Engines. In: New Trends in Technologies: Control, Management, Computational Intelligence and Network Systems, Meng Joo Er (Ed.), Sciyo, Available on: http://www.intechopen.com/articles/show/title/intelligent-technologies-in-modeling-and-control-of-turbojet-engines,. pp. 17-38, ISBN: 978-953-307-213-5, (2010)

[12] ANDOGA, Rudolf – MADARÁSZ, Ladislav – KAROĽ, Tomáš – FŐZŐ, Ladislav – GAŠPAR, Vladimír: Intelligent Supervisory System for small Turbo-jet. In: Aspects of Computational Intelligence: Theory and Applications (2012). Springer Verlag Berlin Heidelberg, 2012, pp. 85-104, ISBN 978-3-642-30667-9, ISSN 2193-9411

[13] KYRIAZIS A., MATHIOUDAKIS K.: Gas Turbine Fault Diagnosis Using Fuzzy-based Decision Fusion, JOURNAL OF PROPULSION AND POWER Vol. 25, No. 2, March–April 2009, ISSN 0748-4658

[14] PEČINKA, Jiří; JÍLEK, Adolf: Preliminary Design of a low-cost Mobile Test Cell for Small Gas Turbine Engines (GT2012-69419) In: Proceedings of ASME Turbo Expo 2012. Copenhagen: ASME, 2012

[15] Yi-Jen Mon, Chih-Min Lin, Imre J. Rudas: Wireless Sensor Network (WSN) Control for Indoor Temperature Monitoring, Acta Polytechnica Hungarica Vol. 9, No. 7, ISSN 1785-8860, 2012

Implementation of a Distributed Genetic Algorithm for Parameter Optimization in a Cell Nuclei Detection Project

Sándor Szénási, Zoltán Vámossy

Óbuda University, Bécsi 96/B, H-1034 Budapest, Hungary
szenasi.sandor@nik.uni-obuda.hu
vamossy.zoltan@nik.uni-obuda.hu

Abstract: The processing of microscopic tissue images and especially the detection of cell nuclei is nowadays done more and more using digital imagery and special immunodiagnostic software products. One of the most promising image segmentation methods is region growing, but this algorithm is very sensitive to the appropriate setting of different parameters, and the long runtime due to its high computing demand reduces its practical usability. As a result of our research, we managed to develop a data-parallel region growing algorithm that is two or three times faster than the original sequential version. The paper summarizes our results: the development of an evolution-based algorithm that was used to successfully determine a set of parameters that could be used to achieve significantly better accuracy than the already existing parameters.

Keywords: tissue image segmentation; data parallel algorithm; GPGPU, genetic algorithm; distributed system

1 Introduction

Nowadays the digital microscope is becoming a more and more popular device among pathologists. In addition to several improvements of the up-to-date devices (good quality, focused photos, the possibilities for objective measurements, etc.) it is worth mentioning that in addition to the suitable IT background, the images gained this way can be subjected to numerous other processes, in addition to simply viewing them once, which can promote later consultation (distribution, categorization [1], remote access, etc.) and can provide for preliminary or post processing of tissue samples.

This kind of processing offers a very promising way of using different segmentation processes with the images received, and in this way the different components of the tissues can be well separated. Appropriately precise recognition of the tissue components (morphologic and morphometric parameters of these components) can provide a safe background for automated status analysis of the examined patients, or at least it can aid the work of the pathologists with this pre-processing. By means of separating the sick and the healthy tissue parts, the Labeling Index of immunohistochemical reactions used for examining the structures and the evaluation (Scoring) parameters in figures can be given more accurately.

In the course of our research we have analyzed tissue samples taken from haematoxylin-eosin stained colonic tissue samples (Figure 1). The most important structures in these cases that are worth separation are the following: cell nuclei, glands and epithelium [2]. Of course there are a lot of methods for their recognition, but most of them begin with the precise determination of the position of cell nuclei and, on this basis, then attempt to state the position of the other components. There are several alternative methods of searching for cell nuclei; we have improved one of the most promising solutions, namely region growing.

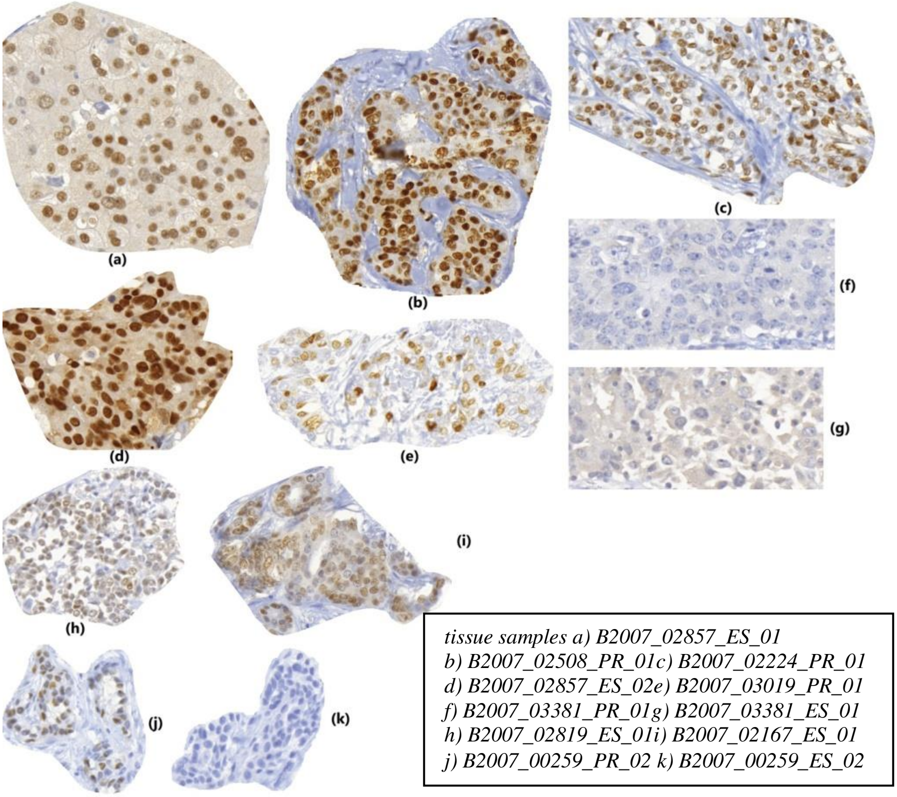
1.1. GPGPU-based Region Growing Algorithm

The process of region growing has already been well proven in practice. It is quite accurate: however, the long runtime due to its high computing demand reduces its practical usability. At the first stage, this could be improved by the development of a new algorithm running in a parallel environment that is implemented on data-parallel GPGPU, resulting in a 3-to-4-fold rate increase [3], which can be naturally further raised by using more GPUs.

The first step in region growing is to select a set of seed points, which requires some idea about the pixels of the required region (we assume that nuclei are usually darker than their environment). In the next step, the neighboring pixels of the initial seed points are examined and it is determined whether the pixel neighbors should be added to the region or not (by minimizing a cost function). This process is iterated until some exit condition is met.

1.1.1. Searching for Seed Points

The search for seed points is a nicely parallelizable task, since our aim is to find the point with the highest intensity that complies with some rules (it cannot be inside a previously found region, etc.). When running a sequential algorithm on the CPU, this means a single point, but in the case of the GPU, this can result in multiple points, because it is possible to execute multiple cell nucleus searches in multiple blocks. In the latter case, the adjacent seed points can cause problems, since the parallelized search of those can result in overlapping cell nuclei, which would require a lot of computational time to administer.



tissue samples a) B2007_02857_ES_01
b) B2007_02508_PR_01c) B2007_02224_PR_01
d) B2007_02857_ES_02e) B2007_03019_PR_01
f) B2007_03381_PR_01g) B2007_03381_ES_01
h) B2007_02819_ES_01i) B2007_02167_ES_01
j) B2007_00259_PR_02 k) B2007_00259_ES_02

Figure 1
The selected tissue samples (masked)

Luckily enough, we know what the maximum radius of a cell nucleus can be in an image with a given zoom; so we can presume that the searches started from two seed points (that are at least four times further apart than this known distance) can be considered as independent searches; so they can be launched in a parallelized way. The iteration is continued until the thread runs out of possible seed points, or the required amount of points is gathered for the starting of the efficient region growing.

1.1.2. Parallel Region Growing

Region growing itself consists of the following consecutive steps [3], which depend on each other: first comes the search for possible new contour expanding points; the next step is the evaluation of the available points; then comes the selection of the best valued point; and the last step is the expansion of the area with the selected point. These steps can be very well parallelized on their own, but every operation needs the output of the previous step, so we definitely need the introduction of synchronization points. This significantly reduces the count of the possible solutions, since when using the CUDA environment, we can achieve synchronization methods only within one single block. Thus, it seems practical to assign a single block to the processing of one single cell nucleus.

Region growing itself iterates three consecutive steps until one of the stop conditions is met [3]: (1) It examines the possible directions in which the contour can be expanded. The full four-neighborhood inspection is evidently only required around the last accepted contour point (when starting the kernel, this means the starting seed point). Since the examinations of the neighboring points do not depend on each other, this can be parallelized as well. (2) The various different contour points must be evaluated to decide in which direction the known region should be expanded. For this, a cost function [4] must be calculated for every point that uses more parameters (intensity of the neighbors, average intensity of the region, distance of the point from the seed, etc.). As the values change at the insertion of every new point, they have to be re-calculated at every iteration for every point. This is however a typical data parallelized calculation, so it can be very well parallelized on the GPU. Every thread counts the cost of a single contour point. (3) The contour point with the smallest cost must be selected.

After every iteration, a fitness function is evaluated that reflects the intensity differences between the region's inner and outer contour, and the region's circularity. The process continues until the region reaches the maximum size (in pixels or in radius), and its result is the state where the maximum fitness was reached.

For the case when two cell nuclei intersect each other, another stop condition is inserted into the algorithm. According to our experience, the overgrowing of a region into another nucleus can be detected from the intensity changes; the constantly decreasing intensity suddenly starts to increase. Due to this phenomenon, we calculate the time differential of the intensity-differences, and if the resulting function passes a given value, then we stop the region growing.

1.2. Parameter Optimization

However, the region growing algorithm prepared this way has several parameters (parameters of the different filters, parameters of intensity-contrast-dimensions, etc.) the fine tuning of which is as important as the previously mentioned speed increase. Due to the large number of parameters and their reasonably large target set, defining the values manually seems hopeless, so we had to develop an optimization algorithm.

The basis of optimization is the comparison of test results gained with the help of the parameter sets and using the region growing algorithm with the reference results approved as good and then making conclusions from the difference in the two samples as to the appropriateness of the tested parameter set. We have, as a reference, 41 pieces of tissue samples, processed by trained pathologists; they marked the precise position of the cell nuclei (later we will refer to them as “gold standards”). The test samples are generated on the basis of the given raw tissue sample with the help of the previously developed region growing algorithm. A detailed description of the genetic algorithm managing the populations, including the individual parameter sets, is included in another paper; here we only wish to detail how searching of the parameters could be accelerated. However, it was immediately revealed in practice that the parameter range to be tested is quite large (27 different, independent parameters), and testing of the individual parameter sets is rather time-consuming (because the region growing must be run on the tissue sample, then the gained results have to be compared to the reference result given by the pathologists), consequently the solution that can be run with the classic sequential 1 processor would not give results within a reasonable period of time, in practice. Due to the fact that time-consuming parts (region growing and comparison) are given, and the genetic operators themselves do not demand considerable time, so the reduction of runtime can only be achieved by means of making the genetic algorithm and the evaluations parallel.

2 The Evolutionary Algorithm for Parameter Optimization

2.1. Initial Generation

The initial generation is usually built up using randomly generated instances. If we need to further refine some previously tested parameters, then we have the possibility to place them into the initial generation, but in our case we usually want to start a completely new search. Table 1 shows some of the parameters that have known bounds [5] with a standard distribution. With some technical parameters, it is not possible to perform such preliminary tests; in these cases, the initial values of the parameters are distributed using the currently known best set of parameters.

2.2. Evaluation of Generations

After we create a new generation, we have to evaluate all instances. Since in our case the fitness value of a given instance is determined based on how well the given parameter set stored in the genes performs with the region growing cell nuclei detector algorithm, this means that we need two steps to determine this fitness value:

Name	lower bound	upper bound
Cell nuclei size	34	882
Cell nuclei radius	4	23
Cell nuclei circularity	27.66	97.1
Cell nuclei average intensity (RGB avg.)	36.59	205.01
Seed point intensity (RGB avg.)	0	251

Table 1
Bounds for some initial parameter

1. First, we must execute the region growing cell nuclei detector algorithm with the given set of parameters on several slides.
2. After this, the results produced by the algorithm with the given parameter set (test result) must be compared with the manual annotations of the Gold Standard slides (the reference result). By averaging the results of the comparisons (for all slides) we get the fitness value for the given parameter set. This value can be used to evaluate the fitness of the given chromosome.

To compare the test and reference results of cell nuclei search methods, we have already developed an evaluation algorithm [6]. There are several approaches of accuracy calculation (e.g. some fuzzy models [7]). Our method is based on the very often used confusion matrix [8] that can be constructed using a comparison of the two result sets. The matrix (assuming we have two possible outcomes) contains the number of true positive (TP), true negative (TN), false negative (FN) and false positive (FP). The accuracy of the region growing is simply the ratio of sum the true values and the sums of all cases.

2.3. Accuracy of One Parameter Set

Our measurement number does not only reflect a pixel-by-pixel comparison; instead it starts by matching the cell nuclei together in the reference results and in the test results. One cell nucleus from the reference result set can only have one matching cell nucleus in the test result set, and this is true the other way around too: one cell nucleus from the test result set can only have one matching cell nucleus in the reference result set. After the matching of the cell nuclei, the next step is the similarity comparison between the paired elements.

The very critical point of the evaluation is how the cell nuclei are matched against each other in the reference and the test result sets, because obviously this greatly affects the final result. Since this pairing can be done in several ways (due to the overlapping cell nuclei), it is important that from the several possible pair combinations we have to use the optimal: the one that gets the highest final score.

During the practical analysis of the results, we found that on areas where cells are located very densely, we have to loop through a very long chain of overlapped cells, which results in groups that contain very many cell nuclei from the test and from the reference sets as well. Since increasing the number of elements in a group exponentially increases the processing time of the group, it is practical to find some efficient algorithm for the matching; we use a modified backtracking [9] search algorithm, which searches and returns the optimal pairing of a group containing test and reference cell nuclei.

The algorithm above should be executed for every group, and in this way the optimally paired elements can be located (including the elements that are alone in their group and the elements that cannot be paired at all). Evaluating every pair (and single element), and summing up the values, we can determine the weighted total *TP*, *FP*, *FN* pixel numbers that represent the whole solution. These can be interpreted on their own or in a simple way using the aforementioned accuracy equation. For the genetic algorithm, this will be the fitness value of the parameter set represented by the given instance.

2.4. Apply Genetic Operators

2.4.1. Selection of Parents and Survivors

To select the parents and the survivors, we use the so-called roulette wheel selection method [10]. After the evaluation of the current generation, we know the fitness values for all the instances. Knowing these, and P_i: the probability to select the instance #i, F_i: the fitness value for the instance #i, Min(F): the smallest fitness value for the current generation. We have developed the following formula:

$$P_i = \frac{F_i - Min(F)}{\sum_j F_j - Min(F)} \tag{1}$$

Due to the large number of parameters, the search space is reasonably large, so the occasionally occurring instances with exceptionally high fitness value can easily disappear in the next generation due to the obligatory random crossovers. For this reason (similar to elitism [11]) the instances with the highest fitness values are carried along (without crossovers or mutations) into the next generation (not only the single best instance, but the top 30 instances, so 10% of the generation is selected this way). This slightly decreases the number of trial runs per generation, but in this way we guarantee that the best chromosomes are kept and their genes remain constant. A side effect of this is that we get a monotonically increasing series of fitness values for the best instances.

2.4.2. Crossover

It is hard to determine the most effective crossover method in advance. It is advised to use the two-point crossover in the case of a large population, the uniform crossover in the case of a smaller population, less cut points in the case of short chromosomes, and more cut points in the case of large chromosomes. In our case, the size of the population can be considered as reasonably small (because the evaluation of the single instances can be very time consuming, and so we cannot use a large population), while the chromosomes are considered to be reasonably large (27 parameters, several hundred bits altogether). Because of these, we clearly need to use the uniform crossover method.

In our case, it is not feasible to perform the uniform crossover for any bit, because there are some parameters that have to comply with some additional rules (e.g. divisibility), and the bitwise mixture of those can easily lead us to values that do not belong to the target set. For these reasons, during the crossover we only combine whole genes; for every gene of a new chromosome we use a random number to determine which parent's gene is inherited. To converge faster, the parent's gene that belongs to the parent with the greater fitness value has bigger priority. Knowing the probability (P_a) that the gene of parent A is inherited, the probability (P_b:) that the gene of parent B is inherited, the fitness value (F_a) for parent A, the fitness value (F_b) for parent B, and min(F) is the smallest fitness value for the current generation, for every gene. We have developed the following formula used to determine which parent's gene is used:

$$P_a = \frac{F_a - Min(F)}{F_a + F_b - 2*Min(F)} \tag{2}$$

$$P_b = \frac{F_b - Min(F)}{F_a + F_b - 2*Min(F)} \tag{3}$$

We have already tried another crossover method (where $P_a = P_b = 0.5$) but it caused significantly slower convergence.

2.4.3. Mutation

It is feasible to use the same type of mutations with genetic algorithms as are used in nature. This is especially important if, after many generations, the changes are reasonably small, and the various parameters have settled around some values. We have defined the mutation arbitrarily (based on the initial test runs) according to the following rules:

- The probability of a mutation is separately 10% in every new generation for every parameter.
- The size of the mutation is random: there is a 60% chance for a small, a 30% probability for a medium-sized, and a 10% probability for a large mutation.

2.5. Implementation

The actual implementation was done according to the following criteria:

- We are searching for the values of 27 parameters.
- The initial generation has 3,000 chromosomes.
- Every following generation has 300 chromosomes.
- Every parameter set (every instance) is tested against 11 representative tissue samples (Figure 1).

3 Distributed Genetic Algorithms

We can find in the professional literature several methods of making genetic algorithms parallel. The most fundamental of them are [12] and [13]:

A. *Systems based on one population*: in these cases the genetic algorithm itself is actually not different from the classic sequential solutions. It is only that implementation is attempted to be modified so that the individual genetic procedures can be performed with the same results, but with a shorter runtime.

  1. *Compiler based automatic parallelism*: Parallelization is implemented in this case at a quite low level; here we try to utilize the possibilities of the hardware. In view of the fact that this is mainly a technical possibility, here we can rely on the help of the different compilers in most of the cases because they automatically perform different optimizations in time of compiling [14].
  2. *One population - parallel evaluation/crossover/mutation*: Parallelism is implemented at a higher level in this case. We try to parallelize not only the individual elementary operations themselves, but larger tasks are considered as atomic and they are solved independently of each other. This coarser granularity makes it possible to extend parallel processing to multi-processor or multi-node systems as well [15].

B. *Systems based on several populations*: in these cases we do not work necessarily with a global population, but we can manage the development of several, independent populations at the same time. Of course different kinds of relations may be constructed among them, and they can communicate and the results can be compared at the end of processing. Granularity in the names of the two methods is nothing else but the ratio of computations and communication. When this value is high, then we can speak about a coarse grain algorithm; when it is low, then the algorithm is fine-grained.

1. *Coarse grain PGAs*: The classic case of a coarse grain is when each executing unit is running its own genetic algorithm in the distributed system on an independent population with all of its operations (selection, crossing, mutation). This can be *mutually exclusive*, if the independent units are completely isolated from each other. Or it can be *non mutually exclusive,* when there is communication at some level among the populations (typically they distribute the chromosomes with best results among each other). Several other implementations can be imagined. The main point is that the principle of the genetic algorithm does not change, and sequential algorithms run again, but in many populations at the same time [16].
2. *Fine grain PGAs*: While we often imagine completely separated populations (islands) in the case of coarse granularity, the typical example of fine grain can be a large-sized global population, whose elements are organized in a grid and every element can communicate only with its neighbors (of course with parallelization at the level of the individual instances). Depending on the size of the executing units and of the population, this can be implemented technically in several ways; the arrangement itself and communication with the neighbors demand individual design in each case [17], [18].

3.1. Methods Working on a Global Population

Methods A.1 and A.2 are fundamentally based on the same principles as the traditional sequential genetic algorithms; however they promise quicker implementation depending on the architecture. Due to the fact that all of the executing units work on the same population, this raises several problems (closures, communication, time loss due to waiting for each other), these losses must be considered by all means before making decision. Based on the recommendations of the professional literature [14], it is clear that they only have raison d'être if the operations with the individual instances are computationally demanding and time consuming.

This is typically true in our case, since full region growing must be run for each set of parameters, and then the evaluation of the gained results must be performed (because we need fairly exact scoring here, it takes longer than region growing itself) and it should be done not only for one but for several tissue samples after each other. As compared to the costs of these operations, the resource demand of the different genetic operators (parent selection, crossing, mutation) themselves are negligible, so it is practically worth choosing from the A versions.

A.2 offers more possibilities for us because local optimization provided by the compilers gives fairly limited possibilities, and here we would like to accelerate not the genetic operators but the evaluation of the fitness function of the individual elements. The master-slave method offers a well-implementable solution for this; it can appropriately and efficiently utilize our available resources. In addition, we have found several where it lives up to the promises [19], [20], [21], [22].

3.2. Methods Working on Several Populations

Method B.1 is very popular, too, though it has several limits [21]:

- Perfect utilization of the processing units demand more attention when dimensioning the populations.
- Certain steps of evolution can be reproduced only with difficulty due to asynchronous nature of population processing and migration.
- Communication among the independent islands makes the model much more complicated.

And even if the benefits are undoubted in certain cases, they are less dominant for us because, due to the above mentioned difficult fitness evaluation, it is clear that scarcity of the resources does not enable us to process a fairly large volume of instances. And because the size of the population will be the bottleneck for the genetic algorithm, maintaining further populations will not give us an advantage.

In the same way, we cannot utilize the benefits of method B.2. It is clear that in our task the evaluation of the fitness function demands the most resources; the other operations (selection, crossing, mutation) are negligible in this respect. However, in this case, the fine-grained method does not show benefit as compared to the client-server implementation, but it involves fairly many limitations. The greatest problem is loss of liberty in parent selection. Since it is evident that when using the traditional sequential methods any two elements of the full population can be chosen as parents, then in the case of the coarse-grained method, we can choose only from the instances of the given sub-population, and in the case of fine-grained method even stricter limitations must be accepted, since we can choose only from the direct neighbors here [23].

Based on the above, although algorithms working with more populations are more popular, their use in our case does not seem to be practical. One of their great benefits (reduction in communication) does not mean essential development, because it has negligible demand of resource besides evaluation of the fitness function on the one hand; and due to the same difficult evaluation our target is to find the quickest result and this can be probably achieved if we can freely choose from every member of the full global population in the course of crossings.

3.3. Hybrid Solutions

Based on the first experiments it is practical to develop a hybrid system [24] later on. Namely it became clear during the first runs that the processing time of certain instances can importantly differ from each other, and with given combinations of the parameters the runtime of region growing can be either multiple of the average value. Consequently, sometimes such an unfavorable condition may occur that, although each instance of the generation with one exception was processed, the parallel slaves are forced to wait for the result of the last processing. These idle times could be surmounted if the waiting clients would be assigned to processing the instances of another population.

3.4. Master-Slave Implementation

Even though there are numerous advantages of master-slave (simple implementation; the principle is practically the same as the sequential genetic algorithms so it can be simply adapted; it provides very good performance in many cases), one of the biggest problems is fairly high communication demand. That is why it is practical to undertake preliminary examinations whether it is a real alternative.

Execution time of a master-slave GA is made up of two components:

- *Time spent on computation*: in this case it includes evaluation of the fitness function in the first place. Based on processing of the available 1,550,318 instances the below average value was measured: 1498 ms is the time of region growing run on one image; 8,249 ms is the average time of the following processing. The time of the different genetic operators is practically negligible (0.16 ms in the case of one instance). In the case of the population chosen by us, it is 50 ms per generation.
- *Time spent on communication*: we can speak about communication in master-slave system, when the master distributes the tasks among the clients and when they return the fitness values computed by them. This value depends on the different hardware devices used (network, interfaces) on the one hand, and on the protocol used for communication on the other hand. In this case we can quite well reduce the time of communication with this latter, since only a minimum volume of data needs to be transferred between the master and the slave. At the time of distributing the tasks, the master transmits the required parameters of the region growing, which is (even in a not very economical plain text format) 70.67 bytes on average. Sending of the results means only the transmission of some numerical values; in case of 11 images, it is 539.32 bytes on average.

4 Development of the Distributed System

In addition to principle implementation we can find several other possibilities in professional literature in respect to the master-slave model; we developed our own model after having studied them.

There are several standardized techniques available today for developing the distributed system, including very complex ones suitable for industrial applications, too [25], [26], [27] (cloud, grid, etc.). Although they could have provided a very fashionable and elegant solution for running of the genetic algorithm, their use would involve disproportionally great extra resources, which is unreasonable in the present experimental phase. The implementation of the genetic algorithm under development could be modified according to the above aspects, but we would have to adapt to the already existing external modules. These, however, would require important alterations for the purpose of cooperation with the above-mentioned standard systems (the region growing algorithm, the evaluation algorithm), and in certain cases this seems to be unfeasible (the GPU-based region growing algorithm).

The emerging special problems demand specialized solutions, and therefore it is worth returning again to the solutions previously applied for the distributed systems. Although they require a bit more work (the framework is not ready, and it has to be established), as a result, the final achievement will meet the demands in every respect.

During our search, we were able to use infrastructure of the Óbuda University and the resources of some remote computers; however, they raise some special demands, which are all supported by the newly developed system:

- The most important aspect is the fact that how many clients can be started changes dynamically over time (we can use for searching only the currently free resources, but they should be completely used if it is possible). The system must support the entering of new clients at optional times, as well as the exiting of the available clients. The system is based on the classical master-worker model [28].
- It is essential that the system should use only the possible simplest and usual communication modes (protocols, ports) since it is possible that some of the clients will try to access the server from behind a fire wall.
- The installation of the required programs must be as easy as possible for the individual clients; or ideally, no installation should be required. In addition, the possibility of automatic updating should be possible without separate, manual access to the individual work stations.

4.1. Communication between the Genetic Base and the Communication Layer

It was an essential aspect when developing the system that it should be flexible thus enabling simple development of further genetic operators (for the genetic algorithm itself) or protocols (which ensure master-slave communication) later on.

The genetic operations can be located practically within one component in optional implementation; the system only requires the realization of the bellow interface (where type *Genom* includes the data of one chromosome, mainly in the form of a vector including integers):

- *Genom CreateRootItem()* - It creates a new instance with random parameters. It is called only in case of creating the instances of the initial, randomly created generation.
- *SelectParents(out Genom g1, out Genom g2)* – It chooses two parents (g1 and g2) randomly from the list of the stored instances. Select operator can be implemented by realization of this in compliance with the condition of how we would prefer the instances with better fitness function.
- *Genom CreateChild(Genom p1, Genom p2)-* Based on two parents transferred as parameters it establishes a third instance and then returns it. Actually it corresponds to the cross operator, and we can determine here which genes should the new instance get from the parents.
- *Mutation(Genom gen)* - Properties of the unit transferred as the parameter can be optionally changed. It complies with the mutation operator usual in genetic algorithms, and implementation is completely optional.
- *PrepareParentSelector()* - It is an auxiliary technical method that runs exactly once before the above methods. It is practical to implement different initializations here (e.g. in the case of choosing the roulette wheel method, the wheel itself has to be created and initialized).
- *CONST_GENOM_ROOT_CNT* – Size of generation #0.
- *CONST_GENOM_CNT* – Size of all generations excluding generation #0.
- *CONST_ELITISM_PRCNT* – Percentage value of elitism. The best instance from each generation that complies with this will automatically be passed to the next generation.

Having implemented the above abstract methods the framework is suitable for development of the initial generation and application of the genetic operator needed later on. Thanks to the above abstraction, the technical details have been well separated from the genetic operations, and thus their complex development becomes possible, e.g. the implementation of the already mentioned distributed operation during evaluation of the fitness function.

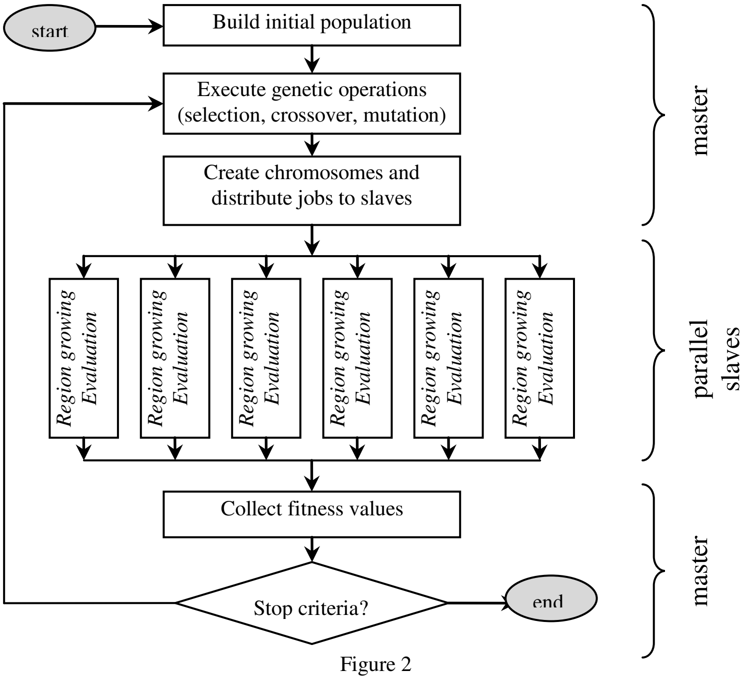

start
Build initial population
Execute genetic operations (selection, crossover, mutation)
Create chromosomes and distribute jobs to slaves
Region growing Evaluation
Region growing Evaluation
Region growing Evaluation
Region growing Evaluation
Region growing Evaluation
Region growing Evaluation
Collect fitness values
Stop criteria?
end
master
parallel slaves
master

Figure 2
Main steps of our distributed genetic algorithm

4.2. Implementation of the Master-Slave Protocol

The operating framework executes the required genetic operators and also manages the evaluation of the fitness function. This latter is executed in distributed mode (Figure 2) and this provides for determination of the protocol to be used. This is important because the operation can be customized in this way depending on how the master-slave settings have been implemented: among processes within one computer (communication actually does not need a network, and it is more practical to use process-level communication supported by the operating system) or among several computers in a local network or among remote computers (it involves significant limits if these computers are behind different firewalls).

Communication between the two sides is realized by an interim component which provides for services with the implementation of the below interface (among others), both towards the master and the slave sides:

Methods to be called from the master side:

- *ProcessNextGeneration(generationID, List of genoms)*: master transfers identification of the next generation as well as the list of instances included with this call. The function of the communication component is to run the fitness function evaluation for each of the chromosome received as parameters.

- *ContinueGeneration(generationID)*: due to technical reasons, the continuation of an already finished generation or one just under processing may become necessary; the master can initiate this by calling this method.

Methods to be called from the slave side:

- *LoadWaitingPackets(generationID)*: a slave can indicate through this method if it would like to join the processing of the given generation.
- *Genom LockNextProcessable()*: it reserves and loads the data of the instance to be evaluated next. Its return value is a *Genom* structure that includes all of the genes (in the present case, the parameters of region growing).
- *FinishAndSaveScore(Genom, List of score)*: after evaluation of the fitness function, the slave can reload the results by this method. Due to the fact that each parameter set has to be run for several tissue samples, the result will be a vector consisting of floating point numbers.

Thanks to the object-oriented approach, the above interface can be implemented in optional form, depending on the condition of how we want to solve actual communication. There are several communication tools to implement parallel genetic algorithms [13]. First, we implemented a version based on FTP protocol [29]. It uses, in fact, a third level for communication, since the server uploads the data on the chromosomes to be processed to a FTP server, the clients download the elements to be processed for themselves and then they upload the results back to the same place.

It has the advantage that no direct contact between the master and slave computers is required; the only key point is that an interim FTP server must be used that is visible from everywhere. Usually FTP protocol itself is allowed on firewalls, so a new client can fairly easily be initiated, if required.

4.3. Ensuring Robustness

During operation of the system robustness is of key importance. Being a distributed system, hundreds of clients can work at the same time, so possibly emerging errors cannot be monitored or repaired manually during runtime; the system should manage every problem. It must be done in such as way that the basic operation will not be affected, but running of the system with the best utilization enjoys priority.

Elimination of the most fundamental problems is simply a task of programming technique. Problems can emerge at any time during operation, e.g. network related errors (network cannot be accessed, server cannot be accessed, etc.) or the outage of hardware devices for shorter or longer periods (e.g. the server is down, certain clients shut down, power failure, etc.). These problems can be eliminated mainly with a well-developed application architecture; the system is prepared for faulty external network (sending of data from the network, receiving of data from the network) or errors in starting further processes (starting of region growing application), so (depending on the nature of the error) it restarts the operation or perhaps tries to repair it.

Inappropriately selected parameters can cause trouble, too, which makes the region growing and the operation of the evaluation algorithm impossible, or make them more difficult (slowing all down). Because we do not know the relations between the parameters, these problematic parameter sets can emerge at any time and the errors caused by them must be prepared for: empty result, error in program run, increased runtime, etc.

4.3.1. Empty Result

Because we try continuously to alter a fairly large number of parameters in loose relation with each other, parameter sets which are inappropriate for the region growing algorithm frequently emerge (e.g. when the minimum region size given as a parameter is larger than the maximum region size as a parameter). Of course parameter hierarchy and their effect on each other is much more complicated in reality than this, so the involvement of a complex, preliminary control system that will filter out all of the practically meaningless parameters would demand a lot of time (if it could be implemented at all).

We applied in practice the following method: we run the region growing for each generated parameter set; in cases when the parameters are in contradiction, then some kind of faulty response is expected as a final result (typically a response that the program did not find any cell nucleus). Although these evaluations also demand resources, they do not affect the path leading to the result, because the false results coming from the faulty parameters are given rather bad scores during evaluation, so these chromosomes will screen themselves out in the next generations.

4.3.2. Error in Program Run

The inappropriate selection of certain parameters will lead to even more critical results; it can cause a shutdown or a complete freeze of the program. Parameters of the different filters are typical, and they can be rather variable, e.g. they can determine certain ratio in case of the window sizes, or width of the window can be only an even number (in case of some filters), etc. In the ideal case, these are already revealed before the running of the algorithm during preliminary checking; sometimes, however, these are revealed only during running and then subsequent freezing of the program (because region growing uses external components such as different filters, which can behave completely unpredictably in case of faulty data, so freezing cannot be prevented in every case.)

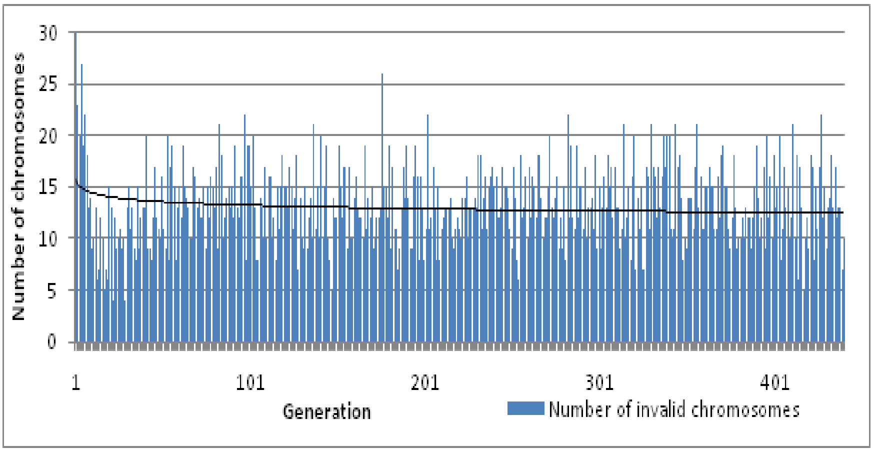


Diagram 1

Number of invalid chromosomes by generation

That is why a special control has been built in; should the program stop with an error code when evaluating any image, or should it freeze during operation of the region growing application, then the client automatically considers it as 0% result, finishes processing, does not continue with the next image belonging to this parameter set, and invalidates the results of the previous images. It is clear that, because we would not like to deal with a set of parameters later that is not capable of completely processing each image, irrespective of the fact what result has been received with the images where no errors emerged. Although a complex pre-filtering module could be done here that would examine the correctness of the input parameters, we followed here the method that we let the genetic algorithm filter out the faulty parameters themselves.

It works and it is well seen from the post-statistics that while a fairly large part of the components in some first generation stored unusable set of parameters (but do not forget that the initial generation has been actually created with random numbers in certain given intervals, so it is evident that too many contradicting parameters have been made), they have been rapidly cleared out in the later generations; after the 5th generation, these kind of parameters were created only rarely (of course they will never disappear due to mutations and crossings without control, but thanks to the above treatment method they will never get into the next generation). This is shown well by Diagram 1.

4.3.3. Increased Runtime

Chromosomes that do not cause any trouble but importantly do increase the resource demand of processing result in a much greater problem than the above problems. The processing of one image took about 10 seconds in practice, and this value can increase many times with certain settings. This would not cause any problem in itself, because the clients work irrespective of each other; the problem occurs because each result of the previous generation must be summarized for starting every generation. That is, if a client gets a relatively long-lasting processing, then the other clients (having finished processing each elements of the given generation, lacking other possibilities) are must only wait. In this period, naturally we can use only a fraction of the available resources, so this idle time must be minimized by any means possible, even if certain inaccuracy occurs. These kind of extremely long processes are worth stopping. In experiments, we limited this time to 1 minute.

Every stoppage due to the time limit involves a negligence of the rules of genetic algorithms. However, it is worth considering that no accuracy is the strength of GA implementation anyway; the more detailed the analysis is, the more important it is that more generations be created and that the parameters which seem to be viable be combined as many ways as is possible. This is fairly resource demanding, so it seems to be practical to try many hundreds of new combinations by using all of our resources, instead of letting the complete system wait for minutes because of the precise evaluation of a single chromosome.

Although the present article speaks only of searching for the ideal parameters, which usually means the set of parameters offering the most accurate result possible, it is worth of considering the fact that region growing is a rather computationally-demanding algorithm, so the incorrect selection of the parameters can either result in a practically unusable algorithm. Considering this aspect, stopping those evaluations that involve unacceptable processing times is even more justified, since they could not be used in practice, irrespective of accuracy. And although it is not the primary aim of this project, one favorable side effect of this method is that it allows for the filtering out of parameters resulting in very bad runtime as well as the elements supplying poor accuracy.

4.4. Cutting in the Evaluation Algorithm

In addition to region growing, the algorithm evaluating the results has a long runtime in certain cases. In the case that too many elements have to be processed (the number of cell nuclei in the original reference slide or found by region growing are too high) and they have the less than a one-to-one overlap, then more trials will be needed by the backtrack search of the evaluation algorithm. Taking into consideration the above aspects, we use an acceleration technique here, too.

The backtrack algorithm described in our previous article [6] was combined with an extra cutting operation: before the search starts for the ideal pairing, it calculates how many combinations exist (although it will not know how many of them will be actually examined by backtrack) and if the number is more than 1,000,000 then it searches the points where evaluation can be simplified without significantly affecting the final result.

1. In practice this means that it investigates the partial tasks of backtrack individually and finds the solution that brings the greatest fitness value in the case of the individual tasks.
2. Based on these local maximum values, it selects the global maximum of reference and the test cell nucleus pairings that show the greatest overlap irrespective of the other pairing possibilities.
3. This paring will be recorded (just as will be done by the backtrack algorithm at the end of the searching).
4. Because the backtrack algorithm does not have anything to do in this partial task, it clears the full partial task. In the case that the partial task would have required examination of K pieces of paring possibilities, then the full size of the problem space to be overlapped by the backtrack algorithm would reduce to the K-th part.
5. Should the number of the possible combinations still be too high, it restarts this operation, beginning from Step 1.

This method significantly reduced waiting time (since searching time exponentially increases with the number of the possible pairings, and therefore some unfortunately selected pairings would result in drastic waiting time for the whole system). Of course the disadvantage is that the accuracy of evaluation reduces, but thanks to the above method, there is no random element in simplification, so the evaluation process can be repeated at any time and it will give the same result.

5 Assessment of Our Results

5.1. Examination of the Best Results for Every Generation

Due to the used elitism technique, the instances with the best fitness values are automatically carried along into the following generation as well. For this reason, the best fitness values give us a monotonically increasing series of numbers. By examining these values, we can make assumptions about the speed of our optimization.

As a comparison, we can use the parameter set used in [4]. After executing the analysis of the representative 11 tissue samples, the usual evaluation gave us the average accuracy of 78.1%. Currently this can be considered as the basis result and this is compared later with our result. Surpassing this value by any extent can clearly be considered as an improvement.

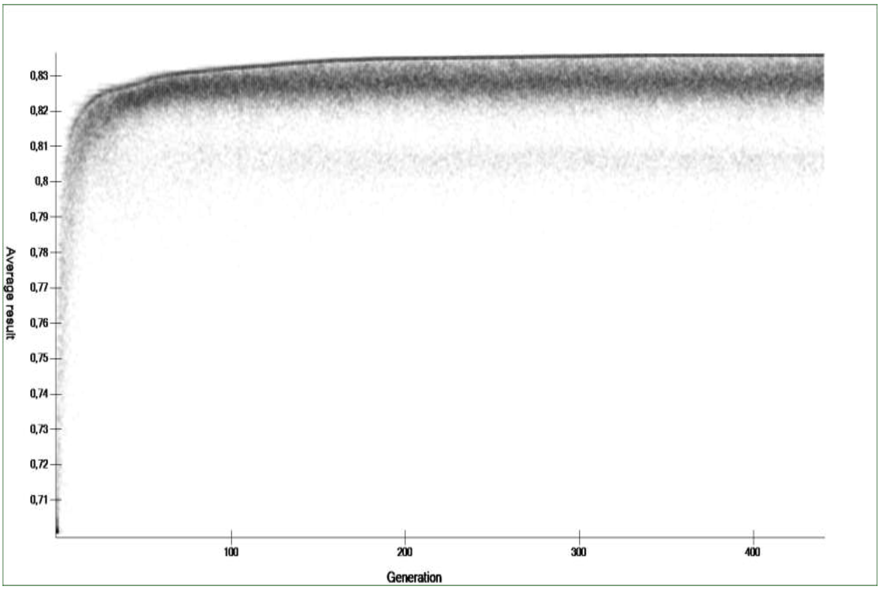


Diagram 2
Distribution of accuracy by generations

Diagram 2 shows the distribution of the different instances within the various generations. For every generation (horizontal axis) we display the accuracy levels (vertical axis) for every instance, using grey dots. Obviously the dark areas represent many chromosomes with similar results.

It is visible that in the first few generations the instances are usually well below the aforementioned 78.1% value, but after some generations the majority of the population had chromosomes that yielded a better result than this value. By generation #273, the best accuracy was reached at 83.6%; this means an improvement of 5.5% compared to the previous best result. We kept the algorithm running for even more time, but until we stopped it (at generation #440) it produced no generation that had better accuracy than this.

5.2. The Convergence of the Generations

The first few populations (until generation #4 - #5) mostly contained chromosomes with weaker results, which did not even reach the previously known best value (though there were always a few that surpassed this level). This number then quickly increased during consecutive generations. For the first few generations there is a darker grey “patch” around the 70% level; this is because there were chromosomes that were technically viable, but the region growing found no cell nuclei (these are the instances that yielded empty results); and those chromosomes got an average accuracy value of 70% (maybe it is strange, but it is not a mistake; most pixels of the tissue images are not parts of any cell nuclei, and therefore a blank result is decidedly better than a result with a lot of false positive hits). The first few generations contained a lot of chromosomes like this, and this is why the darker grey patch appears.

The weak results of first few populations were followed by a dynamic growth phase (until generation #100) where the genes that belonged to the instances with good fitness values start to spread, and so the results improve greatly. It is visible that it took a reasonably small amount of time until almost all instances surpassed the critical 78.1% level. This part shows a process that is expected from a genetic algorithm: the chromosomes that started from a broad spectrum gradually converge towards each other, and since they converge mostly to the chromosomes that yield the best results, this means that their average result is increased as well.

This was followed by the last stage, where this dynamic improvement significantly slows down. For a while the program still found better results, partly because of the crossovers, but later most probably it was only because of the mutations. No improvement was made after generation #273, although it can be noted that even after this point there are many different chromosomes present, and they cover a wide range of the search space; but after this point the search process can be more considered as a simple random search rather than a genetic algorithm.

An interesting pattern appeared in the distribution of the results, which can be described as the following:

- The algorithm uses the elitism method, so the top 30 instances of every generation are automatically carried into the next generation. As is visible, after generation #200 the instances are so close to each other that on the figure they are displayed as a dark horizontal line (where there are a lot of very similar instances in one place). This does not mean that the top 30 elements are totally identical (there are smaller differences in the various genes), but the evaluation yields the same result for all of them.
- Below this line there is a wide light gray strip that contains the newly created instances that (after the crossovers and mutations) yield variously different results. This wide strip is virtually the same during the consecutive generations, and this is true despite the fact that we would expect the crossovers to converge towards some kind of optimum; on the other hand, the mutations are totally random and are more dominant than the crossovers.
- It is worth mentioning the light pale strip below the gray strip, which is the result of a single parameter having a special value that degrades the accuracy of the given chromosome by about 2%. The light strip contains only the instances with this special value (this is clearly shown because if we do not display elements with that value, then this light strip disappears).

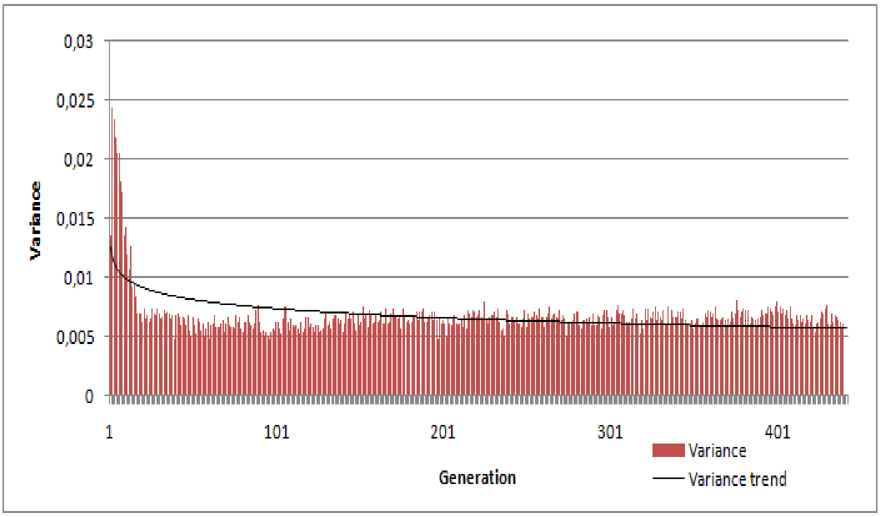


Diagram 3

Distribution of accuracy by generations

We also need to examine our results considering the aspect of how long the algorithm should run. As we stated above, we do not have any expected final results at which point the processing should stop; but there could still be some conditions where the search should be terminated. One such condition could be if the consecutive generations contain instances that have genes that are so close to each other that the crossovers make no substantial changes.

For a better view, it is practical to examine the variance of the accuracy values for the various instances (Diagram 3). As is expected, the variance is pretty high in the beginning for the randomly generated instances and in the first few generations, but as the more viable instances spread, this is quickly decreased.

Then, after a while, this drop of variance stops and the variance remains stable. Although we would expect that with a genetic algorithm the variance would keep dropping (because the elements get more and more closer to an optimal result and to each other), this did not happen in our case. The reason behind this is that we set the size and the probability of the mutations to a pretty high value, because we have a reasonably large search space, and we tried to make the search process faster using a quite high ratio of mutations. But the many mutations prevent the genes from settling at an ideal value, so the variance cannot decrease any more. And even though the variance is high, our results show that the search gradually continued in a good direction, and the elitism method made sure that it did not change course.

It is however clearly visible that we cannot set a stopping condition based on the distances of the chromosome from each other, because due to the mutations we cannot set such a limit. For this reason, we stopped the search when no improvement was shown for 200 generations. Of course, because of the mutations we cannot say that there will be no improvement from the achieved maximum accuracy, but it is probably more practical to use our resources for a new search that is based on our final results used as an initial population.

Diagram 3 shows a curiosity: the variation increases only once, when stepping from the first to the second generation. After examination of every instance it became clear that this happens because of the many non-viable instances in the first generation; and because of the chromosomes that give empty results after the region growing algorithm. The many similar instances (even though their results cannot be considered as usable results) gave us the lower variance value.

5.3. Examination of the Processing Speed

Examination of processing speed is not one of the primary goals of our research (and we could access the required resources for a limited amount of time only, so our aim was to get the best possible performance and not to measure the different performance levels). We performed no separate speed tests, but by subsequently analyzing the log files of the clients we were able to draw some conclusions on this topic as well. During the regular operation of the system, various clients dynamically connected and disconnected, so in this way we could obtain data about how the number of clients affects the processing time required for one generation.

This is shown in Diagram 4, where the horizontal axis represents the generations, the blue area in the background represents the total processing time required for a generation, and the red dots represent the number of clients that took part in the processing of that generation.

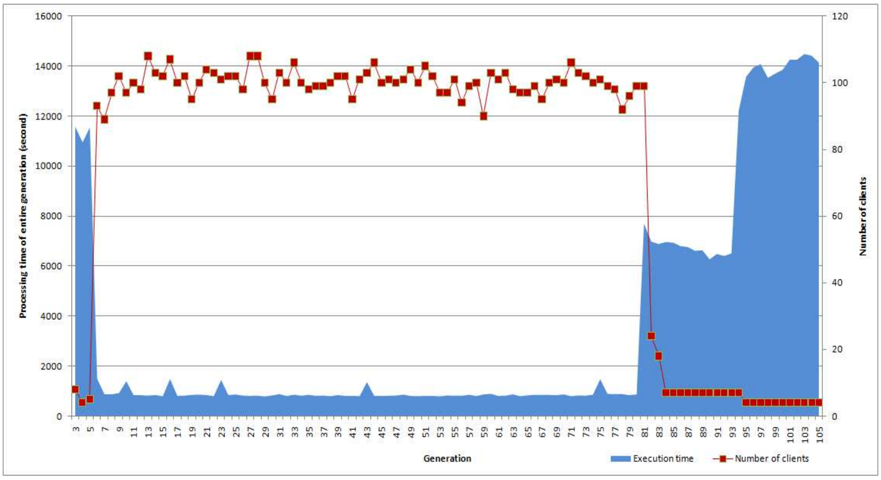


Diagram 4
Execution time and number of clients by generations

It is clearly visible that by increasing the number of clients the required processing time immediately decreased, while during the last generations (as some clients were disconnected) the required time started to increase again. As the individual task units are almost totally independent from each other, and as the processing units are totally independent from each other, we would expect that by increasing the number of clients, the execution time required for a generation would linearly decrease.

This did not happen because we could technically use only 27 workstations, so the only way that we could use more than 100 clients was that the client program was executed in multiple instances; in this way we tried to utilize the full potential of those computers (the computers had quad-core CPUs, an thus according to this, we used 4 instances per workstation).

This slightly increased the processing performance, because the evaluation algorithm only utilizes a single CPU core, so by executing four instances, we did manage to better utilize the available resources during the evaluation phase. As for the cell nuclei detection algorithm this made no changes, because the region growing algorithm was already optimized for a multi-core environment (especially through the intensively used OpenCV function calls), so using multiple instances meant absolutely no benefits in this case.

The linear growth as such cannot be reached in itself for the single reason that the processing time required for the different task units can vary a lot, so there is always a few-minute-long hiatus when one of the clients is still working on the last task unit while all the others are already finished and thus are waiting for the start of the next generation (to execute the genetic operators, the complete results of every task units are required). For this reason, a further improvement could be to estimate the processing time required for the task units (based on the parameters), and in this way it could be possible to implement a more sophisticated scheduler.

Conclusions

As a result of our research, we managed to develop a data-parallel region growing algorithm that is equally as accurate as the sequential version, but its speed is two or three times faster than the original one. The algorithm was implemented on Nvidia Fermi GPU.

The next step was to set the values of the considerably high number of parameters that are required for both the old sequential and the new parallelized versions of the region growing algorithm. We have developed a genetic algorithm for this purpose and implement the framework being in compliance with the above. In addition to the basic classes, a component implementing a communication protocol (applying the already mentioned FTP transmission) has been developed.

As mentioned above, we ran the evaluation for 11 only tissue samples, instead of the available 41 samples, and therefore we could use the remaining samples as a control group. According to this, we executed the evaluation for all of these images: the average accuracy was 76.83% (using the old parameters) and now was 81.15% (using the parameters found by the genetic algorithm).

The genetic algorithm successfully determines a set of parameters that can be used to achieve 81.15% accuracy on the pre-existing reference slides. This, compared to the results with the previously known best set of parameters, means an improvement of 4.32%.

We have already developed a distributed framework for the execution of the genetic algorithm. The framework has lived up to our expectations, and the execution time of 440 generations was fully acceptable.

At present the server gives the tasks for the clients in batches; one batch actually means the complete examination of one chromosome (the running of region growing and the evaluation for the selected images and summary of the results). However, it became clear in practice that further distribution of the batches would be practical (that is, in the case that the evaluation has to be done for several images, then it could be done by different clients); namely it is fairly frequent that several hundred clients are waiting for the next generation while one client is still working with the large-volume job assigned to it.

Due to technical reasons the clients and the server communicate through a FTP server in compliance with the rules of the FTP protocol. This solution was selected because we worked with rather limited instruments and this protocol allowed us simple and problem-free coordination of the remote devices behind different firewalls. However, it became evident in practice that in the case of larger loading (hundreds of clients at a time, concurrent access to the same files) the examined FTP server implementations proved to be bottlenecks, and therefore communication should be practically re-planned with a protocol especially developed for this purpose.

References

[1] K. Palágyi, J. K. Udupa, “Medical Image Registration Based on Fuzzy Objects”, Summer Workshop on Computational Modeling Imaging and Visualization in Biosciences, Sopron, 29-31 Aug. 1996

[2] A. Reményi, S. Szénási, I. Bándi, Z. Vámossy, G. Valcz, P. Bogdanov, Sz. Sergyán, M. Kozlovszky, “Parallel Biomedical Image Processing with GPGPUs in Cancer Research”, 3rd IEEE International Symposium on Logistics and Industrial Informatics (LINDI), 2011, Budapest, 25-27 Aug. 2011, ISBN: 9781457718403, pp. 245-248

[3] S. Szenasi, Z. Vamossy, M. Kozlovszky, “GPGPU-based Data Parallel Region Growing Algorithm for Cell Nuclei Detection“, 2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), 21-22 Nov. 2011, pp. 493-499

[4] Pannon University, “Algoritmus- és forráskódleírás a 3DHistech Kft. számára készített sejtmag-szegmentáló eljáráshoz”, 2009

[5] S. Szenasi, Z. Vamossy, M. Kozlovszky, "Preparing Initial Population of Genetic Algorithm for Region Growing Parameter Optimization", 4th IEEE International Symposium on Logistics and Industrial Informatics (LINDI), 2012, 5-7 Sept. 2012, pp. 47-54

[6] S. Szenasi, Z. Vamossy, M. Kozlovszky, “Evaluation and Comparison of Cell Nuclei Detection Algorithms“, IEEE 16th International Conference on Intelligent Engineering Systems (INES), 13-15 June 2012, pp. 469-475

[7] R. E. Precup, S. Preitl S, J. K. Tar, M. L. Tomescu, M. Takács, P. Korondi, P. Baranyi, “Fuzzy Control System Performance Enhancement by Iterative Learning Control”, IEEE Transactions on Industrial Electronics, 2008, pp. 3461-3475

[8] R. Kohavi, F. Provost, “Glossary of Terms”, Machine Learning, Vol. 30, No. 2, Springer Netherlands, 1998, pp. 271-274., ISSN: 08856125

[9] M. T. Goodrich, R. Tamassia, “Algorithm Design: Foundations, Analysis, and Internet Examples Algorithm Design”, John Wiley & Sons Inc., 2002, ISBN: 0471383651

[10] M. Mitchell, “An Introduction to Genetic Algorithms”, Bradfork Book The MIT Press, Cambridge, 1999, ISBN: 0262133164

[11] D. Gupta, S. Ghafir, “An Overview of Methods Maintaining Diversity in Genetic Algorithms”, International Journal of Emerging Technology and Advanced Engineering, Vol. 2, No. 5, May 2012, ISSN: 22502459, pp. 56-60

[12] R. Murphy, “A Generic Parallel Genetic Algorithm”, M.Sc. Thesis in High Performance Computing, University of Dublin, Department of Mathematics

[13] E. Alba, M. Tomassini, “Parallelism and Evolutionary Algorithms”, IEEE Transactions on Evolutionary Computation, Vol. 6, No. 5, Oct. 2002, pp. 443-462

[14] E. Alba, J. M. Troya, “A Survey of Parallel Distributed Genetic Algorithms”, Complexity, Vol. 4, No. 4, John Wiley & Sons Inc., March/April 1999, pp. 31-52

[15] M. Nowostawski, R. Poli, “Parallel Genetic Algorithm Taxonomy”, Third International Conference Knowledge-based Intelligent Information Engineering Systems, Dec. 1999, pp. 88-92

[16] P. Adamidis, “Parallel Evolutionary Algorithms: A Review”, 4th Hellenic-European Conference on Computer Mathematics and its Applications, 1998

[17] A. Muhammad, A. Bargiela, G. King, "Fine-Grained Parallel Genetic Algorithm: A Stochastic Optimisation Method", In Proceedings of The First World Congress on Systems Simulation, 1997, pp. 199-203

[18] S. Baluja, "Structure and Performance of Fine-Grain Parallelism in Genetic Search", Proceedings of the 5th International Conference on Genetic Algorithms", Morgan Kaufmann, 1993, pp. 155-162

[19] J. D. Lohn, S. P. Colombano, G. L. Haith, D. Stassinopoulos, "A Parallel Genetic Algorithm for Automated Electronic Circuit Design", NASA Ames Research Center, 2000

[20] M. Golub, L. Budin, An Asynchronous Model of Global Parallel Genetic Algorithms, Second ICSC Symposium on Engineering of Intelligent Systems EIS2000, University of Paisley, Scotland, UK, 2000, pp. 353-359

[21] M. Dubreuil, C. Gagne, M. Parizeau, "Analysis of a Master-Slave Architecture for Distributed Evolutionary Computations", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 36, No. 1, Feb. 2006, pp. 229-235

[22] M. Babbar, B. S. Minsker, "A Multiscale Master-Slave Parallel Genetic Algorithm with Application to Groundwater Remediation Design", Late Breaking papers at the Genetic and Evolutionary Computation Conference (GECCO-2002), New York, USA, 9-13 July 2002, pp. 9-16

[23] U. Kohlmorgen, H. Schmeck, K. Haase, "Experiences with Fine-grained Parallel Genetic Algorithms", Annals of Operations Research, 1996, pp. 203-219

[24] I. Rudas, "Hybrid Systems: Integration of Neural Networks, Fuzzy Logic, Expert Systems, and Genetic Algorithms", Encyclopedia of Information Systems, Boston: Academic Press, 2002, pp. 114-1-114-8

[25] I. Foster, Z. Yong, I. Raicu, L. Shiyong, "Cloud Computing and Grid Computing 360-Degree Compared", Grid Computing Environments Workshop (GCE '08), 12-16 Nov. 2008, pp. 1-10

[26] I. Foster, C. Kesselman, S. Tuecke, „The Anatomy of the Grid: Enabling Scalable Virtual Organization", The Intl. Jrnl. of High Performance Computing Applications", Vol. 15, No. 3, 2001, pp. 200-222

[27] M. Sarnovský, P. B., J. Paralič, "Grid-based Support for Different Text Mining Tasks", Acta Polytechnica Hungarica, Vol. 6, No. 4, 2009, pp. 5-27

[28] D. Thain, T. Tannenbaum, M. Livny, "Distributed Computing in Practice: The Condor Experience", Concurrency and Computation: Practice and Experience", Vol. 17, No. 2-4, Feb-Apr. 2005, pp. 323-356

[29] Network Working Group, "FILE TRANSFER PROTOCOL (FTP)", IETF RFC 959

Measuring the Reusable Quality for XML Schema Documents

Tinzar Thaw

University of Computer Studies, Mandalay, P.O. Box: 73 22, Vientiane Lao PDR
Patheingyi Township, Mandalay, Myanmar; e-mail: tinzar.t@gmail.com

Sanjay Misra

Department of Computer Engineering, Atilim University, Kizilcasar Mahallesi,
06836 Incek Golbasi, Ankara, Turkey; e-mail: smisra@atilim.edu.tr

Abstract: eXtensible Markup Language (XML) based web applications are widely used for data describing and providing internet services. The design of XML schema document (XSD) needs to be quantified with software with the reusable nature of XSD. This nature of documents helps software developers to produce software at a lower software development cost. This paper proposes a metric Entropy Measure of Complexity (EMC), which is intended to measure the reusable quality of XML schema documents. A higher EMC value tends to more reusable quality, and as well, a higher EMC value implies that this schema document contains inheritance feature, elements and attributes. For empirical validation, the metric is applied on 70 WSDL schema files. A comparison with similar measures is also performed. The proposed EMC metric is also validated practically and theoretically. Empirical, theoretical and practical validation and a comparative study proves that the EMC metric is a valid metric and capable of measuring the reusable quality of XSD.

Keywords: XML; XSD; WSDL; Software Metrics; Entropy

1 Introduction

The Web Services Description Language (WSDL) is an XML format for describing the functions of web services and network services and defining interfaces between these services and web based applications. A web service is a software system designed to support interoperable machine-to-machine interaction over a network. Within the web services development environments, developers use WSDL language to facilitate web services without understanding the details of network protocols. Any special data types used are embedded in the WSDL file in the form of XML Schema [1]. In the software development process, when considering a Web service design, XML Schema components should be carefully designed for easy reuse for the purpose of software maintainability, the usage of memory and controlling development cost. The inheritance feature of software has a significant impact on software reusable quality.

In object-oriented programming (OOP), for the XML schemas, inheritance is a way to represent in modules (compartmentalization) and reuse schema components by creating collections of structural schema components [2]. A class, a schema type as a collection of elements and attributes, not only inherits elements or attributes from parent elements, but also validates the contents of these components. This means less programming is required when adding functions to complex web applications. The ability to reuse the existing component collections is a major advantage of object-oriented technology [3]. In the World Wide Web Consortium (w3c) standard schema, using extending or restricting keywords in the simple or complex type definitions can provide inheritance features that elements and attributes inherited from parent elements [4].

Reusable quality is important to reduce software development cost. Many metrics help developers and development groups to assess software quality during the software development process. Although not too much effort has been made to develop XML schema quality metrics, entropy-based metrics have been developed for measuring the maintainability and complexity of XML schema documents. Entropy, in information theory, is used to measure the uncertainty associated with a random variable [5]. In the context of an XML schema document, it is difficult to determine that how many inheritance feature components affect the degree of the reusable quality of the XML schema document; the Entropy method is suitable and useful for measuring the complexity.

By considering all the above issues, the Entropy Measure of Complexity (EMC) was proposed and presented at a conference [6]. One of the authors of the present paper proposed a different metric for reusable and extensible quality [7] for XML Schema Documents. The authors have proposed a formula for estimating target quality of XML schema by utilising the extendible quality (EQ) and reusable quality (RQ).The present work is an extension of the entropy measure of Complexity [6]. This metric is based on entropy concept and measures how components of XML schema documents inherit to other schema components. We have extended the conference work and validated EMC through different perspectives which include empirical validation, practical and theoretical evaluation, and a comparison with a similar metric. A rigorous empirical validation is performed by applying EMC on 70 WSDL real files available on the web. A comparison is also performed by applying the metrics on the same 70 WSDL files considered for empirical validation.

The structure of the paper is as follows. The next section represents the related works and metrics applicable for XML schemas. The definition of the EMC metric is summarised in Section 3. The validation of EMC is performed in Section 4. Finally, the conclusion drawn on the work is in Section 5.

2 Related Works

The eXtensible Markup Language (XML) based web applications use XML standard schemas to display information and provide network services. Developing efficient XML web applications requires having good quality XML schema documents. Much research has been done to improve quality in different areas of the software development process and to explore the best practices for knowledge capturing and network services. In addition, many metrics have been proposed to measure the quality of software, but unfortunately, the majority of them are not adopted in industry because of improper empirical validation [8]. Although XML based web applications are important, metrics for XML schema document are scare and there has been very little research to create quality metrics for XML schema documents and thereby improve the web engineering process. Therefore, a mature process can produce high quality schema documents.

McDowell et al. [9] proposed the XML schema analyzer tool to measure two composite indices: the quality and complexity of XML schema documents. This tool was created based on the complexity metrics proposed by Klettke et al. [10]. To ensure the quality of the tool, the ISO 9126 quality model was focused on when developing the tool. Moreover, the tool was an open source tool, to which on could easily add new metrics and change their composite indices according to the requirements of a given application. They concluded that this tool was more important for the XML schema documents than working internal data format for applications. Their future work was the validation of the XML schema analyzer tool.

A schema metric was proposed by Basci and Misra [11] to measure the structure design complexity of the XML schema documents. Their metric was based on the internal structure design components of their schema documents. If their metric value increases, the complexity of the given schema document increases. On the other hand, if the complexity value increases, the quality of the given schema document decreases because of inefficient use of memory and time. They validated their complexity metric theoretically and empirically. To prove the usefulness, they applied well-known structure metrics to XML schema documents and their proposed metric compared with these applied structure metrics.

Basci and Misra [12] have proposed another complexity metric to measure the structural design complexity of the XML schema documents [12]. This metric was developed based on the Shannon entropy function [5], which was suitable for measuring XML schema documents due to having complex structural design of schema components. Their metric provided valuable information for software developers and development groups about the reliability and maintainability of XML schema design. Their proposed metric was analyzed with many examples and empirically validated with test cases [12]. Moreover, to prove the usefulness of their metric, the validation framework and the formal set of nine Weyuker properties were used to evaluate their entropy metric theoretically. The same group of authors has also developed metrics for DTD [13] and Web-services [14]. The authors have proposed to evaluate the structural complexity of the DTD [13] through entropy and estimated the complexity due to repetition of similar structures in schema. A suite of metrics [14] for XML-Web-services maintainability includes five metrics. These metrics evaluate different features of the XML Web-services.

Luo and Shinavier, [15] have proposed a metric to measure schema reuse according to the actor-concept-instance model. Their metric was formulated to calculate the entropy value of simple relationships among actors, concept and instance. In this model, a concept was any one to annotate or describe various data. An actor annotated an instance with a concept. For instance, all user-defined types and build-in types were concepts in a XML schema document and student, teacher and staff types were concepts in the education domain. For example, Rose was an instance of the student type.

The authors [15] used entropy to measure the uncertainty of concepts; the formula would be:

$$H(X) = -\sum_i^n p(c_i)\log p(c_i) \tag{1}$$

where $p(c_i)$ =Pr(X= c_i) = $\frac{|A_{c_i}|}{|A|}$, where $\left|A_{c_i}\right|$ is the total number of annotations using concept c_i, $|A|$ is the total number of annotations and i is the total number of concepts. If the metric value was small, the degree of schema reuse was high. This mean that increasing the metric value tends to decrease the degree of schema reuse. Their metric was evaluated against well-known data sets from the well-known web sites. Their research provided knowledge for users about the usefulness of these data sets to create and reuse popular domains.

3 Entropy Measure of Complexity (EMC) Metric

To formulate the EMC metric, a directed graph is exploited to demonstrate the inheritance structure of XML schema components. Section 3.1 explains how to demonstrate the components of a given schema document into a directed graph representation. The EMC metric is defined and demonstrated with three schema examples in Section 3.2.

3.1 Graph Representation of XML Schema Document

The elements, attributes and types, which are the components of the XML schema, can be inherited from their parent components in the sense of inheritance features that are supported by using restriction or extension keyword implemented within the type definition. The directed graph representation can provide the ability to grasp the complex structure components of the XML schema documents with higher frequencies of occurrences [12]. Before calculating the proposed metric, the inheritance features elements and attributes are counted on the directed graph of a given schema document. For instance, graphs of three schema documents are shown in Figures 1, 2 and 3. In the figures, the root node is a schema element and other circles show either schema elements or attributes. Each Node represents two parts: name and their type with either inheritance features or simple and their name represents in the brackets. The notation of simple type, complex type, simple type with restriction feature, complex type with restriction feature and complex type with extension are ST, CT, STr, CTr and CTe, respectively. Figure 1 contains one element with name CNode and typ: complex type by restriction. It has 4 children: 3 simple type attributes and 1 simple type attribute by restriction.

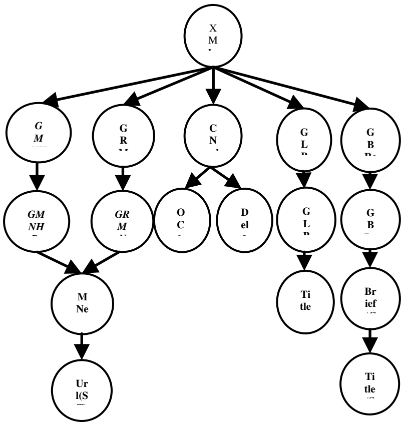


Figure 1
The inheritance feature representation of schema1.xsd

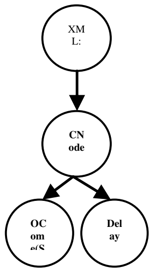


Figure 2
The inheritance feature representation schema2.xsd

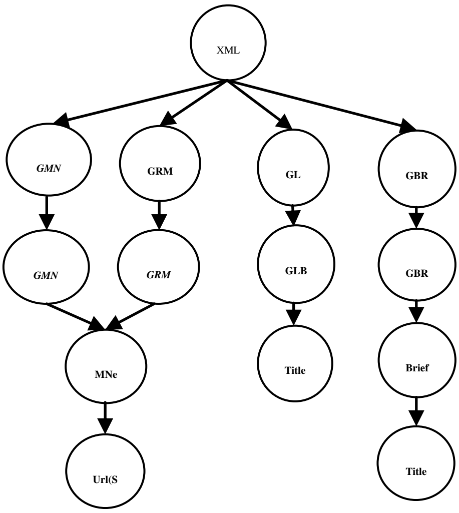


Figure 3

The inheritance feature representation schema3.xsd

3.2 The Metric

The Entropy Measure of Complexity (EMC) measures how XML schema components inherit to other schema components. Increasing the EMC values leads to increasing the reusable quality of XML schema documents. On the other hand, greater EMC values means that the given schema documents have many inheritance feature elements and attributes. The EMC metric is based on the entropy function and the used eight inheritance related metrics. The based eight inheritances related metrics are defined and counted over graph representation of a given XML schema document. These based metrics are shown in Table 1. The metric names are given in the first column and their notations are in the second column of Table.

Table 1

The inheritance feature related metrics over the graph representation of XSD

Metric Name	**Notation**
Total simple type nodes with restriction	$STNode_R$
Total complex type nodes with restriction	$CTNode_R$
Total complex type nodes with extension	$CTNode_E$
Total simple type nodes without restriction	$STNode_{WR}$
Total complex type nodes without restriction	$CTNode_{WR}$
Total complex type nodes without extension	$CTNode_{WE}$
Total complex type nodes	CTNodes
Total simple type nodes	STNodes

In object-oriented programming (OOP), inheritance is a way of organizing and structuring reuse functions and components. In XML schema components, to get this inheritance feature and properties, derive by restriction and derive by extension are used to inheritance schema component structure. The knowledge of the reusable quality of xml schema helps software developers to save time and money in the XML based software system development process [4]. In Table 1, the first three metrics are inheritance feature nodes that support the reusability of schema document. The next three metrics without inheritance feature do not support reusable quality. The CT_{Total} and ST_{Total} are used to get the ratios of above six metrics for the whole document.

Table 2
The based metrics' values of three XML schema documents

Notation	Schema1.xml	Schema2.xml	Schema3.xml
CTNodes	12	1	11
STNodes	24	4	20
$STNode_R$	1	1	0
$CTNode_R$	1	1	0
$CTNode_E$	4	0	4
$STNode_{WR}$	23	3	20
$CTNode_{WR}$	11	0	11
$CTNode_{WE}$	8	1	7
EMC	0.208	0.104	0.237

$STNode_R$, $CTNode_R$ and $CTNode_E$ are inheritance variables, and $STNode_{WR}$, $CTNode_{WR}$ and $CTNode_{WE}$ are non-Inheritance variables. A component contains all elements and attributes. Based on the entropy definition [5], given a schema document (SD), the entropy of a given Schema document has *k* distinct variables (V_k) and k is the total number of inheritance type variables. Each variable contains positive and negative concepts.

To measure the EMC metric, each variable is defined as:

$$V_{STNode_R} \leftarrow [STNode_R, STNode_{WR}]$$

$$V_{CTNode_R} \leftarrow [CTNode_R, CTNode_{WR}] \text{ and}$$

$$V_{CTNode_E} \leftarrow [CTNode_E, CTNode_{WE}]$$

The entropy metric is formulated based on their relative inheritance probabilities of inheritance variables $P(V_k)$. If the XML document does not contain Inheritance features, its complexity will be computed based on its number of types. As a result, the values are negative. For this purpose, before calculating the entropy equation, an algorithm is used. This algorithm is defined as:

Check if a given schema document (SD) contains inheritance variables.

IF not,

For each inheritance variable:

1. multiply the negative value with minus one
2. replace the positive value of variable with the multiply result
3. decrease the total number of particular type nodes by one
4. replace its negative value with the total number of particular type nodes

Accordingly, the EMC metric is defined as:

$$EMC(SD) = -\sum_{k\in(STNode_R, CTNode_R, CTNode_E)} P(V_k)\log_2 P(V_k) ....(2)$$

For example, the EMC metric value for the schema document *schema1.xsd* (the listing of schema1.xsd is in Figure 7) is calculated by using Entropy Equation:

$$V_{STNode_R} \leftarrow [1,23]$$

$V_{CTNode_R} \leftarrow [1,11]$ and $V_{CTNode_E} \leftarrow [4,8]$

$$\begin{aligned} EMC(schema1.xsd) &= -\sum_{k\in(STNode_R, CTNode_R, CTNode_E)} P(V_k)\log_2 P(V_k) \\ &= -P(V_{STNode_R})\log_2 P(V_{STNode_{WR}}) \\ &\quad - P(V_{CTNode_R})\log_2 P(V_{CTNode_{WR}}) \\ &\quad - P(V_{CTNode_E})\log_2 P(V_{CTNode_{WE}}) \\ &= -\frac{1}{24}\log_2\frac{23}{24} - \frac{1}{12}\log_2\frac{11}{12} - \frac{4}{12}\log_2\frac{8}{12} \\ &= 0.208007 \end{aligned}$$

As examples, all inheritance feature related metrics are counted on the graph representation of XML schema documents shown in Table 2. Figure 1, Figure 2 and Figure 3 contain 36, 5 and 31 components, respectively. Figure 3 has the highest ratio of inheritance type variables and the total number of components among them. Therefore, the EMC value also produces the greatest value. A greater EMC metric value means that this XML schema document has many inheritance features, elements and attributes and a high degree of reusable quality.

4 Validation of the Proposed Metric

In this section, the usefulness of the proposed metric will be proved by using the validation process. Software developers and development groups should use only validated metrics to assess product and process quality. The EMC metric is validated empirically and evaluated theoretically in Sections 4.1 and 4.2 respectively.

4.1 Empirical Validation

Empirical validation is the process of proving the practical usefulness of a new metric. To prove the utility of the EMC metric, 70 schema documents from the well-known WSDL files are analyzed, and the analyzed results of the new metric are shown in Appendix A. Figure 4 shows the numbers of nodes with simple types by restriction and with complex types by extension for each schema file. These files have not nodes with complex types by restriction. The comparative results between EMC and H metric values for analyzed schema documents with inheritance features are shown in Figure 5.

The EMC metric can better differentiate the schema files in terms of the inheritance type nodes relationships. Moreover, the two metric values are the ratio to total type nodes. The H value defines and measures the information entropy of actor-concept-instance relationships in a given schema document. According to this Figure, the higher the reuse quality, the higher the EMC values. Inheritance type nodes that contain all elements and attributes are directly related EMC values. The highest EMC value contains more inheritance simple and complex type nodes than others. It is clear that the schema reusable and quality will increase since it has more inheritance feature types of attributes and elements.

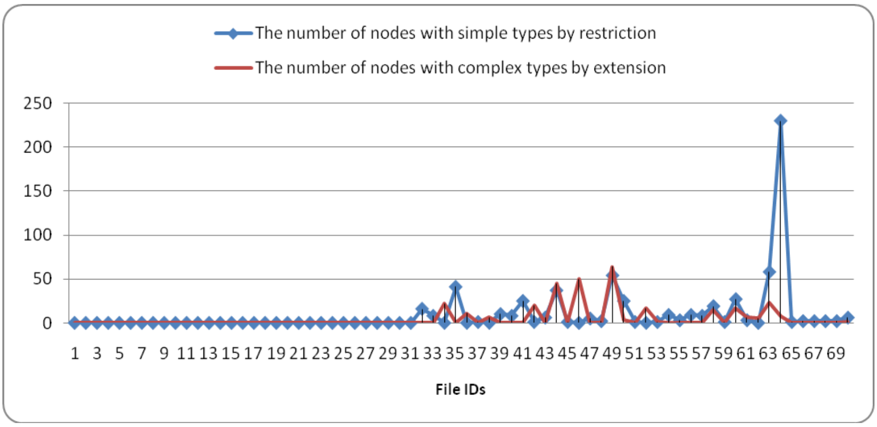


Figure 4
The number of nodes with simple types by restriction and the number of nodes with complex type by extension for each schema file

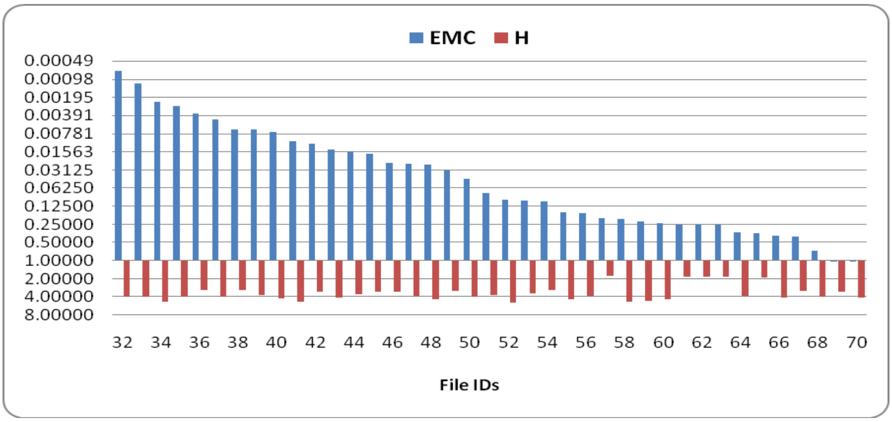


Figure 5
The comparative results between EMC and H values for analyzed WSDL files with inheritance features

Among these 70 schema files, the H metric measures and estimates the schema ID 1 as the most reuse quality at 1.25163 (in Figure 6) and arranges the file IDs 1, 2, 13, 22, 6, 57, 61, 62, 63, 65, 23 and so on in terms of the degree of schema reuse. The file IDs 1, 2, 13, 22, 6, 23 and others greater than ID 31 have not contained any inheritance feature type nodes. Therefore, the H metric does not consider inheritance feature type nodes. The EMC metric measures the 70 schema files and estimates the file ID 70 as having the highest reuse quality. Figure 6 illustrates the comparative results between EMC and H values for analyzed WSDL files without inheritance features. The EMC metric can calculate these files computed based on their simple and complex type nodes in the schema documents.

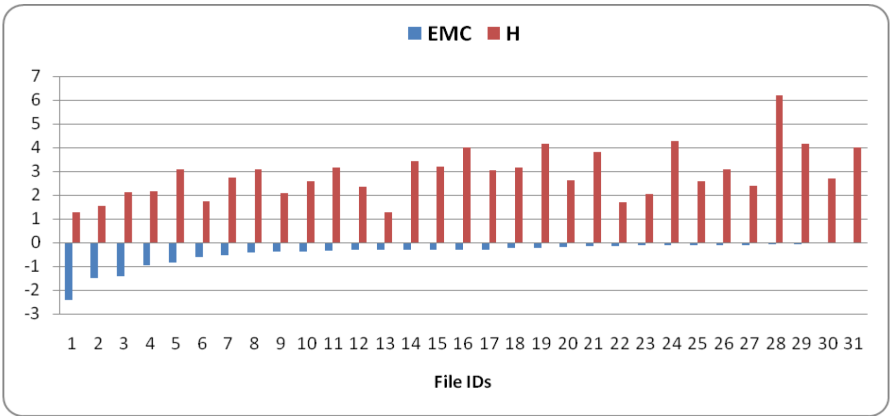


Figure 6
The comparative results between EMC and H values for analyzed WSDL files without inheritance features

4.2 Theoretical Validation of EMC Metric

The usefulness and quality of a new metric is also evaluated by using theoretical validation. In order to perform the validation of the presented metric, the section is organized as follows. EMC is evaluated against Kaner's evaluation framework [16]. Moreover, in section 4.2.2, EMC is also evaluated against the well-known Weyuker's properties [17] through a case study.

4.2.1 Evaluation through a Practical Framework

The practical success of the proposed metric is very important. The metric should be examined formally and practically for its proper validation. When we analyzed the EMC metric according to the practical framework given in [16], EMC is identified as an indirect metric because it depends on many attributes. The EMC is a measure of software reusability and flexibility based on the complexity of schema documents. In the following paragraphs, EMC is evaluated by Kaner's framework.

The purpose of the measure: The purpose of the EMC metric is to help software developers undertake private assessment and to improve their schema based software products.

Scope of usage of the measure: The proposed EMC metric is a reusable quality-measuring tool for software developers and development groups working especially on the XML based applications.

Identified Attribute to measure: The identified attribute measured by EMC is the reusable and flexible quality of the XML schema. A higher complexity value of the schema makes it more reusable and flexible.

Natural scale of the attribute: The natural scale of the attribute is difficult to identify because quality has several definitions, and the reusable quality of XML schema can be measured by several methods.

Definition of metric: The definition of the EMC is given in Section 3.

Measuring instrument to perform the measurement: For inheritance feature metrics of a schema document, the developed oriented model (DOM) parser is used to parse components of this document, and then the system counts these particular components for the particular metric.

Natural scale for the metric: The EMC does not satisfy the additive property so it is not on ratio scale. The exact scale of metric is a task of future work.

Relationship between the attribute to the metric value: The EMC is intend to measure the reusable quality of XML schema, and therefore the metric is directly related to the quality attribute. The experimentations show that an increase in EMC reflects that the schema reusable and flexible quality will increase since it implies having more inheritance feature types of attributes and elements. EMC metric is not a unique indicator of schema reusable and flexible quality.

4.2.2 Evaluation through Weyuker's Properties

In this section, an evaluation of the EMC is also done against Weyuker's properties [17]. Several object oriented metrics are suitable only for the six Weyuker's properties, and other properties are not very useful [18] [19]. The EMC metric is evaluated against 9 properties by using a case study. The evaluations of EMC against the Weyuker's properties are as follows.

***Property 1:* ($\exists$ P) ($\exists$ Q) ($|P| \neq |Q|$) where P and Q are the two different XML schema documents.**

There are different EMC values of all 70 schemas because these different schema documents have different inheritance feature arguments. Hence, the EMC metric satisfies this property.

***Property 2:* Let c be a non-negative number, and then there are only finite numbers of schema documents of complexity c.**

All schema documents consist of only a finite number of inheritance feature based metrics and the EMC metric highly depends on these based metrics. This means that there are only a finite number of XML schema documents of the same complexity if the complexity is a non-negative number. Therefore, EMC satisfies this property.

***Property 3:* There are distinct classes P and Q such that $|P| = |Q|$.**

This property states that there exist many schema documents of the same complexity value. One can find the same EMC values, if different schema documents have the same inheritance feature arguments. Thus, the EMC metric satisfies this property.

***Property 4:* ($\exists$ P) ($\exists$ Q) ($P \equiv Q \,\&\, |P| \neq |Q|$)**

If P and Q are different schema documents having the same functionality, their EMC values can be different because of different implementation. As the EMC metric is based on the internal structure of schema documents, it satisfies this property.

***Property 5:* ($\forall$ P) ($\forall$ Q) ($|P| \leq |P;Q| \,\&\, |Q| \leq |P;Q|$)**

This property states that if the combined schema is constructed from schema P and schema Q, the value of the combined schema document is larger than the value of schema P or schema Q. In Table 3, although Figure 1 is the combination of Figure 2 and Figure 3, the value of Figure 3 is larger than those of the Figure 1. According to this result, the proposed metric cannot satisfy this property.

Property 6: $(\exists P)(\exists Q)(\exists R)(|P| = |Q|) \& (|P;R| \neq |Q;R|)$

This property states that if a new schema document is appended to two schema documents of the same EMC value, the values of the appended documents can be different. For instance, we have two schema documents P, Q and R. These schemas have inheritance feature documents:

P:{ $V_{STNode_R} \leftarrow [2,4]$,

$V_{CTNode_R} \leftarrow [2,14]$,

$V_{CTNode_E} \leftarrow [2,14]$ },

Q:{ $V_{STNode_R} \leftarrow [3,6]$,

$V_{CTNode_R} \leftarrow [1,7]$,

$V_{CTNode_E} \leftarrow [1,7]$ } and

R:{ $V_{STNode_R} \leftarrow [1,13]$,

$V_{CTNode_R} \leftarrow [1,9]$,

$V_{CTNode_E} \leftarrow [2,7]$ }.

P and Q have the same EMC values of 0.243149 and R produces the value of 0.125752. The R schema is then appended to the P and Q schema documents.

(P;Q) :{ $V_{STNode_R} \leftarrow [3,17]$,

$V_{CTNode_R} \leftarrow [3,23]$,

$V_{CTNode_E} \leftarrow [4,22]$ } and

(Q;R):{ $V_{STNode_R} \leftarrow [4,19]$,

$V_{CTNode_R} \leftarrow [2,16]$,

$V_{CTNode_E} \leftarrow [3,15]$ }.

We can observe that the values of the appended P and Q schema documents are different with 0.92657 and 0.110656, respectively. Therefore, this property is also satisfied by the proposed metric.

Property 7: **There are two schema P and Q such that Q is formed by permuting the structure components of P, and** $|P| \neq |Q|$.

The presented metric highly depends on the internal inheritance structure of schema documents and so the EMC metric satisfies this property.

Property 8: **If P is renaming of Q, then** $|P| = |Q|$.

The proposed metric satisfies this property because the value of a given schema is not changed even if the names of the schema and inheritance feature components in this schema are changed.

Property 9: $(\exists P)(\exists Q)(|P| + |Q|) < (|P;Q|)$

According to Table 2, Example1 is the combination schema of P (schema2.xsd) and Q (schema3.xsd). The EMC values of P and Q are 0.104 and 0.207, respectively. The value of the combination schema document is 0.204, and this value is less than the sum of the P and Q values. Therefore, EMC does not satisfy property 9. Further, if two schema are combined, then the complexity of the combined schema will either be less than the sum of the individual ones (due to fact if some modules/elements are in common) or equal (if all modules/elements are different), but in no case will the complexity of the combined schema be less than complexity of the individuals.

In this section, the proposed metric is validated against 9 Weyuker's properties. The EMC metric satisfies 7 properties. It is important to note that it is not necessary to satisfy all the Weyuker's properties [18]. From this point of view, EMC's satisfying seven Weyuker's properties shows that it is a robust measure.

We have followed almost all the steps suggested for the evaluation and validation of software complexity measures [20], except that we have adopted Weyuker's properties in the place of principles of measurement theory for theoretical validation. According to the measurement theory criteria for software complexity measures, a metric should be on ratio scale but is not applicable in majority of object oriented metrics [21]. Our metric is also found not on ratio scale. It is proved via Weyuker's property 9 that EMC is not an additive measure.

Conclusion

The reusable nature of XML schema documents allows developers and software development groups to have the capability of increasing productivity and decreasing development cost of the XML based applications. Increased flexibility and reusability in XML schema documents results in an increased number of inheritance feature elements and attributes in these documents. The EMC metric is developed to achieve these goals. The EMC metric is based on the entropy concept and inheritance feature elements and attributes of XML schema documents. The EMC metric is passed through a rigorous validation process.

EMC is practically evaluated against Kaner's framework. Theoretical evaluation has been done against nine Weyuker's properties. EMC satisfies seven of Weyuker's properties. The practical evaluation and theoretical validation of EMC proves that the metric is developed on scientific principles. The empirical validation is done by applying the metric on 70 real WSDL files. The results and a comparison with the H metric proved the worth and usefulness of the metric. It is found that measuring the reusable quality of XML schema document with the EMC metric will be more useful than via other related metrics. As future work, we aim to explore other factors that are responsible for increasing the complexity of XML schemas. Fixing the threshold values for the EMC metric is also a task of future work.

References

[1] http://www.w3.org/TR/wsdl (Last retrieved on June 2011)

[2] Bieman, J. M., Kang B. K. (1995) Cohesion and Reuse in an Object-Oriented System, In Proceedings of the 1995 Symposium on Software reusability, pp. 259-262

[3] http://java.boot.by/scea5-guide/ch01.html (Last retrieved on June 2011)

[4] http://www.w3c.org/XML/Schema (Last retrieved on June 2011)

[5] Cover T. M., Thomas J. A., Elements of Information Theory, Second Edition (2006)

[6] Thaw T. Z. Khin M. M. (2011) Entropy Measure of XML Schema Document Complexity Metric, Proc. International Conference on Computer Research and Development - ICCRD, pp. 476-479

[7] Thaw T. Z. (2011) Measuring and Evaluation of Reusable Quality and Extensible Quality for XML Schema Documents, In Proc. IEEE Student Conference on Research and Development, pp. 473-478

[8] Misra S. (2011) An Approach for the Empirical Validation of Software Complexity Measures, Acta Polytechnica Hungarica, 8(2), pp. 141-160

[9] McDowell A., Schmidt C., Bun K. Y. (2004) Analysis and Metrics of XML Schema, In Proceedings of Intl Conference on Software Engineering Research and Practice, pp. 538-544

[10] Klettke M., Schneider L., Heuer A. (2002) Metrics for XML Document Collections, In Proc. of XMLDM Workshop, Czech Republic, pp. 162-176

[11] Basci D., Misra S. (2009) Measuring and Evaluating a Design Complexity Metric for XML Schema Documents' Journal of Information Science and Engineering, 25(5) pp.1405-1425

[12] Basci D., Misra S. (2010) Entropy as a Measure of Quality of XML Schema Document, International Arab Journal of Information Technology, 8(1) pp. 16-24

[13] Basci D., Misra S. (2011) Document Type Definition (DTD) Metrics, Romanian Journal of Information Science and Technology, 14(1) pp. 31-50

[14] Basci D., Misra S. (2011) Metrics Suite for Maintainability of XML Web-Services, IET Softwar, 5(3), pp. 320-341

[15] Luo X., Shinavier J. (2009) Entropy-based Metrics for Evaluating Schema Reuse, In Proceedings of the 4th Asian Conference on The Semantic Web (Springer-Verlag Berlin Heidelberg) pp. 321-331

[16] Kaner C. (2004) Software Engineering Metrics: What Do They Measure and How Do We Know? In Proceedings of the 10th International Software Metrics Symposium, 2004, pp. 1-10

[17] Weyuker E. J. (1988) Evaluation Software Complexity Measures, IEEE Transactions on Software Engineering, 14, pp. 1357-1365

[18] Misra S. (2011) Evaluation Criteria for Object-oriented Metric, Acta Polytechnica Hungarica, 8(5) pp. 109-136

[19] Misra S., Akman I. (2008) Applicability of Weyuker's Properties on OO Metrics: Some Misunderstandings, Journal of Computer and Information Sciences, 5(1) pp. 17-24

[20] Misra S., Akman I., Palacios R. C. (2012) A Framework for Evaluation and Validation of Software Complexity Measures, IET Software, 6(4), pp. 323-334

[21] Zuse, H (1996) Foundations of Object-oriented Software Measures. In Proceedings of the 3rd International Symposium on Software Metrics: From Measurement to Empirical Results (METRICS '96) IEEE Computer Society, Washington, DC, USA, pp. 75-84

APPENDIX A

Table 3

The results of EMC metrics for the analyzed schema files. Files are arranged according to the EMC values with ascending.

ID	WEB LINK	EMC	H
1	http://www.webservicex.net/CreditCard.asmx?WSDL	-2.41504	1.25163
2	http://www.thomas-bayer.com/axis2/services/BLZService?wsdl	-1.49185	1.25163
3	http://www.elguille.info/NET/WebServices/HolaMundoWebS.asmx?WSDL	-1.41504	2.12809
4	http://service.ecocoma.com/geo/cityzip.asmx?WSDL	-0.96578	2.16096
5	http://www.yazgelistir.com/YGServices/ArticleService.asmx?wsdl	-0.85982	3.09580
6	http://service.ecocoma.com/geo/distance.asmx?WSDL	-0.63298	1.73136
7	http://www.webservicex.net/*BibleWebservice.asmx*?wsdl	-0.52279	2.73451
8	http://rangiroa.essi.fr:8080/dotnet/evaluation-cours/EvaluationWS.asmx?WSDL	-0.42954	3.06365
9	http://services.nirvanix.com/ws/Authentication.asmx?WSDL	-0.38201	2.07486

10	http://services.argosoft.com/AddressValidation/AddressVerifier.asmx?WSDL	-0.38201	2.56131
11	http://service.ecocoma.com/convert/chinese.asmx?WSDL	-0.34417	3.16608
12	http://www.geoservicios.com/V2.0/sgeo/sgeo.asmx?WSDL	-0.31348	2.35019
13	http://service.ecocoma.com/shipping/fedex.asmx?WSDL	-0.31063	1.26903
14	http://services.test.musiccue.net/rapidcueapplication/WorkManager.asmx?WSDL	-0.30694	3.41259
15	http://ws.cdyne.com/emailverify/Emailvernotestemail.asmx?wsdl	-0.30134	3.20987
16	http://www.mathertel.de/AJAXEngine/S02_AJAXCoreSamples/CalcService.asmx?WSDL	-0.29183	3.99255
17	http://quisque.com/fr/chasses/blasons/search.asmx?WSDL	-0.29170	3.05216
18	http://quiksilver.ws.eto.fr/Connexion.asmx?WSDL	-0.22151	3.14242
19	http://webservice.webxml.com.cn/webservices/ChinaTVprogramWebService.asmx?WSDL	-0.21270	4.16272
20	http://demo.soapam.com/services/FedEpayDirectory/FedEpayDirectoryService?WSDL	-0.17715	2.60746
21	http://www.oorsprong.org/websamples.arendsoog/ArendsoogbooksService.wso?WSDL	-0.16379	3.82050
22	http://trial.serviceobjects.com/pa/phoneappend.asmx?WSDL	-0.15110	1.70601
23	http://ws2.serviceobjects.net/ev/EmailValidate.asmx?WSDL	-0.13412	2.03487
24	http://api.legiomedia.com/Content.asmx?WSDL	-0.13019	4.28320
25	http://secure.adpay.com/affiliate/affiliates.asmx?wsdl	-0.12088	2.57051
26	http://hooch.cis.gsu.edu/bgates/MathStuff/Mathservice.asmx?WSDL	-0.11651	3.08831
27	http://www.sipeaa.it/wset/ServiceET.asmx?WSDL	-0.11507	2.37127
28	http://developer.factiva.com/2.0/wsdl/FDKParsers.wsdl	-0.07789	6.19208
29	http://www.banguat.gob.gt/variables/ws/BDEF.asmx?WSDL	-0.06640	4.13503
30	http://omnovastage.crowechizekasp.com/attributes.asmx?wsdl	-0.04449	2.68170
31	http://ws.eoddata.com/data.asmx?wsdl	-0.00333	4.00160
32	http://www.chemspider.com/MassSpecAPI.asmx?WSDL	0.00072	3.96796
33	http://ws.interfax.net/dfs.asmx?WSDL	0.00117	3.95824
34	http://demo.turtletech.com/latest/webAPI/metering.asmx?WSDL	0.00234	4.77941
35	http://ssl.9squared.com/catalog/catalog.asmx?WSDL	0.00275	3.96266
36	http://www.imagine-r.com/services/WsImagineR.asmx?WSDL	0.00370	3.11048
37	https://api.wildwestdomains.com/wswwdapi/wapi.asmx?wsdl	0.00458	3.95678
38	https://demo.docusign.net/API/3.0/Credential.asmx?WSDL	0.00664	3.08703
39	http://service.thefamousgroup.com/ProjectService.asmx?wsdl	0.00671	3.71067
40	http://www.esendex.com/secure/messenger/soap/ContactService.asmx?WSDL	0.00742	4.28139
41	http://msrmaps.com/TerraService2.asmx?WSDL	0.01034	4.78186
42	http://www.multispeak.org/interface/30j/10_OA_EA.asmx?WSDL	0.01141	3.32276
43	http://terraserver-usa.com/LandmarkService.asmx?WSDL	0.01446	4.11318
44	http://svc.exaphoto.com/eXaPhoto/CollectionServices.asmx?WSDL	0.01601	3.59425
45	http://services.nirvanix.com/ws/Accounting.asmx?WSDL	0.01669	3.28918
46	http://www.phdcc.com/findinsite/SearchService.asmx?wsdl	0.02408	3.22759
47	http://www.partenairedejeu.fr/WebServices/RelationManager.asmx?WSDL	0.02488	3.89019
48	https://www.devcallnow.com/WebService/OneCallNow.asmx?wsdl	0.02583	4.32031
49	https://api.channeladvisor.com/ChannelAdvisorAPI/v5/AdminService.asmx?WSDL	0.03177	3.17985
50	http://114-svc.elong.com/NorthBoundService/V1.1/ NorthBoundAPIService.asmx?WSDL	0.04398	3.91752
51	http://www.hitslink.com/reportws.asmx?WSDL	0.07633	3.71378

52	https://api.channeladvisor.com/ChannelAdvisorAPI/v3/OrderService.asmx?WSDL	0.09683	5.00673
53	http://b3.caspio.com/ws/api.asmx?wsdl	0.09958	3.47567
54	http://gw1.aql.com/soap/sendsmsservice.php?wsdl	0.10376	3.02392
55	http://labs.bandwidth.com/api/public/voip/v1_1/NumberManagementService.asmx?wsdl	0.15874	4.37224
56	http://www.xignite.com/xNews.asmx?WSDL	0.16501	3.89726
57	http://www.webservicex.net/TranslateService.asmx?wsdl	0.19499	1.79248
58	http://www.xignite.com/xCalendar.asmx?WSDL	0.20097	4.83966
59	http://water.sdsc.edu/wateroneflow/EPA/cuahsi_1_0.asmx?WSDL	0.22382	4.71068
60	http://ws.strikeiron.com/MidnightTraderFinancialNews?WSDL	0.24208	4.31742
61	http://www.webservicex.net/ConvertTemperature.asmx?WSDL	0.25428	1.84237
62	http://www.webservicex.net/ConverPower.asmx?WSDL	0.25428	1.84237
63	http://www.webservicex.net/CovertPressure.asmx?WSDL	0.25428	1.84237
64	http://www.xignite.com/xwatchlists.asmx?WSDL	0.33584	3.78410
65	http://www.webservicex.com/CurrencyConvertor.asmx?wsdl	0.34601	1.91830
66	http://www.xignite.com/xQuotes.asmx?WSDL	0.38205	4.14231
67	http://www.xignite.com/xDataSet.asmx?wsdl	0.39984	3.13490
68	http://event.peoplenet.dk/_vti_bin/dspsts.asmx?wsdl	0.69499	4.00041
69	http://www.xignite.com/xNASDAQLastSale.asmx?WSDL	1.00640	3.30242
70	http://www.xignite.com/xMetals.asmx?WSDL	1.02051	4.13493

```
 <?xml version="1.0" encoding="UTF-8" ?>
 <s:schema elementFormDefault="qualified"  targetNamespace="http://www.xignite.com/services/">
 <s:element name="GetBriefings">
    <s:complexType />
</s:element>
 <s:element name="GBResp">
    <s:complexType>
       <s:sequence>
           <s:element minOccurs="0" maxOccurs="1" name="GBResu"
              type="tns:ArrayOfBriefing" />
      </s:sequence>
    </s:complexType>
 </s:element>
 <s:complexType name="ArrayOfBriefing">
     <s:sequence>
       <s:element minOccurs="0" maxOccurs="unbounded" name="Brief" nillable="true"
          type="tns:Briefing" />
     </s:sequence>
 </s:complexType>
 <s:complexType name="Briefing">
    <s:complexContent mixed="false">
    <s:extension base="tns:Common">
    <s:sequence>
      <s:element minOccurs="0" maxOccurs="1" name="Title" type="s:string" />
      <s:element minOccurs="0" maxOccurs="1" name="Time" type="s:string" />
      <s:element minOccurs="0" maxOccurs="1" name="Text" type="s:string" />
      <s:element minOccurs="0" maxOccurs="1" name="Html" type="s:string" />
    </s:sequence>
 </s:extension>
 </s:complexContent>
</s:complexType>
<s:element name="CNode">
<s:complexType name="Common">
<s:restriction base="xsd:anyType">
   <s:sequence>
<s:element minOccurs="1" maxOccurs="1" name="OCome" type="tns:OutcomeTypes" />
  <s:element minOccurs="0" maxOccurs="1" name="Message" type="s:string" />
  <s:element minOccurs="0" maxOccurs="1" name="Identity" type="s:string" />
  <s:element minOccurs="1" maxOccurs="1" name="Delay" type="s:double" />
  </s:sequence>
</s:restriction>
 </s:complexType>
</s:element>
<s:simpleType name="OutcomeTypes">
  <s:restriction base="s:string">
  <s:enumeration value="Success" />
  <s:enumeration value="SystemError" />
  <s:enumeration value="RequestError" />
  <s:enumeration value="RegistrationError" />
    </s:restriction>
 </s:simpleType>
<s:element name="GLBR">
  <s:complexType>
     <s:sequence>
        <s:element minOccurs="0" maxOccurs="1" name="GLBRe"  type="tns:Briefing" />
     </s:sequence>
  </s:complexType>
 </s:element>
<s:element name="GMNHRp">
  <s:complexType>
    <s:sequence>
      <s:element minOccurs="0" maxOccurs="1" name="GMNHRs"
        type="tns:ArrayOfMarketNews" />
    </s:sequence>
  </s:complexType>
 </s:element>
```

```
<s:complexType name="ArrayOfMarketNews">
    <s:sequence>
        <s:element minOccurs="0" maxOccurs="unbounded" name="MNews"  nillable="true"
type="tns:MarketNews"/>
    </s:sequence>
</s:complexType>
<s:complexType name="MarketNews">
   <s:complexContent mixed="false">
   <s:extension base="tns:Common">
      <s:sequence>
         <s:element minOccurs="0" maxOccurs="1" name="Headline" type="s:string"/>
         <s:element minOccurs="0" maxOccurs="1" name="Time" type="s:string"/>
         <s:element minOccurs="0" maxOccurs="1" name="Source" type="s:string"/>
         <s:element minOccurs="0" maxOccurs="1" name="Url" type="s:string"/>
            <s:element minOccurs="0" maxOccurs="1" name="OriginalUrl" type="s:string"/>
            <s:element minOccurs="0" maxOccurs="1" name="Summary" type="s:string"/>
       </s:sequence> </s:extension> </s:complexContent>
  </s:complexType>
<s:element name="GRMNHRp">
   <s:complexType> <s:sequence>
      <s:element minOccurs="0" maxOccurs="1"  name="GRMNHRs” type="tns:ArrayOfMarketNews"/>
      </s:sequence> </s:complexType>
</s:element>
</s:schema>
```

Figure 7

The list of the schema documents schema1.xsd

Navigation of Mobile Robots Using WSN's RSSI Parameter and Potential Field Method

János Simon

Subotica Tech, Department of Informatics
Marka Oreškovića 16, 24000 Subotica, Serbia; e-mail: simon@vts.su.ac.rs

Goran Martinović

Faculty of Electrical Engineering, J. J. Strossmayer University of Osijek Kneza Trpimira 2b, 31000 Osijek, Croatia; e-mail: goran.martinovic@etfos.hr

Abstract: In the current work we present an algorithmic proposal for mobile robot navigation using a Wireless Sensor Network (WSN) for the location of a mobile measuring station in a controlled microclimatic environment. Another point of consideration is determining the navigation strategy. Publications in this field of robotics offer a large number of localization methods, mainly focusing on two fields: navigating locally and globally. Navigating locally will determine the mobile robot's position and orientation by implementing a series of sensors. Once we start from an initial position, the robot's position and orientation are updated continuously through the given time frame. Global navigation will ensure that the robot is able to determine its own position and orientation without having previously studied a map or being given some specific information.

Keywords – Localization;Sun SPOT; WSN; Mobile robot; RSSI

1 Introduction

A number of areas in engineering are centered on the problem of localization, so this field has been the focus of research for quite a long time. This is especially true for robotics, but the greatest challenge with indoor localization is posed by how effectively the GPS (Global Positioning System) can be used. This has been the reason for why WSNs (Wireless Sensor Networks) have been used more and more for the localization of a mobile object when indoors. [18] deals with the question of co-operative trilateration. The approach of the authors in that work used the exact range determination for its basis (implementing the ultrasound ranging technique as presented in [17]) for the solution of a least square problem on a large order system. The information needed for completely formulating the square problem requires the transmission from a possibly remote location through a number of hops. The study had a continuation in the form of the examinationof the scalability of the approach, as described in [19]. The authors of [21] were faced with the problem of how to localize in the absence of nodes. The role of the nodes is to create local coordinate systems and then continue to gather them to form a unique network coordinate system. Directional approaches are the topic of [20], which includes steerable directional antennas. [16] offers a presentation of how to estimate unknown node positions in a sensor network based solely on connectivity-induced constraints. The model of the known peer-to-peer communication within the network takes the form of a set of geometric constraints on the node positions. The global solution of a feasibility problem for these constraints results in estimated values for the unknown node positions within the network.This method [1] has several disadvantages: the central point of computation with the associated traffic overhead, scalability and reliability issues. The following section will give a more elaborate description of global navigation and the role it plays in this work.

2 Global Navigation

One of the central issues in numerous areas of engineering is the issue of global navigation when the indoor environment is not familiar. For robotics, exact location and orientation are of primary importance and have been the focus of studies for a number of years [7]. The key to successful indoor localization is whether or not the Global Positioning System can be used effectively. As a way of circumventing this challenge, the GPS is replaced by the Wireless Sensor Networks (WSNs) for helping with locating a mobile object in an indoor environment.

2.1 Estimating Parameters

This subsection will initially focus on the introduction of the model used for parameter estimation of the wireless sensor network. This paper presents the comparison of twoclassic propagation models. One of them is the least squares model, the other the maximum likelihood model, both used for the estimation of the parameters as it is fairly easy to influence the signal through noise and disturbance.The entire parameter estimation system has the following schematic structure:

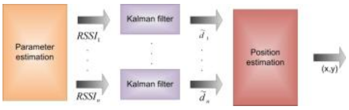


Figure 1
Parameter estimation

As can be seen in this figure above, the extended Kalman filter is implemented, in order to relinearise every estimation once it has been calculated. The results in a new state estimate and the improved better reference state trajectory will subsequently be included in the estimation process, as described in [16].

2.2 Algorithm for Triangulation

A novel version of the Geometric Triangulation algorithm will be the focus of this part of the study, one whichworks withoutnode ordering and covers the whole navigation plane, with the possible exception ofseveral well-determined lines where localization is impossible [23]. Carefully defining the angles used by the algorithm is key to attaining improvements. The angles used are limited only by the common restrictions of all three object triangulation algorithms.

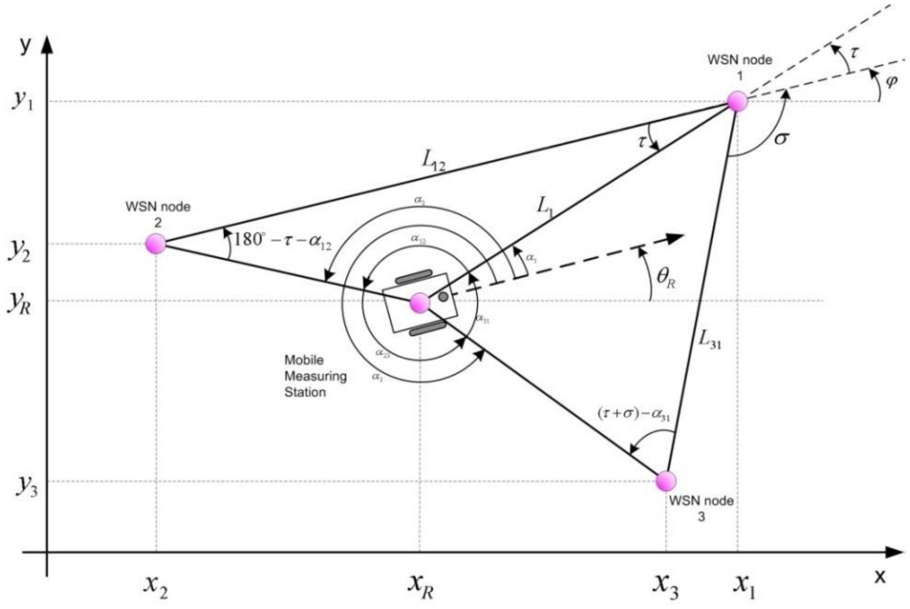


Figure 2

Geometric Triangulation algorithm

Let us take three distinguishable nodes in a Cartesian plane (Fig. 2); we will randomly label them 1, 2 and 3, and give them known positions (x_1, y_1), (x_2, y_2) and (x_3, y_3). L_{12}denotes the distance fromnode 1 to node 2. L_{31}will refer to the distance between nodes 1 and 3. L_1defines how far the robot is fromnode 1. The aim is to determine the robot's position (x_R, y_R) and orientation θ_R. Thus, using the counterclockwise direction, the robot measures the angles α_1, α_2 and α_3, denoting the orientation of the nodes in relation to the robot heading [23].

```
1. If there are less than three visible nodes available, then return a warning message and stop.
2. $\alpha_{12} = \alpha_2 - \alpha_1$
3. If $\alpha_1 > \alpha_2$ then $\alpha_{12} = 360^\circ + (\alpha_2 - \alpha_1)$
4. $\alpha_{31} = \alpha_1 - \alpha_3$
5. If $\alpha_3 > \alpha_1$ then $\alpha_{31} = 360^\circ + (\alpha_1 - \alpha_3)$
6. Compute $L_{12}$ from known positions of WSN node 1 and 2.
7. Compute $L_{31}$ from known positions of WSN node 1 and 3.
8. Let φ be an oriented angle such that -180º < φ ≤ 180º. Its origin side is the image of the positive x semi-axis that results from the translation associated with the vector whose origin is (0, 0) and ends on node 1. The extremity side is part of the straight line defined by nodes 1 and 2 whose origin is node 1 and does not go by node 2.
9. Let σ be an oriented angle such that -180º < σ ≤ 180º. Its origin side is the straight line segment that joins nodes 1 and 3. The extremity side is part of the straight line defined by nodes 1 and 2 whose origin is node 1 and does not go by node 2.
10. $\gamma = \sigma - \alpha_{31}$
11. $\tau = tan^{-1}\left[\frac{sin\alpha_{12}(L_{12}sin\alpha_{31} - L_{31}sin\gamma)}{L_{31}sin\alpha_{12}cos\gamma - L_{12}cos\alpha_{12}sin\alpha_{31}}\right]$
12. If $\left(\begin{array}{c}\alpha_{12} < 180^\circ \\ \tau < 0^\circ\end{array}\right.$ then $\tau = \tau + 180^\circ$
13. If $\left(\begin{array}{c}\alpha_{12} > 180^\circ \\ \tau > 0^\circ\end{array}\right.$ then $\tau = \tau - 180^\circ$
14. If $|sin\alpha_{12}| > |sin\alpha_{31}|$ then $L_1 = \frac{L_{12}\sin(\tau + \alpha_{12})}{sin\alpha_{12}}$
15.     else $L_1 = \frac{L_{31}\sin(\tau + \sigma + \alpha_{31})}{sin\alpha_{31}}$
16.$x_R = x_1 - L_1\cos(\varphi + \tau)$
17. $y_R = y_1 - L_1\sin(\varphi + \tau)$
18. $\theta_R = \varphi + \tau - \alpha_1$
19. If $\theta_R \leq -180^\circ$ then $\theta_R = \theta_R + 360^\circ$
20. If $\theta_R \geq 180^\circ$ then $\theta_R = \theta_R - 360^\circ$
```

Figure 3

Geometric Triangulation algorithm

Fig. 3 presents the algorithmic approach, which is capable of yielding considerable accuracy [23]. In terms of how costly an approach is, the RSS is a rather cheap way of using techniques that measure the power of the received signal strength, which will then transform the measured RSS value into a distance. The RSS method has the added advantage that it does not need any furthermeasuring hardware.

2.3 Measuring and Estimating Parameter Using RSSI

In order to measure the RSSI, a chip with the specification of CC2440 was used. The characteristics include an onboard antenna, and the chip operates within the 2.4 GHz ISM band. This device comprises the node's "heart" used in the current research [11, 12]. The chip supports the IEEE 802.15.4/ZigBee protocol. The relationship between measured RSSI values and distances is shown in the figure below [4, 10].

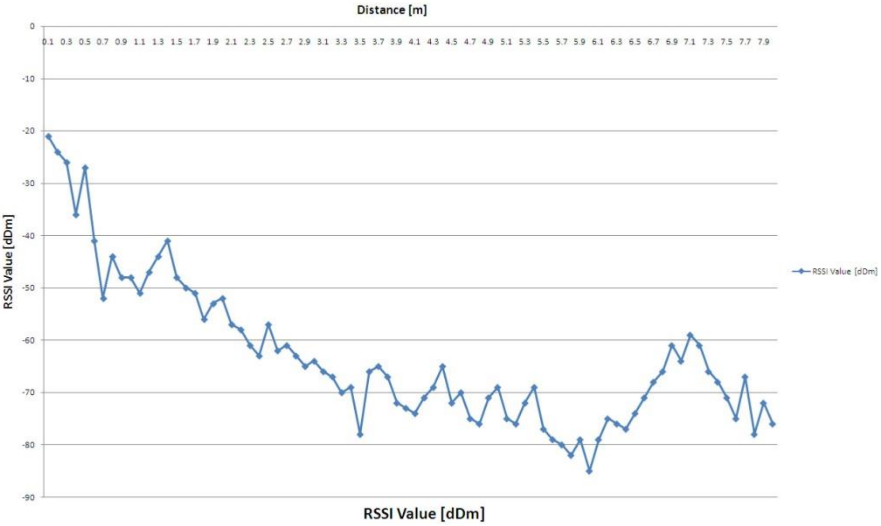


Figure 4
Measured RSSI values

Fig. 4 presents the following: the curve showing the relationship between the RSSI values and the distance is not a smooth one. Due to the multi-path interference, the measured values vary randomly [22, 9]. So the keystone of this model was the rationalization of parameters.

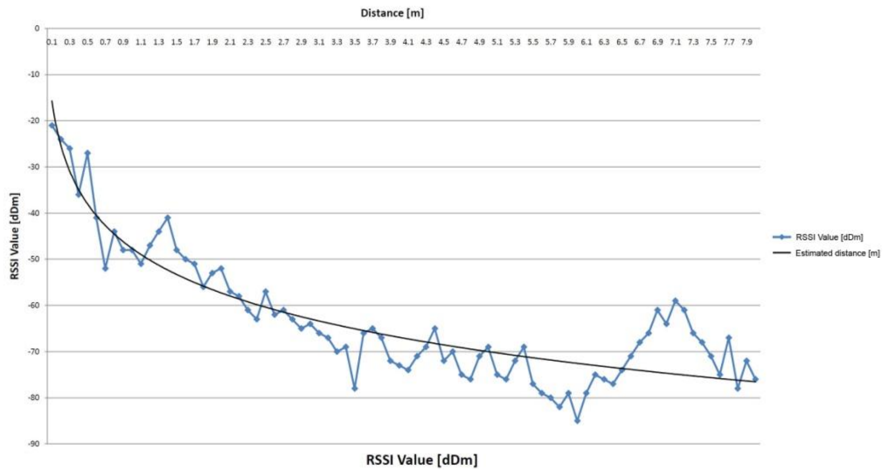


Figure 5

Estimated distance based on measured RSSI

3 Local Navigation

For path planning and collision avoidance, the Potential Fields Approach [15,8] is one of the most often used options, mainly due to how simple and elegant it is in a mathematical sense. It is simple to implement and yields respectable and rapid results for real-time navigation without delay [14]. The objective of any path planning algorithm is to flee from obstacles and move towards a desired goal. The basic idea behind this method is the idea of attractive and repulsive forces. In potential fields, the aim would be defined as a global minimum potential value, with all obstacles seen as high valued potential fields. How the robot moves will be defined by the attained values which appear in its path; in an ideal situation the values will shift from high to low potentials. There are two major formulations, the Local Potential Approachand the Global Potential Approach. In this work we consider only the Local Potential Approach. For this approach, there are several different types of potential field functions, differing from each other by the way the potential is calculated.

3.1 Repulsive Potential Field

The repulsive potential field is the element that keeps the robot away from the obstacles encountered on its path. Every repulsive action vector points away from the obstacle surface, directing the robot to bypass the obstacle.

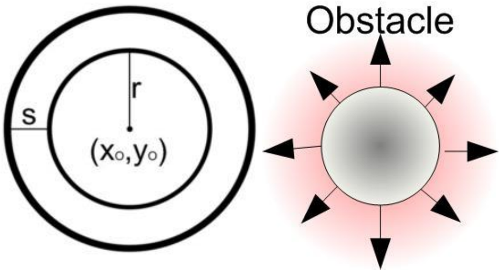


Figure 6
Action vectors of the repulsive potential field [11]

It is possible to calculate the repulsive action vector by applying a scalar potential field function to the robot's position and then calculating the gradient of that function.

$$\nabla = [\nabla U(x, y)] = [\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}] \tag{1}$$

Within the obstacle, the repulsive potential field is infinite and points out from the center of the obstacle. Beyond the circle of influence, the repulsive potential field is zero. The repulsive action vectorcan be calculated.

Having defined [14]

- $[x_O, y_O]$ as the obstacle position;
- r as the obstacle radius;
- $[x_R, y_R]$ as the robot position;
- s as the size of area of influence of the obstacle;
- β as the strength of the repulsive field $(\beta > 0)$

it is possible to compute ∇x and ∇y with the help of the steps given below:

1. Define the distance d between the obstacle and the robot:

$$d = \sqrt{(x_O - x_R)^2 + (y_R - y_O)^2} \tag{2}$$

2. Specify the angle θ between the robot and the obstacle:

$$\theta = \tan^{-1}\left(\frac{y_O - y_R}{x_O - x_R}\right) \tag{3}$$

3. Set ∇x and ∇y based on the rules below:

$$\text{If } d < r \text{ then } \begin{cases} \nabla x = -sign(\cos(\theta))\infty \\ \nabla y = -sign(\sin(\theta))\infty \end{cases} \tag{4}$$

$$\text{If } r \le d \le s + r \text{ then } \begin{cases} \nabla x = -\beta(s + r - d)\cos(\theta) \\ \nabla y = -\beta(s + r - d)\sin(\theta) \end{cases} \tag{5}$$

$$\text{If } d > s + r \text{ then } \nabla x = \nabla y = 0 \tag{6}$$

The attractive potential rules, similarly to the above-listed ones, are simple and describe three different robot behaviors with regards to its positionrelevant to the obstacle. It is essential to notice that all action vectors have to point away from the obstacle, hence the necessity to employ negative values [11].

- In the first rule (step 3), the robot is within the radius of the obstacle, so the action vector needs to be infinite, thus expressing the need to move away from the current location.
- In the second rule, when the robot is not inside the obstacle's radius but inside its area of influence, the action vector is assigned to a high value in order to express the need to move away from the current location.
- The third rule illustrates the case where the robot is outside the area of influence of the obstacle; the action vector is set to zero, and thus no repulsive forces are acting on the robot [14].

Special care must be taken when selecting the value of s, since the repulsive force only acts when the robot is inside the area of the obstacle's influence. If the value of s is small, it may cause trajectory problems by abruptly changing its orientation in its path and leading to speed limitation of the robot. If s has a high value, it can also result in problems in the robot's movement since it might limit its movement in tight spaces where the robot cannot pass.

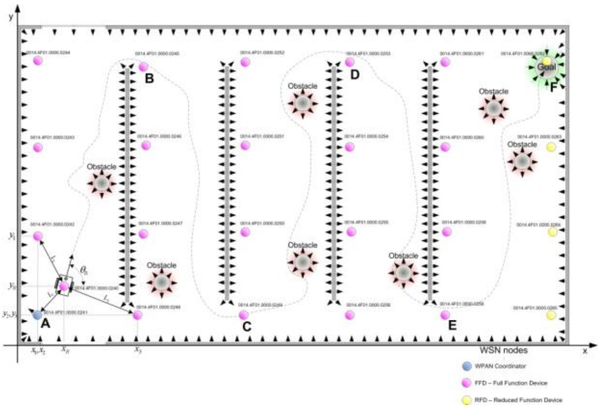


Figure 7

Potential Fields simulation

The repulsive force has the objective of repelling the robot only if it is close to an obstacle and its velocity points towards that obstacle [11]. The robot is able to navigate along the map from point A to point F via points B, C, D, E (see Fig. 7) collecting the environmental parameters.

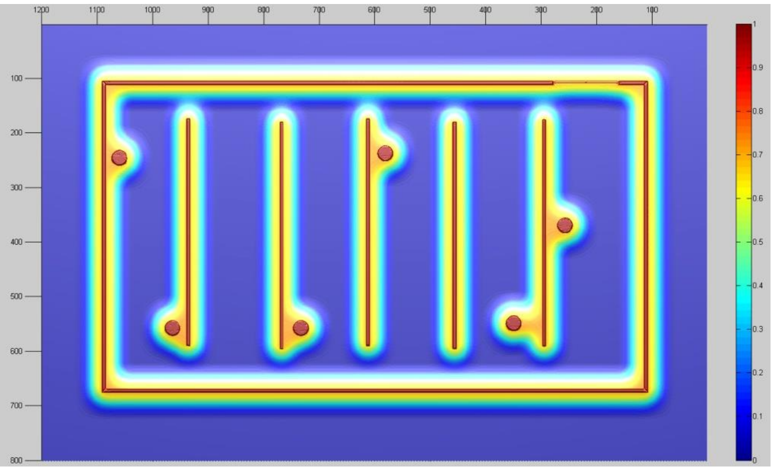


Figure 8

Top view of Potential Fields Values

The 3D surface presented inFig. 9 shows the force vector field of the repelling walls and obstacles, with further references indicated in [2,3]. The surface depicted inFig. 8 presents the repelling forces of the simulated environment. The map shown in Fig. 9 presents the absolute value of the forces of the simulation environment's potential fields.

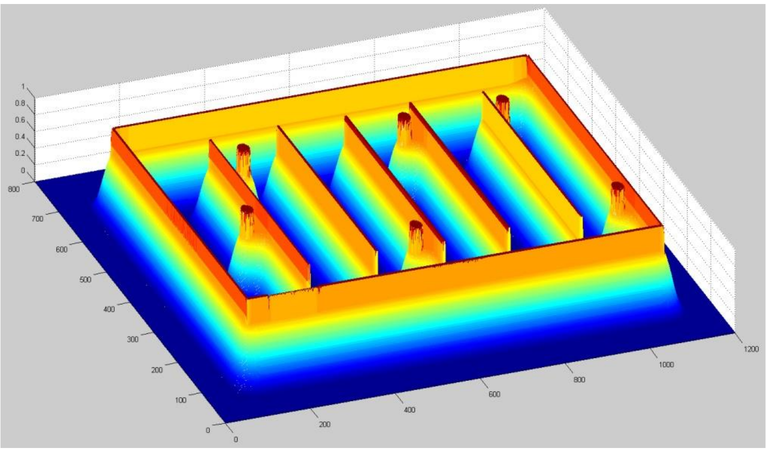


Figure 9

3D view of potential field

It is possible for the robot to be caught in local minima. This simply means that the robot has reached a location where the force is zero, with the opposing forces cancelling each other out. The result will be that the robot ceases to move and its final destination will never be reached [6].

Figure 10
Java based Sun SPOT

The SunSPOT nodes (Fig. 10) radio communication is based on an embedded CC2420 radio circuit which is IEEE 802.15.4 compliant and works in the 2.4 GHz to 2.4835 GHz ISM bands [5]. The CC2420 datasheet specifies the equation to compute the received signal strength (in dbm) using the raw values collected by the sensors.

Conclusions

In order to achieve efficient robot navigation in a controlled micro climatic environment the solution seems to be the application of the WSN as an aid for navigation. In this way we can ensure effective navigation by estimating the robot's position through the exact position of the mobile measuring stations, without the use offurther hardware. As a future extension of this experiment, this application can be implemented in the field of precision agriculture. Another area of possible improvement is the real-time tracking algorithm of the WSN aided navigation in order to increase the precision of the ranges not covered that lie beyond the reference nodes.

References

[1] Andrzej Pawlowski, Jose Luis Guzman, Francisco Rodríguez, Manuel Berenguel, José Sánchez, Sebastián Dormido, 2009, "Simulation Of Greenhouse Climate Monitoring And Control With Wireless Sensor Network And Event-Based Control" pp. 232-252

[2] C. H. Chiang, J. S. Liu and Y. S. Chou. 2009,"Comparing Path Length by Boundary Following Fast Matching Method and Bug Algorithms for Path Planning", Opportunities and Challenges for Next-Generation Artificial Intelligence, Springer, pp. 303-309

[3] Gomide, R. L., Inamasu, R. Y., Queiroz, D. M., Mantovani, E. C., Santos, W. F., 2001,“An Automatic Data Acquisition and Control Mobile Laboratory Network for Crop Production Systems Data Management and Spatial Variability Studies in the Brazilian Center-West Region”, ASAE Paper No.: 01-1046. The American Society of Agriculture Engineers, St. Joseph, Michigan, USA

[4] Gy. Mester, 2009, „Wireless Sensor-based Control of Mobile Robot Motion“, Proceeding of the IEEE SISY 2009, pp 81-84, Subotica, Serbia

[5] István Matijevics and Janos Simon 2010, “Improving Greenhouse's Automation and Data Acquisition with Mobile Robot Controlled System via Wireless Sensor Network”, Wireless Sensor Networks: Application-Centric Design, Geoff V Merrett and Yen Kheng Tan (Ed.), InTech

[6] J. Simon, G. Martinović, 2009, “Web Based Distant Monitoring and Control for Greenhouse Systems Using the Sun SPOT Modules”, Proceedings of the Conference SISY 2009, pp. 1-5, Subotica, Serbia

[7] J. Vasu, L. Shahram, 2008, “Comprehensive Study of Routing Management in Wireless Sensor Networks- Part-1”

[8] K. Kreichbaum. 2006, Tools and Algorithms for Mobile Robot Navigation with Uncertain Localization. PhD thesis, California Institute of Technology

[9] L. Gonda, C. Cugnasca, 2006, “A Proposal of Greenhouse Control Using Wireless Sensor Networks” In Proceedings of 4thWorld Congress Conference on Computers in Agriculture and Natural Resources, Orlando, Florida, USA

[10] Liu, G., Ying, Y., 2003, “Application of Bluetooth Technology in Greenhouse Environment, Monitor and Control”, J. Zhejiang Univ., Agric. Life Sci. 29, 329–334

[11] Luca Bencini, Davide Di Palma, Giovanni Collodi, Antonio Manes and Gianfranco Manes, 2010,“Wireless Sensor Networks for On-Field Agricultural Management Process”, Wireless Sensor Networks: Application-Centric Design, Geoff V Merrett and Yen Kheng Tan (Ed.), InTech

[12] M. J. Matarić, 2007, The Robotics Primer. The MIT Press, 1st edition

[13] Mizunuma, M., Katoh, T., Hata, S., 2003, “Applying IT to farm fields”, A Wireless LAN. NTT Tech. Rev. 1, pp. 56-60

[14] O. Khatib, 1985, “The Potential Field Approach and Operational Space Formulation in Robot Control” Proc. Fourth Yale Workshop on Applications of Adaptive Systems Theory, Yale University, New Haven, Connecticut, pp. 208-214

[15] O. Khatib, 1986, "Real-Time Obstacle Avoidance for Manipulators and Mobile Robots", Int. J. of Robotic Research, Vol. 5, No. 1, p. 60

[16] R. Langer, L. Coelho and G. Oliveira. 2007, "K-Bug, a New Bug Approach for Mobile Robot's Path Planning", IEEE International Conference on Control Applications, pp. 403-408

[17] Roland Siegwart and Illah R., 2004, "Introduction to Autonomous Mobile Robots", Nourbakhsh

[18] Serodio, C., Cunha, J.B., Morais, R., Couto, C. A., Monteiro, J. L., 2001, "A Networked Platform for Agricultural Management Systems", In: Computers and Electronics in Agriculture, vol. 31. Elsevier, pp. 75-90

[19] V. Lumelsky and A. Stepanov. 1987, "Path Planning Strategies for a Point Mobile Automaton Moving Amidst Unknown Obstacles of Arbitrary Shape", Algorithmica, Vol. 2, pp. 403-430

[20] X. Feng, T. Yu-Chu, S. Yanjun, S. Youxian, 2007, "Wireless Sensor/Actuator Network Design for Mobile Control Applications. Sensors" Proceedings of the Conference

[21] Y. Takahashi, T. Komeda, H. Koyama. 2004, "Development of Assistive Mobile Robot System", Amos. Advanced Robotics, Vol. 18, No. 5

[22] Silvester Pletl, Péter Gál, DraganKukolj, László Gogolák, 2010, "An Optimizing Coverage in Mobile Wireless Sensor Networks", 8th International Symposium on Intelligent Systems and Informatics - SISY 2010, Subotica

[23] Joao Sena Esteves, Adriano Carvalho and Carlos Couto, 2003, "Generalized Geometric Triangulation Algorithm for Mobile Robot Absolute Self-Localization", ISIE 2003 - 2003 IEEE International Symposium on Industrial Electronics, Rio de Janeiro, Brazil, pp 1-12

On the Consistency Analyzing of A-SLAM for UAV Navigating in GNSS Denied Environment

A. Ersan Oguz*, Hakan Temeltas**

*Electronics Engineering Department, Turkish Air Force Academy, Yesilyurt, 34149 Istanbul, Turkey (e-mail: ae.oguz@hho.edu.tr)

**Control Engineering Department, Istanbul Technical University, 34469 Istanbul, Turkey (e-mail: hakan.temeltas@elk.itu.edu.tr)

Abstract: In this paper, EKF Based A-SLAM concept is discussed in detail by presenting the formulas and MATLAB Simulink model, along with results. The UAV kinematic model and state-observation models for the EKF Based A-SLAM are developed to analyze the consistency. The covariance value caused by the EKF structure is analyzed. This value was calculated by filtering with error between the UAV's actual and estimated positions. It is concluded that the calculated covariance value diminishes despite error and does not decrease as a result of the inconsistency of the EKF Based A-SLAM structure. The necessary conditions for the consistency of the EKF Based A-SLAM structure are described and the consistency of the EKF Based A-SLAM is consequently investigated by considering these conditions for the first time by using simulation results. The analysis during this landmark observation exhibits that a jagged UAV's trajectory indicates an inconsistency caused by the EKF Based A-SLAM structure with the provocation of error accumulation. Finally, the major reasons of the filter inconsistency can be listed as unobservable subspace and Jacobian matrices used for linearization. As a future work, the methods that can be used to provide consistency of the EKF Based A-SLAM are proposed for better UAV navigation performance.

Keywords: A-SLAM; EKF; SLAM; Navigation

1 Introduction

Unmanned Aerial Vehicles (UAVs) have a quite distinguished role for decision makers and operational agents due to their extensive usage and large variety of applications, such as reconnaissance and surveillance in the military and civilian areas. The main usage purpose of UAVs is autonomous navigation that provides flexible functionality particularly on a stand-alone flight in autonomous operations. It is crucial to determine the UAVs precise position for better navigation. Global Navigation Satellite System (GNSS), which is actually preferred to geolocation tools in the world, is also the most adequate tools for the UAV position determination process. However, it can be hard to have successful results for this process in GNSS denied environments. Thus, there exist a great deal of work in the literature about this problem. In GNSS denied environments, when the map knowledge also does not exist, the problem of simultaneous position determination and production of map information can be solved by Extended Kalman Filter (EKF) Based Airborne Simultaneous Localization And Mapping (A-SLAM), which is the consistency analysis issue in this paper.

The SLAM problem is a hot topic for both industrial and military robotic research. The SLAM problem was introduced by R. Smith and P. Cheesman [1], then was detailed by G. Dissanayake, H. F. Durrant-Whyte and T. Bailey [2], and finally was presented as a whole concept by T. Bailey and H. F. Durrant-Whyte [3, 4]. While the researchers were interested in the SLAM problem for territorial, aerial, and marine vehicles, they examined a way to decrease the error that is caused by filter structure. In these studies, the structure of the Kalman filter-based SLAM effects of partial observability [5-7], stability [8] and the consistency problem [9, 10] were examined to present recommendations for solutions to the problem. Airborne SLAM applications are continued simultaneously [11-14]. It was very important to know the current location in the works on autonomous navigation in aircraft [15, 16] and the creation of an accurate map. The optimal level of mapping method was also examined in L. M. Paz and J. Neira [17].

In the literature, filter consistency can be studied as a separate title. In the case that the estimated covariance of the filter is greater than or equal to the actual covariance that is the covariance difference between actual and estimated positions of UAVs, the filter is called consistent. The filter consistency is very important in terms of performance. When the filter is inconsistent, error accumulation increases, and consequently the vehicle location is determined with some error. The filter consistency is still an attractive research topic and researchers have made numerous attempts to improve consistency. The consistency conditions of the Kalman filter, named the state estimator, were revealed by Y. Bar-Shalom, X. Rong Li and T. Kirubarajan [18]. T. Vidal-Calleja, J. Andrade-Cetto and A. Sanfeliu demonstrated the marginal stability status of the Kalman filter [8], and J. A. Castellenos, J. Neira and J. D. Tardos analyzed the consistency of the suboptimal filter and examined limits to the SLAM consistency [10]. It is shown by T. Bailey, J. Nieto, J. Guivant, M. Stevens and E. Nebot that in the case of two core symptoms, the EKF-based SLAM filter is inconsistent [9].

However, the inconsistency problem for an aerial vehicle has not been identified yet. SLAM filters that are designed for aerial vehicles must include the kinematic model. In this case, the filter consistency problem becomes more complex. In this work, a SLAM filter designed for aircraft and the consistency problem have been analyzed for the aircraft kinematics model. In the literature, the SLAM problem for an aircraft is known as an A-SLAM. The inconsistency in the EFK-based A-SLAM filters is expressed clearly with the proposed analysis methods.

The aircraft kinematic model is shown in detail in Section 2. The structure of the EKF based A-SLAM is described in Section 3. In the subsections, non-linear prediction and the observation model are demonstrated, the procedures in estimation and correction steps are described and the Jacobian matrices used for linearization are established. The problem of consistency and its symptoms are described in Section 4. Simulations for several scenarios are performed and, according to the simulation results, the consistency structure of the EKF based A-SLAM are examined in Section 5. The obtained simulation results were evaluated in Section 6 and finally, the solution methods proposed in future works with emerging inconsistency problems are discussed.

2 The UAV Kinematical Model

The transformation of the body frame motions of an aerial vehicle to the navigation frame plays an important role since the global mapping of the environment requires aerial vehicle equation of motion to the global frame. Euler angle transformations can be used for this purpose. The aerial vehicle body frame accelerations of the directional navigation and angular rates are transferred to the navigation frame, and the position of the aerial vehicle is calculated in navigation frame. The general equation generated during this process is expressed as the kinematic equation of the aerial vehicle. The matrix expression obtained by the transfer of directional acceleration to navigation frame of aerial vehicles is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\Psi & \cos\Psi\sin\theta\sin\phi - \cos\phi\sin\Psi & \cos\Psi\sin\theta\cos\phi + \sin\phi\sin\Psi \\ \cos\theta\sin\Psi & \cos\phi\cos\Psi + \sin\theta\sin\phi\sin\Psi & -\sin\phi\cos\Psi + \sin\theta\cos\phi\sin\Psi \\ -\sin\theta & \cos\theta\sin\phi & \cos\theta\cos\phi \end{bmatrix} . \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad (1)$$

The transformation matrix of angular rates from body frame to navigating frame is

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta \end{bmatrix} . \begin{bmatrix} p \\ q \\ r \end{bmatrix} \qquad (2)$$

These two matrix equations are referred to as the kinematic model of the aerial vehicle. The vector [u,v,w] is the directional acceleration in the body frame and [p,q,r] is the angular rates of the body frame.

3 EKF-based Airborne SLAM

An EKF-based filter was used for the A-SLAM in this section. First, the non-linear model of prediction and observation is given and then the estimate and update of the steps of the EKF are described.

3.1 The Non-Linear Prediction and Observation Model

The aerial vehicle state vector and map vector come from the state vector used in non-linear system model. The aerial vehicle state vector is composed of the position (x, y, z), velocity (Vx, Vy, Vz) and the Euler angles (roll, pitch and yaw). The map vector is made up of the position of each landmark (x_L, y_L, z_L). The length of this vector is 3*n (n: number of landmark).

$$x(k) = \begin{bmatrix} x_{UAV} \\ x_{MAP} \end{bmatrix} \tag{3}$$

EKF is a filter structure in a non-linear system that estimates the next step from the previous state and the observation value. As in the Kalman filter, it has estimation and correction steps. The state space model expression that will be used in filters is

$$x(k+1) = f(x(k), u(k), v(k)) \tag{4}$$

In the state space model, the state vector at step of k+1 depends on the state vector and input data of step k.

The landmark position can be found as the range, bearing, and elevation as in the observation model. The transformation of that value from sensor frame to navigation frame is performed by the transformation matrix. The observation model depends on the position information of the aerial vehicle at step k. It is stated as

$$z(k+1) = h(x(k), \omega(k)) \tag{5}$$

where v(k) and w(k) are zero mean white noise. Their covariance is $Q = \left[v(k).v(k)^T\right]$ and $R = \left[\omega(k).\omega(k)^T\right]$

The noise covariance can be written as

$$E\left[v(k)\right] = 0$$

$$E\left[v(k)v(k)^T\right] = Q(k) \tag{6}$$

and

$$E\left[w(k)\right] = 0$$

$$E\left[w(k)w(k)^T\right]=R(k) \tag{7}$$

In equation (4) the state space model in explicit for can be written as

$$f_{uav}=\begin{bmatrix} P^n(k) \\ V^n(k) \\ \Psi^n(k) \end{bmatrix}=\begin{bmatrix} P^n(k-1)+\Delta t*V^n(k-1) \\ V^n(k-1)+\Delta t*T_b^n(k-1)*f^b(k)] \\ \Psi^n(k-1)+\Delta t*E_b^n(k-1)*w^b(k) \end{bmatrix}+\begin{bmatrix} w_{P^n} \\ w_{V^n} \\ w_{\Psi^n} \end{bmatrix} \tag{8}$$

Here it will be seen that the state space model is composed of the position, directional velocity and Euler angle. The state space model is obtained from the kinematic equation of the aerial vehicle.

3.2 EKF Estimation Step

The A-SLAM needs the stochastic estimation, which is a guess of the UAV position parameter at step k from the previous state with Gaussian noise. The structure is constructed by the kinematic system model as compatible with the EKF structure. As in the EKF-based A-SLAM it is elementarily composed of estimation and update steps. For the model linearization process, the Jocobian matrices are used. In the EKF-based A-SLAM, the state estimation at step k, due to input variables and the state at step-k, can be written as

$$\hat{x}(k+1|k)=f(x(k|k),u(k)) \tag{9}$$

The observation estimation is

$$z(k+1|k)=h(\hat{x}(k+1|k)) \tag{10}$$

The estimated covariance calculated by is

$$P(k+1|k)=\nabla F_x.P(k|k).\nabla F_x^{\ T}+\nabla F_u.Q.\nabla F_u^{\ T} \tag{11}$$

3.3 EKF Update State

The estimation update step is performed next to the EKF estimation step. The updated state estimation expression is

$$x(k+1|k+1)=\hat{x}(k+1|k)+W.\upsilon(k+1) \tag{12}$$

where W is the Kalman gain and $\upsilon(k+1)$ is innovation. The updated covariance is

$$P(k+1|k+1)=P(k+1|k)-W.S(k+1).W^T \tag{13}$$

The calculation of the innovation and its covariance (covariance error) is done by

$$\upsilon(k+1)=z(k+1)-z(k+1|k) \tag{14}$$

$$S(k+1) = \nabla H_x P(k+1|k) \nabla H_x^{\ T} + R \tag{15}$$

Finally the Kalman gain is

$$W = P(k+1|k).\nabla H_x^T .S^{-1}(k+1) \tag{16}$$

The Jacobian matrix $\nabla f_{uav}(k)$ used for the model linearization in A-SLAM that is composed of partial differentials of position, velocity and the Euler angle of the UAV kinematic model is

$$\nabla f_{uav}(k) = \begin{bmatrix} \dfrac{\partial P^n(k)}{\partial P^n(k-1)} & \dfrac{\partial P^n(k)}{\partial V^n(k-1)} & \dfrac{\partial P^n(k)}{\partial \Psi^n(k-1)} \\ \dfrac{\partial V^n(k)}{\partial P^n(k-1)} & \dfrac{\partial V^n(k)}{\partial V^n(k-1)} & \dfrac{\partial V^n(k)}{\partial \Psi^n(k-1)} \\ \dfrac{\partial \Psi^n(k)}{\partial P^n(k-1)} & \dfrac{\partial \Psi^n(k)}{\partial V^n(k-1)} & \dfrac{\partial \Psi^n(k)}{\partial \Psi^n(k-1)} \end{bmatrix} \tag{17}$$

The Jacobian matrix $\nabla f_w(k)$ is composed of the differential equation of the non-linear process gain model input variables is

$$\nabla f_w(k) = \begin{bmatrix} \dfrac{\partial P^n(k)}{\partial f^b(k-1)} & \dfrac{\partial P^n(k)}{\partial w^b(k-1)} \\ \dfrac{\partial V^n(k)}{\partial f^b(k-1)} & \dfrac{\partial V^n(k)}{\partial w^b(k-1)} \\ \dfrac{\partial \psi^n(k)}{\partial f^b(k-1)} & \dfrac{\partial \psi^n(k)}{\partial w^b(k-1)} \end{bmatrix} \tag{18}$$

In the observation model, the landmark object information from the sensor is range, bearing and elevation. The presentation of object information in the Cartesian coordinates is

$$z(k) = [\rho \ \varphi \ \vartheta]^T$$

$$= \begin{bmatrix} \sqrt{x_s^{\ 2} + y_s^{\ 2} + z_s^{\ 2}} \\ a\tan(y_s / x_s) \\ a\tan(z_s / \sqrt{x_s^{\ 2} + y_s^{\ 2}}) \end{bmatrix} \tag{19}$$

The transformation of the landmark object sensor frame position to the navigation frame is

$$Z^n(k) = P^n + (T_b^n * L_b^n) + (T_b^n * T_s^b * z(k)) \tag{20}$$

The symbol α refers to the angle between the body and the observation sensor camera. The transfer matrix is

$$T_s^b = \begin{bmatrix} Cos(-\alpha) & 0 & -Sin(-\alpha) \\ 0 & 1 & 0 \\ Sin(-\alpha) & 0 & Cos(-\alpha) \end{bmatrix} \tag{21}$$

The Jacobian matrix $\nabla h(k)$ used for model linearization in the A-SLAM that is composed of partial differentials of position, velocity and the Euler angle of observation model is

$$\nabla h = H = \begin{bmatrix} \frac{\partial \rho(k)}{\partial P^n(k)} & \frac{\partial \rho(k)}{\partial V^n(k)} & \frac{\partial \rho(k)}{\partial \Psi^n(k)} \\ \frac{\partial \varphi(k)}{\partial P^n(k)} & \frac{\partial \varphi(k)}{\partial V^n(k)} & \frac{\partial \varphi(k)}{\partial \Psi^n(k)} \\ \frac{\partial \vartheta(k)}{\partial P^n(k)} & \frac{\partial \vartheta(k)}{\partial V^n(k)} & \frac{\partial \vartheta(k)}{\partial \Psi^n(k)} \end{bmatrix} \tag{22}$$

In explicit form it is

$$\nabla h = \begin{bmatrix} \frac{x_s}{\sqrt{x_s{}^2 + y_s{}^2 + z_s{}^2}} & \frac{y_s}{\sqrt{x_s{}^2 + y_s{}^2 + z_s{}^2}} & \frac{z_s}{\sqrt{x_s{}^2 + y_s{}^2 + z_s{}^2}} \\ \frac{-y_s}{x_s{}^2 + y_s{}^2} & \frac{x_s}{x_s{}^2 + y_s{}^2} & 0 \\ \frac{-x_s * z_s}{(x_s{}^2 + y_s{}^2 + z_s{}^2) * \sqrt{x_s{}^2 + y_s{}^2}} & \frac{-y_s * z_s}{(x_s{}^2 + y_s{}^2 + z_s{}^2) * \sqrt{x_s{}^2 + y_s{}^2}} & \frac{\sqrt{x_s{}^2 + y_s{}^2}_{\ s}}{x_s{}^2 + y_s{}^2 + z_s{}^2} \end{bmatrix} \tag{23}$$
$$* \begin{bmatrix} -(T_b^n * T_s^b)^T & 0_{3x3} & \frac{\partial T_b^s * T_n^b * (L^n - P^n)}{\partial \Psi^n} & (T_b^n * T_s^b)^T \end{bmatrix}$$

where

$$\frac{\partial T_b^s * T_n^b * (L^n - P^n)}{\partial \Psi^n} = \begin{bmatrix} \frac{\partial x_s}{\partial \phi} & \frac{\partial x_s}{\partial \theta} & \frac{\partial x_s}{\partial \psi} \\ \frac{\partial y_s}{\partial \phi} & \frac{\partial y_s}{\partial \theta} & \frac{\partial y_s}{\partial \psi} \\ \frac{\partial z_s}{\partial \phi} & \frac{\partial z_s}{\partial \theta} & \frac{\partial z_s}{\partial \psi} \end{bmatrix} \tag{24}$$

4 The Inconsistency Problem and Symptoms

The EKF structure is developed to apply in non-linear systems. Jocabian matrices are used to linearize the non-linear systems. Using Jacobian matrices causes new errors in the Kalman filter estimations.

Since the EKF is also applicable in non-linear systems, it is used in the SLAM structure. The simultaneous reduction in the error and filter covariance as in the Kalman filter is expected result of the EKF based A-SLAM structure.

Y. Bar-Shalom et al. [18] stated that a state estimator is consistent in the event that it satisfies the following requirements:

- Estimation error mean is zero,
- Real covariance is less than or equal to the covariance calculated by the filter.

The EKF SLAM filter consistency is analyzed by T. Bailey et al. [9], and it is stated that EFK SLAM is inconsistent in case of the observation of the following symptoms:

- Excessive information gain (estimated error covariance is less than the real covariance),
- Jagged in vehicle movement curve.

5 Numerical Result

The state equation given in section 3 and the A-SLAM structure is used in the simulation. The UAV trajectory and landmark is specified due to a certain scenario. The EKF-based A-SLAM simulation results are transferred to the graphics and, using the results filter consistency, analysis is performed.

5.1 EKF-based A-SLAM

In the simulation, a UAV at a constant velocity and constant altitude is flying at 30m/sec over an area of 1400 m x 1400 m in a closed loop. The landmark detection with camera and position determination is performed. The UAV is supposed to start flying at [0, 0, 0]. The total simulation time is bounded at 240 seconds. An IMU is used as a sensor. The eighth landmark object is depicted in the map. The UAV trajectory and landmarks are depicted in Fig. 1.

The UAV directional velocity in Fig. 2 is given versus time. The UAV localization error is shown in Fig. 3, the directional velocity error in Fig. 4, the Euler angle error in Fig. 5, and the X axis position error and filter covariance in Fig. 6.

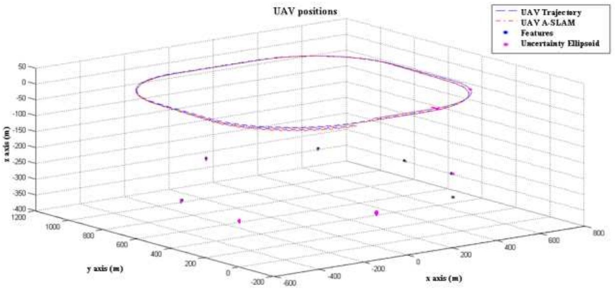


Figure 1
UAV and Landmark Localization. In this plot, the blue dashed line is the UAV trajectory, the red dotted line is the A-SLAM, the blue ellipse is the landmark localization, and the red ellipse is the uncertainty ellipsoid

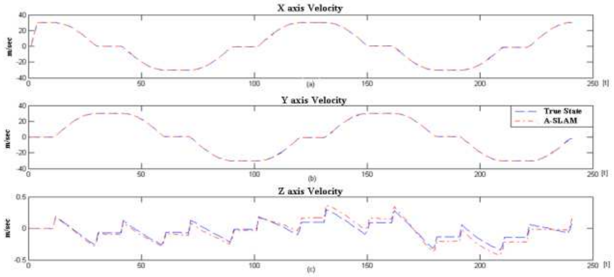


Figure 2

UAV Directional Velocities versus Time. (a) X axis velocity, (b) Y axis velocity, (c) Z axis velocity. In this plot, the blue dashed line is the UAV real velocity, and the red dotted line is the velocity calculated by A-SLAM

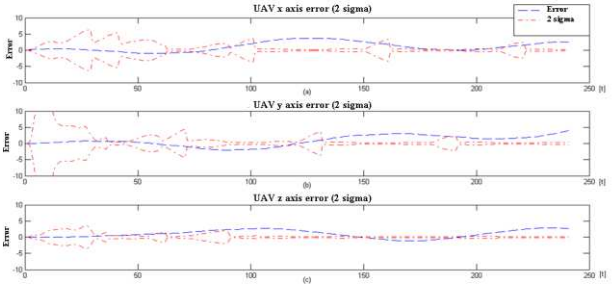


Figure 2

UAV Localization Error. (a) UAV X axis error, (b) UAV Y axis error, (c) UAV Z axis error. In this plot, the blue dashed line is the UAV position error, and the red dotted line is the 2 sigma confidence interval

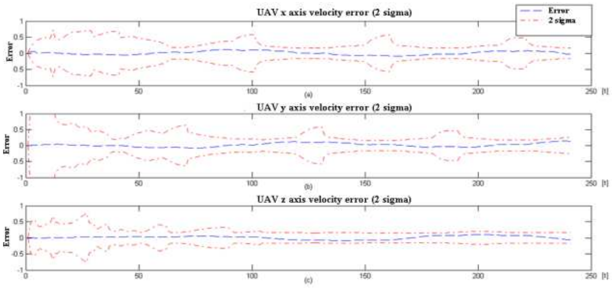


Figure 3

UAV Velocity Error. (a) UAV X axis velocity error, (b) UAV Y axis velocity error, (c) UAV Z axis velocity error. In this plot, the blue dashed line is the UAV velocity error, and the red dotted line is the 2 sigma confidence interval

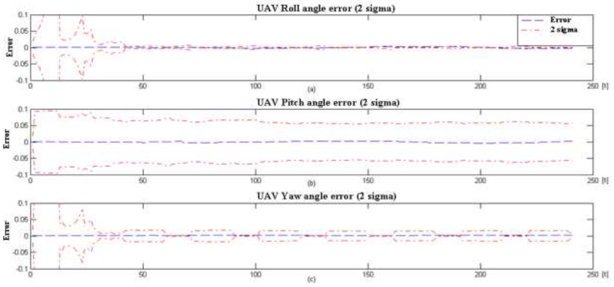


Figure 4

UAV Euler Angle Error. (a) UAV roll angle error, (b) UAV pitch angle error, (c) UAV yaw angle error. In this plot, the blue dashed line is the UAV angle error, and the red dotted line is the 2 sigma confidence interval

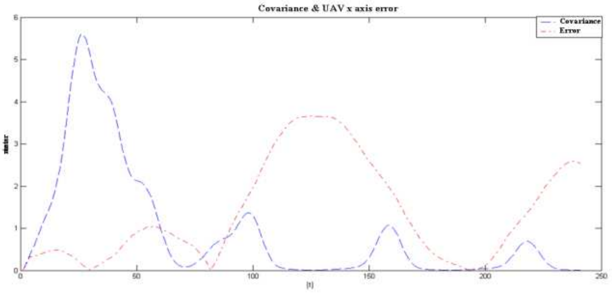


Figure 5

Position Error on X Axis and Filter Coveriance. In this plot, the blue dashed line is the filter X axis covariance, and the red dotted line is the X axis error

5.2 First Symptom: Information Gain

It is necessary that the filter covariance should decrease over time in a consistent filter structure, and the error must be smaller than the estimated covariance. In a similar manner, the consistency of the EKF based A-SLAM structure can be analyzed.

This analysis can be done in two different ways: first by checking the error and covariance relation by using simulation results, and second, by making use of the normalize estimation error square value.

5.2.1 Filter Covariance – Error Relation

According to the values used in the simulation scenario, the second moments of the UAV's [x,y,z] position variables with the Airborne SLAM inspection of the estimated values of covariance are:

As the second moment:

$$P(k) \Rightarrow \mathrm{var}(x) = 0{,}1628; \mathrm{var}(y) = 0{,}6432; \mathrm{var}(z) = 0{,}1751 \cdot \tag{25}$$

The estimated Coveriance:

$$P(k|k) \Rightarrow P_x = 0{,}0051 \quad P_y = 0{,}0258 \quad P_z = 0{,}0063 \tag{26}$$

Since $P(k|k) < P(k)$ the filter is inconsistent.

5.2.2 The Normalized Estimation Error Square

If the actual statistical values are unknown but the true state x(k) is known, the normalised Estimation Error Squared (NEES) can be used.

Let N-dimensional random numbers vector x be in Gaussian distribution, the mean be $\overline{x}$ and covariance be P, then q=$(x-\overline{x})'.P^{-1}.(x-\overline{x})$ can be written in chi-square distributed quadratic form with N dimension of freedom. The Chi-square distribution can be defined as the ratio of the squares of two Gaussian distributed variables.

The chi-square conformity test is a decision as to whether the difference between the expected and obtained value is within a specified limit or not. The process test not only variable distribution but also the whole filter system.

In the EKF-based A-SLAM structure, if the UAV's real position $x(k)$ is known, then the difference between the UAV's real position $x(k)$ and calculated position, the term $x(k|k)$ can be computed. The NEES or filter performance characteristic is:

$$\varepsilon(k) = (x(k) - x(k|k))'.P(k|k)^{-1}.(x(k) - x(k|k)) \tag{27}$$

The NEES distribution is expected to be a chi square, or if the computed distribution is not chi-square, then the filter is not consistent.

The chi-square distribution depends on the degrees of freedom of system. The expected bounds of distribution with degrees of freedom are determined using a table. The UAV's position is specified by nine degrees of a freedom-position vector composed of the velocity and angle values. The NEES, with nine degrees of freedom and its reliability boundary, is given in Fig.6. The computed NEES in a filter should be within the reliability boundary for consistency. Since it is not, the proposed filter of the EKF based A-SLAM is inconsistent.

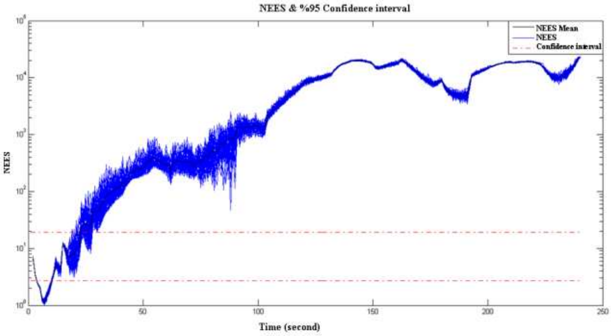


Figure 6
EKF Based A-SLAM, 50 Monte Carlo Simulation, NEES 95% confidence interval. In this plot, the blue line is the NEES, the black line is the NEES mean, and the red dashed lines are the 95% confidence interval

5.3 Second Symptom: Jagged UAV Trajectory

The other results of the filter inconsistency are jumps in the UAV's trajectory. That originates from the UAV's correction when it detects a landmark. This symptom appears by the size of the update, and the UAV's position tends to be much larger than the actual error. When the landmark is observed, the UAV's trajectory is jagged, as shown in Fig. 7.

Both symptoms indicate inconsistency with the A-SLAM consistency investigation. Despite the accumulation of errors, which are the difference between the UAV's real position and the filter estimation, the A-SLAM covariance decreases, as shown in study. The reason for filter inconsistency is the limited observability of the A-SLAM or other observability problems.

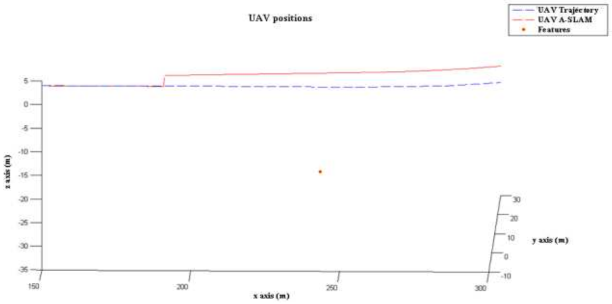


Figure 7
Jagged UAV Trajectory. In this plot, the blue dashed line is the UAV trajectory, the red line is the A-SLAM, and the red ellipse is the landmark localization

Conclusion and Future Work

In this paper, a mathematical model and the simulation results for the EKF-based A-SLAM structure of UAVs are presented in order to investigate the consistency of the EKF-based A-SLAM. It is observed that the error stems from the Jacobian matrices used for linearization. Therefore, the uncertainty emerged from Jacobian matrices and the effects of the landmark uncertainty that causes error accumulation. The filter estimation covariance should be greater or equal to the error, and the consistent filter is supposed to calculate covariance in accordance with error. Since it is stated that the filter is not able to respond to error accumulation, this implies filter inconsistency.

The inconsistency in the EKF-based A-SLAM structure was emphasized and background information for the solution of the problem was provided. It was also observed that main reason for filter inconsistency is the information gain arising from unobservable subspace. The imperfect or incorrect information coming from the unobservable subspace filter does not respond to error accumulation.

In future works, methods such as extending observability, designing an observability constrained filter, or constructing a full observable filter structure can be applied to making the filter consistent. Furthermore, Jacobian matrices that are used in the A-SLAM, can be calculated by methods, except for differential methods.

References

[1] R. Smith and P. Cheesman, On the Representation of Spatial Uncertainty, Int. J. Robot. Res., Vol. 5, No. 4, pp. 56-68, 1987

[2] G. Dissanayake, H. F. Durrant-Whyte and T. Bailey, A Computationally Efficient Solution to the Simultaneous Localisation and Map Building (SLAM) Problem, ICRA, 2000

[3] T. Bailey and H. F. Durrant-Whyte, Simultaneous Localization and Mapping (SLAM): Part I The Essential Algorithms., IEEE Robotics and Automation Magazine, 2006

[4] T. Bailey and H. F. Durrant-Whyte, Simultaneous Localisation and Mapping (SLAM): Part II State of the Art., IEEE Robotics and Automation Magazine, September 2006

[5] J. Andrade-Cetto, and A. Sanfeliu, The Effects of Partial Observability in SLAM, IEEE Int. Conf. Robot. Automat., New Orleans, pp. 397-402, 2004

[6] M. Bryson, and S. Sukkarieh, Observability Analysis and Active Control for Airborne SLAM, IEEE Transaction on Aerospace and Electronic Systems, Vol. 44, No. 4, pp. 261-280, 2008

[7] G. P. Huang, N. Trawny, A. I. Mourikis and S. I. Roumeliotis, Observability-based Consistent EKF Estimators for Multi-Robot

Cooperative Localization, Autonomous Robots, Volume 30, Number 1, pp. 99-122, 2011

[8] T. Vidal-Calleja, J. Andrade-Cetto and A. Sanfeliu, Conditional for Suboptimal Filter Stability in SLAM, IEEE Inernational Conferance on Intelligent Robots and Systems, Sendai Japan, 2004

[9] T. Bailey, J. Nieto, J. Guivant, M. Stevens and E. Nebot, Consistency of the EKF-SLAM Algorithm, IROS'06 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006

[10] J. A. Castellanos, J. Neira, J. Tardos, Limits to The Consistency of EKF-Based SLAM, IAV2004 5th IFAC Symp. on Intelligent Autonomous Vehicles, 2004

[11] J. Kim and S. Sukkarieh, Airborne Simultaneous Localisation and Map Building, IEEE Int. Conf. Robot. Automat., Taipei Taiwan, 2003

[12] A. E. Oguz and H. Temeltas, Extended Kalman Filter Based Airborne Simultaneous Localization And Mapping, Journal of Aeronautics and Space Technologies, Vol. 6, No. 2, pp. 69-74, 2013

[13] J. Kim, S. Sukkarieh, S. Wishart, Real-Time Navigation, Guidance and Control of a UAV Using Low-Cost Sensors, International Conferance of Field and Service Robotics, Yamanashi, Japan, pp. 95-100, 2003

[14] J. Kim, S. Sukkarieh, Autonomous Airborne Navigation in Unknown Terrain Environments, IEEE Transaction on Aerospace and Electronic Systems, Vol. 40, No. 3, pp. 1031-1044, 2004

[15] S. Kurnaz, O. Cetin and O. Kaynak, Fuzzy Logic Based Approach to Design Flight Control and Navigation Tasks for Autonomous UAVs, Journal of Intelligent and Robotic Systems, v. 54, pp. 229-244, 2009

[16] S. Kurnaz, and O. Cetin, Autonomous Navigation and Landing Tasks for Fixed Wing Small Unmanned Aerial Vehicles, Acta Polytechnica Hungarica, Vol. 7, No. 1, 2010

[17] L. M. Paz, and J. Neira, Optimal local map size for GKF-based SLAM, IROS'06 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006

[18] Y. Bar-Shalom, X. Rong Li and T. Kirubarajan, "Estimation with Applications to Tracking and Navigation", John Wiley and Sons, 2001, pp. 232-145

Running Time Comparison of Two Realizations of the Multifractal Network Generator Method

Árpád Horváth

Óbuda University, Alba Regia University Centre, Székesfehérvár, Hungary
horvath.arpad@arek.uni-obuda.hu

Abstract: Over the last decade a lot of common properties were found in complex networks in several fields such as sociology, biology and computer engineering. Recently, the multifractal network generator method has been developed, and it seems to be a promising way to generate networks with prescribed statistical properties. For educational purposes, however, it would be adequate to create an easy-to-use redevelopment framework. Therefore, a software package had been developed in Python language that can generate a network with a given degree distribution or a given average degree using the multifractal generator method. This package is a part of the cxnet framework, which itself is suitable for educational applications. The present paper discusses the reasons why this framework was developed in Python. Those parts of the program that need longer running times were identified and rewritten in C++. Running times of the generations were measured, changing several parameters, and the new version turned out to be an order of magnitude faster.

Keywords: complex network; graph theory; multifractal; software

1 Introduction

Networks have a collection of entities, called nodes. These nodes can be connected or not, so the networks can be described as a graph in every moment. Complex networks are very large networks with a usually different structure from that of the random network. One of the aims of the science of complex networks is to study the general properties of real networks.

There are a lot of networks in the fields of engineering and informatics (the World Wide Web, the Internet), biology and medicine (network of protein interactions, the food chain) and sociology (acquaintances). Over the last decade, many networks and network models have been studied [1, 8].

To study the general properties of networks, one usually needs a method to create networks with prescribed properties. To create such networks, one can use optimization, which means that we change some parameters to approach the properties we want to achieve. A promising optimizing method is the multifractal network generator [9]. This was improved to create less isolated nodes [10]; however at the size of real networks, the original method is reasonable. Using the multifractal network generator, a broad range of networks with arbitrary properties can be generated. The entropy of such generated networks is bigger than that of the other usually-used models, such as the Erdős-Rényi model, the small world model and Barabási-Albert model [4]. With this method, we can set more than one target property, for example a power-law function as the degree distribution and a clustering coefficient decreasing inversely proportional to the degrees of the nodes. These two properties can be found in many real-world networks and in the resulting networks of the hierarchical model [11]. Our goal is to develop an educational framework that is suitable for generating and analyzing networks, and then students would be able to develop standalone functions to extend the possibilities of the framework. The framework has been implemented in Python language, but some parts were written in C++ language as well. In this paper we describe the method and compare the running times of the two versions: the one written in pure Python and the other, where the calculation of the degree distribution is written in C++.

2 The Multifractal Network Generation Method

The method of generating networks with the usage of multifractals is described in detail in the article of Palla et al. [9].

In multifractal network generation, the generating measure is a central concept. The generating measure is a probability measure defined on the $[0,1[\times [0,1[$ unit square. A network can be generated from the generating measure in two steps. The first step is to create a link probability measure with the iteration of the generating measure. In the second step the program creates links between the nodes using the link probability measure. During generation, however, the program does not need to generate any networks; it calculates the estimated properties of the network from the generating measure.

2.1 Generating Measure and Link Probability Measure

Both the x and y axes of the unit squares are divided into m not necessarily equal intervals to define a generating measure. In our version, the x and y axes have the same division points. With this division we created m^2 rectangles on the unit square. Probabilities p_{ij} are assigned to each of the rectangles in a symmetric fashion,

$$p_{ij} = p_{ji}, \quad \sum_{i,j=0}^{m-1} p_{ij} = 1 \tag{1}$$

The probability assigned to the rectangle at the origin is denoted by the p_{00} and the one at the opposite corner is $p_{m-1,m-1}$.

The K^{th} iteration of the generating measure means a unit square divided into rectangles with assigned probabilities, as in the generating measure, but with $m^k \times m^k$ rectangles. The first iteration ($K = 1$) by definition gives the generating measure itself.

For the case of $K > 1$, we obtain the division points from the division points of the (K-1)st iteration by dividing each of its intervals into m subintervals, where the length of subintervals are proportional to the length of intervals of the original generating measure.

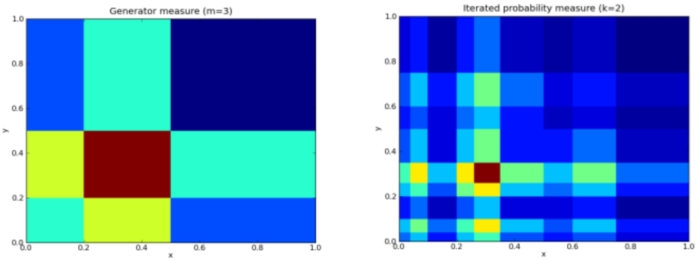
The $p_{ij}(K)$ probabilities of the Kth iteration can be calculated as

$$p_{ij}(k) = \prod_{q=1}^{K} p_{i_q j_q}, \tag{2}$$

where

$$i_q = \lfloor \frac{i \; mod \; m^{K-q+1}}{m^{K-q}} \rfloor \tag{3}$$

The notation (i mod d) means that the remainder of the integer division i/d and $\lfloor x \rfloor$ denotes the floor (integer part) of x. Analogous equation to (3) gives j_q as well.



(a) (b)

Figure 1

A generating measure (a) with the division points 0.2 and 0.5, and the link probability measure resulted by two iterations (b)

2.2 Generating the Network

The generation of networks proceeds in two steps. The first step is the iteration of the generating measure to get the link probability measure. The second one is the generation of the network from the link probability measure obtained after the iterations. The latter goes as follows.

First the number of iterations (K) and the number of nodes (N) in the network need to choose. If one axis of the generating measure is divided into m intervals with the division points, there will be m^K intervals in one axis of the link probability measure. We assign to each node with index l ($l \in [1, N]$ integer) a r_l

random value from a uniform distribution on the [0,1[interval. We determine the i_l index of the interval where r_l is located ($i_l \in [0, m^{K-1}]$).

For all pairs of the nodes we determine whether to connect the nodes or not. If the r_{l1} and r_{l2} random number belonging to the two nodes falls into the intervals i_{l1} and i_{l2}, respectively, we connect the two nodes with the probability $p_{i_{l1},i_{l2}}(K)$, where the values of $p_{ij}(K)$ are the probabilities after the Kth iteration of the generating measure defined in equation (3).

2.3 Adjusting the Generating Measure

The creation of the generating measure can be shown as an annealing process where the energy of the generating measure closes to the minimum as the temperature decreases.

To create a generating measure that gives a network with a given target property, we need to define an energy function (a non-negative function) that measures the goodness of the generating measure. The smaller energy, the closer the network created from the generating measure to the one with the target property.

After giving the m numbers of intervals on one axis and the K number of iteration, our program starts with equal probabilities and equal interval lengths on the axes. In each step it either relocates a division point or changes the probabilities. Then it calculates the energy belonging to the generating measure. If this E' energy is smaller than that belonging to the network of the existing generating measure E, than it changes the generating measure to the new one and stores the energy. If $E' > E$, then the new generating measure will be accepted with the probability

$$P(T) = \exp\left(-\frac{E'-E}{T}\right), \tag{4}$$

and rejected with $1 - P(T)$ probability. The arbitrary parameter T plays the role of temperature (in units of the energy).

Decreasing the temperature slowly, the generating process has the possibility to escape from local minima. The smaller the temperature, the more changes will be rejected, and the network converges to that with the target property.

2.4 Calculating the Degree Distribution

One of the targets of the generation can be the degree distribution. The degree distribution $p(k)$ is a function of degree k giving the probability of a node having the degree k. The expected values of a generating measure can be calculated as

$$p(k) = \sum_{i=0}^{m^k} p_i(k) l_i, \tag{5}$$

where

$$p_i(k) = \frac{<k_i>^k}{k!} e^{-<k_i>}, \tag{6}$$

$$< k_i > = N \sum_j p_{ij} l_j \tag{7}$$

$< k_i >$ is the expected degree of a node in the i$^{\text{th}}$ interval, and $l_i = d_{i+1} - d_i$ is the length of the i^{th} interval.

We can define the energy of a link probability measure as

$$E = \sum_{k=k_{\min}}^{k_{\max}} \frac{p^*(k) - p(k)}{\max(p^*(k), p(k))} \tag{8}$$

where $p^*(k)$ is the degree distribution of the actual link probability measure, and $p(k)$ is the target degree distribution.

3 Results

3.1 The mfng Program

There is an existing implementation of the multifractal network generator written in C++ without the option of setting target properties [9]. Its source code is unfortunately not available. Our earlier works are about the *cxnet* framework (a Python package) we developed to investigate complex networks and bring them into higher education [6-7]. The *mfng* generator does not need the *cxnet* framework, but the analysis of the result needs it. During generation, the program does not create networks. It calculates the expected values of the degree distribution from equation (4) (see below). The *analyser* module of the *mfng* software package provides three main features:

1) It can generate networks from the generating measure constructed by the mfng generator.

2) It can calculate the degree distributions of these networks.

3) It can plot the degree distributions of these networks as well as the degree distribution calculated from equation (4). It can use several binning methods to create clearer plots. Figure 3 (a) is an example plot created using the analyser module.

Earlier the *mfng* generator and *analyser* was a sub-module of the *cxnet* package. To make the installation of the *mfng* easier it has become a standalone package. The documentation of cxnet with the installation of the *mfng* package and a tutorial can be reached from the page [12]. The *mfng* program can be reached from its repository [13] using the git version control system, or can be downloaded as zip or gzipped tar archive from there.

The mfng module includes the *ProbMeasure* class, the *Generator* class and some property classes. An instance of the *ProbMeasure* class contains the probabilities and the division points. It includes a function to iterate the measure, returning with a new *ProbMeasure* instance. This function makes it possible to create the link probability measure from the generating measure using the numpy module. The *mfng* program generates the generating measure for the networks with the given properties. There are two steps with the same temperature T. In one step, the generator changes the probabilities; in the second step it changes the division points. The *Generator* class stores the main parameters of the generation and the property we want to achieve. The main parameters of the generation are the initial and final temperature, the number of steps, the number of intervals in the generating measure and the K number of iteration.

3.2 Changing the Division Points and the Probabilities

In two alternating steps, the program first changes the probabilities and then changes one of the division points. Changing the probabilities is performed in three steps. First, the program chooses one of the elements of the probability matrix randomly. In the second step, it multiplies the probability with a random value from a uniform distribution on the [0.9, 1.1[interval, so the probability will not change more than 10%. For the element not in the main diagonal, the symmetric element needs to be multiplied as well. In the third step, the probability matrix is normalized.

To change the division points, the program adds zero and one to the list of the division points, so the division points will be $d_0, d_1, d_2, \ldots, d_{m-1}, d_m$, where d_0=0 and d_m=1.

Then the program chooses randomly one of the inner division points with the index i∈[1,m-1] and chooses a p random value from the uniform distribution on the [0,1[interval. The program relocates the chosen division point to $d_i + \Delta_\mathrm{i}(p)$, where

$$\Delta_\mathrm{i}(p) = \begin{cases} l\dfrac{\left(p-\frac{a}{l}\right)^n}{\left(1-\frac{a}{l}\right)^{n-1}}, \text{if}\, p > \frac{a}{l} \\ l\dfrac{\left(p-\frac{a}{l}\right)^n}{\left(-\frac{a}{l}\right)^{n-1}}, \text{otherwise} \end{cases} \tag{9}$$

Here, $= d_{i+1} - d_{i-1}$, and $a\ =\ d_i - d_{i-1}$ and $n > 1$ is an arbitrary exponent.

If the $\Delta_\mathrm{i}(p)$ function gives 0, the division point stays in the original place. With the increasing n exponent parameter, the $\Delta_\mathrm{i}(p)$ function is more likely to be close to zero, so the new division point is more likely to stay in the proximity of the original division point (Figure 2).

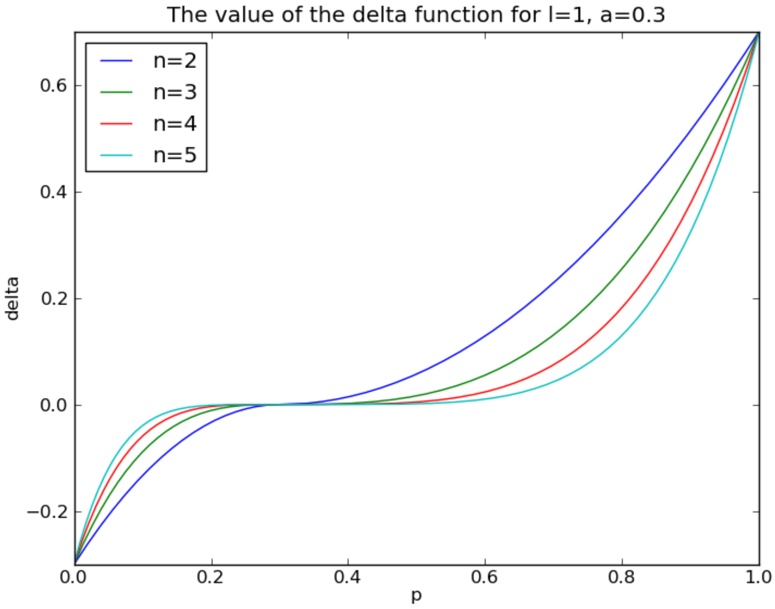


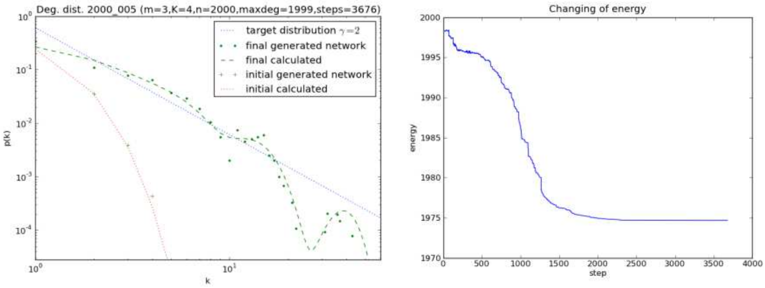
Figure 2

An example for the $\Delta_{\mathrm{i}}(p)$ function used for relocating a division point. This example uses one inner division point ($m = 2$) with the actual value 0.3, so the parameters are $l = 1$ and $a = 0.3$. The function was plotted with these parameters and with the exponents $n = 2,3,4,5$. The value of the function will be in the interval $[-a, l - a[= [-0.3, 0.7[$.

3.3 Generating a Network with Given Degree Distribution

Our program calculates the degree distribution as in equation (5), and in our measurements, it used the energy function in the equation (8).

The two property classes, *DegreeDistribution* and *DegreeDistributionC,* can calculate the degree distribution of a generating measure and can return with the energy of that distribution. In the first one, the calculation of the degree distribution has been written in Python using the *numpy* package. The second one uses C++ functions for that calculation. Each version uses the same Python function to calculate the energy from the degree distribution.



(a) (b)

Figure 3

Results of a generation. The target was a power-law function degree distribution with the exponent -2. In subfigure (a) the target degree distribution was drawn with a blue dotted line, the initial degree distribution calculated from the generating measure with a red dotted line, and the degree distribution calculated from the last accepted generating measure with a green dashed line. The degree distributions of networks generated from the initial and from the last accepted one are drawn with dots. In the subfigure (b) the energy as the function of the step number can be seen.

3.4 The Advantages of Python Programming Language

There are several reasons to use the Python language for the main program. Python itself is a dynamically-typed, object-oriented language with some useful complex data types (list, dictionary, set). These data types and the dynamically-typed property make possible a very flexible argument handling of functions with default argument values and keyword arguments. We frequently use two Python shells (*ipython* and *IDLE*) to run Python commands interactively. *IDLE* is part of most installations, but *ipython* has several useful extra abilities, like the interactive plotting of the functions with the *pylab* package.

The Python language has a huge standard library that can be reached in the standard installations on many operating systems, including Windows, Linux and MacOSX. We used the *shelf* package to store the generated results as well as the energies, the divisions and the probabilities of each step in a binary form.

Another advantage is the many useful free and high quality packages not included in the standard library. One of them is the numpy package that has its own data structures like *array* and *matrix*. An *array* is a sequence of elements of the same type. *Numpy* has mathematical functions like logarithm that can calculate the logarithm of each element of the array or matrix in one step. This calculation is quite fast, because the functions of numpy are written in the C programming language. The calculations can be slow if the calculations have too many steps at the Python level. For example, if we have more additions, subtractions, multiplications, divisions and functions calls, the Python must check whether the factors, the terms or arguments are arrays or not. These steps slow down the calculations.

The other useful package is the *pylab* package, which provides mathematical and plotting functions very similar to that of in MATLAB. This package is based on the *numpy* and *matplotlib* packages. The *pylab*, *numpy* and *matplotlib* packages are not part of the standard library.

To analyse the properties of the network belonging to the generating measure, the program uses the *igraph* complex network analyser package.

A big part of the code is covered with unit tests, which allows us to check easily the functioning our program in several environments (Python versions and operating systems), as the *unittest* package is also part of the standard library.

To identify the most CPU intensive parts of the program the *cProfile* module of the standard library can be used.

3.5 Numpy Version of the Program

The first version of the program used the numpy package of the Python programming language. With this package we can use arrays (row vectors), which can be manipulated more efficiently than Python lists. We used the cProfiler module to determine the parts of the program that needs too much time. We found that the iteration and the calculation of the estimated degree distribution belonging to the link probability measure were two such parts.

We ran the generation with 2000 steps and 2000 nodes. The time of the generation was 5392 seconds. The calculation of the degree distribution from the link probability needed 3625 seconds (67%), and the calculation of the link probability measure took 1748 seconds (32%). The calculation of the energy from the actual degree distribution and the other parts of the program took less than 1% of the running time.

3.6 The C++ Version of the Program

According to the running time measurements, the iteration of the generating measure and the calculation of the degree distribution were rewritten. The C++ program gets the generating measure from Python and writes the degree distribution to the standard output, where the Python collects information. In the future we plan to implement a more appropriate coupling between the C++ and the Python part of the program. We will wrap the C++ code with SWIG or CPython [2] to call the C++ functions easier.

3.7 The Comparison of the Running Times

We carried out a sequence of generations to compare the running time of the pure Python version using the numpy module and the version using C++ code. The program ran on a Debian Linux server installed as a VMware virtual machine.

The target degree distribution of the generations was a power-law degree distribution with the exponent -2. There were three parameters that changed: the number of steps, the number of iterations and the number of nodes in the network. More details about the network generator program can be found in the Appendix. The resultant running times are in Table 1 and are plotted on Figures 4 and 5.

Table 1

The running times of the two versions and ratio of the numpy version and the C++ version. The full running times are in minutes, but the running times of one step are in milliseconds.

		2000 steps						5000 steps					
		K=4		K=5		K=6		K=4		K=5		K=6	
# nodes	***type***	full	one step	full	one step	full	one step	full	one step	full	one step	full	one step
2000	numpy	23,37	701,01	89,88	2696,40	503,47	15104,10	58,35	700,18	224,40	2692,80	1259,86	15118,32
	C++	1,68	50,52	4,44	133,20	13,59	407,64	4,28	51,42	11,10	133,20	33,52	402,19
	ratio	**13,9**		**20,2**		**37,1**		**13,6**		**20,2**		**37,6**	
5000	numpy	53,62	1608,51	175,50	5265,00	759,10	22773,00	133,81	1605,73	439,85	5278,20	1905,34	22864,08
	C++	3,79	113,80	9,75	292,50	27,78	833,49	9,70	116,44	24,20	290,40	69,59	835,08
	ratio	**14,1**		**18,0**		**27,3**		**13,8**		**18,2**		**27,4**	
10000	numpy	102,72	3081,60	319,47	9584,10			257,36	3088,28	801,56	9618,72		
	C++	7,12	213,56	17,60	528,00			16,96	203,54	43,83	525,96		
	ratio	**14,4**		**18,2**				**15,2**		**18,3**			

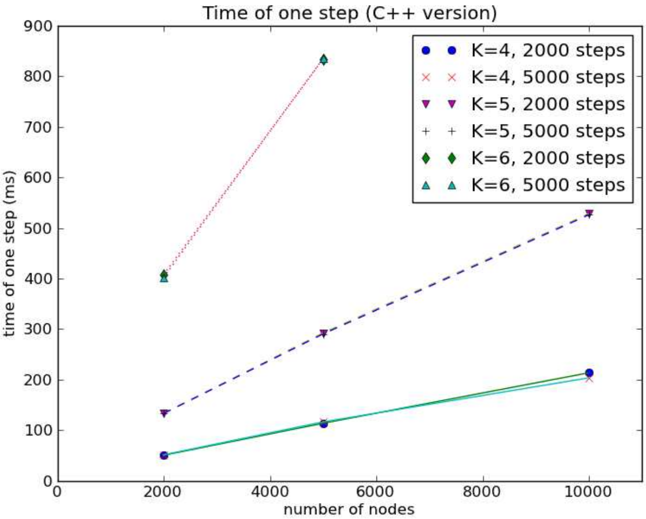


Figure 4

Running time of the C++ version of the generator as a function of number of nodes with a scale-free degree distribution as the target property. The number of iteration (*K*), number of nodes and number of steps have been varied.

This method of network generation makes it impossible for the network to have multiple edges between a node pair or to have self-loops (edges with the same nodes at its two endpoints), so the maximal degree cannot be bigger than the number of nodes in the network. During generation, the maximal degree in equation (8) was set to smaller by one than the number of the nodes, so with an increase in the number of nodes in the generation, the number of the terms in the sum and the running time increases, too. With an increase in the number of iterations, the running time increases fast, because the number of probabilities in the link probability measure increases exponentially with the number of iteration. For larger values of the number of iterations, the running times of the mfng version became too large (bigger times than one day can be found in Table 1), but with the C++ version this time is acceptable. The running time of mfng version compared to that of the C++ version is 13–38 times longer in these generations, and this ratio increases with the increasing number of iterations.

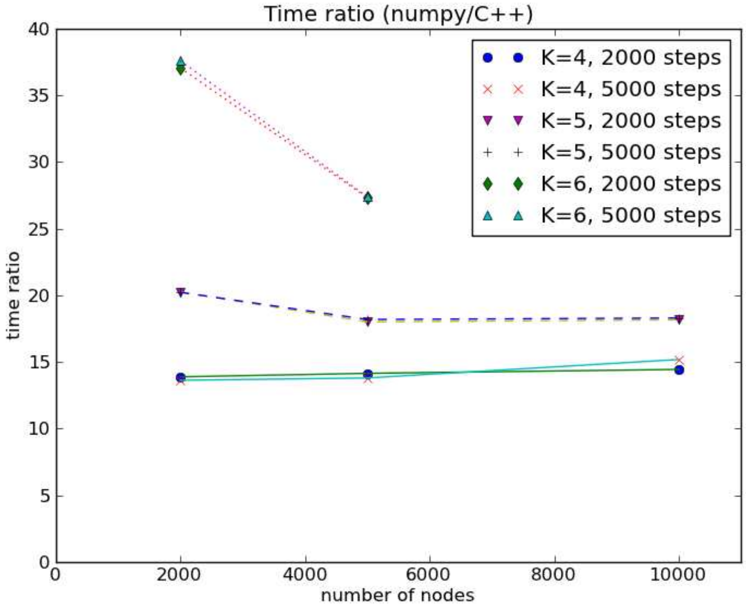


Figure 5

The ratio of the running times of the numpy version and the C++ version.

Conclusions

With the multifractal network generator (MFNG) method one can generate a wide range of networks with prescribed statistical properties. The method uses a mapping between the generator measures (a measure defined on the unit square) and the networks. It simulates an annealing process to get the optimal parameters of the generator measure. If one knows the generator measure, the degree distribution and some other statistical properties of the network can be calculated.

In our *mfng* program there are two realizations of the MFNG method. Our first realization of the MFNG method was written in Python using the *numpy* package. After rewriting some functions of the *mfng* program in C++ language, the running time of the program was reduced significantly, which allows for using a higher iteration number and more steps, so one can create generating measures that generate networks with properties closer to the target property.

Our program is part of the *cxnet* program framework intended to be used in education. As the C++ functions can be reached from Python, the easy-to-use Python framework would not necessarily be dropped. It makes it easier to use the framework in the education, especially if the students are familiar with the *cxnet* framework.

Acknowledgement

I would thank to Péter Orlik for his help in writing the C++ code. I would thank to Marianna Machata, Trócsányi Zoltán and József Lakner for their linguistic help and advice.

Appendix

In the running time comparison the program started the generation with a Python program as follows:

```
import mfng
for steps in [10000, 20000]:
    T0 = 0.2
    generator = mfng.Generator(T0=T0, steps=steps, Tlimit=T0/10000,
            m=3, K=4,
            n=2000,
            divexponent=7,
            project="power_law",
            )
    generator.append_property(
            mfng.DistributionFunction(
                "k**-2",
                maxdeg=n-1, mindeg=1
                )
            )
    generator.go()
```

The meaning of the program is as follows. First the program imports the functions and classes of the mfng module. It carries out two generations creating a generator in both generations. The temperature will decrease from 0.2 to 2×10^{-5} in 10000 and 20000 steps respectively. The generating measure in the generations would have 3×3 probabilities and it would be created for a network with n=2000 nodes. The changing of the division points will use the exponent 7 in equation (9). The result will be saved in the project_power_law directory. There is one target property with a distribution function proportional to the k^{-2} power-law function. The degree distributions in the generation will be compared to the target distribution from the degree 1 to 1999.

This version uses the numpy version to generate the generator function. If we slightly modify this program—we would add DistributionFunctionC (with C in the end) property to the generator instead of DistributionFunction—the generator runs the C++ version.

References

[1] Albert, R., Barabási A.: Statistical Mechanics of Complex Networks, Reviews of Modern Physics, Vol. 74, No. 1, 2002, pp. 47-97, doi:10.1103/RevModPhys.74.47

[2] Behnel, S., Bradshaw, R., Citro, C., Dalacin, L., Seljebotn, D. S., Smith, K.: Cython: The Best of Both Worlds, Computing in Science Engineering, Vol. 13, No. 2, 2011, pp. 31-39

[3] Csárdi, G., Nepusz, T.: The Igraph Software Package for Complex Network Research, InterJournal Complex Systems, 2006, Manuscript Number. 1695

[4] Cardanobile, S., Pernice, V., Deger, M., Rotter S.: Inferring General Relations between Network Characteristics from Specific Network Ensembles. PLoS ONE Vol. 7, Jun 2012, e37911, doi:10.1371/journal.pone.0037911

[5] Horváth, A., Trócsányi, Z.: Multifractal Network Generator with Igraph, Symposium on Applied Informatics and Related Areas, 2010

[6] Horváth, A., Trócsányi, Z.: Complex Networks in the Curriculum of the Computer Engineers, IEEE Proceedings of the 8th International Symposium on Applied Machine Intelligence and Informatics, 2010

[7] Horváth, A.: Studying Complex Networks with cxnet, Acta Physica Debrecina, Vol. XLIV, 2010, pp. 37-47

[8] M. E. J. Newman, The Structure and Function of Complex Networks, SIAM Review, Vol. 45, No. 2, 2003, pp. 167-256

[9] Palla, G., Lovász, L., Vicsek, T.: Multifractal Network Generator Proceeding of the National Academy of Sciences, Vol. 107, No. 17, Apr. 2010, pp. 7640-7645, doi:10.1073/pnas.0912983107

[10] Palla, G., Pollner, P., Vicsek, T.: Rotated Multifractal Network Generator, Journal of Statistical Mechanics: Theory and Experiment, Vol. 2011, No. 02, February 2011, P02003, doi:10.1088/1742-5468/2011/02/P02003

[11] Ravasz, E., Barabási A.: Hieararchical Organization in Complex Networks, Physical Review E, Vol. 67, No. 026112, Feb 2003

[12] The homepage of the cxnet program, http://django.arek.uni-obuda.hu/cxnet

[13] The repository of the mfng program, http://github.com/horvatha/mfng

What about Linear Logic in Computer Science?

Daniel Mihályi, Valerie Novitzká

Department of Computers and Informatics, Technical University of Košice
Košice, Slovakia
E-mail: Daniel.Mihalyi@tuke.sk, Valerie.Novitzka@tuke.sk

Abstract: In this paper we discuss several useful features of linear logic especially from the viewpoint of computing science. We start with a short overview of linear logic with an emphasis on the special properties of linear implication and exponential operators. We present our idea of the possible fragmentation of linear logic and the usefulness of particular fragments in various areas of computing science. Finally, we consider possible extensions of linear logic and we illustrate how an extension with epistemic operators can serve for obtaining knowledge and belief about an intrusion attempt.

Keywords: linear logic; type theory; behavioral theory; modal logics

1 Introduction

Linear logic was introduced by J. Y. Girard in 1987 [6] as a non-classical logic of actions and resources enabling one to describe dynamics of processes and resource handling. This logic can be considered as a suitable interface between logic and computing science because it can manipulate with the events of real world in natural way. Linear logic is a new logic, but the whole classical logic can be translated into linear formulae [3].

From the computing science's point of view, for intuitionistic fragment of linear logic the Curry-Howard correspondence [21] is valid, i.e. the formulae of linear logic correspond with the types of data structures. Similarly, the proof trees of the sequent calculus of linear logic correspond with programs [7]. If we consider formulae as resources, within the realization of a proof they are distributed in time and space [5] in some model of the real world, e.g. a computer machine in a precise and controlled manner. During our research, we have recognized several interesting properties and possibilities of linear logic:

- We have used an intuitionistic fragment of linear logic to formally describe program execution [17], [20], [22], [24];
- We have used linear logic to define linear type theory [15];

- In the sense of Curry-Howard correspondence, functional programming can be regarded as logical reasoning in linear logic. Linear proofs enable us to anticipate computability and correctness of computing [18], [19];
- Formulae are equivalent with some paterns of Petri nets [8], [9], [15];
- Extending linear logic with modal operators of necessity and possibility we have used modal linear logic for reasoning about the observable behavior of programs [13];
- Extending linear logic with epistemic operators of knowledge and belief we obtained epistemic linear logic useful for achieving experiences about incomming network intrusions based on a natural manner of causalities [13], [14].

The aim of this paper is to discuss several interesting features of linear logic and the possible applications of linear logic in several areas of computing science. We consider propositional linear logic. The second section contains a short introduction to linear logic with special emphasis on its modal operators and on the static and dynamic nature of linear implication. In the third section, we present our view of linear logic fragmentation that can serve for different purposes in various areas of computing science. In the fourth section, we show how classical logic can be expressed by linear logic. The fifth section shows the correspondence between linear logic and linear type theory. In the sixth section, we show how an extension of linear logic with epistemic modalities of knowledge and belief can provide useful information about the behaviour of programs.

2 Linear Logic Overview

In this section we introduce the basic notions of linear logic. Let *Props={p$_1$, p$_2$...}* be a countable set of atomic propositions denoted by the letters p_1,p_2.... Any proposition can be considered in two ways: as an action or as a resource. A linear formula φ can be of the form defined by the following BNF rule:

$$\varphi ::= p_n \mid 0 \mid 1 \mid \perp \mid \top \mid !\varphi \mid ?\varphi \mid \wedge\varphi \mid \vee\varphi \mid \varphi^{\perp} \mid \varphi_1 \otimes \varphi_2 \mid \varphi_1 \,\&\, \varphi_2 \mid \varphi_1 \oplus \varphi_2 \quad (1)$$

$$\mid \varphi_1 \wp \varphi_2 \mid \varphi_1 -\!\circ \varphi_2$$

Linear logic has two conjunction operators and two disjunction operators. We describe an informal meaning of linear connectives:

- *Linear implication* φ_1 *—o* φ_2 is causal; - it expresses that an action described by φ_1 is a cause of the (re)action described by φ_2. If we consider resources, a resource φ_1 is consumed after linear implication, i.e. it becomes a linear negation $\varphi_1^{\perp}$;

- *Multiplicative conjunction* (*"times"*) $\varphi_1 \otimes \varphi_2$ has the neutral element **1**. It expresses that both actions φ_1 and φ_2 will be performed simultaneously or that we have both resources φ_1 and φ_2 at once.
- *Additive conjunction* (*"with"*) φ_1 & φ_2 has the neutral element $\top$. It expresses that only one of the actions described by φ_1 and φ_2 will be performed. But we can deduce or anticipate from an environment which of them will be performed. This formula can be considered as an analogy with the statements *if-then-else* and *case* in programming languages. Somethimes it is called *external nondeterminism* (dependent choice);
- *Additive disjunction* (*"plus"*) $\varphi_1 \oplus \varphi_2$ has the neutral element **0**. It expresses that only one of the actions described by φ_1 and φ_2 will be performed (or only one of these resources is available), but we cannot anticipate which one. It can be considered *internal nondeterminism* (free choice);
- *Multiplicative disjunction* (*"par"*) $\varphi_1 \parr \varphi_2$ has the neutral element $\perp$ and its meaning can be expressed as follows: if an action φ_1 is not performed, then an action φ_2 is done or vice versa; if an action φ_2 is not performed, then an action φ_1 is done. Multiplicative disjunction can be regarded as an allegory of the well-known construct *xor* in programming;
- *Linear negation* $\varphi^{\perp}$ denotes a reaction of an action φ or a consumption of a resource φ. Linear negation is involutive, i.e.

$$\varphi^{\perp\perp} \equiv \varphi \tag{2}$$

2.2 Linear Exponentials

Another special property of linear logic represents two unary operators called exponentials. These operators can be considered from two points of view: concerning resources or concerning modalities. If we consider resources, then

- the operator "!" expresses potential resource inexhaustibility and
- the operator "?" expresses the actuality of potential resource inexhaustibility. .

Exponentials are dual, i.e.

$$(!\varphi)^{\perp} \equiv ?\left(\varphi^{\perp}\right) \tag{3}$$

Duality between exponentials can be considered as the difference between actual and potentional infinity [26]. The formula *(!φ)* expresses an unexhausted store of a resource φ and the formula $?(\varphi^{\perp})$ expresses potentional replenishment of exhausted resources. For instance, if we consider a resource φ to be a part of computer memory, we can indicate the potentional need to extend it. The resource character of exponentials are in the Table 1.

Table 1

Linear exponentials dealing with resources

Operator	Resource view	Modal view
!	unexhaustibility	*of course*
?	potential unexhaustibility (depending on actual replenishment)	*why not*

In terms of modality:

- the operator “!” (“*of course*”) expresses obviousity and
- the operator “?” (“*why not*”) expresses polemic.

Linear exponentials can be considered as linear alternatives of traditional modalities of necessity (“□”) and possibility (“◊”), respectively, as is shown in the Table 2.

Table 2

Modal nomenclature in linear logic manner

	∇_1	∇_2
Modal logic	◊	□
	Possibility	Necessity
Linear logic	?	!
	Polemic	Obviousity

Linear exponentials are necessary also for translating classical propositional logic into linear logic. We consider this translation in the Section 4. The exponential “*of course*” can also serve for expressing the repeating of some actions [14].

2.3 Static and Dynamic Nature of Implication

Classical logic has an obvious implication $\varphi_1 \Rightarrow \varphi_2$ with a static character. Ituitionistic propositional logic knows also weaker implication called partial implication $\varphi_1 \Rightarrow_p \varphi$ corresponding with linear partial functions under the Curry-Howard correspondence [4]. Both these implications can be translated to linear formulae $!\varphi_1 \multimap \varphi_2$ and $\varphi_1 \multimap ?\varphi_2$, respectively, thanks to exponentials, as we show in Table 3. Traditional linear implication $\varphi_1 \multimap \varphi_2$ has a dynamic character; its premise φ_1 is consumed after performing the linear implication. If we consider formulae as actions, we can say that an action φ_2 follows an action $\varphi_{1.}$ From Table 3 we can see that linear logic has more forms of implication, and so linear logic has greater expressive power. In addition, if we combine translated forms we can get a generalized form of linear implication $!\varphi_1 \multimap ?\varphi_2$ that can be particularly useful for programming languages with recursion [4].

Table 3
Forms of linear implications

Classical view	Linear view	Kinds of linearity
	$\varphi_1 \multimap \varphi_2$	Linear implication
$\varphi \Rightarrow \varphi$	$!\varphi_1 \multimap \varphi_2$	Unrestricted linear implication
$\varphi \Rightarrow_p \varphi$	$\varphi_1 \multimap ?\varphi_2$	Partial linear implication
	$!\varphi_1 \multimap ?\varphi_2$	Generalized linear implication

In linear logic we can choose whether we would like to work in static mode or in dynamic mode. If we translate classical implication $\varphi_1 \Rightarrow \varphi$ into $!\varphi_1 \multimap \varphi_2$, we work in static mode. We also note that classical implication $\varphi_1 \Rightarrow \varphi_2$ is equivalent with disjunction:

$$\varphi_1 \Rightarrow \varphi_2 \equiv \neg\varphi_1 \vee \varphi_2 \tag{4}$$

Translating the left and right parts of the previous formula into linear logic, we get the following equivalence:

$$\varphi_1 -\circ \varphi_2 \equiv \varphi_1^{\perp} \oplus \varphi_2 \tag{5}$$

but this is not valid by [6] because there exist two proof trees for the linear formula on the right side.

If we would like to consider dynamically, linear implication $\varphi_1 \multimap \varphi_2$ can be understood that an action φ_2 follows an action φ_1 , i.e. an action φ_2 starts after an action φ_1 . In contrast to the previous case, the following equivalence of linear formulae is valid:

$$\varphi_1 -\circ \varphi_2 \equiv \varphi_1^{\perp} \wp \varphi_2 \tag{6}$$

3 Linear Logic Fragmentation

Linear logic can be used as a whole, but in some cases it is appropriate to consider only a fragment of linear logic. In this section, we present an overview of how linear logic can be fragmented into several blocks according to the a nature of the particular fragments [12]. We illustrate our ideas of possible fragmentations in Figure 1.

First, we consider the vertical ellipses. The left one contains the multiplicative fragment of linear logic, and the right one contains the additive fragment of linear logic together with the corresponding constants. Linear implication and linear negation are neutral, they play important role in both fragments.

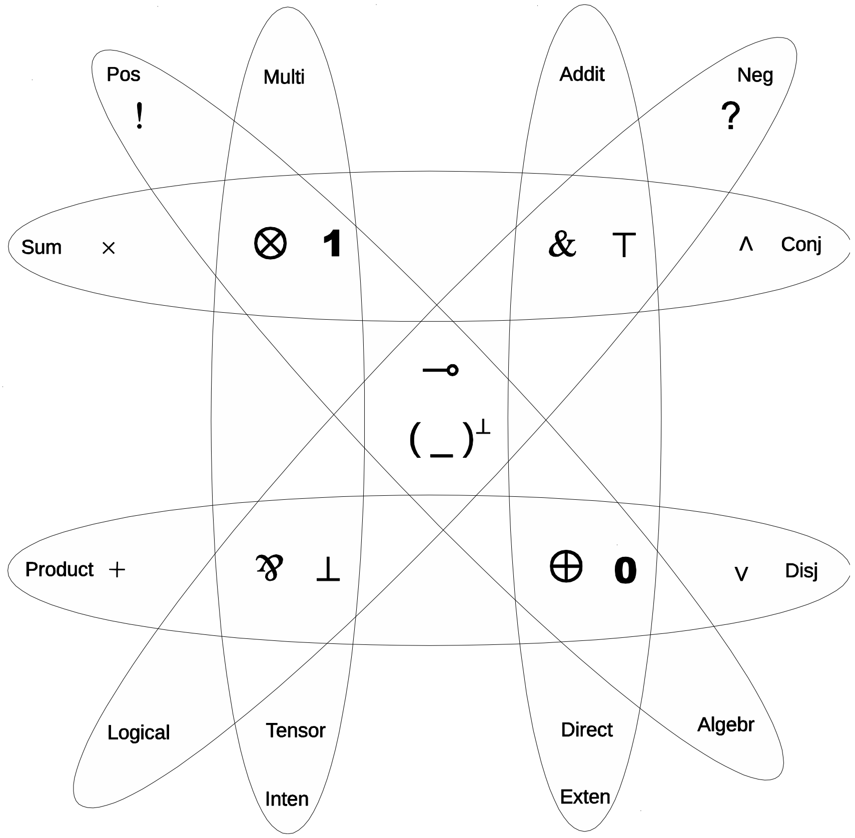


Figure 1
Linear logic fragmentation

From the semantic point of view we can consider the left ellipse as the intensional frangment and the right one as the extensional frangment. This arises from the semantical notions of extension and intension [2]. Whereas the extension of a given concept is its subject or the family of subjects included within it, the intension is the content of it. The extension of a given action is a truth value in the Tarski tradition; – the intension is an idea (sense) expressing it, – in the Heyting tradition. Extension we understand as a denotation and intension we identify as a sense. Traditionally, atomic propositions in (the Tarski tradition) are assertions that have exactly true or false truth values. In the extensional fragment of linear logic we assign to linear formulae the truth values (**1** or ⊥). But in the intensional fragment we consider their sense or nonsense (⊤ or **0**). For instance, if we have atomic proposition *Snowing*, it can be valid (**1**) and it has also sense (⊤). But the atomic proposition *Spowing* has no sense (⊥) and neither can it be valid (**0**). This also demonstrates the greather expressive power of linear logic, which is able to diferentiate between denotation and sense already at the syntactic level.

Vertical fragments play their role also in linear type theory. The left fragment contains the tensor product and sum while the right fragment contains the direct product and sum.

Now we consider the diagonal ellipses, which reflect another kind of fragmentation based on the idea of polarity. All logical connectives and neutral elements can be split into the following groups with:

1 Positive polarities: **0**, **1**, ⊗, ⊕,!;

2 Negative polarities: ⊥, T, &, ℘,?;

3 Dependent polarity: —o;

4 Turn over polarity: $(.)^{\perp}$;

These fragments can be considered from algebraic/logical point of view: the connectives with positive polarities correspond with the *algebraic style* and the connectives with negative polarities correspond with the *logical style*.

Linear negation causes the polarity to be turned over. This means that if an action is positive, its negation becomes negative, and vice versa. Linear implication is neutral again with respect to polarity; it causes the polarity of implication premise to be changed. An action (formula) of linear logic is positive if its outermost logical connective is positive; it is negative if its outermost logical connective is negative.

Finally, we consider the horizontal fragments of linear logic. If we work with the translation of propositional logic into linear logic, the upper fragment contains two linear conjunctions corresponding with classical conjunction and the lower fragment contains two linear disjunctions corresponding with classical disjunction. From the point of view of linear type theory we can regarded the upper fragment as the product type's constructors and the lower fragment as the sum type's constructors.

4 From Classical Logic to Linear Logic

As we mentioned above, linear logic can be considered as a generalization of classical logic. Every formula of classical logic can be unambiguously translated into linear formula. The static character of classical implication in linear logic ensures the exponential "!":

$$\varphi_1 \Rightarrow \varphi_2 \rightarrow !\varphi_1 - \circ \varphi_2 \tag{7}$$

Table 4 consists of the corresponding connectives for translating propositional logic into linear logic.

Table 4

Aristotelian logic to linear logic translation overview

CL to LL	∧	∨	⇒	¬	True	False
	&	⊕	—o	$(.)^{\perp}$	⊤	**0**

Aristotelian logic based on the Tarski semantical tradition can be translated into the fragment of linear logic in the sense of Table 4. In this fragment the linear additive conjunction (&) is a generalized classical conjunction (∧), the linear additive disjunction (⊕) is a generalization of classical the disjunction (∨), the linear implication (—o) is a generalized classical implication (⇒) and classical negation (¬) is expressed by linear negation ($(.)^{\perp}$). Aristotelian truth values, *True/False*, correspond to neutral elements, ⊤/ **0** of additive conjunction and disjunction, respectively.

Table 5

Intuitionistic logic to linear logic translation overview

CL to LL	∧	∨	⇒	¬
	&	⊕	—o	!_ —o 0

When we come out from the Heyting semantical tradition, such generalization leads to intuitionistic linear logic (Table 5). For example, intuitionistic formulae can be translated into linear formulae using the following equivalences:

$$\begin{aligned} &\varphi_1 \wedge \varphi_2 \equiv \varphi_1 \,\&\, \varphi_2 \\ &\varphi_1 \vee \varphi_2 \equiv \varphi_1 \oplus \varphi_2 \\ &\varphi_1 \Rightarrow \varphi_2 \equiv !\varphi_1 -\!\circ \varphi_2 \end{aligned} \tag{8}$$

5 From Linear Logic to Linear Type Theory

Due to the Curry-Howard correspondence between intuitionistic linear logic and type theory [1], any formula φ of linear logic can be interpreted as a linear type denoted e.g. by *A*. Using linear connectives we can formulate a linear type theory in the sense of Table 6. According to the selected fragment of linear logic we can work with tensor fragment and/or direct fragment.

Table 6

Type theory nomenclature in Linear logic manner

Linear type theory			***Linear logic***	
Tfrag	Tensor product	⊗	Multiplicative conjunction	Mfrag
	Tensor sum	℘	Multiplicative disjunction	

Linear type theory			***Linear logic***	
Dfrag	Direct product	&	Additive conjunction	Afrag
	Direct sum	⊕	Additive disjunction	

Every programming language has a collection of predefined types. These types can be considered as basic types forming a set *Btypes={X, Y,...}*. Let *I* be a unit type. We can construct linear Church's types over basic types and unit type using type operators corresponding with linear logic connectives. Then the syntax of the linear types can be defined as:

$$A ::= I \mid X \mid A_1 \otimes A_2 \mid A_1 \oplus A_2 \mid A_1 - \circ A_2 \mid A_1 \,\&\, A_2 \mid A_1 \wp A_2 \tag{9}$$

In this grammar, *I* denotes a linear unit type and *X* denotes a linear basic type. The following constructions are linear Church's types:

- $A_1 \otimes A_2$ is product linear type;
- $A_1 \oplus A_2$ is coproduct (sum) linear type, and
- $A_1 \multimap A_2$ is function linear type as a set of functions from type A_1 to a type A_2.

Binary product/coproduct linear types can be generalized to

- Finite product linear types of the form $A_1 \otimes A_2 \otimes \ldots \otimes A_n$ together with the projections π_i*:* $A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n \rightarrow A_i$*, i=1,...,n*;
- Coproduct linear types of the form $A_1 \oplus A_2 \oplus \ldots \oplus A_n$ together with the coprojections (injections) κ_i*:* $A_i \rightarrow A_1 \wp A_2 \wp \ldots \wp A_n$*, i=1,...,n*.

Correspondence between traditional type theory and linear type theory is shown in Table 7. In linear type theory, any variable can appear in a term only once [1]. Product types (⊗) together with projections (&) are illustrated in the upper horizontal ellipse in Figure 1. Coproduct types (⊗) together with coprojections (℘) are illustrated in the lower horizontal ellipse in Figure 1.

Table 7

Traditional and linear Type theory

Type manipulation	**Type theory**		
	Traditional	***Linear***	***Comment***
Product	×	⊗	constructor
		&	selector
Coproduct	+	⊕	deconstructor
		℘	integrator
Function type	→	—o	constructor

6 Behavior, Knowledge and Belief

The expressive power of linear logic can be increased by various extensions, for instance with modal operators. If we consider the vertical fragments in Figure 1, we can construct various modal extensions of linear logic that enable additional useful applications in computing science.

Firstly, consider the intensional fragment of linear logic (left vertical ellipse). If we extend this fragment with modal operators for necessity and possibility (◊, □), we achieve a new logical system constructed over coalgebra [10], [11] as a resource oriented modal coalgebraic linear logic suitable for describing the observable behavior of running programs [12].

Now we consider the extensional fragment of linear logic (right vertical ellipse). We extend this fragment by epistemic objective knowledge operator K and by epistemic rational belief operator B. Assuming an agent c a formula $K_c\varphi$ expresses that an agent c has a knowledge φ and a formula $B_c\varphi$ expresses that an agent c has belief about φ . This fusion between epistemic and linear logic we have used to construct a Kripke model for acquiring knowledge and empirical belief about incoming network intrusions [12], [13] based on [25]. We shortly describe the main ideas of our approach. We use the following extensional fragment of linear logic extended with epistemic operators:

$$\varphi ::= p_n \mid \varphi_1 \,\&\, \varphi_2 \mid \varphi_1 \oplus \varphi_2 \mid \varphi_1 - \circ \varphi_2 \mid \varphi^{\perp} \mid !\varphi \mid K_c\varphi \mid B_c\varphi \tag{10}$$

Assume a signature based Intrusion Detection System and three possible types of intrusions: *A, B* and *C.* Every type of intrusion attempt has several symptoms that can be described as elementary propositions.

Let *007* be an rational agent, e.g. some program. Let the symptoms of an intrusion attempt of a type *A* be denoted by elementary propositions a_1, a_2, a_3 and a_4. The symptoms of an intrusion of a type *B* are b_1, b_2 and b_3 and the ones of a type *C* are c_1, c_2 and c_3. An intrusion attempt of a particular type occurs only if all its symptoms have occurred. Using additive conjunction we can describe the knowledge about all mentioned types of intrusion attempts by the following formulae:

$$\begin{aligned} K_{007}\varphi &\equiv K_{007}a_1 \,\&\, K_{007}a_2 \,\&\, K_{007}a_3 \,\&\, K_{007}a_4 \\ K_{007}\psi &\equiv K_{007}b_1 \,\&\, K_{007}b_2 \,\&\, K_{007}b_3 \\ K_{007}\theta &\equiv K_{007}c_1 \,\&\, K_{007}c_2 \,\&\, K_{007}c_3 \end{aligned} \tag{11}$$

Let $K_{007}\tau$ be a formula describing the knowledge about a sender, e.g. its IP address. A formula

$$\left(\underbrace{K_{007}\varphi \,\& \ldots \&\, K_{007}\varphi}_{300}\right) \& \, K_{007}\tau \tag{12}$$

describes that we have the knowledge that an intrusion attempt of type *A* occurred three hundred times from the same sender. The following epistemic linear formula

$$K_{007}\chi \equiv \left(\underbrace{K_{007}\varphi \,\& \ldots \&\, K_{007}\varphi}_{300}\right) \& \, K_{007}\tau \multimap \left(\left(K_{007}\psi \,\&\, K_{007}\theta\right) \& \, K_{007}\tau\right) \tag{13}$$

expresses the situation when after three hundred attempts of type *A* the attempts of types *B* and *C* follow immediately. The same situation exists in real IDS, e.g. vertical portscan [23]. If this situation repeats, we can state that our agent *007* has achieved a rational belief about the intrusion attempt expressed by the formula

$$!\left(K_{007}\chi\right) \multimap B_{007}\chi \tag{14}$$

and we can realise some protective actions. Exponential *!* enable us to describe the repetition of attempts, i.e. a real behavior of intrusions by the principle "*Repetitio est mater studiorum*".

In [16] we explained our approach in detail together with a construction of a Kripke model and a definition of the semantics of our epistemic linear logic.

Conclusions

In our paper we presented a few inventions regarding possible areas of applying linear logic in various disciplines of computing science. We considered several criteria for the fragmentation of linear logic and we discussed the known applications of these fragments in type theory and behavioral theory. We also discussed the special properties of linear connectives and exponentials. The static and dynamic properties of linear logic we illustrated in various forms of linear implication. Classical and intuitionistic logic can be translated into linear logic using exponentials. The dynamic character of linear logic enables it to definine linear type theory. Linear logic can be extended by new operators, e.g. modal operators, epistemic operators, etc. These extensions allow for increasing of the expressive power of linear logic and for opening new application domains. The modal intensional fragment of linear logic can be useful for describing the observable behavior of programs, and the epistemic extensional fragment enables us to obtain knowledge and beliefs about intrusion attempts.

The dynamic/static resource oriented character of linear logic destines it for wide usage in computing science. In this paper, we presented only a few possible applications of it based mainly on our research results. We believe that the presented inventions can lead to the discovery of further applications in computing science.

Acknowledgement

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 "Modelling, simulation and implementation of GPGPU-enabled architectures of high-throughput network security tools."

This work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF.

References

[1] Ambler S. J.: First Order Linear Logic in Symmetric Monoidal Closed Categories, PhD. Thesis, University of Edinburgh, 1991

[2] Avron A.: The Semantics and Proof Theory of Linear Logic, Theoretical Computer Science, Vol. 57, 1988, pp. 161-184

[3] Braüner T.: Introduction to Linear Logic, BRICS LS-96-6, Aarhus, 1996

[4] Chang, E. B.-Y., Chaudhuri, K., Pfenning, F.: A Judgmental Analysis of Linear Logic, Carnegie Mellon University, Report CMU-CS-03-131R, 2003

[5] Girard, J.-Y. From foundations to ludics. Bulletin of Symbolic Logic 9, 2 (2003), 131-168

[6] Girard J.-Y.: Linear Logic, Theoretical Computer Science, Vol. 50, 1987, pp. 1-102

[7] Girard J.-Y.: P. Taylor, Y. Lafont, Proofs and Types, Cambridge University Press, New York, NY, USA, 1989

[8] Korečko Š, Sobota B.: Using Coloured Petri Nets for Design of Parallel Raytracing Environment, Acta Universitatis Sapientiae. Vol. 2, No. 1, 2010, pp. 28-39

[9] Korečko Š, Sobota B., Szabó Cs.: Performance Analysis of Processes by Automated Simulation of Coloured Petri Nets, Intelligent Systems Design and Applications: Proceedings of the 10th international conference: 29 Nov.-1 Dec. 2010, Cairo, Egypt

[10] Kurz A.: Coalgebras and Modal Logic, CWI, Amsterdam, Netherlands, 2001

[11] Mihályi D.: Duality Between Formal Description of Program Construction and Program Behaviour, Information Sciences and Technologies Bulletin of the ACM Slovakia, Vol. 1, No. 2, 2010, pp. 1-5

[12] Mihályi D., Jenčík M.: Few Inventions about Utilising Linear Logic in Computer Science, ICTIC 2012 - Information and Communication Technologies – International Conference, Žilina, 19. – 23. 3. 2012, 2012

[13] Mihályi D., Novitzká V., Ľaľová M.: Intrusion Detection System Epistème, Central European Journal of Computer Science, Vol. 2, No. 3, 2012, pp. 214-221

[14] Mihályi D., Novitzká V., Ľaľová M.: Intrusion Detection System Epistème, Proceedings of the International Scientific Conference Informatics'2011, Rožňava, 16.-18.11.2011, Košice, Equilibria, 2011, 11., pp. 61-65

[15] Mihályi D., Novitzká V., Slodičák V.: From Petri Nets to Linear Logic, CSE'2008, Fifth International Scientific Conference on Electronic Computers and Informatics, Vysoké Tatry - Stará Lesná, 24. - 26. 9. 2008, Košice, 2008, pp. 48-56

[16] Mihályi D., Novitzká V.: Towards to the Knowledge in Coalgebraic Model of IDS, Computing and Informatics, 2012 (accepted)

[17] Novitzká V., Mihályi D., Slodičák V.: Categorical Models of Logical Systems in the Mathematical Theory of Programming, Journal of Pure Mathematics and Applications, 17, 3-4, 2006, pp. 367-378

[18] Novitzká V., Mihályi D., Slodičák V.: How to Combine Church's and Linear Types, ECI'2006 - Seventh International Scientific Conference on Electronic Computers and Informatics, Košice - Herľany, 20.-22.9.2006, Košice, 2006, pp. 128-133

[19] Novitzká V., Mihályi D., Slodičák V.: Linear Logical Reasoning on Programming, Acta Electrotechnica et Informatica, Vol. 6, No. 3, 2006, pp. 34-39

[20] Novitzká V.: Logical Reasoning about Programming of Mathematical Machines, Acta Electrotechnica et Informatica, Vol. 3, No. 3, 2005, pp. 50-55

[21] Sørensen M. H., Urzyczyn P.: Lectures on the Curry-Howard isomorphism, DIKU Rapport 98/14, 1998

[22] Slodičák, V.: Some Useful Structures for Categorical Approach for Program Behavior, Journal of Information and Organizational Sciences, Vol. 35, No. 1, 2011, pp. 99-109

[23] Snort web site. Availaible on: http://www.snort.org

[24] Szabó Cs., Slodičák V.: Software Engineering Tasks Instrumentation by Category Theory, Proceedings of the 9th IEEE International Symposium on Applied Machine Intelligence and Informatics SAMI 2011, 27-29.1.2011, Košice, Elfa s.r.o., 2011, pp. 195-199

[25] Vokorokos L. Baláž A.: Distributed Detection System of Security Intrusions Based on Partially Ordered Events and Patterns, Towards Intelligent Engineering and Information Technology, Studies in Computational Intelligence, Vol. 243, Springer, 2009, pp. 389-403

[26] Zlatoš P.: Ani matematika si nemôže byť istá sama sebou, Iris, Bratislava, 1995

Development Stages of Intelligent Parking Information Systems for Trucks

Zsolt Péter Sándor, Csaba Csiszár

Budapest University of Technology and Economics (BME),
Faculty of Transportation Engineering and Vehicle Engineering,
Department of Transport Technology and Transport Economics
Műegyetem rakpart 3, H-1111 Budapest, Hungary
E-mail: sandorzs@kku.bme.hu, csiszar@kku.bme.hu

Abstract: The route planning of trucks also includes the planning of the parking. Research has already shown that parking demands often exceed capacities. This problem can be managed by appropriate information provision. It is also discussed by European strategic documents, the ITS Directive and the ITS Action Plan, as a priority assignment. Advanced intelligent parking management systems for trucks, planned and partially installed ones, provide real-time information and central (system optimum based) navigation with automatic parking place booking. In this way, the capacity utilization of parking facilities can be maximized. Nowadays such complex systems are still not available, only some part-functions, but modern information and communication technology offers new opportunities in transportation applications. The article classifies the parking information systems into five plus one (5+1) service levels according to their functions. On the highest service level, the information system offers individual route plans for every user while it takes drivers' working hours, actual traffic conditions and personal preferences into consideration. In the future, mobility and parking demands can be influenced by real-time and interactive information management.

Keywords: intelligent truck parking; integrated information system; parking management; navigation; intelligent transport systems and services

1 Introduction

The route planning of trucks includes also planning for parking. As a result, detailed route plans of vehicles are determined according to their departure place and time, as well as to the destination and the scheduled time of arrival. Route plans contain scheduled stops, too. Intermediate halts are specified by European rules and regulations [8]. Truck drivers spend this mandatory time at parking facilities located along the motorways.

Parking demand usually exceeds capacity. The aim of the intelligent parking information system (which is a part of ITS services, i.e., intelligent transport systems) is to maximize the capacity utilization of the limited available parking spaces by coordination of parking demands and capacities. The services of complex information systems go beyond parking management, and they realize network management functions, too.

2 Intelligent Parking Management for Trucks

Intelligent parking management systems are new not only in Hungary but also in Europe, and they have high priority [1], [3]. Management services combine:

1. Information provision about parking facilities and free capacity,
2. Complex route guidance and
3. Preliminary parking place booking.

Services are available pre-trip and on-trip through different communication channels. Systems can provide individual (personalized) and collective information.

In this way, it is possible to avoid utilization over 100% caused by irregular parking when a few kms away facilities are unexploited in space and time. This phenomena may have several causes [6], [7], [10], [11]:

- Truck drivers do not have any information about actual occupancy, services and locations of the next few facilities;
- Some of the parking places do not meet the basic requirements (safety, security, sanitary facilities and restaurants, etc.);
- Experience shows that in the case of no free parking places, drivers would rather park in a dangerous, not designated area of an already full truck park than pass on and break the rules of the obligatory rest time.

Preliminary parking space booking supports the work of truck drivers and dispatchers of haulier companies.

Effects of the service [2], [4], [5]:

- **Safety**: parking in dangerous and not designated areas (due to congestion) can be avoided. Drivers can have rest in time, taking rules and regulations into consideration. Secure parking places protect the cargo against vandalism and theft.
- **Network efficiency**: Capacity utilization can be maximized and unnecessary traffic searching for parking places can be decreased. With the use of real-time traffic data, route guidance becomes more efficient and more comfortable. Thus, congestions can be avoided.

- **Environmental impact**: As a consequence of the above, air pollution can be mitigated. With optimal capacity exploitation there is no need for new infrastructure.
- **Decrease in operational damages**: The amount of infrastructure damage can be decreased.

In regions and/or cities where mobility management is well-developed, operators of parking facilities exploit the potential of smartphones and the internet. They can do it independently or jointly. Websites and smartphone applications (route planning and information provision applications) have been developed. With these applications, users can select facilities, they can browse static and dynamic information, and they can pre-book parking lots.

The penetration of smartphones provides a new information provision platform for system developers and service providers, especially in the matter of individual information. Data exchange protocols of on-board navigation systems provide the opportunity to transmit parking information data through RDS-TMC[1] and TPEG[2]. Pre-booking of parking places has already been available in some pilot projects and it can be presumed that new projects will also be launched in the near future.

Safe and secure parking does not belong closely to the core domain of parking management systems. However, this topic is also a high priority because 60% of the road transport attacks occur in unguarded parking places. The value of these damages is about 8 billion Euros per year [12].

3 Levels of the Parking Management Systems

Parking management systems can be classified by their functions and operations. In the following, a classification containing 5+1 categories (service levels) is been presented. Service levels – according to solutions with different information provision and different intervention functions – are built up modularly. The lower level services are supplemented by additional services at the higher levels. Fig. 1 illustrates the structure of levels. Level "0" can be separated from the dynamic parking management systems.

[1] **R**adio **D**ata **S**ystems - **T**raffic **M**essage **C**hannel: technology for delivering traffic and travel information to motor vehicle drivers

[2] **T**ransport **P**rotocol **E**xperts **G**roup – They developed the TPEG specifications for transmission of language independent multi-modal Traffic and Travel Information. TPEG data are human understandable as well as machine readable.

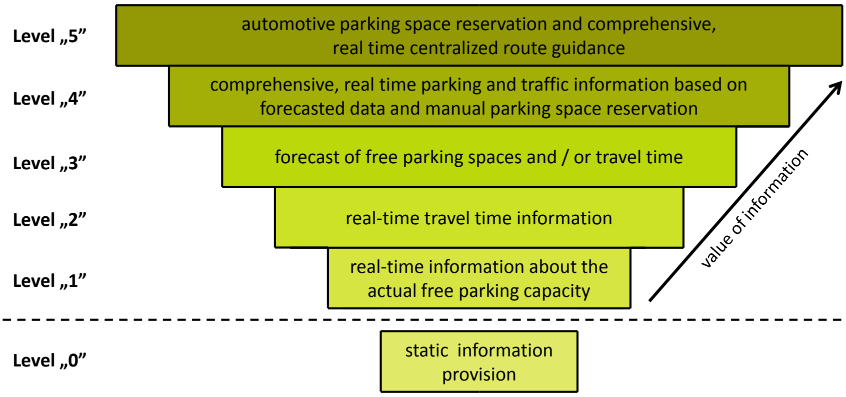

Level „5"
automotive parking space reservation and comprehensive, real time centralized route guidance
Level „4"
comprehensive, real time parking and traffic information based on forecasted data and manual parking space reservation
Level „3"
forecast of free parking spaces and / or travel time
Level „2"
real-time travel time information
Level „1"
real-time information about the actual free parking capacity
Level „0"
static information provision
value of information

Figure 1
Classification of service levels

Table 1 illustrates the functions and features of each service level based on leading European examples. Currently, the applied parking and information management systems can be categorized into Level "0" to Level "4".

As a consequence of new technologies, the management of traffic data has become much more integrated. The roles of traffic management centres are being revaluated and their functions extended. Central, interactive (two-way communication), integrated, real-time navigation solutions have increasingly come increasingly into praxis. These new solutions enable system optimum based traffic control, which provides smooth and maximized capacity utilization of parking facilities, while also taking users preferences into consideration.

The growing penetration of mobile communication and mobile internet is resulting in the development of new applications that can be used in the transport sector. Thus, the route planning procedures of haulier companies are changing.

The features of the certain service levels have been summarized. Figures illustrate the top level functions.

Table 1

Service levels of parking management systems (functions and features)

		Level "0"	Level "1"	Level "2"	Level "3"	Level "4"	Level "5"
Information	**Static / Dynamic**	static	dynamic				
	Content	location, capacity, services, operator of parking facilities, etc.	free capacities	actual travel times	expected travel times and / or expected free capacity	comprehensive, real-time parking and traffic information based on forecasted data	comprehensive, real-time centralized route guidance
	Communication channel	leaflets, flyers, brochures, handbooks, internet, static road signs, etc.	roadside variable message signs, internet based applications (one way communication)		smartphone applications (one way communication)	smartphone application (two way communication)	user or vehicle device (two way communication)
	Covered area	not specific	following one or more facilities		motorway sections or networks		
Operational features	**Update frequency**	rarely	event controlled	sampling time cycle is 5-10 min for travel times	in every 5-10 minutes for forecast	real-time even bookings	real-time
	Forecast horizon	not available			app. 15-60 minutes		even 24 hours
	Location-based services	not available			based on cellular network		satellite
	Route planning	not available				previous + route guidance software	centrally, based on user preferences
	Booking	not available				manually	automatically with route planning
Example		**Printed**: IRU Truck parking areas handbook **Web**: IRU TransPARK www.iru.org	**On the spot**: French A13 motorway **Web**: Rheinland-Pfalz verkehr.rlp.de	**On the spot**: Germany, A3 motorway (Frankfurt) **Web**: Traffic Scotland trafficscotland.org	**Web**: Bayern Info www.bayerninfo.de	Germany, Highway Park www.highway-park.de	this service is not yet available

4 Structure of the Parking Management System

The main components (elements, subsystems) of a complex, integrated parking management system are:

- **Users**: truck drivers and/or haulier companies with their own preferences, parking booking demands; with personal devices that support two-way communication. This equipment can be applied as on-board units as well, e.g., *navigation devices, smart phones, palmtops, laptops, (even PCs)*;
- **Trucks**: may have on-board computers that provide two-way communication and positioning. OBUs[3] can replace personal devices.
- **Motorway networks**: with human (*road operator dispatcher*) and machine components. The latter ones include: automatic traffic and environment detectors (*traffic counting stations, CCTVs*[4] *with license plate recognition, weather stations, air pollution sensors, etc.*), roadside information devices (*VMSs*[5]*, displays*) and communication equipment.
- **Parking facilities**: with human (*dispatcher, security guard*) and machine components. The latter ones include: devices for data acquisition [*occupancy detectors, ultrasonic sensors, vehicle identification and surveillance systems, etc*.], devices for local data procession, devices for information provision [*VMSs, displays, road markers, etc*.] and communication equipment. Image processing and identification equipments (*CCTV cameras, licence plate recognition*), access control systems. In certain cases the human component may be omitted, for example when the facility can operate without direct human supervision.
- **Parking management centre**: contains human (*operator*) and machine components. It processes the incoming data and determines the information transmitted to devices for information provision. One centre can supervise one region or even one country. Centres are located close to other traffic management centres or may even be a part of them.

[3] **o**n-**b**oard **u**nit

[4] Closed-circuit television

[5] **V**ariable **M**essage **S**ign

Information terminals:

- **Mobile devices**: personal and/or vehicle on-board devices with their own operating systems (iOS, Android, Symbian, etc.), with multi-channel communication units (GPS, GSM, Wi-Fi, etc.), and with route planning and navigation programs that provide personalized, real-time information and parking place booking.
- **Immobile devices**: road traffic signs, variable message signs, controllable displays, LED markers at parking lots (lighting equipment or controllable prisms built in the road surface).

The listed elements are connected in the indicated relations by telecommunication channels (*wired or wireless data networks*). A continuous line indicates wired communication, and broken lines indicate wireless communication. The structure of the system is illustrated in Figure 2. The figure corresponds with the build-up of the highest service level. In the case of the lower levels, certain components are to be omitted. Only one element of each component is represented.

5 Operation of the Parking Management System

The most important objectives of parking management and route planning are [2], [4], [12]:

- the optimization of the capacity utilization of parking places,
- low time consumption (for parking place searching, entrance and exit), reliable, calculable and predictable travel time,
- the decrease of unnecessary travel distances and emissions, and the mitigation of noise pollution,
- a decrease in stress, an increase in user comfort,
- a decrease in the risk of accidents,
- the protection of vehicles,
- the maximizing of driving times while considering the generic rules,
- the minimizing of travel costs (overrun costs, facility usage fees, etc.).

The operational processes of the telematics system are summarized in Figure 3, which is coherent with the structural architecture. Arrows indicate the direction of data transfers. The figure corresponds with the operation of the highest service level. At lower levels, certain functions are not available; however, at higher levels the management system provides network management functions as well. Table 2 contains a description of certain operations. Level "4" and level "5" realize closed-loop control because real-time information provision (parking place booking and navigation) feeds back to the users' pre-trip decisions and thus influences demand.

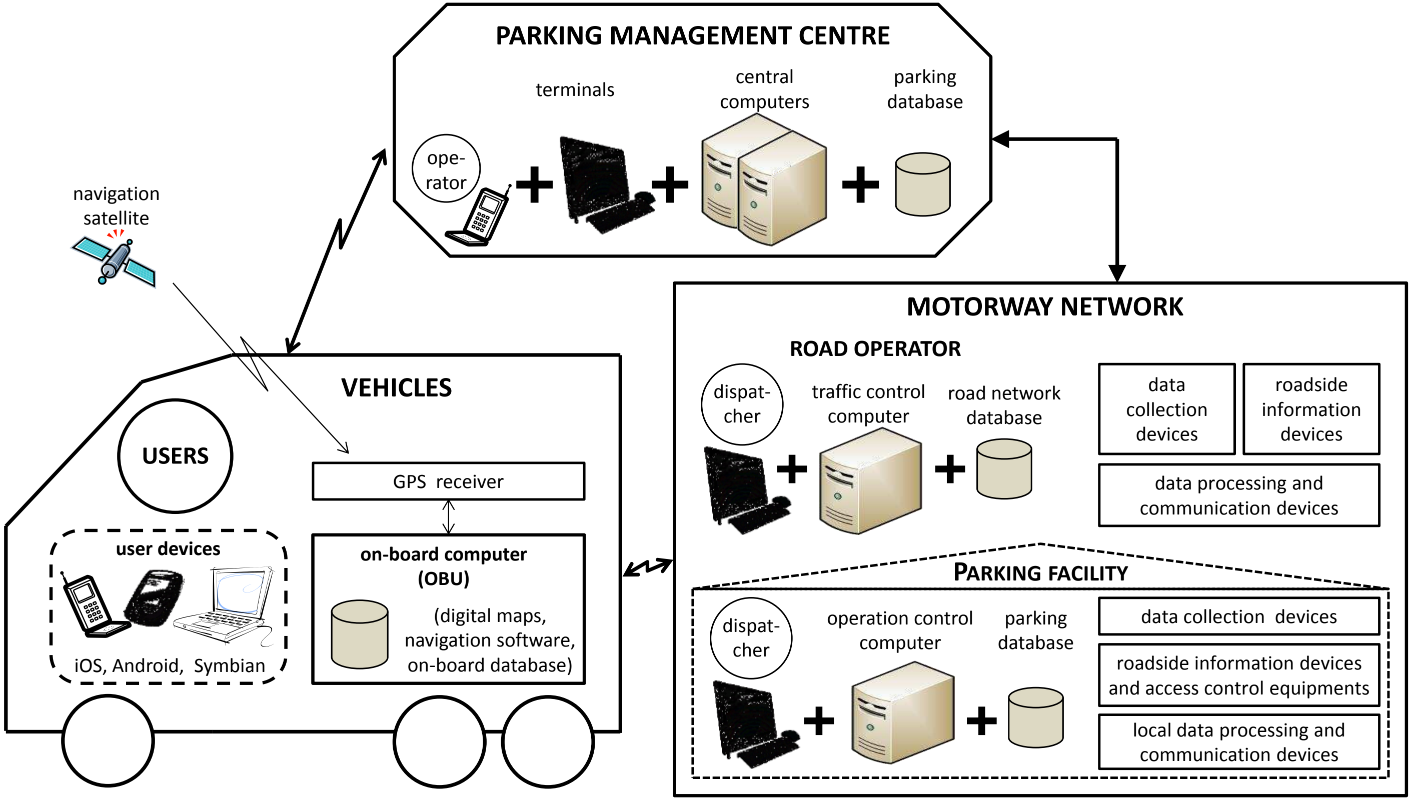


Figure 2

Structure of the integrated parking management system

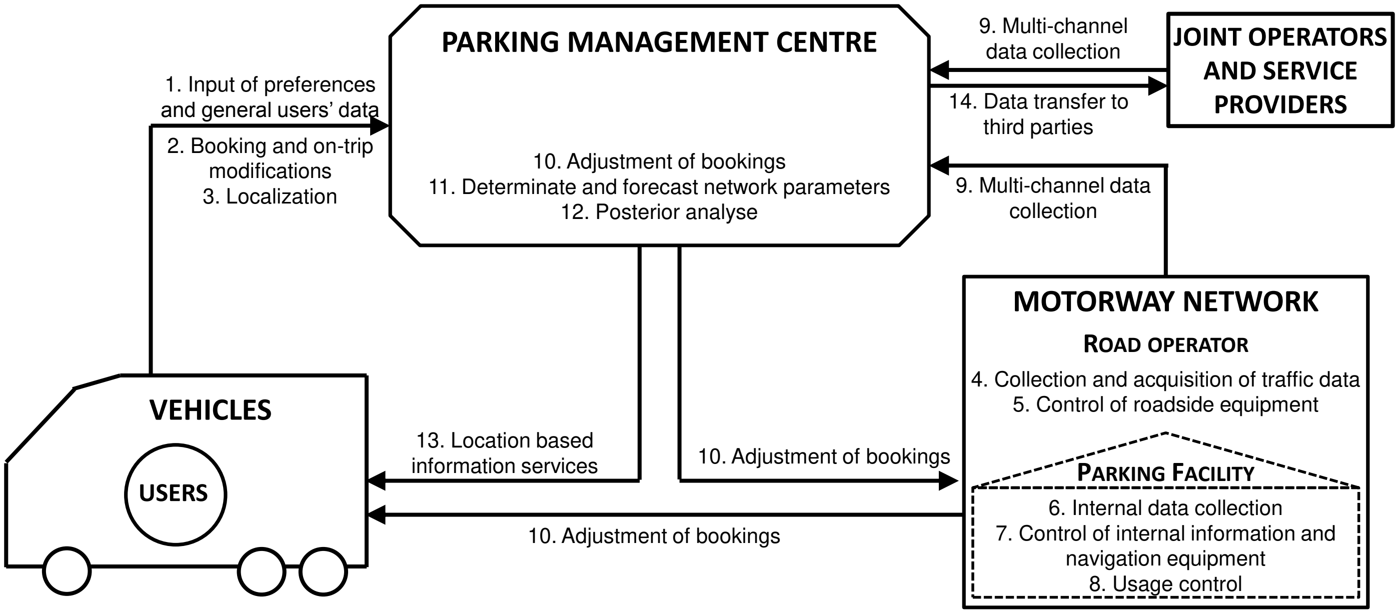

PARKING MANAGEMENT CENTRE
10. Adjustment of bookings
11. Determinate and forecast network parameters
12. Posterior analyse
1. Input of preferences and general users' data
2. Booking and on-trip modifications
3. Localization
9. Multi-channel data collection
14. Data transfer to third parties
JOINT OPERATORS AND SERVICE PROVIDERS
9. Multi-channel data collection
MOTORWAY NETWORK
ROAD OPERATOR
4. Collection and acquisition of traffic data
5. Control of roadside equipment
PARKING FACILITY
6. Internal data collection
7. Control of internal information and navigation equipment
8. Usage control
VEHICLES
USERS
13. Location based information services
10. Adjustment of bookings
10. Adjustment of bookings

Figure 3

Operation model of the integrated parking management system

Every user has their own user account at the parking management centre. Before route planning, they log in and record every route specific preference in connection with the given journey. During route planning, the incoming demands are assigned to the parking facilities, taking personal preferences and actual traffic parameters into consideration (*departure time, route, driving times, stop-points, services for sale*). The users are in contact with the management centre, and they transmit all such data to it. Booking records are confirmed by the parking facility and the management centre as well, so it is the only case when users get information in addition to that from the management centre. The route and the bookings can be modified based on actual conditions and the vehicle position.

Vehicles can be identified by their licence plates at the facilities. The data is also used to control facility usage, and it is possible to control driving times with or without the tachograph.

There is hierarchical data storage with duplicated/multiplied data elements. Data are stored in the management centre and also in the facilities. Facilities receive only the parking-related information about the users. Road operators do not have any route-specific information about the users and vehicles. Parking facilities are controlled by the parking management centre, but they are can operate autonomously. The management centre and some special parking facilities can be operated with non-stop human supervision.

During parking place selection, drivers take the following factors with decreasing importance into consideration: [9]

- Security;
- Previous own experiences;
- Additional services of the facility (*e.g. sanitary facilities, catering, grocery shopping, petrol station, ATM, etc.*);
- The number of available parking places.

Personalized parking place booking and intelligent parking management with dynamic navigation are value-added services. Thus, service providers may impose service fees and sanctions to motivate to the proper use (vehicle arrives in time, cancel bookings in time, etc.). The incurred expenses are paid on the spot or based on contracted agreements.

Table 2

Parking-related information management operations

groups of operation		no.	name	description
USERS' OPERATION	Pre-trip	1	Input of preferences and general users' data	Input of data regarding customers, vehicles and personal preferences. Record static data regarding route and route-specific personal preferences.
	On-trip	2	Booking and on-trip modifications	Route planning and booking/parking place selection according to personal preferences and actual traffic situation. Modifications during journeys.
		3	Localization	Transmitting current vehicle position to the information management centre.
OPERATIONS OF ROAD (MOTORWAY) OPERATOR		4	Collection and acquisition of traffic data	Measurement of traffic, weather and air pollution parameters.
		5	Control of roadside equipment	Control of VMSs; display actual and forecasted travel times and capacities.
OPERATIONS OF PARKING FACILITY		6	Internal data collection	Acquirement of actual occupancy data, identification of entering and exiting vehicles.
		7	Control of internal equipment	Operation of facility information and navigation displays, VMSs and lighting equipment built in the road surface.
		8	Usage control	Monitoring users and operation of safety & security equipment (e.g. access control and image processing systems).
OPERATIONS OF PARKING MANAGEMENT CENTRE	Data collection	9	Multi-channel data collection	Data come from several automatic data collector elements (from traffic counting stations, weather stations, occupancy detectors, etc.). Data from joint operators and service providers are also incoming data.
	Data processing	10	Adjustment of bookings	Procession of incoming travel and parking demands; assignment of free parking lots to booking demands.
		11	Determinate and forecast network parameters	Determinate actual network parameters based on incoming traffic parameters (free parking capacity, travel times). Determinate expectable travel times and free parking capacity based on actual traffic parameters and historical data. (Forecast horizon depends on the applied algorithm.)
		12	Posterior analyse	Analysis and evaluation of users' behaviour and operations; creation of statistics and traffic predictions.
	Information provision	13	Location based information services	Real-time route guidance and navigation for vehicles with combination of individual and system optimum (*dynamic modification of released route plan and automatic rebooking*), emergency management. Control of user and vehicle (OBU) equipment.
		14	Data transfer to third parties	Data sharing and transfer to joint operators and service providers.

The main task of the parking management system is to redirect the exceeding demands automatically to previous or – when driving time allows it – following facilities by taking driving times and user preferences into consideration. When it is not possible, demands can be realigned by changing departure times. The management system can modify only pre-planned parking demands. It does not have any effect on ad hoc users and those who do not have a booking. They can be influenced by roadside occupancy information. The service offers secondary, complementary guidance for trucks via the transmission of traffic and parking related information (Dispatchers from the transport company give the primary guidance via the completion of route plans).

At the top level of service, complex traffic management centres will be established, where all traffic, parking and transportation related information are available. Management centres can serve comprehensive traffic management and information provision tasks as well based on the available data. The traffic management system can handle emergencies (it alerts the emergency services, coordinates the rescue, provides real-time traffic control and management, etc.), provides real-time information for vehicle drivers, and if necessary – after user acceptance –makes changes automatically related to alternative routes.

6 Implementation Possibilities

The opportunities for such a complex information provision system can be fully realized when services are developed in an interoperable way, allowing the latter network connection for further operators. The use of a common data structure (e.g. DATEX II[6]) supports the former aims and the harmonized information provision, as well as cross border operation. The partners of the EasyWay consortium – established by the EU – are currently developing their own parking information services, but these project are only in the pilot phase and the questions of cross border data transmission are still open [15].

It is an obvious solution to provide information via on-board equipment that is already built-in. Devices of satellite-based tracking (GPS, GNSS[7]) electronic fee toll collection systems are well-suited for these solutions that can realize location based services and further information provision services in an integrated manner.

In Hungary, the parking management system would have the greatest impact on the 4th corridor of the Trans-European Transport Network (M1-M5 motorways) where transit and freight transport is greatest. Capacity utilization of the parking areas would be better. Parking needs can be allocated according to supply and

[6] **DAT**a **EX**change – improves data exchange between countries and organisations

[7] **G**lobal **N**avigation **S**atellite **S**ystem

demand. Nowadays in Hungary the implementation of a satellite-based electronic fee toll collection system is quite current. During the system design, experts should endeavour to create and realize a complex toll collection system that is appropriate for other functions as well, like the dissemination of traffic and parking related information.

Conclusion

The route planning of trucks includes also the planning of parking, and parking management must be a part of traffic management. The intelligent parking management system for trucks is a rather new solution for traffic control and information provision services, as part of intelligent transport systems. The feasibility and the benefit of this service has been analysed in several studies and pilot projects. With advanced information provision, capacity utilization can be improved and, at the same time, congestions can be avoided. All of these require investments mainly in the information infrastructure, not in the parking facilities themselves. The information system supports the better capacity utilization by demand and capacity assignment. As every ITS application, this service also allows for better traffic circulation without significant infrastructure investment by detailed information services. Parking management systems allow for minimizing traffic searching for parking places, and the services contribute to more efficient and effective route planning and driving.

The main contribution of the article is the creation of a model including 5+1 service levels. It outlines the future incremental developments. For the top service level, detailed structural and operational models have been elaborated. The presented integrated information system provides comprehensive, central navigation that supports every participant of the transport and haulier sector. The drivers' work becomes safer and more comfortable, while the supply chain becomes more predictable.

References

[1] COM (2008) 886 *Action Plan for the Deployment of Intelligent Transport Systems in Europe*. Brussels 2008

[2] Core European ITS Services and Actions: *Guideline for the deployment of Intelligent Truck Parking*, 2010

[3] Directive 2010/40/EU of the European Parliament and of the Council*, on the Framework for the Deployment of Intelligent Transport Systems in the Field of Road Transport and for Interfaces with other Modes of Transport*, Brussels, 2010

[4] Freight & Logistics Services: *Intelligent Truck Parking and Secure Truck Parking*. Deployment guideline, FLS-DG01, Version 01-02-00, Jan. 2012

[5] Gongjun, Yan: SmartParking - *A Secure and Intelligent Parking System*. Intelligent Transportation Systems Magazine, Vol. 3, Issue 1, pp. 18-30, 2011

[6] *Real-Time Dynamic Information Services 2009, Action Plan for the Real-Time Dynamic Information System on Motorway M1*, EasyWay Project co-financed by TEN-T, Project Number: 2007-EU-50010-P

[7] *Real-Time Dynamic Information Services/Development of the Monitoring System of Parking Areas, Occupancy Monitoring, Suggestions, Navigation, etc / for the Heavy Goods Transport on the Main Road Corridors,* Feasibility Study, EasyWay Project co-financed by TEN-T, Project Number: 2007-EU-50010-P

[8] Regulation EC No 561/2006 of the European Parliament and of the Council*, on the Harmonisation of Certain Social Legislation Relating to Road Transport*, Brussels, 2006

[9] Sándor, Zsolt - Nagy, Enikő: *Development Possibilities of the Intelligent Truck Parking System in Hungary.* TDK paper, BME, Department of Transport Technology, Budapest, 2011

[10] Sándor, Zsolt - Nagy, Enikő: *Foreign Solution for the Area of Intelligent Truck Parking and National Application Possibility*. Conference edition IFFK-2011(Paper 42), ISBN 978-963-88875-3-5 and 978-963-88875-2-8

[11] Sándor, Zsolt - Nagy, Enikő*: Intelligent Truck Parking on the Hungarian Motorway Network.* Pollack Periodica, Vol. 7, No. 2, August 2012

[12] *Secured Truck Parking in Europe*. ITP Workshop, Montabaur, 07.-08.11.2010

[13] Szabó, Péter: *Observing Possibilities of the Rules of Driving Times on the Hungarian Motorway Network and Necessary Developments*. Thesis work, BME, Department of Transport Technology, Budapest, 2010

[14] Freight & Logistics Services: *Intelligent Truck Parking and Secure Truck Parking*. Deployment guideline, FLS-DG01, Version 02-00-00, Dec. 2012

Estimation of Recycling Capacity of Multi-storey Building Structures Using Artificial Neural Networks

Vladimir Mučenski[1], Milan Trivunić[1], Goran Ćirović[2], Igor Peško[1], Jasmina Dražić[1]

[1] University of Novi Sad, Faculty of Technical Sciences, Department of Civil Engineering and Geodesy, Trg Dositeja Obradovića 6, 21000 Novi Sad, Republic of Serbia, mucenskiv@uns.ac.rs, trule@uns.ac.rs, igorbp@uns.ac.rs, dramina@uns.ac.rs

[3] Belgrade University, College of Applied Studies in Civil Engineering and Geodesy, Department of Civil Engineering, Hajduk Stankova 2, 11000 Belgrade Republic of Serbia, cirovic@sezampro.rs

Abstract: In recent years, we are witnessing a greater tendency towards the use of existing construction waste, in order to reduce the amount of material being disposed of on the one hand, and to limit the exploitation of natural resources necessary for the production of construction materials on the other hand. This paper provides an outline of a process for predicting the recyclable amount of concrete and reinforcement built in structures of residential buildings based on artificial neural networks (ANN). The following analyses are included in the process: an analysis of the optimal network structure, analysis of the effect of training algorithms and a network sensitivity analysis. While analyzing these, networks with one and two hidden layers trained with 5 algorithms (Gradient descent with adaptive lr backpropagation, Levenberg-Marquardt backpropagation, quasi-Newton backpropagation, Bayesian regularization and Powell-Beale conjugate gradient backpropagation) for neural network training were observed. The research was carried out with the purpose of observing ANN that will quickly and with adequate precision provide information regarding the amounts of concrete and reinforcement that can be recycled.

Keywords: Recycling; concrete; reinforcement; prediction of the quantity; artificial neural network; training algorithm; sensitivity analysis

1 Introduction

An increase in the use of recycled materials has become an imperative in the process of environmental protection. The need to create a sustainable production system within which the exploitation of natural resources and the amounts of waste materials will be reduced to a minimum has been present in construction industry for years.

The issue of estimating recycling capacities of building constructions is of crucial importance for establishing financially justifiable recycling processes and for the re-use of construction materials. In the Republic of Serbia, 22,272,500 m^2 of flats are older than 65 years, and 74,053,973 m^2 are older than 40 years, which presents a significant recycling potential when it comes to construction materials. In order to estimate the amount of concrete and reinforcement to be recycled in relation to the characteristics of a building as accurately as possible, it is necessary to analyze several parameters which describe the building, with parameters of both quantitative and qualitative nature. In such cases, statistical methods do not provide sufficiently accurate results.

With the development of software for solving mathematical problems, but also with the development of a totally new concept of programming and calculation within the same, known as "soft computing", the opportunity became available for using particular mathematical concepts, the realization of which, up to that point, had not been possible for very complex problems.

One of these concepts is that of artificial neural networks, which attempts to simulate the working of the human brain in order to solve particular mathematical problems. As the amount of research using neural networks has increased, so has the number of attempts to apply them in the construction industry, on the basis of which it can be confirmed that their use is more than justified.

The aim of this paper is to develop such a model for the estimation of the amount of recyclable concrete and reinforcement, one which does not require the use of project documentation or data which cannot be collected by visual examination of a building. The reason for this lies in a lack of projects for a large number of building constructions within the archive.

For this reason, research was carried out into the use of artificial neural networks for predicting the amount of recyclable concrete and reinforcement built in the skeletal structure of residential buildings. The prediction of the amount of materials was done on the basis of a database formed for the purposes of this research; the database included 9 parameters: the (complexity of the building, the total gross area of the building, the average gross floor area, the height of the building, the number of stiffening walls, the longitudinal and transverse raster of the construction, the type of floor structure and the type of floor support structure. All are available or can be easily defined based on project documentation. The output values of the database are the amounts of concrete and reinforcement.

In this paper, a brief overview of the concept of artificial neural networks is given, as well as an overview of the current situation regarding their application in solving a given problem. Additionally, a detailed methodology is given for the implementation of research into predicting the amount of materials required, which includes: the process of forming a database, the process of finding the optimal network architecture, the process of finding an optimal training algorithm and the process of sensitivity analysis of the network on the input data.

2 The Fundamentals of Artificial Neural Networks

The basic logical scheme for the imitation of the biological nervous system was formed by McCulloch and Pitts, who defined the mathematical principles which enabled the formation of artificial neural networks [1]. The principle of artificial neural networks is based on the attempt to imitate the biological nervous system in which artificial neural networks support the recognition of particular regularities and their memorization. In addition, gradual learning is possible during their application, i.e. adapting already established rules within the given network. Because of their flexibility in finding dependence, artificial neural networks are suitable for analyzing problems for which there is no clear describable mathematical dependence. There are many definitions of artificial neural networks. Hajkin defines them as huge distributed parallel processors [2]; Zurada regards them as physical cellular systems which can learn, memorize and use experimental knowledge [3]; and Nigrin defines them as systems which consist of a large number of simple elements for processing information [4].

If one wants to form a mathematical model of a biological neuron, particular respect must be paid to its structure. Dendrites, the body of the neuron and the axon must be formed. Figure 1 shows the McCulloch-Pitts general model of a mathematical neuron, the so-called M-P neuron, which has been used for the purpose of this research [5]. The weighted input section of the neuron represents the dendrites. In the body of the neuron, the summing of the signal occurs, on the basis of which the neuron is activated or not. If activation occurs, a signal is sent via the output (axon) to the neurons to which it is connected.

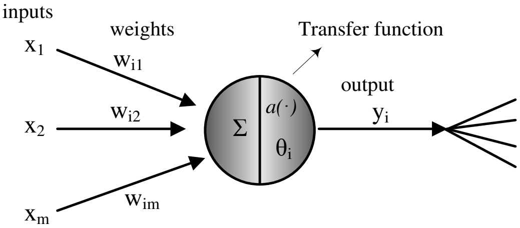


Figure 1
The McCulloch-Pitts model of an artificial neuron

Neurons are interconnected, forming a network where they can be arranged either within a single or several layers. The number of neurons can vary from one layer to another. The acquired network structure has a great impact on the speed and quality of neural network training. The process of finding the optimal structure is iterative.

In addition to defining the network structure and the number of neurons, it is necessary to choose the manner of network training. Network training is based upon finding regularities by the network based on a sufficient amount of data given in order to determine interdependence. The result of a network training process is the values of weight coefficients which describe links between neurons, as was explained earlier, and/or the change in the neural network structure. Considering all of this, we can distinguish between two basic types of networks [6]:

- fixed networks, in which the weights cannot be changed, i.e. dW/dt=0. In such networks, the weights are fixed a priori according to the problem to solved and
- adaptive networks, which are able to change their weights, i.e. dW/dt≠0. The weights change according to adopted learning rules, such as the Hebbian learning rule, the correlation learning rule, the instar learning rule, winner takes all, the outstar learning rule, the Widrow-Hoff LMS learning rule, linear regresion and the delta learning rule.

The neural network learning process is carried out by using learning algorithms. All learning algorithms used for adaptive neural networks can be classified into two major categories:

- supervised learning, which incorporates an external teacher, so that each output unit is told what its desired response to input signals ought to be, and
- unsupervised learning, which uses no external teacher and is based upon only local information. It is also referred to as self-organisation, in the sense that it self-organises data presented to the network and detects their emergent collective properties.

Combining and summing of the input data and the weight coefficient occur in the input of the neuron. Output of the neuron is defined by the transfer function "a" which activates or prevents activation of the observed neuron depending on the output value of the function. This function typically falls into one of three categories: linear (or ramp), threshold or sigmoid.

For our research purposes, a supervised adaptive network and two types of transfer functions were used:

- hyperbolic tangent sigmoid function

$$a(f) = \frac{2}{1+e^{-\lambda f}} - 1 \tag{1}$$

- and linear function.

$$a(f) = f \tag{2}$$

3 Review of Relevant Literature

Research up to now has been based on the use of both statistical models and ANN when establishing dependence between parameters which describe a building construction and the building costs. Kim et al. [7] compared the efficiency of regression analysis and ANN when faced with the problem of predicting building costs, and concluded that ANNs offer a more precise estimate of the required data. Wang and Gibson [8] carried out similar research by comparing the effectiveness of applying neural networks with regression analysis when predicting the success of a construction project in which one of the parameters of success was the project budget. Gunaydin and Dogan [9] analyzed the application of ANN in estimating the cost of building an RC construction in which costs were defined per m^2 of the area of the building. The analysis was conducted on the basis of 30 projects completed in Turkey, and the average error was 7%.

In addition, a large number of analyses were carried out using hybrid ANN models and fuzzy logic and/or genetic algorithms. Yu and Skibniewski [10] formed a hybrid neuro-fuzzy model for finding the optimal technology for constructing buildings based on technologicity. Kim et al. [11] analyzed the use of ANN with optimization of the same using genetic algorithms during the estimation of construction costs. For this they used a database containing information on 530 residential buildings. The result of using a hybrid model was that 80% of the data for validating the network was found in a error interval of up to 5%. Cheng et al. [12] analyzed the application of fuzzy ANN for predicting conceptual construction costs, whereby they evaluated the significance of 47 building parameters. The research resulted in an average error estimate of 5.9%. On the other hand, Cheng et al. [13] formed the Evolutionary Web-based Conceptual Cost Estimator (EWCCE), a hybrid model including WWW, genetic algorithms, neural networks and fuzzy logic with the purpose of estimating construction costs in the early stages of a project. The accuracy of the model was greater than 75%. It should be noted that research dealing with the prediction of the quantities of materials that could be recycled is rare.

4 Data Collection and Database Creation

The quality of a neural network depends on the amount and quality of the data on the basis of which the neural network is trained. For this reason, for the purposes of this research, a database was established containing information from major residential building construction projects in Novi Sad, Republic of Serbia. The data was randomly divided into two groups, namely: a data set for training the neural network (95 projects) and a data set for evaluating the quality of the analyzed network (15 projects). The parameters chosen for describing the characteristics of the structure are shown in Table 1 and include the geometric and structural characteristics of the building.

Table 1
Contains the result of comparing in pairs with the final result

Type	Building parameter	Definition	Parameter type	Interval or parameter definition
Input data	x_1	Complexity of the building	Discrete	Simple (1), medium (2), complex (3), very complex (4)
	x_2	Total gross area	Numeric	$1000m^2$ - $8000m^2$
	x_3	Average gross floor area	Numeric	$200m^2$ - $2000m^2$
	x_4	Building height	Numeric	13m – 27m
	x_5	Number of stiffening walls	Numeric	0 – 13
	x_6	Longitudinal raster	Discrete	1.00m-1.99m (1); 2.00m-2.99m (2); 3.00m-3.99m (3); 4.00m-4.99m (4); 5.00m-5.99m (5); 6.00m-6.99m (6); 7.00m-7.99m (7)
	x_7	Transverse raster	Discrete	
	x_8	Type of floor structure	Discrete	Full RC slab (1), Semi-prefabricated ceiling type "FERT" (2)
	x_9	Type of supporting floor structure	Discrete	Direct support(1), girder support (2)
Output data	y_1	Quantity of concrete	Numeric	$420m^3$ - $4500\ m^3$
	y_2	Quantity of reinforcement	Numeric	28500kg – 310000kg

It should be noted that all the analyzed buildings have base slab support. In addition to the above, the database also includes buildings with one or without any dilation since this is the case in over 95% of residential buildings in the analyzed area.

The complexity of the building was adopted because of attempts to define the influence of the building's characteristics from the aspect of the complexity of the construction and the shape of the building on the output values. Included in simple buildings are buildings with a rectangular base and are without any changes in the construction of the floor. Medium complex buildings are characterized by particular changes in the construction of the floor or an approximately rectangular base with fewer deviations (L base). In the category of complex buildings are those with an indented base (Π base, H base and so on), while very complex buildings are characterized by an indented base and/or atypical changes in the floor construction such as a reduction in the floor area with a growth in its height and an atypical shape of the skeletal construction.

The total area of the building is a parameter which is expected to have a great influence on the quantity of materials. Here, the gross area was taken into account to make prediction easier.

Including the floor area is an additional attempt to establish a correlation between the shape, i.e. the footprint of the building and the output values. In doing this, the gross floor area was also adopted.

The definition adopted for the height of the building was the distance from the ground surface to the highest point of the building.

Given that the seismic resistance of the building (and with that the amount of concrete and reinforcement) above all depends on the number and the distribution of stiffening walls, the influence of the same on the output values was adopted. The size of the longitudinal and transverse raster of structure has a direct influence on the span of the girders and the ceiling. The interval range and adopted parameters are shown in Table 1.

During data collection, two types of floor structure and two types of supporting floor structure were dominant within the project, as shown in Table 1. Table 2 shows the segments of the data base on the basis of which training and validation of the ANN were carried out.

Table 2

Segments of the input and output data sets for the training and validation of neural networks

Input data									**Output data**	
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	y_1	y_2
3	2000	250	23	5	5	3	1	1	800	86000
2	4700	790	17	6	4	5	1	1	1870	148000
2	4350	830	18	8	5	5	1	1	2100	152000
2	2250	340	22	5	3	3	2	2	1050	66500

When considering the limitations of the database, it should be noted that it includes only buildings which have a concrete skeletal system. In addition to the above, the limits of the established base represent the minimum and maximum data values on the basis of which the ANN is trained. That is, when carrying out an estimate of the amount of materials for a new building, the parameter values must be found within the interval of the data collected so that it is possible to analyze only the options with the analyzed types of base structure, floor structure and the support of the same.

Taking into account that there is a clear distinction in the order of magnitude of the data from 0 to 10^5 (see Table 2), it is necessary to prepare the data in order for it all to be analyzed equally, i.e. it is necessary to carry out normalization of the data. Normalization of the data leads to an increase in the performance of the trained ANN [9]. Based on the above, normalization of the whole database was carried out, i.e. the input and output data both in the training set and in the testing set for the ANN. Normalization of the data was performed using “Z-Score” [14] transformation in the distribution where the mean is (μ) 0, and the standard deviation (σ) 1 using the following expression:

$$S_{ij} = \frac{X_{ij} - \mu_i}{\sigma_i} \tag{3}$$

where: S_{ij} – is the normalized data value

X_{ij} – is the actual data value

μ_i – is the mean distribution (data set for training)

σ_i – is the standard deviation of the distribution (data set for training)

i – is the input ($i = i_1, i_2,..., i_9$) or output ($i = o_1, o_2$) data

j – is the number of combinations (j = 1, 2,…, n); n – is the number of data sets.

Of course, given that training of the ANN is carried out based on normalized data, the output from the ANN is also normalized. It is necessary to transform the output in order to obtain real values comparable with the expected values from the data sets for testing the ANN on the basis of which error is determined [14]. Transformation is performed using the following expression:

$$X^{real}{}_{ij} = S^{NN}{}_{ij} \bullet \sigma_i + \mu_i \tag{4}$$

where: $S^{NN}{}_{ij}$ – is the normalized data value obtained as output from the ANN

$X^{real}{}_{ij}$ – is the real data value obtained on the basis of $S^{NN}{}_{ij}$

μ_i – is the mean distribution (data set for training)

σ_i – is the standard deviation of the distribution (data set for training)

i – is the output data ($i = o_1, o_2$)

j – is the number of combinations (j = 1, 2,..., m); m – is the number of data sets for testing.

5 ANN Analyses for Predicting the Quantity of Recycling Material

In this part of the study we present a detailed overview of the process of using neural networks for predicting the quantity of recyclable concrete and reinforcement. The process is presented for finding the optimal network architecture and optimal algorithm for training the network, along with a sensitivity analysis of the network on the input data used for training the network.

5.1 Modelling the ANN

The type and the structure of a neural network have a significant effect on the quality and efficiency of the network. ANNs consist of neurons in layers, the number of which depends primarily on the problem being solved by the network. Based on research, a conclusion was drawn that networks with a small number of neurons (in relation to the optimal number of neurons for the given problem) offer solutions with a rough approximation, i.e. with large deviations. On the other hand, too large a number of neurons gives too precise an approximation taking into account the small deviation when searching for dependence [15]. In addition to the number of neurons, the manner of grouping them in so-called layers of neurons is also significant. Therefore, we distinguish between single-layered and multi-layered networks.

In the phase of defining the ANN model, the required amount of input and output data is defined first. The number of input parameters determines the spatial dimensions of the network and the number of output parameters determines the number of solution surfaces generated by the network [16, 17]. The amount of input data is defined by the formation of a database based on the analysis carried out as to the significance of individual data, while the amount of output data is defined by the amount of information required as the end result of the application of a neural network.

An important characteristic of neural networks, in addition to the number of neurons and layers, is the method of data processing, i.e. the transfer flow of information between neurons. Therefore, we distinguish between forward oriented networks (the transfer of information takes place in one direction, forwards) and networks oriented backwards (the transfer of information takes place in both directions, forwards and backwards). In these, the networks mostly used are those with back propagation of errors where the signal is transmitted forwards, while the error is transmitted backwards in order to minimize it, and the whole process is repeated until the error reaches its minimum [9, 11, 18-21]. In view of the above, an algorithm with error backpropagation was used for the purposes of this research.

The process of finding the optimal structure of neural networks in its essence is examining different structures on the basis of the same set of data where the investigation involves varying the number of layers and the number of neurons in the network. The process of determining the quality of the network structure in relation to the given problem is based on determining the size of the error which is obtained as an output result after the network training process. Neural networks were applied to the given problem using the Matlab R2007b software package, on which the analysis of the structure, the training of the network and the simulation of the working of the same were carried out. The network type was not varied, since it was shown that networks with error backpropagation were optimal for the problem of prediction [7-9, 11, 19].

In the research presented in this paper, different training functions were used (see Table 3) in order to see their effect on the ANN modelling.

Table 3
Used training functions

Training Functions	
traingda	Gradient descent with adaptive lr backpropagation.
trainlm	Levenberg-Marquardt backpropagation.
trainbfg	BFGS quasi-Newton backpropagation.
trainbr	Bayesian regularization.
traincgb	Powell-Beale conjugate gradient backpropagation

In addition to defining the network type, it is necessary to define the number of layers and the number of neurons along with the transfer function in the neurons, which define the method of data transmission between the same.

For the observed network, the hyperbolic tangent sigmoid transfer function was used in the hidden layers except for in the output layers of the network for which a linear transfer function was used.

The research was carried out on two types of neural networks. The first type included networks with one hidden layer, whereas the second consisted of networks with two hidden layers. For each type, 4 versions of a neural network with different number and ordering of neurons in layers were analyzed (see Table 4). Previous research has shown that networks more complex than networks presented in table 4 are unstable in prediction and less accurate [22]. In Figure 2 some of the analyzed networks are shown. Figure 2a shows the network 2-2, which contains one hidden layer with two neurons, and the output layer, which also contains two neurons considering 2 outputs, whereas Figure 2c presents a network with three neurons in its hidden layer. Figure 2c shows the network 2-2-2 containing two hidden layers each with two neurons, as well as the output layer containing two neurons considering 2 outputs.

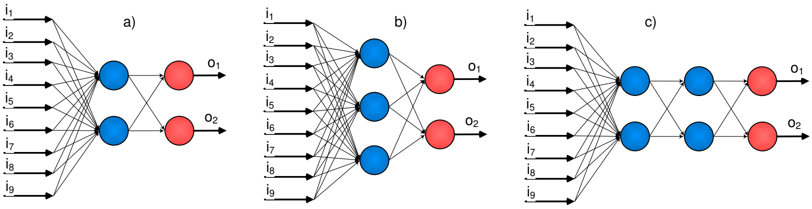


Figure 2
Models of ANN used with 9 inputs; (a) network 2-2, (b) network 3-2, (c) network 2-2-2

Testing the ANN, i.e. the evaluation of its performance, can be done on the basis of different criteria. In this research the performance of the ANN was evaluated on the basis of MAPE (mean absolute percent error) using the following expression:

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|actual_i\ \text{-}predicted_i\right|}{actual_i}\bullet 100\% \tag{5}$$

where n – is the number of data sets for testing.

Testing was conducted on 15 sets of input and output data which were not taken into account when the network was being trained. In addition to testing all the neural networks on the basis of MAPE, their stability was also tested. The control of the network stability was realized by comparing MAPE output results obtained through testing networks for 4 consecutive testing iterations; i.e. for each version of the network, 4 tests were carried out. If MAPE differs from the iterations of prediction the network is considered unstable, i.e. it will not always carry out a prediction with the same accuracy. Apart from MAPE, control of stability of network sensitivity on input parameters was also carried out, defining the impact of input parameters on the result. If the significance of input data varied within the four analyzed iterations, the network was considered to be unstable; i.e. the network was only considered to be stable provided that it gave the same results of MAPE and the significance of input parameters for all four testing iterations.

Table 4

MAPE and stability of analyzed networks with nine input parameters

Training functions	Output	one hidden layer				two hidden layers			
		number of neurons in hidden layer - number of neurons in output layer				number of neurons in hidden layer 1 - number of neurons in hidden layer 2 - number of neurons in output layer			
		2-2	3-2	4-2	5-2	2-1-2	2-2-2	3-2-2	2-3-2
traingda	concrete	unstable	unstable	unstable	unstable	unstable	unstable	unstable	unstable
	reinforcement								
trainlm	concrete	unstable	unstable	unstable	unstable	unstable	unstable	unstable	unstable
	reinforcement								
trainbfg	concrete	unstable	unstable	unstable	unstable	unstable	unstable	unstable	unstable
	reinforcement								
trainbr	concrete	**11.55%**	**14.22%**	unstable	unstable	unstable	unstable	unstable	unstable
	reinforcement	**8.93%**	**13.73%**						
	average	**10.24%**	**13.98%**						
traincgb	concrete	unstable	unstable	unstable	unstable	unstable	unstable	unstable	unstable
	reinforcement								

As can be seen in Table 4, stable networks were only present in the case when there was one hidden layer of neurons that contains two or three neurons, where the network was trained by trainbr function (Bayesian regularization). At the same time, the network with 2 neurons in the hidden layer provides higher accuracy of prediction compared with the network with 3 neurons in the hidden layer. For the network trainbr 2-2, $MAPE_{average}$=10,24% , whereas for the network trainbr 3-2, $MAPE_{average}$=13,98%. Figures 3 and 4 show the values of PE (percentage error, equation 6) for the two chosen networks (trainbr 2-2 and trainbr 3-2).

$$PE=\frac{predicted_i\text{-}actual_i}{actual_i}\bullet 100\% \tag{6}$$

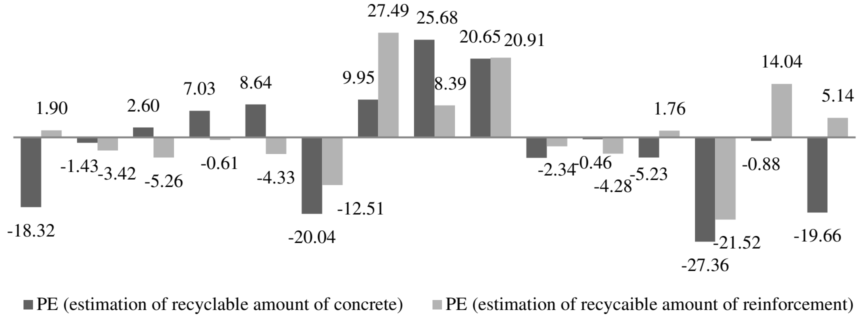


Figure 3

The PE graphic of each individual piece of data for testing trainbr 2-2 (network with nine inputs)

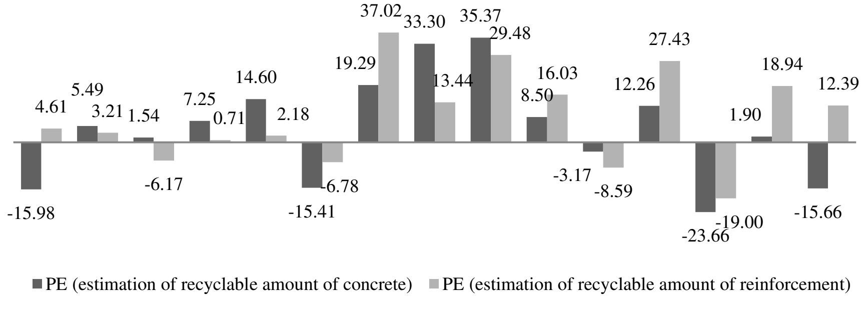


Figure 4

The PE graphic of each individual piece of data for testing trainbr 3-2 (network with nine inputs)

Observing Figures 3 and 4, a conclusion can be drawn that trainbr 2-2 network provides smaller errors of the output data, where the maximum error in predicting the amount of concrete is 27.36%, and the maximum error in the prediction of the amount of reinforcement is 27.49%.

In order to realize the significance of input parameters on the output results, the sensitivity analyses of the observed networks was carried out. Sensitivity analysis provides vital insights into the usefulness of individual input variables. Through sensitivity analysis, variables that do not have significant effect can be taken out of the neural network model and key variables can be identified [20]. Sensitivity analysis results are shown within Figure 5.

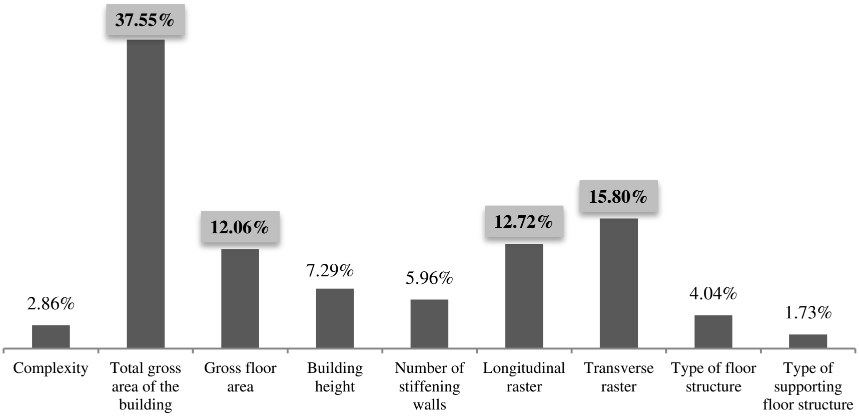


Figure 5

Sensitivity analysis (network with nine inputs)

Based on Figure 5, it is possible to conclude that in the case of the observed networks, the most significant parameters are total gross area of the building (37.55%), transverse raster (15.80%), longitudinal raster (12.72%) and gross floor area (12.06%). These parameters are adopted for the further analysis since their significance exceeds 10%. Training and testing of the ANN were carried out again, but instead of using nine inputs, only the four were used

An analysis was carried out identical to the one for the previous input data. The analysis results are presented in Table 5. Fig. 6 shows a PE graph for the data set for testing the ANN (network with four inputs) obtained on the basis of PE (percentage error).

Table 5

MAPE of analyzed networks with four input parameters

Training functions	Output	One hidden layer	
		Number of neurons in hidden layer - number of neurons in output layer	
		2-2	3-2
trainbr	concrete	9.32%	9.47%
	reinforcement	8.87%	10.82%
	average	9.10%	10.15%

If tables 4 and 5 are compared, it is possible to draw the conclusion that the network trainbr 2-2 trained with 4 input parameters provides the most accurate prediction of the recyclable amount of concrete and reinforcement ($\text{MAPE}_{\text{average}}$=9.10%). At the same time, the network trainbr 2-2 provides the most accurate prediction of the amount of concrete ($\text{MAPE}_{\text{concrete}}$=9.32%), as well as the most accurate prediction of the amount of reinforcement ($\text{MAPE}_{\text{reinforcement}}$=8.87%).

In Fig. 6 and Fig. 7 is a PE graph for the data set for testing the ANN (network with four inputs) obtained on the basis of PE (percentage error).

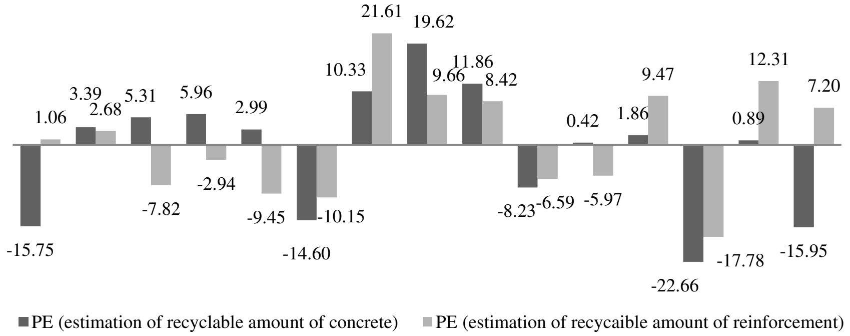


Figure 6
The Graphic PE of each individual piece of data for testing trainbr 2-2 (network with four inputs)

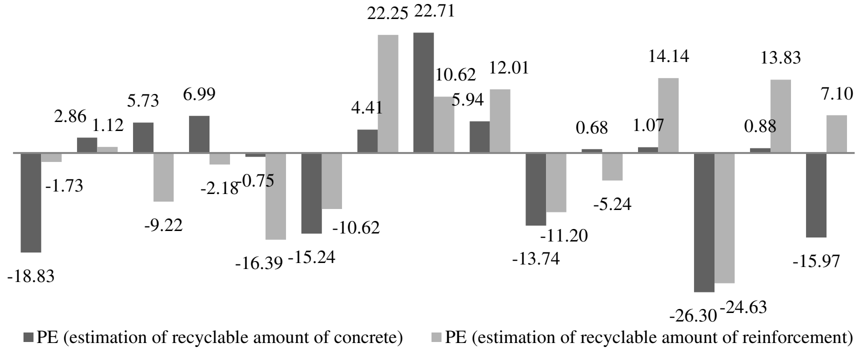


Figure 7
The Graphic PE of each individual piece of data for testing trainbr 3-2 (network with four inputs)

If Figures 3 and 6 are compared, it is possible to conclude that after removing 5 input parameters for the network trainbr 2-2, the maximum error of prediction of the amount of reinforcement was reduced from 27.49% to 22.66%, as well as from 27.36% to 21.61% regarding the amount of concrete. Considerable progress was made for the network trainbr 3-2, resulting in a significant reduction in the maximum error for the prediction of the amounts of both concrete and reinforcement. Maximum prediction error for the amount of concrete was reduced from 34.37% to 26.30%, whereas the maximum prediction error for the amount of reinforcement was reduced from 37.02% to 24.63%. Figure 8 shows the significance of input parameters for networks with 4 input parameters.

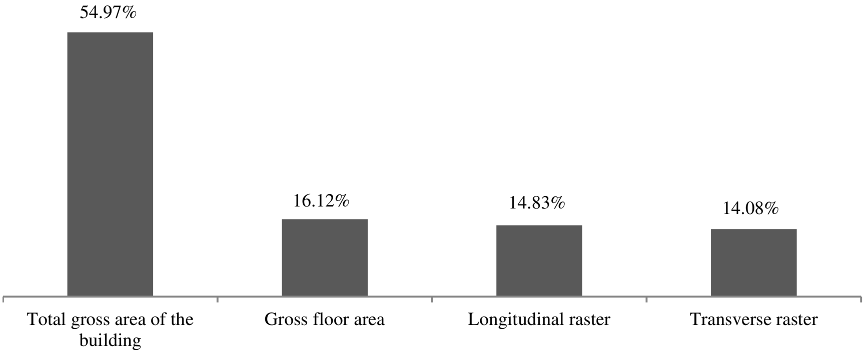


Figure 8
Sensitivity analysis (network with four inputs)

As can be seen in Figure 8, there was an increase in the significance of total gross area of the building and average gross floor area parameters, setting the balance between the significance of parameters relating to skeletal construction raster (longitudinal raster and transverse raster).

6 Discussion of the Results

During research into the applicability of ANN to the problem of predicting the quantities of concrete and reinforcement that can be recycled, a database of 110 residential building projects was created with 9 input parameters and 2 output parameters. The adopted input parameters are the following: complexity of the building, total gross area of the building, average gross floor area, building height, number of stiffening walls, longitudinal raster, transverse raster, type of floor structure, type of supporting floor structure. Before neural network training, preparation of data through "Z-Score" normalization was carried out.

Within the analysis, the networks with error backpropagation algorithms were observed, containing at the same time one hidden layer with 2 to 5 neurons as well as two hidden layers with 1 to 3 neurons. In network training, 5 training functions were used: traingda (Gradient descent with adaptive lr backpropagation), trainlm (Levenberg-Marquardt backpropagation), trainbfg (BFGS quasi-Newton backpropagation), trainbr (Bayesian regularization) and traincgb (Powell-Beale conjugate gradient backpropagation). As the first criterion for network validity, its stability was adopted, where the observed results were obtained through 4 consecutive iterations for the same values of data. In doing so, stability of the obtained values of MAPE was observed, as well as the stability of sensitivity analysis on the input parameters. The network was regarded as stable only if it provided the same results of MAPE and input data significance for all 4 testing iterations. After analyzing 40 networks in 160 training and testing processes, a conclusion was reached that only networks with one hidden layer containing 2 or 3 neurons and trained with the trainbr function (Bayesian regularization) proved to be sable. The network trainbr 2-2 containing 2 neurons in its hidden layer and two in its output layer has an average error of prediction of the amount of concrete and reinforcement of 10.24%, whereas the network trainbr 3-2 containing 2 neurons in its hidden layer and two in its output layer displays an average error of 13.98%. For the network trainbr 2-2, the individual error never exceeded the value of 27.49%, while in the case of trainbr 3-2 network the individual error never exceeded the value of 37.02%.

When analyzing the stability of the observed networks on the input parameters, a conclusion was drawn that four parameters have the strongest impact on the output results: total gross area of the building, average gross floor area, longitudinal raster and transverse raster. For this reason, the training process of the two observed networks was repeated with these 4 input parameters. By doing so, the average error of the output results for the network trainbr 2-2 was reduced from 10.24% to 9.10%, whereas for the network trainbr 3-2 the average error was reduced from 13.985 to 10.15%. In addition to reducing the average error, the maximum value of individual errors was also reduced. For the network trainbr 2-2, the individual error did not exceed the value of 22.66%, whereas in the case of trainbr 3-2 network, the individual error did not exceed the value of 26.30%. Based on all of the above, it can be concluded that for the observed problem, it is justified to train the network based on the 4 defined parameters. At the same time, the obtained results indicate that when collecting the data with the aim of data base expansion, particular attention should be paid to these input data.

Conclusions and Future Research

In this study is a presentation of the analysis and formation of networks for the purpose of predicting the amounts of recyclable concrete and reinforcement based on a database formed for the purposes of this research which contains data from 110 major residential building projects. It was concluded that for the given database, the best results are offered by a network with error backpropagation, and with one hidden layer of 2 neurons. When analyzing the results, a conclusion was reached that a network provides higher accuracy when it is trained with the 4 most significant parameters out of 9 which are defined within the base. The value of the average error of predicted amounts of concrete and reinforcement, compared with actual values, amounts to 9.10%.

Future research should be based on an expansion of the data base to other types of building constructions, as well as on other types of materials, such as brick, wood, ceramics, etc. In this way, a larger number of constructions suitable for recycling could be comprised, where it would be possible to determine a more comprehensive recycling capacity of a building construction, due to being able to predict the amounts of all the recyclable materials.

Acknowledgement

The work reported in this paper is a part of the investigation within the research project TR 36017 "Utilization of by-products and recycled waste materials in concrete composites in the scope of sustainable construction development in Serbia: investigation and environmental assessment of possible applications", supported by the Ministry for Science and Technology, Republic of Serbia. This support is gratefully acknowledged.

References

[1] W. McCulloch, W. Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, Vol. 5 (1943) 115-133 http://www.cse.chalmers.se/~coquand/AUTOMATA/mcp.pdf (last visited on Oct. 17. 2011)

[2] S. Haykin, Neural Networks, Prentice Hall, New Jersy, 2005

[3] J. M. Zurada, Introduction to Neural Systems, West Publishing Company, St. Paul, 1992

[4] A. Nigrin, Neural Networks for Pattern Recognition, Cambridge MA: The MIT Press, 1993

[5] K. Kawaguchi, The McCulloch-Pitts Model of Neuron, http://wwwold.ece.utep.edu/research/webfuzzy/docs/kk-thesis/kk-thesis-html/node12.html (last visited on Feb. 15, 2011)

[6] http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html (last visited on Feb. 01. 2013)

[7] G. H. Kim, S. H. An, K. I. Kang, Comparison of Construction Cost Estimating Models Based on Regression Analysis, Neural Networks, and Case-based Reasoning, Building and Environment, 39 (2004) 1235-1242

[8] Y. R. Wang, E. Gibson Jr., A Study of Preproject Planning and Project Success Using ANNs and Regression Models, Automation in Construction 19 (2010) 341-346

[9] H. M. Gunaydın, S. Z. Dogan, A Neural Network Approach for Early Cost Estimation of Structural Systems of Buildings, International Journal of Project Management 22 (2004) 595-602

[10] W. Yu, M. J. Skibniewski, A Neuro-Fuzzy Computational Approach to Constructability Knowledge Acquisition for Construction Technology Evaluation, Automation in Construction 8 (1999) 539-552

[11] G. H. Kim, J. E. Yoon, S. H. An, H. H. Cho, K. I. Kang, Neural Network Model Incorporating a Genetic Algorithm in Estimating Construction Costs, Building and Environment 39 (2004) 1333-1340

[12] M. Z. Cheng, H. C. Tsai, E. Sudjono, Conceptual Cost Estimates Using Evolutionary Fuzzy Hybrid Neural Network for Projects in Construction Industry, Expert Systems with Applications 37 (2010) 4224-4231

[13] M. Y. Cheng, H. C. Tsai, W. S. Hsieh, Web-based Conceptual Cost Estimates for Construction Projects Using Evolutionary Fuzzy Neural Inference Model, Automation in Construction 18 (2009) 164-172

[14] J. S. Wu, J. Han, S. Annambhotla, S. Bryant, Artificial Neural Networks for Forecasting Watershed Runoff and Stream Flows. ASCE, Journal of Hydrologic Engineering, may/june (2005) 216-222

[15] H. Adeli, M. Wu, Regularization Neural Network for Construction Cost Estimation, Journal of Construction Engineering and Management, January/February (1998) 18-24

[16] I. Flood, N. Kartam, Neural Network in Civil Engineering I. ASCE, Journal of Computing in Civil Engineering, 8(2) (1994) 131-148

[17] I. Flood I, N. Kartam, Neural Network in Civil Engineering II. ASCE, Journal of Computing in Civil Engineering, 8(2) (1994) 149-162

[18] R. Rohas, Neural Networks, A Systematic Introduction, Springer-Verlag, Berlin, 1996, pp. 151-184

[19] S. Bhokha, S. Ogunlana, Application of Artificial Neural Network to Forecast Construction Duration of Buildings at the Predesign Stage Engineering, Construction and Architectural Management 6/2 (1999) 133-144

[20] O. Tatari, M. Kucukvar, Cost Premium Prediction of Certified Green Buildings: A Neural Network Approach, Building and Environment 46 (2011) 1081-1086

[21] S. Rebano-Edwards, Modelling Perceptions of Building Quality—A Neural Network Approach, Building and Environment 42 (2007) 2762–2777

[22] V. Mučenski, I. Peško, M. Trivunić, J. Dražić, G., Ćirović, Optimization for Estimating the Amount of Concrete and Reinforcement Required for Multi-storey Buildings, Building Materials and Structures 55(2) (2012) 27-46

Gender Classification using Multi-Level Wavelets on Real World Face Images

Sajid Ali Khan, Muhammad Nazir, Naveed Riaz

Department of Computer Science, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Plot # 67, Street # 9, H/8-4 Islamabad, 44000, Pakistan
sajid.ali@szabist-isb.edu.pk, nazir@szabist-isb.edu.pk, n.r.ansari@szabist-isb.edu.pk

Abstract: Gender classification is a major area of classification that has generated a lot of academic and research interest over the past decade or so. Being a recent area of interest in classification, there is still a lot of opportunity for further improvements in the existing techniques and their capabilities. In this paper, an attempt has been made to cover some of the limitations that the associated research community has faced by proposing a novel gender classification technique. In this technique, discrete wavelet transform has been used up to five levels for the purpose of feature extraction. To accommodate pose and expression variations, the energies of sub-bands are calculated and combined at the end. Only those features are used which are considered significant, and this significance is measured using Particle Swarm Optimization (PSO). The experimentation performed on real world images has shown a significant classification improvement and accuracy to the tune of 97%. The results also reveal the superiority of the proposed technique over others in its robustness, efficiency, illumination and pose change variation detection.

Keywords: Gender Classification; Discrete Wavelet Transform; Particle Swarm Optimization; Feature Selection; Real World Face Images

1 Introduction

In today's technological world Gender Classification plays a vital role. It is widely used in applications such as customer-oriented advertising, visual surveillance, and intelligent user interfaces and demographics.

With the evolution of human-computer interaction (HCI), in order to meet the growing demands for secure, reliable and convenient services, computer vision approaches like face identification, gesture recognition and gender classification will play an important role in our lives.

Features are generally classified into two categories 1) Appearance-based (Global) and Geometric-based (Local) features. In the appearance-based feature extraction technique, an image is considered as a high dimensional vector and features are extracted from its statistical information without any dependence on extensive knowledge about the object of interest. This technique is simple and fast but unreliable, especially when variations in local appearance occur. In the geometric-based feature extraction approach, geometric features like nose, mouth and eyes are extracted from the face portion. This approach has the advantage of rotation and variation invariability but generally misses a lot of helpful information.

In the early 1990s, Golomb et al. [1] trained a two-layer neural network called SEX-NET and achieved 91.9% classification accuracy by using 90 frontal face images. In 1995, Brunelli and Poggio et. al. [2] extracted geometric features and used them to train the networks, achieving a 79% classification accuracy rate. Sun et. al. (2002) [3] claimed that Genetic Algorithm (GA) performs well for important feature's selection task. They used Principal Component Analysis (PCA) to create the features' vector and GA to select the important features. They achieved 95.3% accuracy after training the Support Vector Machine classifier by using those important features. In 2004, Jain and Huang et. al. [4] extracted facial features using Independent Component Analysis (ICA) and then classified gender using Liner Discriminant Analysis (LDA). They performed experiments on a normalized FERET database and achieved a 99.3 % classification accuracy rate. In 2006, Sun et al. [5] used Local Binary Pattern (LBP) to create features for the input of AdaBoost and achieved 95.75% accuracy in terms of the classification rate. In 2007, Baluja and Rowley et al. [6] achieved 93% classification accuracy. They used Pixel comparison operators with an Adaboost classifier. In 2010, Nazir et al. [7] used Discrete Cosine Transform (DCT) to extract the important facial features and then used a K-nearest neighbor classifier to classify gender. In 2011, Sajid et al. [8] used Discrete Wavelet Transform (DWT) to extract facial features. They claimed that classifier performance was better after different classifiers ensemble using the weight majority technique. They performed experiments on a Stanford University Medical Students (SUMS) face database and achieved 95.63% classification accuracy rate.

A public setback alongside the above studies is that they have utilized the frontal face under controlled environments (e.g. SUMS, FERET). The images in these databases usually contain images that have a clean background, are occlusions free, provide only frontal faces, contain limited facial expressions and have consistent lighting effects. However, real-life images are usually captured in unconstrained environments and conditions. A real-life image usually contains significant appearance variation, such as illumination change, poor image quality, makeup or occlusions and different facial expressions. The images in Fig. 1 demonstrate these facts. Gender recognition in an unconstrained environment is a more challenging task compared to faces captured in constrained environments like the FERET and SUMS face databases. In literature, this problem has been addressed by few studies. Shakharovich et al. (2002) [9] collected 3,500 face images from the web. They used Haar-like features and obtained a 79% accuracy rate after using Adaboost and 75.5% after using SVM. Gao and Ai (2009) et al. [10] reported 95.5% classification accuracy after performing experiments on 10,100 real-life images. They used Haar-like features with a probabilistic boosting tree. It is difficult to use these results as benchmarks as their database are not publicly available. Kumar et al. (2009) [11] performed experiments on real-life images. They reported 81.22% classification accuracy after training many binary "attributes" classifiers. They focused more on face verification rather than gender classification.

Figure 1

Sample Labeled Faces in the Wild (LFW) Face database images

In our proposed technique, we extract important facial features using 5-Level discrete wavelet transform technique. Then, the energy of each sub-band is calculated and combined in the last phase. As features of all sub-bands are combined, the proposed system supports variations in facial expression and poses changes. After facial feature extraction, we implemented PSO for the selection of important features and dimension reduction. The Support Vector Machine classifier is trained and tested by using PSO-based selected features.

We have organized the paper in such a way that in Section 2, the proposed methodology is presented. In the next section, experimental setting and results are discussed and compared with state of the art existing techniques. Conclusion and future work are discussed in Section 4.

2 Proposed Methodology

Our proposed technique has the following steps, given below, and Fig. 2 depicts these steps:

Preprocessing: First the sample images are aligned using commercial software [12], then histogram equalization is applied to normalize the face.

Face Extraction: Facial portion is extracted and removal of the unwanted area using a spatial coordinate system is performed.

Features Extraction: Facial features in vertical and horizontal direction are extracted using 5-level DWT.

Optimized Features Selection: Feature sets with high importance and high accuracy are selected using PSO.

Classification: SVM is trained and tested by using optimized features.

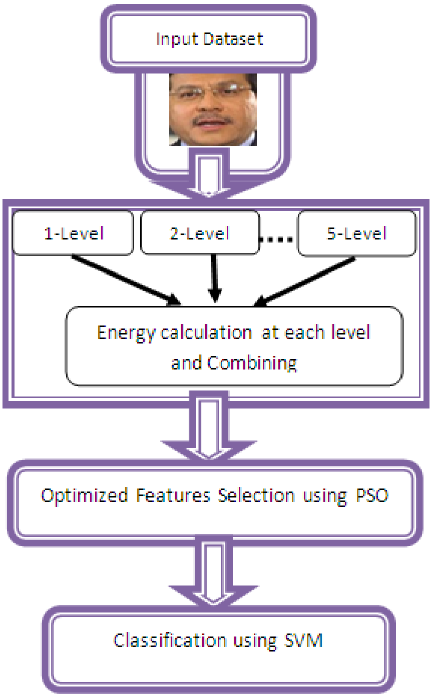

Input Dataset
1-Level
2-Level
....
5-Level
Energy calculation at each level and Combining
Optimized Features Selection using PSO
Classification using SVM

Figure 2
Proposed System Architecture

2.1 Facial Portion Extraction

We have preferred to use the system with continous varying coordinates on discerete indices. In spatial coordinate system, a position of an image is described in terms of x and y and not in row and column format. Fig. 3 dipicts the spatial coordinate system.

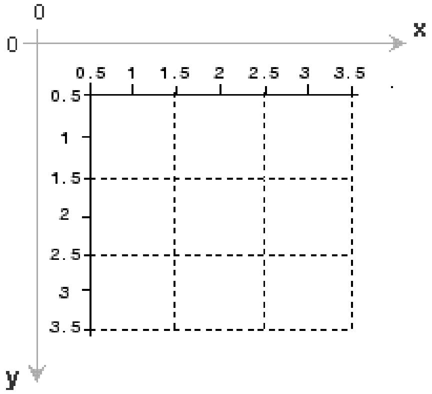


Figure 3
Spatial Coordinate System

As a single small image may contain thousands of pixels, high dimensional data affects the computational time and makes the system slow. To remove the unwanted area and to extract only the facial portion, we have used spatial coordinate system. By a hit and miss method, we set the X and Y coordinate values.

Fig. 4, depicts the extracted faces from original images.

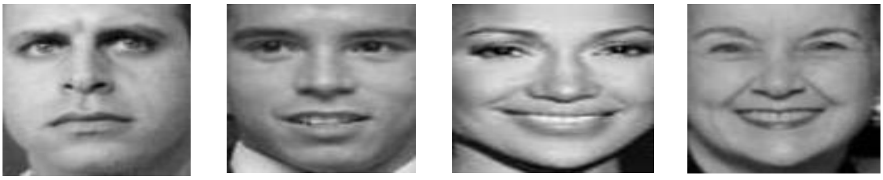


Figure 4
Extracted faces using Spatial coordinates

2.2 Feature Extraction

Wavelets are comparatively more beneficial than other mathematical transformations such as Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT). Functions with interruptions and functions with sharp spikes generally follow wavelet basis functions to a lesser degree than sine cosine functions to obtain similar resemblance. Wavelets have been utilized in different ways in images processing since 1985 [13]. Its potential to furnish spatial and frequency representations of the image at the same time influences its use for feature extraction. The dissolution of the input data into many layers of division in space and frequency permits us to separate the frequency components presented by inherent or natural damage. Wavelet-based methods reduce or cut off these variable sub-bands and concentrate on the space frequency sub-bands that include the most appropriate information to show the data in a better way and assist in the classification task. A huge selection of wavelet families exists. These rely on the selection of the mother wavelet. In 1986, Sergent [14] stated that the low-frequency band and high-frequency band play different roles. The low-frequency band provides the universal description while the high-frequency component provides the first-rate characteristics needed in the identification task. Sergent argues that as a human face is a flexible object, it has countless facial expressions, and expressions tend to affect the local spatial components of the face. Wavelets are very helpful in enhancing the authenticity of image registration. For this purpose, wavelets take into account both spatial and spectral information by yielding a multi-resolution representation and keeping away from wandering to any global or local information. The other advantages of using wavelets include bringing data with different spatial resolution to a common resolution using the low-frequency sub-bands while providing access to edge features using the high-frequency sub-bands. As depicted in Fig. 5, four new images are formed at each level of wavelet decomposition from the original images (N x N pixels). The four new images have a reduced size. 1/4 of the original image. Filters are applied to the images in horizontal and vertical directions; that is why the four new images are given names according to the filters. The four decreased images are LL, HL, LH and HH. The LL contains the most information of the image information, and it is also the reduced version. The HH image is noisy because it contains high-frequency information, which is why it is not useful for image registration application. Theh LH image represents horizontal edge features while HL represents vertical edge features.

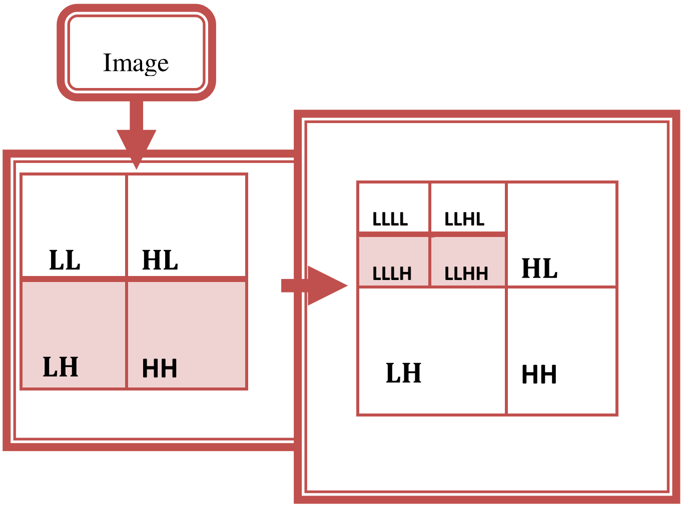


Figure 5
2-Level 2-D wavelet transform

2.3 Optimized Feature Selection

PSO was put forward by Dr. Eberhart and Dr. Kennedy in 1995 as a computational model based on the idea of cooperative behavior and swarming in biological populations influenced by the social behavior of fish schooling [15]. At present, PSO has been utilized as a prosperous optimizer in countless areas, for example training man-made neural webs, constrained purpose optimization, wireless web optimization, and data clustering.

Computation in PSO is derived from a population (Swarm) of processing elements called particles in which each particle symbolizes a candidate solution. PSO has a strong resemblance to evolutionary computation techniques like those of GA. This system is commenced with a population of unplanned solutions and seeks for optima by modernizing generations. The seeking process makes use of a combination of deterministic and probabilistic rules that depend on information sharing among their population fellows to boost their search procedures. Nevertheless, in contrast to GA, PSO has no evolution operators such as crossover and mutation. Every particle in the search space develops its candidate solution over time, utilizing its individual memory and knowledge acquired by the swarm as a whole. The information sharing mechanism in PSO is altogether different compared to GA. In GA, chromosomes contribute information to each other. As a result, the entire population advances as one group toward an optimal area. In PSO, the word finest particle discovered among the swarm is the sole information contributed among the particles. It is a one-sided information sharing mechanism. In PSO, the computation time is considerably less compared to GA because all the particles in PSO have the tendency to coincide with the finest solution swiftly.

PSO is used for problem optimization. Particles, also known as swarms, are used to search for the optimal solution in the search space. Each particle represents a candidate solution in the search space and is represented by particular coordinates.

$$X_i = (x_{i1}, x_{i2}, \ldots\ldots x_{iD}) \quad (1)$$

Where X_i represents the *eighth* position of the particle. The velocity which is the rate of change of current and the new position is denoted by the equation 2.

$$V_i = (v_{i1}, v_{i2}, \ldots\ldots\ldots v_{iD}) \quad (2)$$

The fitness function for each particle is determined after comparison with the previous best result. It is also compared with the best result in the search space. Equations 2 and 3 are used to update the position and velocity of the particle after finding the best values.

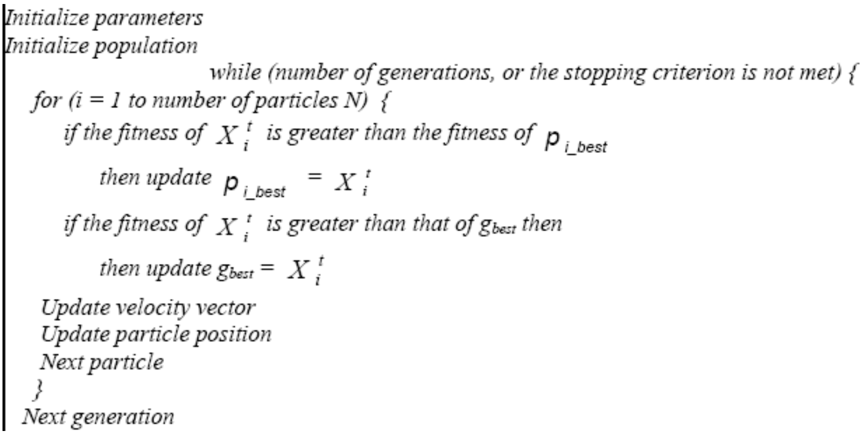
$$V_i^{t+1} = \omega * V_i^t + c_1 * rand_1 * (p_{i\_best} - X_i^t) + c_2 * rand_2 * (g_{best} - X_i^t) \quad (3)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1}) \tag{4}$$

Where,

- $i = (1,2,3.......N)$
- N is the size of particle (swarm)
- p_{i_best} and g_{best} are the local and global best solution in the search space.
- C1 and C2 are cognitive (represent the particle private itself experience) and social (represents the cooporation between particles) parameters having values between 0 and 2.
- In the first part of equation number 3, *w* represents the inertia weight that is used to control the search algorithm balance between exploitation and exploration.

Fig. 6 represents the pseudo code of PSO.



```
Initialize parameters
Initialize population
                        while (number of generations, or the stopping criterion is not met) {
  for (i = 1 to number of particles N) {
      if the fitness of $X_i^t$ is greater than the fitness of $p_{i\_best}$
          then update $p_{i\_best}$ = $X_i^t$
      if the fitness of $X_i^t$ is greater than that of $g_{best}$ then
          then update $g_{best}$ = $X_i^t$
    Update velocity vector
    Update particle position
    Next particle
  }
 Next generation
```

Figure 6
PSO technique pseudo code

2.3.1 Binary Particle Swarm Optimization (BPSO)

A binary PSO algorithm was developed in [16]. In this edition, the particle locale is coded as 1 or 0 and the velocity function is used as the probability distribution. The following is the equation which updates the particle position.

$$if\ rand < \frac{1}{1+e^{-V_i^{t+1}}}\ then\ X_i^{t+1} = 1\ else\ X_i^{t+1} = 0 \quad (5)$$

The bit value ‘1’ indicates that the particle (feature) is used for the next generation and ‘0’ indicates that particle (feature) is terminated for the next generation.

Each particle is searched for the best solution (Optimal feature) in the search space. Evolution is driven by fitness function. Each particle is coded as P = F1 F2 F3……Fn, where ‘n’ is the size of the feature set. If a feature set size is 20, i.e. P = F1 F2 F3…… F20, then theh BPSO selects 1, 4, 5, 6, 7 and 9 as the best features using the fitness function. Thus, the accuracy of those optimized six features is greater than the combined 20 features.

Fitness Function: In every iteration, each particle is assessed employing fitness purpose and worth of the optimal particle returned as the result. This evaluation is driven by the fitness function ‘F’ that evaluates the quality of the evolved particles in terms of their ability to minimize the class separation term indicated by the scatter index among the different classes [17]. Let k1 and k2 denote the classes of male and female and N1 N2…..Nn represent the number of images each class had.

Let's define M1, M2…… Mn, me as;

$$M_i = \frac{1}{N_i}\sum_{j=1}^{N_i} K_j^{(i)}, i = 1, 2, 3......n \quad (6)$$

$K_j^{(i)}$, J = 1,2,3........ N_i, indicate the image from class k_i.

$$M_0 = \frac{1}{N}\sum_{i=1}^{n} N_i M_i$$

Here, M_0 is the grand mean and N is the total number of images for both classes.

Fitness function ‘F’ can be computed by equation 7;

$$F = \sum_{i=1}^{n} (M_i - M_0)^t (M_i - M_0) \quad (7)$$

2.4 Classification using Support Vector Machine (SVM)

The binary classification is a two class (0 or 1) problem, and the goal of this problem is to separate the two classes by mean of function. SVM is a useful technique for data classification and is easier to use compared to Neural Networks.

SVM takes data (features) as input and predicts which data belongs to which class. The goal of SVM is to find the optimal hyper plan such that it minimizes the error rate for an unseen test sample. According to the structural risk minimization principals and VC dimension minimization principle [18] a linear SVM uses a systematic approach to find a linear function with the lowest capacity.

The SVM classifier correctly separates the training data of labeled sets of M training samples (Xi, Yi), where Xi ε RN and Yi are the associated label i.e. (Yi ε {-1,1}).

The hyper plan is defined by equation 8;

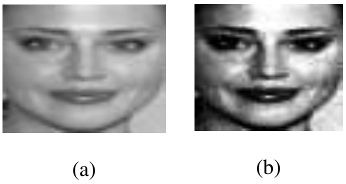
$$f(x) = \sum_{i=1}^{M} y_i \alpha_i K(x, x_i) + b \tag{8}$$

3 Experimetal Results and Discussion

We used the MATLAB 2009a environment for our experiments. Labeled faces in the wild (LFW) [19] face database were used in our experiment. This database contains 13,233 images collected from the web. To remove unwanted areas, a face portion is extracted using a spatial coordinate system and then, as shown in Fig. 7, histogram equalization is applied to normalize the face image. We selected 400 face images, 200 male and 200 female, for the experiments. All the faces are aligned using commercial align software (Wolf et al 2009) [12]. We avoided the selection of those images for which it is difficult to establish the ground truth (such as a hidden face). 5-fold cross validation is used in all experiments. All the images are converted into gray scale from the RGB color space to retain the luminance factor and to eliminate the hue and saturation information. An 8 bit grayscale image contains 256 gray levels which have values between 0 and 255.

The image is resized to the size of 32x32 pixels and the spatial coordinate system is used to extract the facial portion. A hit and miss method is used to find the x1, x2, y1 and y2 coordinates of an image. As all the images are of the same size, the same coordinate values are used for all images to extract the face portion. Figure 8 depicts this process.

Five level decomposition is performed using a 2-dimensional Haar wavelet transform. First, details coefficients of all levels are combined, and their energy is found out. Then horizontal and vertical coefficients are combined and their energy is calculated. The combined detail co-efficient and horizontal and vertical co-efficient of DWT are then passed to PSO. PSO evaluates the features and provides results as optimized features. Features' vectors of size 10, 20, 30, 40 and 50 are obtained after PSO implementation. These features are then passed to SVM with a train-to test-ratio of 1:3 and 3:1 using 5-fold cross validation. We have also used some other state-of-the-art classifiers for testing resultant feature sets and compared their results with SVM classifier. Table 1 shows that the use of the 30-feature SVM classifier outperforms other classifiers. The accuracy of BPNN and KNN is same, with both using 10 and 20 features, but the accuracy increases when the features set size is set to 30. We have also noticed that the accuracy rate degrades when the features set size is more than 30.



(a) (b)

Figure 7

(a) Face before normalization

(b) Face after normalization

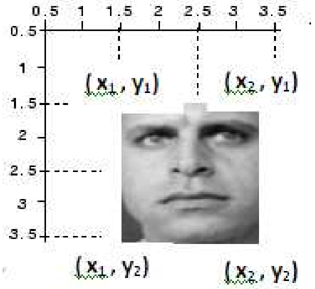


Figure 8
Face image extraction

Table 1
Comparison of SVM classifier accuracy rate with other state-of-the-art classifiers accuracy using different number of features

Classifier/ Features	10	20	30	40	50
BPNN	0.7433	0.7433	0.8683	0.8683	0.7017
KNN	0.7433	0.7433	0.8051	0.7433	0.7017
SVM	0.785	0.785	**0.973**	0.8267	0.7433
FLDA	0.7433	0.785	0.8192	0.8267	0.6183
NMS	0.7017	0.6183	0.7513	0.785	0.785
LDA	0.7017	0.5767	0.8447	0.785	0.7433
NMC	0.7017	0.7433	0.8135	0.785	0.7017

Accuracy is evaluated after passing the 5-level DWT combined features to SVM. We compared it with the proposed techniques as shown in Table 1. Table 1 clearly shows that the proposed technique’s accuracy rate is high with reduced dimensions (i.e. it utilizes a minimum number of features). In Fig. 9, the Feature set of size FS-250, FS-300, FS-450 and FS-500 are used to train and test SVM to evaluate the DWT+SVM accuracy rate. On the other hand, high classification accuracy of 97% is obtained with Feature set size reduced (i.e. FS-20, FS-30, FS-40, FS-50) after optimizing the features using PSO (i.e. using our proposed technique).

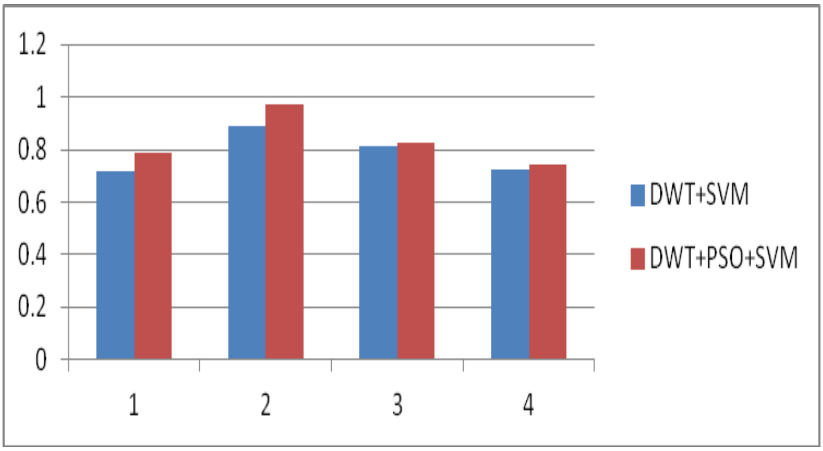


Figure 9

Proposed technique (DWT+PSO+SVM) comparison with another mechanism (DWT+SVM) using the different number of features.

Table 2

Parameter settings for PSO

Parameter Name	Value
Swarm Size (N)	100
Cognitive Parameter (C1)	2
Social Parameter (C2)	2
Inertia weight (ω)	0.6
Iterations	100

Table 3

Parameter settings for GA

Parameter Name	Value
Population Size (N)	100
Cross Over Probability (Pc)	0.5
Mutation Probability (Pm)	1
Iterations	100

Fig. 10 presents the comparison of the PSO-based optimized feature's classification accuracy rate with GA-based optimized feature's classification accuracy rate using a features set size of 20, 30, 40 and 50. The GA-based optimized features' accuracy rate is high only when we use a features set of size 50; however, in all other cases (i.e. Feature set of size 20, 30 and 40), the PSO-based optimized feature's accuracy rate is higher. In PSO, the computation time is substantially less as compared to GAs because all the particles in PSO have the tendency to coincide to the finest solution swiftly. So the PSO-based optimization mechanism is considerably more accurate and fast as compared to the GA-based optimization mechanism in the case of gender classification.

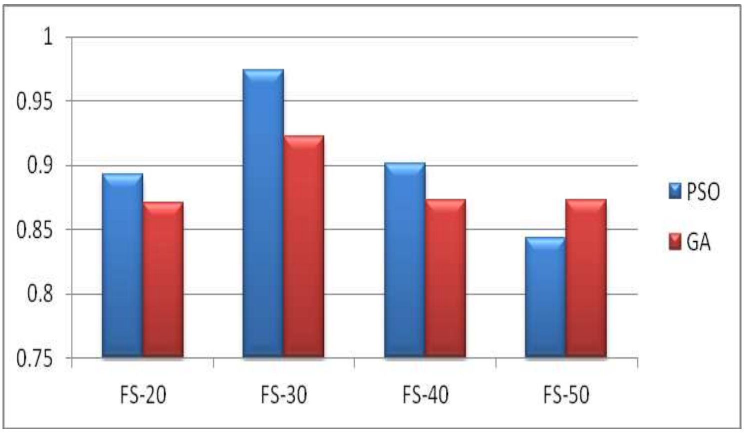


Figure 10

PSO and GA-based classification accuracy rate comparisons

Table 4 presents a comparison of our proposed technique with other gender classification techniques. In our proposed technique, we overcome problems such as high dimensions (by utilizing the minimum number of features) and variation in pose or occlusions (by combining different level's DWT coefficients). We have used real-life images that have a large amount of variations in facial expression and pose. The database that we have used is also publicly available, which helps to benchmark for the future.

Table 4

Comparison of SVM classifier accuracy rate with other

Methods		**Database**	**Data Dimensions**	**Real Life**	**Publically available**	**Recognition Rate**
Proposed		**LFW**	**30**	**Yes**	**Yes**	**97%**
Sun et al [3]	PCA,SVM	General Faces	150	No	No	95.3%
Jain and Huang [4]	ICA,LDA	FERET	200	No	Yes	96%
Baluja et al [6]	Pixels Comparisons	FERET	2409	No	Yes	94.3%
Nazir et al [7]	DCT,K-NN	SUMS	256	No	Yes	99.3%
Shakhnarovich et al[9]	Haar-like features	World Wide Web	35,00	Yes	No	79%
Gao et al [10]	Haar-like features	Consumer Images	10,100	Yes	No	95.5%
Shan, C [20]	LBP, SVM	LFW	2891	Yes	Yes	95%
Sajid et al [8]	DWT,SVM	SUMS	20	No	Yes	95.63%

Computational Complexity

In this section, performance is measured in terms of time complexity. Table 5 presents the computational time comparison of our proposed technique with other techniques. Most of the researchers have used geometric based feature extraction techniques to extract the face features and classify gender, which makes their technique computationally more expensive. We can see from the table that [6] used Adaboost and Support Vector Machine (SVM) to classify gender, and their technique consumed more time as compared with our proposed technique to classify gender. Similarly [7] used the Voila and Jones technique to detect the face portion first and then DCT to get the efficient face features. The Voila and Jones technique takes more time to extract the face portion as compared to the Spatial coordinate system. In [10] Boosting tree and Active shape model (ASM) are used to locate the facial points. The computational time of ASM increases with an increase in face image sets. In [20] the same ASM geometric feature extraction based approach is used, which is more time consuming compared to our proposed appearance-based approach.

Table 5

Computational time comparison with other techniques

Technique			Total Time (Sec)
Proposed		20 features	40
		30 features	60
		40 features	70
Baluja et al. [6]	AdaBoost+SVM	20 features	47
		30 features	68
		40 features	78
Nazir et al. [7]	Voila & Jones + DCT+KNN	20 features	130
		30 features	160
		40 features	183
Gao et al. [10]	Boosting Tree +ASM	20 features	240
		30 features	300
		40 features	360
Shan, C [20]	ASM+LBP+SVM	20 features	165
		30 features	220
		40 features	268

Conclusions and Future Work

Gender classification is considered one of the most active research areas in pattern recognition and image processing. Currently most of the acclaimed work in this domain revolves around a frontal facial image based classification. In this paper, we have focused on reducing the data dimensions and have tried to produce a more optimal feature set that more accurately represents a gender face. If the inadequate features are used, then even the best classifiers usually fail to achieve higher accuracy. Therefore, in this paper, we have tried to optimize the features by using PSO algorithms. After the optimization phase, a large number of redundant and irrelevant features are eliminated, resulting in a reduction in data dimensions. We have achieved a 97% classification accuracy rate after performing experiments on real-world face images (LFW) and have utilized the minimum number of features. We have combined different level's DWT detail coefficients, making our system more stable and supportive of variations in facial expressions, illumination and poses. The results of proposed technique when compared to other state-of-the-art gender classification techniques show that the proposed technique provides higher classification accuracy by utilizing the minimum number of features. This also reduces time complexity and makes the system fast. Furthermore, we intend to explore more Swarm-based optimization algorithms (SOAs), such as Ant colony optimization, and to make our system more accurate and stable.

References

[1] B. Golomb, D. Lawrence, T. Sejnowski: Sexnet: a Neural Network Identifies Sex from Human Faces, In: Advances in Neural Information Processing Systems, 1991, pp. 572-577

[2] R. Brunelli, T. Poggio: HyberBF Networks for Gender Classification, in Proceedings DARPA image understanding workshop, 1995, pp. 311-314

[3] Z. Sun, G. Bebis, X.Yuan, S. J. Louis: Genetic Feature Subset Selection for Gender Classification: A Comparison Study, In: proceeding 6th IEEE workshop on Applications of computer vision, Orlando, FL, USA, 2002, pp. 165-170

[4] A. Jain, J. Huang, S. Fang: Gender Classification using Frontal Facial Images, In: IEEE international conference on pattern Recognition, Tampa, Florida, 2008, pp. 1-4

[5] N. Sun, W. Zheng, C. Sun, C. Zou, L. Zhao: Gender Classification Based on Boosting Local BINARY pattern, In: Proceedings of the Third international conference on Advances in Neural Networks, Verlag Berlin, Heidelberg, 2006, pp. 194-201

[6] S. Baluja and H. Rowley: Boosting Sex Identification Performance, International Journal of computer vision, Vol. 71, 2007, pp. 111-119

[7] M. Nazir, M. Ishtiaq, A. Batool, A. Jaffar and M. Mirza: Feature Selection for efficient gender classification, In: 11th WSEAS International conference, Wisconsin, USA, 2010, pp. 70-75

[8] S. A. Khan, M. Nazir, N. Naveed and N. Riaz: Efficient Gender Classification Methodology using DWT and PCA, In: 14th IEEE international Multi-topic conference, Karachi, Pakistan, 2011, pp. 155-158

[9] G. Shakhnarovich, P. Viola and B. Moghaddam: A Unified Learning Framework for Real Time Face Detection and Classification, In: IEEE Internet Conference on Automatic Face & Gesture Recognition (FG), 2002, pp. 14-21

[10] W. Gao and H. Ai: Face Gender Classification on Consumer Images in a Multiethnic Environment, In: International Conference on Biometrics (ICB), Verlag Berlin, Heidelberg, 2009, pp. 169-168

[11] N. Kumar, P. N. Belhumeur and S. K. Nayar: Face Tracer: A Search Engine for Large Collections of Images with Faces, In: European Conference on Computer Vision (ECCV), 2009, pp. 340-353

[12] L. Wolf, T. Hassner and Y. Taigman, "Similarity Scores Based on Background Samples", In: Asian Conf. on Computer Vision (ACCV), 2009

[13] A. S. Samra, S. E. Gad Allah, R. M. Ibrahim: Face Recognition Using Wavelet Transform, Fast Fourier Transform and Discrete Cosine Transform, In: Proc. 46th IEEE International Midwest Symp. Circuits and Systems (MWSCAS'03), Vol. 1, 2003, pp. 272- 275

[14] J. Sergent: Micro Genesis of Face Perception, In: H. D. Ellis, M. A. Jeeves, F. Newcombe and A. Young, Editors, Aspects of Face Processing, Nijhoff, Dordrecht, 1986

[15] J. Kennedy and R. Eberhart: Particle Swarm Optimization, In: Proc. IEEE International Conference on Neural Networks, 1995, pp. 1942-1948

[16] J. Kennedy and R. C. Eberhart: A Discrete Binary Version of the Particle Swarm Algorithm, In: Proc. IEEE International Conference on Systems, Man, and Cybernetics, Vol. 5, 1997, pp. 4104-4108

[17] C. Liu and H. Wechsler: Evolutionary Pursuit and Its Application to Face Recognition, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 6, 2000, pp. 570-582

[18] V. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995

[19] G. Huang, M. Ramesh, T. Berg "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments", Tech. Rep. 07-49, University of Massachusetts, Amherst, 2007

[20] C. Shan, "Learning Local Binary Pattern for Gender Classification on Real-World Face Images", Pattern Recognition Letter, 2011

The Role of Legal Forms in Professional Football

Zoltán Imre Nagy

Óbuda University
Népszínház u. 8, H-1081 Budapest, Hungary
nagy.imre.zoltan@kgk.uni-obuda.hu

Abstract: In the event of a professional football club the fundamental sports objectives can get into contradiction with the intentions of the investor, who might wish, in many cases, to take care primarily for the increase of profits instead of serving the aims of football and the most important designation of a football club. In my study you will find an analysis of legal forms for football clubs, as well as the method of the management by owner determined primarily by such forms. We keep in mind primarily the aspect how these forms influence the possibilities of financing and the acquisition of capital. A form of entrepreneurship and management which takes into account the possibilities of a football club allows for the implementation of the most important objectives of sports, even if it is necessary to involve a considerable amount of capital. Through transportation into an incorporated firm with the appropriate shaping and operation of a group of companies, the aims of capitalisation will also be pursued, while the sports objectives remain the primary ones. (Such forms are limited liability companies, limited partnerships, share companies, registered associations, etc.) On the basis of foreign experience our study intends to make findings and proposals in order to serve establishing and operating such forms of enterprises, which will create – in addition to the priority of sports objectives – also broad possibilities for capitalisation in professional football.

Keywords: cost, football capital-company; capital; legal form of Business; Registered Association; Ltd; Share Company

1 Introduction

A rather big part of professional football clubs may be characterised by a permanent process of structural and management transformation, which comes to expression in the variation of the legal forms of undertaking in professional football, and, moreover, in the phenomena of up-to-date controlling, planning, risk management and financial management. ([7] Nagy, Z. I. 2011). With the help of such modern elements, football clubs start building up a specialised branch, up-to-date and complex from an economic viewpoint, in which the most important requirement will be to secure the requisites of licence until the end of the championship. Here, liquidity reserves have high priority importance, while the analysis of indebtedness is relegated to the background even in the course of judging the licence requirements. In accordance with the regulations of UEFA, any possible insolvency must be prevented, for many years, in each country, through making very high requirements, during the season and championship. These strict regulations pose a challenge even for the most highly developed football countries; thus, for example such regulations are not entirely enforced even in Germany. Thus, the equality of financial chances highlighted by UEFA is realised only to a restricted extent in respect of all European football clubs. Based on the always expanding business activity of various football clubs, the permanent constraint of adapting themselves to the requirements of global economy generates more and more significant financial and economic restrictions against football clubs. Thus it is to be stressed that the Basel II Directive considerably influences the funding of football clubs on the side of credit institutions, although professional football is impacted in this process only indirectly. At the same time, the evaluation and examination of credit standing and the credit rating have become highly important for business enterprises operating in professional football. Of course only those football clubs functioning with permanently positive results can meet such an acute challenge, since they are capable to improve their credit rating on a regular basis. In contrast, the football clubs having an unfavourable credit rating might get into very difficult circumstances in consequence of such a credit rating. This may even result in their getting no credit at all. For example, the strict regulations do not cause any problem for the football clubs with licence; in their case an essential improvement of credit rating can be found on a continuous basis, which is a true index of economic consolidation.

2 Importance of the Legal Form of Football Clubs

In line with the international outlook I wish to rely primarily on the peculiarities of German professional football, since it shows most of all the presumable mainstream of the development of football. Let us overview which are the forms of undertaking in the German League football, and what the changes in trends are.

On the basis of the 1998 Resolution of the German Federal Parliament, it is possible to operate football undertakings also in the form of an incorporated firm [which means that the members are not responsible for the liabilities of the company; these are share companies, limited liability companies, and limited share partnerships], in addition to the classical form of association ([3] Bundesliga Offizielles WEB-Seite 2011). The main point here is that such part of the association participating in League football can be outsourced into the form of incorporated firm. It is interesting that in Germany 21 teams from League I and 36 teams from League II are operating currently in the form of incorporated firms.

The change in legal forms, i.e. the transformation from the traditional form of association into an incorporated firm, has accelerated and has become prominent over the past number of years. Nevertheless, registered associations continue to play an important part in the Leagues, since this form is functioning even directly in the Leagues, but it continues to remain in the foreground also as parent company of incorporated firms.

Legal Form	Frequency	Legal Form	Frequency
Association	15	LLC with right	5
Share comp.	2	LLC & Co.LP.	10
LLC	4	I.-II Liga , **Total**	**36**

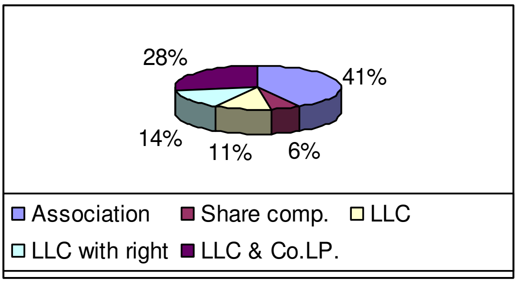


Figure 1

Distribution of legal status in Germany (I and II. League clubs)

Source: [5] Dworak, A.: Finanzierung für Fußballunternehmen, Erick Schmidt Verlag 2010, p. 52

3 Advantages and Disadvantages of Registered Association

The determining element of this legal form is the public purpose (non-profit goal) even in Germany. The associations have a public target even in the field of football, and thus they are subject to the effect of the law of associations. The management and the supervisory board are elected by the members of the association at the members' meeting. In Germany, 90,000 associations are functioning with a membership of 24 million, mainly in the fields of mass (recreational) sports/amateur sports. The side of revenues of the annual budget consists mainly of the payments of membership dues, state subsidies and donations, but even business activity may be pursued as an ancillary activity in the interest of increasing the revenues, provided this serves the public purpose. In Germany, associations were able to manage their finances at an average of EUR 49,000 in the past years at the annual level; however, half of the sports associations were operating with a budget of EUR 14,000 on average [7] Nagy, Z.

I., 2011). The more important clubs manage a significantly higher budget, and this results, in the case of small clubs, in a considerably negative deviation from the average and in the high frequency of such deviation.

Funding, registering and operating an association is relatively simple, and the rules are transparent. All strata of society like to make use of the advantages of this form of non-profit company having legal personality, which provides tax benefits. The registered associations fill such an important social role in Germany that they are exempted, in respect of the public purpose activity, from corporate income tax, industrial tax, real estate and property tax, and in addition, they are only charged by a beneficial 7% general turnover tax (VAT). Otherwise, the general rate of VAT is 19%. However, these tax benefits touch professional football undertakings to only a minor extent as these professional associations may obtain only a small amount of revenues exempt from taxation. The tax benefit helps mostly amateur and mass sports ([4] Bundesligareport (2011).

It is a disadvantage that the members of the association may exercise rights of control to a limited extent only, since the management is only obliged to satisfy the members' need for information essentially at the members' meetings ([8] Schmid, V. Unternehmensführung im Profifußball). Also the depth of information and orientation is insufficient. Apart from the above, the members have practically no possibility for obtaining information. The associations' supply of business information may not be compared even remotely with the requirements governing incorporated firms. In the event of a bad impression obtained by a member of the association at the members' meeting in connection with business management, he will have first of all the possibility of withdrawal, and, if the dissatisfaction of members becomes overwhelming, the issue of winding up of the association may be raised. However, the member of an association will never reacquire, in any of the cases above, his or her paid-in contributions, and thus the only possibility remaining for him or her is to trust blindly in the work of the association's management.

It is a further deficiency in the operation of the form of association that there are no regulations for the application of profits. Consequently, the earned profits are spent on increasing the value of the roster of players in most of the football clubs, and they pay much less attention to forming financial reserves, although this could be very useful later on from the point of view of long-term operations.

The rules relating to associations do not provide satisfactory protection even for lenders. Associations are answerable with their own assets for their liabilities, but there are no detailed legal prescriptions in the relevant law of associations in Germany. Of course, in principle, regulations may be set up relating to publicity, application of profits and collateralising the loans in the statutes of associations, which is, however, scarcely feasible. We can hardly suppose that the elected management of an association would set up restrictions on itself. Logically, all this has a negative impact on the opportunities for financing with outside capital.

It follows from all this that a registered association may only make use to a restricted extent of any potential possibilities of financing. An increase in equity is allowed primarily through the retention of profits, but we must add that the profits may only be increased up to a certain limit per year, in order to maintain the non-profit status of the company. The hosting of shares and the involvement of contributed capital may only be secured after transformation into an incorporated firm. Therefore, in the event of associations, funding by banks or private individuals who are committed to the given association will be possible in line with own revenues and subsidies. The first-mentioned resources may cover current expenditures to a minor extent; they play a part mostly in the financing of long-term projects. Due to the aforementioned lack of publicity and of mechanisms of control and protection, the issuance of bonds or of the so-called participation certificates poses difficulties to associations.

Altogether we can state that the registered association will continue to function as one of the most significant legal forms of professional football in Germany as well, in spite of the fact that the law of associations is far from providing satisfactory conditions for the operations of football clubs carrying out modern business activity. Thus, it is no wonder that in Germany there appears a strong inclination towards the transformation from the form of association into the form of incorporated firm. This does not mean that the clubs (mainly the amateur and mass sports clubs) will turn away from the form of association. Here we should rather mention the trend that the fields and branches, where up-to-date conditions of business management make this necessary, must be subject to the process of transformation (through becoming an incorporated firm via outsourcing from the association). Professional football is such a field.

In Hungary, there is a need for finding the appropriate form of motion in order to secure that vision, and the practice of entrepreneurship should gain ground in spite of the fact that sports leaders, audiences and authorities have grown up in an attitude ruled by the form of association ([1] Bács, Z.-Szima, G. 2012).

4 Registered Association: VfB Stuttgart, Germany (Very Successful Functioning)

This famous club was founded in 1893 and is one of the founders of the German Professional League, where, in 1975, huge financial tensions were generated in consequence of the then president's exaggerated and over-dimensioned concept of investment and purchase of players. As a result, the club dropped out from the First League, but it came back in 1976, and it reached even the fourth place. In parallel with the successful participation in the championship, they succeeded in enlarging also the infrastructure, with a new club centre built by the middle of the 80s.

The club gained the championship four times, in Season 2006/07 for the last time, when the team also won the German Football Cup. In contrast to many professional football clubs, Stuttgart did not build a structure of an incorporated firm, but, essentially, it preserved the form of registered association. Several business areas were outsourced over the past years, but these remained profit centres operating under the direction of the association. The licence right remained with the association.

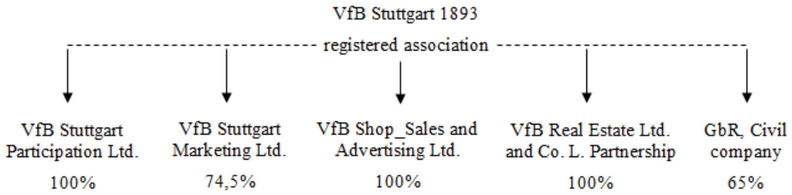

VfB Stuttgart 1893
registered association
VfB Stuttgart Participation Ltd.
100%
VfB Stuttgart Marketing Ltd.
74,5%
VfB Shop_Sales and Advertising Ltd.
100%
VfB Real Estate Ltd. and Co. L. Partnership
100%
GbR, Civil company
65%

Figure 2
Organizational scheme VfB Stuttgart

Source: [8] V. Schmid, Unternehmensführung im Profifußball, Berlin 2004

Proprietary association: VfB Stuttgart Registered Association. The task of VfB Stuttgart Holding LLC is the development of alternative solutions for financing, e.g. the involvement of investors functioning as dormant partners. The task of VfB Stuttgart Marketing LLC is operating the professional football division, and as a compensation it receives an agent's commission from the association, in a manner basically similar to an external agent selling external rights. The task of VfB Shop Sales and Advertisement LLC is organisation and arrangement of merchandising. The task of VfB Property GmbH & Co. KG [limited partnership with an LLC as general partner] is the management of leaseholds and real estate for the association. The GbR (civil law association) is a simple form of undertaking wide-spread in Germany, based on a specific contract of agency.

The traditional form of association as a basic model of management may also be successful in professional football, but it is true that this is realised also thanks to Stuttgart's main sponsor, its exclusive partners and its team partners. The list is deserving of attention: Gazi, Merzedes Benz, Puma EnBW, Kronbacher, BW Bank, Breuninger, Kärcher, and Coca Cola. Therefore, is no wonder that VfB Stuttgart was and is very successful.

Records of VfB Stuttgart:

German championship:

- Winners (5): 1950, 1952, 1984, 1992, 2007
- Runners-up (4): 1935, 1953, 1979, 2003

German Cup:

- Winners (3): 1954, 1958, 1997
- Runners-up (2): 1986, 2007

German Super Cup:

- Winners (1): 1992

German League Cup:

- Runners-up (3): 1997, 1998, 2005

Oberliga Süd I:

- Winners (3): 1945–46, 1951–52, 1953–54
- Runners-up (3): 1949–50, 1952–53, 1955–56

2nd Bundesliga Süd II:

- Winners (1): 1977

Bezirksliga Württemberg-Baden:

- Winners (2): 1926–27, 1929–30
- Runners-up (1): 1925–26

Gauliga Württemberg:

- Winners (4): 1934–35, 1936–37, 1937–38, 1942–43
- Runners-up (4): 1938–39, 1939–40, 1940–41, 1941–42

UEFA Cup:

- Runners-up (1): 1988–89

UEFA Cup Winners' Cup:

- Runners-up (1): 1997–98

UEFA Intertoto Cup:

- Winners (2): 2000, 2002

Accordingly, the corporate philosophy of the association is conservative and successful, if we consider the achievement of sports results to be the main objective. This creates the foundations for success also in the economic field, i.e. a safe economic background must be underpinned by successful sports results. This statement raises the question as to the proper tangible and intangible results of sports.

If we have a closer look, however, we can see that this rather simplified objective cannot be a strategic guideline in this new dimension of modern football, which in recent years has shown extraordinary development and today confronts football enterprises with global challenges ([2] Bott, S. 2007). Naturally, the main objective of football enterprises is also complex, made up of specific systems of sub-goals. A thorough examination and analysis of subordinated goals contribute to determining the ultimate objective itself and provides particular help in achieving this. The aggregate system of objectives in the case of a football enterprise is made up of three components: *sports results, economic results and* *intangible results (value).*These three components can substitute each other only to a limited extent and strongly correlate with each other. By breaking down the objectives that form the basis of enterprises into components, they can be determined and optimised more precisely, as separate calculations can be prepared for each and every result component and each element therein ([1] Bács, Z.-Szima, G. 2012, [11] Borbély, A 2112).

Sports results, which might mean winning international and national titles, qualifying for various international tournaments, avoiding relegation or being promoted to a higher division, are undoubtedly the most important components of the aggregate result, as these are what various factors of both economic and intangible results depend on.

It is clear that an unforeseen wave of injuries could foil the plans to achieve sports results, just as other unexpected factors could. In other words, by maximising one factor we will not necessarily get closer to our goal, as reaching this goal is influenced by a whole series of factors. This means that other short and long term economic result factors must also be taken into account.

Since the improvement of the financial situation usually has a positive effect on aggregate result, we can assume that the management of any given football enterprise (particularly its financial management) wishes to maximise this and will consider the maximisation of profit the focus of entrepreneurial activity ([10] Zieschang, K.– Klimmer, Ch. 2004). Economic results are closely connected to other result factors. Radical changes of the various result factors could cause undesired effects that might in turn lead to an unforeseeable financial crisis. As a result, players' wages might have to be significantly reduced or assets might have to be sold. In the end, these will have a negative effect on the sports results as well. In the case of relegation to a lower division (which is a significant change), the given football enterprise is forced to face serious economic consequences, which can be lightened through prize indemnity insurance. Even a remarkable sports result, such as winning a championship, could become a negative influencing factor if the club pays extraordinarily high bonuses to players for the title, which could overload the enterprise's budget. Naturally, insurance can be taken out for such cases (which of course costs money).

The *intangible result*: the success of a football enterprise entails all value factors that are very difficult to quantify using financial-economic or sports result aspects. This includes the image of the football enterprise, the improvement of which usually results in an increased media presence or the improvement of financial positions or the growth of the market value of the enterprise. Of course this also impacts economic results, as an improved image could lead to increased sponsorship. When speaking of intangible results, we should also mention the supporters' strong identification with the club. This strong relationship has a positive impact not only on economic results, but could also boost the achievement of better sports results.

A football enterprise naturally strives to maximise these intangible results; however, these can only be achieved at the expense of other result factors. Primarily these results could clash with the economic result. If the club's management and owner decide to sign valuable star players and managers regardless of the size of the expenditure, this will obviously have a positive effect on the intangible results and, naturally, sports results as well. It is very likely that in the long-term the enterprise's revenues will also increase. In the mid-term, however, such decisions impose significant financial burdens, which in turn could lead to a serious financial crisis for the club. In professional football – as opposed to other economic fields – it is not customary to prepare a separate plan with respect to image impact for example. Today, the change of image is still an indirect effect of investment. The way of the future is that football enterprises shall directly plan to increase intangible results through certain targeted promotions, e.g. by planning and managing member recruitment, loyalty promotions, etc.

The three main components of the aggregate result and their factors are closely interrelated. For example, the improvement of sports results has a positive impact on economic result, which in turn leads to an increase in television and ticket revenues and the sale of promotional products, but member contributions and revenues from sponsorship deals could also grow. Improving sports results also affects intangible results: the club becomes more respected, supporter ties become stronger and media interest increases. A good economic result has a favourable effect on sports results as the sufficient liquidity provided through profit and appropriate management allows the club to fulfill incentive wages and premium payments without any problems. We should also emphasise that a good economic result has a positive effect on intangible results as well. Opportunities open up to finance supporter projects and clubs as well as significant advertising activity [7] Nagy, Z. I., 2011). A football enterprise acts correctly if it strives to achieve the maximum of the composite result made up of the aforementioned three components.

5 Legal Forms of Football Clubs

In Germany, the relationship of associations with incorporated firms may be deemed to be customary already in professional football, promoting the arrangement of business activity. These so-called marketing limited liability companies are essentially similar to business undertakings and they are in general the subsidiaries of associations. However, the real concept of an incorporated firm of football refers more to the outsourcing of football undertakings and of the fields neighbouring football into incorporated firms, or, respectively, to the transformation thereof into incorporated firms. The surviving association operates as a so-called parent association, and it disposes usually over the majority of votes in the incorporated football firm.

In the event of transformation into or the foundation of an incorporated firm the obtaining operating licence requires compliance also with the requirements made by the League Association, in addition to any other statutory provisions in force. In this respect, first of all the rule of 50%+1 votes must be highlighted in Germany, which must be asserted of course both in the case of the original undertaking and the outsourced marketing company. It is a further important prescription that employees of the business undertaking may not be representatives in the bodies of any other football undertaking. The parent company must dispose over majority representation in the organ of control of the incorporated firm. The right of delegation is within the competence of the parent company. In addition, the statutes of the League Association require from the football companies also a minimum capital amounting to EUR 2.5 million.

However, the election of the legal form of a football association contemplating outsourcing and transformation has not only legal but also economic consequences. Here, particularly, the possibilities of financing should be stressed, which will be enlarged for the football undertakings after the transformation.

In Germany the most frequent incorporated football firms are in the form of the share company, limited liability company or limited partnership, with restricted liability of shareholders.

6 The Importance of the Share Company in Football

This is a company having an independent legal personality, in which the company is answerable for its liabilities with the assets (capital wealth) of the company, and which must function in accordance with the Act on Share Companies also in the field of professional football. Its bodies consist of the board of directors, the supervisory board and the general meeting. Any other regulations are asserted, mutatis mutandis, also in the case of football companies, and thus the rules relating to accounting, publicity and the utilisation of profits. The shareholder's equity may be increased, among other ways, through the surrender and exchange of shares, and, respectively, through the issuance of shares and the increase of capital at the stock exchange or over the counter.

Share companies are advantageous, if there is a great need for capital and if it is possible to involve further financers into joint financing. However, in the course of the acquisition of capital, the majority of votes is in each case an indispensable requirement for the parent association, which can be secured through the issuance of registered shares and/or preference shares with restricted transferability [7] Nagy, Z. I., 2011). In the first case the permission of the management of a share company is also required for the acquisition of shares, while the latter ones do not secure any votes for the shareholders, and their issuance may not exceed the sum of the shares already issued (ordinary shares and shares with restricted transferability). In this way the formation of an undesirable majority may be prevented; but obviously, the involvement of capital will also be narrowed due to the restrictions, since in the event of contribution of further capital required due to the increase of capital or losses the parent association may incur obligations of payment in order to secure the majority, because the licence rights are possessed in this case by the parent association. The League Association strictly controls compliance with each regulation.

7 Successful Football Club: Bayern München Share Company

In the field of German professional football, FC Bayern München is an association, among others, having opted for the form of a share company as a type of incorporated firm. In 2002, the team disposing over the play rights licence, the amateur team No. 1, as well as junior teams A and B and the women's team, were outsourced from the main association, and they were integrated into the marketing subsidiary having been previously founded, and this was then transformed into FC Bayern München AG. Thereby all fields important from the economic viewpoint were integrated into the incorporated firm of football. The non-economic fields and several participations remained with the current company. After that, 10% of shares were surrendered to a so-called strategic partner (Adidas AG) via share exchange.

The association of FC Bayern München, being the most successful football club of Germany and belonging to the first 5 teams of Europe, can be deemed a model regarding both the transformation of professional football undertakings into market-oriented servicing undertakings, and its significant and salient business and sports results. The club exceeded a turnover of EUR 300 million in the championship 2009/10 for the first time. From among the teams of the German Premier League, it was Bayern München to realise the business opportunities hidden in targeted "merchandising", and it created also the required infrastructure. The leaders who were considered professional managers in the 70s had already decided to search for new resources of revenues, e.g. via opening VIP-boxes. In the 80s they increased revenues under the leadership of Hoeneß in the fields of merchandising and advertisement, as well as through sharing in the revenues of the German federal television. They modernised the functioning of the club and set up a new organisational structure. The club was a pioneer in the application of merchandising in Germany, but later on this business activity was outsourced into the Sportwerbe LLC, established for this purpose. For the supervisory board an institutionalised system of control was established. In the board of directors the competencies for trade and sports were strictly separated.

In 2002, with the purpose of establishing an appropriate professional organisation and, last but not least, in order to construct a new stadium, they outsourced the professional football division into the already existing FC Bayern München share company. At that time, the shares were not floated on the stock exchange, but this idea was coming into spotlight more and more (Borussia Dortmund is the single stock exchange company active in the field of football in Germany). Figure 3 shows the ownership structure.

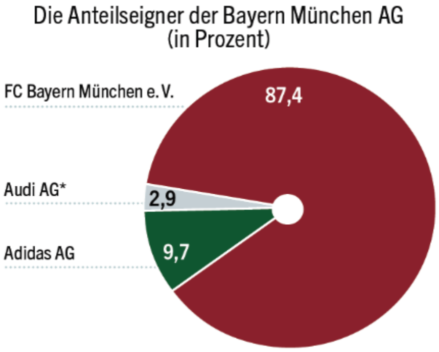


Figure No. 3

Ownership Structure

Source: [6] FC Bayern München WEB site

In accordance with the plans agreed on in advance, Audi AG increased its shareholding in three phases to 9.09% by July 2011, in this way overtaking the shareholding of Adidas AG. 81.2% of the shares continued to be held by the association and the registered capital increased to EUR 27.5 million. The share company is the sole owner of Allianz Arena München Stadion LLC, which constructed and is operating the stadium. TSV 1860 München had been the 50% owner of this LLC until 2006, but it sold its share to FC Bayern München Share Company for EUR 11 million, due to financial problems ([6] FC Bayern München Offizielle WEB-Seite, [7] Nagy, Z. I. 1012).

Records of Bayern München:

- Most Bundesliga titles won: 21
- Most Bundesliga games won (907) and points achieved (3095)
- Most game days on top of the league: 592
- Most average points per game in the Bundesliga: 1.93
- Most Bundesliga goals scored: 3412
- Most consecutive wins in the Bundesliga (19 March–20 September 2005): 15
- Most games won in a club's first Bundesliga season (1965–66): 20
- Biggest lead over second-place finisher (2002–03): 16 points

- Championship with fewest points under the 3-point rule (2000–01): 63
- Championship with the most losses in a season (2000–01): 9
- As a negative record, Bayern's match in Dortmund in the 2000–01 season was the most unfair match in Bundesliga history, with 15 cards having been shown (10 yellow, 1 yellow-red, 2 red), and of those, 12 (8, 1, 1) were shown to Bayern players, which is also a record in Bundesliga history.
- Most championships won: 22
- Most cups won: 15
- Most league cups won: 6
- Most doubles won: 8
- Only club to win the double (Bundesliga and cup) twice in a row (2005–06 and 2006–07)
- Only club to win a championship with the men's and women's football department
- Fastest goal in Champions League history: After 10 seconds by Roy Makaay on 7 March 2007 against Real Madrid
- Last club to win the Champions League/European Cup three times in a row: 1974–76
- Highest aggregate win in the UEFA Champions League knockout stage: 12–1 on 24 February 2009 (5–0) and 11 March 2009 (7–1) against Sporting CP

8 The Importance of the Limited Liability Company in Football

The LLC is a preferred form of incorporated firm having legal personality in the field of professional football in Germany, whose regulations are set out in the LLC Act and in commercial law. The registered capital of an LLC is composed of business quotas and this must total at least EUR 25,000. Any other rules are asserted in the same manner as in the case of other LLC's. (In accordance with the prescription of the League Association, the minimum capital amounts to EUR 2.5 million in football!) The equity may be raised through the utilisation of annual profits (in keeping with the restriction of profits linked to a non-profit company, as well as through the increase of the primary stakes and the involvement of new quota holders. The form of LLC is much preferred in other sports, and thus it is wide-spread in the fields of handball and ice hockey. In spite of the fact that the LLC is applied in two main forms, it does not have such an important role in football. One of these forms is a football LLC, which covers all undertakings linked to football and operates as a subsidiary of the parent company with no affection for business. The other form is the LLC having the right of participation in the championship, in which case the direction by the parent company is also asserted; but in this case the full right of play (licence) has been assigned to the subsidiary. However, all other fields (marketing, merchandising, etc.) remain with the parent company, or potentially they are outsourced into other companies. The advantage of the second form of LLC is that the permission of play (licence) will be only applied for by the company to use the right of play in the future, and thus it will not be the whole parent company to be forced to comply with the strict licence regulations, in which case such a company may be in a much worse situation than the LLC outsourced for professional football. Let us mention as an example Bayer 04 Leverkusen Fußball – GmbH and Vfl. Wolfsburg-Fußball – GmbH. In the case of the first club Bayer AG is the 100% shareholder, while in the other club Volkswagen AG has a 90% share. The transformation served in both cases the connection of the football clubs to the parent companies, since in the case of both clubs a long-term linkage to the aforementioned reputable concerns can be experienced. Substantially, these football clubs developed and functioned in parallel with the aforementioned undertakings, and they emerged from their respective corporate sports clubs into League Clubs. This special link between a football club and an undertaking is expressly recognised also by the statutes of the League Association. Accordingly, the presidency of a League Association may resolve outsourcing with 50%+1 votes if given the business undertaking (in our case the aforementioned concerns) had been supporting their branch of football for more than 20 years prior to 1 January 1999 without interruption and to a considerable extent. This support existed in both cases, and in the case of Bayer Leverkusen this had been prevailing already as of 1904 (!).

Notwithstanding the aforementioned special cases there is also a more commercialised and frequent form of football LLCs, which can be resolved by a parent association as majority owner. This solution means essentially the establishment of a subsidiary LLC. An example for this is Borussia Mönchengladbach. This solution may be contemplated if the football undertaking waives the solution of procurement of capital through the stock exchange. Here the number of potential capital owners may be restricted to a predetermined scope of owners, and in this way it is possible to comply with the prescription relating to 50%+1 votes.

9 A Highly Efficient Legal Form, BVB Dortmund in Germany

The BVB Borussia Dortmund is a successful limited partnership with shareholder liability. The club was founded by a junior group of 18 of the Catholic Holy Trinity Congregation in 1909. Ball Society, Borussia (BVB), became slowly a modern, professionally managed football club from an amateur football society of peasants and workers.

Nowadays it is one of the most successful clubs in Germany. It has won the German Championship six times, and it has been the winner of the German Cup two times. In 1968 it has won the Cup of Cup Winners, and in 1997 it was the winner of the Champions League. After that the club decided to transform most parts of the association participating in League I into an incorporated firm quoted also on the stock exchange. As a first step they outsourced in 1999 the business fields, and they founded Borussia Dortmund GmbH & Co. KG incorporated firm operating through the issuance of shares. Borussia Dortmund Management LLC became the general partner, which was the fully-owned subsidiary of the association. They secured their position in the League Association through the GmbH & Co. KG in such a manner that the association continued to exercise control over the licence division. In 2002 to the float onto the stock exchange was implemented under the leadership of Deutsche Bank AG. Thus Borussia Dortmund is the single German club admitted to the stock exchange. The price of issuance of the shares was around EUR 11, and the issuance covered 13.5 million shares, which generated revenues amounting to EUR 131 million after the deduction of direct costs amounting to EUR 7.5 million. In addition, other ancillary costs which are difficult to quantify were also been incurred.

At the beginning Borussia played successfully in the championship started in the year of transformation (1999/2000), but later on it almost dropped out for the second time from League I. In order to avoid dropping out they made decisions regarding the football club entailing changes in the staff including the hiring of a coach having achieved successes earlier, as well as further old and committed players. By the end of the championship, the team reached the 11th place. After that the club started intensely purchasing players. They acquired star players through transfer for amounts of double-digit millions in each case. As a result, the team won the Championship for the 6th time in 2002. A decline occurred both in financial stability and the sports results, primarily due to exaggerate purchase of players. The League division operated with a loss of more than EUR 65 million in Season 2003/4 ([1] Bács, Z. - Szima, G., 2012 analyzed the revenues of DVSC (Hungarian club)). The overall debts amounted to EUR 118 million. At that time the GmbH & Co. KG tried to improve its financial situation through the sale of key players (resulting in significant criticism from the management of the Association), and it sold Westfalenstadion to Commerzbank AG, which had still been 75% owned upon introduction to the stock exchange. The shares fell deeply below the price of issuance and the experts did not see any hope for an increase in the price even in the long term.

Based on the demand of investors a new management was elected both to the top of the Association and that of the GmbH & Co. KG. This latter office was vested in the treasurer of the Association. The debts amounted to EUR 98 million in total, threatening seriously the mere existence of Borussia. The restoration was successful. Increases of capital were made several times, the stadium sold was repurchased, all this with the strong support by Morgan Stanley, an American investment bank. The net amount of liabilities was reduced to EUR 27 million, and, simultaneously, the proportion of equity increased from 20.7% to 34.5% (percentage of equity / balance sheet total).

Under the influence of the experience obtained during the most recent crisis the management of the association – having learnt the consequences of earlier mistakes – reduced the budget of the League team, contributing thereby to the stability of the financial situation ([9] Stöhr, M. 2010). They replaced the old star players transferred abroad in the meantime with young players developed in Borussia's own association. After that, the team played in the medium field of the championship, up to the season 2008/09, when a new coach was engaged. The club rose to the 6th place in the First League, and, in the following Championship, it was the 5th best team, and, as a surprise, the team won the German Championship of 2010/11.

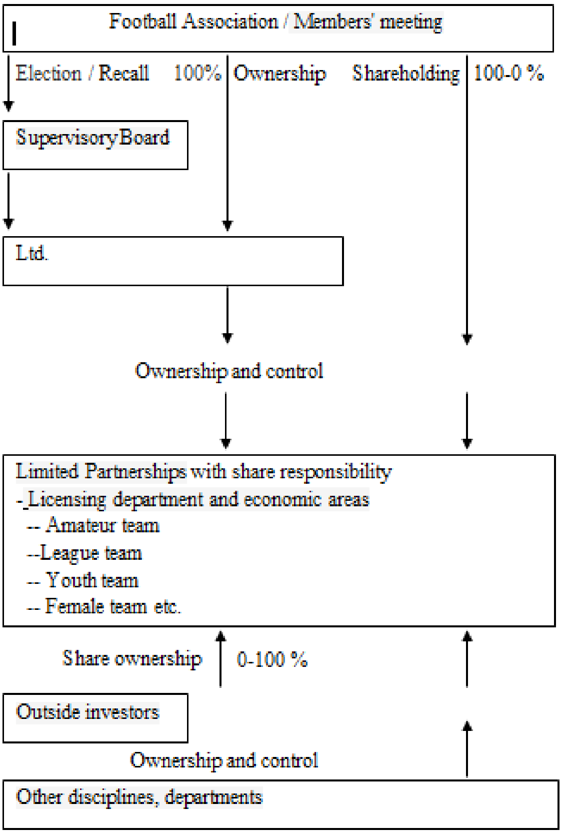


Figure 4

Organisational Schema "Limited Partnerships with share responsibility"

Source: [5] A. Dworak : Finanzierung für Fußballunternehmen

Conclusions

In conclusion, the best solution is a Limited Partnership with Restricted Liability of Partners. This is a company having legal personality, in which, as is well-known, at least one of the partners (the general partner) must bear unlimited and joint and several liability vis-à-vis the creditors. The other owners, who have otherwise interest in the registered capital embodied by the share capital, need not warrant for the liabilities of the company (dormant partners). This is an intermediate legal form that shows the marks of a share company and the personal liability existing in the case of limited partnerships, where the general partner provides for the management of the undertaking. The supervisory board elected by the general meeting of the dormant partners supervises the work of the general partner acting with personal liability, on the one hand, and executes the resolutions of the general meeting, on the other hand, which includes the dormant partners' declaration of acceptance and agreement in the case of more important decisions. Within the scope of legal forms such solutions are interesting for football clubs, since they are vested with the marks of restricted liability covering the general partner of a limited partnership, and at the same time they open a lot of alternative opportunities for finding equity. This limited partnership with restricted shareholders' liability also exists in the case of the forms of GmbH & Co. KG.

In the case of the legal form operating with the limited liability of a limited partnership, the parent company as general partner may not restrict its personal liability. Thus the association should act simultaneously as general partner and dormant partner if it wishes to retain the aforementioned limited partnership directly as its own property, which is not at all permitted from the legal aspect. Therefore, it is required to insert a limited liability company between the existing association and the limited partnership. This subsidiary LLC is a personally answerable general partner, and it functions as managing director at the same time. The shareholders as dormant partners may consist of the parent company itself and of external investors. This has the advantage, on the one hand, that the parent company is secured by an LLC of limited liability, and, on the other hand, that the possibility exists to increase the dormant partners' share in the limited partnership up to almost 100%, without violating the rules of the League Association relating to majority share. In Germany, 10 GmbH & Co. KG companies were operating in 2009 in the Licence Leagues (League 1 and 2). See Figure No. 1, page 2.

This legal form offers a reasonable solution for the football clubs for outsourcing the respective fields of licence and business from the form of association into incorporated firms.

Should a football GmbH & Co. KG be floated to the stock exchange, it must tolerate a significant handicap as opposed to the companies operating in the form of share company, if it has issued also preference shares without voting right. It is a further problem that the possibility of controlling efficiently the leading bodies of an incorporated firm is missing also in the form of football GmbH & Co. KG for the potential investors; they must satisfy themselves with the role of dormant partner if they are not members of the parent association. Unfortunately, this may repel potential investors or, at least, it may be taken into account as a significant negative factor upon the valuation of the participation. It is very important in this solution that separation between the rights of general partners and dormant partners should be complied with strictly in practice. It occurs that potential investors frequently link their possible commitments to being involved into the direction and operation of the undertaking, which will of course appear understandable from the aspect of taking care for the safety of investment. In the event of a football GmbH & Co. KG, any influence by investors may only be implemented within certain limits. By the way, this is shown by the objection made by the League Association against an investor agreement entered into by TSV 1860 München in the 2009 spring season. The investor agreement contained decisions in respect of persons as conditions for contract, which would have for consequence illicit influence on the side of dormant partner, and/or it would have secured the right of veto in respect of the decisions of the associations.

We can finally state that the outsourcing of significant branches of a football undertaking into the form of an incorporated firm should be deemed as a considerable step forward in comparison with the deficient regulations of the operations of associations, regarding the economic professionality of football clubs, partially from the side of commercial law and partially from the side of shareholding law. Through the introduction of organisational and legal structures that are necessary for the football clubs to be considered professional, there are further positive effects generated for football companies: for example, the football undertaking has to reevaluate its assets due to the transformation, which usually has a positive effect on the proportion of equity. This is still complemented by the fact that in the event of incorporation, the equity will increase through the accumulation of profits based on the compulsory regulations of the application of profits. Incorporation allows external investors to get involved in football undertakings, which may be implemented through investment into the respective business quotas of an LLC or a share company or a limited partnership. Here we must stress above all the obligation of publicity and information which makes the incorporated firms transparent for the potential investors of debt capital, and allow for reliable credit ratings and risk assessments. Furthermore, it is also advantageous that incorporated football firms operate similarly to incorporated firms functioning in other fields, from the point of view of the protection of creditors. Finally, a large palette of alternative financing is available for incorporated firms which which are not available in the event of the original legal form and organisational structure of associations. Nevertheless, it should be underlined that in the event of the use of the form of the incorporated football firm, it is about the transformation and change of the organisational unit of execution and not about a total abandonment of the form of association. The form of association will continue to be the main form for football undertakings and a fundamental solution for the procurement of capital. Due to the reasons above, professional football is partially an exemption in this respect.

References

[1] Bács, Z. - Szima, G.: Values of the Club Building (and What is Behind It). International Conference on Tourism and Spoertmanagement. Universitas Debreceniensis. 2012, pp. 1-12

[2] Bott, S.: Internationalisierungsstrategien von nationalen Fußballigen, 1. Auflage, Books on Demand GmbH, Norderstedt. 2007, pp. 1-143

[3] Bundesliga Offizielles WEB-Seite www.bundesliga.de, 2011

[4] Bundesligareport 2011: http://www.bundesliga.de/media/native/autosync/dfl_bl_report_2011_fin_150dpi_deutsch.pdf

[5] Dworak, A.: Finanzierung für Fußballunternehmen, Erich Schmidt Verlag, Berlin 2010, pp. 1-387

[6] FC Bayern München, WEB-Seite 2011

[7] Nagy, Z. I.: Finances of the Professional Football Enterpises. Journal Article, Danube/Law and Economics Review. Volume 3, Issue Number 1, 2012. 03. 31. pp. 53-69

[8] Schmid, V.: Unternehmensführung im Profifußball, Erich Schmidt Verlag Berlin 2004

[9] Stöhr, M.: Unternehmensführung im deutschen Profifußball, Seminararbeit bei Dr. Nagy, Z. I. Obudai Egyetem, Budapest 2010, pp. 1-26

[10] Zieschang, K. / Klimmer, Ch. (Hrsg.): Unternehmensführung im Profifußball, Erich Schmidt Verlag, Berlin 2004, pp. 1-229

[11] Borbély, A.: Országos Sporttudományi Kongresszus, Magyar Sporttudományi Szemle, 2012/2. Budapest 2012, pp. 6-12

A Multi-Variable LQG Controller-based Robust Control Strategy Applied to an Advanced Static VAR Compensator

Ali Tahri, Houari Merabet Boulouiha, Ahmed Allali and Tahri Fatima

Electrical Engineering Faculty, Electrotechnics Department, Laboratory of Electrical Engineering of Oran,

University of Sciences and Technology of Oran USTOMB, BP 1505 El Mnaouar 31000 Oran, ALGERIA

E-mail: ali.tahri@univ-usto.dz, houari.merabet@univ-usto.dz, ahmed.allali@univ-usto.dz, fatima.tahri@univ-usto.dz

Abstract: In this paper, an Advanced Static VAR Compensator (ASVC) based on a voltage type inverter is presented and analysed. The ASVC, which uses the Pulse Width Modulation (PWM) technique, will be modelled in the Park system. Moreover, a low pass filter is introduced at the output of the inverter in order to eliminate the high order harmonics. The proposed system may compensate inductive or capacitive reactive power. The mathematical model thus obtained will be used for the design of a linear quadratic Gaussian (LQG) multivariable (MIMO: Multiple Input Multiple Output) control strategy. The simulation results obtained with MATLAB software are included and will be thoroughly discussed in this paper.

Keywords: LQG controller; multi-variable control; VAR; Programmed PWM; Low pass filter; ASVC; Matlab; Simulink; SimPowerSystems toolbox

1 Introduction

Reactive power is known to be responsible for frequency changes in the voltage of the network and a weakening of the power factor.

The increase in the absorption of this reactive energy creates disturbances on the transmission line.

In an ideal AC system, the voltage at each nodal point should be constant and free of harmonics, with a power factor close to unity; however these parameters must be independent of the size and characteristics of the load of the consumers. This can only be guaranteed if the loads are equipped with a means of reactive power compensation, and it is a must that this has very fast dynamic compensation. Various classical reactive power compensators have been in use for years, particularly those based on thyristors, such as TCR, TSC and SVC.

The TCR compensator (Thyristor Controlled Reactor) is composed of an inductance connected with a Dual switch made up of two thyristors in antiparallel. Each thyristor is controlled every half cycle. The control of the reactance depends on the load demand by varying the instant of the switching of thyristors [1]. The circulating current in the TCR is thus highly inductive due to the presence of the reactance; i.e. the compensator can only absorb the reactive power [2] – [6].

The TSC Compensators (Thyristor Switched Reactor) are generally composed of dual thyristors, each of which controls a capacitor bank. With the presence of capacitors, the current always leads the voltage, which will undoubtedly result in a flow of reactive power to the network [7].

The combination of TCR and TSC devices result in a hybrid compensator called an SVC (Static Var Compensator), playing on the thyristor firing instant to obtain an equivalent of a continuously variable reactive power source [2] [8] [9]. However, this type of compensators exhibits a major drawback, that of being a source of harmonics [6].

Taking advantage of advances in the field of power electronics, devices of self-commutation have been adopted, such as the Advanced Static Var Compensator (ASVC), which is based on a voltage source inverter. This compensator is very flexible to any fluctuation or changes of the loads; it can either deliver or absorb reactive power instantaneously to improve the stability and dynamic performances of the grid network.

The signals from the states of the switches of the three-phase inverter are obtained through the use of a programmed PWM technique. To improve the quality of energy, we have added a resonant filter based on a low pass filter RLC.

The synthesis of the controller LQG (Linear Quadratic Gaussian) multi-variable is based on a mathematical model of multi-state variable of the ASVC for simultaneously controlling the voltage in the DC side and the reactive power.

The control strategy applied to this type of compensator will be validated by simulation results using MATLAB software

2 Main Circuit Configuration of the ASVC

The proposed circuit of the ASVC (Advanced Static Var Compensator) using a programmed PWM technique for DC-AC conversion is illustrated in Figure 1.

Figure 1 shows the different blocks comprising the circuit configuration of the multi-variable control system of the ASVC compensator. Measurements of the currents of Park in both d and q axes and the DC voltage will be added to a white noise (this noise is generated by a white noise generator).

Fluctuations in the DC side voltage and the reactive power may be possible because of the multi-variable controller, where the output could provide two outputs. The first output is the phase α which is added to ωt. The second output is the conversion ratio between the inverter fundamental output voltage and input DC voltage (D). With the latter, we can calculate the modulation index (MI) by the following equation:

$$MI = \sqrt{\frac{2}{3}} D \tag{1}$$

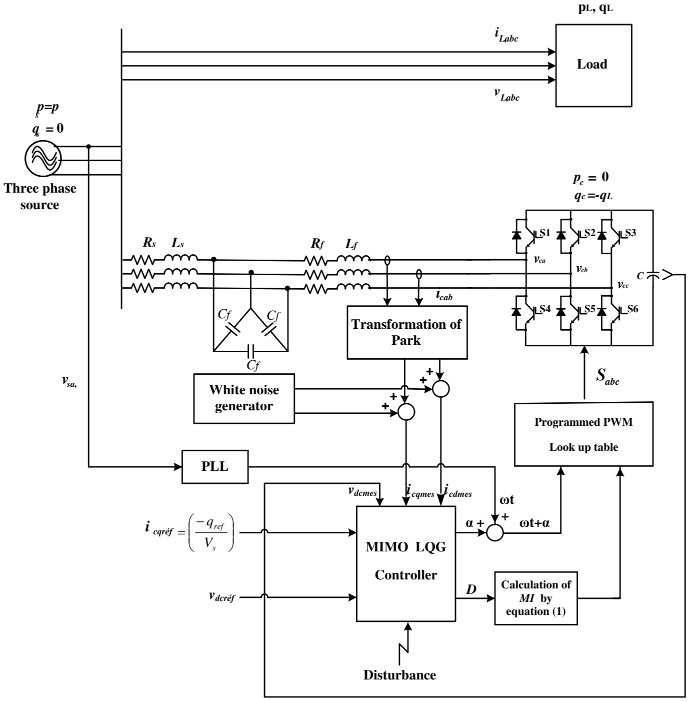


Figure 1
Multi-variable control of the ASVC

3 Programmed PWM

3.1 The Main Objectives of a PWM

We obtain load currents whose waveform is close to a sine wave by controlling the evolution of the duty cycles and thanks to a high switching frequency of the devices with respect to frequency of the output voltages.

To allow fine control of the amplitude of the fundamental output voltages generally to a large extent and for a widely variable output frequency.

3.2 Generation of the Programmed PWM Look Up Table

This PWM technique is used to calculate the switching instants of the devices so as to match certain criteria on the frequency spectrum of the resulting wave. These sequences are then stored and used cyclically to ensure the control of the switches.

The criteria commonly used are:

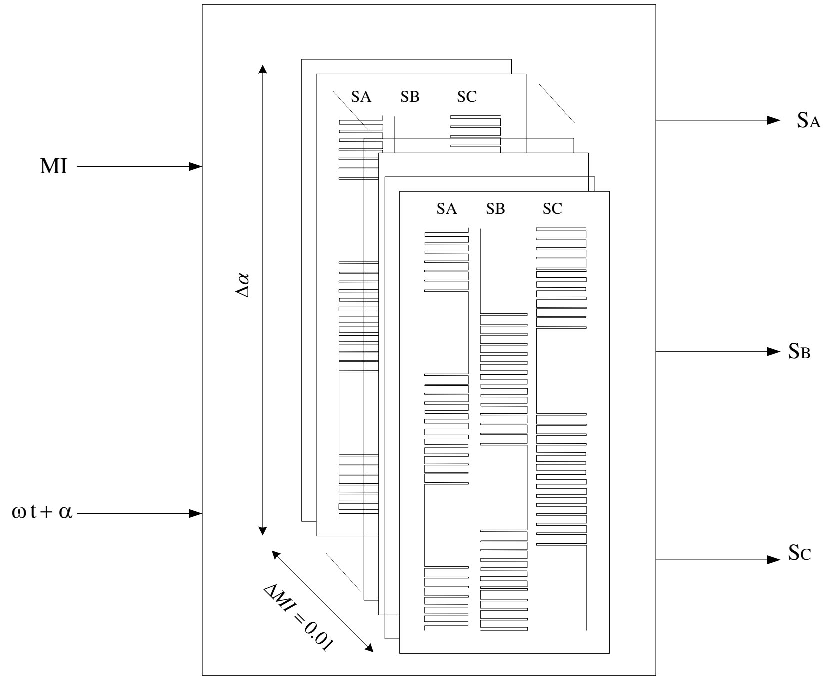
- Elimination of harmonics of a certain range.
- Elimination of harmonics in a specified frequency band.

In this paper, we have used the tools of Matlab/ Simulink in order to implement this programmed PWM [10] technique for generating control signals for switches as there is not any block in the SimPowerSystems toolbox.

The modulation index is varied from 0 to 1.15 in steps of 0.01 (that is to say 115 is the value of the modulation index). Each modulation index is shown in a table of ωt for a period of 0.02 sec with a sampling time of 0.02/4096.

The solutions obtained by the Newton-Raphson method to eliminate 10 harmonics for each modulation index is stored in an array of size 4 Kbytes (4 × 1024 = 4096), that is to say 4096 points per cycle.

Then, two counters were designed (one horizontally to detect the modulation index and the other vertically to detect the instant ωt + α and to count from that instant for a period of 0.02 sec in steps of 0.02 / 4096), where the first counter pointer detects the modulation index (MI). Once the modulation index is detected, the pointer of the second counter is attached to the table for the modulation index obtained by the first counter from the time ωt + α, as shown in Figure 2.



Programmed PWM Look up Table

Figure 2
Programmed PWM look-up table generation method

A low pass filter is added to the output of the voltage type inverter in order to eliminate high order harmonics produced by the inverter and which are not eliminated by the programmed PWM.

Figure 3 shows how the programmed PWM is implemented in Matlab/Simulink. It is composed of two major parts, the first part is the counter which itself is made of two counters, one horizontal to detect the modulation index and the other vertical to detect the instant ωt + α, the second part is the PMW look up table.

Figure 4 shows the details of the counter block, where an Mfile was developed and implemented as a MATLAB function.

Finally, Figure 5 shows the generation of the programmed PWM block diagrams where three 2 dimensional direct look up tables SIMULINK blocks are used, each for one phase.

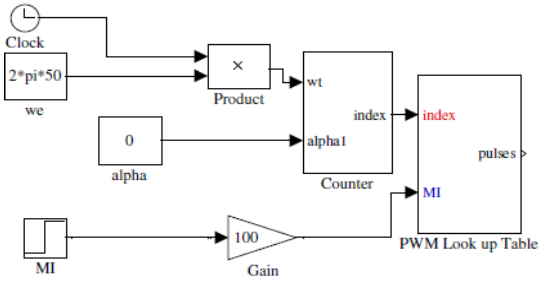


Figure 3
Programmed PWM block in MATLAB/SIMULINK

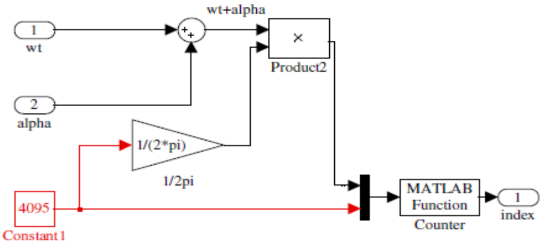


Figure 4
Counter Bloc in MATLAB/SIMULINK

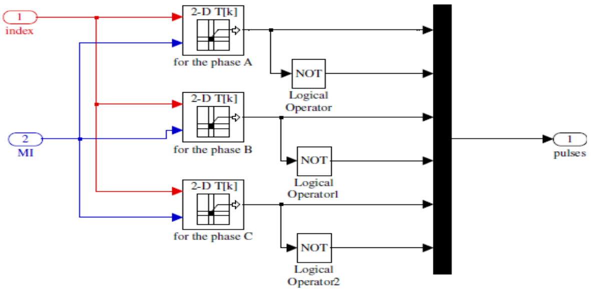


Figure 5
Inverter Pulses Generating Bloc in MATLAB/SIMULINK

4 Modelling and Analysis of the ASVC System

Functions of the Pulses are given by the following equation:

$$S = \begin{bmatrix} S_a \\ S_b \\ S_c \end{bmatrix} = MI \begin{bmatrix} \sin(\omega t + \alpha) \\ \sin\left(\omega t - \frac{2\pi}{3} + \alpha\right) \\ \sin\left(\omega t + \frac{2\pi}{3} + \alpha\right) \end{bmatrix} \tag{2}$$

The mathematical model of the ASVC currents in the axes of Park (d-q) is:

$$\begin{bmatrix} L_s p + R_s & \omega L_s \\ -\omega L_s & L_s p + R_s \end{bmatrix} \begin{bmatrix} i_{cq} \\ i_{cd} \end{bmatrix} = \begin{bmatrix} -V_s \sin\alpha \\ V_s \cos\alpha - Dv_{dc} \end{bmatrix} \tag{3}$$

Where,

V_s : is the effective value of the source voltage.

D: is the conversion ratio of the inverter between the fundamental of the output voltage and the DC input current. α: is the phase angle between the network voltage and the output voltages of the inverter.

p: Laplace Operator.

The equation in the DC side in d-q is given by:

$$\frac{dv_{dc}}{dt} = \frac{D}{C_s} i_{cd} \tag{4}$$

After manipulation, we find the system state multivariable of the ASVC in Park [7], [11]:

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y = Cx(t) \end{cases} \tag{5}$$

With:

$$A=\begin{pmatrix} -\frac{R_s}{L_s} & -\omega & 0 \\ \omega & -\frac{R_s}{L_s} & -\frac{D}{L_s} \\ 0 & \frac{D}{C_s} & 0 \end{pmatrix},\ B=\frac{1}{L_s}\begin{pmatrix} -V_s & 0 \\ 0 & v_{dc0} \\ 0 & 0 \end{pmatrix},\ C=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$x=\begin{bmatrix} i_{cq} & i_{cd} & i_{dc} \end{bmatrix}^{\mathrm{T}},\ u=\begin{bmatrix} \alpha & D \end{bmatrix}^{\mathrm{T}}$$

Where:

A is the state matrix, B is the input matrix, C is the output matrix, x is the state vector, u is the input vector and y is the output vector.

5 Synthesis of the Controller LQG MIMO

The control method LQG consists of an LQR (Linear Quadratic Regulator) with the observer states of the system via the method of the Kalman filter. In the case where the system state (5) is a linear one or linear around an operating point, we can represent it as in the following form:

$$\begin{cases} x(t+1)=Ax(t)+Bu(t)+M\,w(t) \\ y=Cx(t)+v(t) \end{cases} \tag{6}$$

With:

$$M=\begin{bmatrix} 10.124 & 7.33 \\ 2.31 & 5.12 \\ 10.555 & 20.66 \end{bmatrix}$$

Where w(t) and v(t) represent white noise with zero as mean value, independent, respectively, with covariance matrix W and V.

$$E=\begin{bmatrix} w(t) & w^T(t) \end{bmatrix}=\begin{bmatrix} \beta_1^2 \\ \beta_2^2 \end{bmatrix}=W\delta(t)=\begin{bmatrix} 9 \\ 100 \end{bmatrix}$$

$$E=\left[v(t)\quad v^{T}(t)\right]=\begin{bmatrix}1\\1\end{bmatrix}=V\,\delta(t)$$

Where:

E : mathematical esperance.

δ : is the Kronecker Symbol.

The Kalman filter matrices are supposed to be symetrical:

$$Q_f=Q_f^T\geq 0 \text{ and } R_f=R_f^T>0 \tag{7}$$

Where such weighting matrices are given by:

$$\begin{cases}Q_f=MWM^T\\ R_f=V\end{cases} \tag{8}$$

With:

$$\begin{cases}W=\begin{bmatrix}9&0\\0&100\end{bmatrix}\\ V=I_2\end{cases} \tag{9}$$

The estimating gain K_e of Kalman is obtained by the following relation:

$$K_e(t)=P_f(t)C^TR^{-1} \tag{10}$$

Where P_f is the solution matrix of Riccati equation:

$$P_fA+A^{\mathrm{T}}P_f-P_fBR_f^{-1}B^{\mathrm{T}}P_f+Q_f=0 \tag{11}$$

The synthesis of the LQR controller consists then in finding a matrix with gain L and where the optimal feedback control is given by [12]:

$$u_{opt}=-Lx(t) \tag{12}$$

The minimizing quadratic criterion for obtaining the control law (12) is:

$$J_{LQR}=\int_0^{\infty}\left(x^{\mathrm{T}}Qx+u^{\mathrm{T}}Ru\right) \tag{13}$$

Where the weighting matrices Q and R are given by:

$$\begin{cases} Q = C^{\mathrm{T}} C \\ R = \rho I_2 \end{cases} \tag{14}$$

Where:

I_2: is the identity matrix of dimension 2.

The necessary condition for optimum of the derivative being zero of the cost function J_{LQR} leads to the solution:

$$L = R^{-1} B^{\mathrm{T}} P \tag{15}$$

Where P obeys the algebraic equation of Riccati [13]:

$$\dot{P} = -PA - A^{\mathrm{T}} P + PBR^{-1} B^{\mathrm{T}} P - Q \tag{16}$$

Considering the steady state solution of this equation P, which is done in most cases in the literature, it is sufficient to solve the following equation:

$$PA + A^{\mathrm{T}} P - PBR^{-1} B^{\mathrm{T}} P + Q = 0 \tag{17}$$

Figure 6 shows the schematic diagram of the LQG control method, where the term N is calculated in steady state to eliminate the static error by the following equation [14]:

$$N = \left[C \left[I - A + BL \right]^{-1} B \right]^{-1} \tag{18}$$

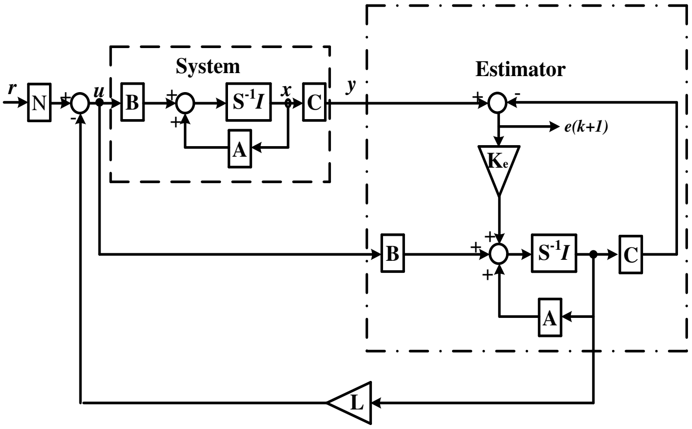


Figure 6
The structure of LQG Controller

6 System Blocs Design in Matlab/Simulink

The proposed system is composed of two main parts: one consists of the control circuit and the second part is concerned with the power side.

The bloc representing the controller (control section) MIMO LQG in Matlab Simulink is shown in Figure 7.

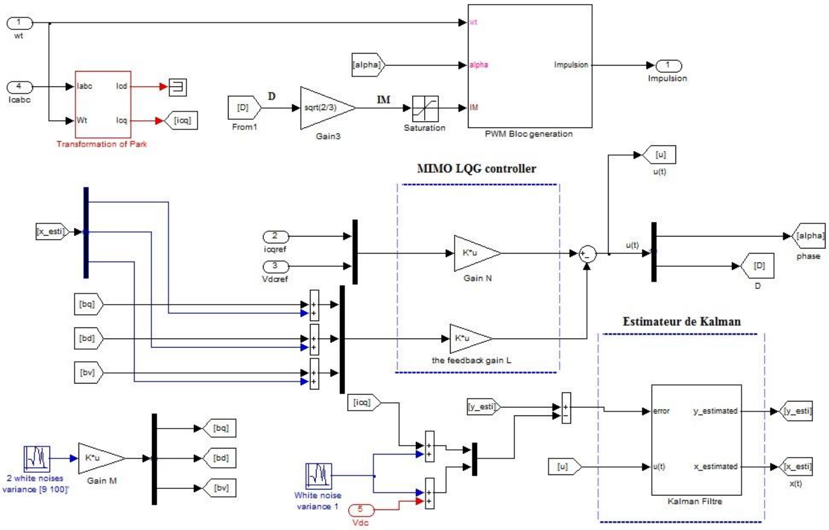


Figure 7
Detailed Circuit Diagram of the LQG MIMO

The simulation model of the ASVC is implemented in Matlab / Simulink as shown in Figure 8 using SimPowerSystems toolbox blocs. However, the ASVC is implemented by the universal bridge block with IGBT switches. In order to synchronize the ASVC with the supply a discrete three phase PLL is used.

In order to simulate the ASVC operating from $10KVAR$ capacitive to $10KVAR$ inductive two 3 phase RLC parallel loads are used and switched on and off with two 3 phase breaker blocks.

To run the simulation quickly, a discrete mode is chosen in powergui block, which is the environment block for SimPowerSystems models. The sample time for the simulation is set to the same sampling time (0.02 / 4096) as calculated for the PWM look up table as given in section 3.2.

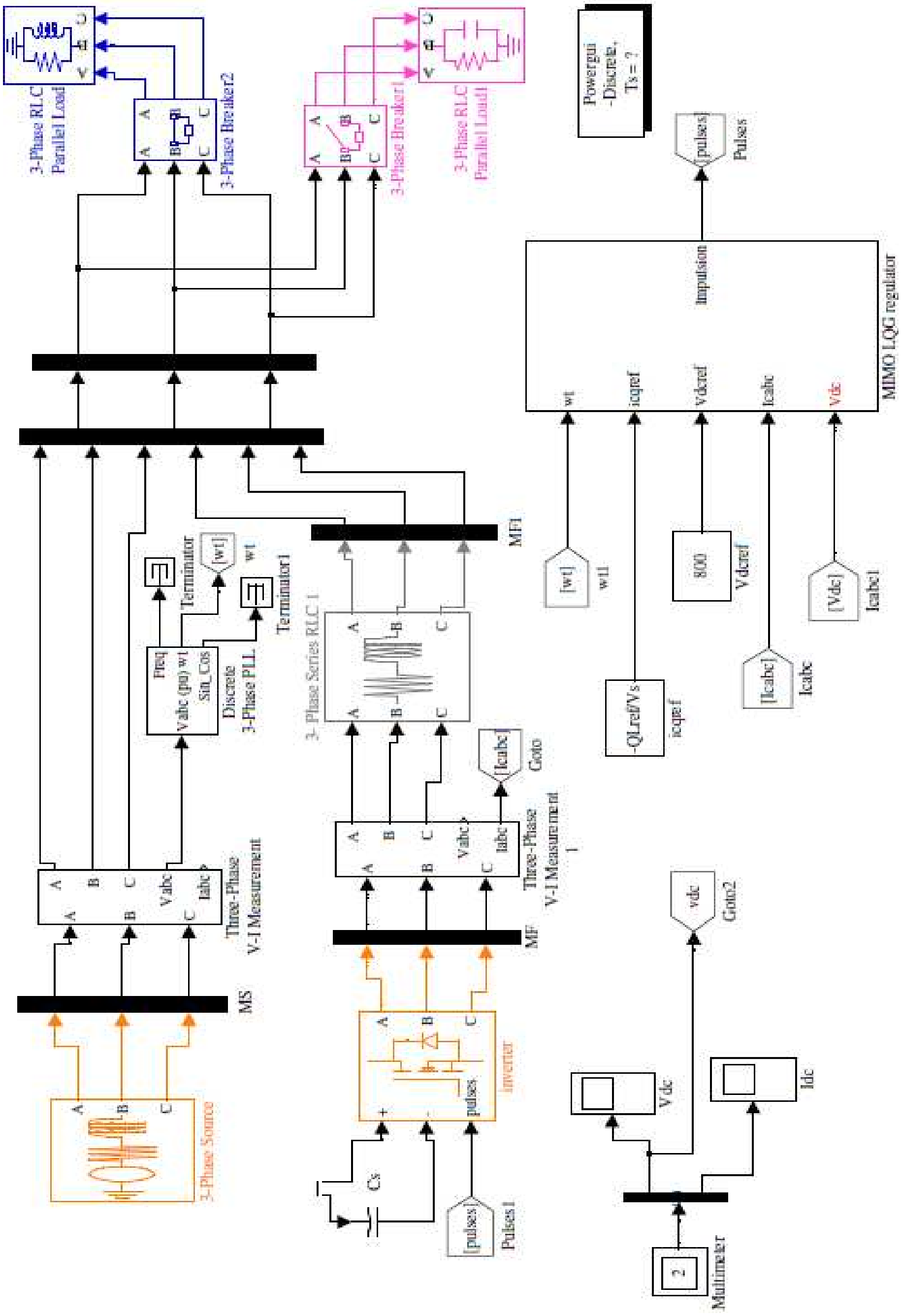


Figure 8
Equivalent Circuit of the ASVC Compensator on Simulink SimPowerSystems toolbox

Block Kalman filter in MATLAB / SIMULINK is shown in the following figure:

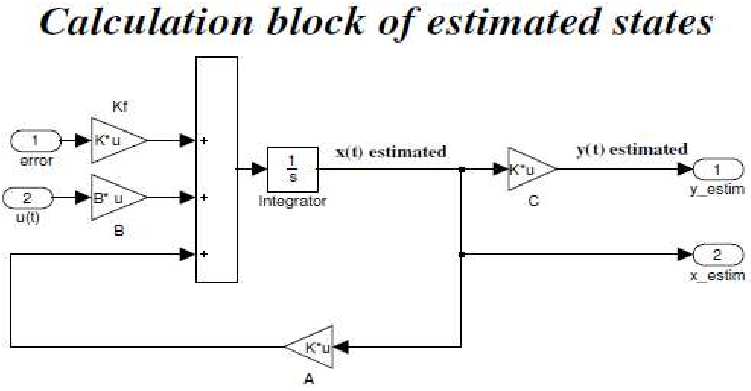


Figure 9

Detailed Circuit Diagram of the Kalman filter

7 RLC Filter Synthesis

A low pass filter has been chosen to eliminate the high order harmonics produced by the inverter which are not eliminated by the programmed PWM technique, where the capacitors of the filter are Delta connected. The transfer function of the filter is given by:

$$G_f\left(j\frac{\omega}{\omega_0}\right)=\frac{1}{1-2L_fC_f\left(\frac{\omega^2}{\omega_0^2}\right)+j2R_fC_f\left(\frac{\omega}{\omega_0}\right)} \tag{19}$$

Where:

ω: is the pulsating frequency.

ω_o: is the particular pulsating frequency.

This function has a dynamic of the second order. By identifying the denominator to the canonical form as follows:

$$\left(j\frac{\omega}{\omega_0}\right)+2\xi\left(\frac{\omega}{\omega_0}\right)+1 \tag{20}$$

After calculation we find:

$$\begin{cases} \omega_0 = \dfrac{1}{\sqrt{2L_f C_f}} \\ \xi = \dfrac{R_f}{\sqrt{2}} \sqrt{\dfrac{C_f}{L_f}} \end{cases} \tag{21}$$

With:

ξ: is the damping coefficient.

By fixing $R_f = 1\Omega$, ω_o=6×ωs (where $\omega_s = 2\times\pi\times 50$) and $\xi = 0.083$. After calculation we obtain:

$$\begin{cases} L_f = \dfrac{R_f}{2\omega_0 \xi} \approx 3.2mH \\ C_f = \dfrac{1}{2\omega_0^2 L_f} \approx 43\mu F \end{cases}$$

We choose the commercially available value, thus the values of the filter are:

$L_f = 3.3mH$ and $C_f = 40\mu F$

8 Simulation Results

The simulation results are obtained by using the following parameters:

Network :

$$R_s = 3.4\Omega,\ L_s = 3.3\, mH$$

$$V_s = 380V,\ \ f = 50\, Hz$$

DC side:

$C_s = 500\mu F$

$D = 0.8\sqrt{\dfrac{3}{2}},\quad v = 800\, V$

Parameters of the LQR MIMO controller:

The weighting matrices:

$$Q=\begin{pmatrix}1&0&0\\0&0&0\\0&0&1\end{pmatrix},\ R=\begin{pmatrix}0.1&0\\0&0.1\end{pmatrix}$$

The gains of the LQR multi-variable controller:

$$L=-\begin{pmatrix}1.77\times10&\text{-}6.6\times10&9.74\times10\\1.042\times10&\text{-}1.421\times10&\text{-}9.85\times10\end{pmatrix}$$

$$N=\begin{pmatrix}9.562\times10^{\text{-}3}&\text{-}2.05\times10^{\text{-}5}\\\text{-}8.25\times10^{\text{-}3}&8.93\times10^{\text{-}5}\end{pmatrix}$$

Parameters of Kalman filter

The weighting matrices:

$$Q_f=\begin{pmatrix}6295.348384&3963.43796&16105.50938\\3963.43796&2669.4649&10797.358445\\16105.50938&10797.358445&43686.232225\end{pmatrix},R_f=I_2$$

The gain of Kalman:

$$K_e=\begin{bmatrix}3559.25&-11252.76\\-233364.46&728641.38\\-11252.76&36072.36\end{bmatrix}$$

Reactive power load varies between $\pm10KVAR$.

Figure 10 shows the waveforms of the load current i_{La}, the compensator current i_{ca} and that of the source i_{sa}, respectively, with respect to the supply voltage v_{sa} for both inductive and capacitive modes.

At the time of the operation of the compensator at t = 0.02 sec, we notice that the supply current becomes in phase with its voltage. At time t = 0.1 sec, a disturbance was applied to the system states, and we notice that despite the disruption the current i_{sa} stays in phase with its voltage v_{sa}. The DC voltage v_{dc}, measured and estimated with its reference and current i_{dc} on the DC side as well as the shape of the supply and compensator voltages and the modulation index MI, are illustrated in Figure 11. Moreover, we noticed that the DC side voltage follows perfectly its reference of 800V before and after the disturbance.

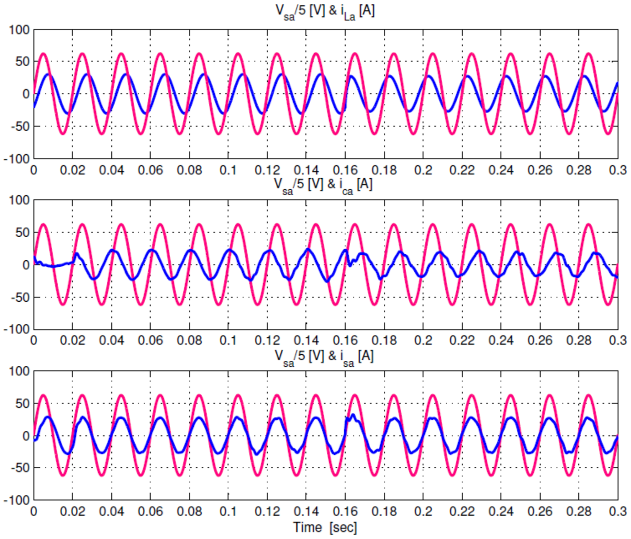


Figure10

Waveforms of the source voltage, respectively, with the load, compensator, and source currents

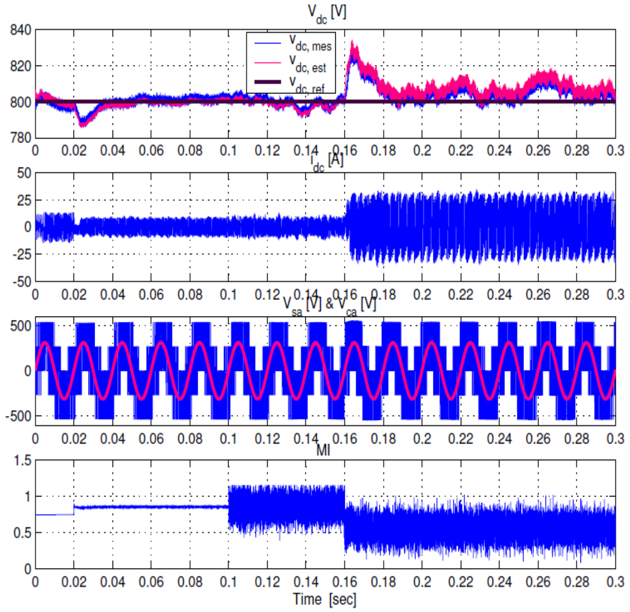


Figure11

Responses of the DC side voltage and current with network and compensator voltage and modulation index MI for v_{dcref}*=800V*

Figure 12 shows from top to bottom the active and reactive power at supply bus bar. We notice that the ASVC from t=0.02 sec to 0.16 sec generates $10KVAR$ capacitive reactive power to compensate for the inductive reactive power absorbed by the load, and from 0.16 sec it absorbs the reactive power generated by the capacitive load.

The estimated and measured reactive currents both follow the reference of the reactive current. Finally, the phase angle α between the network voltage and the output voltages of the inverter is depicted; it is clear that when the ASVC is in capacitive mode α is negative and when the ASVC is in inductive mode it is positive.

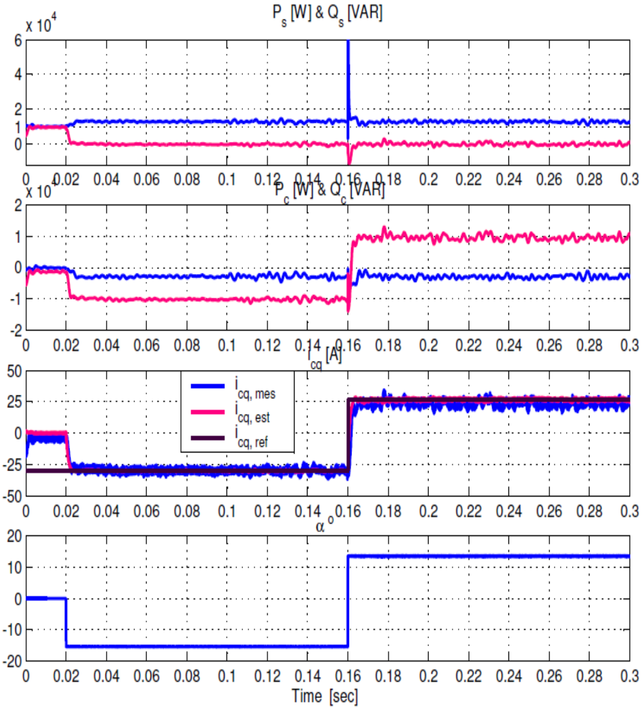


Figure 12

Responses of the active and reactive power respectively of the network and compensator, and the reactive component of the current of the compensator, and control law α

Conclusions

In this paper, an Advanced Static Var Compensator (ASVC) using a voltage type inverter has been thoroughly discussed.

The LQG controller (Linear Quadratic Gaussian) is composed of a regulator, LQR MIMO (Multi Input Multi Output), with an observer of the system states through a Kalman estimator.

The synthesis of the MIMO LQR controller is based on a mathematical model of the multi-variable state of the static reactive power compensator in the dq axes of Park. The synthesis of the Kalman filter is based on the perturbation matrix (M) and by white noise chosen by the manufacturer.

This control strategy exhibits better speed and good filtering of the white noise, with the possibility of controlling simultaneously the reactive component of the compensator current i_{cq} and the DC side voltage v_{dc}..

The switching signals from the semiconductor devices of the three-phase inverter are obtained through the use of a programmed PWM technique.

References

[1] E. Acha, V. Agelidis, O. Anaya, TJE Miller, "Power Electronic Control in Electrical Systems", Newnes, 2002

[2] S. Khalid, N. Vyas, "Applications of Power Electronics in Power System", Laxmi Publications Pvt limited, 2009

[3] G. W. E. Moon, "Predictive Current Control of Distribution Static Compensator for Reactive Power Compensation", IEEE Proc-Gener. Tunsm. Distrib, Vol. I46, N^{o}5,September 1999

[4] E. Acha, C. R. Fuerte-Esquirel, H. Ambriz-Perez, C. Angeles. Camacho, "FACTS Modelling and Simulation in Power Networks", John Wiley and Sons, Ltd. 1004

[5] N. Mohan, T. M. Undeland, W. P. Robbins "Power Electronics, Converters, Application, and Design", Second Edition, John Wiley and Sons, Inc. 1989

[6] Fatima Tahri, "Contribution à la commande d'un compensateur statique d'énergie réactive", Thèse de Magister, USTO, 11 Juillet 2006

[7] Houari Merabet Boulouiha, "Commande adaptative d'un compensateur statique de puissance réactive », Thèse de Magister, USTO, 17 Juin 2009

[8] V. K. Sood, "HVDC and FACTS Controllers Applications of Static Converters in Power Systems", Kluwer Academic Publishers, Boston. 2004

[9] A. Tahri, "Analyse et Développement d'un Compensateur Statique Avancé Pour L'amélioration de la Stabilité transitoire des systèmes Electro-Energétiques", Thèse de Magister, USTO, Juillet 1999

[10] SimPowerSystems User's Guide, the Mathworks Inc 1990

[11] Guk C. Cho, Gu H. Jung, Nam S. Cho and Gyu H. Cho, "Control of Static VAR Compensator (SVC) With DC Voltage Regulation and Fast Dynamique By Feedforward and Feedback loop", Conférence IEEE Conf. Pub, 367-374 Vol. 1, 26th annual meeting, pp. 7803-2730-6, Korea, 18-22 Jun 1995

[12] André. Noth, “Synthèse et Implémentation d’un Controleur pour Micro Hélicoptère à 4 Rotors”, Ecole Polytechnique Fédérale de Lausanne, Institut d’Ingénierie des Systèmes, I2S Autonomous Systems Lab, ASL, Février 2004

[13] Daniel. Alazard, “Régulation LQ/LQG, Notes de cours”

[14] Houari. Merabet Boulouiha, Ali. Tahri et Ahmed. Allali “Commande Multi-Variable d’un Compensateur Statique de Puissance Réactive Par L’utilisation d’un Régulateur Linéaire Quadratique LQR”, $2^{\text{èmes}}$ Journées Internationales d’Electrotechnique, de Maintenance et de Compatibilité Electromagnétique à l’ENSET Oran, Conférence Internationale JIEMCEM’ 2010, 25-27 Mai 2010