On the Use of the Boundary Element Method for Electromagnetic Numerical Dosimetry

L. Zilberti[(1)], O. Bottauscio[(1)], M. Chiampi[(2)]

[(1)]Istituto Nazionale di Ricerca Metrologica, Strada delle Cacce 91, I-10135 Torino, Italy
[(2)]Dip. Ingegneria Elettrica, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy
Phone: +390113919848, fax: +390113919849
e-mail: l.zilberti@inrim.it, o.bottauscio@inrim.it, mario.chiampi@polito.it

Abstract: This paper proposes a summary of a research activity devoted to the development of computational tools useful for the numerical investigation of human exposure to electromagnetic fields, with particular reference to the point of view raised by some standards in force. The subject has been faced for both radiofrequency and low-frequency fields and so we discuss how to set the different problems in order to limit the computational burden. All the numerical approaches that we propose are based on the Boundary Element Method and make use of homogeneous human models. After a general introduction, the paper is subdivided into two main chapters focused on radiofrequency and low-frequency fields respectively. Each chapter includes the description of the adopted formulations as well as some explanatory examples.

Keywords: boundary element method, human exposure to electromagnetic fields, dosimetry

1. Introduction

The protection of human health against possible harmful effects coming from the exposure to electromagnetic fields is a subject faced by some international organizations, such as the International Commission on Non-Ionizing Radiation Protection (ICNIRP) and the Institute of Electrical and Electronics Engineers (IEEE), which, in the last years, established specific exposure limits on the induced quantities (current density and specific absorption rate, SAR) [17-20].

At European level, the ICNIRP Guidelines devoted to time-varying electromagnetic fields have been adopted as the basis for the Recommendation 1999/519/EC [20] addressed to the protection of the general public, as well as for the Directive 2004/40/EC [1] which, in a next future, will make the ICNIRP provisions a legal requirement for the exposure of workers. Owing to their economical and social outcomes, both the Recommendation and the Directive have caused a strong need for reliable assessment procedures. High resolution human models have been developed in the last decade [13-16,22] and can be used together with valuable electromagnetic

solvers in order to investigate exposure situations [8,9,10,2]. Nevertheless, in some technical standards (see for example [5]) such models are still considered to be in the field of scientific research, because they make use of dedicated software, not widely available, which require highly specialized competences. Thus, at present time, the aforementioned techniques are not relevant with regard to standardization objectives, where it is essential to have effective procedures with high efficiency and versatility.

For the above-mentioned reasons, in this paper we propose an overview of some computational tools, numerically handled according to the Boundary Element Method (BEM), that we have conceived to be quite versatile and suitable for the evaluation of the induced quantities during the inspection of standard observance under verisimilar exposure conditions to radiofrequency and low-frequency electromagnetic fields [23]. In particular, thanks to the analysis of the peculiarities related to the different sources/environments, the proposed approaches have been based on selected formulations, specifically studied in order to optimize the computational burden.

For the representation of the human body we will make use of some models, suggested by the relevant standards, having simplified shapes and homogeneous properties. In the light of this latter feature, the choice of BEM is due to the possibility of representing the objects in the considered domain only through their external surfaces; moreover, since the propagation of the field is synthesized by the Green function, BEM does not require a description of the space interposed between the different objects; this fact represents an important advantage from a computational side, mainly when the sources are located quite far from the subject exposed to the field. Finally, inside each region of the domain, the field quantities computed according to the proposed approach are continuous and not bound by the volume meshes typical of other techniques, like the Finite Difference in Time Domain, often adopted to face the complexity of the heterogeneous human models.

2. Modelling of exposure to radiofrequency fields

This section describes a BEM model able to simulate the human exposure to radiofrequency (RF) electromagnetic fields. With the term "radiofrequency", here we will consider approximately the range 100 kHz – 500 MHz. Since the electric and magnetic components of the field are quite strongly linked each other, the formulation is based on the complete set of Maxwell's equations; therefore, in principle nothing prevents its use outside of the RF band, even if, in this case, the solver is not optimized and its performances could appear not completely adequate.

2.1. Formulation for RF problems

In the analysis of human exposure to RF fields, we assume the reasonable hypothesis of propagation through linear media and we work in the frequency domain. It must be noted that this approach does not affect the generality of the problem formulation, because the angular frequency can identify a component of a Fourier series (or a Fourier integral) and it will be possible to apply the superposition successively. Furthermore, if we consider homogeneous and isotropic materials, Maxwell's equations can be written as

$$\nabla \times \boldsymbol{E} = -j\omega\mu\boldsymbol{H} \tag{1}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J}_s + (\sigma + j\omega\varepsilon)\boldsymbol{E} \tag{2}$$

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon} \tag{3}$$

$$\nabla \cdot \boldsymbol{H} = 0 \tag{4}$$

where $\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{J}_s$ and ρ are respectively the electric and magnetic field strength, the impressed current density and the volume charge density. The phasor representation is exploited and so all these field quantities (both scalars and vectors) are complex. The parameters ω, μ, σ and ε are the angular frequency, the magnetic permeability, the electrical conductivity and the dielectric permittivity, respectively; j is the imaginary unit.

If one takes the curl of both equations (1) and (2), with a cross-substitution it is possible to obtain

$$\nabla \times \nabla \times \boldsymbol{E} = \omega^2\mu\left(\varepsilon - j\frac{\sigma}{\omega}\right)\boldsymbol{E} - j\omega\mu\boldsymbol{J}_s = k^2\boldsymbol{E} - j\omega\mu\boldsymbol{J}_s \tag{5}$$

$$\nabla \times \nabla \times \boldsymbol{H} = \omega^2\mu\left(\varepsilon - j\frac{\sigma}{\omega}\right)\boldsymbol{H} + \nabla \times \boldsymbol{J}_s = k^2\boldsymbol{H} + \nabla \times \boldsymbol{J}_s \tag{6}$$

where $k = \omega\sqrt{\mu\left(\varepsilon - j\frac{\sigma}{\omega}\right)}$ is the complex propagation coefficient.

In [21] Stratton and Chu showed how, thanks to the vectorial form of Green's theorem, equations (5) and (6) inside a region Ω bounded by a closed surface $\partial\Omega$ have the following integral solutions

$$\boldsymbol{E} = \int_\Omega \left(\frac{\rho}{\varepsilon}\nabla\Psi - j\omega\mu\Psi\boldsymbol{J}_s\right)dv + \oint_{\partial\Omega}\left[(\boldsymbol{n}\cdot\boldsymbol{E})\nabla\Psi + (\boldsymbol{n}\times\boldsymbol{E})\times\nabla\Psi - j\omega\mu\Psi(\boldsymbol{n}\times\boldsymbol{H})\right]ds \tag{7}$$

$$\boldsymbol{H} = \int_\Omega (\boldsymbol{J}_s \times \nabla\Psi)dv + \oint_{\partial\Omega}\left[j\omega\left(\varepsilon - j\frac{\sigma}{\omega}\right)\Psi(\boldsymbol{n}\times\boldsymbol{E}) + (\boldsymbol{n}\times\boldsymbol{H})\times\nabla\Psi + (\boldsymbol{n}\cdot\boldsymbol{H})\nabla\Psi\right]ds \tag{8}$$

Here $\boldsymbol{n}$ is the normal unit vector directed inwards the volume Ω. If r is the distance between the computation and source points, the appropriate Green function is

$$\Psi = \frac{e^{-jkr}}{4\pi r} \tag{9}$$

Equations (7) and (8) allow the computation of an electromagnetic field directly from a given distribution of sources without the introduction of vector and scalar potential or of a Hertzian vector. Here the surface integrals have a direct interpretation in terms of equivalent sources of charges and currents (including fictitious magnetic sources), while the volume integrals are particular solutions of the corresponding differential equations due to the actual sources. It is therefore clear how the function of the integrals is that of combining together the elementary spherical wave contributions of all the sources to get their total interference in the computational point. If there are no actual sources inside Ω, the remaining surface integrals carry out the "surface equivalent principle" [7].

It must be noted that equation (7) involves the knowledge of the charge density, which is often hard to obtain. However, the continuity equation may be used to replace this term with another one which is proportional to the divergence of $\boldsymbol{J}_s$ [12], so that the only known terms are represented by the impressed current density and its divergence. On the basis of these latter considerations, it is interesting to note that for a divergence-free source (e.g. a current loop) the contribution related to the charge density can be directly omitted.

Now we can set the numerical treatment of equations (7) and (8), that we can call respectively Electric Field Integral Equation, or EFIE, and Magnetic Field Integral Equation, or MFIE. A quite natural choice to solve numerically the proposed formulation seems to be the Boundary Element Method. According to this technique, the surfaces of the computational domain are discretized into a number M of quite small triangular elements; on each single triangle m the field quantities $\boldsymbol{E}$ and $\boldsymbol{H}$ are assumed to be uniform. We start the analysis of the problem considering the simple case of a homogeneous and isotropic test object, radiated by some impressed current densities which are confined inside external sub-regions Ω_s. In both the external and internal volumes, for each i-th element of the mesh, the following discretized versions of the EFIE and MFIE have to be verified

$$\begin{aligned}\xi \boldsymbol{E}_i = &-j\omega\mu\int_{\Omega_s}\Psi \boldsymbol{J}_s dv + \sum_m^M (\boldsymbol{n}\times\boldsymbol{E})_m \times \int_{\partial\Omega_m}\nabla\Psi_{i,m} ds \\ &+ \sum_m^M (\boldsymbol{n}\cdot\boldsymbol{E})_m \int_{\partial\Omega_m}\nabla\Psi_{i,m} ds - j\omega\mu\sum_m^M (\boldsymbol{n}\times\boldsymbol{H})_m \int_{\partial\Omega_m}\Psi_{i,m} ds\end{aligned} \tag{10}$$

$$\begin{aligned}\xi \boldsymbol{H}_i = &\int_{\Omega_s}(\boldsymbol{J}_s\times\nabla\Psi) dv + \sum_m^M (\boldsymbol{n}\times\boldsymbol{H})_m \times \int_{\partial\Omega_m}\nabla\Psi_{i,m} ds \\ &+ \sum_m^M (\boldsymbol{n}\cdot\boldsymbol{H})_m \int_{\partial\Omega_m}\nabla\Psi_{i,m} ds + j\omega\left(\varepsilon - j\frac{\sigma}{\omega}\right)\sum_m^M (\boldsymbol{n}\times\boldsymbol{E})_m \int_{\partial\Omega_m}\Psi_{i,m} ds\end{aligned} \tag{11}$$

In the numerical point of view, the i-th element plays the role of the computational point, while all the m-th elements represent the source points (for this reason it is specified that the Green function is computed taking into account the distance between the i-th element and the considered m-th triangle). In the previous equations the singularity factor ξ, necessary to avoid the singularity arising when $r = 0$, is equal to 0.5. Both the EFIE and the MFIE are vectorial equations and involve six unknowns (one normal and two tangential components for the electric and the magnetic field, for a total of $6M$ unknowns). If we project them on the normal and on the two tangential unit vectors defined for the i-th element, we obtain six scalar relations. Applying this procedure to all the M triangles, we get a linear algebraic system of $6M$ equations with $6M$ unknowns; this means $36M^2$ complex terms, or rather, $72M^2$ real terms. Generally the BEM matrix is full and this feature apparently increases the computational burden with respect to other techniques involving volume meshes (e.g. the Finite Element Method). However, it is important to remember that since only the surfaces in the domain are discretized, the number of matrix elements is not so large in comparison with the methods requiring a volume discretization, especially when the sources are located quite far from the test objects. It is fundamental to note that, in order to take into

account the properties of both the external and internal regions, the six scalar equations coming from the EFIE and MFIE cannot be all written with reference to the same volume (the interior of the body or the external air), but it is necessary to mix the two points of view. To face this problem there are many different choices, but probably the simplest one is to write the EFIE for the internal volume and the MFIE for the air. Seemingly the number of unknowns becomes 12*M*, but actually the classical continuity conditions relating the external/internal field components across the interface can be used to keep their number to 6*M*.

The above described procedure has been implemented in a computational code. Besides the sources located in Ω_s, the code allows the introduction of external impressed fields as sources. In particular, it is possible to define plane waves colliding with the test objects (in this case the volume integrals are directly replaced by the values of the unperturbed fields over the external surfaces of the body). The solution of the BEM matrix provides the field values on the bounding surfaces, which in turn are used to compute the electromagnetic state of the system in any point of the domain. This can be done by exploiting relationships analogous to (10) and (11) where the computational points do not belong to the mesh, so that the singularity factor is not necessary. Finally, the knowledge of the electric field inside the radiated object, of mass density δ, allows the computation of the induced quantities to be checked:

$$\text{“Total” induced current density:} \quad \boldsymbol{J}_T = (\sigma + j\omega\varepsilon)\boldsymbol{E} \tag{12}$$

$$\text{Specific absorption rate:} \quad SAR = \frac{\Re\left[\boldsymbol{E}\cdot\boldsymbol{J}_T^{*}\right]}{\delta} = \frac{\sigma E^2}{\delta} \tag{13}$$

2.2. Application to RF exposure assessment

This paragraph shows an example of computations performed with our code [4]. The radiated body is the Specific Anthropomorphic Mannequin (SAM) proposed by [6]; the mesh adopted during the simulations is shown in Figure 1a. and involves about 2000 triangular elements. The phantom is here exposed to the near field generated by a small loop antenna (average radius: 25 mm) located laterally to the head, at a distance of about 4 cm (see Figure 1b.). A unitary current (peak value) flowing in the loop produces an electromagnetic field at 100 MHz; at this frequency the electric conductivity, relative permittivity and mass density are assumed to be 0.33 S/m, 76 and 1000 kg/m^3 respectively.

In Figure 2. the reader can find a chromatic map representing the distribution of the local SAR over the middle plane of the phantom (where a preliminary analysis has identified the highest SAR values). The “hot-spots” are located near the nape and the root of the nose. It must be noted how the absence of a volume discretization allows us to compute with continuity the SAR in an arbitrarily high number of points.

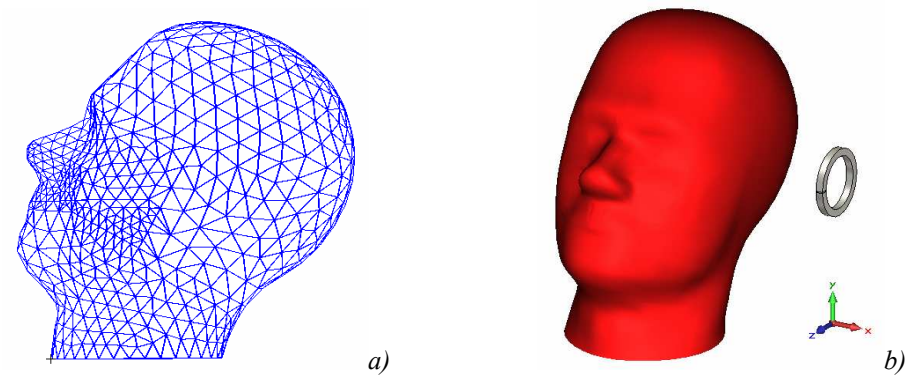


Figure 1. SAM phantom: a) BEM representation; b) exposure conditions

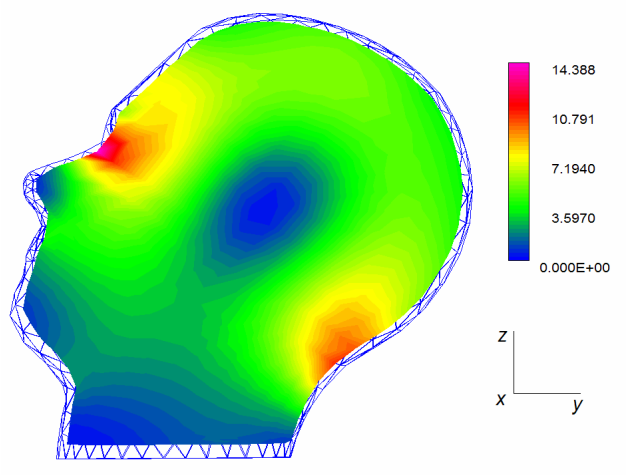


Figure 2. Chromatic map of the SAR distribution (values in mW/kg)

3. Modelling of exposure to low-frequency fields

This section is devoted to the discussion of formulations useful to simulate the human exposure to low-frequency (LF) quasi-stationary magnetic and electric fields. Under this classification we refer to time-varying fields with a frequency up to 100 kHz approximately, for which the propagation phenomena in general can be neglected and the two components (magnetic and electric) of the field seemingly act as two completely independent entities. We can exploit these features to obtain two different computational tools, one fit for the magnetic field and the other for the electric field, optimized for the specific problem we would like to solve. Obviously, this approach makes the formulations unsuited for the analysis of more general cases, but their performances increase and however they can be used to study any kind of situation that complies with the simplifying hypothesis.

3.1. Formulation for LF magnetic field problems

A formulation useful for the analysis of exposure to LF magnetic fields can be obtained, thanks to a deductive process, directly from the more general formulation discussed in paragraph 2.1; this latter, by virtue of suitable simplifying hypotheses, undergoes here some important changes.

At low frequency it is reasonable to disregard the dielectric currents with respect to the conduction currents. Moreover, owing to the conductivity of the biological tissues, the currents induced in the human body are too weak to modify the magnetic fields produced by external sources. So, if we consider a domain without any element able to perturb the magnetic field produced by the sources, we can assume it as an impressed quantity. Moreover, we have a simplification in the Green function defined through expression (9). In fact, if we consider the propagation of the field in the air, we have

$$kr = 2\pi f\sqrt{\mu_0\varepsilon_0}\ r = \frac{2\pi c_0}{\lambda}\sqrt{\mu_0\varepsilon_0}\ r = \frac{2\pi r}{\lambda} \tag{14}$$

where the relation has been reduced thanks to the link between the speed of the electromagnetic field in the vacuum (c_0) with the permeability (μ_0) and the permittivity (ε_0) of the vacuum. Now, taking into account that the frequency f is limited to 100 kHz, the shortest wavelength is $\lambda = 3000$ m and even if we contemplate a quite long distance r (for example 100 m) the previous expression becomes almost null. Therefore, the Green function is no longer a complex quantity but just

$$\Psi = \frac{1}{4\pi r} \tag{15}$$

Since the magnetic field is an impressed quantity, relation (8) reduces to the well-known *Biot-Savart law*:

$$\boldsymbol{H} = \int_{\Omega_s} (\boldsymbol{J}_s \times \nabla\Psi) dv \tag{16}$$

where Ω_s is the volume taken up by the known impressed current density.

The computation of the electric field induced inside the human body (which is a source-free region) can be performed on the basis of a modified form of equation (7):

$$\boldsymbol{E} = \oint_{\partial\Omega}\left[(\boldsymbol{n}\cdot\boldsymbol{E})\nabla\Psi + (\boldsymbol{n}\times\boldsymbol{E})\times\nabla\Psi - j\omega\mu\Psi(\boldsymbol{n}\times\boldsymbol{H})\right] ds\,, \tag{17}$$

It must be remarked that we keep on working in frequency domain adopting the phasor representation, but, according to the traditional treatment of LF phenomena, the magnitude of the field quantities is here assumed equal to the root mean square (rms) value of the corresponding sinusoidal signal.

In order to introduce the Boundary Element Method, we consider a homogeneous human model, whose surface is discretized into a number M of quite small triangular elements; on each single triangle m the field quantities are assumed to be uniform. The discretized version of equation (17) for a generic i-th element which lies over the bounding surface becomes

$$\xi \boldsymbol{E}_i = \sum_m^M \left(\boldsymbol{n}\times\boldsymbol{E}\right)_m \times \int_{\partial\Omega_m} \nabla\Psi_{i,m} ds + \sum_m^M \left(\boldsymbol{n}\cdot\boldsymbol{E}\right)_m \int_{\partial\Omega_m} \nabla\Psi_{i,m} ds - j\omega\mu \sum_m^M \left(\boldsymbol{n}\times\boldsymbol{H}\right)_m \int_{\partial\Omega_m} \Psi_{i,m} ds \quad (18)$$

In equation (18) we find again the singularity factor (ξ = 0.5) that is necessary to avoid in a proper way the singularity of the Green function. It is worth noting how the decoupling between the magnetic and the electric field exposures allows us to ignore the presence of the ground plane (whose properties are such that it is not able to perturb the magnetic field distribution); so, we avoid its introduction in the domain and the consequent mesh.

Equation (18) is a vectorial relation that in general involves three unknowns (the normal and the two tangential components of the electric field) because the magnetic field acts as a known term. Under the hypothesis of disregarding the dielectric currents, the continuity of the normal component of the conduction current density (that in air is null) implies that the electric field normal component is null everywhere along the surface. So, it is possible to project equation (18) only on the two tangential unit vectors; in this way the algebraic system will be constituted by $2M$ equations with $2M$ unknowns, giving rise to a square matrix of $4M^2$ real elements. Also in this case, after the solution of the BEM matrix, the electric field induced inside the body is computed with a relation similar to (18), in which the singularity factor is not present. The induced current density is finally obtained from the electric field simply through a multiplication by the proper conductivity.

3.2. Formulation for LF electric field problems

At low frequency, the effects of the magnetic field and of the impressed current densities on the environmental electric field are often negligible, and therefore the only source is represented by the charges. So, we can consider that the electric field is generated by charged bodies having a stated voltage. Usually the voltage is known (the power lines are a typical example) while the charges can only be obtained by solving the capacitance matrix. For this reason, it is reasonable to set a formulation in which the known terms are expressed directly in terms of the voltages. On the basis of these considerations, we will discuss a procedure that exploits the previous experiences with the BEM but, at the same time, diverges from them because it is focused on the electric scalar potential, instead of on the field quantities.

Under quasi-stationary conditions, the equation (1) assumes the approximate form

$$\nabla\times\boldsymbol{E} = 0 \quad (19)$$

so that it is possible to introduce the electric scalar potential V

$$\boldsymbol{E} = -\nabla V \quad (20)$$

Since in our analysis we do not consider the presence of charges (the charged bodies will be characterized through their voltage), the electric scalar potential satisfies the classical Laplace's equation. In the light of the Green theorem, for a homogeneous region Ω, bounded by a set of surfaces $\partial\Omega$, this kind of problem can be solved by the following integral equation

$$V = \oint_{\partial\Omega} \left(V \frac{d\Psi}{dn} - \Psi \frac{dV}{dn} \right) ds \tag{21}$$

where the Green function Ψ is the same defined with relation (15).

In order to introduce the Boundary Element Method, we discretized each surface into quite small triangular elements and on each single triangle the potential and its normal derivative are assumed to be uniform. For each volume in the domain, the potential of a generic *i*-th element, which lies over a bounding surface, receives the contribution of all the *M* triangles which overlook the volume:

$$\xi V_i = \sum_m^M V_m \int_{\partial\Omega_m} \left(\nabla \Psi_i \cdot \boldsymbol{n} \right)_m ds - \sum_m^M \left(\nabla V \cdot \boldsymbol{n} \right)_m \int_{\partial\Omega_m} \Psi_{i,m} ds \tag{22}$$

The factor $\xi = 0.5$ is used again to avoid the singularity arising when $r = 0$.

Each element allows the setting up of a scalar equation that involves two unknowns (the potential and its normal derivative). However, for the analysis of the human exposure to LF electric field, we have to write relation (22) for a domain constituted by the air (in which we find the sources) and the human model. Consequently, in order to set the matrix which describes the whole system, relation (22) is applied to both the air region and the interior of the body, so that each triangle introduces two scalar equations. The number of the unknowns does not increase thanks to the following interface conditions applied to each surface element:

$$\begin{aligned} V^{(a)} &= V^{(b)} \\ \left(\nabla V^{(a)} \cdot \boldsymbol{n}^{(a)} \right) \varepsilon^{(a)} &= -\left(\nabla V^{(b)} \cdot \boldsymbol{n}^{(b)} \right) \varepsilon^{(b)} \end{aligned} \tag{23}$$

where the minus sign in the second interface condition is used to take into account that the normal unit vectors are in opposite directions for the two volumes.

In general we have a number of unknowns lower than 2*M*, because over the surfaces of the sources the potential is imposed (Dirichlet condition); in this case, since each element introduces just one unknown, the BEM equation is only written for the external volume.

Unlike what we said in paragraph 2.1, for electric field problems the ground plane plays a fundamental role because it behaves as a good conductor and its presence cannot be ignored. In order to avoid the corresponding mesh (that would make the dimension of the matrix increase), we consider the ground as a perfect conductor and exploit the images principle. So, when we compute the contribution of the *m*-th element on the *i*-th one, we also introduce the contribution of the corresponding image element *m*'. This latter has opposite potential and normal derivative with respect to the actual triangle and involves a proper value for the Green function (on the basis of its coordinates). Since we do not write any equation for the images, the dimension of the matrix does not change and we need to store the same number of terms (with a consequent advantage on the overall computational burden).

After the solution of the matrix, the computation of the potential in any point of the considered volume can be performed thanks to relation (22) by setting $\xi = 1$. Finally, the

electric field can be obtained through the numerical computation of the gradient of the potential and allows an immediate calculation of the induced currents.

3.3. Application to LF exposure assessment

As an example of exposure assessment, in this section we show some results obtained during the analysis of workers' exposure in an electric substation operating at power frequency [3]. For the representation of the human body we make use of a 3D version of the "reference man" proposed in [5], which is here discretized with a surface mesh of about 2300 triangles (see Figure 3). As suggested by [5], the conductivity and relative permittivity of the human model are set to 0.2 S/m and 10^5 respectively.

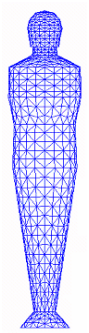


Figure 3. BEM mesh of the 3D reference man

The case under study assumes that an operator is placed nearby a three-phase 380 kV, 50 Hz line voltage busbar system. The supplied components are disposed at 6.5 m above the ground, the distance between the conductors is 5.5 m, and the distance between the insulators is 3.5 m. Three-phase balanced currents (1500 A) with a power factor $\cos\varphi$ = 0.9 are assumed to flow through the busbars. The body is placed both at 2 m from the closer phase (position A), where the maximum electric field is experienced, and below the central bar (position B), where the magnetic field reaches its maximum value. The two positions are indicated in Figure 4., which shows how each busbar is suspended by vertical insulators located at the top of basements at ground potential. This situation implies a mixed exposure to LF electric and magnetic fields and so we can take advantage of the two specific formulations previously described. The effects of the electric and magnetic fields are separately evaluated and then superimposed thanks to the problem linearity. The results are summarized in Figure 5, reporting the rms value of the induced current density along the *x*-axis, at the forehead height. The effect of the electric field exposure is always predominant; the presence of the magnetic field gives rise to an increase of the maximum value of about 20% in position A and about 30% in position B. The results obtained assuming $\cos\varphi = 1$ are also reported for position B, where it is evident that the value of the power factor has a weak influence. It is interesting to note that the effects of the two fields can interfere both in an additive and in a subtractive way. This result is due to the vectorial nature of the induced quantities and in general depends on the phase-shift of the sources.

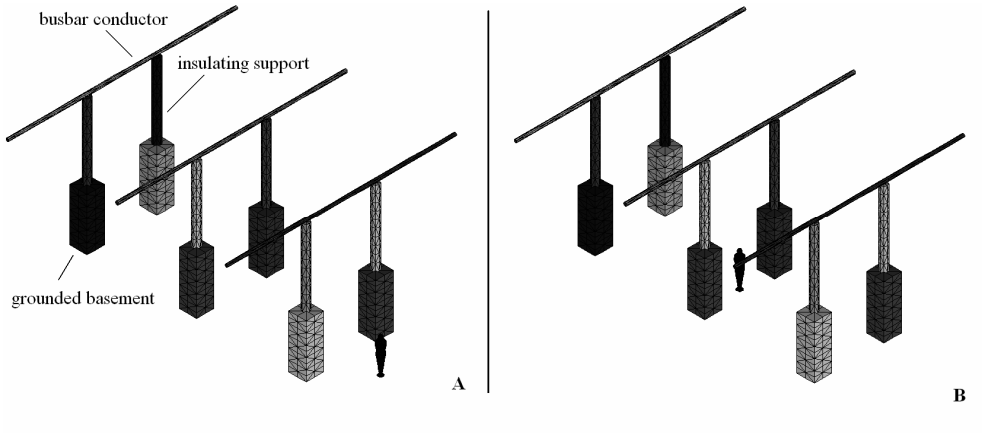


Figure 4. Position A and B for the exposure evaluation in an electric substation

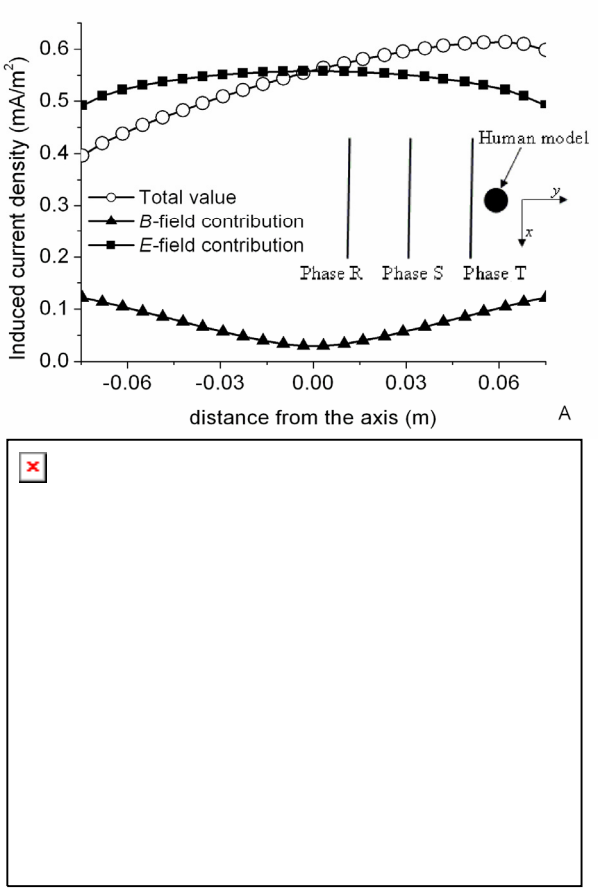


Figure 5. Distribution of the induced current density along the x-axis at the forehead (up-position A, down-position B)

4. Conclusions

In this paper we have shown how the Boundary Element Method can be usefully employed in order to set efficient computational tools devoted to the study of human exposure to electromagnetic fields from low to radiofrequency. The preliminary analysis of the physics involved by the exposure conditions has allowed the formulations of optimized procedures, general enough to be used for the numerical assessment of compliance with the limits in force as requested by the current technical standards.

References

[1] *1999/519/EC Council Recommendation on the limitation of exposure of the general public to electromagnetic fields (0 Hz to 300 GHz)*, (1999).

[2] Bernardi P.: *Specific absorption rate and temperature elevation in a subject exposed in the far-field of radio-frequency sources operating in the 10-900 MHz range*, IEEE Trans. on Biomedical Engineering, Vol. 50, No. 3, (2003).

[3] Bottauscio O., Chiampi M., Zilberti L.: *Boundary element approaches for the evaluation of human exposure to low frequency electromagnetic fields*, IEEE Transaction on Magnetics, Vol. 45, No. 3, (2009).

[4] Bottauscio O., Chiampi M., Zilberti L.: *A model to relate SAR to surface field measurements in human phantoms*, in Proceedings of the17th Conference on the Computation of Electromagnetic Fields, Florianopolis, Brazil, (2009), pp. 646-647.

[5] CENELEC, *Exposure to electric or magnetic fields in the low and intermediate frequency range - Methods for calculating the current density and internal electric field induced in the human body, Part 3-1: Exposure to electric fields – Analytical and 2D numerical models*, EN 62226-3-1, (2007).

[6] CENELEC, *Human exposure to radio frequency fields from hand-held and body-mounted wireless communication devices - Human models, instrumentation, and procedures - Part 1: Procedure to determine the specific absorption rate (SAR) for hand-held devices used in close proximity to the ear (frequency range of 300 MHz to 3 GHz)*, EN 62209-1, (2006).

[7] Chatterjee A., Kempel L. C., Volakis J. L.: *Finite element method for electromagnetic – Antennas, microwave circuits and scattering applications*, IEEE Press, New York, (1998).

[8] Dawson T. W. et al.: *High-resolution organ dosimetry for human exposure to low-frequency magnetic fields*, IEEE Trans. Mag., Vol. 34, No. 3, (1998).

[9] Dawson T. W. et al.: *High-resolution organ dosimetry for human exposure to low-frequency electric fields*, IEEE Trans. on Power Delivery, Vol. 13, No. 2, (1998).

[10] Dawson T. W. et al.: *High-resolution magnetic field numerical dosimetry for live-line workers*, IEEE Trans. Mag., Vol. 35, No. 3, (1999).

[11] *Directive 2004/40/EC of the European Parliament and of the Council of 29 April 2004 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (electromagnetic fields)*, (2004).

[12] Elliott R. S.: *Electromagnetics – History, theory and applications*, IEEE Press, New York, (1993).

[13] Gjonaj E. et al.: *High-resolution human anatomy models for advanced electromagnetic field computations*, IEEE Trans. Mag., Vol.38, No. 2, (2002).
[14] http://www.nlm.nih.gov/research/visible/visible_human.html.
[15] http://www.itis.ethz.ch/index/index_humanmodels.html.
[16] http://www2.nict.go.jp/w/w122/old/mt/b186/bio/bio_human_model_E.html.
[17] ICNIRP, *Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz)*, (1998).
[18] ICNIRP, *Guidelines on limits of exposure to static magnetic fields*, (2009).
[19] IEEE Std C95.6, *Standard for safety levels with respect to human exposure to electromagnetic fields, 0 - 3 kHz*, (2002).
[20] IEEE Std C95.1, *Standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz*, (2005).
[21] Stratton J. A. and Chu L. J.: *Diffraction theory of electromagnetic waves*, Phys. Rev., 56, 99-107, (1939).
[22] Zaidi H., Xu G.: *Computational anthropomorphic models of the human anatomy: the path to realistic Monte Carlo modeling in radiological science*, Annual Review of Biomedical Engineering, Vol. 9, (2007).
[23] L. Zilberti, PhD Dissertation: *Computational models for the evaluation of human exposure to electromagnetic fields*, Politecnico di Torino, Dipartimento di Ingegneria Elettrica (2010).

The first Experiments in Magnetic Stimulation – a History of Discoveries within two Parallel Lives

A. Krawczyk, E. Łada-Tondyra

Czestochowa University of Technology,
Al. Armii Krajowej 17, 42-200 Częstochowa
e-mail: a.krawczyk@el.pcz.czest.pl, e.lada-tondyra@el.pcz.czest.pl

Abstract: The paper contains a history of the first effective experiments in magnetic stimulations of human tissues, which started in the second half of XIX century due to two scientists. They were Jacques-Arsene d'Arsonval and Silvanus P. Thompson, both of them born in 1851. The experiments which were carried out by these pioneers are of great importance for nowadays activity in medical applications of electromagnetic field.

Keywords: Jacques Arsene d'Arsonval, Silvanus P. Thompson, bioelestromagnetics, magnetophosphenes

1. Introduction

The real and effective experiments in magnetic stimulations of human tissues and nerves started in the second half of XIX century due to two scientists – both were born in 1851. They were:

- Jacques-Arsene d'Arsonval, French physician and physicist (1851–1940)
- Silvanus P. Thompson, British engineer (1851–1916) (Figure1).

Figure 1. Jacques-Arsene d'Arsonval and Silvanus P. Thompson

Both of them created the fundaments for therapeutical action of electromagnetic field as well as found the phenomenon called magnetophosphene (gr. magnet + *phōs,* light +

phainein, to show). The experiments which were carried out by these pioneers are of great importance for nowadays activity in bioelectromagnetics.

2. Jacques Arsene d'Arsonval

Jacques-Arsene d'Arsonval was born at La Borie, France on June 8, 1851. He came from a noble French family. His grandmother's godfather was Napoleon I. His grandfather and father were physicians. The paternal line of d'Arsonval's family for seven centuries bore the title of Count. He studied classics at the Lycée Imperial de Limoges and later at the Collège St.-Barbe. Having obtained a baccalaureate degree from the Université de Poitiers in 1869, d'Arsonval decided upon a career in medicine and commenced his studies at Limoges. He attended the universities at Limoges and Paris – at the College de France he was under influence of Claude Bernard (1813–1878), the great experimental physiologist and creator of the concept of a constant internal environment (*milieu interior*), which leads to the notion of homeostasis. As Claude Bernard was also the great methodologist of science, d'Arsonval could grasp from him the spirit of research experiments. D'Arsonval got his medical degree in Paris in 1877. After Bernard's death he assisted Charles-Édouard Brown-Séquard (1817–1894), giving the latter's winter courses, and replaced him at the Collège de France when Brown-Séquard died in 1894. In the same year he became a professor and was elected to the Academy of Sciences where in 1917 was nominated as its president.

Minister of public education enabled Collège de France to establish a laboratory for biophysics at rue St.-Jacques. D'Arsonval directed the laboratory until 1910, when he moved to the new laboratory at Nogent-sur-Marne directed by him till his retirement in 1931. He died on December 31, 1940 at the same place he was born almost ninety years earlier.

D'Arsonval was a member of the Society of Biologists, the Academy of Medicine and the Academy of Sciences. He also was an active member of a few other societies dealing with:

- electrotherapy,
- physics,
- electronics,
- engineering.

D'Arsonval's most outstanding scientific contributions, however, involved the biological and technological applications of electricity. Much of this work concerned muscle contractions. In 1882 d'Arsonvel together with French scientist Étienne-Jules Marey (1830–1904) and French electrical engineer Marcel Deprez (1843–1918) invebted a new device, called later the Deprez-d'Arsonval galvanometer. The invention came after he had studied muscle contractions in frogs using an extremaly weak current comparable with physiological current. In 1881, Arsène d'Arsonval suggested for the first time harnessing the temperature difference in the tropical seas for the generation of electricity. In 1902 d'Arsonval worked with his student Georges Claude (1870–1960) on industrial methods for the liquefaction of gases.

In his scientific activity d'Arsonval applied the laws of physics to solve bioelectromagnetic problems, such as:

- inductive and capacitive heating human body (diathermy),
- magnetic stimulation of excitable tissues by eddy currents.

In 1891 he created nis first variable-frequebcy stimulator (Figure 2.), consisted on potentiometer (P), series of electrolytic cells (V) and a graphic recorder (K-kymograph) on which muscle (M) contraction and the frequency (f) were displayed. His first results showed that as frequency was increased, the intensity had to be higher to obtain a muscular response. Because of the frequency limitation, he used an alternator that provided sinusoid voltage up to 10 kHz. With this stimulator he obtained similar results.

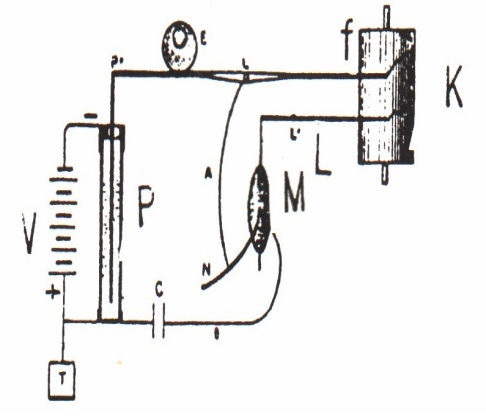


Figure 2. Variable-frequency stimulator

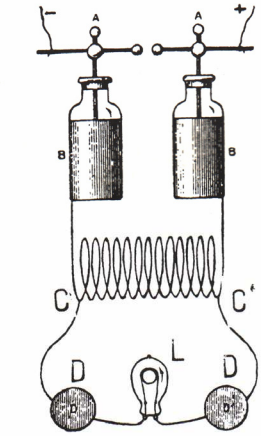


Figure 3. High frequency stimulator

One of the most important experiment in his life was that showed in Figure 3. He used a Hertz oscillator, two Leyden jars an air-core coil. The ends of coil were connected to a circuit in series, which consisted of a 100 W light and two subjects (humans). The current flowing through the subjects was about 1 A. D'Arsonval noticed that the subjects felt thermal sensation, but they did not feel pain. He demonstrated that a human organism could conduct an alternating current strong enough to switch on an electric lamp.

D'Arsonval showed that current could be passed through the human body by capacitive coupling (Figure 4.). Coil of the high-frequency oscillator was connected to an electrode in the human hand. D'Asonval noted that more than 300 mA can be passed through the body with this method.

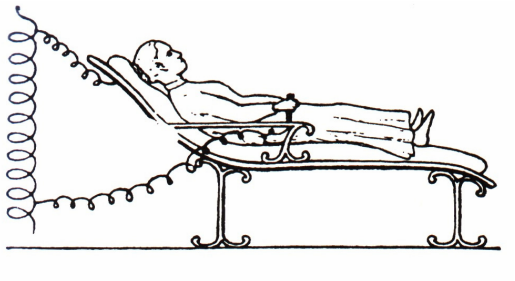


Figure 4. Method of capacitive coupling

D'Arsonval was also interested in an inductive heating (Figure 5.). A man or a rabbit were placed inside long solenoid with current. He realized himself that the current induced in the body was not confined to the surface. This experiment shows that d'Arsonval performed first medical diathermy.

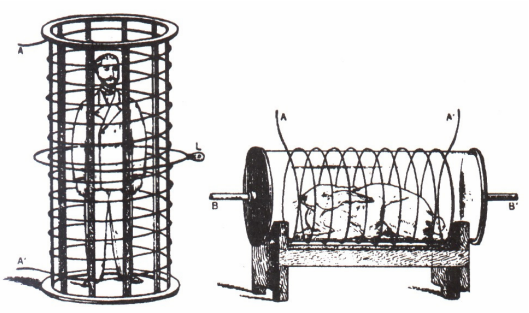


Figure 5. Inductive heating of a man and rabbit

In the twentieth century the techniques, recently named *darsonvalisation* was performed using two methods: general and local. General treatments bases on inductive heating, using a stationary solenoid with a height of ca. 2 meters, formed as a thick coil of copper wire. Patient during the procedure remains inside the solenoid using a wooden chair. The procedure is realized every two days, approximately the treatment is twenty-fold one. For cosmetic purposes one uses local treatments.

D'Arsonval showed in his demonstration that a strong, low frequency electromagnetic field can induce eddy currents into living tissue and thereby cause stimulation. Using generator of 42 Hz, the frequency used at that time for energy transmission, he discovered the phenomena which links electromagnetic field with visual sensations which have been called magnetophosphenes . He wrote about his discovery:

„There occurs, when one plunges the head into the coil, phosphenes and vertigo, and in some person, syncope... Alternating magnetic field modifies the form of muscular contraction and produces in living beings other effect…"

In 1933 the Ministry of Education held an official jubilee for d'Arsonval at the Sorbonne. He was created knight of the Legion of Honour in 1884, and received the Grand Cross in 1931.

3. Silvanus P. Thompson

Silvanus Philips Thompson was born on June 19, 1851 in Settle, UK. The family were Quakers and consisted of mother Bridget and 8 children, 5 boys and 3 daughters with Silvanus Philips being their second child. After schooling he went to the Quakers Training College at Pontefract, where he studied to be a teacher and obtained a baccalaureate degree at the age of 19. After trying his hand as a teacher, he studied chemistry and physic in Royal School of Mines art South Kensington. After graduation he was appointed at Bristol University Collage and very soon he has got the scholarship at the University of Heidelberg, Germany. In 1878 he was elected to Chair of Physics. At age of 27 he had assumed higher academic responsibility, that of a professorship. His first publication appeared in 1876 in "Philosophical Magazine" and was titled "On some

phenomenon of induced electric spark". A year later he published the paper "On improved lantern galvanoscope" and the paper was highly evaluated by William Thomson, later Lord Kelvin. Starting with these two papers he published every year a couple of papers as well as he presented lectures to special auditoria.

He was described as a competent and skilled speaker both in scientific debate and social occasions. William Hale White said about him: *"The audience was thrilled... Silvanus Thompson was a prince among lecturers. I have never heard a better demonstration or attended a more memorable medical meeting."*

Thompson was the first President of the Roentgen Society (now The British Institute of Radiology) founded in 1897. He described the society as being between medicine, physics and photography. Since 1918 they have an annual Silvanus Thompson memorial lecture for which they award a silver medal.

He was also the President of the IEE (Institution of Electrical Engineers) and he was a member of the Royal Society and of the Royal Swedish Academy of Sciences.

Figure 6. Silver medal of Silvanus P. Thompson

He was a recognised authority upon electricity, magnetism and acoustics.
Thompson wrote a few scientific books on electricity, such as:

- Elementary Lesson in Electricity and Magnetism (1881)
- Dynamo Electrical Machinery (1896)
- Calculus Made Easy (1910),

which ran through some 40 editions and reprints.

He also wrote biographies of Michael Faraday and Lord Kelvin as well as biography of William Gilbert, the physician of Queen Elizabeth I. He prepared an edition of Gilbert's opus vitae *De Magnete* at the Chiswick Press in 1900. In 1912 Thompson published the first English translation of *Treatise on Light* by Christian Huygens.

All his activity was concentrated on, using the contemporary notions, bioelectromagnetic phenomena. In the next 30 years Thompson's research was prolific, mainly in the fields of electricity, magnetism and optics. He was deeply involved in exploring Roentgen's X Rays and fluorescence phenomenon. He is also considered to be the precursor of transcranial magnetic stimulation (TMS) – in 1910 he started to investigate magnetic field effects on the brain.

Silvanus P. Thompson was admired in academic world for his outstanding way of lecturing . His lectures gathered a lot of students and other listeners.

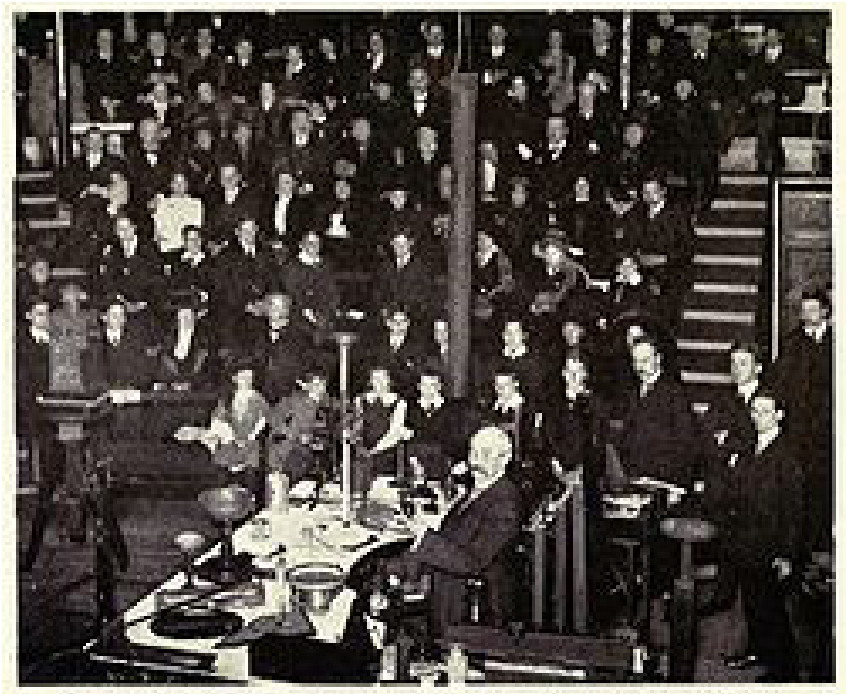


Figure 7. Silvanus P. Thompson during lecture

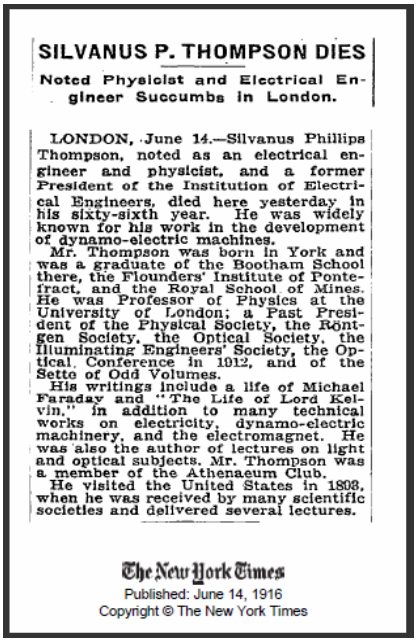
William Hale White (1857–1949), a distinguished British physician and medical biographer, after one of Thompson's lectures said: *The audience was thrilled... Silvanus Thompson was a prince among lecturers. I have never heard a better demonstration or attended a more memorable medical meeting.*

In 1910 he published the article *A Physiological Effect of an Alternating Magnetic Field* (Proceedings of Royal Society B July 21, 1910 82:396-398) in which, independently of d'Arsonval, the visual sensations of electromagnetic field were described. They were magnetophosphenes again. It is very interesting to notice that he made experiments on himself (Figure 8.). The parameters of his experiment were as follows: peak value of magnetic flux density up to 140 mT, frequency 50 Hz. The contemporary studies show that retina is the structure which can be stimulated by much weaker magnetic field – magnetophosphenes can be obtained with 10 mT (rms) and frequency 20 Hz.

Figure 8. Stimulating retinal light by magnetism

The last years of the Silvanus P. Thompson life were dominated by the First World War. As he was a Quaker he had difficulties with all the kinds of violence. The first problem he had was the Boer War, but the time of the First World War was more difficult for him. He gave up any technical and scientific activity, which he believed to be militaristic. He looked at mainland Europe with great anxiety and sorrow, but he carried on working and death overtook him in 1916 at the relatively young age of 65.

Silvanus P. Thompson was known all over the world – just in a couple of days after his death the special obituary announcement was published in The New York Times (Figure 9.).



SILVANUS P. THOMPSON DIES

Noted Physicist and Electrical Engineer Succumbs in London.

LONDON, June 14.—Silvanus Phillips Thompson, noted as an electrical engineer and physicist, and a former President of the Institution of Electrical Engineers, died here yesterday in his sixty-sixth year. He was widely known for his work in the development of dynamo-electric machines.

Mr. Thompson was born in York and was a graduate of the Bootham School there, the Flounders' Institute of Pontefract, and the Royal School of Mines. He was Professor of Physics at the University of London; a Past President of the Physical Society, the Röntgen Society, the Optical Society, the Illuminating Engineers' Society, the Optical Conference in 1912, and of the Sette of Odd Volumes.

His writings include a life of Michael Faraday and "The Life of Lord Kelvin," in addition to many technical works on electricity, dynamo-electric machinery, and the electromagnet. He was also the author of lectures on light and optical subjects. Mr. Thompson was a member of the Athenaeum Club.

He visited the United States in 1908, when he was received by many scientific societies and delivered several lectures.

The New York Times
Published: June 14, 1916
Copyright © The New York Times

Figure 9. Obituary published in NYT

4. Conclusion

The magnetic stimulation of tissues proposed and investigated by Jacques Arsene d'Arsonval and Silvanus P. Thompson was the precursor of the magnetotherapy and transcranial magnetic stimulation, the electromagnetic techniques used in contemporary medicine. Therefore, it is very important for the researchers who deal with electromagnetic medicine nowadays to know how the problem began more than 100 years ago.

References

[1] J. S. Thompson, H. G. Thompson, Silvanus Phillips Thompson – His Life and Letters, London, T.Fisher Unwin, Ltd., (1920).
[2] L. A. Geddes, d'Arsonval' Physician and Inventor, IEEE Engineering in Medicine and Biology, vol.18, 4, (1999).
[3] D. R. Justesen, A. W. Guy, Arsene Jacques d'Arsonval: A Brief History, Bioelectromagnetics, 6, (1985).
[4] www.rpec.co.uk/engineerswalk/st_walk.html

Improving the Spatial Distribution Algorithm of the Traffic-Flow Analysis

L. Muka and G. Lencse

Elassys Consulting Ltd., H-1026 Budapest, Törökvész lejtő 10/a.
Phone: + 36 1 346-8814, fax: + 361 346-8814
e-mail: muka.laszlo@elassys.hu, lencse@sze.hu

Abstract: This paper investigates and improves an important algorithm that is used in the Traffic-Flow Analysis (TFA). TFA can be used for the fast and approximate (performance) analysis of Information and Communication Technology (ICT) systems. The method contains an algorithm for the spatial distribution of the traffic in the system. It is shown how the error of the spatial distribution can be measured, and the effect of the so called *size of routing unit* parameter (S_{RU}) of the algorithm is investigated. In order to define S_{RU}, the Aggregated Traffic Tree-Model (ATTM) is introduced. Using step-by-step Discrete-Event Simulation (DES) and appropriate statistical algorithms in ATTM a method for determination of S_{RU} is described. A method for the refinement of the S_{RU} – based on the metrics defined for the measurement of TFA results – is introduced. A method taking into account Quality of Service (QoS) requirements in S_{RU} refinement is also described in an example.

Keywords: discrete-event simulation, traffic-flow analysis, information and communication systems

1. Introduction

1.1. Performance analysis methods

Discrete-Event Simulation (DES) is a widely used method for the performance analysis [3] of ICT and BP systems. The simulation of large and complex systems requires a large amount of memory and computing power that is often available only on a supercomputer. Efforts are made to use multiprocessor systems or clusters of workstations. The conventional synchronisation methods for parallel simulation (e.g. conservative, optimistic) [1] use event-by-event synchronisation and they are unfortunately not applicable to all cases, or do not provide the desirable speedup. The Statistical Synchronisation Method proposed by Pongor [11] does not exchange individual messages between the segments but rather the statistical characteristics of the message flow. The method can produce excellent speed-up [4] but has a limited area of application [5].

The Traffic-Flow Analysis [6] was proposed for the rapid performance estimation of ICT systems. Its spatial distribution algorithm was studied first in [9]. The applicability

of Traffic-Flow Analysis for large networks and for ICT and BP systems was examined in [7] and [10].

1.2. The Traffic-Flow Analysis

The Traffic-Flow Analysis (TFA) is a combination of simulation and analytical and/or numerical methods. While the traditional discrete-event simulation models the travelling of each packet through the network, TFA uses statistics to model the networking load of applications [6], [8]. TFA works in two stages:

In the *first stage,* the method *distributes the traffic* (the statistics) in the network, using routing rules and routing units.

In the *second stage,* the influence of the *finite capacities* (line and switching-node capacities) is calculated.

The important features of TFA:

- The results are approximate but the absence and the place of *bottlenecks* is shown by the method.
- The *execution time* of TFA is expected to be significantly less than the execution time of the detailed simulation of the system.
- TFA describes the *steady state* behaviour of the network (there is no need for warm-up time definition).

2. Stating the problem of spatial distribution

TFA is a general method, and can be used with any traffic model that satisfies the requirements of TFA for the traffic model. In [6], there were proposed *bit-throughput distribution* and *packet-throughput distribution* (practically histograms) as traffic models to model the traffic on the lines and in the nodes, respectively.

The traffic model is always an *aggregated traffic model*, that is, it represents the *complete traffic of a given type of applications* that are *connected to the given node*. For example it represents the full traffic (*in both directions*) of 35 *FTP* applications that are connected to a router (by switches). If *static* routing is used, we can handle the complete traffic of the before mentioned 35 FTP applications (or 100 *web browsers* or any other type of applications) together: we must route only one statistics package through the network (containing the two types of histograms). However, if we have *adaptive* routing, then the traffic of a given type of application should not be handled together, rather it must be routed in multiple packets (routing units), each of which *represent a given portion of the traffic of the given type of application connected to the given node*. According to the previous definition, the number of routing units may be a number between one and the number of the units of traffic generated during the detailed simulation of the system.

When determining the *size of the routing unit* (S_{RU}), we must consider the following issues:

The larger S_{RU} we choose the fewer statistics packages are to be routed in the first phase and the less traffic model addition operations are to be performed in the second phase of TFA. However, if S_{RU} is too large, the spatial distribution of the traffic may considerably differ from the one that is formed in the detailed simulation of the system (and from the one in the real system). If S_{RU} is small, the spatial distribution of the traffic may be quite precise, but the larger amount of messages to be routed and traffic models to be added slow down the analysis. *The choice of S_{RU}[1] must be a reasonable compromise (between the contradicting requirements) that is made in the knowledge of the whole system modelled.*

3. Defining the Aggregated Traffic Tree-Model

First, for the analysis in order to define S_{RU}, let us introduce the tree model of the aggregated traffic. The Aggregated Traffic Tree-Model (ATTM, shown in Figure 1) is a *fragment* of the whole examined network model which consists of *nodes* and *lines* with *application models.* The ATTM has the following features:

- The application model a_y^x generates the aggregated traffic of all the applications of the *y*-th class (e.g. VoIP, FTP, web browsing, etc.) connected to node h_x, that is a_y^x contains all the application models of the *y*-th class for node h_x and a_y^x generates the full traffic between the *source and destination node pairs (h_x - h_j, h_x - h_{j+1}, ..., h_x - h_{j+d})* in *both directions*. (In Figure 1, the application model a_y^x is in a rectangle drawened with thick line, the logical connection of a_y^x to node h_x is represented by thick line, nodes are shown with thick-line circles and lines of the ATTM model are drawened by thick lines between circles.)

[1] Examples for the influence of the selection of different values for S_{RU} are described in [6].

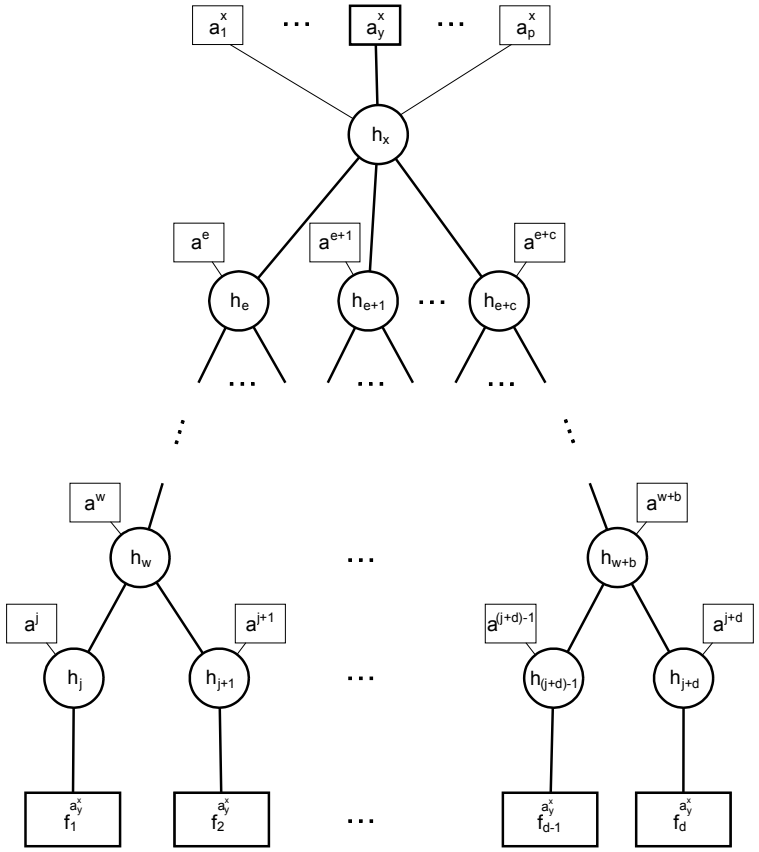


Figure 1. The Aggregated Traffic Tree-Model of an Application

- Other application models generate the traffic *independently* from a_y^x: the application models connected to h_x (a_1^x, ... , a_p^x) represent the traffic of other classes; application models connected to other nodes (for example a^e, a^w, and a^j connected to nodes h_e, h_w *and* h_j respectively) represent the application models of all classes. (In Figure 1, all these applications are shown with thin-line rectangles and their logical connections to nodes are also represented by thin lines.)

- Any other application model connected to any other node in the model of the *examined network* can generate traffic for the nodes and lines of ATTM.

- ATTM contains only the nodes and lines used by the traffic generated by a_y^x.

- The *routing* in the *examined network* is an *adaptive* routing with unknown features.

- In ATTM there is only *one route* between the source node and a destination node. (A route is a sequence of *nodes* and *lines* between the source node and a destination node related to a_y^x). If there are more routes between the source and the same destination used by the traffic generated by a_y^x then each route is handled separately as if it has a different destination. For these "*split destinations*", the traffic (for each unit of the traffic) is traced and recorded separately for all the elements of the route. Thus, it also means, that more split destination of ATTM will represent one original, non-split destination. (Splitting of destinations is performed according to the traffic generated by event-by-event simulation in the examined network.)

- In the ATTM, the traffic generated by a_y^x is observed at destination nodes h_j, h_{j+1}, ..., h_{j+d}. The frequency distribution of the traffic observed at destination node h_{j+d} is denoted by $f_d^{a_y^x}$ and shown in a rectangle drawn by thick line, and the logical connection of which to node h_{d+1} is also represented by a thick line. (The exact description of the measurement of this frequency distribution is given in point 4. of this paper.)

4. Statistical analysis of the aggregated traffic

Now, using the ATTM, let us examine the *behaviour of the aggregated traffic* generated by the application model a_y connected to the node h_x (that is by a_y^x) between the *source* and *destination* node pairs (h_x - h_j, h_x - h_{j+1}, ...,h_x - h_{j+d}) in both directions at destinations. The traffic between the node pairs is *influenced* by node and line capacities, by the traffic generated of other application models and by the routing (which is *adaptive routing* for now) in the examined network. All the application models are detailed DES (Discrete-Event Simulation) models and generate the traffic according to the event-by-event simulation method. The aggregated traffic at destinations is examined in time intervals of T length. The T[2] interval is selected of the same length as the throughput collection interval in TFA. For the observation, in ATTM, the aggregated traffic is *recorded* for each destination of a_y in the subsequent T long time intervals.

To *observe* the aggregated traffic at destinations of a_y in ATTM, the *statistical sampling* approach will be applied: the statistical sampling method is used to define the *mean aggregated traffic* at a destination in T long time intervals. There are observed frequency distributions generated by a_y: $f_i^{a_y^x}$ denotes the frequency distribution at destination i

$$(f_i^{a_y^x} \in \left\{f_1^{a_y^x}, f_2^{a_y^x}, \dots, f_d^{a_y^x}\right\}).$$

The mean $\bar{k}_i$ of the observed frequency distribution $f_i^{a_y^x}$ is calculated by the formula:

[2] The influence of the selection of the value for the T parameter is analysed in [6].

$$\overline{k_\iota} = \frac{\sum_j f_{i_j} k_{i_j}}{\sum_j f_{i_j}} \quad (1)$$

where k_{i_j} denotes the number *units of the aggregated traffic* observed in *T* time interval at destination *i* and f_{i_j} is the observed frequency of k_{i_j}.

In using the statistical sampling method, both *infinite* and *finite population* approach may be applied. (In case of simulation of the traffic for long-run, the observation is characterised by a large number of *T* intervals and may be approximated by an *infinite population*, but the observation of simulation of one day probably may be modelled by using the *finite population* approach.)

The mean aggregated traffic, m_i, at a destination *i* can be calculated using the formula:

$$m_i = \overline{k_\iota} \pm B_i, m_i \in \{m_1, m_2, \dots, m_d\} \quad (2)$$

where $\overline{k_\iota}$ is the *mean aggregated traffic of the sample* (frequency distribution), B_i is the allowable error for determination of the mean m_i at destination *i*.

The sample size which is necessary to define the m_i with the precision B_i depends on the *method of sampling* too.

Let us examine the application of the *simple random sampling (SRS)* method which has the advantages that requires minimum knowledge of *population* and free of possible classification errors. The *sample size* for *SRS* samples *without replacement from an infinite population* may be calculated according to the formula:

$$n = \frac{z^2\sigma^2}{B^2} \quad (3)$$

where *n* is the required sample size, σ denotes population standard deviation, *z* is the *z* score determined for a specific confidence level and *B* is the allowable error.

The *sample size for SRS* samples without replacement from a *finite population* with *finite population correction factor* may be determined as:

$$n = \frac{\sigma^2 N}{\frac{(N-1)B^2}{z^2} + \sigma^2} \quad (4)$$

where *N* is the population size.

Let us see an *example* of sampling analysis of the aggregated traffic at a destination: the aggregated traffic is measured in packets, the length of *T* interval is 5sec, the required confidence level is 95% ($z_{0.05}$ = 1.96), the allowable error *B* = 10 packets, the finite population sizes are *N*=17280 and *N*=120960 (according to the 1-day and 7-day long observation intervals).

The value of the population standard deviation is determined using an *estimate*: using the *GAM* (General Application Model) *transaction parameters of TFA* [6] *for* a_y connected to the node h_x the *range* of the aggregated traffic may be determined. (The GAM parameters which can be used for the determination of the range are: $N_p^{\{s/r\}}(tt)$

the number of packets sent or received by the application during a transaction of the given type, $N_{TrPD}(tt, day)$ the number of transactions performed by an application from the given transaction type on a given day of the week, $I_{Tr}(tt, day, time)$ the probability density function describing the time distribution of the transactions of the given type on the given day of the week.)

For the example, the range of the aggregated traffic determined this way is 290 packets for 5 sec, thus – according to the *estimation* based on *Chebyshev's theorem – s* (sample standard deviation) ≅ *range*/4 ≅ 72.5 packets ≅ σ.

Thus, the sample sizes for SRS sampling from finite populations are:

$n_{N=17280} = 199.60 \cong 200$ and $n_{N=120960} = 201.59 \cong 202$ (after rounding it is the same size as the sample size for SRS sampling from *infinite* population: $n = 201.92 \cong 202$). After performing the sampling according to the determined sample size, it can be asserted with 95% confidence that the mean aggregated traffic *m* at the examined destination can be calculated as $m = \bar{k} \pm 10$ packets, where $\bar{k}$ is the mean aggregated traffic of the sample.

The sample sizes should be determined for each destination of a_y according to the method described in the previous example and the *largest sample size* should be used for sampling:

$$n = sup\{n_1, n_2, \dots, n_d\} \tag{5}$$

The use of *other sampling methods* may also be considered:

If there is more available preliminary knowledge about the behaviour of the aggregated traffic generated by the application model, then, for example, the *stratified sampling* method may be used. Applying this method the *variability* of data within strata, as well as their *size* may be taken into account.

The approach of *sampling based on the expert judgment* and on *convenience* – that is on non-probability considerations – may simplify the execution of sampling. (Using the expert judgement approach, for example, one or more *typical* or *critical* intervals for observation are selected and the sampling is executed taking into account the necessary warm-up intervals for the simulation. The approach using the convenience considerations may mean, for example, using only the intervals for which data have been collected.) The disadvantage of this approach is the occurrence of uncontrolled *bias* in sampling.

5. Determining the size of the routing unit

Now, let us formulate the requirements for S_{RU}. Let us use *TFA application models* in the *ATTM*: the aggregated traffic in the examined network is generated by the *GAM* of TFA. Let S_{RU} be the *size of the unit of the traffic* that can be handled together for the given *type of applications* connected to the *given node.* If the S_{RU} is increased then the amount of the necessary computations will be decreased but the resolution of the traffic is decreased thus the *uncertainty* will be increased and the reliability of the results will be decreased.

How to find a compromise? Let us call for help the *chi-squared test* for "goodness of fit".

The calculation for the χ^2-test is expressed by the formula:

$$\chi^2 = \sum_{j=1}^{r} \frac{(k_j - n^* p_j)^2}{np_j} \tag{6}$$

where k_j *is the observed frequency in class j,* p_j is the *expected hypothetical probability* for class *j, n* is the *number of observations*, *r* is the *number of classes*, the *grade of freedom* is *r-1* and of course $\sum_{j=1}^{r} p_j = 1$ and $\sum_{j=1}^{r} k_j = n^*$.

According to the logic of the test, a hypothesis is evaluated: for example if the value of the test statistic is large enough (at a determined level of significance) to lie in the critical region of the test the null hypothesis is rejected and it is concluded that the theoretical expected distribution is not a good fit to the observed distribution.

The necessary sample size – in order to have a good approximation to a chi-squared distribution – may be determined according to the *inequality* $np_j \geq 10$ (which is well known from statistics) where p_j is the minimum of the expected theoretical probability.

For further consideration, let us use the *frequency distributions observed* at destinations during the *ATTM DES work sampling* described in the previous point.

The observed *DES-sampling* frequency distribution at destination *i* is $f_i^{a_y^x}$ ($f_i^{a_y^x} \in \left\{f_1^{a_y^x}, f_2^{a_y^x}, \dots, f_d^{a_y^x}\right\}$), the sampling mean aggregated traffic at destination *i* is m_i ($m_i \in \{m_1, m_2, \dots, m_d\}$) Let us calculate the expected probabilities using the observed DES-sampling frequency distributions taking into account the number of *classes* for the expected hypothetical distribution. The probability of class *j* calculated this way at destination *i* is p_{i_j}, $p_{i_j} \in \left\{p_{i_1}, p_{i_2}, \dots, p_{i_r}\right\}$. Now, let us use these probabilities to determine the necessary sample sizes. The minimum probability at the destination *i* may be expressed as:

$$inf\left\{p_{i_1}, p_{i_2}, \dots, p_{i_j}, \dots, p_{i_r}\right\} \tag{7}$$

Thus – using the *inequality* $n^* p_j \geq 10$ (that is $n^* p_{i_j} \geq 10$) – the *necessary* sample size at destination *i* may be expressed as follows:

$$n_i^* \geq \frac{10}{inf\left\{p_{i_1}, p_{i_2}, \dots, p_{i_j}, \dots, p_{i_r}\right\}},\ n_i^* \in \left\{n_1^*, n_2^*, \dots, n_d^*\right\},\ (j = 1, 2, \dots, r) \tag{8}$$

Now, taking into account the observed means of the aggregated traffic at destinations of a_y^x, the number of required *RUs* for a_y^x may be determined using the following formula:

$$\mathbb{N}_{RU}^{a_y^x} \geq sup_i \left\{ \frac{\frac{1}{inf\left\{p_{i_1}, p_{i_2}, \dots, p_{i_j}, \dots, p_{i_r}\right\}} \sum_{s=1}^{d} m_s}{m_i} \right\} \tag{9}$$

where $m_i \in \{m_1, m_2, \dots, m_d\}$, $(i = 1, 2, \dots, d)$.

In $\mathbb{N}_{RU}^{a_y^x}$ is taken into account – because of the definition of ATTM – the weight (m_i) and the shape ($n_{i\,i}^*$) of every route for a_y^x destinations.

The required S_{RU} for the given type of application a_y connected to the node h_x in ATTM may be calculated according to the *inequality*:

$$S_{RU}^{a_y^x} \leq \frac{\mathbb{N}_u^{a_y^x}}{\mathbb{N}_{RU}^{a_y^x}} \tag{10}$$

where $\mathbb{N}_u^{a_y^x}$ is the amount of traffic (measured in *packets* or in *bits*) that has to be transferred in transactions generated by a_y.

6. Refining the routing unit size determination

The traffic model is always an aggregated traffic model, that is, it represents the complete traffic of a given type of applications that are connected to the given node. In the aggregated traffic model of TFA, the traffic of all application models of the same class – that is the *traffic of the same class* – connected to the *given node* is modelled together.

In order to *refine* the determination of S_{RU}, let us examine the determination of S_{RU} *together with the classes of traffic* of TFA.

The traffic in TFA may be expressed using the following formula:

$$\mathbb{T} = \bigcup_{q=1}^{Q} \bigcup_{v=1}^{V} \mathbb{T}_{qv} \tag{11}$$

where V is the number of nodes, Q is the number of traffic classes, $\mathbb{T}_{qv}$ is the traffic generated by the application model a_q^v, and $\mathbb{T}$ is the whole traffic generated by the *GAM* of TFA. To increase the number of RUs for traffic $\mathbb{T}_{qv}$ ($\mathbb{N}_{RU}^{\mathbb{T}_{qv}}$) means the change of the S_{RU} parameter of the application model a_q^v generating the traffic of class q for node v ($\mathbb{N}_{RU}^{\mathbb{T}_{qv}} = \mathbb{N}_{RU}^{a_q^v}$).

6.1. Introducing metrics for evaluation

To be able to determine good enough values for S_{RU} that take into account the traffic classes too it is necessary to introduce tools to measure the features of a given spatial distribution of the traffic in TFA.

The *capacity matrix* $\underline{\underline{K}} = [k_{ij}]$ describes the capacity of *nodes* and *lines* in the *examined network*. The capacity matrix is a *V*V* matrix, where *V* is the number of nodes in the examined network. Matrix element k_{ii} is the routing capacity of node *i* (measured in packets per second), and matrix element k_{ij}, where $i \neq j$ is the transmission capacity of the line from node *i* to node *j*, measured in Mbit/s. If there is no transmission line from node *i* to node *j* then $k_{ij} = 0$.

The results of the TFA are described by the *traffic matrix* $\underline{T} = [t_{ij}]$. Matrix element t_{ii} describes the result of TFA for node *i*: both the packet-throughput distribution (P_{ii}) and the delay distribution (D_{ii}) of TFA resulted in node *i*, and matrix element t_{ij} describes the result of TFA for the line from node *i* to node *j*: both the bit-throughput distribution (P_{ij}) and the delay distribution (D_{ij}) of the line from node *i* to node *j*.

To evaluate *TFA distribution procedure* from the point of view of utilisation of the network elements, let us determine the *utilisation matrix* $\underline{U} = [u_{ij}]$ that gives us the evaluation of the load of every node and link:

u_{ii} = (the average number of packets - defined by the arithmetic mean of the load in the steady state - for the node *i*)/ k_{ii} ,

u_{ij} = (the average number of Mbits - defined by the arithmetic mean of the load in the steady state – for the line from node *i* to node *j*)/ k_{ij} .

For the evaluation of the *uncertainty* of results of the given spatial distribution, the *sample evaluation matrix* $\underline{S} = [s_{ij}]$ is introduced. Matrix elements may have the following values:

s_{ii}, s_{ij} = 3, *small sample*, if the number of RUs through a given or line < S_S

s_{ii}, s_{ij} = 2, *medium sample*, if the number of RUs through a given node or line is between S_S and S_L, $S_S \leq$ the number of RUs $\leq S_L$

s_{ii}, s_{ij} = 1, *large sample*, if the number of RUs through a given node or line > S_L

The matrix element s_{ij} = 0 for an application if there was no RU sent through line from node *i* to node *j* and s_{ij} = 0 for all applications if there is no line from node *i* to node *j* as it is sure that there is no traffic on a non-existing line.

For the $\underline{S}$ matrix to be used, the *limits* S_S and S_L should be determined for applications: S_L may be got from the RU defining procedure described in the previous points, the value S_S may for example be determined according to the sample size to be a *small sample* for sampling (*30*).

6.2. Increasing the number of routing units

Let us examine some typical *conditions* where it is necessary to *increase* the number of RUs for traffic $\mathbb{T}_{qv}$ ($N_{RU}^{\mathbb{T}_{qv}}$).

Condition of *high uncertainty* (*HU*) of traffic distribution: after evaluation of the $\underline{S}$ *matrices,* it may be found that:

$$\exists_{ii}(s_{ii(\mathbb{T}_{qv})} = 3\,) \vee \exists_{ij(i\neq j)}(s_{ij(\mathbb{T}_{qv})} = 3\,) \qquad (12)$$

which expresses that there has been measured *small sample* in the distribution of traffic $\mathbb{T}_{qv}$.

Condition of *uncertain utilisation* (*UU*): after evaluation of the $\underline{U}$ *matrix* and the $\underline{S}$ *matrices,* it may be found that:

$$\exists_{ii}(u_{ii} > U_{node} \wedge (\exists_{\mathbb{T}_{qv}}(s_{ll(\mathbb{T}_{qv})} = 2) \vee \exists_{\mathbb{T}_{qv}}(s_{ln(l\neq n)(\mathbb{T}_{qv})} = 2\,)) \vee$$

$$\vee \exists_{ij(i\neq j)}(u_{ij} > U_{line} \wedge (\exists_{\mathbb{T}_{qv}}(s_{ll(\mathbb{T}_{qv})} = 2) \vee \exists_{\mathbb{T}_{qv}}(s_{ln(l\neq n)(\mathbb{T}_{qv})} = 2\ )) \qquad (13)$$

which means that the *limit* defined for utilisation of the network elements has been exceeded (limit for nodes U_{node} or limit for lines U_{line}) together with a *medium sample* of $\mathbb{T}_{qv}$ in the network.

Condition of *uncertain delay* (*UD*): after evaluation of the $\underline{T}$ *matrix* and the $\underline{S}$ *matrices,* it may be found that:

$$\exists_{\mathbb{T}_{qv}}(D_{\mathbb{T}_{qv}} > D_{limit(\mathbb{T}_{qv})} \wedge (\exists_{\mathbb{T}_{gh}}(s_{ii(\mathbb{T}_{gh})} = 2) \vee \exists_{\mathbb{T}_{gh}}(s_{ij(i\neq j)(\mathbb{T}_{gh})} = 2) \qquad (14)$$

which expresses that the delay *limit* defined for $\mathbb{T}_{qv}$ ($D_{limit(\mathbb{T}_{qv})}$) has been exceeded together with a *medium sample* of $\mathbb{T}_{gh}$ in the examined network.

6.3. Routing unit determination with the analysis of traffic classes

Now, let us examine an example of the routing unit determination based on the analysis of traffic classes. In the example, the capacities of a network should be examined from the point of view of capacity needs of different classes of traffic generated by IP communication services on the network, taking into account the *QoS* (Quality of Service) requirements for IP communication services.

The traffic classification used in the example is similar to the classification described in IP Capacity Planning (IPCP) Framework [2] – the traffic is classified depending on the *degree of delay tolerance* – with the difference that the *class for real-time traffic has been divided into two classes* for the examination, according to the *traffic volume* level that has to be propagated:

Traffic class $\mathbb{T}_1$ – high-volume real-time traffic (e.g. video services)

Traffic class $\mathbb{T}_2$ – low-volume real-time traffic (e.g. voice)

Traffic class $\mathbb{T}_3$ – low-delay tolerant traffic (e.g. web browsing)

Traffic class $\mathbb{T}_4$ – high-delay tolerant traffic (e.g. FTP)

In the TFA modelling, the traffic classes $\mathbb{T}_1$ and $\mathbb{T}_2$ are examined with the condition of high network availability (because of their strict delay requirements) and traffic classes $\mathbb{T}_3$ and $\mathbb{T}_4$ are examined with *decreased* available network capacities (providing the network is capable to manage the situation). The fulfilment of QoS requirements (specified by the ITU (International Telecommunication Union) for each service class) is examined through the allowable delay for the traffic class using the resulting delay distributions of TFA.

The examination *starts* with the *statistical* determination of S_{RU}s, decision about S_S and S_L values and determination of *U* and *D limitations* (the capacity of the network is defined by the $\underline{K}$ matrix). The following steps of the *TFA examination* (*Step-1 - Step-4*) are not precise descriptions of the examination steps but they tend to be an illustration for the use of routing unit number-decisions in the process of examination:

Step-1

TFA examination of the network with traffic load $\mathbb{T}_{Step-1} = \bigcup_{v=1}^{V} \mathbb{T}_{1v}$ (i.e. the high-volume real-time traffic)

Execution of the spatial distribution: S_{RU} decisions with the examination of *HU*, *UU* and *UD* conditions

Delay evaluation according to QoS requirements

Step-2

TFA examination of the network with traffic load $\mathbb{T}_{Step-2} = \bigcup_{q=1}^{2} \bigcup_{v=1}^{V} \mathbb{T}_{qv}$ (i.e. the high-volume and the low-volume real-time traffic together)

Execution of the spatial distribution: S_{RU} decisions with the examination of *HU*, *UU* and *UD* conditions

Delay evaluation according to QoS requirements

Step-3

TFA examination of the network with traffic load $\mathbb{T}_{Step-3} = \bigcup_{v=1}^{V} \mathbb{T}_{3v}$ (i.e. the low-delay tolerant traffic)

The *decreased available network capacity* for the traffic class $\mathbb{T}_3$ is calculated as follows:

$$k_{ii(\mathbb{T}_{Step-3})} = k_{ii}\ (1\text{-}u_{ii\,(\mathbb{T}_{Step-2})}),\ k_{ij(\mathbb{T}_{Step-3})} = k_{ij}\ (1\text{-}u_{ij\,(\mathbb{T}_{Step-2})}) \qquad (15)$$

Execution of the spatial distribution: S_{RU} decisions with the examination of *HU*, *UU* and *UD* conditions

Delay evaluation according to QoS requirements

Step-4

TFA examination of the network with traffic load $\mathbb{T}_{Step-4} = \bigcup_{v=1}^{V} \mathbb{T}_{4v}$ (i.e. the high-delay tolerant traffic)

The *decreased available network capacity* for the traffic class $\mathbb{T}_4$ is calculated as follows:

$$k_{ii(\mathbb{T}_{Step-4})} = k_{ii(\mathbb{T}_{Step-3})}\,(1\text{-}u_{ii(\mathbb{T}_{Step-3})}),\ k_{ij(\mathbb{T}_{Step-4})} = k_{ij(\mathbb{T}_{Step-3})}\ (1\text{-}u_{ij(\mathbb{T}_{Step-3})}) \qquad (16)$$

Execution of the spatial distribution: S_{RU} decisions with the examination of *HU*, *UU* and *UD* conditions

Delay evaluation according to QoS requirements

Remark: Of course, a better utilisation of the network may be achieved with the same QoS parameters if the traffic in the network (with all the traffic classes) is handled together.

7. Conclusions

We have defined a tree model of the aggregated traffic (ATTM) in the examined network and based on this model, a statistical algorithm – which serves for the determination of the size of the routing unit of TFA – has been introduced.

We have introduced formal description for the networks and traffic conditions such as: capacity, cost, traffic and utilization matrices as well as metrics for the difference of the traffic and utilization matrices.

On the basis of the statistical constraints on sample size, we have introduced the sample evaluation matrix, the elements of which express if the number of RUs are high enough for a given node or line.

We have shown how the S_{RU} can be dynamically controlled during the spatial distribution phase of TFA.

We have also shown on an example how the QoS parameters of traffic classes can be taken into account in determining S_{RU}.

We conclude that with our results on the appropriate choice of the S_{RU} of TFA, these methods have been matured for implementation.

References

[1] Fujimoto, R. M.: *Parallel Discrete Event Simulation*, Communications of the ACM, Vol. 33, No. 10, (1990), pp. 31-53.

[2] Gareth, D., Hardt, M., Kelly, F.: *Come the Revolution – Network Dimensioning, Service Costing and Pricing in a Packet Switched Environment*, Telecommunications Policy, Vol. 28, No. 5-6, (2004), pp. 391-412

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

[4] Lencse, G.: *Efficient Parallel Simulation with the Statistical Synchronization Method*, in Communication Networks and Distributed Systems Conference, San Diego, (1998), pp. 3-8

[5] Lencse, G.: *Applicability Criteria of the Statistical Synchronization Method,* in Communication Networks and Distributed Systems Conference, San Francisco, (1999), pp. 159-164

[6] Lencse, G.: *Traffic-Flow Analysis for Fast Performance Estimation of Communication Systems*, Journal of Computing and Information Technology, Vol. 9, No. 1, (2001), pp. 15-27.

[7] Lencse, G.: *Speeding up the Performance Analysis of Communication Systems*, in 2005 European Simulation and Modelling Conference, Porto, Portugal, (2005) pp. 329-333

[8] Lencse, G., Muka, L.: *Convergence of the Key Algorithm of Traffic-Flow Analysis*, Journal of Computing and Information Technology, Vol. 14, No 2, (2006), pp. 133-139

[9] Lencse, G., Muka, L.: *Investigation of the Spatial Distribution Algorithm of the Traffic Flow Analysis and of the Entity Flow-Phase Analysis*, in 2007 European Simulation and Modelling Conference, St. Julians, Malta, (2007) pp. 574-581

[10] Muka, L., Lencse, G.: *Cooperating Modelling Methods for Performance Evaluation of Interconnected Infocommunication and Business Process Systems*, in 2008 European Simulation and Modelling Conference, Le Havre, France, (2008) pp. 404-411

[11] Pongor, Gy.: *Statistical Synchronisation: a Different Approach to Parallel Discrete Event Simulation*, in 1992 European Simulation Symposium, Dresden, (1992), pp. 125-129.

Laser Beam Scanner for Quality Evaluation of Pre-sensitized Offset Plates

A. Skibniewski[(1)], K. Kluszczyński[(2)], T. Trawiński[(2)], Z. Pilch[(2)], M. Szczygieł[(2)], P. Kielan[(2)]

[(1)]FUJIFILM Manufacturing Europe B.V., Tilburg, the Netherlands.
[(2)]Silesian University of Technology, Faculty of Electrical Engineering, Department of Mechatronics, Gliwice, Poland.
e-mail: andre_skibniewski@fujifilm.eu, tomasz.trawinski@polsl.pl

Abstract: The paper contains description of a newly built prototype of a laser scanner for evaluation of PS plate quality. Detailed information about the construction and operation algorithm of the apparatus is given, together with measurement results of scanned plates. Finally conclusions and recommendations are given.

Keywords: pre-sensitized plates, offset technology, guillotine cutting, laser scanning.

1. Introduction

Pre-sensitized plates are very commonly used for all printing works in the offset technology. They are made of thin, mostly 0.3 mm aluminium sheets, covered by multiple layers of light sensitive substances. Hundreds of millions square meters of this material are produced worldwide and used for printing books, newspapers, billboards etc. Offset technology is applicable always where the highest printing quality has to be combined with long run lengths of printed materials.

Some finishing operations in the production process require cutting larger sheets of PS plates into specified sizes on guillotines. Compliance with the ISO standard [1] requires from manufacturers not only meeting strict requirements regarding the plate dimensions. Also plate's edge quality is of a crucial importance. Size of burrs on the edge is limited to 45 μm, when the best manufacturers strive to deliver products with defects substantially below this limit. Deformations to the plate's edges caused by cutting can have detrimental influence on the plate's performance (e.g. "Edge toning") can contaminate or damage the rollers of the printing press. Sharp, uneven edges can also cause injuries to the press operators. Especially the newest plates made for Computer to Plate (CtP) printing technology require almost ideal, undisturbed edge quality of the finished plate. The cutting process on guillotines had been investigated extensively in recent years [2],[3], however still much more research is required to resolve many issues related to a very specific behaviour of the PS plates.

One of the side effects of guillotine cutting is a specific "edge roughness", caused by small, brittle particles originating from the treated aluminium surfaces. They adhere to the plate's edge (Figure 1) and can relatively easily contaminate rollers of the printing

press or a plate setter. Severity of this defect cannot be measured with commonly used surface roughness testers, due to fragile bonds between the particles and the plate's surface (particles can be removed by the stylus tip). Therefore, the new test stand should be built to allow contactless scanning of the products. It should not only be affordable and reliable – but above all it should take into account all specific properties of PS plates, should work in difficult light conditions, giving accurate and reliable data describing the plate quality. In addition, it should take into account previous experiences of the manufacturer, compiled in many years of PS plates manufacturing.

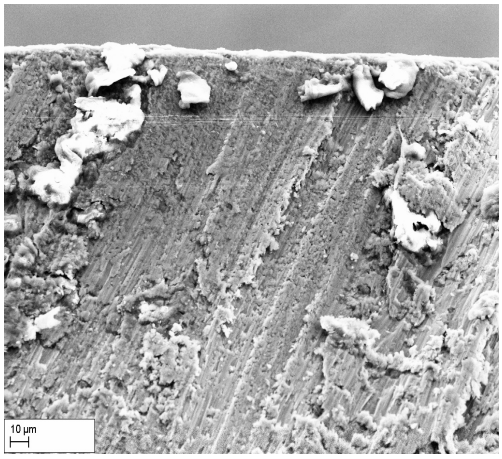


Figure 1. Microscopic view of the PS plate edge after guillotine cutting

As a result of close cooperation between the Silesian University of Technology and the FUJIFILM Manufacturing Europe Company, a prototype of such a test stand had been built recently. It already proved its usefulness in industrial conditions, enabling more effective quality control of pre-sensitized plates.

2. Components of the PS Plate Edge Scanner

2.1. Description of mechanical components of the test stand for evaluation of PS plate edges

The test stand had been constructed with the steel elements, with the steel frame DIN EN 10210-2 - 100 × 50 × 4 to which the linear Festo module is fixed.

A step motor, with a gearbox and a tooth belt drives the linear module. Support for the plate samples and their fixation gives a sample carrier, fixed to the dolly of the linear drive. Sample plates are fixed to the carrier by a magnetic gripper. The scanning heads are placed against each other in such a way that the samples are placed against measuring head close to its measuring range. Four legs, adjustable in two directions support the stand. A general view of the apparatus is shown in Figure 1.

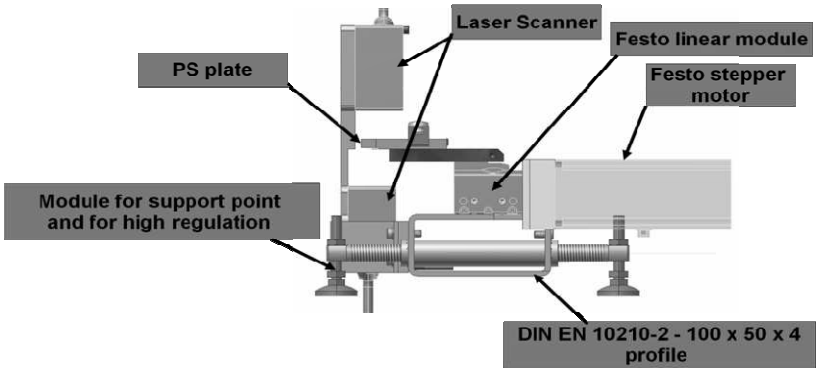


Figure 2. Schematic view of the test stand for scanning PS plates

For the bearing structure of the stand (bearing frame and the legs) basic frequencies of free vibrations and corresponding to them shape of vibrations have been determined. Results of these calculations are shown in Figure 2. and Figure 3.

Figure 3. Shape of the eigen vibrations at frequency 147 Hz

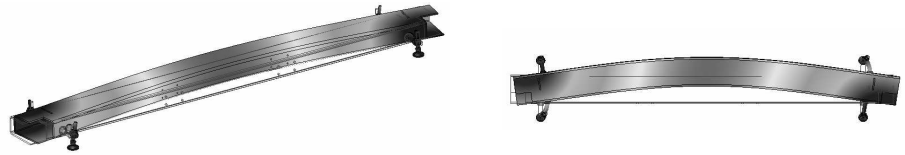


Figure 4. Shape of the eigen vibrations at frequency 246 Hz

2.2. Algorithm of input and processing measurement data

Transducer of the Keyence laser head stores scanning results in the base memory. After the scanning process is completed, the user reads the data with the use of LJ – Navigator software and saves the "*.csv" file at the required destination.

It is assumed that the "*.csv" data file stores at least 256 scanning results, containing data of scanning from the upper and lower scanning head.

Algorithm of the primary data processing **AWO** is as follows: after the data input, fragment of data related to the shape of samples is filtered out and results are divided into two matrixes representing scanned shape of the upper and lower surface of the plate. The algorithm in its graphical form is shown in Figure 5.

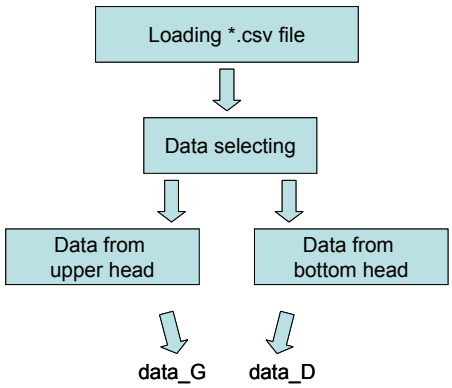


Figure 5. Algorithm of primary data processing AWO

As a result of the above mentioned algorithm two matrixes: "data_D" (bottom) and "data_G" (top) are obtained. Size of one matrix is limited to 200 columns and 256 rows, while recommended number of scans is 200. An example of a matrix representing scanning results (distances from laser heads) is shown in Figure 6 and Figure 7. One corner of the sample was placed approximately 120 μm higher than the other one, what is clearly visible on the graphs.

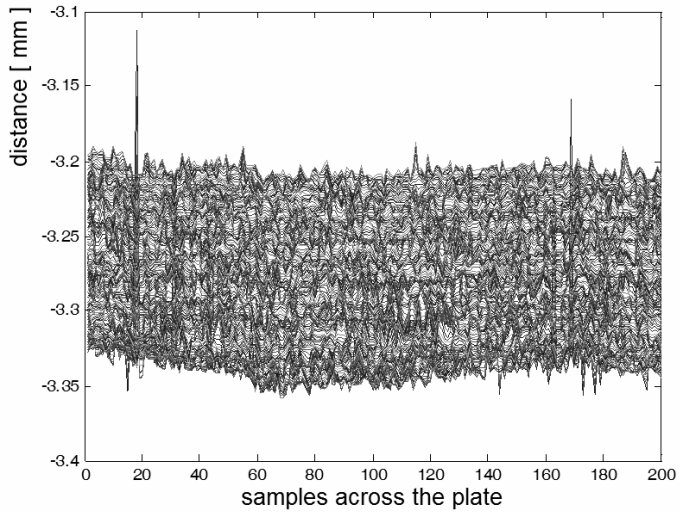


Figure 6. Example of data contained in the matrix "data_D", generated by the bottom laser head

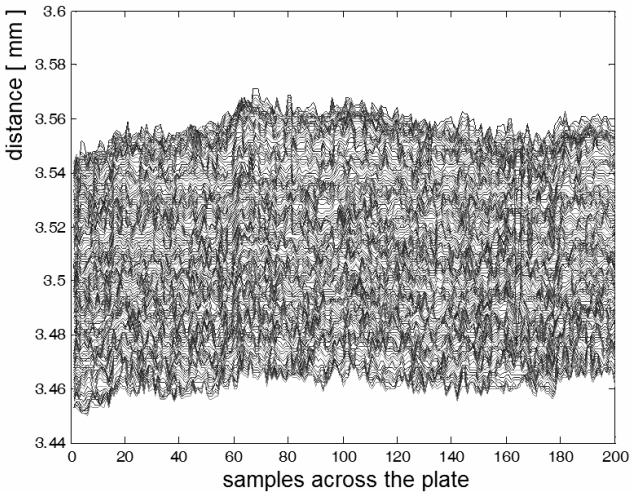


Figure 7. Example of data contained in the matrix results "data_G" from the upper laser head

Data contained in "data_D" and "data_G" matrixes can be represented also by a surface graph, shown as an example in Figure 8. and Figure 9. Structures of the lower and upper surfaces of the plate can be seen very clearly.

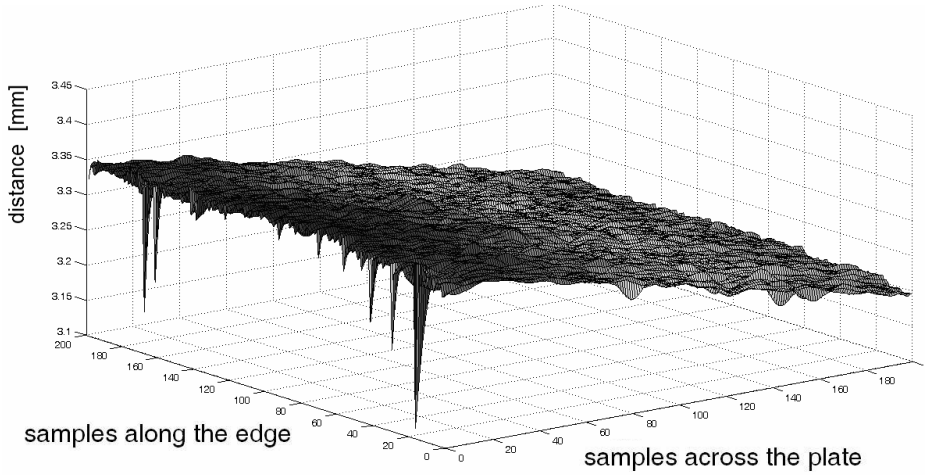


Figure 8. Upper surface of the PS plate

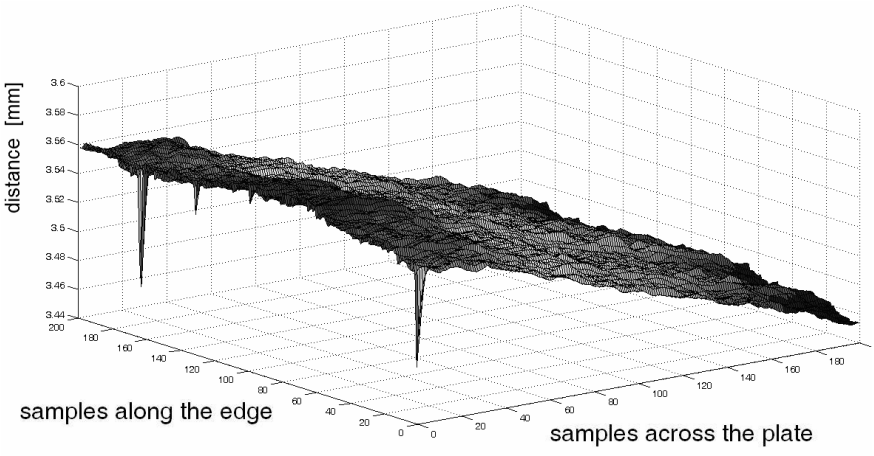


Figure 9. Lower surface of the PS plate

It is possible to "add" these two parts together and as a result to obtain an accurate estimation of the plate thickness. „Assembly" of the lower and upper surface of the plate is shown in Figure 10.

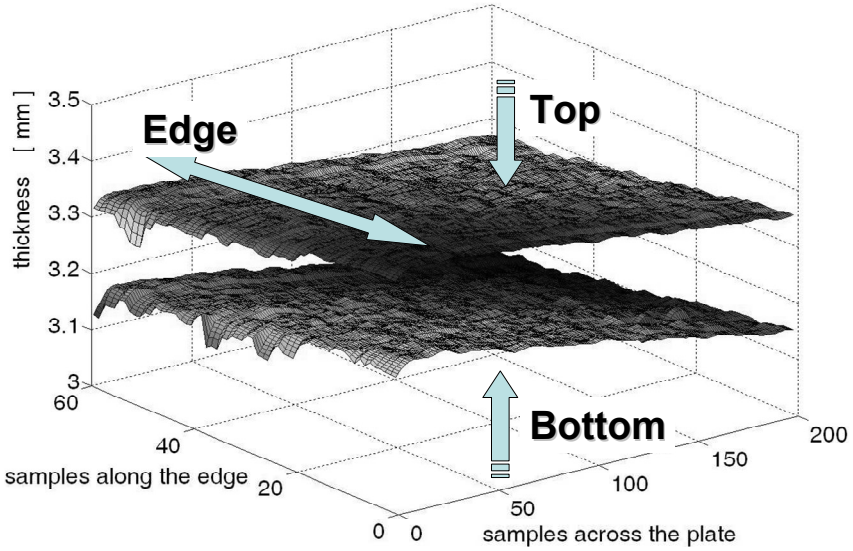


Figure 10. Fragment of the scanned plate

The next algorithm used for data processing is **Crossections_1**. Its goal is to retrieve data from the "data_G" and "data_D", files and to discard from analysis 30 measurement points taken at the vicinity of the plate's edge (measurements can be influenced by edge folding or other defects). It also adds linear regression of the remaining points for every scan of each side of the plate. Linear regression is obtained through an auxiliary function called **reg_lin_2**. Areas influenced by the regression, for a single scan of upper and lower plate surface are shown in Figure 10.

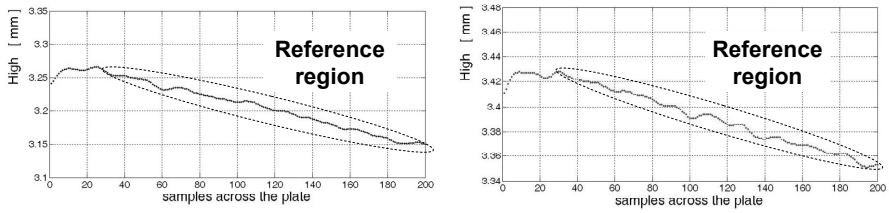


Figure 11. Areas where linear regression is applied (upper part of the plate – drawing on the LHS, lower part of the plate – RHS)

Straight line interpolating the results from the selected area is extrapolated to measurement points outside the interpolated area, close to the plate's edge. Distances between this straight line and actual measurement points are calculated, as shown as bar plots in Figure 12. and Figure 13. Consequently it is possible to evaluate deformation of the whole area close to the edge of the scanned plate. The described above procedure is repeated for each registered scans.

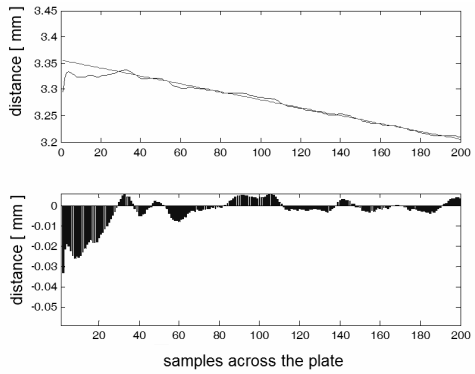


Figure 12. Results of interpolation and extrapolation of measurement data (single scan) for the lower part of the plate

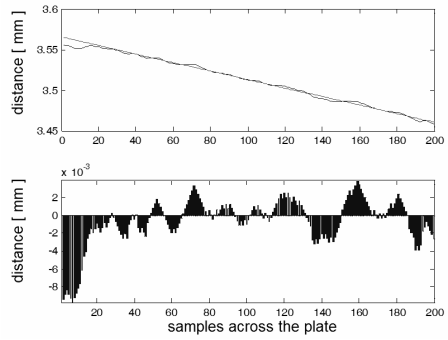


Figure 13. Results of interpolation and extrapolation of the measurement data (single scan) for the upper side of the plate

Similarly data taken at the plate's edge are interpolated, however only 30 measurement points are taken into account. Also in this case distances between actual scan measurements and interpolating straight line are calculated. Example of scan measurements taken at the edge area is shown in Figure 14. and in Figure 15.

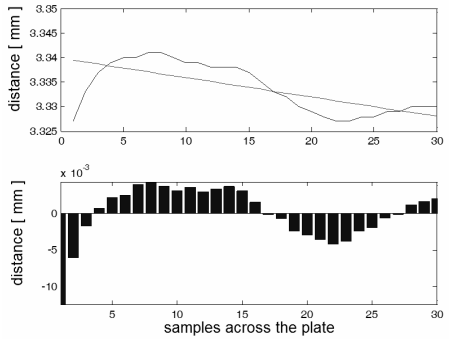


Figure 14. Results of interpolation for measurements taken at the vicinity of the plate's edge (single scan), lower side of the plate

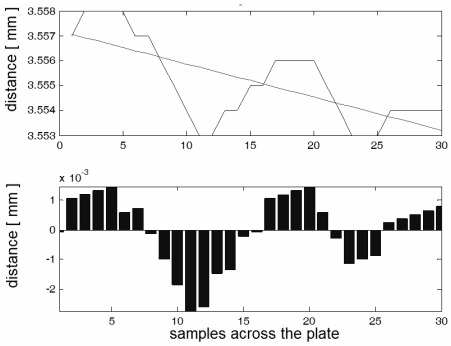


Figure 15. Results of interpolation for measurements taken at the vicinity of the plate's edge (single scan), upper side of the plate

The next important part of the algorithm, enabling us to evaluate the quality of the plate after cutting is a method of representing obtained results in the form of matrices named "bars_D" and "bars_G". First columns of the said matrixes represent the shape of the plate's edge (lower and upper, respectively). For unbiased evaluation of the quality of the plate the following formula have been established:

$$RF = \frac{\sum_k S_{Ak} + \sum_k S_{Bk}}{Z - k} k \cdot 100\,, \tag{1}$$

where: RF – empirical "Roughness Factor", S_{Ak} and S_{Bk} – height of the single bar from the bottom and the upper side, Z – total number of scans (i.e. 200), k – number of measurements with bars higher than 10 μm,

Taken into account measurement results of over 500 plate samples, the following quality criteria have been established (Table 1):

Table 1. Classification of the PS plate quality

RF - scale	Class of defect	Description of edge quality
<=10	A	Very good
>10 and <= 40	B	Good
>40 and <= 100	C	Questionable
>100	D	Bad

2.3. Software for evaluation the PS plate's edge quality

The main program for evaluation of the PS plate roughness is PSPA (Pre-Sensitized Plates Analysis). The program analyses the data, taking into account deviations from a straight line of the plates edge after cutting. Information regarding the class of irregularities at the plates edge is given as a class of defect (A, B, C, D), together with a figure – roughness coefficient ("Roughness Factor").

The program had been created in the MATLAB environment, with the use of the GUI (The Graphical User Interface) module. The PSPA software is dedicated to users of various level of advancement. Figure 16 shows the start window of the program. The first "Basic Level" is dedicated to the guillotine operators. Due to the character of their work it is not necessary to show all data related to the tested sample, except the most important and basic information. In this case the user is not able to display the graphs related to the measurement data. Only information regarding the value of the "Roughness Factor" and classification of the severity of the defect is displayed.

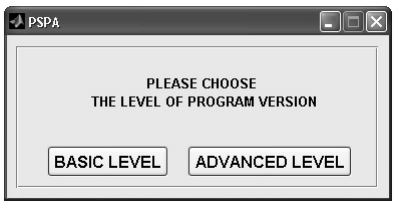


Figure 16. Start window of the PSPA program

The operator can insert his personal data, like his (nick)name, the team/shift he is working for, the guillotine number. After the test is completed he can see the measurement results (Figure 17.).

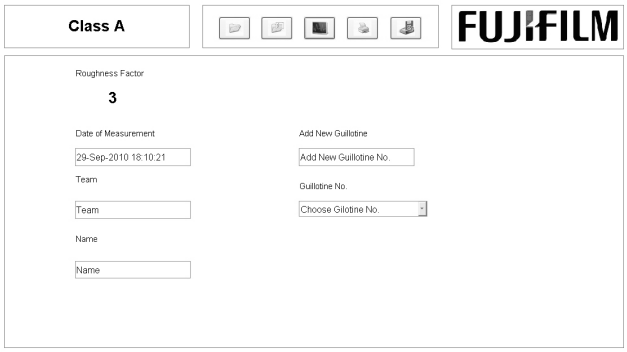


Figure 17. View of the Main window for the "Basic Level" version

The "Advanced Level" version, dedicated to the management staff and engineers allows the user to access much wider range of the program tools. The measurement results can be shown as linear graphs, bar or surface charts (Figure 18–21.). As in case of the "Basic Level" version, the value of the Roughness Factor is displayed as well.

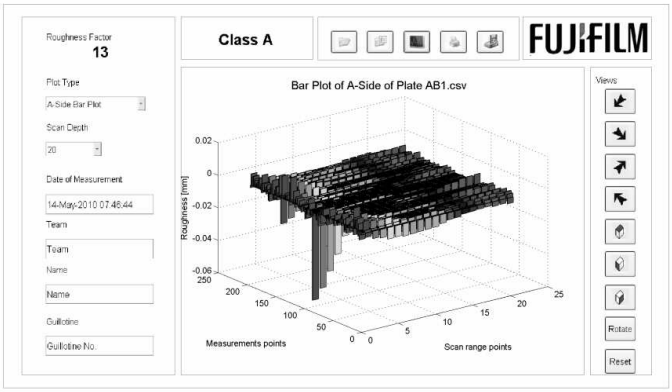


Figure 18. The view of the main window of the "Advanced Level" program

The main window displayed on the screen for the "advanced" user the can be divided into the four basic areas. The first one is located in the central part of the window and enables visualization of the measurement results in the form of graphs. On the left hand side of the window the "User's panel" had been placed (Figure 19.), allowing data input by the operator. The operator can choose the type of graph and the scope of measurements in the sample's cross direction. In addition he must put in his personal data, like his name, team guillotine number etc. in the same way as the Basic Level user.

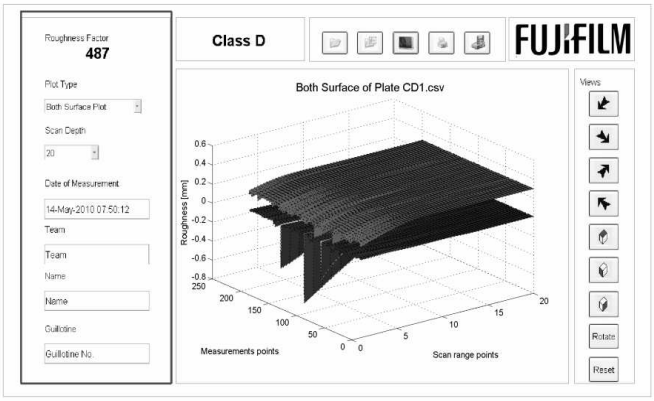


Figure 19. View of the window "Advanced Level" program – user's panel

On the right hand side of the screen the "View panel" had been placed (Figure 20.). It allows a free change between the two dimensional and isometric 3D views of graphs, especially useful when a very detailed analysis of the plate's condition is required. The last element of the program is the "Data Analysis Panel", shown in Figure 20. and Figure 21. This panel contains tools for settings of file folders, charts, print out formats and file saving options. The software does not require entry of new data, because the procedure analyses the data contained in "*.csv" files stored in the hard disk drive. If necessary the old files can be analyzed again and compared with the more recent ones.

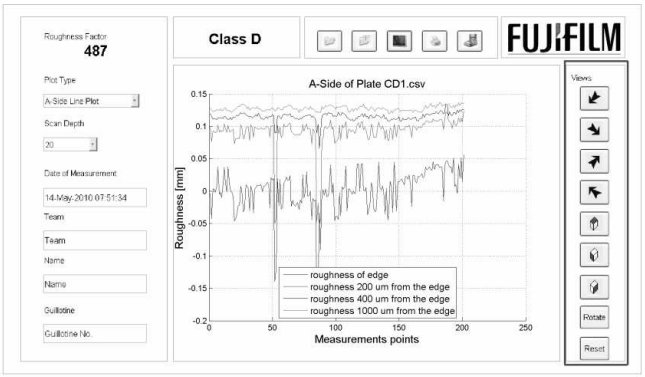


Figure 20. View of the "Advanced Level" panel window – options for data analysis

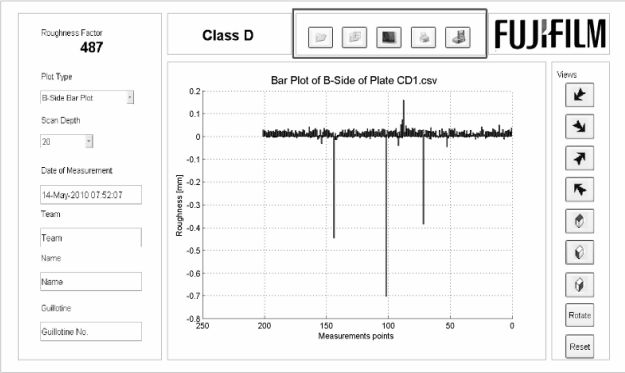


Figure 21. View of the "Advanced Level" panel window – options for printing and file saving

The software had been designed with the aim to simplify the usage of the graphic interface by the operator and to maximize the clarity and legibility of presented information. The necessary disk memory is limited due to the small size of the text file.

3. Conclusions

Test results performed on over 500 material samples had proven good performance of the test stand prototype. Classification of the severity of defect (Class of Roughness) by the apparatus and the organoleptic judgment by the operators in over 85% of cases correspond with each other. It creates a very good starting point for further development of the apparatus, aiming at easy and precise evaluation of other aspects the plate's quality like edge folding and cutting burrs.

It will not only help to assure product quality during the production process, but will open new perspectives in research of the pre-sensitized offset plates, adding a very useful tool where non destructive testing is required.

References

[1] ISO 12635:2008. *Graphic technology – Plates for offset printing – Dimensions* International Standards Organization (2008).

[2] Kaczmarczyk J., Gąsiorek D., Mężyk A., Skibniewski A.: A*nalysis of rising defects causes in an established cutting process of plates using guillotines*. Modelowanie Inżynierskie No 34 (2007), Gliwice, pp. 61-66, (in polish)

[3] Mężyk A., Gąsiorek D., Kaczmarczyk J., Rak Z., Skibniewski A.: *Research of cutting process of plates sets on guillotines,* Przegląd Mechaniczny Vol.5/10 (2010), pp. 36-38, (in polish),

[4] Polozy G. N. et al.: *Proximate computing methods*, Wydawnictwa Naukowo Techniczne, WNT Warszawa (1966), (in polish),

Youla-Parametrized Controllers for Stable Multivariable Processes

L. Keviczky, Cs. Bányász

Széchenyi University, Győr
Computer and Automation Research Institute
and Control Engineering Research Group Hungarian Academy of Sciences
H-1111 Budapest, Kende u 13-17, HUNGARY
Phone: +361-466-5435; Fax: +361-466-7503
e-mail: keviczky@sztaki.hu; banyasz@sztaki.hu

Abstract: The *Youla-parametrization* is a useful method to design controllers for open-loop stable plants. The *KB-parametrization* is a successful extension of this method for two-degree-of freedom (*2DOF*) systems. The paper extends this methodology for multivariable case.

Keywords: Youla-parametrization, multivariable process, MIMO controller

1. Introduction to Multivariable Process Models

Consider a *M*ulti-*I*nput-*M*ulti-*O*utput (*MIMO*) continuous time *L*inear *T*ime *I*nvariant (*LTI*) dynamic plant described by the state variable representation (*SVR*)

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \dot{\boldsymbol{x}} &= \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u} \\ \boldsymbol{y} &= \mathbf{C}^{\mathrm{T}}\boldsymbol{x} + \mathbf{D}\boldsymbol{u} \end{aligned} \tag{1}$$

Here $\boldsymbol{u}$, $\boldsymbol{y}$ and $\boldsymbol{x}$ are the input, output and state variables of the process to be controlled and $^{\mathrm{T}}$ stands for transposition.

The transfer function representation (*TFR*) of the *MIMO* plant can be calculated by

$$\mathbf{P}(s) = \mathbf{C}^{\mathrm{T}}\left(s\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} = \mathbf{C}^{\mathrm{T}}\boldsymbol{\Phi}(s)\mathbf{B} + \mathbf{D} = \frac{1}{\mathcal{A}(s)}\mathcal{B}(s) \tag{2}$$

where

$$\boldsymbol{\Phi}(s) = \left(s\mathbf{I} - \mathbf{A}\right)^{-1} = \frac{1}{\mathcal{A}(s)}\boldsymbol{\Psi}(s) \quad ; \quad \boldsymbol{\Psi}(s) = \mathbf{adj}\left(s\mathbf{I} - \mathbf{A}\right) \tag{3}$$

and the scalar denominator polynomial

$$\mathcal{A}(s) = \det\left(s\mathbf{I} - \mathbf{A}\right) \tag{4}$$

is the characteristic polynomial. Assume that the process is square, i.e. the number of input and output variables is the same: p. So $\mathbf{C}$ is a $p \times n$, $\mathbf{B}$ is an $n \times p$, $\mathbf{D}$ is a $p \times p$, $\mathbf{A}$ is an $n \times n$ constant parameter matrix, furthermore $\mathbf{I}$ is the $n \times n$ unit matrix. Consequently the degree of $\mathcal{A}(s)$ is also n. The $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are $n \times n$ polynomial matrices, $\boldsymbol{P}(s)$ is an $n \times n$ transfer function matrix (TFM).

The formal "numerator" matrix polynomial

$$\mathcal{B}(s) = \mathbf{C}^{\mathrm{T}}\mathbf{\Psi}(s)\mathbf{B} + \mathcal{A}(s)\mathbf{D}\Big|_{\mathbf{D}=0} = \mathbf{C}^{\mathrm{T}}\mathbf{\Psi}(s)\mathbf{B} \tag{5}$$

is commutative with the scalar denominator $\mathcal{A}(s)$. So the form (2) of $\boldsymbol{P}(s)$ is the simplest *MIMO* model, which is not irreducible, however does not involve the more complex forms of the left and right *MFD*'s (*M*atrix *F*ractional *D*escription, see later). It is sometimes called "naive" model representation.

A *MIMO* TFM $\boldsymbol{P}(s)$ can be transformed into the so-called SMITH-form

$$\boldsymbol{P}(s) = \mathcal{U}(s)\,\boldsymbol{D}(s)\,\mathcal{V}(s) \tag{6}$$

where $\mathcal{U}(s)$ and $\mathcal{V}(s)$ are unimodular polynomial matrices. (A matrix is unimodular, if its determinant is a scalar constant. The inverses of unimodular matrices are also polynomial ones.) The matrix $\boldsymbol{D}(s)$ is a special diagonal TFM

$$\begin{aligned}\boldsymbol{D}(s) &= \boldsymbol{D}_{\mathrm{D}}(s)\mathcal{D}_{\mathrm{N}}(s) = \mathcal{D}_{\mathrm{N}}(s)\boldsymbol{D}_{\mathrm{D}}(s) = \mathcal{D}_{\mathrm{D}}^{-1}(s)\mathcal{D}_{\mathrm{N}}(s) = \mathcal{D}_{\mathrm{N}}(s)\mathcal{D}_{\mathrm{D}}^{-1}(s) = \\ &= \mathbf{diag}\left\langle \frac{\varepsilon_1}{\psi_1}, \frac{\varepsilon_2}{\psi_2}, \ldots, \frac{\varepsilon_{\mathrm{r}}}{\psi_{\mathrm{r}}}, 0, \ldots, 0 \right\rangle = \mathbf{diag}\left\langle \lambda_1, \lambda_2, \ldots, \lambda_{\mathrm{r}}, 0, \ldots, 0 \right\rangle\end{aligned} \tag{7}$$

where the diagonal matrices are commutative. Here

$$\mathcal{D}_{\mathrm{N}}(s) = \mathbf{diag}\left\langle \varepsilon_1(s), \varepsilon_2(s), \ldots, \varepsilon_{\mathrm{r}}(s), 0, \ldots, 0 \right\rangle ; \lambda_i(s) = \frac{\varepsilon_i(s)}{\psi_i(s)} \tag{8}$$

and

$$\boldsymbol{D}_{\mathrm{D}}(s) = \mathcal{D}_{\mathrm{D}}^{-1}(s) \quad ; \quad \mathcal{D}_{\mathrm{D}}(s) = \mathbf{diag}\left\langle \psi_1(s), \psi_2(s), \ldots, \psi_{\mathrm{r}}(s), 0, \ldots, 0 \right\rangle \tag{9}$$

The $\varepsilon_i(s)$ and $\psi_i(s)$ are coprime polynomials. The diagonal $\boldsymbol{D}(s)$ TFM can be expressed from $\boldsymbol{P}(s)$ as

$$\boldsymbol{D}(s) = \mathcal{U}^{-1}(s)\,\boldsymbol{P}(s)\,\mathcal{V}^{-1}(s) \tag{10}$$

and is called invariant equivalent (SMITH-MCMILLAN) form of the *MIMO* process. The number of nonzero element r in the main diagonal of these matrices is the so-called

MCMILLAN-*degree*. It is obvious that $\mathcal{A}(s)$ is the least common multiple of the denominators and $n \geq r$. (The typesetting italic bold means *MIMO* TFM's, italic bold Euclid fonts mean *MIMO* polynomial matrices, whose product is generally not interchangeable !!!)

The SMITH-form (6) is a good basis to rewrite the TFM $\boldsymbol{P}(s)$ into equivalent *MFD*'s, where the parts of the representation are matrix polynomial matrices only. The TFM $\boldsymbol{P}(s)$ has two equivalent *MFD* forms

$$\boldsymbol{P}(s) = \boldsymbol{\mathcal{N}}_{\mathrm{R}}(s)\boldsymbol{\mathcal{D}}_{\mathrm{R}}^{-1}(s) = \boldsymbol{\mathcal{D}}_{\mathrm{L}}^{-1}(s)\boldsymbol{\mathcal{N}}_{\mathrm{L}}(s) \tag{11}$$

where $\boldsymbol{\mathcal{N}}_{\mathrm{R}}(s)$ and $\boldsymbol{\mathcal{D}}_{\mathrm{R}}(s)$ are the right, $\boldsymbol{\mathcal{N}}_{\mathrm{L}}(s)$ and $\boldsymbol{\mathcal{D}}_{\mathrm{L}}(s)$ are the left *MFD* polynomials, which can be – at least formally – easily computed

$$\boldsymbol{\mathcal{D}}_{\mathrm{L}}(s) = \boldsymbol{\mathcal{D}}_{\mathrm{D}}(s)\boldsymbol{\mathcal{U}}^{-1}(s) \quad ; \quad \boldsymbol{\mathcal{N}}_{\mathrm{L}}(s) = \boldsymbol{\mathcal{D}}_{\mathrm{N}}(s)\boldsymbol{\mathcal{V}}(s) \tag{12}$$

and

$$\boldsymbol{\mathcal{D}}_{\mathrm{R}}(s) = \boldsymbol{\mathcal{V}}^{-1}(s)\boldsymbol{\mathcal{D}}_{\mathrm{D}}(s) \quad ; \quad \boldsymbol{\mathcal{N}}_{\mathrm{R}}(s) = \boldsymbol{\mathcal{U}}(s)\boldsymbol{\mathcal{D}}_{\mathrm{N}}(s) \tag{13}$$

Note that the so-called transmission zeros of $\boldsymbol{P}(s)$ are the roots of $\prod_{i=1}^{r} \varepsilon_i(s)$.

In practical controller design it is reasonable to assume that the multivariable process $\boldsymbol{P}$ is factorable to an inverse stable $\boldsymbol{P}_+$ (*IS*) and to an inverse unstable $\boldsymbol{P}_-$ (*IU*) matrix operator

$$\boldsymbol{P} = \boldsymbol{P}_-\boldsymbol{P}_+ \neq \boldsymbol{P}_+\boldsymbol{P}_- \tag{14}$$

One must know that the *MIMO* plant form (14) can be rewrite into the equivalent

$$\boldsymbol{P} = \bar{\boldsymbol{P}}_+\bar{\boldsymbol{P}}_- \neq \bar{\boldsymbol{P}}_-\bar{\boldsymbol{P}}_+ \tag{15}$$

form. Based on the left and right *MFD* forms of the *MIMO* process the (14) and (15) representations can be written as

$$\boldsymbol{P}(s) = \boldsymbol{\mathcal{N}}_{\mathrm{R}}^{-}(s)\boldsymbol{\mathcal{N}}_{\mathrm{R}}^{+}(s)\boldsymbol{\mathcal{D}}_{\mathrm{R}}^{-1}(s) = \boldsymbol{\mathcal{D}}_{\mathrm{L}}^{-1}(s)\boldsymbol{\mathcal{N}}_{\mathrm{L}}^{+}(s)\boldsymbol{\mathcal{N}}_{\mathrm{L}}^{-}(s) \tag{16}$$

i.e., the factorable parts are

$$\boldsymbol{P}_-(s) = \boldsymbol{\mathcal{N}}_{\mathrm{R}}^{-}(s) \;;\; \boldsymbol{P}_+(s) = \boldsymbol{\mathcal{N}}_{\mathrm{R}}^{+}(s)\boldsymbol{\mathcal{D}}_{\mathrm{R}}^{-1}(s) \;;\; \bar{\boldsymbol{P}}_-(s) = \boldsymbol{\mathcal{N}}_{\mathrm{L}}^{-}(s) \;;\; \bar{\boldsymbol{P}}_+(s) = \boldsymbol{\mathcal{D}}_{\mathrm{L}}^{-1}(s)\boldsymbol{\mathcal{N}}_{\mathrm{L}}^{+}(s) \tag{17}$$

Here we assumed that the above factorization can always be expressed in the "numerator" matrix polynomials, i.e.

$$\boldsymbol{\mathcal{N}}_{\mathrm{R}}(s) = \boldsymbol{\mathcal{N}}_{\mathrm{R}}^{-}(s)\boldsymbol{\mathcal{N}}_{\mathrm{R}}^{+}(s) \quad \text{and} \quad \boldsymbol{\mathcal{N}}_{\mathrm{L}}(s) = \boldsymbol{\mathcal{N}}_{\mathrm{L}}^{+}(s)\boldsymbol{\mathcal{N}}_{\mathrm{L}}^{-}(s) \tag{18}$$

2. *Youla-Parametrization* of *MIMO* Closed-Loop Systems

The formal generalization of the *Y-parametrization* (*YP*) [5] for linear dynamic *MIMO* control systems is relatively simple, introducing the TFM

$$\boldsymbol{Q} = \boldsymbol{C}\left(\mathbf{I} - \boldsymbol{PC}\right)^{-1} = \left(\mathbf{I} - \boldsymbol{CP}\right)^{-1}\boldsymbol{C} \tag{19}$$

and using the matrix regulator

$$\boldsymbol{C} = \left(\mathbf{I} - \boldsymbol{QP}\right)^{-1}\boldsymbol{Q} = \boldsymbol{Q}\left(\mathbf{I} - \boldsymbol{PQ}\right)^{-1} \tag{20}$$

Here $\boldsymbol{P}$ is a stable *MIMO* dynamic process (see Figure 1.), furthermore the *Y-parameter* $\boldsymbol{Q}$ and regulator $\boldsymbol{C}$ are also matrix operators (TFM's). It is not very difficult to prove that the left and right side formulation of $\boldsymbol{Q}$ and $\boldsymbol{C}$ are equivalent.

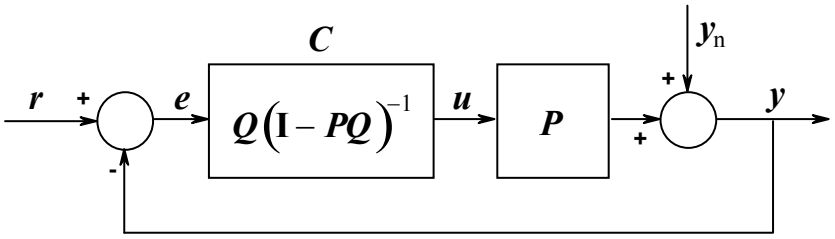


Figure 1. Generalization of the Y-parametrization

The next two relationships play important role in the closed-loop characteristics of the systems:

$$\left(\mathbf{I} + \boldsymbol{PC}\right)^{-1} = \left[\mathbf{I} + \boldsymbol{PQ}\left(\mathbf{I} - \boldsymbol{PQ}\right)^{-1}\right]^{-1} = \mathbf{I} - \boldsymbol{PQ} \quad ;$$

$$\left(\mathbf{I} + \boldsymbol{CP}\right)^{-1} = \left[\mathbf{I} + \left(\mathbf{I} - \boldsymbol{QP}\right)^{-1}\boldsymbol{QP}\right]^{-1} = \mathbf{I} - \boldsymbol{QP} \tag{21}$$

Here the following matrix identities were used:

$$\left[\mathbf{I} + \left(\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{A}\right]^{-1} = \mathbf{I} - \mathbf{A} \quad ; \quad \left[\mathbf{I} + \mathbf{B}\left(\mathbf{I} - \mathbf{B}\right)^{-1}\right]^{-1} = \mathbf{I} - \mathbf{B} \tag{22}$$

After some straightforward (do not miss the interchangeability in matrix operator products) manipulations

$$\boldsymbol{y} = \boldsymbol{P}\,\boldsymbol{Q}\boldsymbol{r} - \left(\mathbf{I} - \boldsymbol{PQ}\right)\boldsymbol{y}_n \tag{23}$$

is obtained for the *YP* closed-loop system in Figure 1. Constructing a special two-degree-of-freedom (*2DOF*) scheme, shown in Figure 2 the closed-loop is "virtually" opened as

$$y = P\,r - \left(\mathrm{I} - PQ\right)y_{\mathrm{n}} \tag{24}$$

This special scheme is called the *KB-parametrization* of a *2DOF* system. This parametrization applies internally a Q^{-1} prefilter to the *YP* scheme in Figure 1. The presented scheme and principle completely corresponds to the methods developed for *SISO* systems [1], [2]. Note that the *KB-parametrization* works for any controller C, not only for YP, and provides $P\,r$ for the tracking properties of the *2DOF* system. The disturbance rejection properties $\left(\mathrm{I} - PQ\right)$ can be obtained only for *YP* loop.

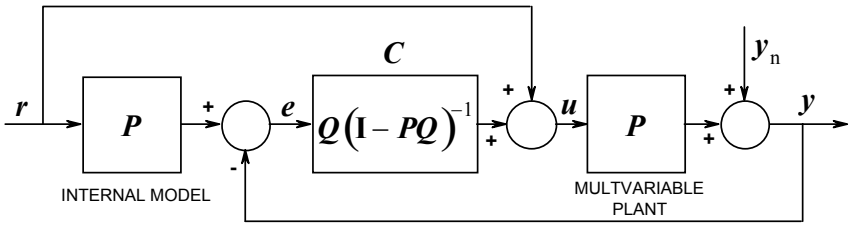


Figure 2. The general form of the K-B-parametrized MIMO 2DOF control system

3. A Generic Controller Scheme for Multivariable Processes

The general form of the *K-B-parametrized MIMO 2DOF* control system makes possible to construct a simple and effective controller scheme, which is called *generic 2DOF* (*G2DOF*) control and shown in Figure 3. The *K-B parametrization* for closed-loop control is not so widely known as the *Youla-Kucera-* (*Y-K*) *parametrization*, however, it is much closer to a control engineering view and its most important advantage in *2DF* systems is that it virtually opens the closed-loop [3], [4]. However, this parametrization can only be applied for open-loop stable processes. The introduced scheme can be referred as a *MIMO G2DF* control system.

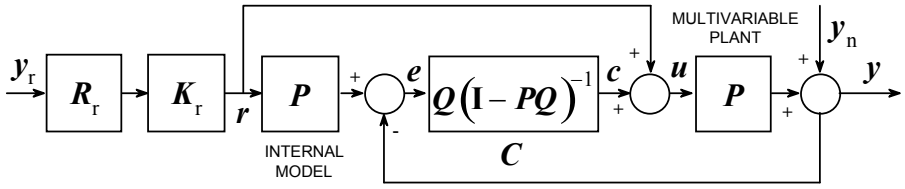


Figure 3. Generalization of the generic 2DOF control system for MIMO processes

It is interesting to see how the transfer characteristics of this *G2DOF MIMO* control system look like:

$$y = PQ_{\mathrm{r}}\,y_{\mathrm{r}} - \left(I - PQ_{\mathrm{n}}\right)y_{\mathrm{n}} = PK_{\mathrm{r}}R_{\mathrm{r}}\,y_{\mathrm{r}} - \left(I - PK_{\mathrm{n}}R_{\mathrm{n}}\right)y_{\mathrm{n}} = y_{\mathrm{t}} + y_{\mathrm{d}} \tag{25}$$

If the multivariable process P is factorable according to (14) then using the proper

selection of

$$\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{n}} = \boldsymbol{K}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}} = \boldsymbol{P}_{+}^{-1}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}} \tag{26}$$

and

$$\boldsymbol{Q}_{\mathrm{r}} = \boldsymbol{K}_{\mathrm{r}}\boldsymbol{R}_{\mathrm{r}} = \boldsymbol{P}_{+}^{-1}\boldsymbol{W}_{\mathrm{r}}\boldsymbol{R}_{\mathrm{r}} \quad ; \quad \boldsymbol{K}_{\mathrm{n}} = \boldsymbol{P}_{+}^{-1}\boldsymbol{W}_{\mathrm{n}} \quad ; \quad \boldsymbol{K}_{\mathrm{r}} = \boldsymbol{P}_{+}^{-1}\boldsymbol{W}_{\mathrm{r}} \tag{27}$$

the following *MIMO* regulator is obtained

$$\boldsymbol{C} = \boldsymbol{Q}\left(\mathbf{I} - \boldsymbol{P}\boldsymbol{Q}\right)^{-1} = \boldsymbol{K}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}}\left(\mathbf{I} - \boldsymbol{P}\boldsymbol{K}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}}\right)^{-1} = \boldsymbol{P}_{+}^{-1}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}}\left(\mathbf{I} - \boldsymbol{P}_{-}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}}\right)^{-1} \tag{28}$$

Based on the previous formulas the transfer characteristics of the closed-loop is now obtained as:

$$\boldsymbol{y} = \boldsymbol{P}_{-}\boldsymbol{W}_{\mathrm{r}}\boldsymbol{R}_{\mathrm{r}}\,\boldsymbol{y}_{\mathrm{r}} - \left(\mathbf{I} - \boldsymbol{P}_{-}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{R}_{\mathrm{n}}\right)\boldsymbol{y}_{\mathrm{n}} = \boldsymbol{y}_{\mathrm{t}} + \boldsymbol{y}_{\mathrm{d}} \tag{29}$$

where $\boldsymbol{y}_{\mathrm{t}}$ means tracking, $\boldsymbol{y}_{\mathrm{d}}$ mean the regulating properties of the *2DOF* closed-loop system. Here $\boldsymbol{K}_{\mathrm{r}}$ and $\boldsymbol{K}_{\mathrm{n}}$ contain the $\boldsymbol{P}_{+}^{-1}$ inverse of $\boldsymbol{P}_{+}$, i.e., the invertible part of $\boldsymbol{P}$, and apply $\boldsymbol{W}_{\mathrm{r}}$ and $\boldsymbol{W}_{\mathrm{n}}$ to attenuate the influence of the invariant *MIMO* process factor $\boldsymbol{P}_{-}$ [2]. (Note that the inverses mean functional inverses and they are not simple fractions.) The above method is a straightforward generalization of the *G2DOF* control system developed for *SISO* processes [4].

If the multivariable process $\boldsymbol{P}$ is factorable according to (15) then an equivalent regulator

$$\bar{\boldsymbol{C}} = \left(\mathbf{I} - \bar{\boldsymbol{Q}}\bar{\boldsymbol{P}}\right)^{-1}\bar{\boldsymbol{Q}} = \left(\mathbf{I} - \boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{K}}_{\mathrm{n}}\bar{\boldsymbol{P}}\right)^{-1}\boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{K}}_{\mathrm{n}} = \left(\mathbf{I} - \boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{W}}_{\mathrm{n}}\bar{\boldsymbol{P}}_{-}\right)^{-1}\boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{W}}_{\mathrm{n}}\bar{\boldsymbol{P}}_{+}^{-1} \tag{30}$$

can also be derived, where

$$\bar{\boldsymbol{K}}_{\mathrm{n}} = \bar{\boldsymbol{W}}_{\mathrm{n}}\,\bar{\boldsymbol{P}}_{+}^{-1} \tag{31}$$

The transfer characteristics of the closed-loop with controller $\bar{\boldsymbol{C}}$ is the following

$$\boldsymbol{y} = \boldsymbol{P}_{-}\boldsymbol{W}_{\mathrm{r}}\boldsymbol{R}_{\mathrm{r}}\,\boldsymbol{y}_{\mathrm{r}} - \left(\mathbf{I} - \boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{W}}_{\mathrm{n}}\bar{\boldsymbol{P}}_{-}\right)\boldsymbol{y}_{\mathrm{n}} = \boldsymbol{y}_{\mathrm{t}} + \bar{\boldsymbol{y}}_{\mathrm{d}} \tag{32}$$

It is easy to see that the $\boldsymbol{y}_{\mathrm{t}}$ is unchanged (because $\boldsymbol{K}_{\mathrm{r}}$ is unchanged), however $\bar{\boldsymbol{y}}_{\mathrm{d}}$ is different.

In the sequel two useful applications, which are connected to further simplified *MIMO* models, will be presented.

4. A *G2DOF* Controller Using the "Naive" *MIMO* Process Model

The *MIMO* process model (2) is the simplest (some times called *naive*) *MIMO* model, which is not irreducible, however does not involve the more complex forms of the left and right *MFD*'s. Its factorable form corresponding to (21) is

$$\boldsymbol{P}(s)=\boldsymbol{P}_{-}\boldsymbol{P}_{+}=\frac{1}{\mathcal{A}(s)}\boldsymbol{\mathcal{B}}(s)=\frac{1}{\mathcal{A}(s)}\boldsymbol{\mathcal{B}}_{-}(s)\boldsymbol{\mathcal{B}}_{+}(s) \tag{33}$$

If we assume that the plant is the simplest "invertible" then the model is given by

$$\boldsymbol{P}=\frac{1}{\mathcal{A}(s)}\boldsymbol{\mathcal{B}}_{+}(s) \quad ; \quad \boldsymbol{\mathcal{B}}_{+}=\boldsymbol{\mathcal{B}} \quad ; \quad \boldsymbol{\mathcal{B}}_{-}=\mathbf{I} \tag{34}$$

Introduce the following reference model transfer matrices

$$\boldsymbol{R}_{\mathrm{r}}=\frac{1}{\mathcal{A}_{\mathrm{r}}(s)}\boldsymbol{\mathcal{B}}_{\mathrm{r}}(s) \quad \text{and} \quad \boldsymbol{R}_{\mathrm{n}}=\frac{1}{\mathcal{A}_{\mathrm{n}}(s)}\boldsymbol{\mathcal{B}}_{\mathrm{n}}(s) \tag{35}$$

If $\boldsymbol{\mathcal{B}}_{-}=\mathbf{I}$ and $\boldsymbol{\mathcal{B}}_{+}=\boldsymbol{\mathcal{B}}$, in (33) then there is no need for further optimization [2], [4], therefore it is reasonable to use the selections $\boldsymbol{W}_{\mathrm{r}}=\boldsymbol{W}_{\mathrm{n}}=\mathbf{I}$, similarly to *SISO* systems. Now using the *Youla-parameter* matrix (22)

$$\boldsymbol{Q}=\boldsymbol{Q}_{\mathrm{n}}=\mathcal{A}(s)\boldsymbol{\mathcal{B}}^{-1}(s)\boldsymbol{R}_{\mathrm{n}}=\frac{\mathcal{A}(s)}{\mathcal{A}_{\mathrm{n}}(s)}\boldsymbol{\mathcal{B}}^{-1}(s)\boldsymbol{\mathcal{B}}_{\mathrm{n}}(s) \tag{36}$$

the optimal *YP* regulator is obtained after some straightforward manipulations as

$$\boldsymbol{C}(s)=\mathrm{A}(s)\mathbf{B}^{-1}(s)\boldsymbol{R}_{\mathrm{n}}(s)\left[\mathbf{I}-\boldsymbol{R}_{\mathrm{n}}(s)\right]^{-1}=\mathrm{A}(s)\mathbf{B}_{+}^{-1}(s)\mathbf{B}_{\mathrm{n}}(s)\left[\mathrm{A}_{\mathrm{n}}(s)\mathbf{I}-\mathbf{B}_{\mathrm{n}}(s)\right]^{-1} \tag{37}$$

For discrete-time system the naive plant model is

$$\boldsymbol{G}(z)=\frac{1}{\mathrm{A}(z)}\mathbf{B}(z) \quad ; \quad \mathbf{B}_{+}=\mathbf{B} \quad ; \quad \mathbf{B}_{-}=z^{-d}\mathbf{I} \tag{38}$$

which results in the following matrix regulator

$$\begin{aligned}\boldsymbol{C}(z)&=\mathcal{A}(z)\boldsymbol{\mathcal{B}}^{-1}(z)\boldsymbol{\mathcal{B}}_{\mathrm{n}}(z)\left[\square_{\mathrm{n}}(z)\mathbf{I}-z^{-d}\boldsymbol{\mathcal{B}}_{\mathrm{n}}(z)\right]^{-1}=\\&=\mathcal{A}(z)\boldsymbol{\mathcal{B}}_{+}^{-1}(z)\boldsymbol{\mathcal{B}}_{\mathrm{n}}(z)\left[\mathcal{A}_{\mathrm{n}}(z)\mathbf{I}-z^{-d}\boldsymbol{\mathcal{B}}_{\mathrm{n}}(z)\right]^{-1}\end{aligned} \tag{39}$$

It is interesting to note that the similarly computed *SISO* discrete-time controller is

$$\begin{aligned}C&=A(z)B^{-1}(z)B_{\mathrm{n}}(z)\left[A_{\mathrm{n}}(z)-z^{-d}B_{\mathrm{n}}(z)\right]^{-1}=\\&=A(z)B_{+}^{-1}(z)B_{\mathrm{n}}(z)\left[A_{\mathrm{n}}(z)-z^{-d}B_{\mathrm{n}}(z)\right]^{-1}\end{aligned} \tag{40}$$

5. A *MIMO* Process Model Linear in Parameter Matrices

It is an important observation that the left *MFD* form of a discrete-time pulse TFM

$$\boldsymbol{G}(z)=\boldsymbol{\mathcal{D}}_{\mathrm{L}}^{-1}(z)\boldsymbol{\mathcal{N}}_{\mathrm{L}}(z) \tag{41}$$

can be easily used for constructing a vector difference equation by computing the left *MFD* of backward shift operator z^{-1} and normalizing by the leading matrix coefficient $\mathbf{D}_{\mathrm{L}}^{0}$, i.e.

$$\begin{aligned} \mathcal{D}_{\mathrm{L}}\left(z^{-1}\right) &= \mathbf{I} + \mathbf{D}_{\mathrm{L}}^{1} z^{-1} + \mathbf{D}_{\mathrm{L}}^{2} z^{-2} + \ldots + \mathbf{D}_{\mathrm{L}}^{m} z^{-m} = \mathbf{I} + \tilde{\mathcal{D}}_{\mathrm{L}}\left(z^{-1}\right) \\ \mathcal{N}_{\mathrm{L}}\left(z^{-1}\right) &= \mathbf{N}_{\mathrm{L}}^{1} z^{-1} + \mathbf{N}_{\mathrm{L}}^{2} z^{-2} + \ldots + \mathbf{N}_{\mathrm{L}}^{m} z^{-m} \end{aligned} \tag{42}$$

The m-order vector difference equation using (42)

$$\begin{aligned} \boldsymbol{y}[k] = \mathcal{N}_{\mathrm{L}}\left(z^{-1}\right)\boldsymbol{u}[k] - \tilde{\mathcal{D}}_{\mathrm{L}}\left(z^{-1}\right)\boldsymbol{y}[k] = \mathbf{N}_{\mathrm{L}}^{1}\,\boldsymbol{u}[k-1] + \mathbf{N}_{\mathrm{L}}^{2}\,\boldsymbol{u}[k-2] + \ldots + \mathbf{N}_{\mathrm{L}}^{m}\,\boldsymbol{u}[k-m] - \\ -\mathbf{D}_{\mathrm{L}}^{1}\,\boldsymbol{y}[k-1] + \mathbf{D}_{\mathrm{L}}^{2}\,\boldsymbol{y}[k-2] + \ldots + \mathbf{D}_{\mathrm{L}}^{m}\,\boldsymbol{y}[k-m] \end{aligned} \tag{43}$$

is linear in the matrix parameters $\mathbf{N}_{\mathrm{L}}^{i}$ and $\mathbf{D}_{\mathrm{L}}^{i}$, therefore it is easy to construct a *Least-Squares* parameter estimation method to identify the *MIMO* process model [3]. The *MIMO* controller (27) can be used to form a two-step procedure where the output $\boldsymbol{c}[k]$ of the controller can be computed via an auxiliary variable $\boldsymbol{x}[k]$.

Assuming the simpler naive process model (34) and a left *MFD* reference model $\boldsymbol{R}_{\mathrm{n}}$ as

$$\boldsymbol{R}_{\mathrm{n}} = \mathcal{A}_{\mathrm{n}}^{-1}(z)\mathcal{B}_{\mathrm{n}}(z) = \left[\mathbf{I} + \tilde{\mathcal{A}}_{\mathrm{n}}\left(z^{-1}\right)\right]^{-1}\mathcal{B}_{\mathrm{n}}\left(z^{-1}\right) \tag{44}$$

the basic controller equations can be obtained in the forms of

$$\boldsymbol{c} = \bar{\boldsymbol{C}}\,\boldsymbol{e} = \left(\mathbf{I} - \boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{P}}_{-}\right)^{-1}\boldsymbol{R}_{\mathrm{n}}\bar{\boldsymbol{P}}_{+}^{-1}\boldsymbol{x} \quad ; \quad \boldsymbol{x} = \bar{\boldsymbol{P}}_{+}^{-1}\boldsymbol{e} \tag{45}$$

These equations are easy to rewrite into the vector difference equations

$$\boldsymbol{c} = \left(\mathcal{B}_{\mathrm{n}}\mathcal{N}_{\mathrm{L}}^{-} - \tilde{\mathcal{A}}_{\mathrm{n}}\right)\boldsymbol{c} + \mathcal{B}_{\mathrm{n}}\,\boldsymbol{x} \tag{46}$$

and

$$\boldsymbol{x} = \mathcal{D}_{\mathrm{L}}\,\boldsymbol{e} - \tilde{\mathcal{N}}_{\mathrm{L}}^{+}\,\boldsymbol{x} \tag{47}$$

Here the following equations were used

$$\bar{\boldsymbol{P}}_{+}^{-1} = \left[\mathcal{N}_{\mathrm{L}}^{+}(s)\right]^{-1}\mathcal{D}_{\mathrm{L}}(s) \quad ; \quad \bar{\boldsymbol{P}}_{-} = \mathcal{N}_{\mathrm{L}}^{-} \quad ; \quad \mathcal{N}_{\mathrm{L}}^{+} = \mathbf{I} + \tilde{\mathcal{N}}_{\mathrm{L}}^{+} \tag{48}$$

6. Examples

Example 1. Let us consider a very simple example first, when the plant transfer function matrix is

$$\boldsymbol{P}(s)=\frac{1}{\mathcal{A}(s)}\mathcal{B}(s)=\begin{bmatrix}\frac{1}{1+s} & \frac{1}{1+2s}\\ 0 & \frac{1}{1+4s}\end{bmatrix}=\frac{1}{(1+s)(1+2s)(1+4s)}\begin{bmatrix}(1+2s)(1+4s) & (1+s)(1+4s)\\ 0 & (1+s)(1+2s)\end{bmatrix} \tag{49}$$

Select a speeding up and decoupling reference model

$$\boldsymbol{R}_{\mathrm{n}}(s)=\frac{1}{\mathcal{A}_{\mathrm{n}}(s)}\mathcal{B}_{\mathrm{n}}(s)=\frac{1}{(1+0.5s)}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\frac{1}{(1+0.5s)}\mathbf{I} \tag{50}$$

The detailed computation of the transfer function matrix of the *MIMO* regulator using equation (37) gives

$$\boldsymbol{C}(s)=\begin{bmatrix}\frac{1+s}{0.5s} & -\frac{(1+s)(1+4s)}{0.5s(1+2s)}\\ 0 & \frac{1+4s}{0.5s}\end{bmatrix} \tag{51}$$

So the obtained regulator contains *PI* like regulators in two and a *PID* like in one element.

Example 2. Investigate a discrete-time process now. Let the matrix pulse transfer function of the plant be

$$\boldsymbol{G}(z)=\begin{bmatrix}\frac{0.5z^{-1}}{1-0.5z^{-1}} & \frac{0.2z^{-1}}{1-0.8z^{-1}}\\ 0 & \frac{z^{-1}-0.5z^{-2}}{1-1.7z^{-1}+0.2z^{-1}}\end{bmatrix} \tag{52}$$

and apply again a speeding up and decoupling reference model

$$\boldsymbol{R}_{\mathrm{n}}(z)=\begin{bmatrix}\frac{0.8z^{-1}}{1-0.2z^{-1}} & 0\\ 0 & \left(\frac{0.9z^{-1}}{1-0.1z^{-1}}\right)^2\end{bmatrix}=\frac{\begin{bmatrix}0.8z^{-1}\left(1-0.1z^{-1}\right)^2 & 0\\ 0 & \left(0.9z^{-1}\right)^2\left(1-0.2z^{-1}\right)\end{bmatrix}}{\left(1-0.2z^{-1}\right)\left(1-0.1z^{-1}\right)^2} \tag{53}$$

The detailed computation of the pulse transfer function matrix of the *MIMO* regulator using equation (37) gives

$$\boldsymbol{C}(z)=\begin{bmatrix}\boldsymbol{C}_{11}(z) & \boldsymbol{C}_{12}(z)\\ \boldsymbol{C}_{21}(z) & \boldsymbol{C}_{22}(z)\end{bmatrix} \tag{54}$$

where

$$\boldsymbol{C}_{11}(z)=\frac{1.6\left(1-0.5z^{-1}\right)}{1-z^{-1}} \quad ; \quad \boldsymbol{C}_{12}(z)=\frac{-0.32\left(1-1.7z^{-1}+0.2z^{-2}\right)}{\left(1-z^{-1}\right)\left(1-0.8z^{-1}\right)} \tag{55}$$

$$\boldsymbol{C}_{21}(z)=0 \quad , \quad \boldsymbol{C}_{22}(z)=\frac{0.81z^{-1}\left(1-1.7z^{-1}+0.2z^{-2}\right)}{\left(1-z^{-1}\right)\left(1+0.8z^{-1}\right)\left(1-0.5z^{-1}\right)} \tag{56}$$

So all elements of the regulator are realizable, because of the special selection of the matrix reference model elements. These sub-regulators are integrating, because the matrix reference model $\boldsymbol{R}_{\mathrm{n}}$ has unity gain nontrivial elements. (These examples are very simple low order ones, because the typesetting for higher order plants is very difficult, the regulator elements are much higher order then.)

7. Conclusions

A formal generalization of the *Youla-parametrization* for *MIMO* systems is presented. The matrix transfer function form of the *Youla-parameter* was introduced and the design of a *MIMO* regulator was also developed for the *KB-parametrization* based generic *2DOF* scheme. The optimal regulator is given for continuous and discrete-time case, too. The controller equations are further simplified for a commutative naive *MIMO* plant model.

A continuous-time and a discrete-time simple example is shown to ease the understanding of the methodology.

References

[1] Keviczky, L.: *Combined identification and control: another way*. (Invited plenary paper.) 5th IFAC Symp. on Adaptive Control and Signal Processing, ACASP'95, Budapest, H, (1995), pp. 13-30.

[2] Keviczky, L., Bányász, CS.: *Optimality of two-degree of freedom controllers in* H_2 *- and* $\boldsymbol{H}_\infty$ *-norm space, their robustness and minimal sensitivity*. 14th IFAC World Congress, **F**, Beijing, PRC, (1999), pp. 331-336.

[3] Keviczky, L., Bányász, CS.: *Closed-loop parametrization schemes: Identification of the K-B and d-Y parameters*. 3rd IFAC Symposium on Robust Control Design, ROCOND'2000, CD printing, Prague, CZ, (2000).

[4] Keviczky, L., Bányász, CS.: *On the* H_2, $\boldsymbol{L}_2$ *and* $\boldsymbol{H}_\infty$, $\boldsymbol{L}_\infty$ *optimality of some two-degree of freedom control systems*. Journal of Systems Science, 33(3), (2008), pp. 39-49.

[5] Maciejowski, J. M.: *Multivariable Feedback Design*, Addison Wesley, (1989).

This work was supported in part by the Control Engineering Research Group of the HAS, at the Budapest University of Technology and Economics and by the project TAMOP 421B, at the Széchenyi University of Győr.

Finite Element Analysis of a Solid-Rotor Induction Machine

D. Marcsa

Laboratory of Electromagnetic Fields, Department of Telecommunications
Széchenyi István University, Egyetem tér 1, H-9026, Győr, Hungary
marcsadaniel@yahoo.co.uk

Abstract: The paper deals with the analysis of the single-phase induction machine of the Problem No. 30a of the COMPUMAG TEAM Workshop. The problem has been solved by the motional two-dimensional time-harmonic Finite Element Method (FEM) using different potential formulations, the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation and the $\boldsymbol{T},\Phi-\Phi$-potential formulation. Here the problem is a linear eddy current field problem.

Keywords: Finite element method, Induction machine, Eddy current field potential formulations, Motion voltage term.

1. Introduction

The paper presents the solution of the single-phase induction machine of the Problem No. 30a of COMPUMAG TEAM Workshop, [1] by using different potential formulations and the Finite Element Method (FEM) [2], [3], [4]. The linear problem has been solved in the frequency domain by the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation and by the $\boldsymbol{T},\Phi-\Phi$-potential formulation and the solutions of them have been compared. The investigated problem is considered to be a linear eddy current field problem.

The used potential formulations of the motional two dimensional time-harmonic Finite Element Method [4], [5], [6], [7] by means of the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation with motion [4], [5] and by the $\boldsymbol{T},\Phi-\Phi$-potential formulation with motion [8], [9] have been implemented. The two-dimensional linear time varying eddy current field problem has been solved in the frequency domain [10] by the above two potential formulations.

The introduced methods are applied to compute the electromagnetic torque [4], [5], the power dissipation of the rotor, and the rotor steel [11], [12] in the single-phase solid-rotor machine. These quantities are computed on a per unit depths (1m) basis.

The induction machine under investigation is considered to be an eddy current field problem. The eddy current field of the machine is assumed to be two-dimensional, i.e. the influence of the end-region fields have has been neglected.

2. Model of induction machine

The arrangement of the problem is shown in *Figure 1*, which is a single-phase induction machine with solid-rotor body, where σ is the conductivity, μ_r is the relative permeability. In this figure, the current excitation of the windings is represented by the source current density $\boldsymbol{J}_0$.

In the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation the inverse model of constitutive relations is used. In the $\boldsymbol{T},\Phi-\Phi$-potential formulation the direct model of constitutive relations is applied.

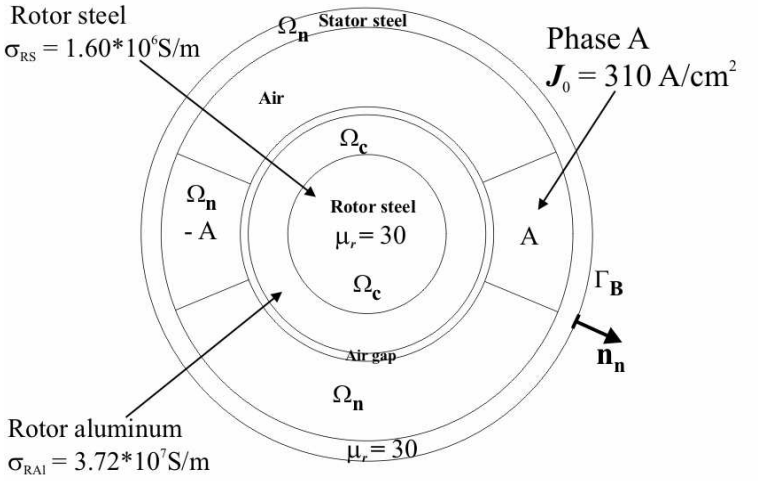


Figure 2. The arrangement of the analyzed induction machine

The rotor is made of rotor steel and rotor aluminum, i.e. the rotor steel is surrounded by the rotor aluminum. The stator steel is laminated, and its conductivity has been selected as $\sigma = 0$, i.e. eddy current is neglected there. The windings are not embedded in slots. The rotation of the rotor is counterclockwise. The range of computation of the single-phase induction machine for rotor angular velocities is from 0 to 358 rad/s (0.95%) of the peak field speed) The synchronous speed is $\omega_S = 2\pi \cdot f = 377\,\text{rad/s}$, because the winding is excited by a current with the frequency $f = 60\,\text{Hz}$.

The studied eddy current field problem is separated into two parts. The first is the conductor region, i.e. the rotor, denoted by Ω_C, and the second is the nonmagnetic and non-conductive region, the windings, the air, and the laminated stator steel, symbolized by Ω_n. Maxwell's equations in the eddy current free region are modeling a static magnetic field, while in the eddy current region the magneto-dynamic Maxwell's equations are valid. The boundary Γ_B is the closing boundary of the problem region, where the normal component of magnetic flux density is neglected.

The arrangement of the problem has been discretized by a second order triangular finite element mesh.

3. Methods of analysis

Assuming that the analyzed model has a moving part, it is necessary to take movement into account. This can be performed by using the motion voltage term $\boldsymbol{v} \times \boldsymbol{B}$ [4], [5], [6], [7], [8], [9], where $\boldsymbol{v}$ is the velocity and $\boldsymbol{B}$ is the magnetic flux density. In this case, the spatial discretization of the field quantities results in unsymmetrical global matrix [5].

3.1. Modeling motion

The rotating induction machines are separated into two reference frames, a fixed reference frame $O(x,y)$ and a moving reference frame $O'(x',y')$, which is moving with the rotor [4], [7].

The displacement currents are negligible, because of the working low frequencies in the electrical machines, so, Maxwell's equations can be written in the fixed reference frame as

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} \,, \tag{1}$$

$$\nabla \times \boldsymbol{E} = -\mathrm{j}\omega \boldsymbol{B} \,, \tag{2}$$

$$\nabla \cdot \boldsymbol{B} = 0 \,, \tag{3}$$

$$\boldsymbol{J} = \begin{cases} \boldsymbol{J}_0, & \text{in } \Omega_n, \\ \sigma \boldsymbol{E}, & \text{in } \Omega_c, \end{cases} \tag{4}$$

$$\boldsymbol{B} = \mu \boldsymbol{H} \,. \tag{5}$$

Electromagnetic phenomena are described by the same Maxwell's equations in the fixed, and in the moving reference frames. These equations in the moving reference frame are the following:

$$\boldsymbol{B}' = \boldsymbol{B} \,, \tag{6}$$

$$\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} \,, \tag{7}$$

$$\boldsymbol{H}' = \boldsymbol{H} \,, \tag{8}$$

$$\boldsymbol{J}' = \boldsymbol{J} \,, \tag{9}$$

where the quantities observed from the moving reference frame O' are marked by apostrophes.

Only the electric field intensity vector $\boldsymbol{E}'$ is modified, because of the additional motion voltage term $\boldsymbol{v} \times \boldsymbol{B}$ to $\boldsymbol{E}$.

Substituting the equation of $\boldsymbol{E}'$ into (4) gives the next relation:

$$\boldsymbol{J} = \sigma \boldsymbol{E}' = \sigma\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right), \tag{10}$$

because eddy currents have been induced in the motional part of the conductor region of the induction machine, and these eddy currents are depending on the velocity.

3.2. The $A,V-A$ - potential formulation with motion

This potential formulation coupled with moving velocity seems to be the most widely used formulation in the simulation of electrical machines [4], [5], [6], [7].

The $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation is based on the magnetic vector potential $\boldsymbol{A}$ and the electric scalar potential V. Both of them are approximated by nodal finite elements [2], [3], [13].

Basically, the induction machine is considered to be a two-dimensional problem, and this results in $V=0$ for electric scalar potential in the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation [2], [3].

The divergence-free magnetic flux density $\boldsymbol{B}$ can be described by the curl of the magnetic vector potential $\boldsymbol{A}$, i.e.

$$\boldsymbol{B}=\nabla\times\boldsymbol{A}. \tag{11}$$

In this two-dimensional problem the Coulomb gauge, $\nabla\cdot\boldsymbol{A}=0$, is satisfied automatically.

Substituting (11) into Faraday's law (2), it results in the equation

$$\nabla\times\boldsymbol{E}=-j\omega\nabla\times\boldsymbol{A}. \tag{12}$$

Substituting (10) into (1), and the use of constitutive relation (5), relation (12) and (11) result in the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation with motion in the conducting region Ω_c in the frequency domain,

$$\nabla\times(\nu\nabla\times\boldsymbol{A})+\sigma(j\omega\boldsymbol{A}-\boldsymbol{v}\times\nabla\times\boldsymbol{A})=\boldsymbol{0}, \tag{13}$$

where the reluctivity is $\nu=1/(\mu_0\mu_r)$.

3.3. The $\boldsymbol{T},\boldsymbol{\Phi}-\boldsymbol{\Phi}$ – potential formulation with motion

Basically, this potential formulation has not been used in the simulation of induction machines, however some papers in the literature can be found for FEM analysis [8], [9], [14].

The $\boldsymbol{T},\Phi-\Phi$-potential formulation is based on the current vector potential $\boldsymbol{T}$ with edge element approximation and on the reduced magnetic scalar potential Φ with nodal element approximation [2], [3], [13], [14].

The source current density $\boldsymbol{J}_0$ can be written by the curl of the impressed current vector potential $\boldsymbol{T}_0$ [2], [3], [13], [14],

$$\boldsymbol{J}_0=\nabla\times\boldsymbol{T}_0. \tag{14}$$

The expression of the magnetic field intensity $\boldsymbol{H}$ and the expression of the eddy current density $\boldsymbol{J}$ are as follows from (1):

$$\boldsymbol{H}=\boldsymbol{T}_0+\boldsymbol{T}-\nabla\Phi, \tag{15}$$

$$\boldsymbol{J} = \nabla \times \boldsymbol{T}\,, \tag{16}$$

according to the fact that the eddy current density is divergence-free, i.e. $\nabla \cdot \boldsymbol{J} = 0$.

The basic equation of this formulation with motion voltage term can be formulated in the following way [8], [9]. In the moving reference frame (2) is the following

$$\nabla \times \boldsymbol{E}' = -\mathrm{j}\omega \boldsymbol{B}'\,, \tag{17}$$

and in this equation the magnetic flux density $\boldsymbol{B}'$ is unchanged (see (6)), only the electric field intensity $\boldsymbol{E}'$ is modified (see (7)).

Substitute the (7) into (17) gives the next relation:

$$\nabla \times \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) = -\mathrm{j}\omega \boldsymbol{B}. \tag{18}$$

Combining the above equation (18), the constitutive relations (4) and (5), (15) and (16) result in the $\boldsymbol{T}, \Phi - \Phi$ **-**potential formulation with motion in the conducting region Ω_c in the frequency domain,

$$\nabla \times \left(\rho \nabla \times \boldsymbol{T} + \boldsymbol{v} \times \mu\left(\boldsymbol{T}_0 + \boldsymbol{T} - \nabla\Phi\right)\right) + j\omega\mu\boldsymbol{T} - j\omega\mu\nabla\Phi = -j\omega\mu\boldsymbol{T}_0\,, \tag{19}$$

where $\mu = \mu_0\mu_r$ is the permeability, and $\rho = 1/\sigma$ is the resistivity.

3.4. Computed quentities

3.4.1. Torque

The method of Maxwell's stress tensor is commonly used in the calculation of forces and torque in the numerical analysis of electrical devices [4], [5].

The electromagnetic torque is obtained as a surface integral, but in two-dimensional case the surface integral is reduced to a line integral along the air gap [4], [5].

The electromagnetic torque is obtained as a line integral [4], [5],

$$\boldsymbol{T}_e = L\int_{\Gamma}\left(\boldsymbol{r} \times \mathrm{d}\boldsymbol{F}\right)d\Gamma = L\int_{\Gamma}\left(\boldsymbol{r} \times \left[\frac{1}{\mu_0}\left(\boldsymbol{B}\cdot\boldsymbol{n}\right)\boldsymbol{B} - \frac{1}{2\mu_0}B^2\boldsymbol{n}\right]\right)d\Gamma\,, \tag{20}$$

where L is a unit depth of the domain ($L = 1\mathrm{m}$ in the *z*-direction), and $\boldsymbol{r}$ is the position vector linking the rotation axis to the element $d\Gamma$, and Γ is a surface, which places around the air gap, $\mathrm{d}\boldsymbol{F}$ is the so-called Maxwell's stress tensor, $\boldsymbol{B}$ is the magnetic flux density, and $B = |\boldsymbol{B}|$ is the absolute value of magnetic flux density.

3.4.2. Losses

The eddy current losses are determined by using a method based on the integration of the power density loss over the rotor surface. The equation of eddy current losses is obtained from [11], [12],

$$P_d = L\int_{\Omega} \mathrm{Re}\left\{\frac{1}{2}\sigma \boldsymbol{E}\cdot\boldsymbol{E}^*\right\} d\Omega\,, \tag{21}$$

where the real part of power density loss $\sigma \boldsymbol{E}\cdot\boldsymbol{E}^*/2$ can be integrated over the surface of the body $\Omega = \Omega_c \cup \Omega_n$, L is a unit depth of the domain, and $\boldsymbol{E}$ is the electric field intensity, while $\boldsymbol{E}^*$ is the conjugate of the electric field intensity.

The difference between the above two potential formulations is the handle of the moving velocity. In the $\boldsymbol{T},\Phi-\Phi$-formulation, after the simulation the electric field intensity $\boldsymbol{E}=\boldsymbol{E}'$, whereas the $\boldsymbol{A},V-\boldsymbol{A}$-formulation $\boldsymbol{E}=\boldsymbol{E}'-\boldsymbol{v}\times\boldsymbol{B}$.

4. Results

The problem No. 30a of the COPMUMAG TEAM Workshop is a benchmark problem, and the efficiency of different methods; formulations have been compared by this problem.

This problem has been solved by analytical [14], and by time-harmonic finite element method [6], [15].

The finite element simulation is based on the solution of the weak form of the presented potential formulations [2], [3]. The weak form of a partial differential equation can be obtained by using the weighted residual method and the Galerkin's method [2], [3].

The above introduced $\boldsymbol{A},V-\boldsymbol{A}$ and $\boldsymbol{T},\Phi-\Phi$ potential formulations are applied with motion voltage term in the frequency domain.

Furthermore, in this section, the results of the commercial finite element code, the COMSOL Multiphysics [16] are presented as well. These results also have been computed in the frequency domain.

The results of the formulations have been compared with the analytical solution, resulted in [1], [14], and to each other's.

The numerical computation has been performed by applying the procedure developed by using the functions of COMSOL Multiphysics. The procedures have been run on an Intel(R) Core(TM)2 Duo CPU T9300 with the clock frequency of the processor 2.50 GHz, with 2GByte RAM.

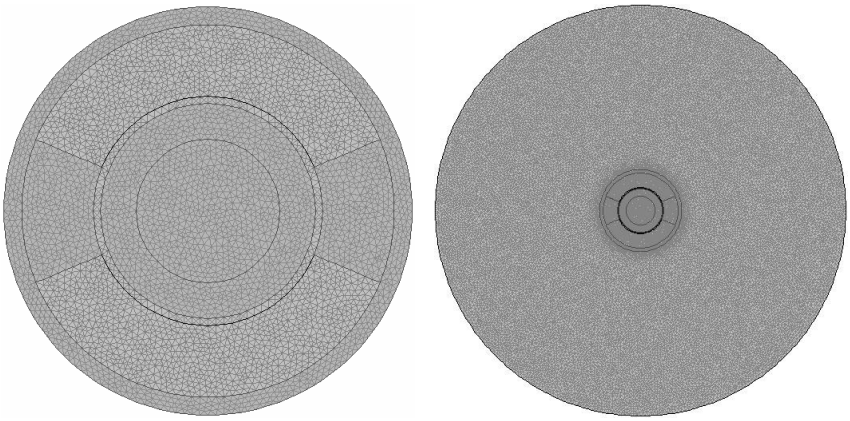
The arrangement of the single-phase machine (*Figure 2a*) has been discretized by a triangular mesh, as it is shown in *Figure 2*.

In the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation, the used mesh consists of 43522 second-order triangular elements, and the number of unknowns is 87405.

In the $\boldsymbol{T},\Phi-\Phi$-potential formulation, the used mesh of machine with closing boundary consists of 223426 second-order triangular elements (*Figure 2b*), and the total number of unknowns is 404286. The computation time is 388s because the impressed current vector potential has been calculated by iterative solver, and the number of

iterations is 1157. The used mesh consists of 43522, and the number of unknowns is 147107 when computing $\boldsymbol{T}$ and Φ,.

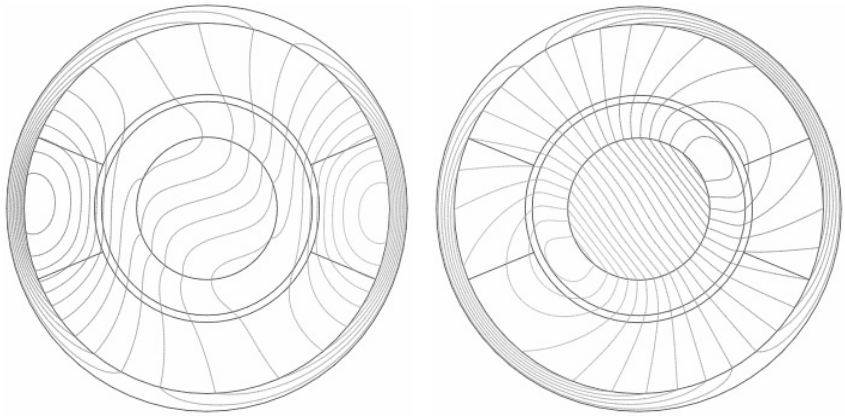
The $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation is faster, because the computation time of this formulation is ranging from 5.48s to 6.61s, whereas the CPU time of the $\boldsymbol{T},\Phi-\Phi$-potential formulation is ranging from 13.62s to 23.05s. Furthermore, the impressed current vector potential $\boldsymbol{T}_0$ has been computed when the edge-element based $\boldsymbol{T},\Phi-\Phi$-potential formulation is applied, and this adds about 388s in every simulation.



a) The single-phase motor *b*) The motor with closing boundary

Figure 2. The single-phase induction motors are meshed by triangles

Figure 3. shows the computed equipotential lines of the magnetic vector potential at 278 rad/s angular velocity. *Figure 3a.* shows the real part of the magnetic vector potential, and *Figure 3b.* presents the imaginary part of the magnetic vector potential. These figures are almost the same as the presented figures in the paper [6].



a) The real part of $\boldsymbol{A}$ *b*) The imaginary part of $\boldsymbol{A}$

Figure 3. The magnetic vector potential lines of the analyzed induction machine

Figure 4. presents the curves of the analytical and the computed electromagnetic torque by the presented potential formulations. In this figure the three different finite element approaches are practically the same. The difference between the analytical and the finite element results is increased with the increase of velocity, and the difference is the largest at the maximum of the torque or the point of breakdown torque. After the point of breakdown torque, the torque is decreasing. Furthermore the curves show the typical torque-speed characteristics of single-phase induction machines, and the point of full load torque are nearly 350 rad/s.

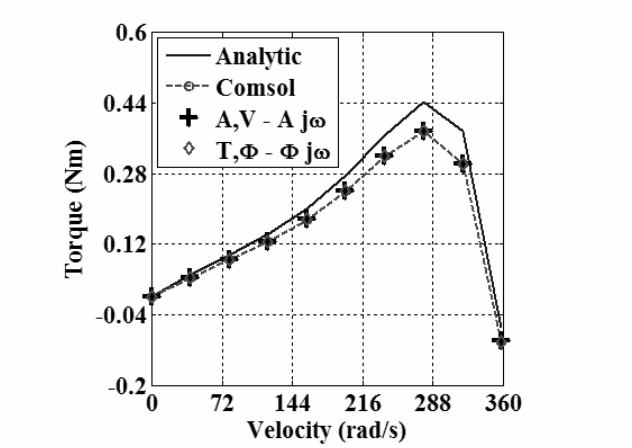


Figure 4. Torque-speed characteristics

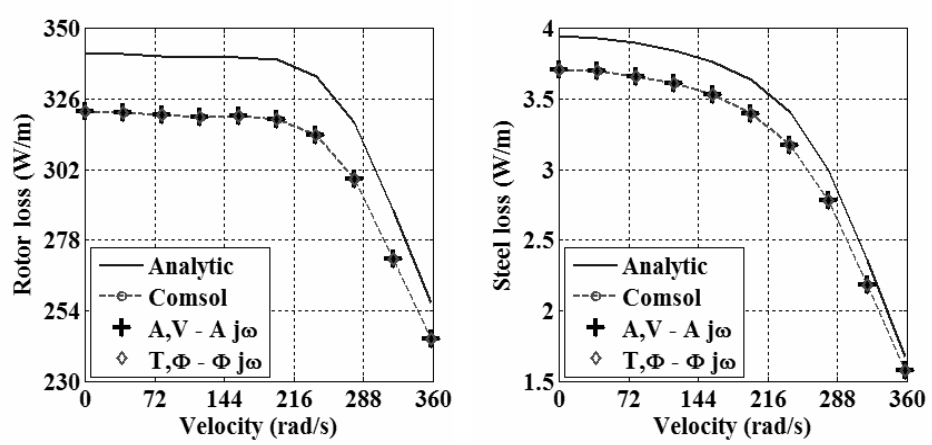
The next figure (*Figure 5a.*) presents the curves of the total rotor loss in both the rotor aluminum and the rotor steel. The solution of both potential formulations and the commercial FEM code seem to be equal at all angular velocities. The character of the finite element results is similar to the analytical solution. The difference is the highest when the machine is standstill, i.e. at 0 rad/s angular velocity.

Figure 5b. presents just the rotor steel loss, which due to the I^2R dissipation. The curves of the steel loss are similar to the curves of the rotor loss, however the steel loss is much smaller. Furthermore, the difference between analytic and finite element results is the largest at the standstill motor. The difference between the curve of analytical results and of the FEM results is decreasing when the speed becomes higher.

5. Conclusion

The analysis of induction machines is based on the numerical solution of the different potential formulations. The $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation coupled with moving velocity seems to be the most widely used formulation of electrical machines. This formulation has been worked out with motional voltage term in numberless articles and papers, and quantities of machines, e.g. torque, losses have been realized by this formulation. The $\boldsymbol{T},\Phi-\Phi$-potential formulation coupled with moving velocity has never been used for solving induction machine problem. However, it has been applied for simulation non-destructive eddy current testing with motion [17].

The introduced numerical procedure of eddy current field calculation results in the expected data. The two different potential formulations have been compared and the results seem to be close to each other. The numerical results have been compared to the



a) The rotor loss of single phase motor *b*) The steel loss of single-phase motor

Figure 5. The power dissipation of rotor and rotor steel

analytical results and the used potential formulation is allowing a fine approximation. The computed and analytical results agree within 10%.

In the two-dimensional case, the $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation seems to be the better solver with motion voltage term. The $\boldsymbol{A},V-\boldsymbol{A}$-potential formulation with motion voltage term is faster, the number of elements and unknowns are less than in the $\boldsymbol{T},\Phi-\Phi$-potential formulation with motion voltage term.

Acknowledgements

This paper was supported by the Hungarian Scientific Research Fund (OTKA PD 73242), and by the Széchenyi István University (15-3210-02).

References

[1] *Induction Motor Analysis*, - International TEAM Benchmark Problem No. 30a, Testing Electromagnetic Analysis Methods (T.E.A.M.), http://www.compumag.co.uk/

[2] Kuczmann, M., Iványi, A.: *The Finite Element Method in Magnetics*, Akadémiai Kiadó, (2008).

[3] Bíró, O., Richter, K. R.: *CAD in Electromagnetism*, In Series Advances in Electronics and Electron Physics, Vol. 82, Academic Press, (1991), pp. 1-92.

[4] Arkkio, A.: *Analysis of Induction Motors Based on the Numerical Solution of the Magnetic Field and Circuit Equations*, Acta Polytechnica Scandinavica, Series Electrical Engineering, No. 59, Helsinki University of Technology, (1987).

[5] Bastos, J. P. A., Sadowski N.: *Electromagnetic Modeling by Finite Element Methods*, Marcel Dekker Inc., (2003).
[6] Kuczmann, M.: *Measurement and Simulation of Vector Hysteresis Characteristics*, IEEE Transaction on Magnetics, Vol. 45, No.11, (2009), pp. 5188−5191.
[7] Burais, N., Foggai, A., Nicolas, A., Pascal, J. P. Sabonnadiere, J. C.: *Numerical Solution of Eddy Currents Problems Including Moving Conducting Parts*, IEEE Transaction on Magnetics, Vol. MAG-20, No. 5, (1984), pp. 1995−1997.
[8] Slama, A., Mazauric, V., Maréchal, Y., Meunier, G.: *Electric Railgun 3D Modeling, Computation of Eddy Currents and Lorenz Force*, IEEE Transaction on Magnetics, Vol. 37, No. 1, (2001), pp. 139−142.
[9] Chadebec, O., Meunier, G., Mazauric, V. G., Le Floch, Y., Labie, P.: *Eddy-Current Effects in Circuit Breakers During Arc Displacemet Phase*, IEEE Transaction on Magnetics, Vol. 40, No. 2, (2004), pp. 1358−1361.
[10] Biro, O.: *Edge Element Formulations of Eddy Current Problems*, Computer methods in Applied Mechanics and Engineering, Vol. 169, (1999), pp. 391−405.
[11] Davey, K. R.: *Rotating Field Analysis Using Boundary Element Methods*, IEEE Transaction on Magnetics,, Vol. 35, No. 3, (1999), pp. 1402−1405.
[12] Iványi, A.: *Magnetic Field Computation with R-Functions*, Akadémiai Kiadó, (1998).
[13] Kuczmann, M.: *Nodal and Vector Finite Elements in Static and Eddy Current Field Problems*, Pollack Periodica, Vol. 3, No. 2, (2008), pp. 85-96.
[14] Kuczmann, M.: *Design of a 2D RRSST System by FEM and the **T,Φ-Φ** Potential Formulation*, Pollack Periodica, Vol. 3, No. 1, (2008), pp. 67–80.
[15] Davey, K. R.: *Analytic Analysis of Single- and Three-Phase Induction Motors*, IEEE Transaction on Magnetics,, Vol. 34, No. 5, (1998), pp. 3721−3727.
[16] Marcsa, D., Kuczmann, M.: *Numerical Analysis of Single- and Three-Phase Induction Motors*, in Proceedings of the 2nd Symposium on Applied Electromagnetics – SAEM'08, Zamość, Poland, (2008), pp. 121-124.
[17] Zhiye, D., Jiangjun, R., Ying, P., Shifeng, Y., Yu, Z., Yan, G., Tianwei, L.: *3-D FEM Simulation of Velocity Effect on Magnetic Flux Leakage Testing Signals*, IEEE Transaction on Magnetics, Vol. 44, No. 6, (2008), pp. 1642-1645.

Entity Flow-phase Analysis for Fast Performance Estimation of Organisational Process Systems

G. Lencse and L. Muka

SZE, 9026 Győr, Egyetem tér 1.
Phone: + 36 96 613-665, fax: + 36 96 613-646
e-mail: lencse@sze.hu, muka.laszlo@elassys.hu

Abstract: This paper describes Entity Flow-phase Analysis (EFA) which is a method for fast performance estimation of organisational process systems. Entity Flow-phase Analysis extends the approach of the promising Traffic-Flow Analysis (TFA) method – created for the rapid analysis of communications networks – for BP (Business Process) systems. EFA operates on the possibilities determined by the *similarities* between communications networks and BP systems and introduces new solutions for the problems that occur because of the differences between ICT and BP systems. EFA, similarly to TFA, uses the combined approach of simulation and numerical methods. In simulation projects initiated to support the design of Information and Communication Technology (ICT) system and Business Process (BP) system in an organisation the parallel analysis of different systems may be efficient. EFA is a promising evaluation method to be applied for systems with determined BP and ICT subsystems in an organisational environment.

Keywords: Entity Flow-phase Analysis, Traffic-Flow Analysis, Information and Communication Technology, Business Process, process modelling, parallel simulation, discrete event simulation

1. Introduction

1.1. Mixed simulation projects

Simulation projects aimed to support the design of Information and Communication Technology (ICT) and Business Process (BP) systems in an organisation traditionally are independent, separate projects, in spite of the fact that these systems may have significant influence on each other. Common analysis of these systems may have advantages but in this case we need to have methods appropriate for both types of systems.

The Entity Flow-phase Analysis (EFA, first defined in [9]) serves for the fast performance evaluation of BP systems (more generally: organisational process systems) has been derived from Traffic-Flow Analysis (TFA, [7]). This derivation of EFA is based on the formal similarity of DES models of ICT and BP systems. EFA uses the

same two phase method as TFA, only the interpretation of the model elements is different.

The *expected advantages* of the new modelling method are:

- EFA – which is a fast (and approximate) simulation like method – is *expected to require significantly less* time for execution then the event-by-event simulation of the BP. This feature, allowing generating a large number of experiments, may help to use the EFA method for increasing the efficiency and effectiveness of the simulation process**,** by understanding the *complex* features of BP systems and by defining the features and *limitations* of *Discrete-Event Simulation models* of BPs.

- EFA may allow of building *online simulation models of BPs* by using data of some enterprise ICT systems.

- EFA together with TFA is a set of methods which may form a homogenous fast modelling and simulation method-environment for *common analysis* of interconnected *ICT and BP systems*.

1.2. Business processes and business process systems

There are some known, basic definitions of business processes:

By the definition given in [4], *processes* are structured, measured sets of activities designed to produce a specified output for a particular customer or market. According to another definition, a *business process* is a partially ordered set of Enterprise Activities which can be executed to realise a given objective of an enterprise or a part of an enterprise to achieve some desired end-result [11]. In enterprise (organisational) modelling, a business process is defined as a network of actions performed in context of one or more organisational roles in order to achieve some goal [6].

According to the descriptions above, for modelling purposes, the business processes may be defined as follows:

Business processes are related to enterprises and they define the way in which the goals of the enterprise are achieved. A *Business Process (BP)* is a set of Enterprise Activities linked together to form a process with one or more kinds of input to produce outputs. A *process system* (BP system) is a set of business processes linked together to perform some Enterprise Function or Subfunction [9].

Business Processes may be characterised by the character of their *structure* (the *ad-hoc* and *permanent* processes) and by the character of their *frequency of execution* (the *cyclic* type processes are executed repeatedly, the *non-cyclic* type processes are executed once). The functional BPs in the decreasing order of their frequency of execution are the following ones: the processes of operations, finance, marketing and R&D [10].

The object of examination is the BP with permanent (or changing in a probability manner) structure and with cyclic execution. (For the analysis of the BPs with ad-hoc structure and non-cyclic execution may be examined using for example the PERT (Program Evaluation and Review Technique) method.

The BP systems usually are *complex systems* in their nature. In order to exceed naive, poor modelling in *current* (*traditional*) *simulation* of BPs, the *advanced simulation* approach of BPs should be based on the thorough analysis of *secondary (historical)* data (data, logs, models, etc.) and *primary* data on *structures*, *resources, loads* and *timing* relations of BPs too [1]. The *static model* of BPs – containing information on the structure of BPs with the definition of activities – may be described and found in many forms: in forms of Flow Charts, IDEF0 (Integrated Definition for Function Modelling) models, UML (Unified Modelling Language) diagrams, or in form of a model of a modelling and simulation tool (for example ARIS) [11]. Information on *resource*s – both on human and technical resources – allocated to BPs may be found in some systems of MIS (Management Information System) [3]: for example in the ERP (Enterprise Resource Planning) system and in the ABC (Activity Based Costing) system. Data on *load* and *timing* relations for BPs may be got the from CRM (Customer Relationship Management) system, from the WFMS (Workflow Management System) or from the ABC system.

In order to collect the sufficient information for modelling and simulation of BPs, usually primary information should also be collected (it may require, for example, measuring in some system, making interviews, etc.).

1.3. The Traffic-Flow Analysis

The *Traffic-Flow Analysis* (TFA) [7] (about the convergence of TFA: [8]) is a simulation-like method for the fast performance analysis of communication systems. TFA uses statistics to model the networking load of applications.

In the *first part* the method distributes the traffic (the statistics) in the network, using routing rules and routing units.

In the *second part* the influence of the finite capacities (line and switching-node capacities) is calculated.

The important features of TFA:

- The results are approximate but the absence and the place of *bottlenecks* is shown by the method.
- The *execution time* of TFA is expected to be significantly less than execution time of detailed simulation of the system.
- TFA describes the *steady state behaviour* of the network (there is no need for warm-up time definition).

2. DES-P Model Elements in EFA Models

The Discrete-Event Simulation (DES) [2] model of a BP (we call it DES-P model) is for the event-by-event simulation of the BP. The EFA model – which is intended to serve for fast, approximate modelling and simulation of a BP – is a set of functions within the DES program and uses the DES model and the DES engine. Therefore, at this point, the features of DES-P models (important for EFA modelling) are overviewed.

2.1. Activities, links and entitiy generation

The Business Process (BP) model is a set of *activities* and *links* connecting the activities together.

The *entity* represents an *elementary task* (*case*) that is executed by an activity. An entity could be (or could not be) executed by an activity depending on the *type* of the entity.

The *links* are *logical connections* between activities with a *direction* showing the performance order, the direction of the travelling of entities. The links have no capacity limit because of their logical character.

The *entity-load* of the BP is produced by *entity-generators*. The different types of entities are produced by different entity-generators.

The entities enter the BP model at its *entry* points and leave the BP model at its *exit* (*result*) points. The destinations of entities are the *exit* points of the BP model.

Between the entry and exit points of the BP model, *exit-entry* point pairs can be inserted into the BP model for an entity type.

Two or more BP can be connected by links to form a *BP system*. *Internal links* are connecting the activities inside a BP; *external links* are connecting activities of a BP system.

Remark: The entity generation in DES-P may be implemented in a *General Entity Generation Model*. This model should describe the structural and time features of entity generation by a set of *profiles* (for all the *typical* end *extreme* situations) for different entity types. The entity generation may have *offline* and *online* (having input from some ICT system) components.

2.2. Entity routing decisions

Similarly to communication systems, there is a routing of entities in the BP model, according to activities and the links connecting the activities. Entity routing depends on business process decisions. A routing decision may be made using different algorithms:

- *Percentage distribution*: the destination of an entity is decided according to the probability distribution of the possible entity outputs.

- *Entity-feature distribution*: output for an entity is chosen by some *feature of an entity* (for example type, priority, etc.).

- *Load-balancing distribution*: output for an entity is chosen on the basis of some load-balancing algorithm (including some quantity-limit consideration too).

Remark: The modelling of routing in BP systems may be different from those in ICT systems, because routing rules frequently can be defined less exactly, with more uncertainty (for example links between activities may show an uncertain character).

2.3. Special elements for entity routing and generation

There may be some other *special elements* in the model influencing the generation and routing of entities:

- *Fork* element makes copies of an entity (this is a parent-child relationship) and routes the copies to outputs in a parallel way. The pair of the Fork is the *Join* element that collects the entities divided by Fork element into one entity. The delay of an entity collected by Join equals the maximum delay of entities routed by Fork.

- *Split* element also makes copies of an entity and routes the copies according to the output links of the element but the split entities will not be collected into one entity again. Split generates entities, which will have separate ways in the process.

- *Transform* element may change the features of an entity in the process.

Remark: The influence of the Fork-Join, Split and Transform elements may be taken into account by inserting an *exit-entry point pair* for the entity.

3. EFA model elements

The EFA method uses the same *two-stage* organisation of work like TFA:

- In the *first stage, the entity-load* (the statistics) is distributed (using routing rules) in the business process.

- In the *second stage,* the influence of the *capacities of activities* is calculated.

3.1. Modelling the entity-load in EFA

In EFA, the *entity-load* is modelled by *throughput* and *delay* distributions: the *entity-load* is described by the throughput distributions of the *original entity-throughput* ($p_T^{\{o\}}[l]$), the *steady state* entity-throughput ($p_T^{\{s\}}[l]$) and of the *resulting* entity-throughput ($p_T^{\{c\}}[l]$) and by the delay distribution of the *entity-delay* ($p_T^{\{d\}}[i]$).

The *original* entity-throughput distribution (1) denotes the probability that l entities *arrive* to the activity in T time where L is the random variable of the *load* which assumes values denoted by l ($l = 0,1,2, \ldots$).

$$p_T^{\{o\}}[l] = P_T^{\{o\}}(L = l),\ \sum_{l=0,1,2,\ldots} p_T^{\{o\}}[l] = 1 \tag{1}$$

The entity-throughput distribution of *steady state* (2) denotes the probability that l entities *require execution* (*service*) by the activity in T time after the distribution of the entity-load. The entity-throughput distribution of *steady state* shows the result of summation of the distributed entity-load.

$$p_T^{\{s\}}[l] = P_T^{\{s\}}(L = l),\ \sum_{l=0,1,2,\ldots} p_T^{\{s\}}[l] = 1 \tag{2}$$

The *resulting* entity-throughput distribution (3) (calculated by taking into account the finite capacity of the subsystem performing the activity) denotes the probability that l entities are *executed* (*serviced*) by the activity in T time.

$$p_T^{\{c\}}[l] = P_T^{\{c\}}(L = l),\ \sum_{l=0,1,2,\dots} p_T^{\{c\}}[l] = 1 \tag{3}$$

(The three throughput distributions defined above – similarly to TFA[1] – represent different stages of the execution of the EFA method but there is only one type of throughput used in EFA while TFA has bit throughput for lines and packet throughput for nodes.)

The *delay* distribution denotes the probability that entities – taking into account the finite capacity of the subsystem performing the activity – are *delayed* by i number of T long time intervals where D denotes the random variable of entity-delay.

$$p_T^{\{d\}}[i] = P_T^{\{d\}}(D = iT),\ \sum_{i=0,1,2,\dots} p_T^{\{d\}}[i] = 1 \tag{4}$$

The *T parameter*[2] in the formulas above describes the size of the *throughput collection interval* and gives the *resolution* of the examination. If the value of T is defined too big compared to the change of the speed of entity generation then it will not show the change within the interval and if it is defined too small than it may result in too big amount of data for modelling. In the process analysis, a typical value to describe the speed of entity generation is an hour, but for example in case of Callcenters it may be one minute and in case of ICT-BP connections the typical value may be measured in seconds too.

The entity-load model of EFA should meet the following requirements:

- The *entity-load model* should describe the intensity and the time distribution of the entity-load.
- The entity-load model should have the feature of to be *closed* for the operation of *addition* that is the result of addition is of the same type as of its components.
- The entities of the same type, *between the entry and exit points* of the given entities can be handled together, in arbitrary order, in units of arbitrary size, grouped in an arbitrary way as *aggregated entity-load*.

Remark: The Size of Routing Unit (S_{RU}) is a feature of the entity-load model and describes the size of the unit of entity-load of a given type of entity that can be handled together.

[1] TFA uses a *continuous* approach in which the PDF (probability density function) of throughput (and PDF of delay) is approximated by a histogram, while according to the *discrete* approach in EFA, throughput (or delay) distribution is a PMF (probability mass function).

[2] The considerations about the value of the T parameter for TFA can be read in [7].

3.2. Modelling the capacity of activities in EFA

3.2.1. Activities, resources, capacities

Now, let us examine the *capacity of activities* from the point of view resources assigned to the activity.

Activity in the BP model may be viewed as a *subsystem* with its assigned resources performing the activity. In general, for the execution of the elementary task (represented by an entity) there should be allocated the necessary *technical resources* and a *set of human resources* to the subsystem (activity). The resources are used in a specific way determined by the activity to create the capability – which can be characterised by the capacity of activity – for the activity to be able to execute the elementary tasks.

The *capacity* of an activity is the *measure* of the *performance* of the subsystem in the execution of the elementary tasks.

The set of human resources act as a *group* – using the necessary and provided technical resources assigned to the activity – cooperating in a way depending on the activity in order to perform the activity. The *capacity of an activity* depends on the set of human resources allocated to the activity supposing that the *necessary technical resources are assigned* to the activity. The set of human resources of the activity may be characterised by the *type* and *quantity* of resources (as general features) and also by *activity-specific features* both on personal and on group level (for example features that can be described by *learning curves*). The *influence* on the capacity of the same type and quantity of human resource may be different in different activities.

The *transformation* of *a set of human resources* with given *types* and *quantities* into the capacity measure of a given activity (supposing that the required condition of technical resources for the performance of human resources is provided) requires a careful approach. The *calculation* of the capacity of an activity can be based on the analysis of *historical data* about the performance of the given activity with its assigned resources. (The historical data for the analysis may be for example, data collected by the ABC (Activity Based Costing) and CRM (Customer Relationship Management) systems.)

The *availability* of human resources required by an activity is changing in time and may be described in a probability manner in the following way:

- $P_{R_n}(R_n, week, day)$ – is the *Available Resource Value* of *type n* – for a given week of the year, and for the given day of the week, (this probability may be determined by the analysis of historical data for a longer time interval)
- $E(R_n)$ – *Expected Value* of the *Available Resource* of type *n* – for a given period

3.2.2. Analysing the behaviour of resources in BPs

The behaviour of resources assigned to the activities from the point of view of human resources will be examined below.

The BP is *embedded* into the environment formed by the *organisational structure*. Let us examine the *resource assignment* of an enterprise department which is a structural unit of the organisation (*Example 1*). In this case let the department be the *subsystem* dedicated for performing an activity of a BP. A frequent way of improving the efficiency is that a department – assumed critical – may *borrow* resources *temporarily* from other department(s) assumed non-critical. The result is that the amount of resources of the non-critical department(s) temporarily will *decrease*. In this case, it may be *observed* – if the loads of departments (entity-loads of activities) are independent – that the amount of resources of the non-critical department(s) will change *independently* from its load.

Let us examine the *change of resources* assigned to activities as the result of influences between business processes (*Example 2*). Let us analyse the following example: There are two *business processes* which have a common part in the set of resources (resources of their activities), some common elements in the set of loads, and may have different priorities for the access to the common set of *resources*. In this case, there may be *observed* different types of *correlation* between the change of the load and the amount of resources of the processes (thus in the amount of resources of their *activities*): for the process with *higher* priority for the access to the common set of resources there may be a *positive correlation* between the load and the amount of its resources, and at the same time for the process with *lower* priority for the access to the common set of resources there can be observed a *negative correlation* between the load and the amount of resources that can be used. The mutual influences between these processes may lead even to *curvilinear correlation* between the load and the amount of resources: negative type correlation and positive type correlation may exist between the load and the amount of resources used by a process for subsequent periods if the *priority* for the access to the common set of resources is changing (for example depending on some state of the BP system).

3.2.3. Requirements for the resource and capacity model of activities

The *resource and capacity model* of EFA should satisfy the following requirements:

- The effect of a given value of the *capacity of an activity* (capacity of the subsystem performing the activity) on the *time distribution of the entity-load* can be calculated.
- The *capacity of an activity* is defined by and can be calculated taking into account the *resources* assigned to the activity (to the subsystem performing the activity).
- The model of *resources* (measuring the quantity of resources) should be closed for the operation of addition in the case of resources of the same type.
- The *resource quantity* assigned to an activity can be decreased and the *subtracted* resource can be added to the resource of the same type assigned to other activity of the BP model.
- The quantity of a resource assigned to an activity cannot be negative.

- The same resource cannot be assigned to two activities at the same time.

Remark: The requirements above concerning the resources are formulated with the assumption that the *necessary technical resources are provided for the activities* thus they describe the requirements for the *human resources*.

3.3. Activity-Capacity Model

The *capacity of an activity* may be defined in the following way: The *capacity C* of an *activity* takes the value *K* if the (*maximum*) number of the *elementary tasks* (entities) that can be *executed* by the activity in the *T* interval is *K*.

3.3.1. Activity-capacity modelling: Constant capacity model

If the value of the *capacity* of an activity is a *constant* then the *resulting entity-throughput* ($p_T^{\{c\}}[l]$) and *delay* distribution ($p_T^{\{d\}}[i]$) can be calculated according to the formulas described for TFA:

$$p_T^{\{c\}}[l] = \begin{cases} p_T^{\{s\}}[l], \ l < K \\ \sum_{l=K}^{\infty} p_T^{\{s\}}[l] \ , l = K \\ 0, \ l > K \end{cases} \tag{5}$$

$$p_T^{\{d\}}[i] = \begin{cases} \sum_{l=0}^{K} p_T^{\{s\}}[l] \ , i = 0 \\ \sum_{l=iK+1}^{(i+1)K} p_T^{\{s\}}[l], i \geq 1 \end{cases} \tag{6}$$

where *K* is the value of the constant capacity of the activity

3.3.2. Activity-capacity modelling: Probability distribution model

Let us examine the possibility when *the capacity of the activity* is *not a constant* and the *capacity* and of the activity is *changing independently* from its load (see Example 1 in 3.2.2.).

In this case, a *probability distribution of the capacity of activity* may be used to model the changing capacity of the activity:

$$p_j = P(C = K_j), \ \sum_{j=1}^{N} p_j = 1 \tag{7}$$

where p_j is the probability that the random variable of the capacity *C* assumes the value K_j, *N* is the number of different values that the capacity of the activity can take, and the set of events that the capacity of the activity takes one of the values K_1, K_2, ..., K_N is *jointly exhaustive*. For the modelling, the value of *N* should be limited to a reasonably small number. For this purpose, the capacity values with low probability may be replaced, for example, according to the following formulas (supposing that the capacity values $K_1, K_2, \dots, K_{N-1}, K_N, K_{N+1}, \dots, K_n$ are ordered according to the decreasing probabilities ($p_1 \geq p_2 \geq \dots \geq p_{N-1} \geq p_N \geq p_{N+1} \dots \geq p_n$) :

$$K_N^* = \frac{\sum_{j=N}^{n} p_j K_j}{\sum_{j=N}^{n} p_j}, \ p_N^* = \sum_{j=N}^{n} p_j \tag{8}$$

where K_N^* is the value replacing the low probability values $K_N, K_{N+1}, \dots, K_n$ and p_N^* is the probability of K_N^*. (Thus, the expected value of the capacity will not change.) If the replaced capacities have very different values then there can be used different replacing values for different intervals (clusters).

Depending on the *speed of the change of C (quick or slow*) there can be used different approaches. The change of *C* is *quick* if *C may* take different values in the examined time interval of the throughput distribution. For the case of the *quick* change, the *total probability theorem* may be applied for the *resulting entity-throughput* distribution:

$$\left[p_T^{\{c\}}[l]\right]_{K=K_j} = \begin{cases} p_T^{\{s\}}[l], \ l < K_j \\ \sum_{l=K_j}^{\infty} p_T^{\{s\}}[l] \ , l = K_j \\ 0, \ l > K_j \end{cases} \tag{9}$$

$$p_T^{\{c\}}[l] = \sum_{j=1}^{N} \left[p_T^{\{c\}}[l]\right]_{K=K_j} p_j \tag{10}$$

(where N is the number of different values that the capacity of the activity can take in case of the quick change)

and for *the delay* distribution too:

$$\left[p_T^{\{d\}}[i]\right]_{K=K_j} = \begin{cases} \sum_{l=0}^{K_j} p_T^{\{s\}}[l] \ , i = 0 \\ \sum_{l=iK_j+1}^{(i+1)K_j} p_T^{\{s\}}[l], i \geq 1 \end{cases} \tag{11}$$

$$p_T^{\{d\}}[i] = \sum_{j=1}^{N} \left[p_T^{\{d\}}[i]\right]_{K=K_j} p_j \tag{12}$$

If the change of the value of the capacity of an activity is *slow* then the capacity of the activity may be determined for each simulation run using the *probability distribution of the capacity of activity* distribution. (The change of *C* is *slow* if *C* is changing but it may be approximated by different constant values for the examined time intervals of the throughput distribution.)

If the *variance* of *the capacity of an activity* is small, then the expected value of the capacity (*m*) may be used:

$$m = E(C) = \sum_{j=1}^{N} p_j K_j \tag{13}$$

In this case, for example, the distribution of the *delay* may be calculated as:

$$p_T^{\{d\}}[i] = \begin{cases} \sum_{l=0}^{[m+0.5]} p_T^{\{s\}}[l] \ , i = 0 \\ \sum_{l=[mi+0.5]+1}^{[m(i+1)+0.5]} p_T^{\{s\}}[l], i \geq 1 \end{cases} \tag{14}$$

(In the formula, the value [*m+0.5*] is used which is the value of *m* rounded to an integer.)

3.3.3. Activity-capacity modelling: Regression model

For the case, described in *Example 2* (in 3.2.2.), a *regression model* – based on the *analysis* of *historical data* – may be used to establish the relationship between the capacity and the load of the activity.

For this case, the *simple regression linear model* has the following form:

$$C = (\alpha + \beta L) + \varepsilon \tag{15}$$

where C is the capacity of activity (dependent variable), L is the load of the activity (independent variable), α is the population intercept, β is the population regression coefficient, and ε is the error term.

The observed value of activity capacity in *j*-th observation may be expressed by the equation:

$$c_j = \hat{c}_j + e_j \tag{16}$$

where c_j is the observed (measured) value of the capacity of activity, $\hat{c}_j$ is the *estimated* (calculated by regression) capacity for the given (observed, measured) load of activity, and e_j is the random error. The *estimated* capacity may be expressed as:

$$\hat{c}_j = a + b\, l_j \tag{17}$$

where *a* is the least-squares intercept (*estimator* of α), *b* is the least-squares regression coefficient (*estimator* of β), and l_j is the *j-t*h observed value of the load of activity ($j=1,2, \ldots, n$, where *n* is the number of observations).

For the values of *l* different from the observed values, the capacity of the activity may be *predicted* using the regression for interpolation or extrapolation.

A possible way of improving the quality of the regression model of the activity-capacity may be to include the load of other activities into the set of independent variables in order to take into account influences between activities. For this case, the *multiple regression linear model* may be described as follows:

$$C_i = (\alpha + \beta_1 L_1 + \beta_2 L_2 + \ \ldots + \beta_v L_v + \ \ldots + \ \beta_k L_k) + \varepsilon \tag{18}$$

where C_i is the capacity of activity *i* (dependent variable) in the BP model, $L_1, L_2, \ldots, L_v, \ldots, L_k$ are the independent variables (where L_v is the load of activity *v*, $v=1,2, \ldots, k$, *k* is the number of independent variables), α is the population intercept, $\beta_1, \beta_2, \ldots, \beta_v, \ldots, \beta_k$ are the population regression coefficients, indicating the mean effect of each *L* on *C* (in the population), holding all other *L* constant, and ε is the error term.

The set of *necessary* independent variables is proposed to include the loads of activities in the model of BP (system). The number of the *necessary and sufficient* explanatory variables can be determined in the *regression diagnostics* procedure. In the regression diagnostics, the *goodness of fit* of the regression model and the *statistical*

significance of parameters is examined, for example, using the following methods (Siegel 1994): *R-squared analysis* to determine the variability in *C* explained by the *L* variables as a group, *F-test* to check the statistical significance of the overall fit. If the *F-test* is significant then *t-test* can be used to examine individual regression coefficients, to see whether a particular *L* variable has an effect on *C*, holding the other *L* variables constant, to decide about the load of a particular activity as an explanatory variable to remain in the regression model.

In case of *non-linearity* in regression, in order to decrease the amount of computing, a *segmented linear regression* method – for example the method of the *multivariate adaptive regression splines* [5] – is proposed to be used.

Remark: In the capacity models, the availability of resources of a given activity should be taken into account too (see 3.2.1.).

4. Executing the EFA method

4.1. Stages of the EFA method

The work of the EFA method is divided into *two stages*:

Stage I: Distribution of the entity-load in the process (*spatial distribution*)

- Generation of entity-throughput statistics and sending them to the activities according to the routing decisions of the process in units determined by the S_{RU}

Stage II: Calculation of the influences of capacities of activities

- Summation of the distributed statistics
- Determination of the capacities of activities (Depending on the behaviour of resources in the process there may be used *constant* or *changing* capacities of activities. To model the changing capacity the *probability distribution* model or the *regression* model may be appropriate.)
- Calculation of the influences of the defined capacities (calculation of the resulting entity-throughput and delay distributions)

4.2. One-phase and multi-phase execution of EFA

The *spatial distribution in EFA* (*Stage I*) may be executed in two ways:

According to the *One-Phase Method,* the entity-load is distributed to every activity in one phase

According to the *Multi-phase Method,* in one phase the distribution of the entity-load is performed for activity group featured with equal distance from the entry point (The distance is determined by the number of activities on the route.) In Multi-phase Method, the step *Stage II* is executed for every phase. (The *Multi-phase Method* can be an efficient, for example, in case when the routing in BPs should take into account the changing capacities of activities.)

For the *Multi-phase Method*, it is proposed to take into account some constraints: There should be one entry point for the entity generation, the BP model proposed to be built in a way that there could be no big differences between the execution times of activities in one phase and the activities of one phase should get all their distributed entity-load.

Remark: The feed-back in the model may be taken into account as an entity-load proportional to the input of the related activity.

4.3. Features of the EFA method

- The method has the possibility of *parallelisation*: the calculations for different activities may be executed parallel by multiple processors.
- The speed of EFA can be increased using Routing Units. Changing S_{RU} is a tool to find the compromise between *accuracy* and *speed.*
- The results of EFA are approximate but the *critical paths* may be determined by the method.
- EFA shows the *steady state work* of the process (thus there is no need for warm-up time definition). The transient behaviour of processes may also be examined using a series of EFA runs.

4.4. The power of the EFA method

The combination and interworking of EFA method with TFA and DES models of ICT and BP systems may result in more accurate modelling, better simulation results and may also give a good chance to efficient parallel execution of the simulation of a system containing ICT and BP parts, both in the phases of preliminary and detailed modelling and both for critical and non-critical parts of systems.

5. Conclusions

In this paper, we have introduced a new method for the fast performance analysis, EFA. The EFA method extends the TFA approach for business processes.

We have overviewed and analysed the usual DES elements in EFA models: activities, links, entity generation and entity routing.

We have defined model elements for EFA:

- we have defined entity-load model as the entity-throughput model,
- we have given formulas and methods for delay-time calculation by introducing *a new model of the capacities of activities of BPs*.

Using the introduced elements we have outlined two versions of EFA method: One-phase method for rapid analysis and Multi-phase method for a more precise fast evaluation.

We have given solution to the problem of possible feed-back loops in the examined process.

References

[1] Aalst, W. M. P. van der, Nakatumba, J., Rozinat, A., Russel, N.: *Business Process Simulation: How to get it right?*, External research report, Eindhoven University of Technology (2008).

[2] Banks, J., Carson, J. S., Nelson, B. L.: *Discrete-Event System Simulation*, Prentice Hall, Upper Saddle River, New Jersey, (1996).

[3] Curtis, G.: *Business Information Systems*, Addison-Wesley, Wokingham, UK, (1989).

[4] Davenport, T. H.: *Process innovation: Reengeneering work through information technology*, Harvard Business School Press, Boston, Massachusetts, (1993).

[5] Friedman, J. H.: *Multivariate Adaptive Regression Splines,* Annals of Statistics, Vol. 19, (1991), pp. 1–141.

[6] Koubarakis, M., Plexousakis, D.: *Business process modelling and design – a formal model and methodology*, BT Technol. J. Vol. 17, No. 4, (1999), pp. 23-35

[7] Lencse, G.: *Traffic-Flow Analysis for Fast Performance Estimation of Communication Systems*, Journal of Computing and Information Technology, Vol. 9, No. 1, (2001), pp. 15–27

[8] Lencse, G., Muka, L.: *Convergence of the Key Algorithm of Traffic-Flow Analysis*, Journal of Computing and Information Technology, Vol. 14, No 2, (2006), pp. 133–139

[9] Lencse, G., Muka, L.: *Expanded Scope of Traffic-Flow Analysis: Entity Flow-Phase Analysis for Rapid Performance Evaluation of Enterprise Process Systems*, in 2006 European Simulation and Modelling Conference, Toulouse, (2006), pp. 94–98

[10] Pearce, J.A., Robinson, R. B.: *Formulation, implementation and control of competitive strategy*, Richard D. Irwin Inc, Boston, (1991).

[11] Savén, R.: *Process Modelling for Enterprise Integration: review and framework*, Department of Production Economics, Linköping Institute of Technology, Linköping, Sweden, (2002).

[12] Siegel, F. A.: *Practical Business Statistics*, Irwin, Burr Ridge, Illinois, (1994).

Transient Analysis Module from an Object Oriented Electrical Circuit Designer Application

E. Katona

**Széchenyi István University, Department of Telecommunication,
Laboratory of Electromagnetic Fields, H-9026, Egyetem tér 1, Győr, Hungary
email: eva.katona@inbox.com**

Abstract: The full paper deals with a software module, developed by the first Author, from a complex electrical circuit designer application. This analysis module presents the transient response of the designed network to the excitation applied to it. The results can be calculated by some different type ordinary differential equation solvers. The most important aim of this part of the research is creating a new algorithm, which the all different type ordinary differential equation solvers can be used with one algorithm.

Keywords: application, circuit analyzer, transient analysis, differential equation solver

1. Introduction

This paper deals with a new software module from a complex electrical circuit designer application [5,6] shown in Figure 1, developed by the first Author, what have been improved for two and a half years.

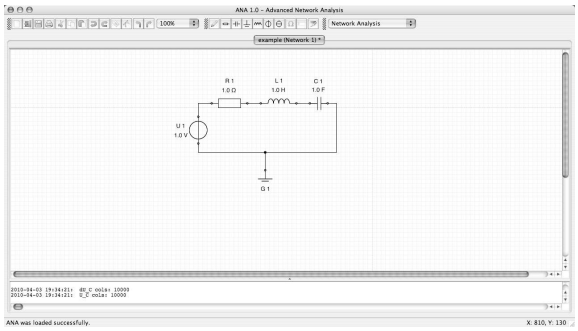


Figure 1. The graphical interface of the software, developed by the first Author

This analysis module presents the transient response of the designed network to the excitation applied to it. The result can be voltage, current or power function of the time [7].

The excitations can be different type of time functions, due to all, unit step, Dirac impulse, square, triangle, sine and cosine wave functions.

The results can be calculated by some different type ordinary differential equation solvers, due to all by the Euler algorithms, by the Galerkin algorithm [8], by the Cranck-Nicholson algorithm and by the Predict-Correct scheme [9].

The most important aim of this part of the research is creating a new algorithm, which the all different type ordinary differential equation solvers can be used with one algorithm.

2. About the analysis, resulting the time function

The transient analysis is based on the matrix equation [3]:

$$\begin{bmatrix} \underline{\underline{0}} & \underline{\underline{A}}_U^T & \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{A}}_U & \underline{\underline{A}}_R \underline{\underline{G}} \underline{\underline{A}}_R^T & \underline{\underline{A}}_C \underline{\underline{C}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{A}}_C^T & \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{A}}_L^T & \underline{\underline{0}} & \underline{\underline{-L}} \end{bmatrix} \begin{bmatrix} \underline{i}_U \\ \underline{\varphi}_1 \\ \frac{d\underline{u}_C}{dt} \\ \frac{d\underline{i}_L}{dt} \end{bmatrix} = \begin{bmatrix} \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & -\underline{\underline{A}}_L \\ \underline{\underline{1}}_C & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix} \begin{bmatrix} \underline{u}_c \\ \underline{i}_L \end{bmatrix} + \begin{bmatrix} \underline{\underline{1}}_U & \underline{\underline{0}} \\ \underline{\underline{0}} & -\underline{\underline{A}}_I \\ \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix} \begin{bmatrix} \underline{u}_s \\ \underline{i}_s \end{bmatrix}. \quad (1)$$

The solution of the matrix equation system gives the vectors of voltages $\underline{u}_C$ and $\underline{u}_s$ currents $\underline{i}_L$ and $\underline{i}_s$ and the vectors of $\frac{d\underline{u}_C}{dt}$ and $\frac{d\underline{i}_L}{dt}$, from which the normal form of the state equation can be obtained [9,4].

The results can be calculated by some different type of ordinary differential equation solvers, which can be chosen in a dialog box shown in Figure 2, after starting the analysis.

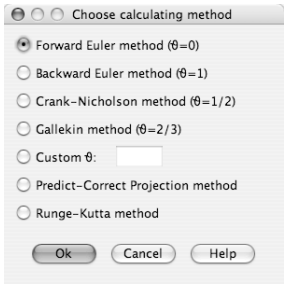


Figure 2. Changing attributes of sine wave

Ordered and simplified the equation by using $\dot{x} = \frac{x_{n+1} - x_n}{\Delta t}$, the next results can be given:

$$\begin{bmatrix} \dot{\underline{u}}_c \\ \dot{\underline{i}}_L \end{bmatrix} = \left[\underline{\underline{A}}\right] \begin{bmatrix} \underline{u}_c \\ \underline{i}_L \end{bmatrix} + \left[\underline{\underline{B}}\right] \begin{bmatrix} \underline{u}_s \\ \underline{i}_s \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} \dfrac{u_{C,\,n+1}-u_{C,\,n}}{\Delta t} \\ \dfrac{i_{L,\,n+1}-i_{L,\,n}}{\Delta t} \end{bmatrix} = \left[\underline{\underline{A}}\right] \begin{bmatrix} u_{C,\,n} \\ i_{L,\,n} \end{bmatrix} + \left[\underline{\underline{B}}\right] \begin{bmatrix} u_{s,\,n+1} \\ i_{s,\,n+1} \end{bmatrix} \tag{3}$$

From the form (3) can be seen, that it is a same equation form like

$$\frac{x_{n+1}-x_n}{\Delta t} \cong f(x_n, t_n). \tag{4}$$

(4) can be solved by using the Forward Euler method [4]:

$$x_{n+1} \cong x_n + \Delta t \cdot f(x_n, t_n), \tag{5}$$

by the Backward Euler method (solved by the Predict-Correct scheme) [4]:

$$x_{n+1} \cong x_n + \Delta t \cdot f(x_{n+1}, t_{n+1}) \tag{6}$$

$$x_{n+1}^{(0)} \approx x_n + \Delta t \cdot f(x_{n,} t_n) \tag{7}$$

$$x_{n+1}^{(j+1)} \approx x_n + \Delta t \cdot f(x_{n+1}^{(j)}, t_n) \tag{8}$$

$$\left\| \frac{x_{n+1}^{(j+1)} - x_{n+1}^{(j)}}{x_{n+1}^{(j+1)}} \right\| < 10^{-8}, \tag{9}$$

by Crank-Nicholson method [4]:

$$x_{n+1} \cong x_n + \frac{1}{2}\Delta t \cdot \left[f(x_n, t_n) + f(x_{n+1}, t_{n+1}) \right], \tag{10}$$

and by Galerkin method [8]:

$$x_{n+1} \cong x_n + \frac{2}{3}\Delta t \cdot f(x_n, t_n) + \frac{1}{3}\Delta t \cdot f(x_{n+1}, t_{n+1}). \tag{11}$$

The results can be compared to thanks for the different solvers. The time functions can be analyzed with a little function viewer application, shown in Figure 3, developed by the first Author. After clicking on the analyzable dipole of the current network, the voltage, current or the power time function of the actually component as it can be seen in the viewer.

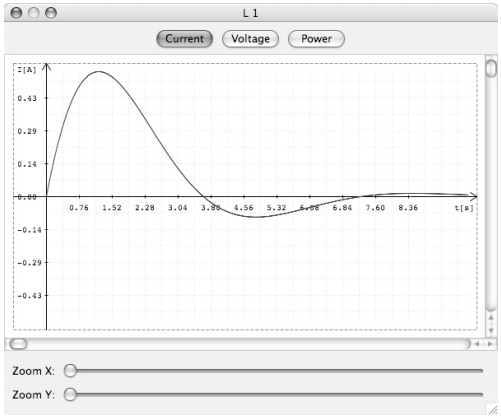


Figure 3. Function viewer application

3. About the new algorithm

With a ϑ constant - what must be between 0 and 1 – and with the

$$\frac{x_{n+1}-x_n}{\Delta t}=\vartheta\frac{dx_{n+1}}{dt}+(1-\vartheta)\frac{dx_n}{dt} \tag{12}$$

$$x_{n+1}\approx x_n+(1-\vartheta)\Delta t\cdot f(x_n,t_n)+\vartheta\Delta t\cdot f(x_{n+1},t_{n+1}) \tag{13}$$

equations [3], using the Predict-Correct scheme, a new simplified algorithm can be given, which what, all the different type of ordinary differential equation solvers can be used with one algorithm.

The type of the solvers depended by the value of ϑ. If $\vartheta = 0$, the Backward Euler method, if $\vartheta = 1$, the Forward Euler method, if $\vartheta = 1/2$, the Crank-Nicholson method, and if the $\vartheta = 2/3$, the Galerkin method can be given.

The algorithm can be seen in Figure 4.

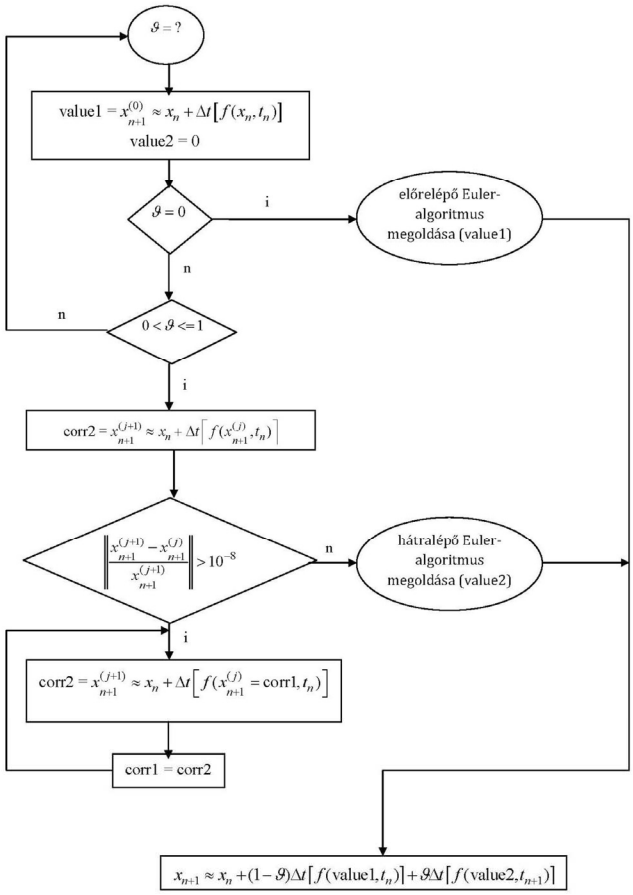


Figure 4. The new solver algorithm

4. Different type of excitations

The excitations can be different type of time functions, what can be chosen – after double-clicking on the graphic of the source shown in Figure 5. – in a dialog box what can be seen in Figure 6.

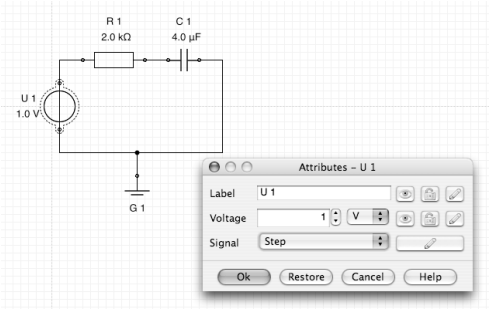


Figure 5. The attributes dialog box of U1 voltage source

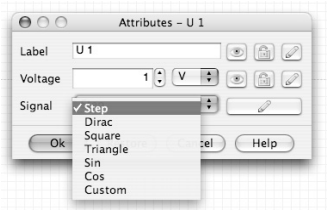


Figure 6. Changing the types of the excitation

The time function can be unit step, Dirac impulse, square, triangle, sine and cosine wave functions. Every type of functions has a lot of attributes that can be changed by the decision of the user:

- unit step: amplitude (A), start of edge (T)
- Dirac Pulse: amplitude (A), width of pulse (T),
- square wave: amplitude (A), frequency (f), rise/fall time (τ), period (T),
- triangle wave: amplitude (A), frequency (f), period (T),
- sine wave: amplitude (A), frequency (f), phase (θ), period (T),
- cosine wave: amplitude (A), frequency (f), phase (θ), period (T).

These attributes can be chosen in a dialog box shown in Figure 7.

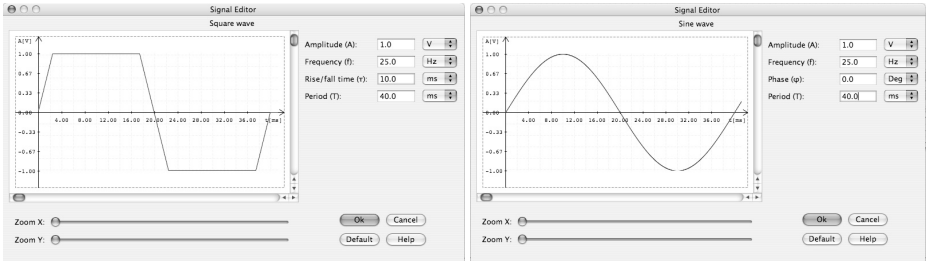


Figure 7. Changing attributes of a triangle and a sine wave

5. Example 1

In small circuits, with the different type of solvers, the same results must be obtained. After testing the new algorithm, the next same results could be given to the all type of solvers.

The example circuit and the start of the analysis can be seen in Figure 8. The RLC circuit parameters are: $U_1 = 1V$, $R_1 = 2k\Omega$, $L_1 = 4mH$ and $C_1 = 2\mu F$ [2].

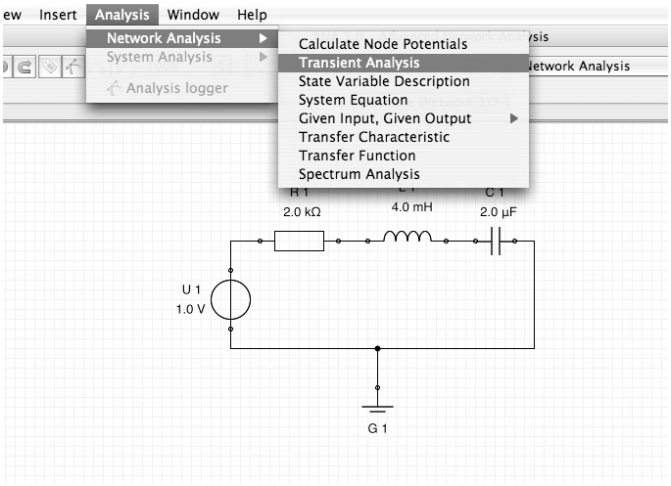


Figure 8. Starting the analysis

The results can be seen in Figure 9 – Figure 17.

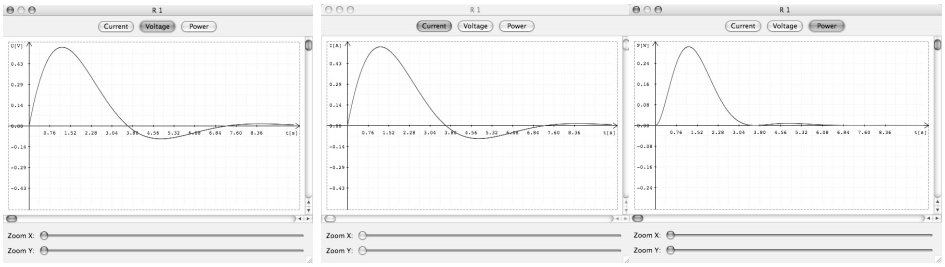


Figure 9 – 11. Voltage, current and power of R1 resistor

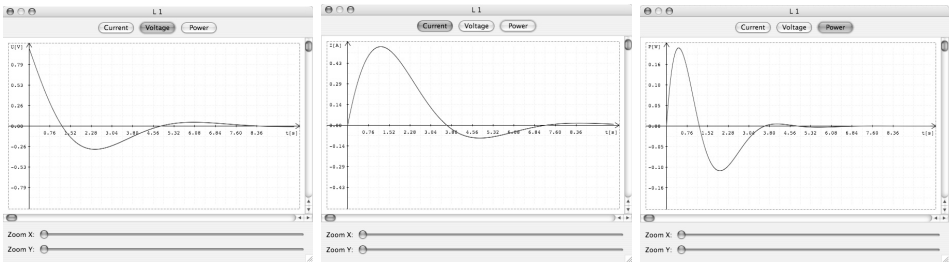


Figure 12 – 14. Voltage, current and power of L1 coil

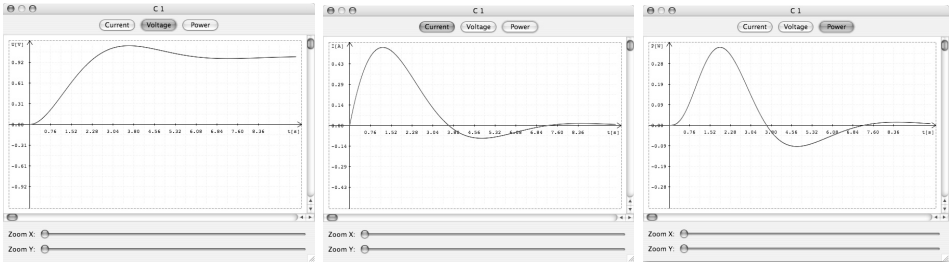


Figure 15 – 17. Voltage, current and power of C1 capacitor

6. Example 2

The second example shows the results of the example RC circuit – what can be seen in Figure 18. – to different type excitations [1].

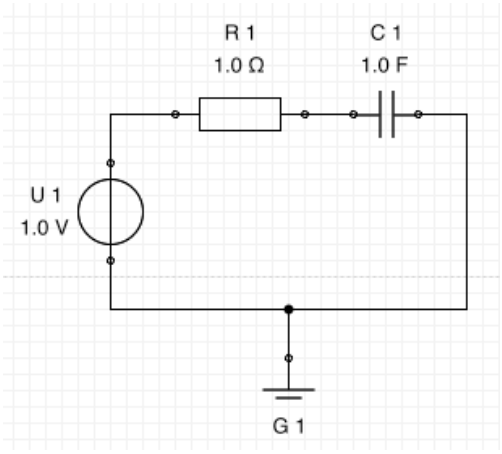


Figure 18. The example RC circuit

In every case – shown in Figure 19. - Figure 28. – the first picture shows the excitations, and the second shows one result time function from a component of the current network.

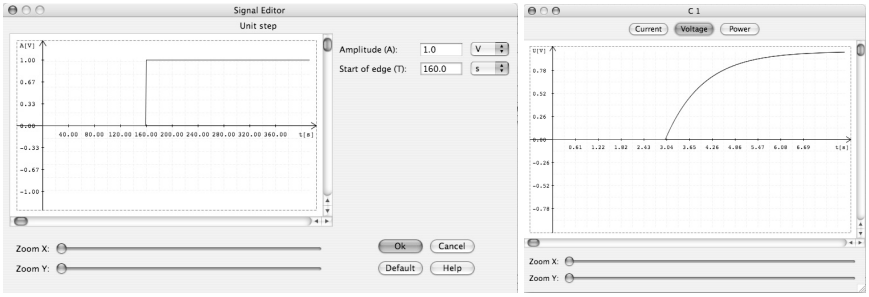


Figure. 19 – 20. Unit step excitation with the voltage result of C1 capacitor

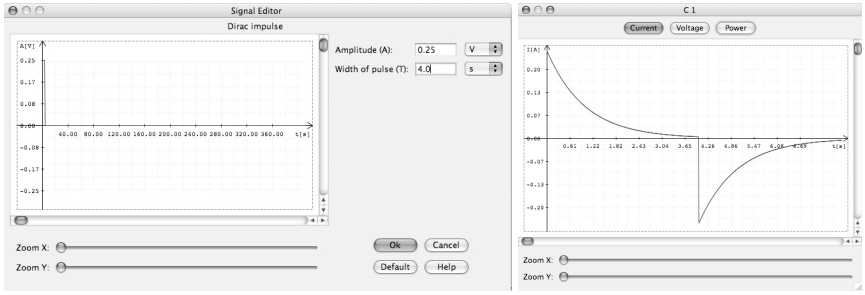


Figure 21 – 22. T=4s with of pulse excitation with the current result of C1 capacitor

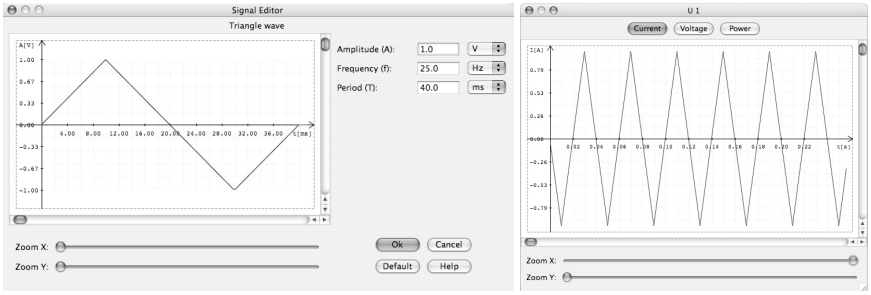


Figure 23 – 24. T=40ms triangle wave excitation with the 100% zoomed current result of U1 voltage source

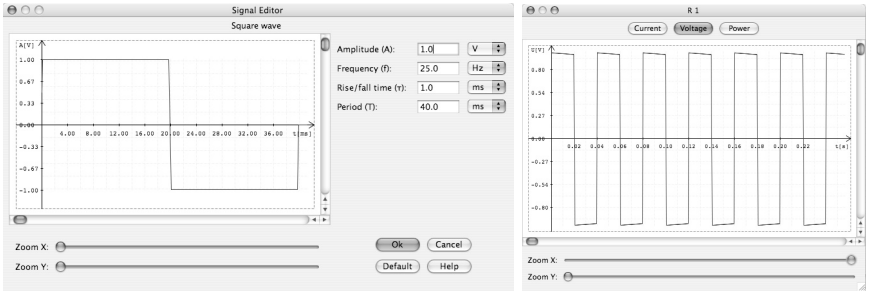


Figure 25 – 26. T=40ms square wave excitation with the 100% zoomed voltage result of R1 resistor

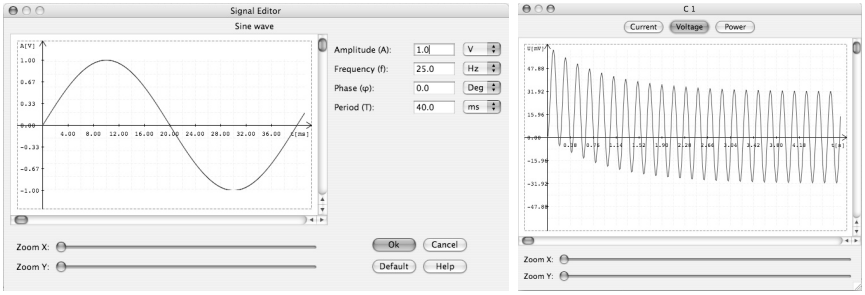


Figure 27 – 28. T=40ms sine wave excitation with the voltage result of C1 capacitor

7. Conclusion

In conclusion, the deal of this paper is to show a little part of a huge application. The aim of this research is to implement a lot of analysis modules, like this.

The most important part of the work of last period, was creating a new algorithm, which can be solving the different type differential equation solvers with one step. The other part of the work was implementing different type excitations to the application, and testing the given time results with lot of different circuits.

Until our days, the software has three implemented and tested complex analysis module. With the application the nodal potentials, the currents, voltages, powers and the time results of dipoles, the eigenvalues, and the network equations can be calculated. Lot of type networks can be edited, analyzed, and tested and their work can be followed with some other module of the program.

The next part of the work is the Fourier-transform, and the Spectral Analysis module.

The other future plans are due to all, calculate and display transfer characteristics, transfer function, shape preserving signal transmission, Laplace-, z-transform, Bode plot, and Nyquist plot. After all, the application will be improved with many other services [2].

References

[1] Fodor, Gy.: *Networks and Systems*, Műegyetemi Kiadó, Budapest, (2006).
[2] Fodor, Gy.: *Nodal Analysis of Electrical Networks* (Studies in Electrical and Electronic Engineering), Elsevier Science Ltd, (1988).
[3] Fodor, Gy.: *Signals, Systems and Networks*, Akadémiai Kiadó, Budapest, (2006).
[4] Géher, K.: *Linear Networks*, Műszaki Könyvkiadó, Budapest, (1972).
[5] Katona, É. Kuczmann, M.: *ANA – Advenced Network Analysis Java Software Package for Analizing, Designing, and Real Time Testing Networks and System*, Proceedings of the 2ND Symposium on Applied Electromagnetics, SAEM'08, ZAMOŚĆ, Poland, (1-4. June, 2008) PP. 83-86, CD Proceedings
[6] Katona, É. Kuczmann, M.: *Analysis and Design of Electrical Circuits*, Proceedings of the XIX PSAE Symposium Applied Electromagnetism in Modern Technologies and Informatics, PSAE 2009, Worliny, (21-24 June, 2009).
[7] Kuczmann, M.: *Signals and Systems*, Universitas-Győr Kht. Győr (1999).
[8] Kuczmann, M. Iványi, A: *The Finite Element Method in Magnetics* , Akadémiai Kiadó. Budapest (2008).
[9] Simonyi, K. Zombory, L.: *Theoretical Electromagnetism*, Műszaki Könyvkiadó, Budapest, (2000).

Acknowledgments: This paper was supported by the Hungarian Scientific Research Fund (OTKA PD 73242), and by the Széchenyi István University (15-3210-02).