ITHEA

International Journal INFORMATION TECHNOLOGIES & KNOWLEDGE

2009 Volume 3 Number 3

International Journal

INFORMATION TECHNOLOGIES & KNOWLEDGE

Volume 3 / 2009, Number 3

IJ ITK is official publisher of the scientific papers of the members of the ITHEA International Scientific Society

IJ ITK rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITK as well as the **subscription fees** are given on ***www.ithea.org***
The camera-ready copy of the paper should be received by e-mail: info@foibg.com.
Responsibility for papers published in IJ ITK belongs to authors.

General Sponsor of IJ ITK is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Vol.3, Number 3, 2009

Edited by the **Institute of Information Theories and Applications FOI ITHEA®**, Bulgaria,
in collaboration with the **V.M.Glushkov Institute of Cybernetics of NAS**, Ukraine,
and the **Institute of Mathematics and Informatics**, BAS, Bulgaria.

Publisher: ITHEA®

Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: info@foibg.com

Printed in Bulgaria

ISSN 1313-0455 (printed) ISSN 1313-048X (online) ISSN 1313-0501 (CD/DVD)

GENERALIZED INTERVAL ESTIMATIONS IN DECISION MAKING AND SCENARIO ANALYSIS

Gennady Shepelyov, Michael Sternin

Abstract: *Main features of a new method of expert knowledge elicitation – method of generalized interval estimations – and peculiarities of its using in decision making are presented. Expert estimation of an analyzed quantitative parameter is defined in the framework of the method by a set of intervals. This set may be shown on a plane (X, Y) by a curvilinear trapezoid. If densities of probability distributions were defined on Y-axis ("weights" of different interval estimations at their set) and on X-axis (for all interval estimations at the set) we received generalized interval estimation (GIE) of the parameter. GIE method may be used at different directions. Firstly, we may reduce GIE to mono-interval estimation averaging distributions on X-axis, taking into account their weights, and use then famous probabilistic methods to analyze problem. The average distribution is in fact a probability mixture of distributions on intervals of their set. Secondly, we may study complete structure of interval estimations that reflects expert knowledge in details. Here we automatically receive for resulting diagrams (X - parameter value, Y - level of probability) such probability distributions that are boundaries of all distributions of GIE (probability tube, or box). Besides we may receive probability distributions for different sections of p-tube both for parameter values and for probability levels and use these curves in the process of decision-making. At last, we may use GIE method as an instrument of scenario analysis in decision-making. Illustrative examples are given in the paper to demonstrate the decision-making, expert knowledge and scenario analysis aspects of the proposed approach.*

Keywords: *Generalized interval estimations, scenario analysis, generalized probability distributions, probability tubes.*

ACM Classification Keywords: *H.4.2 Types of Systems – Decision support; G.3 Probability and Statistics – Distribution functions.*

Introduction

During analysis of many practical problems there are rather often situations when decisions are made based on certain quantitative indicators, or criteria, important for decision-makers (e.g. net present value of investment projects in project analysis; present and future exchange rates, interest rates and stock prices in financial markets analysis; volumes of oil/gas reserves and prices in oil/gas property evaluation and others). The values of these indicators are usually the outcomes of various computational schemes (mathematical models). As a rule such models are already well known for many areas of human activity. However the complexity of real-life problems inevitably brings uncertainties into the values of input data (parameters) of models. In many situations, especially for multidisciplinary tasks, expert knowledge, in fact, could be the only way to overcome these difficulties and provide reasonable estimates for parameters. Thus to analyze many problems under uncertainty we should combine methods of expert knowledge elicitation, representation and methods of decision making. This stimulates the development of various approaches to expert judgment formalization, of methods and tools to elicit and represent expert knowledge in an adequate way.

In this paper, we present a survey of main results of a method to elicit and represent expert knowledge on quantitative parameters in problems under uncertainty – the method of Generalized Interval Estimations (GIE) [Shepelyov and Sternin, 2003; Chugunov et al., 2006; Chugunov et al., 2008 A; Chugunov et al., 2008 B; Shepelyov and Sternin, 2008]. We provide several illustrative examples to demonstrate both the decision-making, expert knowledge and scenario analysis aspects of the proposed approach.

Interval and Probabilistic Estimations

Until now, in spite of presence of other tools (e.g., fuzzy sets and possibility theory, rough sets), the main language that experts of real economy branches use to characterize uncertainties in input parameters of models is based on interval analysis and probability theory. Methods of interval analysis are adequate in situations under maximal uncertainty, when the expert is able to provide only the minimum and maximum value for the estimated parameter. But if one applies only interval methods to quantify uncertainty in the model input parameters, the width of the interval for the model outcome is likely to be very wide, and the value of such information often can hardly help one to make decisions. The combination of interval analysis and probability theory could be advantageous when the expert has more precise knowledge on the estimated parameters. In this case, the input data of models is represented by experts as random variables defined on certain intervals by their probability distribution functions that characterize uncertainty in the framework of interval estimations. If interval and probabilistic expert estimations for parameters of used models were received, final results of models could be calculated with methods of the Monte-Carlo family. Results of probabilistic calculations may be represented on a graph displaying the probability “not less than” (probability of guaranteed result) as a function of analyzed criteria.

Figure 1 shows an example of such graph for an oil reservoir evaluation for field that is in the predrilling (initial) stages of exploitation. Monte-Carlo calculations were made on base of the standard volumetric model [Welsh et al., 2004]. Estimated reserves Q are calculated as the product of input parameters of the volumetric model representing reservoir properties and the expert specified the input parameters as random variables.

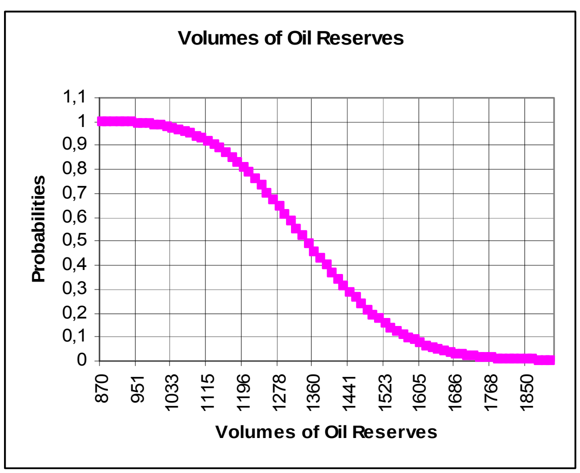


Fig. 1. An example of probabilistic reserves estimate

The decision-maker (DM) may analyze this graph from two points of view: what is the level of probability (chances on the realization of a chosen outcome) that corresponds to a desired value of oil reserves in place, or what is the value of the analyzed indicator for a given level of confidence? DM can then make a reasonable decision based on this interpretation. Based on the international classification of oil and gas reserves, the DM (investor) may obtain the reserves estimates (proved, probable and possible reserves categories) needed for a final decision.

For example, according to the classification, proved reserves would correspond to the probability level of 0.9; proved + probable reserves to the probability level of 0.6; proved + probable + possible ones to the probability level of 0.2. Let us note that these estimations are point ones though initial parameters had interval estimations. Various measures of risk connected with different possible decisions may be introduced and used in the framework of this approach.

Poly-interval and Generalized Interval Estimations

It may seem that the “mono-interval” approach could be enough to quantify uncertainty in decision making problems of mentioned above class. However, our joint work with geologists as experts showed some narrowness of such way. Experts, giving interval estimation of a parameter, may meet some difficulties both on length and on location of estimation. Attempts to eliminate these difficulties defining very wide interval estimations decrease usefulness of expert knowledge. On the other hand, if an expert tries to be too precise (i. e. the interval estimation is too narrow) there is a high probability that the true value of the parameter could be outside of the given range. Psychologists [Kahneman et al., 1981] observed this effect for a more general environment, not only related to oil and gas industries. Thus the problem has sufficiently universal character.

This motivated us to develop a method that could allow expressing an expert knowledge on an estimated parameter in a less restrictive way and thus quantify expert’s knowledge more precisely. Namely, we started with the idea to represent possible variability of interval estimation length and location by letting the expert to specify a set of intervals, rather than just one interval. We named such construction poly-interval estimation (PIE) of a parameter. In many situations experts would feel more comfortable thinking of several possible scenarios represented by several corresponding intervals, rather than assessing only one interval to describe “overall” uncertainty in parameter. Thinking in scenario-oriented framework would also make expert pay more attention to each possible scenario and might reduce effects from heuristics [Morgan and Henrion, 1992].

Expert defines in the framework of the method, as an initial estimation of parameter D, an interval of minimal, in his opinion, length (or “up” interval – “best guess”) $[D_{lu}, D_{ru}]$ – here l, r mean left, right boundaries of the interval - and then, taking into account all possible uncertainties, an interval of maximal, in his opinion, length (“down”, or base, interval – “safe bet”) $[D_{ld}, Drd]$. Now, one may further assume that a subset of intermediate intervals between up and base intervals also belongs to the PIE, although this is not necessary, since the subset of intermediate intervals may be selected from all possible ones based on the expert knowledge. PIE may be represented on a plane (X, Y) by a curvilinear trapezoid. The simplest example of a PIE for nested set of intervals is shown in figure 2.

Here Y-axis (α-axis in our notation) is the axis of marks of intervals and we suppose that $0 \le \alpha \le 1$, and X-axis (D-axis in our notation) is the axis of values of parameter under consideration. Side boundaries of trapezoid could be also specified by the expert when he/she defines the subset of intervals in PIE. Each interval of the trapezoid represents a possible scenario of parameter realization and the expert may also specify chances or probability for each scenario (and thus probability distribution on α–axis). If the expert defines distributions both on α- and D-axes we obtain generalized interval estimation (GIE) corresponding to the PIE. Probability distribution on α-axis represents, in a sense, “weights” of scenario-bands. Probability distributions on D-axis represent chances, according to expert knowledge, on realization of possible values of parameter for a given scenario-interval. Thus GIE of a parameter is characterized by joint distribution function of two random variables α and D with density function $\Psi(\alpha, D) = f_1(\alpha)f_2(D|\alpha)$, defined on PIE of the parameter. Hence, there are several ways to represent uncertainty in the framework of GIE method: specification of initial (characteristic) interval estimations; shape of PIE; specification of distributions on scenario-intervals; specification of "chances" or "weights" for scenario-intervals. This also means that generalized interval estimations are able to represent the following two aspects of uncertainty. On the one hand, the expert may express certain confidence in that some subset of intervals from PIE adequately represents his/her understanding of the estimated parameter's value (this is described by $f_1(\alpha)$). On the other hand, each possible scenario implies inherent variability of parameter's value bound by the corresponding interval (this is described by $f_2(D|\alpha)$).

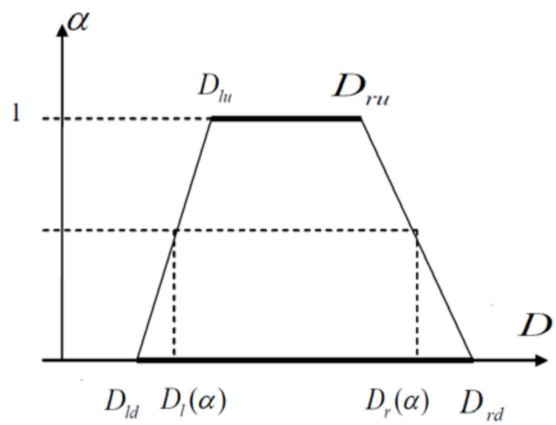


Fig. 2. Poly-interval estimation of an input parameter for nested set of intervals

One may build more general PIE and corresponding GIE, for example if an expert defines initial characteristic interval estimations to represent optimistic, realistic and pessimistic estimations corresponding to the bottom, middle and top intervals shown on Figure 3. Here we have shifted, not embedded as above, set of intervals.

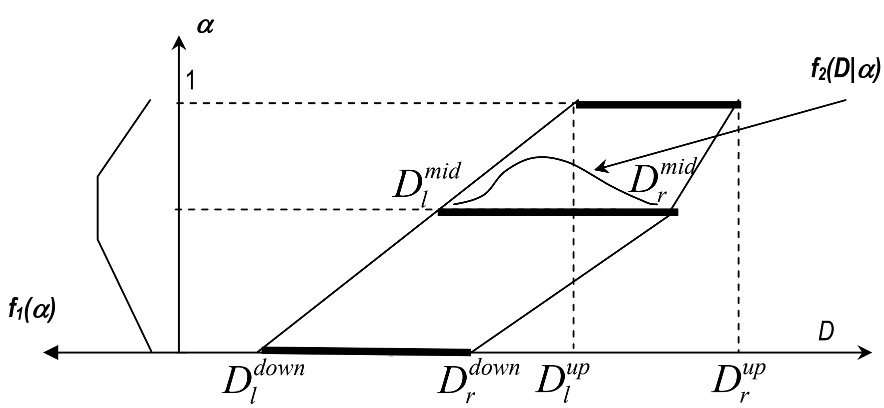


Fig. 3. Generalized interval estimation of an input parameter for shifted set of intervals

Another example of "shifted" PIE may arise in estimating volumes of oil recoverable reserves for ill-investigated fields. Volumes of recoverable reserves would depend not only on geology of the reservoir and technologic conditions but also on the oil prices. As the oil prices goes up, so does the estimation of recoverable reserves volumes. Let us construct the following PIE: Y-axis is axis of the oil prices, X-axis is axis of volumes of recoverable reserves. For each point estimate of the oil price there exists an interval estimation of recoverable reserves. The shape of such PIE would resemble a shape of the tower at Italian town Pisa, i.e. PIE is built on shifted intervals. This example also demonstrates additional possibilities of PIE/GIE method to quantify uncertainty in estimation of dependent parameters. Let also note that Y-axis is not here axis of interval marks but has real meaning.

Now there are three ways to use information provided by the expert in the framework of GIE approach for decision-making and expert analysis. GIE method allows one to obtain an aggregated (averaged) probabilistic distribution, which we also refer to as GIE projection, defined on the base interval of PIE (an interval, containing all possible values for D) and reduces calculation of the model output to a well-known statistical simulation procedure, which can be performed by Monte-Carlo family methods [Sternin and Shepelyov, 2003]. Alternatively, GIE can be treated as a set of intervals with a certain internal structure that represents uncertainty in expert knowledge concerning in estimated parameter [Chugunov et al, 2006]. At last GIE approach may be used as tool of scenario analysis in decision making [Shepelyov and Sternin, 2008]. All cases will be considered in the next sections.

GIE Projection: Mono-interval Approach in GIE Method

For simple case of nested intervals in PIE with straight boundaries (fig. 2) we can obtain the following expression for a probability density function of averaged distribution function $f_{av}(D)$:

$$f_{av}(D) = \begin{cases} \int\limits_0^{\alpha_l(D)} f_1(\alpha) f_2(D|\alpha) d\alpha, & D \in [D_{ld}, D_{lu}) \\ \int\limits_0^1 f_1(\alpha) f_2(D|\alpha) d\alpha, & D \in [D_{lu}, D_{ru}) \\ \int\limits_0^{\alpha_r(D)} f_1(\alpha) f_2(D|\alpha) d\alpha, & D \in [D_{ru}, D_{rd}]. \end{cases} \quad (1A)$$

Here $\alpha_l(D) = \dfrac{D - D_{ld}}{D_{lu} - D_{ld}}$, $\alpha_r(D) = \dfrac{D_{rd} - D}{D_{rd} - D_{ru}}$.

Then for averaged probability distribution function $P_{av}(D < D_S)$ we have:

$$P_{av}(D < D_S) = \begin{cases} \int\limits_0^{\alpha_l(D_S)} \int\limits_{D_l(\alpha)}^{D_S} f_1(\alpha) f_2(D|\alpha) d\alpha dD, & D \in [D_{ld}, D_{lu}) \\ \int\limits_0^1 \int\limits_{D_l(\alpha)}^{D_S} f_1(\alpha) f_2(D|\alpha) d\alpha dD, & D \in [D_{lu}, D_{ru}) \\ 1 - \int\limits_0^{\alpha_r(D_S)} \int\limits_{D_S}^{D_r(\alpha)} f_1(\alpha) f_2(D|\alpha) d\alpha dD, & D \in [D_{ru}, D_{rd}]. \end{cases} \quad (1B)$$

Here $D_l(\alpha) = \alpha D_{lu} + (1 - \alpha)D_{ld}$, $D_r(\alpha) = \alpha D_{ru} + (1 - \alpha)D_{rd}$. Similar but more complicated expressions could be derived for GIE with PIE of more general shapes, e.g., as the one in figure 3.

For simple distributions defined by functions f_1 and f_2 (e.g., uniform and triangular distributions) one can obtain direct analytical expression for GIE projection [Shepelyov and Sternin, 2003]. We will consider the case of uniform distributions in the paper some later.

Let us consider an illustrative example for a more complicated case, when GIE has different distributions on different scenario-intervals [Chugunov et al, 2006]. During the evaluation of investment projects in real-estate development business, the final decisions could be made only after detailed analysis of the lot under consideration and business perspectives of the area. Due to uncertainties of various natures, the values of some project's parameters are unknown at the moment of the analysis and are estimated by the expert. In case of business center development, such parameters would include the future price for square meter.

Based on the analysis and data available, the expert is asked to estimate this parameter. GIE method is used to elicit and quantify expert judgments. The results of elicitation process is given in figure 4, where we show density and cumulative probability functions of guaranteed results for the price for sq. m. as random variable. We also provide probability distributions for case where expert use mono-interval approach.

Interval from 2 to 11 thousand money units (m.u.) was specified by the expert as a base interval. Top interval was defined symmetrically – from 5 to 8 thousands m.u. Initially joint distribution function of GIE was built based on uniform distributions on both α and D.

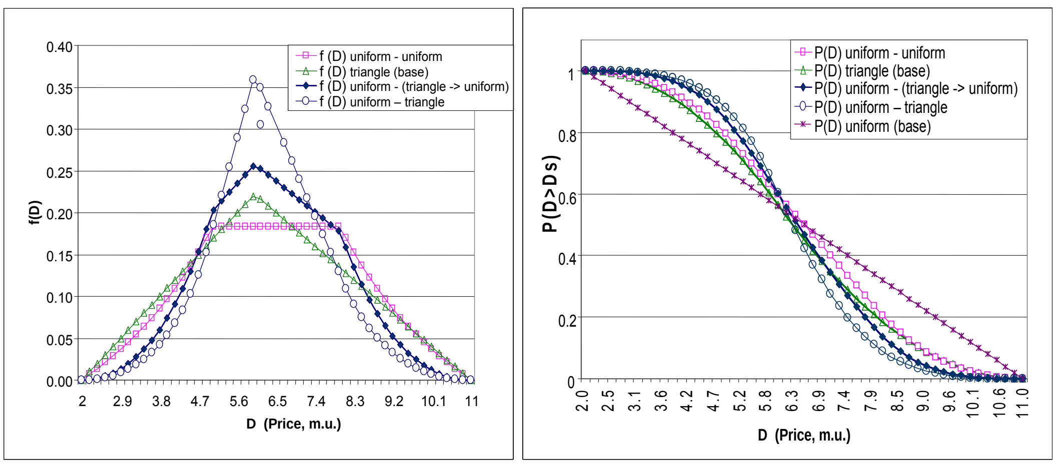


Fig. 4. Density function and cumulative probability function of guaranteed results for future price of sq. m.

Then expert tested a hypothesis with selection of the most preferable value for the price. The expert defined GIE as a combination of uniform distribution on α and triangle one on D with modal value at 6 thousands m.u. (this value is shifted to the left relative to the center of mini interval). This setup led to unreasonably high, in expert's opinion, chances for values in a neighborhood of the modal value. Since, according to the expert, the narrower the interval the more equiprobable the values are, the next hypothesis included uniform distribution on α and triangle distribution on D, with the same as earlier modal value on base interval, but it steadily transformed into uniform distribution on top (mini) interval.

Below we provide expression (2) for density $f(D|\alpha)$ of similar distribution, which could resemble one a shape of country house with roof and walls. On the base interval, the "house" has only "roof" (triangular distribution) and on mini interval it has only "walls" (uniform distribution).

$$f(D|\alpha)=\frac{1}{(D_r-D_l)}\begin{cases}\dfrac{2D-2D_l+\alpha(D_l+D_p-2D)}{D_p-D_l}, & D\in\left[D_l,D_p\right]\\ \dfrac{2D_r-2D-\alpha(D_r+D_p-2D)}{D_r-D_p}, & D\in\left(D_p,D_r\right]\end{cases}\qquad(2)$$

Here D_r, D_l are respectively right and left boundaries of corresponding intervals of PIE (for a straight line case D_r, D_l are given by (1B)), D_p is modal value of distribution.

According to fig. 4, one can see that at level P = 0.85 specified by the expert as his “comfort level”, the difference in estimations of prices for sq. m., corresponding to the results from using mono-interval triangular distribution and expert-defined GIE, is more than 10%: 4340 m.u. and 4880 m.u. respectively.

GIE approach allows one to generate new mathematical objects - probability distributions that are generalizations of traditional distributions. For example, one can construct generalized uniform distribution (GUD) as a projection of GIE when uniform distributions are defined on both axes. Let us consider this simple but important for practice case in more details. In fact a class of GUD may be constructed. Some representatives of GUD will be used later in the process of discussing relationship of GIE method and scenario analysis. Now we consider the simplest case assuming the straight line boundaries of PIE on embedded intervals. Uniform distributions on both α and D axes are defined by $f_1(\alpha) = 1$, $f_2(D|\alpha) = 1/[D_r(\alpha) – D_l(\alpha)]$. The integrals in (1A, B) may integrate by parts to obtain the probability density function for the GUD:

$$f(D)=\frac{1}{B-M}\begin{cases}Ln\dfrac{B(D_{lu}-D_{ld})}{B(D_{lu}-D)+M(D-D_{ld})}, & D\in\left[D_{ld},D_{lu}\right]\\ Ln\dfrac{B}{M}, & D\in\left(D_{lu},D_{ru}\right)\\ Ln\dfrac{B(D_{ru}-D_{rd})}{B(D_{ru}-D)+M(D-D_{rd})}, & D\in\left[D_{ru},D_{rd}\right],\end{cases}\qquad(3)$$

where $B = D_{rd} – D_{ld}$ and $M = D_{ru} - D_{lu}$. In the limit $M \to B$ this distribution transforms to traditional uniform one: $f(D) = 1/M$.

Corresponding expressions for cumulative probability of generalized uniform distribution are:

$$P(D<D_S)=\frac{1}{B-M}\begin{cases}\left(D_{lu}-D_{ld}\right)\alpha_l+\dfrac{B(D_{lu}-D_S)+M(D_S-D_{ld})}{B-M}Ln\dfrac{B-\alpha_l(B-M)}{B}, & \mathrm{D}\in\left[D_{ld},D_{lu}\right]\\ D_{lu}-D_{ld}+\dfrac{B(D_{lu}-D_S)+M(D_S-D_{ld})}{B-M}Ln\dfrac{M}{B}, & D\in\left(D_{lu},D_{ru}\right)\\ \left(D_{lu}-D_{ld}\right)\alpha_r+\dfrac{B(D_{lu}-D_S)+M(D_S-D_{ld})}{B-M}Ln\dfrac{B-\alpha_r(B-M)}{B}+ & \\ +(B-M)(1-\alpha_r), & D\in\left[D_{ru},D_{rd}\right]\end{cases}\qquad(4)$$

Here α_l and α_r are defined by (1A). It should be mentioned, that GUD may be interpreted as a probability mixture of uniform distributions with uniform mixing function. Specific distributions on the intervals in the mixture are defined by the shape of PIE.

It is interesting to compare behavior of uniform distributions defined on the base and mini intervals and GUD on the base interval. All three functions would intersect in one point. This point is defined by the following coordinates: $D_i = (BD_{lu} - MD_{ld})/(B - M)$, $P_i = (D_{lu} - D_{ld})/(B - M)$. Figure 5 shows three distribution functions for the following case: $D_{ld} = 10$, $D_{rd} = 50$, $D_{lu} = 30$, $D_{ru} = 35$.

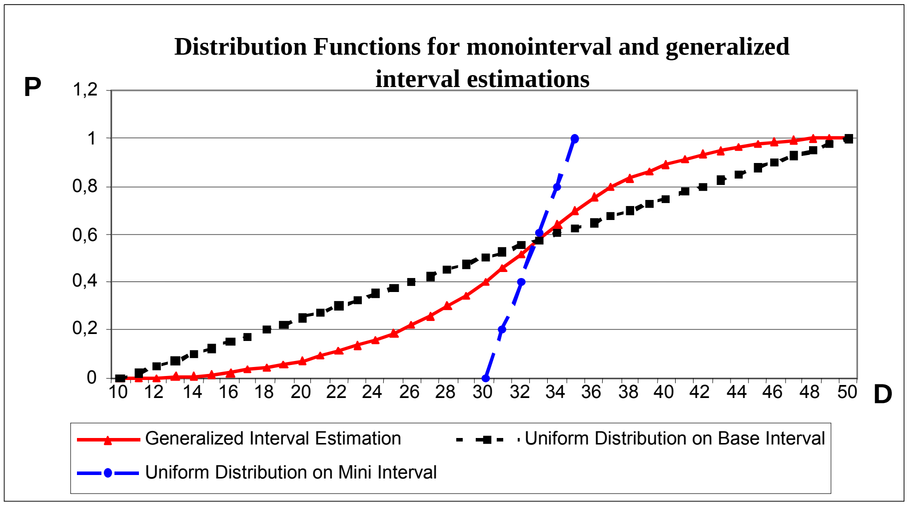


Fig. 5. Cumulative uniform probability functions for mono-interval and generalized cases

This interpretation of the GIE method may be used to analyze the decision making problems that earlier were approached with probabilistic mono-interval methods. In particular, oil reserves evaluation may be performed by GIE method when some or all of the input parameters of volumetric model are described by the expert as GIE estimations. Comparison of two final probabilistic curves for the oil reserves estimations shows that behavior of cumulative curve for GIE case is, in a sense, more balanced and reliable than the outcome for probabilistic mono-interval case. Specifically, GIE increases probabilities of small volumes of reserves and decreases probabilities of large volumes.

Let us now analyze the moments of GUD. Following traditional calculations for expectation $\langle D \rangle$, we obtain: $\langle D \rangle = (D_{ld} + D_{rd} + D_{lu} + D_{ru})/4$. Properties of GDU allow one to derive simple recurrence formula for central moments of n-th $\langle D_n \rangle_c$ and (n-1)-th $\langle D_{n-1} \rangle_c$ orders. To derive this expression, let us use integral definition of central moments and formulae (2) for probability density of GUD. After integration by parts, one can see that algebraic sum of logarithm terms equals zero, and residual can be expressed through polynomial term and moment $\langle D_{n-1} \rangle_c$. Because $BDlu - MD_{ld} = BD_{ru} - MD_{rd}$, we finally obtain:

$$\begin{aligned} <D_n>_c = [& \frac{(D_{lu}-<D>)^{n+1}-(D_{ld}-<D>)^{n+1}+(D_{rd}-<D>)^{n+1}-(D_{ru}-<D>)^{n+1}}{n+1} - \\ & -n\frac{(D_{rd}+D_{ld}-D_{ru}-D_{lu})(B+M)}{4}<D_{n-1}>_c]/[(n+1)(B-M)], \quad n \ge 2. \end{aligned} \tag{5}$$

Note that for symmetric PIE $D_{ld} + D_{rd} = D_{lu} + D_{ru}$, and the second term in (5) is equal to zero for this case. Formula (5) allows one to calculate central moment of (n+1)-th order, if one has central moment of n-th order and expectation for GUD.

From the theory of probability we know that the shape of probability distributions can be well characterized by the first four moments. In case of GUD, these moments can be calculated from (5). But for the first three of them we may obtain exact analytical expressions. Let us introduce the following notation $<D_d> = (D_{ld} + D_{rd})/2$, $<D_u> = (D_{lu} + D_{ru})/2$, $L_r = D_{rd} – D_{ru}$, $L_l = D_{lu} – D_{ld}$. Note that $L_r = L_l$ for symmetric PIE.

Then for the variance $<D_2>_c$ of GUD we have:

$$< D_2 >_c = \frac{B^2 + M^2 + MB}{36} + \frac{(< D_d > - < D_u >)^2}{12} = \frac{B^2 + M^2 + MB}{36} + \frac{(L_r - L_l)^2}{48}.$$

The first term of this expression shows dependence of variance on the lengths of intervals in PIE, the second one shows dependence of variance on a possible asymmetry of PIE.

For the coefficient of skewness $<D_3>_c$ we obtain:

$$< D_3 >_c = \frac{(B^2 - M^2)(< D_d > - < D_u >)}{48} = \frac{(< D_{2d} >_c - < D_{2u} >_c)(< D_d > - < D_u >)}{4}.$$

Here we take into account the fact that variances of mono-interval uniform distributions defined on the base and mini intervals are equal to $B^2/12$ and $M^2/12$ respectively.

The sign of $<D_3>_c$ is determined by the sign of a difference $<D_d> - <D_u>$, because $B^2 > M^2$. This means that GUD has negative asymmetry (shifted to the left) if $<D_d>$ is less than $<D_d>$, and it has positive asymmetry (shifted to the right) if $<D_d>$ is larger than $<D_u>$.

Values of kurtosis may be also calculated from (5). If we compare standardized abrupt normal distribution defined on finite interval with GUD, we see that there are allokurtic GUD: specifically, kurtosis for symmetric GUD may have both positive values and negative values and also zero value (Figure 6).

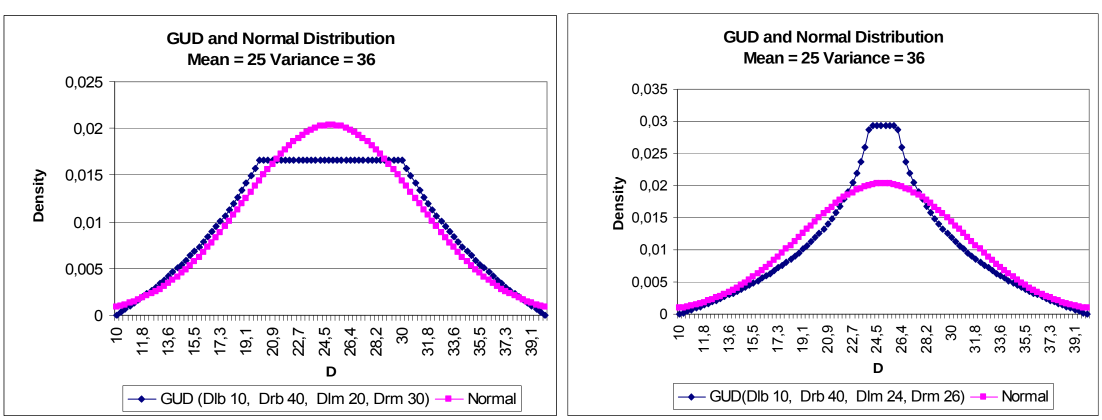


Fig. 6. Densities of generalized uniform and normal distributions for different values of kurtosis

Another set of decision-making problems that can be approached in the framework of the GIE method by GIE projection calculation includes dynamic (time-dependent) problems and problems with dependent parameters.

Figure 7 illustrates situation when the expert describes dynamic behavior of parameter D_2 as a pair of intervals for each time (parameter D_1). This setup could be represented by a “tube” with walls of finite variable thickness. Now we may construct GIE for each value of time if expert could also specify probability distributions on t-sections. The base intervals of the GIE for each time step are defined by outer boundaries of the tube and mini intervals are defined by inner ones. Since time dependence could be considered as just a case of general parameter dependence, GIE for dealing with uncertainty in dependent parameters may be built in a similar way.

When dealing with dynamic tasks in the GIE approach, an expert may define GIE for several characteristic sections (time-steps) and then “propagate” GIE for other time values by interpolation. If an expert thinks that distributions on different time sections belong to the same distribution type, one can derive analytic expressions for the projections of joint distribution functions of GIE. Specifically, this can be done for various combinations of uniform and triangle distributions on α and *D*.

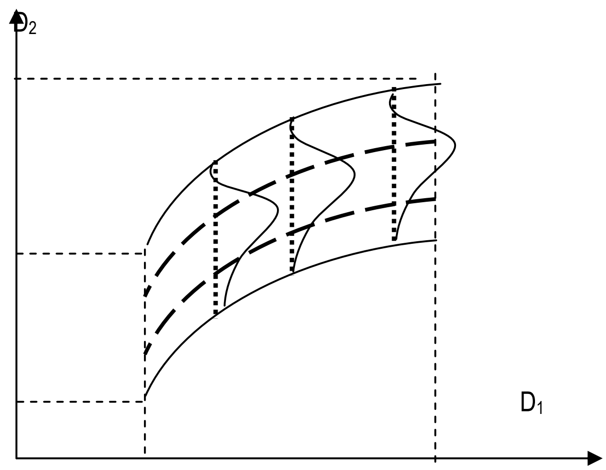


Fig. 7. GIE for parameter D_2 changing with time or parameter D_1

Internal Structure of GIE: Probability Tubes

This way is to reject averaging but analyze inner structure of GIE and use results of such analysis during decision-making (Chugunov et al., 2008). Let us fix some value Ds from the base interval. Since the expert defines probability distribution for each interval (or α-scenario) of GIE, we can try to analyze distribution of cumulative probability functions P(D<Ds|α) with respect to α. Marginal probability distribution function for random variable D is as follows:

$$P(D < Ds) = \int_0^1 \int_{D_{ld}}^{Ds} f_1(\alpha) f_2(D \mid \alpha) dD d\alpha = \int_0^1 f_1(\alpha) P(D < Ds \mid \alpha) d\alpha. \tag{6}$$

We wish to derive up and down bounds for P(D<DS):

$$P(D<Ds)=\int_0^1 f_1(\alpha)P(Ds\,|\,\alpha)d\alpha \le \tilde{P}_{\max}(Ds)\int_{\alpha_d}^{\alpha_u} f_1(\alpha)d\alpha + P_u(Ds)+P_d(Ds)=P_{\max}(Ds) \quad (7)$$

$$P(D<Ds)=\int_0^1 f_1(\alpha)P(Ds\,|\,\alpha)d\alpha \ge \tilde{P}_{\min}(Ds)\int_{\alpha_d}^{\alpha_u} f_1(\alpha)d\alpha + P_u(Ds)+P_d(Ds)=P_{\min}(Ds). \quad (8)$$

Here α_d and α_u correspond respectively to down and up intervals from GIE (figures 2, 3), which contain the fixed value Ds,

$$\tilde{P}_{\max}(Ds)=\max_{\alpha\in[\alpha_d;\alpha_u]} P(D<Ds\,|\,\alpha),\ \tilde{P}_{\min}(Ds)=\min_{\alpha\in[\alpha_d;\alpha_u]} P(D<Ds\,|\,\alpha),$$

$$P_d(Ds)=P(D<Ds\,|\,\alpha_d)\int_0^{\alpha_d} f_1(\alpha)d\alpha,\quad P_u(Ds)=P(D<Ds\,|\,\alpha_u)\int_{\alpha_u}^{1} f_1(\alpha)d\alpha.$$

Note, that functions P_{min}(Ds), P_{max}(Ds) are non decreasing functions defined on the base interval; they both always equal zero at the left endpoint of this interval, and equal one at its right endpoint. Hence, these functions are some probability distribution functions. Unknown "true" probability distribution function lies between these functions. So we may receive triple of probability distribution functions P_{min}(Ds), P_{max}(Ds) and P(D<Ds) on the base interval. Pair P_{min}(Ds), P_{max}(Ds) constitutes probability tube (p-box) [Dempster, 1967; Williamson and Downs, 1990], which represents uncertainty in assessment of "true" distribution (figure 8).

Let us now go back to distribution of P(D<Ds|α) values for some fixed Ds. Let Ds belong to each interval from the subset, described by the intervals corresponding to α_u and α_d. Then, if all values P(D<Ds|α) are unique, we can define probability distribution for values P(D<Ds|α) with probability density g(P(D<Ds|α)):

$$g(P(D<Ds\,|\,\alpha))=f_1(\alpha)/\int_{\alpha_d}^{\alpha_u} f_1(\alpha)d\alpha\ .$$

Probability tubes approach is widely used in real-life problem analysis under uncertainty. Note, that each boundary p-tube is not always correspond to a single possible scenario (e.g. optimistic or pessimistic), but could be specified by a combination of those, with each scenario playing its role on a certain subinterval of the base interval. Therefore, in some complex situations it might be difficult for an expert to specify the bounds for a whole base interval at once. GIE approach allows an expert to construct boundary p-tubes P_{min}(Ds) and P_{max}(Ds) automatically based on naturally driven reasoning about possible scenarios to describe values of estimated parameter and provides comprehensive description of the p-tubes obtained.

It should underline that GIE method permits for each value Ds of some estimated parameter D obtain on the base interval not only boundary tubes [P_{min}(Ds), P_{max}(Ds)], but also a probability distribution on section D = D_S of boundary tubes. In a similar way, for each fixed probability value P_s, one can obtain a distribution for values of estimated parameter D related to this fixed probability level. These two distributions represent in GIE formalism uncertainty in expert judgments on both estimated parameter values and corresponding probability distributions.

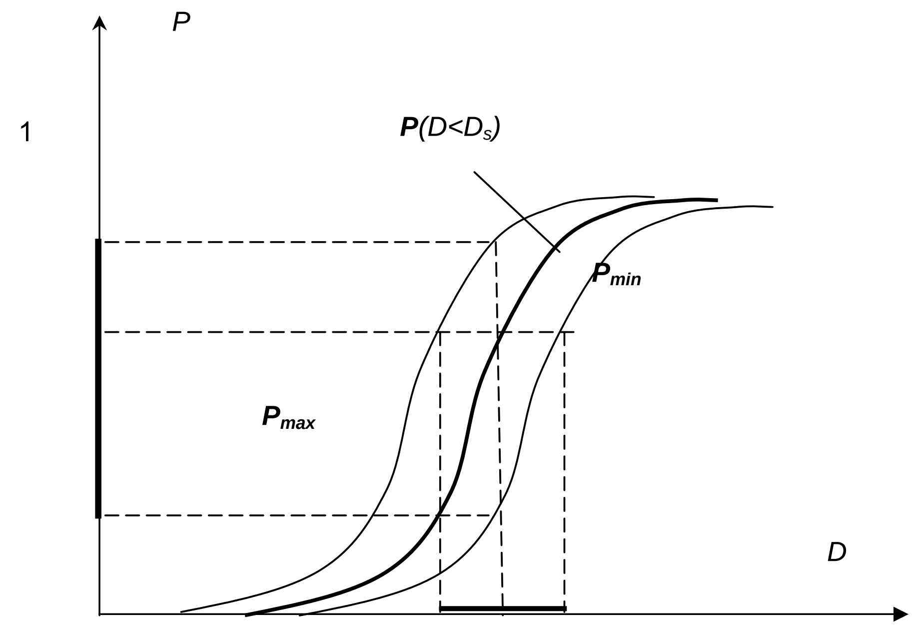


Fig. 8. Probability tubes generated by GIE.

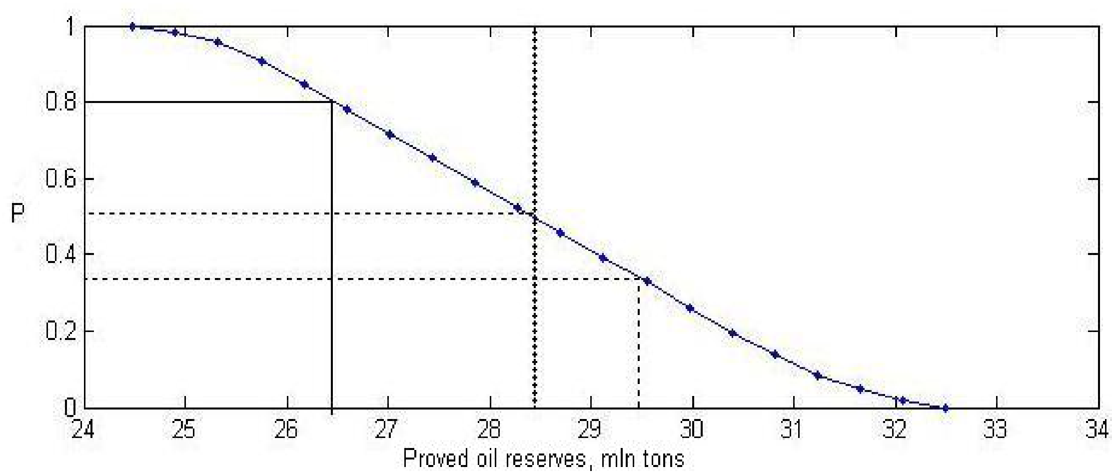


Fig. 9. Probability distribution for proved reserves

Let's again give an illustrative example of oil reserves estimation in the framework of probabilistic approach. Above we do so for mono-interval picture, there point estimations for all categories of oil reserves were received (see fig. 1). Now we present some results of calculations for volumetric model where input parameters have GIE estimations. One can see the results for proved reserves (probability level 0.9) on figure 9. Here values of proved reserves lay in the range from 24.5 to 32.5 ml tons. For the same input parameters mono-interval approach would give us 29.5 ml tons for this reserves category. From figure 9 one can see that this value corresponds to 0.35

confidence level on probabilistic curve, which may be unacceptably low for decision maker. Mean value is 28.4 ml tons. For acceptable confidence level 0.8 the value for proved reserves is 26.4 ml tons, that less than result for mono-interval approach. Similar analysis may be performed for probable and possible reserves categories.

Thus GIE method allow decision maker to obtain not only ranges, but, what is more important for justifiable decision making, cumulative probabilities for each reserves category.

GIE Method in Scenario Analysis

Scenario approach is rather often used in theory and practice of decision-making for the analysis of complex ill-structured problems. The methods of scenario analysis ensure the information-analytical support of the processes of decision making under the conditions of uncertainty. Advantages of scenario approach are accompanied now with a disadvantage due to the need for preliminary labor-consuming preparation of a set of scenarios that should be analyzed. Therefore, at present a finite number of scenarios are usually used for analysis that significantly limits a set of analyzed alternatives and the most rational ones may be exclude in advance. It means that the development of rapid methods for the elaborating alternative scenarios is important. An approach, which gives the set of scenarios by means of indicating its boundaries, may be promising approach. This can be done in the framework of the GIE method. Mentioned situation is a similar to situation of comparison of discrete optimization and a linear programming approach. At the first case choice of the “best” alternative takes place from alternatives prepared earlier.

Above it was noted that the intervals of PIE in the GIE method might be interpreted as the possible scenarios of the realization of the initial parameters or resultant index, and distribution on the Y-axis of PIE defined “the weights” of scenarios. However, for the systematic use of GIE method in the scenario analysis the method should be fitted out for the solution of problems with such dependent parameters where one of the parameters reflects the state of “external environment”, and the second, depending on the first, is the initial parameter or the resultant index of model. Let us note also that in contrast to the original GIE scheme where the Y-axis of PIE was the axis “of the marks” of intervals and had not identifiable “physical” sense, in scenario approach both axes of PIE had real meaning. Thus the expansion of GIE method to the tasks of the scenario analysis requires the modification of mathematical tools developed previously.

Table 1. Configurations of PIE for D ≠ 0, U ≠ 0 (trapezoidal shape), D = 0 или U = 0 (triangular shape)

Shapes of PIE	PIE of trapezoidal shape	PIE of triangular shape
1	$V_{ld} < V_{lu} < V_{rd} < V_{ru}$	$V_{ld} < V_{rd} < V_{u}$
2	$V_{ld} < V_{rd} < V_{lu} < V_{ru}$	$V_{ld} < V_{u} < V_{rd}$
3	$V_{ld} < V_{lu} < V_{ru} < V_{rd}$	$V_{u} < V_{ld} < V_{rd}$
4	$V_{lu} < V_{ld} < V_{rd} < V_{ru}$	$V_{lu} < V_{ru} < V_{d}$
5	$V_{lu} < V_{ld} < V_{ru} < V_{rd}$	$V_{lu} < V_{d} < V_{ru}$
6	$V_{lu} < V_{ru} < V_{ld} < V_{rd}$	$V_{d} < V_{lu} < V_{ru}$

Let us discuss the modifications in PIE shapes that need in scenario analysis. Note by V the values of the test parameters or indicators in problems with dependent variables and α - values of the external factors affecting the occurrence of possible values for V, $\alpha \in [\alpha_m, \alpha_M]$, where α_m (α_M) - the minimum (maximum) value of α respectively. Now PIE is set by the quartet V_{ld} (left-lower angel of PIE), V_{rd} (right lower angel), V_{lu} (left upper angel of PIE), V_{ru} (right upper angel), $D = V_{rd} - V_{ld}$, $U = V_{ru} - V_{lu}$. The relationships between members of the quartet determine the shape of PIE.

All possible shapes of PIE for the case $D \neq 0$, $U \neq 0$ (PIE of trapezoidal shape), and $D = 0$, $V_{ld} = V_{rd} = V_d$, or $U = 0$, $V_{lu} = V_{ru} = V_u$ (PIE of triangular shape) are presented in Table 1.

Let's pay attention to the fact that if the original pattern of PIE is often a set of nested intervals for the problems with dependent parameters is not the case.

Earlier we studied generalized uniform distribution of probabilities for PIE on nested intervals. In problems with dependent parameters variety of possible shapes of PIE leads to the appearance of a family of generalized uniform distributions. The corresponding relations are presented in Tables 2, 3.

Table 2. Generalized uniform probabilities distributions for trapezoidal PIE

Sectors of PIE	Densities for D≠U	Densities for D = U	Distributions: D≠U	Distributions: D=U
1.1. $V_{ld} \le V < V_{lu}$	$f(V) = I_1$	$f(V) = L_1$	$P(V < V_S) = F_1$	$P(V < V_S) = G_1$
1.2. $V_{lu} \le V \le V_{rd}$	$f(V) = I_2$	$f(V) = L_2$	$P(V < V_S) = F_2$	$P(V < V_S) = G_2$
1.3. $V_{rd} < V \le V_{ru}$	$f(V) = I_3$	$f(V) = L_3$	$P(V < V_S) = F_3$	$P(V < V_S) = G_3$
2.1. $V_{ld} \le V < V_{rd}$	$f(V) = I_1$	$f(V) = L_1$	$P(V < V_S) = F_1$	$P(V < V_S) = G_1$
2.2. $V_{rd} \le V \le V_{lu}$	$f(V) = I_4$	$f(V) = L_4$	$P(V < V_S) = F_4$	$P(V < V_S) = G_4$
2.3. $V_{lu} < V \le V_{ru}$	$f(V) = I_3$	$f(V) = L_3$	$P(V < V_S) = F_3$	$P(V < V_S) = G_3$
3.1. $V_{ld} \le V < V_{lu}$	$f(V) = I_1$	$f(V) = L_1$	$P(V < V_S) = F_1$	$P(V < V_S) = 0$
3.2. $V_{lu} \le V \le V_{ru}$	$f(V) = I_2$	$f(V) = L_2$	$P(V < V_S) = F_2$	$P(V < V_S) = G_5$
3.3. $V_{ru} < V \le V_{rd}$	$f(V) = I_5$	$f(V) = L_5$	$P(V < V_S) = F_5$	$P(V < V_S) = 0$
4.1. $V_{lu} \le V < V_{ld}$	$f(V) = I_6$	$f(V) = L_6$	$P(V < V_S) = F_6$	$P(V < V_S) = 0$
4.2. $V_{ld} \le V \le V_{rd}$	$f(V) = I_2$	$f(V) = L_2$	$P(V < V_S) = F_2$	$P(V < V_S) = G_5$
4.3. $V_{rd} < V \le V_{ru}$	$f(V) = I_3$	$f(V) = L_3$	$P(V < V_S) = F_3$	$P(V < V_S) = 0$
5.1. $V_{lu} \le V < V_{ld}$	$f(V) = I_6$	$f(V) = L_6$	$P(V < V_S) = F_6$	$P(V < V_S) = G_6$
5.2. $V_{ld} \le V \le V_{ru}$	$f(V) = I_2$	$f(V) = L_2$	$P(V < V_S) = F_2$	$P(V < V_S) = G_2$
5.3. $V_{ru} < V \le V_{rd}$	$f(V) = I_5$	$f(V) = L_5$	$P(V < V_S) = F_5$	$P(V < V_S) = G_7$
6.1. $V_{lu} \le V < V_{ru}$	$f(V) = I_6$	$f(V) = L_6$	$P(V < V_S) = F_6$	$P(V < V_S) = G_6$
6.2. $V_{ru} \le V \le V_{ld}$	$f(V) = -I_4$	$f(V) = -L_4$	$P(V < V_S) = F_7$	$P(V < V_S) = G_8$
6.3. $V_{ld} < V \le V_{rd}$	$f(V) = I_5$	$f(V) = L_5$	$P(V < V_S) = F_5$	$P(V < V_S) = G_7$

Here

$$I_1 = \frac{1}{U-D} Ln \frac{D(V_{lu}-V)+U(V-V_{ld})}{D(V_{lu}-V_{ld})}, \quad I_2 = \frac{1}{U-D} Ln \frac{U}{D}, \quad I_3 = \frac{1}{U-D} Ln \frac{U(V_{ru}-V_{rd})}{U(V-V_{rd})+D(V_{ru}-V)},$$

$$I_4 = \frac{1}{U-D} Ln \frac{[D(V_{lu}-V)+U(V-V_{ld})](V_{ru}-V_{rd})}{[D(V_{ru}-V)+U(V-V_{rd})](V_{lu}-V_{ld})}, \quad I_5 = \frac{1}{U-D} Ln \frac{D(V_{ru}-V)+U(V-V_{rd})}{D(V_{ru}-V_{rd})},$$

$$I_6 = \frac{1}{U-D} Ln \frac{U(V_{lu}-V_{ld})}{D(V_{lu}-V)+U(V-V_{ld})}.$$

$$L_1 = \frac{V-V_{ld}}{D(V_{lu}-V_{ld})}, \quad L_2 = \frac{1}{D}, \quad L_3 = \frac{V_{ru}-V}{D(V_{ru}-V_{rd})}, \quad L_4 = \frac{1}{V_{lu}-V_{ld}}, \quad L_5 = \frac{V-V_{rd}}{D(V_{ru}-V_{rd})}, \quad L_6 = \frac{V_{lu}-V}{D(V_{lu}-V_{ld})}.$$

$$F_1 = \frac{1}{D-U}\left[V_S - V_{ld} + \frac{D(V_{lu}-V_S)+U(V_S-V_{ld})}{D-U} Ln \frac{D(V_{lu}-V_S)+U(V_S-V_{ld})}{D(V_{lu}-V_{ld})}\right],$$

$$F_2 = \frac{1}{D-U}\left[V_{lu} - V_{ld} + \frac{D(V_{lu}-V_S)+U(V_S-V_{ld})}{D-U} Ln \frac{U}{D}\right],$$

$$F_3 = \frac{1}{D-U}\left[D + V_{lu} - V_S + \frac{D(V_S-V_{ru})+U(V_{rd}-V_S)}{D-U} Ln \frac{D(V_S-V_{ru})+U(V_{rd}-V_S)}{D(V_{rd}-V_{ru})}\right],$$

$$F_4 = \frac{1}{D-U}\left[D + \frac{D(V_{ru}-V_S)+U(V_S-V_{rd})}{D-U} Ln \frac{V_{ru}-V_{rd}}{V_{lu}-V_{ld}}\right],$$

$$F_5 = \frac{1}{D-U}\left[V_S - V_{ld} - U + \frac{D(V_{ru}-V_S)+U(V_S-V_{rd})}{D-U} Ln \frac{D(V_{ru}-V_S)+U(V_S-V_{rd})}{D(V_{ru}-V_{rd})}\right],$$

$$F_6 = \frac{1}{D-U}\left[V_{lu} - V_S + \frac{D(V_S-V_{lu})+U(V_{ld}-V_S)}{D-U} Ln \frac{D(V_S-V_{lu})+U(V_{ld}-V_S)}{U(V_{ld}-V_{lu})}\right],$$

$$F_7 = \frac{1}{D-U}\left[-U + \frac{D(V_S-V_{ru})+U(V_{rd}-V_S)}{D-U} Ln \frac{V_{rd}-V_{ru}}{V_{ld}-V_{lu}}\right],$$

$$G_1 = (V_S - V_{ld})^2 / [2D(V_{lu}-V_{ld})],$$
$$G_2 = (2V_S - V_{ld} - V_{lu})/(2D),$$
$$G_3 = 1 - (V_{ru}-V_S)^2 / [2D(V_{ru}-V_{rd})],$$
$$G_4 = (2V_S - V_{rd} - V_{ld})/[2(V_{lu}-V_{ld})],$$
$$G_5 = (V_S - V_{ld})/D,$$
$$G_6 = (V_S - V_{lu})^2 / [2D(V_{ld}-V_{lu})],$$
$$G_7 = 1 - (V_S - V_{rd})^2 / [2D(V_{rd}-V_{ru})],$$
$$G_8 = (2V_S - V_{ru} - V_{lu})/[2(V_{ld}-V_{lu})]$$

Table 3. Generalized uniform probabilities distributions for triangular PIE

Sectors of PIE	Densities	Distributions functions
1.1: $V_{ld} \le V < V_{rd}$	$f(V) = K_1$	$P(V < V_S) = H_1$
1.2: $V_{rd} \le V \le V_u$	$f(V) = K_2$	$P(V < V_S) = H_2$
2.1: $V_{ld} \le V < V_u$	$f(V) = K_1$	$P(V < V_S) = H_1$
2.2: $V_u < V \le V_{rd}$	$f(V) = K_3$	$P(V < V_S) = H_3$
3.1: $V_u \le V < V_{ld}$	$f(V) = K_2$	$P(V < V_S) = H_4$
3.2: $V_{ld} \le V \le V_{rd}$	$f(V) = K_3$	$P(V < V_S) = H_3$
4.1: $V_{lu} \le V < V_{ru}$	$f(V) = K_4$	$P(V < V_S) = H_5$
4.2: $V_{ru} \le V \le V_d$	$f(V) = K_5$	$P(V < V_S) = H_6$
5.1: $V_{lu} \le V < V_d$	$f(V) = K_4$	$P(V < V_S) = H_5$
5.2: $V_d < V \le V_{ru}$	$f(V) = K_6$	$P(V < V_S) = H_7$
6.1: $V_d \le V < V_{lu}$	$f(V) = K_5$	$P(V < V_S) = H_8$
6.2: $V_{lu} \le V \le V_{ru}$	$f(V) = K_6$	$P(V < V_S) = H_5$

Here

$$K_1 = \frac{1}{D} Ln \frac{V_u - V_{ld}}{V_u - V},\ K_2 = \frac{1}{D} Ln \frac{V_u - V_{rd}}{V_u - V_{ld}},\ K_3 = \frac{1}{D} Ln \frac{V_u - V_{rd}}{V_u - V},\ K_4 = \frac{1}{U} Ln \frac{V_{lu} - V_d}{V - V_d},$$

$$K_5 = \frac{1}{U} Ln \frac{V_{lu} - V_d}{V_{ru} - V_d},\ K_6 = \frac{1}{U} Ln \frac{V_{ru} - V_d}{V - V_d},\ K_7 = \frac{1}{U} Ln \frac{V_{lu} - V_d}{V_{ru} - V_d},$$

$$H_1 = \frac{1}{D}[V - V_{ld} + (V - V_u) Ln \frac{V_u - V_{ld}}{V_u - V}],\ H_2 = 1 - \frac{V_u - V}{D} Ln \frac{V_u - V_{ld}}{V_u - V_{rd}},\ H_4 = \frac{V - V_u}{D} Ln \frac{V_{rd} - V_u}{V_{ld} - V_u},$$

$$H_3 = \frac{1}{D}[V - V_{ld} + (V_u - V) Ln \frac{V - V_u}{V_{rd} - V_u}],\ H_5 = \frac{1}{U}[V - V_{lu} + (V_d - V) Ln \frac{V_d - V}{V_d - V_{lu}}],$$

$$H_6 = \frac{1}{U}[U + (V_d - V) Ln \frac{V_d - V_{ru}}{V_d - V_{lu}}],\ H_7 = \frac{1}{U}[V - V_{lu} + (V_d - V) Ln \frac{V_d - V}{V_d - V_{ru}}],$$

$$H_8 = \frac{V - V_d}{U} Ln \frac{V_{ru} - V_d}{V_{lu} - V_d},$$

New possibilities that appear in the tasks with the dependent variables due to “the equality of rights” of PIE axes may use in two directions. To more completely consider possible uncertainty of the estimation “of external factors” (e. g. forecast of prices in the above mentioned task about recoverable reserves of oil) one may build on the Y-axis (axis of prices in the task about the reserves) additional GIE besides initial one. Furthermore, in some tasks of the type of the task concerning in the dependence of the volumes of recoverable reserves on the price of the hydrocarbons, may be useful constructing for each level of the prices not mono-interval, but GIE that leads to the appearance “multidimensional” GIE.

Probabilistic Forecast for World Oil Production: Illustrative Example

Here we will give an illustrative example to demonstrate application of GIE method, and, specifically, generalized uniform distributions, in scenario approach to analyze probabilistic forecast for world oil production. There is a wide spectrum of estimations for world oil production and oil reserves base ranging from pessimistic ones that predicts the end of oil era in several years, up to those optimistic that argues there are sufficient oil reserves, at least, up to the end of this century.

The pessimistic scenario arises from the theory of “peak of oil” which is quite popular among many oil analysts. According to this theory consumption of oil will overtake, and then will outstrip rates of discovering new reserves whereas already known reserves will start to be exhausted.

In contrast to a widely discussed theory that world oil production will soon reach a peak and go into sharp decline, a new analysis of the subject by Cambridge Energy Research Associates (CERA) finds that the remaining global oil resource base is actually 3.74 trillion barrels -- three times as large as the 1.2 trillion barrels estimated by the theory’s proponents -- and that the “peak oil” argument is based on faulty analysis [CERA, 2006]. According to the study, the global production profile will not be a simple logistic or bell curve postulated by geologist M. King Hubbert, but it will be asymmetrical – with the slope of decline more gradual and not mirroring the rapid rate of increase - and strongly skewed past the geometric peak. It will be an undulating plateau that may well last for decades [CERA, 2006].

There are two important categories of oil reserves used in these studies: conventional and unconventional oil. Conventional oil refers to a reserves category that has supplied most to date and will likely to dominate oil supply far into the future. These reserves are relatively easy, cheap and fast to produce. Unconventional oil includes heavy oil category, dense and extremely viscous oils, as well as those produced from coal and immature source rocks. This category is characterized by a high resource base but a low extraction rate and net energy yield.

Figure 10 shows CERA’s outlook for conventional and unconventional oil including several possible scenarios and peak oil theory.

Here we are not interested in the accuracy and reliability of these forecasts; we are more interested in providing consistent mathematical tools for an independent expert to develop an individual prediction based on those from the study. These tools should allow the expert to express his/her preferences that would be reflected in the final scenario for the future dynamics of world oil reserves. We will use GIE method as such tools. The GIE framework will allow the expert to include the full spectrum of possible scenarios into the final estimation besides initial ones. The existence of those “intermediate” scenarios could be attributed to the uncertainties of the starting predictions as well as to limited number of models or experts involved in generation of those initial scenarios.

To use GIE method an expert should construct PIEs for each time points of forecasts. According to figure 10, there are three initial forecasts – “Peak Oil”, “Conventional Oil”, “Unconventional Oil”, the forecasts sharply differ after year 2010. Therefore we quantified data from the figure to use them in a subsequent scenario analysis based on GIE, starting from 2010. At the preliminary stage of constructing GIE the expert can define at an every time point (year) the base interval being determined by the corresponding values of the most optimistic scenario (CERA, unconventional oil) - right point of the base interval, and the pessimistic (peak oil) scenario – left point of the base interval. The mini interval could be defined by introducing a 10% spread (± 5%) for the most realistic, from expert’s point of view, forecast (CERA, conventional oil) (figure 11). Naturally we might use PIE of triangle shape but the first option gives more diversity of initial scenarios.

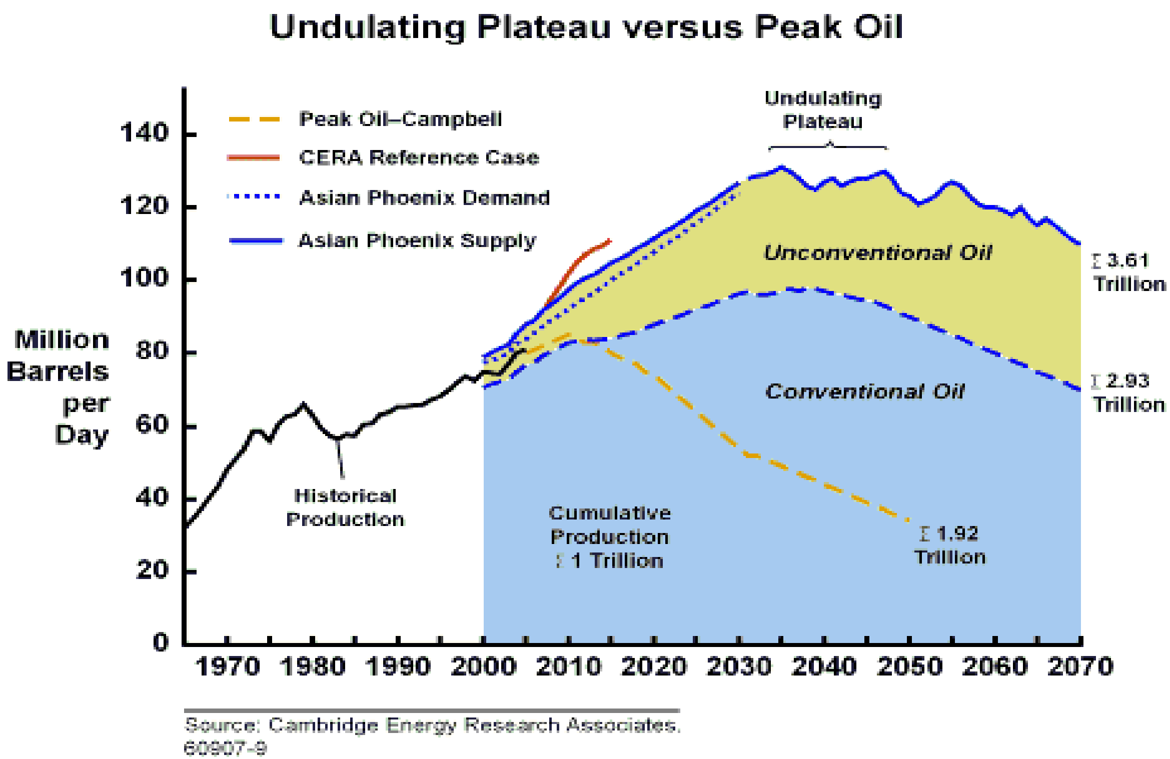


Fig. 10. Wave plateau versus peak theory

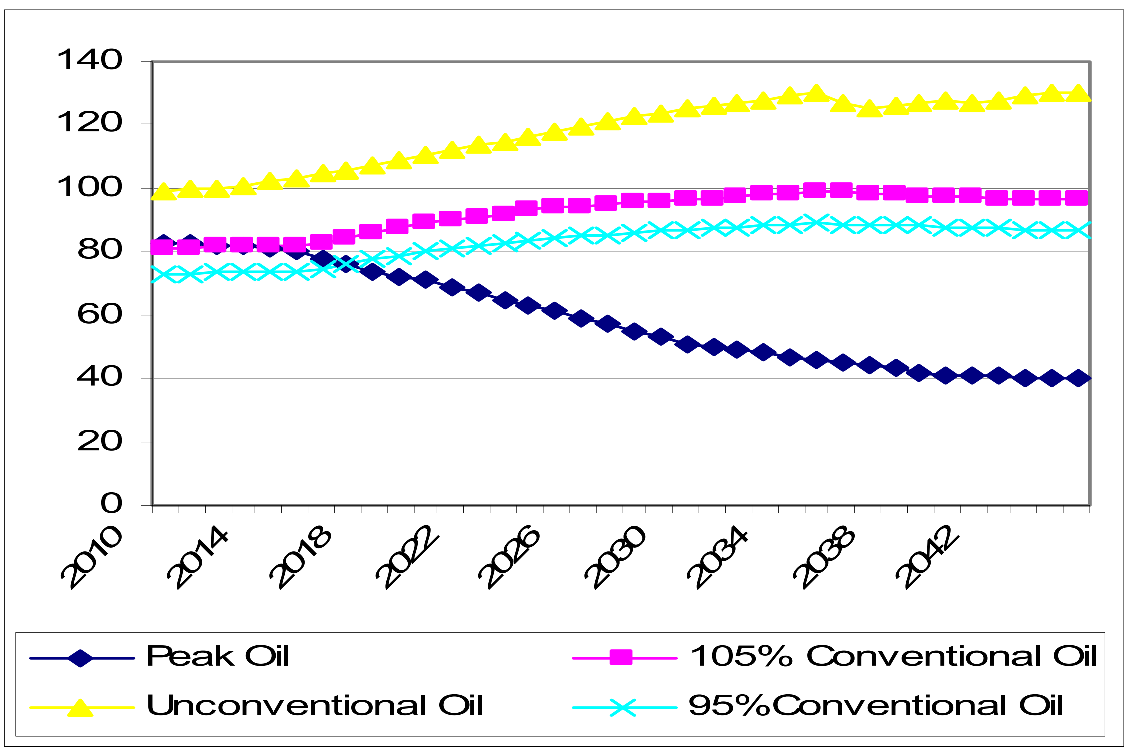
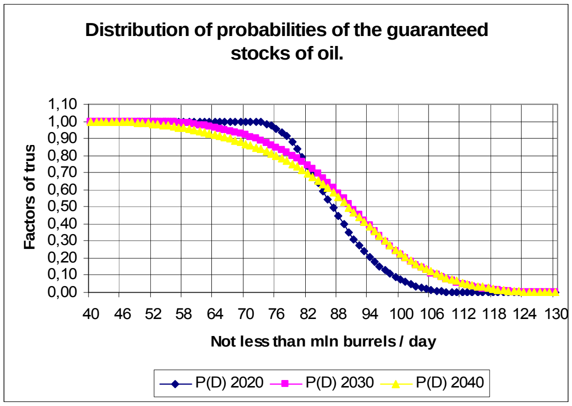


Figure 11. Expert's perspective on Peak or Plato

So for example at the point t = 2020 we have that the base interval is [71; 110.5] while the mini interval is [84.55; 93.45]. Once the PIE structure is defined, the expert can express his/her beliefs by specifying probability density for distributions on the base and mini intervals for a current point t. Let us consider the case of rapid analysis when expert uses generalized uniform distribution as projection distribution.



Confidence Level	Production not less than (Mln barrels per day)		
	2020	2030	2040
0.9	78	72	67
0.8	81	79	76
0.7	83	83	82
0.6	85	87.5	87
0.5	87	90.5	90
0.4	89	93.5	93.5
0.3	91.5	97	97
0.2	94	101	102
0.1	99	107	108

Figure 12. Distributions of probabilities for the guaranteed levels of oil production

Figure 12 shows curves of probabilities of the guaranteed values for world oil production for three time points - the years 2020, 2030, and 2040. Analysis of these curves shows that for a confidence level of 0.9 the estimation of oil production is at least 80 million barrels per day for the year 2020, at least 73 million barrels per day for 2030, and 67 million barrels per day for the year 2040. Estimation of the world oil production at the confidence level of 0.5 gives at least 87 million barrels per day for the year 2020, at least 90.5 million barrels per day for 2030, and 90 million barrels per day for the year 2040. It is interesting to note that for the confidence level around 0.7 the guaranteed value of the estimation for the world oil production is almost constant for years 2020-2040 and does not go bellow 82 - 83 million barrels of oil per day.

Conclusion

Adequate solutions of many complex real-life problems under uncertainty significantly depend on effective elicitation and formalization of expert knowledge on input parameters of used models. This requires specific methods for expert knowledge elicitation, interpretation and processing.

Poly-interval representation and, specifically, Generalized Interval Estimations method seems to be promising approach to elicit and formalize expert knowledge on quantitative parameters in problems under uncertainty. This survey summarizes and further extends theoretical and practical aspects of GIE approach.

Generalized Interval Estimation structure, being scenario-based expert knowledge representation, is analysis-oriented, since it provides flexibility to add/remove/modify scenarios from expert description of estimated quantity. The practical example on projecting the dynamics of the world oil production illustrates possibilities that an independent expert could use to develop and analyze new scenarios based on initial set of forecasts available from relevant studies and research.

Generalized Probability Tubes structure is more decision-oriented interpretation of expert knowledge, since for given decision makers' confidence level it is able to represent the probability for estimated quantity to have a certain value. Oil reservoir evaluation example demonstrates this property of GIE approach, which, we believe, could be important for a wide variety of interdisciplinary problems including investment project evaluation, technology assessment, development of policies and regulations.

We hope that GIE method may be applied also for analysis of other problems. Among them aggregation information concerning in input data of used models, comparison of PIEs and GIEs and others.

Acknowledgments

This work is partially supported by the Russian Academy of Science Research Programs "Basic Problems of Informatics and Information Technologies", "Basic Foundation of Information Technologies and Systems", the Russian Foundation for Basic Research (projects 07-01-00515, 08-01-00247, 09-07-00009).

Bibliography

[CERA, 2006] URL: http://www.cera.com/aspx/cda/public1/news/pressReleases/pressRelease Details.aspx? CID=8444 «Peak Oil Theory – "World Running Out of Oil Soon" – Is Faulty; Could Distort Policy & Energy Debate». 14/11/ 2006.

[Chugunov et al., 2006] N. Chugunov, G. Shepelyov and M. Sternin. A method for uncertainty quantification in expert decision support systems, in F. Adam, P. Brezillon, S. Carlsson and P. Humphreys (Eds.) Creativity and Innovation in Decision Making and Decision Support, London: Decision Support Press, 2006, Vol. 2, pp.851–865.

[Chugunov et al., 2008 A] N. Chugunov, G. Shepelyov and M. Sternin. Probabilistic Methods for Uncertainty Quantification. In: Encyclopedia of Decision Making and Decision Support Technologies. Edited By: Frederic Adam, University College Cork, Ireland; Patrick Humphreys, London School of Economics, UK. Publisher: Information Science Reference, - April 2008, Vol. 1, pp. 732 – 742.

[Chugunov et al., 2008 B] N. Chugunov, G. Shepelyov and M. Sternin. The generalized interval estimations in decision making under uncertainty. International Journal of Technology, Policy and Management, 2008, Vol.8, pp.298–321.

[Dempster, 1967] A.P. Dempster. Upper and Lower Probabilities Induced by a Multivalued Mapping. // The Annals of Statistics, 1967, Vol. 38, pp.325-339.

[Kahneman et al., 1982] D. Kahneman, P. Slovic, A. Tversky. Judgment under uncertainty: heuristics and biases. Cambridge: Cambridge university press. 1982.

[Morgan and Henrion, 1992] G.M. Morgan, M. Henrion. Uncertainty – A guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, New York. 1992.

[Shepelyov and Sternin, 2003] G. Shepelyov and M. Sternin. Method of generalized interval estimations for intelligent DSS // "DSS in the uncertainty of the internet age", Proceedings of the Conference, Katowice, University of Economics, 2003, pp.367–377.

[Shepelyov and Sternin, 2008] G. Shepelyov and M. Sternin. Обобщенные интервальные оценки в сценарном анализе. // Supplement to International Journal Information Technologies and Knowledge. – 2008 – Vol 2 – pp. 87-94. (in Russian).

[Welsh et al., 2004] M.B. Welsh, S.H. Begg, R.B. Bratvold and M.D. Lee. Problems with the Elicitation of Uncertainty. / SPE 90338 paper, SPE Annual Technical Conference, Houston, Texas: Society of Petroleum Engineers. 2004.

[Williamson and Downs, 1990] R.C. Williamson, T. Downs. Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. // International Journal of Approximate Reasoning, 1990, Vol. 4, pp.89–158.

Authors' Information

Gennady Shepelyov – head of the laboratory 'Knowledge-based Computer Systems' of the Institute for Systems Analysis of the Russian Academy of Sciences, 60-letiya Oktyabrya Prosp., 9, Moscow-117312, Russia; e-mail: gis@isa.ru.

Major Fields of Scientific Research: system analysis, mathematical modeling and methods of decision making.

Michael Sternin – senior scientific researcher of the laboratory 'Knowledge-based Computer Systems' of the Institute for Systems Analysis of the Russian Academy of Sciences, 60-letiya Oktyabrya Prosp., 9, Moscow-117312, Russia; e-mail: mister@isa.ru.

Major Fields of Scientific Research: mathematical modeling, decision support systems and knowledge-based systems.

WIRELESS COMMUNICATION AND CONTROL SYSTEM FOR PORTABLE MICRO-ELECTROMECHANICAL DEVICE FOR REAL-TIME BLOOD SAMPLING AND GLUCOSE ANALYSIS (ELECTRONIC MOSQUITO)

Tristan D. Jones, Karan Kaler, Michel Fattouche, Martin P. Mintchev

Abstract: *This study is a continuation of the development of the "Electronic Mosquito," a minimally-invasive blood sampling and analysis device. Wireless communication is a pivotal feature of the Electronic Mosquito in terms of usability and future applicability. The present article describes the design process used to construct a working wireless communication link using an established and reliable protocol. The resulting wireless link performed to specifications, with an empirically determined range of about 20 feet. Power consumption analysis indicated that the wireless system could operate for 100 days off of a CR2032 battery without any power control, and likely longer if power management is implemented in the software. The adopted communication protocol also appears to be a likely candidate for the creation of a Personal Area Medical Network standard which could greatly benefit modern medicine as well as the Electronic Mosquito.*

Keywords: *Electronic Mosquito, Diabetes, Non-Invasive Monitoring, Control Systems, Wireless, ANT*

ACM Keywords: *Medical Information Systems*

List of Abbreviations

ADC	*Analog to Digital Converter*
BG	*Blood Glucose*
GOx	*Glucose Oxidase*
JTAG	*Joint Test Action Group port*
MEMS	*Micro-electromechanical system*
PAMN	*Personal Area Medical Network*
RX	*Receive*
SBGM	*Self Blood Glucose Monitoring*
TX	*Transmit*

1. Introduction

1.1. Diabetes Mellitus

Diabetes Mellitus is an affliction of pandemic proportions, affecting an estimated 171 million people worldwide in the year 2000. Furthermore, that number is expected to rise to 366 million by the year 2030 [1]. Taking into account historical and projected world population dynamics yields an increase in prevalence from 2.8% to 4.4% by the year 2030 [2], clearly making diabetes an important and immediate health concern.

Diabetes is characterized by abnormal blood glucose levels resulting from either insulin deficiency (Type 1 diabetes) or insulin insensitivity (Type 2 diabetes) [3]. Both conditions commonly result in dangerously high blood glucose levels and this seems to be the cause of the various diabetic side effects such as retinopathy leading to blindness, various neuropathies that can lead to limb amputation, nephropathy (kidney damage), and eventually death [4,5].

Type 1 diabetes is often referred to as juvenile diabetes as it commonly arises in infants and is generally thought to be largely a genetic condition with a minor environmental component. In this form of the disease, the β-cells in the pancreas that produce insulin are destroyed via an autoimmune process [6, 7]. Without insulin the human body cannot absorb glucose and as a result it builds up to toxic levels in the blood stream [3]. Type 1 diabetics must monitor their blood glucose levels in order to administer insulin and keep these concentrations at non-toxic levels.

Sufferers of Type 2 diabetes typically develop the disease later in life and as a result it is known colloquially as adult-onset diabetes. While the pancreatic β-cells of sufferers are functioning correctly, the cells in the body develop a resistance to the insulin hormone with the result that glucose once again can accumulate to toxic levels in the blood [7, 8]. While causation has not yet been shown, there is a strong correlation between obesity (especially childhood adiposity) and development of type 2 diabetes later in life [9, 10, 11].

1.2. Current Treatment

Current treatment for diabetes consists of increased exercise and dietary control, as well as self blood glucose monitoring (SBGM) [3, 12]. Current leading-edge treatment/management regimes for diabetes typically consist of two approaches: implanted, continuous glucose monitoring devices [13, 14] and non-invasive monitoring devices typically using reverse iontophoresis [13, 15]. Unfortunately all of these methods suffer from several severe drawbacks which are discussed below.

1.3. “Finger-poking” Self-Monitoring Method

The “finger-poking” method of SBGM is presently the most common technology used by diabetics to monitor their condition. It consists of a spring-loaded needling device and a reusable digital sensor for glucose. The needle is used to create a small wound, typically on the fingertip, from which a small blood sample can be harvested and placed in the glucose meter. The glucose meter then uses a chemical reaction involving Glucose Oxidase (GOx) to generate and display the concentration in units of mmol/L [13]. Two major brands of this type of system are OneTouch (Johnson & Johnson, New Brunswick, NJ, USA) and Accu-Chek (Roche Diagnostics, Laval, QC, Canada).

Unfortunately this technique suffers from the fairly major drawback that it is invasive and painful for patients, which results in significant problems with patient compliance [16]. Furthermore, finger-poking only produces 2-8 glucose measurements a day, which is quite low and limits the information available about the patient’s glucose dynamics. It has been shown in several studies that increasing the rate of sampling improves glycemic control and hence slows progression of the disease and its symptoms [17,18].

1.4. Non-Invasive Glucose Sensors

The only non-invasive glucose-sensing technology to have gained market approval in the United States has been the Gluco-Watch (Cygnus Inc, Redwood City, CA, USA). This device uses reverse-iontophoresis, the application of electrical fields to the skin, to draw glucose out of the interstitial fluid to the electrode where its concentration can be measured.

Reverse-iontophoretic glucose sensing methods draw glucose from the interstitial fluid, which lags behind the glucose concentration in the blood by 18-20 minutes [15]. Furthermore, the ions are not drawn to the electrode at the same concentration as they exist in the interstitial fluid. As a result of these two properties, reverse-iontophoretic devices (including the gluco-watch) need to be regularly calibrated against a known blood glucose concentration, normally obtained by the finger-poking, SBGM technique. Therefore, this painful procedure is still present, albeit slightly less frequently [15, 19].

1.5. Implantable Glucose Sensors

The most popular commercially available implantable continuous glucose monitoring device is Medtronic's MiniMed Paradigm (Medtronic, Minneapolis, MN, USA) that uses a 3-day subcutaneously implanted tube connected to an external sensor to monitor BG levels. Although offering the capability of "continuous glucose monitoring," these solutions cannot avoid problems with calibration. Once again the sampled fluid is the interstitial fluid and thus the system suffers from the same time delays and non-linearities as reverse iontophoretic sensors.

2. The Electronic Mosquito

The Electronic Mosquito (e-Mosquito) introduces several novel technologies and concepts to overcome all of the shortcomings noted in the approaches above [20].

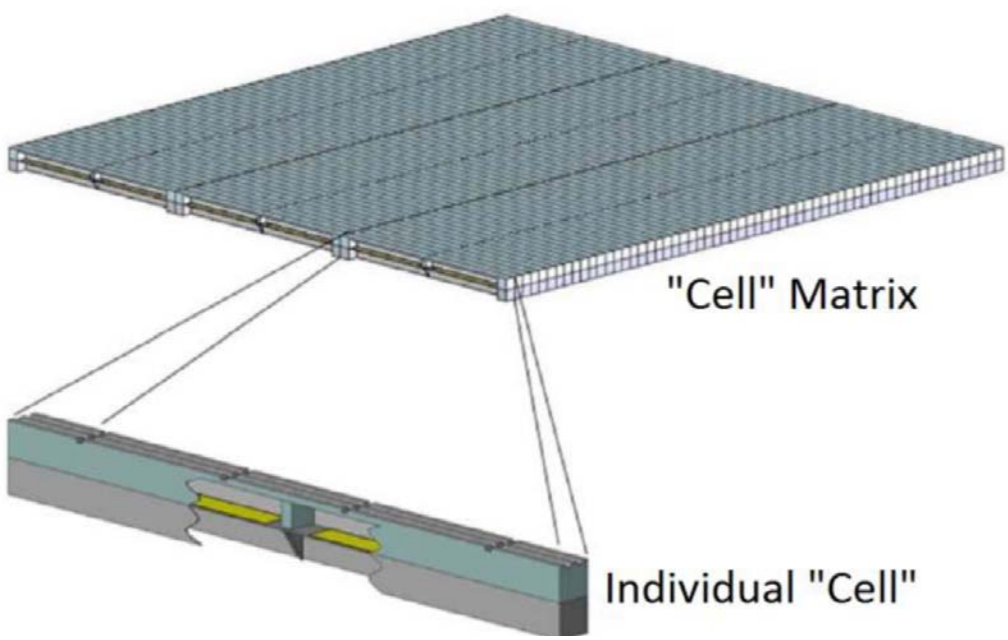


Figure 1. Electronic mosquito "cell" matrix structure showing a closeup of an individual cell as well as the entire matrix of repeated elements [20].

The vision for the device consists of a Band-Aid-like patch (roughly 3" x 3") with a large number (>100) micro-scale "cells" built with MEMS technology (Figure 1). Each "cell" contains of a MEMS microneedle, piezo-electric actuators, and a glucose sensor (Figure 2). Each "cell" is single use only, as the sensor in each "cell" will not be cleaned after use and therefore cannot be reused without contamination of the following sample. The device can then be programmed to use a cell to sample blood glucose levels at an arbitrary interval, for instance every 5 minutes or every 10 minutes. While this is not entirely continuous, the rate of change of glucose levels in the blood are slow enough that sample rates in this range should be essentially lossless. Moreover, with further miniaturization of this technology, the sampling interval can be reduced to well below a minute, if need be.

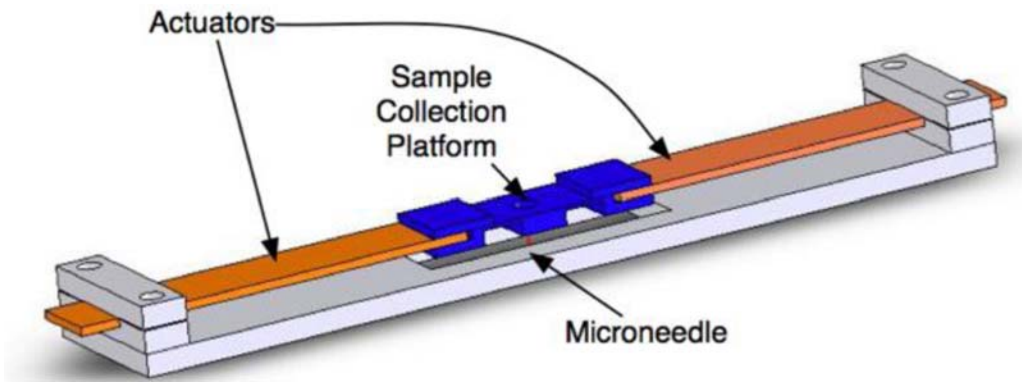


Figure 2. A single "cell" of the e-Mosquito, showing the actuators, microneedle, and sample collection platform where the actual glucose sensor is located [20].

The electrodes in the glucose sensor allow the e-Mosquito to detect the presence of a blood sample. This capability is then coupled to a control system which controls the deflection of the needle actuators. The control system descends the needle by a fixed step amount, and then waits to ensure that no blood is forthcoming. The needle then descends by another step and again waits. This process repeats until blood is detected on the electrodes at which point the needle withdraws (Figure 3). Due to the fact that capillaries lie above nerve endings in many locations in the body (Figure 4), the sampling process should be almost entirely painless.

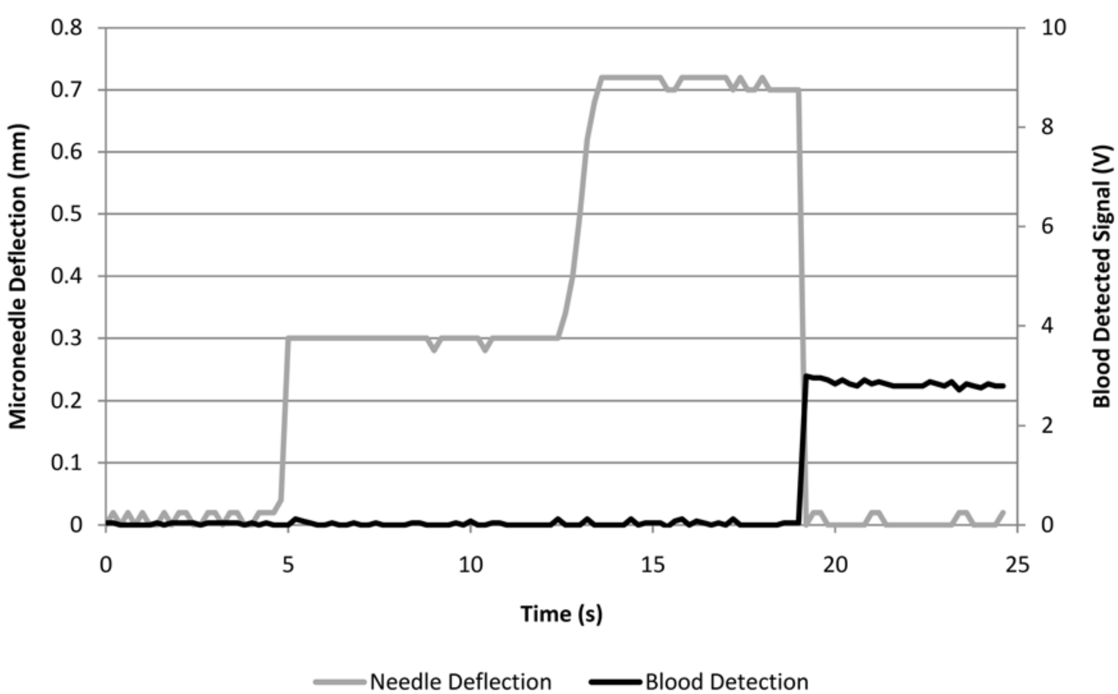


Figure 3. Example of controlled descent of the e-Mosquito needle in an ex-vitro setting. The voltage is proportional to needle deflection downwards. The small step up signal at lower right indicates blood detection on the electrodes [21].

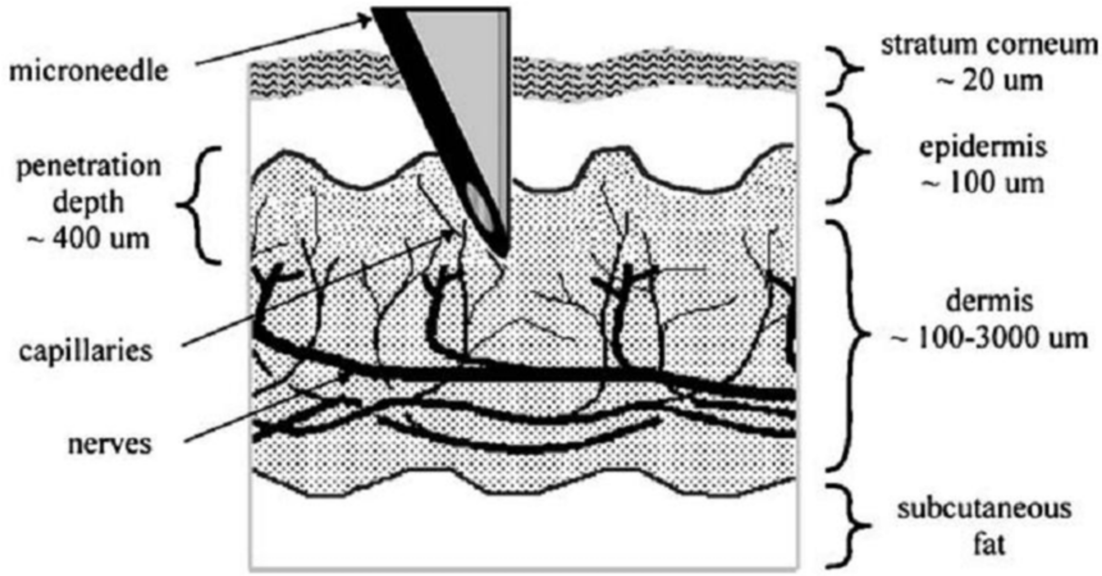


Figure 4. Profile of normal skin at various depths, showing capillary extension above nerve endings. A MEMS microneedle is also shown [21].

The e-Mosquito will permit both high temporal resolution glucose level tracking and in the long-term, a complete artificial pancreas that would allow diabetics to lead normal lives.

2.1. Wireless Communication and Control

In order for the e-Mosquito patch to be convenient for patients to us, it should not require plugging in, and should not have any loose wires to constrict movement or tangle. Furthermore, if the patient were required to be physically tethered to a computer this would severely restrict the patient's freedom of movement.

Wireless data technology enables all of the goals of the e-Mosquito project. Radio-frequency connection to a device such as a health watch can provide near real-time display and recording of data. Connection to an insulin infusion pump will facilitate the creation of a complete artificial pancreas that is almost completely unobtrusive to the patient.

Based on the requirements for the device and the realities of day-to-day living, a bi-directional wireless data link becomes of pivotal importance to the e-Mosquito.

3. Aim of the Study

The primary goal of this study is the development, implementation, and testing of a wireless link for the e-Mosquito. Furthermore, the evaluation of wireless link topographies, security measures, and protocol requirements for the development of a modular personal area medical network (PAMN) is examined and a set of design constraints for a second generation wireless link is proposed in this context.

4. Methods

4.1. Design Methods

The design of the wireless interface was accomplished by using the classical design method. The system constraints were first enumerated and then a design was developed that met or attempted to meet each constraint. The first generation wireless interface outlined here was designed to provide control signals to the e-Mosquito device and to receive data and control responses from the device. At the current stage of the development of the e-Mosquito prototype, the control signals and responses relate primarily to the control of needle penetration and the detection of blood in the sample volume. However, with the addition of a functional glucose sensor, the wireless link is intended to handle glucose data transmission as well.

While the final e-Mosquito would be programmed initially to sample at a specific interval, the current prototype receives its control signals from a connected computer. This computer connects to the e-Mosquito device and receives a configuration message detailing the various analog-digital converter parameters, as well as the current state of every “cell” (blood detected or not and current deflection). The computer is then able to cause sequential “cells” in the e-Mosquito to begin blood sampling by sending a start command (Figure 5).

The list of first generation design constraints is found in Table 1.

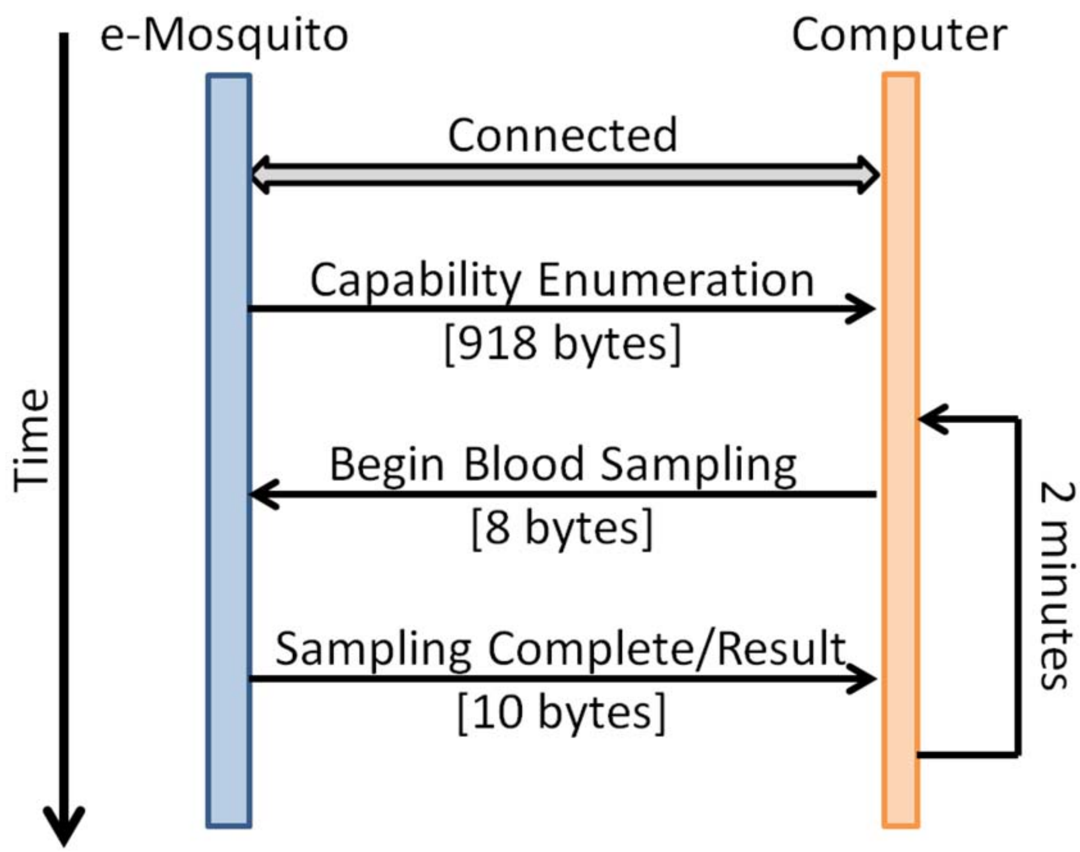


Figure 5. Communication protocol for a computer controlling the e-Mosquito. This case assumes 300 "cells" in the e-Mosquito and a sample interval of 2 minutes.

Table 1. List of design constraints for the first-generation e-Mosquito wireless link.

Constraint	Description
1.1	***Bidirectionality:*** The wireless link must be able to pass both control signals and data in both directions.
1.2	***Low Power:*** The e-Mosquito is designed to use CR-2032 and must operate continuously for a period of 2 days.
1.3	***Transmission Rate:*** The minimum transmission rate required can be calculated from #FIGURE#: $$v_{min} = \frac{(10 \text{ bytes}) + (8 \text{ bytes})}{2 \text{ minutes}} = 0.15 \text{ bps}$$ Of course this rate would cause the initial enumeration to take $0.15 \text{ bytes/s} \times 918 \text{ bytes} = 2.30 \text{ minutes}$ to transmit, which is tolerable, but not desirable. In the interest of maintaining future expandability of the wireless link, it was decided that the bandwidth of the link should far exceed current requirements. A factor of 20,000 times would allow a large number of devices to coexist on the same network simultaneously, as well as provide ample room for expansions to the protocol later. This results in a desired transmission rate of 3kbps.
1.4	***Transmission Fidelity:*** The range of the wireless link should be at least 20 feet so that the e-Mosquito is able to communicate with devices in the same room as the user. Within this range, 99% of packets should be successfully received.
1.5	***PC Connection:*** For the purposes of developing the e-Mosquito, and for end users to archive or transmit their BG data, the wireless link should be able to interface with a standard windows PC.

These design constraints were then used to guide the design process and arrive at a final first-generation link hardware implementation.

4.2. Testing Methods

The prototype wireless communication system underwent testing in a variety of ways. The primary interest was in wireless range and fidelity. In order to quantify this, the wireless-link was tested in two generic hallways as well as outdoors. The number of packets that were successfully received vs. the number of packets that were lost during transmission was taken to be an easy way to quantify signal strength. A laptop was used to collect the transmitted data while the e-Mosquito prototype with an attached wireless chip was used to generate data for transmission. The e-Mosquito prototype was powered by a +5V power supply provided through the JTAG programming interface.

Furthermore, current consumption of the wireless chip was measured by placing an ammeter on the positive supply to the wireless interface. The current consumption was plotted against time over a short period in order to quantify both the idle current draw and the transmitting current-draw. From this an average in-use current was calculated and compared to the capacity of a standard CR2032 battery.

Finally, the wireless link was used to operate the e-Mosquito prototype permanently, constituting an *in situ* test of the system.

5. Results

5.1. Design Results

The design of the wireless system was performed directly from the constraints found in Table 1. The nRF24AP1 wireless chip from Nordic Semiconductor implementing the ANT digital wireless protocol developed by Dynastream (Cochrane, AB, Canada) was chosen based on constraints 1.1 – 1.4, while the high-level architecture was designed based on constraint 1.5.

The nRF24AP1 chip is a transceiver chip, meaning that it incorporates both an RF transmitter circuit (TX) and an RF receiver circuit (RX). This allows it to both receive and transmit messages, fulfilling constrain 1.1 that the wireless link should be bidirectional [21].

The nRF24AP1 has an average current draw of 30 μA [21]. A typical CR2032 button cell battery has a capacity of 220 mAh [19] and assuming ideal characteristics, could provide an average current of $220\ \mathrm{mAh}/48\ \mathrm{h} = 4.58\ \mathrm{mA}$. Thus the wireless link will be consuming on average only 0.655% of the total current available to the e-Mosquito. Conversely, a CR2032 could operate only the wireless link for a period of about 305 days. The current consumption of the nRF24AP1 chip was considered to be small enough that it would not have a significant effect on the performance of the device, satisfying constraint 1.2.

The maximum sustained data rate of the nRF24AP1 is 20 kbps [21], which exceeds the requirement of 3 kbps by a factor of 6.7 times (satisfying constraint 1.3). This also provides significant headroom for future enhancements of the protocol to allow more data to be transmitted.

Unfortunately, the nRF24AP1 documentation does not give any sample range information to allow easy validation of this constraint. However, the sample devices provided with the development kit were tested and demonstrated ranges far in excess of 20 feet with no noticeably missed packets. This allowed constraint 1.4 to be satisfied, with the stipulation that the actual range of the device should be tested and validated later.

The overall architecture was determined by the need for the e-Mosquito to communicate with a computer controller (constrain 1.5). USB was chosen as the interface method on the computer end, as a wide variety of USB interface chips (such as those made by FTDI) and software libraries are available. In order to interface with the wireless chip, a microprocessor (the Texas Instruments MSP430F5438) was used to translate the wireless messages received by the ANT into messages suitable for transmission over the USB port to the computer (Figure 6).

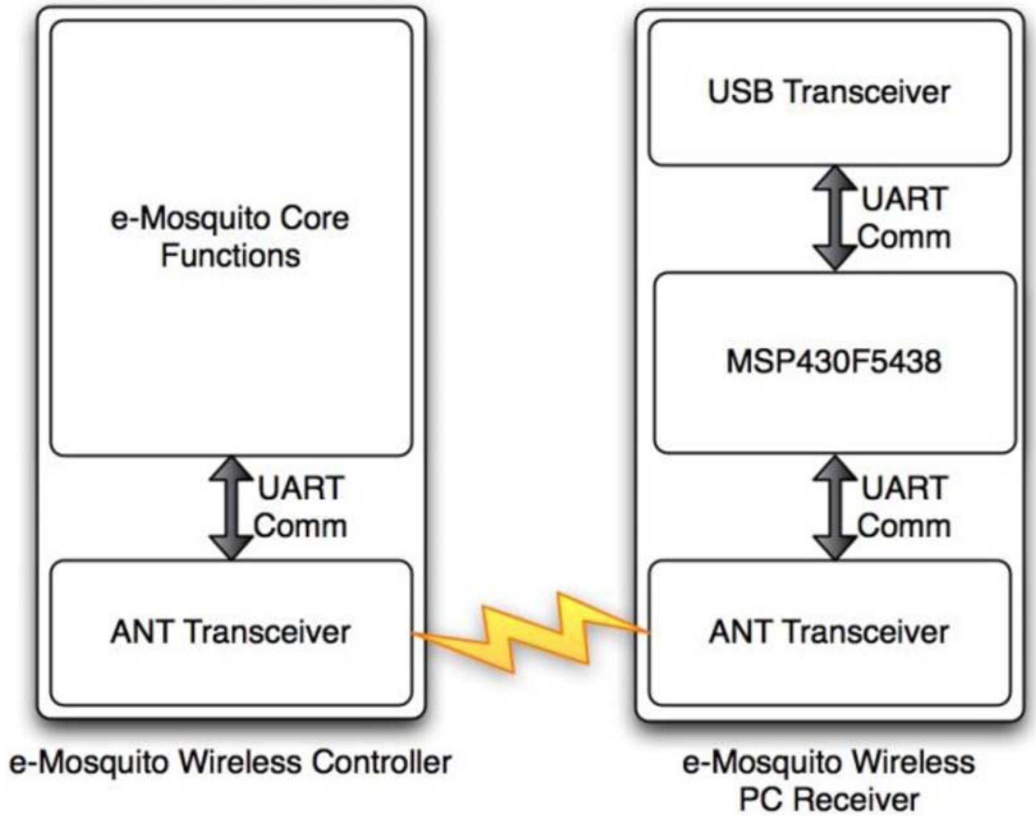


Figure 6. Overall architecture of the first-generation e-Mosquito wireless link. The e-Mosquito transmits data wirelessly to the ANT transceiver, which then translates and delivers the data to a computer via USB [21].

The hardware itself was constructed on a small PCB with discrete surface-mount components for ease of prototyping. The antenna used had an impedance of $50\ \Omega$ and was matched to the nRF24AP1 chip using the design provided in the datasheet (Figure 7) [21].

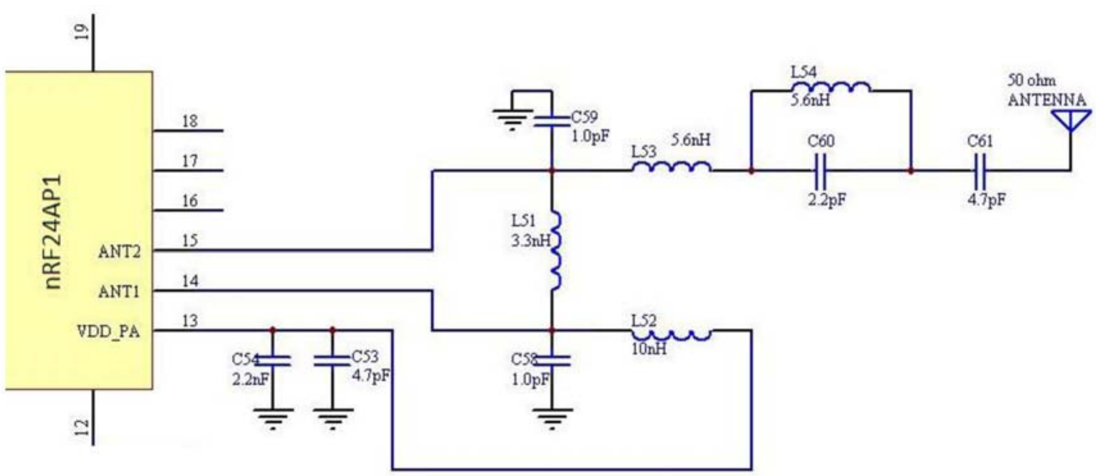


Figure 7. Antenna matching circuitry attached to the nRF24AP1 circuit. Modified from [21].

5.2. Testing Results

The primary concern in testing the wireless link was the range at which it can operate reliably. The wireless communication was tested in three separate environments, all of which had ambient radio noise present. The first two were generic hallways with differing geometry, while the third test took place in a large open outdoor field. The fidelity of the resulting transmission at a variety of distances was approximated as the percentage of packets that were successfully received in an uncorrupted state. The resulting signal fidelity (Figure 8) indicates that the wireless link is 99% reliable to twenty feet in most situations, however hallway #2 had some fidelity issues beginning at roughly 18 feet (70% at 20 feet).

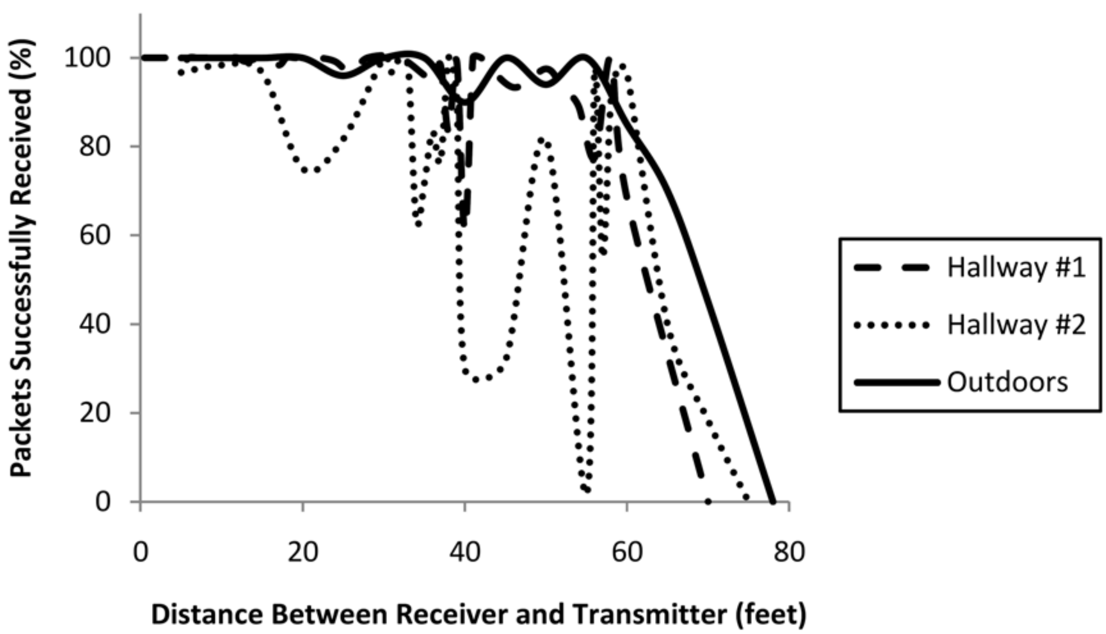


Figure 8. Wireless ANT chip range and signal fidelity in three radio-noisy environments.

Looking at the entire transmission profile out to approximately 80 feet, it becomes apparent that the transmission quality falls rather dramatically at certain points (20 feet, 40 feet, 55 feet) and then recovers. These transmission nodes, coupled with the fact that they are not uniform in location or magnitude across the different settings, indicates that the wireless link is prone to self-destructive multipath interference effects. It was felt during the design stage that the nRF24AP1 chip used would be able to handle this short distance without any issues. However as this is not the case, a packet acknowledge/retransmit system will be implemented (see Section 6 below).

The electrical current consumption of the wireless link is also a primary concern, as this relates directly to the lifetime of the batteries. The e-Mosquito will be disposable and will operate over the course of 1-2 days before replacement, meaning that the demands on the battery are not extreme. Testing has determined that the average current consumption of the wireless link is about $100\ \mu\mathrm{A}$, although the peak current is approximately $15\ \mathrm{mA}$ for $\sim 500\ \mu s$. While the average current draw is about 3 times the $30\ \mu A$ indicated in the datasheet (likely due to additional components such as a power-indicator LED etc.), this still means that the current consumption of the device is still only 2.18% of the total power available to the device (assuming operation over two days. However the peak current does exceed the maximum sustained current draw of the battery and it is unknown presently what effect this will have.

Finally, the wireless link is under continuous use in the laboratory as the primary means of controlling the e-Mosquito prototype. It has been used for a period of 10 months and in that time has shown no noticeable issues. While this is a qualitative measure, it is a strong indicator of the success of the design.

6. Future Directions

The first step to be undertaken to improve the design of the wireless link is to remedy the range issues by implementing a simple acknowledgement response to the protocol. Any message sent by the device will be acknowledged on receipt by a return message being sent. If the original sender does not receive an acknowledgement it will attempt to resend the message at an increasing interval until the message is either received and acknowledged, or times out. This model is quite well known in the communications industry and should easily meet the target of 99% fidelity in a 20 foot range.

The end goal of the e-Mosquito and hence this wireless link is of course an automatic external artificial pancreas (Figure 9). The wireless link allows the e-Mosquito to exist conveniently separate from the insulin infusion pump.

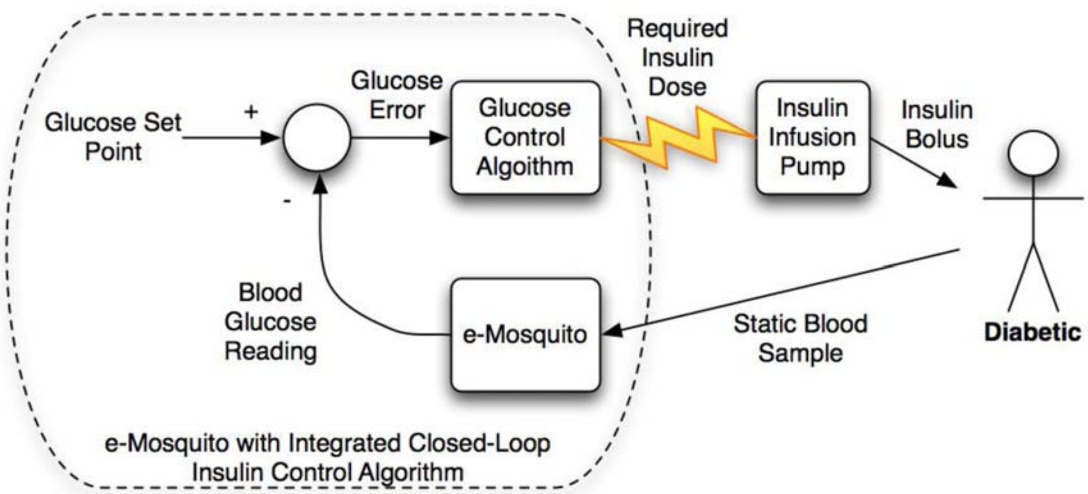


Figure 9. High-level model of the e-Mosquito integrated into a complete artificial pancreas. The required insulin dose is transmitted to the insulin pump wirelessly [21].

Allowing the e-Mosquito to be controlled wirelessly as well as deliver glucose readings also opens up the exciting potential for personal area medical networks (PAMNs) and the second-generation wireless link is being designed to operate in such a manner. The same design process was used as the first-generation link and the constraints table is presented below as Table 2.

Table 2. Design constraints for a second-generation, PAMN-capable wireless link for the electronic mosquito.

Constraint	Description
2.1	***Standard Protocol:*** A standardized communication protocol needs to be developed that can be implemented by a variety of different devices to allow them to interoperate. This protocol design will need to take into account the rest of the constraints.
2.2	***Security Encryption:*** Medical data is subject to strict privacy laws and any devices broadcasting medical data on an uncontrollable medium such as a wireless link. Furthermore, the health and safety of the user could be affected if the device was controlled by an unauthorized signal. RSA encryption using a sufficiently long key that is unique to the patient should be sufficient to maintain the privacy of the network.
2.3	***Dynamic Add and Drop Behaviour:*** Devices coming into range of the network (PAMN) should be able to add themselves to the network if they are keyed to the appropriate patient, and should be able to leave the network if they go out of range. The topology of the network needs to be malleable in order to rearrange from the arrival and departure of devices.
2.4	***Path Redundancy:*** The topology of the network should be redundant, for example organized as a self-healing ring or a mesh network. If a communication channel between two devices is disrupted, the message should be deliverable in another way.
2.5	***Distributed Storage:*** Due to the completely decentralized design of the PAMN necessary by constraints 2.3 and 2.4, there is no centralized repository of data. Instead the devices in the network must store their own data and archives and make this available to other devices in the network. Conversely, the devices must have a manner in determining which device contains the data they are requesting.

7. Discussion

The wireless system designed in this study was sufficient for the continuing development of the e-Mosquito prototype in the lab. However, the multipath echo interference issues need to be addressed. Fortunately, the proposed acknowledge-response method does not require any hardware changes and has been implemented in the software. The transmission rate of the device provides sufficient headroom for many more features to be implemented and many devices to be operated on the same network simultaneously. The power consumption was also minimal and on average should not have any adverse effect on the battery life of the device. However the TX/RX current draw of the link exceeded 13 mA for a very short period, which is greater than the CR2032 maximum sustained current of 3.5 mA. It is not known what effect this might have on the battery life, but no problems have been noted to present, and it is likely the short period of time mitigates any problems.

The more exciting aspect of the proposed wireless link is its future application in an artificial pancreas. An artificial pancreas is the goal of a huge amount of research at present, but the e-Mosquito is ideal for this application since it is painless, hassle-free, and able to sample almost continuously at a high rate. The wireless link is another step in this direction, by untethering the disposable sampling patch from the more permanent implanted insulin pump.

8. Conclusions

The wireless system that was developed for the e-Mosquito prototype uses the nRF24AP1 ANT wireless chip and protocols provided by Nordic and Dynastream (Cochrane, Alberta, Canada). This chip was utilized to develop a first-generation wireless link for the e-Mosquito that fulfilled a variety of design constraints. This wireless system was a basic test-bed for use in the lab and did not implement many of the advanced features that are being implemented in the second-generation setup, including cryptographic security and message acknowledgement. However the device satisfied the majority of constraints placed on the design, and has functioned perfectly in a lab setting for a period of 10 months.

Medical data is slowly migrating to electronic forms in the increased pervasiveness of electronic medical records and electronic imaging and storage. A PAMN is another step along this path with a myriad of potential applications. Patients in range of an open wireless internet linkup could provide data on their health and wellbeing to healthcare providers remotely, and more detailed data will allow in depth and precise personalized diagnosis and treatment. Communication of medical devices permits systems such as paired glucose monitors and insulin pumps, or a heart monitor able to dial a cell phone in case of emergency. The information from multiple devices intercommunicating could be integrated to provide a holistic view of overall patient health, providing a more complete picture of the ongoing disease processes to physicians and caregivers.

Bibliography

[1] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030," Diabetes Care, vol. 27, no. 5, pp. 1047-1053, May 2004.

[2] U.S. Census Bureau. (2009, September) Total Midyear Population for the World: 1950-2050. [Online]. HYPERLINK "http://www.census.gov/ipc/www/idb/worldpop.php" http://www.census.gov/ipc/www/idb/worldpop.php

[3] U. Masharani, "Diabetes Mellitus & Hypoglycemia," in Current Medical Diagnosis and Treatment, 45th ed., Lawrence M. Tierney, Stephen J. McPhee, and Maxine A. Papadakis, Eds.: McGraw-Hill Medical Publishing Division, 2005, vol. 1, pp. 1194-1239.

[4] M. J. Sheetz and G. L. King, "Molecular Understanding of Hyperglycemia's Adverse Effects for Diabetic Complications," Journal of the American Medical Association, vol. 288, no. 20, pp. 2579-2588, 2002.

[5] D. M. Nathan, "Long-Term Complications of Diabetes Mellitus," The New England Journal of Medicine, vol. 328, no. 23, pp. 1676-1685, June 1993.

[6] American Diabetes Association. (2010, January) Diabetes Basics: Type 1. [Online]. HYPERLINK "http://www.diabetes.org/diabetes-basics/type-1/" http://www.diabetes.org/diabetes-basics/type-1/

[7] R. J. Williams, Ed., Biology, 2nd ed. Redwood City, United States of America: The Benjamin/Cummings Publishing Company, Inc., 1990.

[8] American Diabetes Association. (2010, January) Diabetes Basics: Type 2. [Online]. HYPERLINK "http://www.diabetes.org/diabetes-basics/type-2/" http://www.diabetes.org/diabetes-basics/type-2/

[9] J. Steinberger, A. Moran, C. Hong, D. R. Jr Jacobs, and S. R. Alan, "Adiposity in childhood predicts obesity and insulin resistance in young adulthood," The Journal of Pediatrics, vol. 138, no. 4, pp. 469-473, April 2001.

[10] S. Klein et al., "Waist circumference and cardiometabolic risk: a consensus statement from Shaping America's Health: Association for Weight Management and Obesity Prevention; NAASO, The Obesity Society; The American Society for Nutrition; and the ADA," The American Journal of Clinical Nutrition, vol. 85, pp. 1197-1202, 2007.

[11] G. A. Colditz, W. C. Willett, A. Rotnitzky, and J. E. Manson, "Weight Gain as a Risk Factor for Clinical Diabetes Mellitus in Women," Annals of Internal Medicine, vol. 122, no. 7, pp. 481-486, April 1995.

[12] H. Bui, K. Perlman, and D. Daneman, "Self-monitoring of blood glucose in children and teens with diabetes," Pediatric Diabetes, vol. 6, pp. 50-62, 2005.

[13] J. Wang, "Glucose Biosensors: 40 Years of Advances and Challenges," Electroanalysis, vol. 13, no. 12, pp. 983-988, August 2001.

[14] D. S. Bindra, Y. Zhang, G. S. Wilson, and R. Sternberg, "Design and in vitro studies of a needle-type glucose sensor for subcutaneous monitoring," Analytical Chemistry, vol. 63, no. 17, pp. 1692-1696, September 1991.

[15] R. O. Potts, J. A. Tamada, and M. J. Tierney, "Glucose monitoring by reverse iontophoresis," Diabetes Metabolism Research Review, vol. 18, pp. S49-S53, 2002.

[16] J. Wagner, C. Malchoff, and G. Abbott, "Invasiveness as a Barrier to Self-Monitoring of Blood Glucose in Diabetes," Diabetes Technology & Therapeutics, vol. 7, no. 4, pp. 612-619, 2005.

[17] J. M. M. Evans et al., "Frequency of blood glucose monitoring in relation to glycaemic control: observational study with diabetes database," BMJ, vol. 319, pp. 83-86, 1999.

[18] A. J. Karter et al., "Self-monitoring of blood glucose levels and glycemic control: the Northern California Kaiser Permanente Diabetes registry," The American Journal of Medicine, vol. 111, no. 1, pp. 1-9, July 2001.

[19] H. W. Vesper, P. M. Wang, E. Archibold, M. R. Prausnitz, and G. L. Myers, "Assessment of Trueness of a Glucose Monitor Using Interstitial Fluid and Whole Blood as Specimen Matrix," Diabetes Technology & Therapeutics, vol. 8, no. 1, pp. 76-80, 2006.

[20] G. E. Gattiker, K. V. I. S. Kaler, and M. P. Mintchev, "Electronic Mosquito: designing a semi-invasive Microsystem for blood sampling, analysis and drug delivery applications," Microsystem Technologies, vol. 12, no. 1-2, pp. 44-51, December 2005.

[21] G. Thomas, "Electronic Mosquito: A Feedback-Controlled Semi-Invasive Microsystem for Glucose Monitoring," Department of Electrical and Computer Engineering, University of Calgary, Calgary, MSc Thesis 2009.

[22] Nordic Semiconductor, nRF24AP1 Product Specification, September 2006, http://www.thisisant.com/images/Resources/PDF/nrf24ap1%20product%20specification.pdf.

[23] Renata Batteries, CR2032 Datasheet, October 2001.

Authors' Information

Tristan Jones *– M.Sc. Student, Department of Biomedical Engineering, University of Calgary, Alberta, Canada, T2N 1N4. Phone: (403) 210-9553; e-mail:* HYPERLINK "mailto:jones.tristand@gmail.com" jones.tristand@gmail.com

Karan V.I.S. Kaler *– Prof., Dr., Department of Electrical and Computer Engineering, University of Calgary, Alberta, Canada, T2N 1N4. Phone: (403) 220-5809; e-mail:* HYPERLINK "mailto:kaler@ucalgary.ca" kaler@ucalgary.ca

Michel Fattouche – *Prof., Dr., Department of Electrical and Computer Engineering, University of Calgary, Alberta, Canada, T2N 1N4. Phone: (403) 210-8498; e-mail:* HYPERLINK "mailto:%20mtfattou@ucalgary.ca" mtfattou@ucalgary.ca

Martin P. Mintchev *– Prof., Dr., Department of Electrical and Computer Engineering; University of Calgary, Alberta, Canada, T2N 1N4. Department of Surgery, University of Alberta; Edmonton, Alberta T6G 2B7. Phone: (403) 220-5309; Fax (403) 282-6855; e-mail: HYPERLINK "mailto:mintchev@ucalgary.ca" mintchev@ucalgary.ca*

IMPROVING OF RECOGNITION ACCURACY OF ECG-SIGNAL IN VARIOS DISORDERS OF HEART AND OPTIMIZATION OF TREATMENT BY DRUGS

Nataliya Bilous, Elena Visotskaja, Olga Kozina, Andrey Porvan, Gleb Kobzar, Alexey Krasov

Abstract: Application of ECG scale-space representation and its derivatives for clarification of point positions on results of wavelet detector for increase of boundary points determination accuracy in the automatic mode at higher rapidness while providing maximal reliability of recognition of ECG elements is offered. In addition, possibility of scale-space representation application is investigated for the selection of cardiac cycles for averaging signal. A new method of optimal drug selection is also proposed that reduces recipe space more than doubled taking into account required for treatment of patient symptoms pharmacological actions of drugs. Using the obtained function as function of goal for deterministic parametric model of optimal selection of multi-direction activity drugs allows adequately and systematically complete process of selecting of effective treatment of all symptoms of patient..

Keywords: electrocardiogram, ECG analysis, diagnostics, curvature, scale-space, dynamic programming, м-rate of curvature, optimal drug selection, coefficient of efficiency of impact.

ACM Classification Keywords*: I.5 PATTERN RECOGNITION and J.3 LIFE AND MEDICAL SCIENCES*

Introduction

Cardiovascular diseases are the leading cause of death in almost all developed countries. The absolute mortality rate in Ukraine from cardiovascular diseases in the 6-8 times higher than those in developed Western countries. It becomes clear to what extent is the issue of early diagnosis, prevention and treatment of cardiovascular diseases in our country.

Diseases of heart - a bright example of pathology, course and outcome of which is directly dependent on the timing of treatment to the doctor, timely diagnosis and start appropriate treatment. Many functions of cardiovascular system can be monitored by various devices and give important information about the status of the organism and the possible deviations from the normal regulation of functions. Electrocardiography - method for studying heart which does not lose its status over time. It remains one of the most common and integral methods of cardiac diagnosis and has continued to develop and improve. Achievements of recent years show that ECG provides information not only on the electrical and anatomical characteristics of heart, but also on changes in heart at the molecular level.

Modern possibilities of computer processing of signals enable rapid processing of large data sets. The combination of these capabilities and traditional methods of analysis of ECG allows to create computational cardiomonitory which complete automatic analysis of time and frequency parameters, storing electrocardiograms (including compression and transfer) during full cycle of the survey from accumulation of initial data to obtain a qualified medical decision.

Complication of automatic analysis of ECG consists in signal structure ambiguousness. Forms of ECG depends on both the using leads and the patients pathologies. In addition, ECG signal is exposed to baseline wander noise, electromyographic interference, electrodes pop or contact noise, patient-electrode motion artifacts and 50 or 60 Hz power line interference. Thus, the stage of contour analysis is critically important for the correct raising of diagnosis, as logic of diagnostic algorithms works exactly on the basis of recognized and measured parameters of ECG. That is why from 80^{th} and till now automatic analysis of ECG is one of the most actual tasks in the area of biomedical technologies which not have a definite and complete decision. Moreover investigations in telemedicine, development of projects for remote processing of diagnostic data by modern communication means brings to forming new requirements into tasks of ECG automatic analysis (EAA). One most important requirement is an increase of rapidness of EAA for decreasing of a server's loading up to work in the mode of the real time.

The task of EAA has wide set of decisions: structural methods, correlation methods, Wavelet-transform approach, methods on the basis of analysis of derivatives. However, it will shown farther, methods which possess high recognition of ECG elements distinguishing, do not provide sufficient accuracy of elements boundaries determination. And on the contrary, methods providing high accuracy of elements boundaries determination on condition of correct detection of electrocardiosignal are less steady to the presence of noises and weakening of signal, i.e. do not possess sufficient ECG elements recognition reliability. These premises allow talking about the necessity of development of complex approach for recognition of ECG elements and determination of their boundaries.

At the calculation of basic ECG parameters with assistance of characteristic points inside 5-15 secs record arise up tasks of choice of informational cardiac cycle and construct of averaged signal. Providing of correct choice of informational cardiac cycle allows to avoid errors related to arising of high-peak interferences on frequencies of useful signal, and constructing of average signal - to reduce noise level without considerable distortions of initial ECG.

In this article the application of scale-space signal representation and its derivatives for clarification of end-points of characteristic elements positions by results of Wavelet-detector work is offered. Such approach provides maximal reliability of recognition of ECG elements and improves accuracy of characteristic elements end-points positions determination in the automatic mode with increased rapidness.

In addition, using scale-space representation (SSR) of ECG for selecting both representative (informational) cardiac cycle and cardiac cycles at calculation of average signal is offered. Using of SSR allows to remove possibility of selecting of cycles with interferences and with single artefacts as informational cardiac cycle and also enables to find pathological rejections which can arise up in single cardiac cycle. In the article the developed methods of rapid construction and comparison of SSR for providing of increased rapidness during work with plenty of records of ECG are offered.

Adverse changes in the cardiovascular system require the adoption of preventive measures and mild correction.

The complexity of the processes that determine the optimal adaptation of heart to the needs of the organism is reason of not a simple choice of efforts for correcting the state of the organism even for qualified doctor. Drugs that can have a useful effect in one case may give the opposite effect with similar symptoms to cardiac irregularities in the other case, when the disruption of heart based on other causes. In addition, all drugs are affecting to more than one physiological mechanism. That is why the most effective means of improving the state of heart, correcting violations in its work is the optimum set of plant origin drugs.

In this paper, the using of linear and nonlinear programming methods to solve the problem of optimal selection of drugs is proposed. This approach allows to select adequate appointing due to effective treatment of all patient symptoms taking into account his individual characteristics of the organism and the pharmacological characteristics of using drugs.

Also the mathematical model of the theory of Markov chains and numerical methods for evaluating the judgments of experts to determine the severity of the state of the human body is invited.

Introduction

Complication of automatic analysis of ECG consists in signal structure ambiguousness. Forms of ECG depends on both the using leads and the patients pathologies. In addition, ECG signal is exposed to baseline wander noise, electromyographic interference, electrodes pop or contact noise, patient-electrode motion artifacts and 50 or 60 Hz power line interference. Thus, the stage of contour analysis is critically important for the correct raising of diagnosis, as logic of diagnostic algorithms works exactly on the basis of recognized and measured parameters of ECG. That is why from 80th and till now automatic analysis of ECG is one of the most actual tasks in the area of biomedical technologies which not have a definite and complete decision. Moreover investigations in telemedicine, development of projects for remote processing of diagnostic data by modern communication means brings to forming new requirements into tasks of ECG automatic analysis (EAA). One most important requirement is an increase of rapidness of EAA for decreasing of a server's loading up to work in the mode of the real time.

The task of EAA has wide set of decisions: structural methods, correlation methods, Wavelet-transform approach, methods on the basis of analysis of derivatives. However, it will shown farther, methods which possess high recognition of ECG elements distinguishing, do not provide sufficient accuracy of elements boundaries determination. And on the contrary, methods providing high accuracy of elements boundaries determination on condition of correct detection of electrocardiosignal are less steady to the presence of noises and weakening of signal, i.e. do not possess sufficient ECG elements recognition reliability. These premises allow talking about the necessity of development of complex approach for recognition of ECG elements and determination of their boundaries.

At the calculation of basic ECG parameters with assistance of characteristic points inside 5-15 secs record arise up tasks of choice of informational cardiac cycle and construct of averaged signal. Providing of correct choice of informational cardiac cycle allows to avoid errors related to arising of high-peak interferences on frequencies of useful signal, and constructing of average signal - to reduce noise level without considerable distortions of initial ECG.

In this article the application of scale-space signal representation and its derivatives for clarification of end-points of characteristic elements positions by results of Wavelet-detector work is offered. Such approach provides maximal reliability of recognition of ECG elements and improves accuracy of characteristic elements end-points positions determination in the automatic mode with increased rapidness.

In addition, using scale-space representation (SSR) of ECG for selecting both representative (informational) cardiac cycle and cardiac cycles at calculation of average signal is offered. Using of SSR allows to remove possibility of selecting of cycles with interferences and with single artefacts as informational cardiac cycle and also enables to find pathological rejections which can arise up in single cardiac cycle. In the article the developed methods of rapid construction and comparison of SSR for providing of increased rapidness during work with plenty of records of ECG are offered.

1 Automatic Analysis Of ECG. Tasks And Problems

ECG are differed due to a large variety, which is conditioned by both distinction of the decided tasks and specific of the investigated parameters of signal. The automatic analysis of ECG is usually conducted in two stages: preprocessing and extraction of characteristic elements ECG. As marked above, the most essential stage of ECG signal processing is extract of its important elements by recognition a QRS- complex, by selecting of its characteristic points (tops of Q, R, S waves, end points of QRSb and QRSe complexes and Pb, Pe, Tb, Te characteristic elements), by determination of some supporting point for relative start of measuring of RR-intervals durations (Figure 1.1).

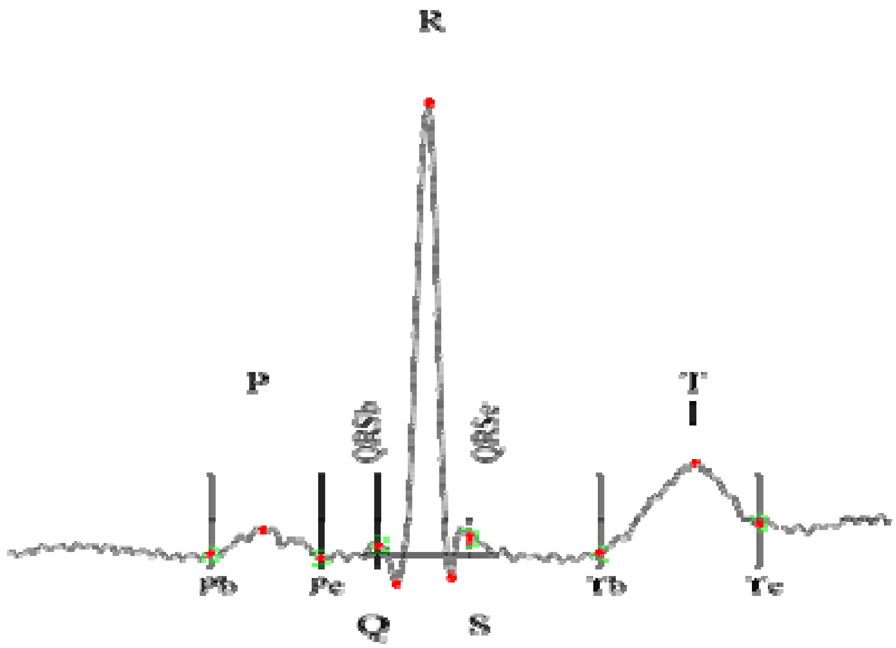


Figure 1.1 – characteristic points of ECG

It is possible to select a few groups of methods of EAA, basic from which are the followings:

–methods which use the analysis of derivatives of signal;

–structural methods that based on preliminary segmentation of signal, representation of signal as a sequence of the simplest elements and subsequent grammatical analysis of the received chain of characters;

–correlation methods which use the analysis of function of correlations between entrance ECG and by a severeal templates of ventricular complex;

- methods which based of wavelet transformations.

The most typical method of recognition of ECG elements is based on analysis of derivatives and consists of three stages: recognition of peaks with additional filtration for the removal of false peaks, a differentiating of peaks or in other words a finding of boundaries between real peaks, and also recognition of characteristic elements of ECG (complexes and waves).

Group of ECG derivative analysis methods is the most widespread as it was a starting point in AA of ECG. One of methods in the group of differentiating of peaks is based on the criterion of maximal curvature (which is formulated through the first and the second order derivatives). Due to this criterion, boundary points of peaks are points of maximal curvature. The other method uses derivatives in a different way analyzing the maximal slope of ECG curve which provides better result. The problem of this group of methods is low stability to the interferences and a necessity of preliminary filtration where authenticity of elements recognition will depend on the filtration efficiency. Nevertheless multidimensional analysis and more sophisticated rules for decision making provides here very good results [Laguna 1994].

The syntactic methods of automatic analysis of ECG are based on grammatical description of the segmented signal. Used segmentation has a few differences from offered Einthoven. Description of ECG is made from four characters — p, r, b, t each of which corresponds to the certain area of curve. In syntactic method, at first, localization of one R wave find by searching of simple maximum, then parametric description of QRS-complex (amplitude, duration of front and back front) create and after scanning of whole initial array of selections for localization of analogical alike areas which are other QRS-complexes is carried out. Methods of this group are showing good stability to the vibrations of baseline, however give errors at close amplitudes of R and T waves and at considerable noise level of initial signal of ECG.

A significant amount of research effort has been devoted to the automated detection of the fiducial (reference) points of the ECG characteristic waves [MINAMI 1992]. Most of these methods are filtering or adaptive thresholding based, which exhibit limitation in real application. Very few algorithms work well for the detection of all fiducial points such as the onsets and offsets of the P wave, T wave and the QRS complex (also known as the ECG wave boundaries). The main drawback of filtering-based approach is that frequency variations in the characteristic waves often adversely affect its performance. The frequency distribution of QRS complexes generally overlaps with that of the noise, resulting in both false positive and false negative detections. The main problems of the thresholding techniques are their high noise sensitivity and their low efficiency when dealing with odd morphologies. Therefore, more sophisticated signal processing techniques are needed to facilitate the development of new detection schemes with higher detection accuracy.

The other group of methods is based on the use of certain sequence of wavelet transformations for finding out complexes and waves. This group of methods appeared and developed comparatively recently – during last 10 years – however stability to noises and efficiency of finding out ECG elements had allowed the wavelet methods to take one of dominant places in the area of EAA [Al-Fahoum 1999]. Wavelet method, possessing most authenticity of ECG complexes recognition, waves and determination of their boundaries is based on consecutive application of CWT and FWT transforms which give abilities to separate noise and P, T waves from the QRS complex. This provides better results compared to one stage filtering of the QRS complex with only FWT or CWT transforms, if P and T waves are of high amplitude. The main advantage of the method is that it can be adjusted by varying the CWT transform frequency in the range of 1–3 Hz to particular T wave morphologies with clinically accepted precision which resulted in improvement of our score [Chesnokov 2006].

This method based on Wavelet transforms gained the lead on the official tests of methods of EAA in 2006 year by providing maximal authenticity of recognition and accuracy of measurings. However the necessity of the manual tuning of frequency for correct determination of P and T waves boundaries for most tasks of EAA allows to talk about the incomplete automatic work of the method and limits spheres of its application. In addition the fast-acting of method is not high because sequence of Wavelet transformations is used.

Thus, the method on Wavelet transforms provides maximal reliability of ECG complexes and waves recognition and determination of their boundaries. Demerit is only a semi-automatic process of measuring of P and T waves boundaries. The removal of this failing and increasing of rapidness will allow solving of given tasks.

2 Complex Approach Of ECG Elements Recognition And Boundary Determination Based On Wavelet Analysis And Scale-Space Representation of ECG

For increase of accuracy of boundary points determination with high reliability of ECG elements recognition the following complex approach is offered:

- QRS complex recognition by the method of wavelet transformations (WT).

- P and T waves recognition and determination of boundary points of QRSb, QRSe and Pb, Pe, Tb, Te characteristic elements by a method based on SSR of cardiosignal.

Thus, the method based on WT will provide maximal authenticity of recognition and the method based on SSR of cardiosignal will provide correct determinations of P and T waves and their boundaries in the fully automatic mode.

The method based on SSR of cardiosignal is offered for P and T waves recognition and determination of their boundaries in place of initial step of WT method.

We will consider formulation of scale-space representation for a continuous signal (in one-dimensional case).

Determination 2.1. For signal $f:\mathbb{R}\to\mathbb{R}$ scale-space representation $L:\mathbb{R}\times\mathbb{R}_{+}\to\mathbb{R}$ is certain so, that reflection of signal at zero level is equivalent to the initial signal [6]:

$$L(x;0)=f(x),$$

and reflection on more wide scales is expressed by convolution of the initial signal and gaussian kernels with increasing width:

$$L(x;\sigma)=g(x;\sigma)*f(x).$$

By using integrals of implicit functions, the result of convolution «*» is like:

$$L(x;\sigma)=\int_{\xi=-\infty}^{+\infty}g(\xi;\sigma)f(x-\xi)d\xi,$$

where $g:\mathbb{R}+\mathbb{R}_{+}\setminus\{0\}\to\mathbb{R}$ is the Gaussian kernel (one-dimensional):

$$g(x;\sigma)=\frac{1}{\sqrt{2\pi\sigma}}e^{-x^2/2\sigma}.$$

As known exactly Gaussian SSR - unique continuous linear SSR which satisfy to all necessary requirements: linearity (regarding to product), invariance to the change, scale invariance, property of semi-group, property of positivity.

For the construction of images of SSR of curvature for selected cardiac cycle the calculation of function of curvature is needed. The formula of curvature for the parametric given closed curve $\Gamma=\{(x(u),y(u)\,|\,u\in[0,1]\}$ has the following view:

$$k(u)=\frac{\dot{x}(u)\ddot{y}(u)-\ddot{x}(u)\dot{y}(u)}{(\dot{x}^2(u)+\dot{y}^2(u))^{3/2}}$$

Curvature Scale-Space (CSS) representation and existent algorithms of comparison of CSS is developed for closed curves and not used for the analysis of signals such as ECG before [Mokhtarian 1996]. However, it is possible to show that it is possible to adapt SSR for representation and comparison of ECG [[Belous 2008].

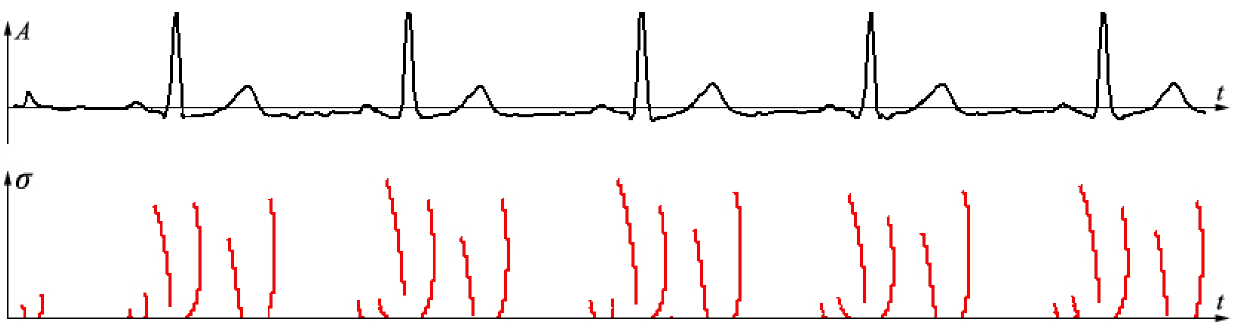
As a round of closed curve can be begun from any point, then CSS is cyclic. Representation of ECG begins and finishes in a certain point, therefore images of CSS of cardiosignal and, accordingly, set of maximums do not require shifting during construction and comparison. The other obvious problem is that curvature scale-space in its initial formulation do not provide convergence. That is why the other SSR is proposed to represent ECG – tangent angle SSR:

$$
\begin{aligned}
&\alpha_i^\sigma = arctg\left(\sum_{k=i-3\sigma}^{i+3\sigma} \alpha_i \cdot g(i-k,\sigma)\right) \\
&g(x,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\cdot\sigma}} \\
&\alpha_i^\sigma \geq 30^\circ, \alpha_{i-1}^\sigma < 30^\circ \\
&\alpha_i^\sigma < 30^\circ, \alpha_{i-1}^\sigma \geq 30^\circ
\end{aligned}
\qquad (2.1)
$$

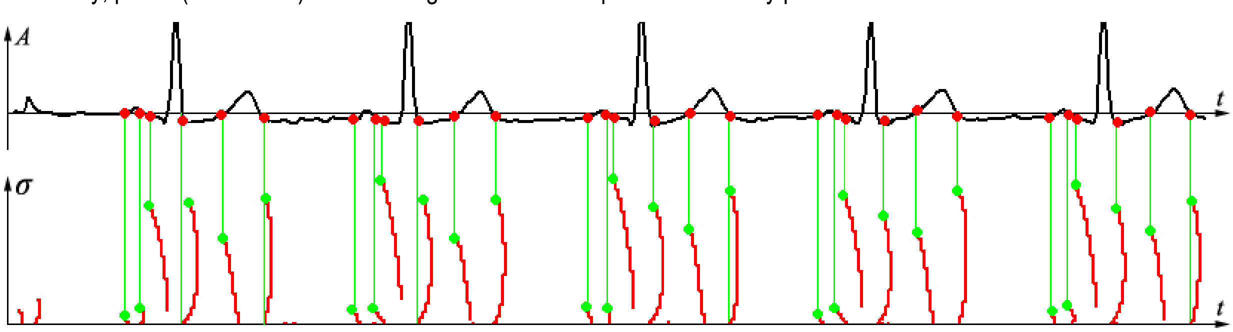
where α_i - tangent angle on the lowest scale, which can be simply calculated having curvature,

α_i^σ - discrete tangent angle representation on the σ -th scale.

As a result of SSR of ECG we get the next image of SSR:



Obviously, peaks (maximums) of SSR image of ECG correspond to boundary points of ECG elements:



3 Selection Of Informative Cardio Cycle, Construction Of Averaged Signal By Comparisons Of ECG Scale-Space representations

The task of informative cardiac cycle selection in the process of EAA appeared as a result of program realization of manual and development of semi-automatic systems of ECG analysis. In the similar systems the operator himself selects most informative cardiac cycle in which parameters of ECG are being calculated (manually or automatically). Operator follows next rules for selection of cardiac cycles:

a) cardiac cycle must be minimum exposed to high-frequency distortions;

b) cardiac cycle must not contain high amplitude interferences on the useful signal frequency (usually, reason of appearance of such interferences is motion of electrodes);

c) cardiac cycle doesn’t have to differ on duration, intensity and also on wave shape from the other ECG cardiac cycles.

The observance of these rules provides the most exact measurings. However with development of the systems of EAA, process of forming of average cardiac cycle begins to replace process of choice of informative cardiac cycle in some systems. Averaging of signal which possible due to periodicity of ECG, provides suppression of noises without distortion of initial signal, especially in small details.

At the choice of cardiac cycles for averaging, algorithms of EAA systems use rules (a) and (b), but don’t take into account that ECG signal is not fully periodic. In addition, published methods of averaging are based on the calculation of correlation on existent cycles which give EAA slower considerably [Strik 1988]. On fig. 3.1 cardiosignal is showed. It contains three cardiac cycles (PP1, PP3, PP4) which will be summarized and one cadiac cycle (PP2) which can result in two possible errors of averaging:

- inclusion of high amplitude interference with useful signal frequency in averaged signal;

- inclusion of two false cardiac cycles FPP2b and FPP2a in averaged signal that will result in considerable distortions.

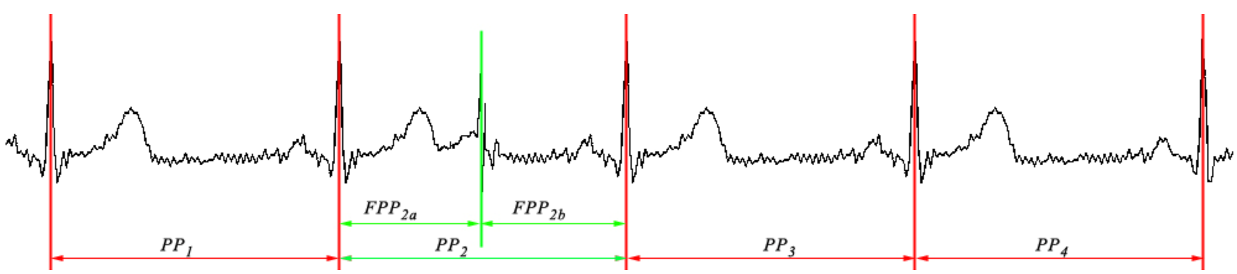


Figure 3.1 Cardiac Cycle Distortion and Results

A method which allows finding the most informative waves in the automatic mode based on comparison of scale-space representations of different cardiac cycles of one signal is considered in this section.

Definition 3.1. Let A - is a set of SSR (maximums of arcs). Then every point $\grave{a} \in A,\ \ i=1,...,N$ are put in accordance two real numbers $u(a) \in \mathbb{R}$ – position of maximum on a curve - and $\sigma(a) \in \mathbb{R}$ is scale level which in pair determine position of point a in scale space.

Comparison of point images or, that equivalence, points set recognition (PSR), is one of the most fundamental problems in area of structural pattern recognition. It arises up in many areas such as 2D and 3D images analysis, treatment of documents, biometric authentication, databases of images, analysis of video, and also biological and biomedical applications. To the same task, obviously, comparison of PSR is related.

A task consists in finding of such accordance between every point of one set and points of other set for which some limitations are carried out, and also some global measure of likeness is optimized. Types of limitations and measures of likeness are determined due to according concrete problem.

Task 3.1 Accordance of two well-organized curve sets A and B of points of PSR can be found in a numeral kind by finding of one-power well-organized subsets $A' \subseteq A$ and $B' \subseteq B$ sets A and B and reflection $f(A') : A' \rightarrow B'$ which minimize the following function:

$$P(A,B,f) = \sum_{i=2}^{|A'|} \sqrt{(\left|u(a_i') - u(a_{i-1}')\right| - \left|u(b_i') - u(b_{i-1}')\right|)^2 + (\sigma(a_i') - \sigma(b_i'))^2} + \\ + \sum_{a \in A/A'} \sigma(a) + \sum_{b \in B/B'} \sigma(b) = P_{dp}^{ex}(A', f(A')) + P^{inex}(A/A' \cup B/B'), \quad (3.1)$$

where $P^{ex}(A', f(A'))$ - is a cost of complete comparison of subsets A', B':

$$P^{ex}(A', f(A')) = \sum_{i=2}^{|A'|} \sqrt{(\left|u(a_i') - u(a_{i-1}')\right| - \left|u(b_i') - u(b_{i-1}')\right|)^2 + (\sigma(a_i') - \sigma(b_i'))^2} \quad (3.2)$$

$P^{inex}(A/A' \cup B/B')$ - is a cost of eliminate of elements of sets A, B at comparison (analogical to the formula 4.3).

Formulation $P^{ex}(A', f(A'))$ in (4.5) requires one clarification. At comparison it is necessary in the beginning to select the first pair of points (a_0, b_0) in relative to which formula (3.2) calculates.

This task may be solved efficiently using dynamic programming [Belous 2008]. Thus, rapid and exact finding of accordance of not signals of cardiac cycles but their scale-space representations is available. This provides a choice only signals with similar form while averaging or selecting informative cycle. Result of ECG averaging that was presented on fig.3.1 is shown on fig 3.3

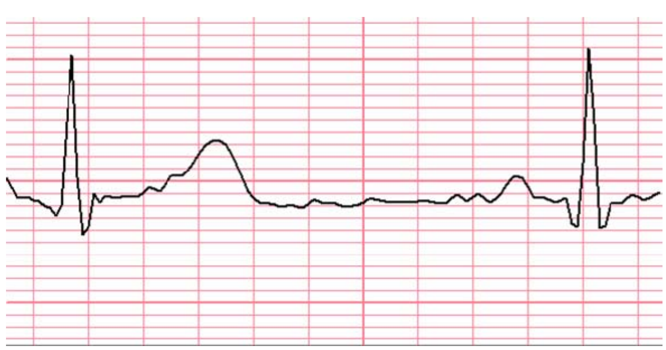


Figure 3.3 Averaging cardiac cycle

4 Method Of Rapid Forming Of Scale-Space Representations Of ECG

As it was mentioned to build SSR of ECG curvature estimation is neded.

As differential characteristic, curvature is very unsteady to discretisation and noising of images. Therefore, for providing of acceptable noise immunity at construction scale-space representations the function of curvature must be estimated indirectly.

Next formulation of m-weight of curve by which it is possible to formulate the estimation of curvature and satisfying to the requirement $\lim_{m \to +0} k_m = k$ (requirement for all estimations of curvature) is offered. For the calculation of such m-weight of curve the points of curve are needed only. The initial geometric curvature estimation method was proposed in [Karkischenko 1998].

It is proposed to use m-estimation of curavature to build curvature base scale-space representations to increase the rapidness of such representation construction. Proposed m-estimation is given in 2-dimentional parametric form of curve representation because it is easily converted to 1D form of ECG but may have a lot of other application in 2D.

Lemma 1. Let m is neighbourhood of point $\gamma(u) = (x(u), y(u))$, into which m-weight of curve given in parametrical view $\Gamma = \{(x(u), y(u) \,|\, u \in [0, L]\}$ is estimated. Then the following asymptotic formula $v_m = v_m^o + O(m), m \to 0$ takes place, thus v_m^o calculated on the following formula:

$$v_m^o = \left| S_m^o - \overline{S}_m^o \right| \Big/ \max(S_m^o, \overline{S}_m^o) \tag{4.1}$$

where $S_m^o, \overline{S}_m^o$ – are areas of sectors of circumference with a radius $\triangle u$:

$$\begin{aligned} S_m^o &= \pi m^2 \angle(\overrightarrow{v^-}(u), \overrightarrow{v^+}(u)) / 360^o \\ \overline{S_m^o} &= \pi m^2 \angle(\overrightarrow{v^+}(u), \overrightarrow{v^-}(u)) / 360^o \end{aligned} \tag{4.2}$$

where $\overrightarrow{v^-}(u)$ - a vector connecting the point of curve $\gamma(u) = (x(u), y(u))$ and $\gamma^-(u) = (x(u-m), y(u-m))$;

$\overrightarrow{v^+}(u)$ - a vector connecting the point of curve $\gamma(u) = (x(u), y(u))$ and $\gamma^+(u) = (x(u+m), y(u+m))$ (fig.4.1).

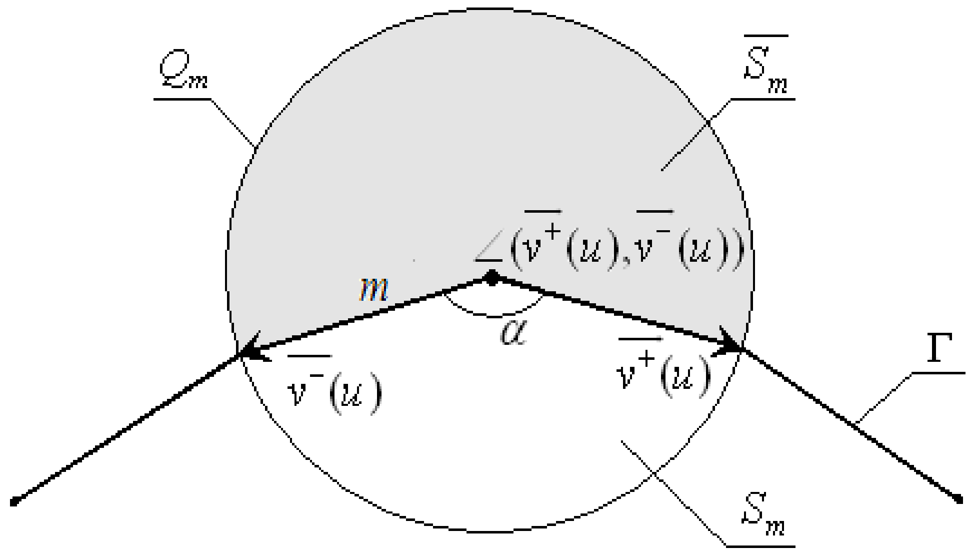


Figure 4.1 Neighbourhood for m-rate of function of curvature of contour curve

Lemma 2.. Let $m = \triangle u$ is Neighbourhood of point into which curvature of curve Γ given in parametrical view $\Gamma = \{(x(u), y(u) \mid u \in [0, L]\}$ is estimated. Then the following asymptotic formula takes place:

$$k = 3\pi v_m^o / (4m) + O_m^o(m),\ m \to 0 \tag{4.3}$$

Lemma 3. It is possible to assert that it is possible to obtain finding of estimation k_m^o with required accuracy at *estimating of curvature to sectoral method by introduction of the following limitation:*

$$v_m^o < \xi_m^o \tag{4.4}$$

where ξ_m^o is some cut-off of m-weight.

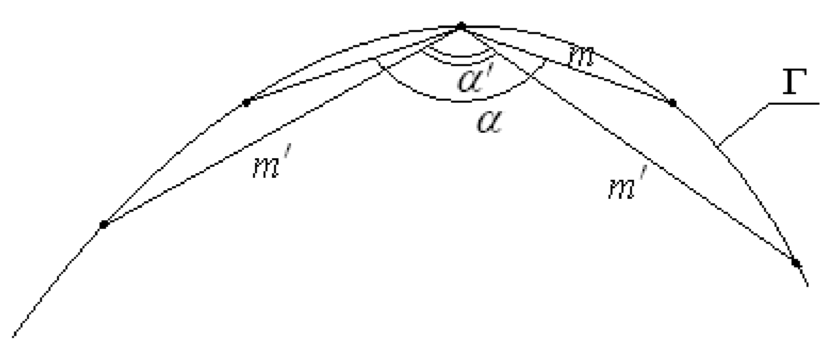


Figure 4.2 Neighbourhood of sectoral m-rate of function of curvature of contour curve

Lemma 3 has a double value for the construction of SSR of curvature and proper handle. At first, changing of neighbourhood m it is possible to check accuracy of m-rate of curvature in points of contour curve.

Secondly, as at $k \rightarrow 0$ error of sectoral rate $O_m^o \rightarrow 0$, it is possible to assert that for finding of transit points of curvature through a zero it is possible initially to take minimum size neighbourhood m . In addition, for points with small curvature it is possible to take small neighbourhoods m with guaranties here finding of estimation of curvature with given accuracy and fast-acting.

For receiving of rapid estimation of curvature with given accuracy it is possible to link neighbourhood of m and step $\triangle u$ of discretisation of contour curve $\Gamma = \{(x(u), y(u) \mid u \in [0, L]\}$.

A curve is presented like set of points $\Lambda = \{(x_i, y_i) \mid i \in [0, M_u]\}, M_u = L_u / \triangle u$, distance between any two points λ_i and λ_{i+1} of set Λ is equal thus

$$\left\| \lambda_i - \lambda_{i+1} \right\| = \triangle u = 1,$$

where $\|\cdot\|$ is the Euclidean metrics.

For the construction of discrete scale-space representation with invariant to scaling, ration the amount M_u of points of curve Λ with losing or adding new points to some general value M , I.e. lead $|\Lambda| = M$ [85, 86]. Thus, accept $L = M$ as new length of contour.

In addition, choose the step of discretisation on a scale $\triangle\sigma$. Then representation of points of curve Λ on a scale σ , such that $\sigma(\mathrm{mod})\triangle\sigma \equiv 0$ (or $\sigma = \triangle\sigma \cdot j, j \in N$), can be easy found on the following formula of the discrete approaching of Gaussian convolution:

$$x_i^\sigma = \sum_{j=-[M/2]}^{j<[M/2]} x_{i-j} g(i-j, \sigma), \quad y_i^\sigma = \sum_{j=-[M/2]}^{j<[M/2]} y_{i-j} g(i-j, \sigma),$$

4

$$g(x;\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/2\sigma} \tag{4.4}$$

Thus, $x_{i-j}^0 = x_{i-j}, \;\; y_{i-j}^0 = y_{i-j}$.

For the construction of SSR of curvature it is necessary to find the values of function of curvature in all M points of discrete curve for each of N_σ levels of scale. The calculation of position of some λ_i point on the levels of scale $\sigma > 0$ according (4.4) requires $O(M^2)$ operations, while due to rule 3σ only points in neighbourhood 3σ bring in a meaningful contribution to the sums (4.4). It is possible to transform formulas (4.4) to the following kind for considerable decreasing of amount of operations, that is ordinary practice at the use of Gaussian convolution in discrete spaces:

$$x_i^\sigma = \sum_{j=-[3\sigma]}^{j<[3\sigma]} x_{i-j} g(i-j,\sigma), \;\; y_i^\sigma = \sum_{j=-[3\sigma]}^{j<[3\sigma]} y_{i-j} g(i-j,\sigma) \quad (4.5)$$

Thus, at the maximal amount of levels on which the calculation of curve points is needed (and this amount in practice depends on how quickly on a curve will be not a single zero transit point of curvature) $\sigma = M/3$ for some $\triangle\sigma$ middle complication of calculations of every level will be order $O(M^2/2)$.

As for the exposure of maximums of SSR of curvature all levels of scale are needed, it is possible to apply another property of Gaussian convolution:

$$(f(x) * g(x,\sigma_1)) * g(x,\sigma_2) = f(x) * g(x, \sqrt{\sigma_1{}^2 + \sigma_2{}^2}).$$

Let σ_i is a level of scale for which the discrete curve $\Lambda^{\sigma(i)}$ of which is already calculated. For the calculation of curve $\Lambda^{\sigma(i+1)}$ of the following level of scale in accordance with chosen $\triangle\sigma$ it is necessary to execute gaussian convolution of curve $\Lambda^{\sigma(i)}$ with a kernel σ^+ size of which can be calculated on the following simple formula:

$$\sigma^+ = \sqrt{\sigma_{i+1}{}^2 - \sigma_i{}^2} \quad (4.6)$$

It allows to take middle complication of calculations of every level to the order $O(M \cdot \sqrt{M})$ at the maximal amount of levels (till to $\sigma = L/3$) on which the calculation of points of curve is needed.

After the calculation of curve of every scale, directly the estimation of curvature can be expected by proposed sectoral m-weight of curve. Corner α_i^σ for the point $\lambda_i \in \Lambda^\sigma$ of curve at some level of scale $\sigma = \triangle\sigma \cdot j, j \in N$ it is possible to calculate as:

$$\alpha_i^\sigma = \arccos\Big((x_i^\sigma - x_{i-1}^\sigma)(x_{i+1}^\sigma - x_i^\sigma) + (y_i^\sigma - y_{i-1}^\sigma)(y_{i+1}^\sigma - y_i^\sigma)\Big) / |\overrightarrow{v_+^\sigma}||\overrightarrow{v_-^\sigma}|, \quad (4.7)$$

where $\overrightarrow{v_+^\sigma} = \lambda_{i+1}^\sigma - \lambda_i^\sigma$;

$\overrightarrow{v_-^\sigma} = \lambda_i^\sigma - \lambda_{i-1}^\sigma$.

After this it is simple to receive sectoral m- weight v_m^o and estimation of curvature k_m^o of point λ_i^σ due to formulas (4.3) and (4.4).

The next method of construction of images of SSR of curvature on the basis of adaptive binary simplification of contour curve and calculation of sectoral m- weight neighbourhood m of which can be related to the level of simplification of curve is offered.

For finding of estimations of curvature of contour function at all levels of scale on the first step it is required to ration the contour curve so that length of discrete curve and amount of points were multiple 2: $L = M = 2^N, \triangle u = 1$.

Determination 4.2.. Binary simplification of discrete curve is an operation of exception of every second point from set Λ^σ *of points of curve at some level of scale* σ *. In other words set* Λ^σ_η *of points of simplified curve can be found as:*

$$\begin{gathered}\Lambda^\sigma_\eta = \{\lambda_i \in \Lambda^\sigma_{\eta-1} \,|\, i = [0, M_{\eta-1}], i (\mathrm{mod} 2 \equiv 0\} \\ \Lambda^\sigma_0 = \Lambda^\sigma\end{gathered} \tag{4.8}$$

Determination 4.3. The level of simplification η *of discrete curve* Λ^σ_η *on a scale* σ *is equal to the amount of simplifications in accordance with (4.8) created for receiving of curve* Λ^σ_η *from a curve* Λ *at all levels of scale* $\sigma' < \sigma$ *. Obviously* $\eta \le \log_2(M)$*.*

It is suggested to calculate the curves of the followings levels depending on the level of simplification, using for development of curve only those points which was saved after binary simplifications on previous levels:

$$x_i^\sigma = \sum_{i \in \{\lambda_j \in \Lambda_\eta^{\sigma-\triangle\sigma} \| \|x_i^{\sigma-\triangle\sigma} - x_j\| < 3\sigma^+\}} x \cdot g(\| x - x_i^{\sigma-\triangle\sigma} \|, \sigma^+), \tag{4.9}$$

where x_i^σ is the first coordinate of finded point $\lambda_i \in \Lambda^\sigma_\eta$ of curve on a scale σ at the level of simplification η;

σ^+ - kernel of convolution which necessary for receiving of curve Λ^σ_η from curve $\Lambda^{\sigma-\triangle\sigma}_\eta$ calculated due to formula (4.8).

By analogy y_i^σ can be calculated .

We will enter cut-off ξ_m^o for m- weight of curves. We will simplify a curve each time during iterative development in accordance with a formula (4.9), when maximal weight of points of curve appears below than some cut-off :

$$\max_{\lambda_i \in \Lambda^\sigma_\eta} v_m^o(\lambda_i) < \xi_m^o + \triangle\xi, \tag{4.10}$$

where $v_m^o(\lambda_i)$ is a value of m- weight of curve Λ^σ_η in a point λ_i ;

$\triangle\xi$ - additional element which guarantees implementation of condition $v_m^o(\lambda_i') < \xi_m^o, \lambda_i' \in \Lambda_{\eta+\Delta\eta}^{\sigma+\Delta\sigma}$ on the curve of next scale and level of simplification.

Due to Lemma 3 it guarantees the calculation of estimation of curvature by sectoral m- weight of curve Λ_η^σ with given accuracy level. Actually, simplification of curve takes place when a curve is smoothed out so that the maximal estimation of curvature $v_m^o(\lambda), \lambda \in \Lambda_\eta^\sigma$ of points of discrete curve decreases to some apriory given level $C \cdot \xi_m^o$ and simplification of curve does not bring to the increase of maximal estimation of curve curvature $v_m^o(\lambda), \lambda \in \Lambda_{\eta+\Delta\eta}^{\sigma+\Delta\sigma}$ higher than some other level $C \cdot (\xi_m^o + \triangle\xi)$, $C = 4m/3\pi$ on next scale.

Application of method of adaptive discretisation allows to considerably accelerate the process of construction of scale-space representations, that will influence on increasing of velocity of as process of recognition of ECG elements and determination of their boundaries so and the process of forming of average signal which was offered in this work.

5 Experiments

For experimental verification of the developed methods selection of control ECG of healthy people, selection of ECG from patients by the heart attack, cardiomiopathy, arrhythmia of different kinds and other of diseases was formed. General amount of records is equal 30 containing on average for 6 leads. Thus amount of records – 180.

3 types of measuring of basic parameters of ECG (determined due to positions of boundary points and peaks of ECG elements) inside every selection, necessary for further diagnostics were conducted: manual measuring by specialists-cardiologists, measuring by the method of WT and offered method. The next tables are got.

Comparing of the automatic and manual measuring:

automatic measuring manual measuring

	Автоматический										Ручной									
№ П	L(P)	A(P)	QRS	A(Q)	A(R)	A(S)	LGK	PGK	L(T)	A(T)	L(P)	A(P)	QRS	A(Q)	A(R)	A(S)	LGK	PGK	L(T)	A(T)
24	0,14	0,058	0,05	0	0,635	-0,81	0,02	0,03	0,26	0,577	0,11	0,057	0,1	-0,03	0,633	-0,81	0,04	0,06	0,23	0,338
	0,14	0,138	0,07	0	1,388	-1,29	0,03	0,04	0,26	0,533	0,12	0,06	0,1	-0,1	1,283	-1,35	0,05	0,05	0,25	0,411
	0,13	0,075	0,09	-0,14	1,008	-1,01	0,06	0,03	0,27	0,247	0,13	0,075	0,1	-0,14	1,008	-1,07	0,06	0,04	0,26	0,151
	0,15	0,132	0,05	0	0	-1,32	0	0,05	0,27	-0,38	0,15	0,132	0,06	0	0	-1,32	0	0,06	0,26	-0,49
	0,17	0,318	0,05	0	0	-1,06	0	0,05	0,31	-0,07	0,15	-0,1	0,08	0	0,337	-1,06	0,03	0,05	0,31	-0,07
	0,13	0,076	0,09	-0,13	1,175	-1,2	0,05	0,04	0,26	0,354	0,13	0,076	0,1	-0,13	1,175	-1,25	0,05	0,05	0,25	0,267
	0,11	-0,05	0,09	0	0,604	-0,64	0,04	0,05	0,23	0,229	0,11	0,055	0,11	-0,03	0,715	-0,87	0,05	0,06	0,22	0,142
34	0,12	0,074	0,11	-0,15	1,323	-1,45	0,05	0,06	0,27	0,44	0,12	0,074	0,11	-0,15	1,323	-1,45	0,05	0,06	0,29	0,44
	0,16	0,109	0,09	0	0,567	-0,78	0,02	0,07	0,29	0,338	0,16	0,109	0,09	0	0,567	-0,78	0,02	0,07	0,29	0,338
	0,09	0,039	0,09	0	0,328	-1,48	0,02	0,07	0,3	-0,26	0,1	-0,1	0,09	0	0,328	-1,48	0,02	0,07	0,32	-0,26
	0,12	0,03	0,07	-0,1	0,857	-0,67	0,04	0,03	0,31	0,121	0,13	0,054	0,08	-0,1	0,857	-0,83	0,04	0,04	0,3	0,213
	0,14	-0,06	0,05	0	1,341	-1,35	0,02	0,03	0,32	0,465	0,14	-0,06	0,05	0	1,341	-1,35	0,02	0,03	0,34	0,465

Rejections according to types:

Отклонения									
L(P)	A(P)	QRS	A(Q)	A(R)	A(S)	LGK	PGK	L(T)	A(T)
0,03	0,001	0,05	0,033	0,002	0,001	0,02	0,03	0,03	0,239
0,02	0,078	0,03	0,104	0,105	0,051	0,02	0,01	0,01	0,122
0	0	0,01	0	0	0,063	0	0,01	0,01	0,096
---	---	---	---	---	---	---	---	---	---
0,008	0,032	0,012	0,017	0,025	0,124	0,004	0,008	0,094	0,228

Mean WT deviation from the manual measuring due to:

- amplitude values: SA=0,085 MB
- intervals values: ST=0,032C

Middle deviation the offered combined method from manual measuring due to:

- amplitude values: SA=0,086MB
- intervals values: ST=0,027C

Rapidness experiments had shown that comparison of ECG (using DTW) is N/E slower than proposed technique based on elastic comparison of SSR of ECG, where N – number of measurings in a cardiac cycle of the initial signal and E – number of points in SSR (~7 for ECG). Using m-estimations and binary simplifications during SSR construction speeds up the process up to log(N) times.

6 The relevance and the problem of optimal selection of drugs

To date a vast clinical experience of joint appointments of several drugs are accumulated. However, the understanding of the main interaction of drugs is not enables to predict the common effect on the body.

Thus, at cardiopatologies with different etiologies if the drug has low bioavailability as a result of high presystem metabolism then the concomitant appointment of drugs or other substances that are its inhibitors, may significantly shift its bioavailability, as well as its effect, and leads to undesirable effects. Conversely, a drug with high bioavailability can be with smaller risk of such interactions because its concentration in the blood under normal conditions closes to maximum. In modern literature described many cases of serious unwanted medical interactions between digidroperidins and immune blockers with low bioavailability.

There are some approaches for finding the optimal set of drugs based on concept decision-making procedures. Decision-making procedures consist on general steps: defining goals; the allocation of many possible ways to achieve it (the set of possible solutions); formation evaluation to determine order's level on the set of solutions (objective estimation); choosing the best solution (optimization problem).

These approaches provide to receive optimal recipe by three-way ranging system of effectiveness evaluating of drug action on the human body due to diagnoses according to respective functions of Gibson-Miller and Clark [Piotrovsky, 1987, Wanger, 2004]. The disadvantage of these approaches is losing of registration of pharmacological interaction between drugs.

There is simplex method for searching of optimal set of drugs [Borzenkov, 1987, Porvan, 2003] which does not take into account the diversity of pharmacological actions of drugs and constructing an isomorphism of graphs. In additional this method use non-linear logarithmic function of extremum search [Vysotskaya, 2004] which does not guarantee the completeness of the intersection of multiple pharmacological actions in finite set of drugs.

It is known the approach of finding the components of medicinal collection from plants with different mechanisms of achieving uniaxial pharmacological effect [Yoneyama, 2000], in which the quantitative ratio of the components of final recipe is defined by minimizing the function $F(X)$ by the formula:

$$F(X)=\sum_{I=1}^{m}[y_I^N(1-\Delta y_I(x))-y_I^{\circ}(x)]^2 \to \min_X \tag{6.1}$$

where $x=(x_1 \dots x_n)$ – drugs composed assumed plant's collection;

$y_I^{\circ}(x)$ – initial value of index I ;

$y_I^N(x)$ – physiological value of index I ;

$\Delta y(x)$ – relative value of pharmacological effect;

m – number of analyzing parameters;

n – number of components in medical collection.

However this approach does not allow carrying out choice of plants because here there is no registration of individual features of the restoration of human body balance and in contrary assume using up the medicines one axial pharmacological action. Implicit indication of analyzed parameters in minimized function can lead to errors of calculation, and thus to inefficient treatment.

Since the human body is an open system, which is influenced by various external factors, and processes in the body are partly probabilistic nature, it can be assumed that the change in the organism homeostasis will also be probabilistic in nature. In connection with this application of all the above described methods and models to adequately describe the evolution of homeostasis as a result of a comprehensive treatment of heart disease drugs is not possible.

7 Simulation of the optimal selection of drugs

Suppose there is a finite set of drugs that constitute the union T_U and set of pharmacological actions of drugs (DS). Let IT_U and IDE_U are sets of their indexes, respectively. And let the patient after surveys receive number of diagnoses from list of diagnosis and symptoms D_U due to international classificatory of diseases (ICD). Let set of indexes ID_U indices D_U .

Assume that the patient have set diagnoses $D = \{d_k\}$, $k \in ID$; $ID \subset ID_U$, where d_k - k-th diagnosis/symptom of the patient. In the treatment of the k-th diagnosis/symptom is recommended using set of DS $T_k = \{t^k_j\}$, $j \in IT_k$; $IT_k \subset IT_U$, where t^k_j – j-th DS useful in the treatment of the k-th diagnosis/symptom. Thus, we can determine set of DS that are recommended for this patient's symptoms, such as:

$$T_X = \bigcup_{k \in ID} T_k \qquad (7.1)$$

At selection of treatment the age, immune status and degree of overall state of the patient must take into account. Therefore expanded base of diagnoses will have view:

$$D' = \{D, G_B, V, A\} = \left\{ \{d_k\}, d_{M[D_U]+1}, d_{M[D_U]+2}, d_{M[D_U]+3} \right\},$$

where $d_{M[D_U]+1} = G_B$ – degree of overall state of the patient;

$d_{M[D_U]+2} = V$ – index of patient age: «0» for age 14 – 60 years old, «1» for 14 years old and after 60 years old patients;

$d_{M[D_U]+3}$ =*A* – allergies to certain drugs value of which can be «0» if the patient have not allergies, «1» – if patient have allergies.

Thus, assume that G_B, V and A can be attributed to the discharge diagnoses, set *k transform in next view*:

$$k' \in ID', ID' = ID \bigcup \{ M[D_U]+1, M[D_U]+2, M[D_U]+3 \},$$

where $M[D_U]$ - the number of diagnoses and symptoms in ICD.

As known, every diagnosis from D' has according set of contraindicated drugs $TP_{k'}$.

$$TP_{k'} = \{t_j^{k'}\}, j \in ITP_{k'}; ITP_{k'} \subset IT_U ,$$

where $ITP_{k'}$- the set of indexes of contraindicated drugs at k'-th diagnosis/symptoms.
The total value of contraindicated drugs is equal:

$$TP = \bigcup_{k' \in ID} TP_{k'} \tag{7.2}$$

Then the set of drugs selected for treatment of the patient using (9) and (10) will defined as:

$$\Omega = T_X \setminus TP \tag{7.3}$$

with its set of indexes – IT_Ω..
Denote power of set of recommended drugs as $M_T = M[\Omega]$, then set of incompatible to j-th drug is:

$$TN_j = \{tn_g^j\}; j, g \in IT_U ,$$

where - g-th incompatible drugs with j-m drug.

Conclusion

Thus, application of scale-space representation of signal and its derivatives for clarification of positions of boundary points on results of wavelet detector provides authenticity of ECG elements recognition at the level of method of wavelet transformations and increase of determination accuracy of positions of bondary points in the automatic mode on 18%.

In addition, offered method for selection of informative cardiac cycle and also selection of cardiac cycles at the calculation of average signal on the basis of scale-space representation, allows to remove possibility of choice of cycles with distortions while averaging the signal, and also enables to recognize pathological rejections arising up in single cardiac cycles. Thus, the method works in N/E times faster than correlation method due to binary simplification (where N is an amount of measurings in a second, E – number of points in SSR).

Proposed method of optimal drug selection reduces recipe space more than doubled taking into account required for treatment of patient symptoms pharmacological actions of drugs. Analyzing the type of function of goal it is possible to get conclusion that in order to find extremum standard mathematical optimization methods can be used, such as method of variation of weight coefficients at private criteria and method of steepest descent. Using the obtained function as function of goal for deterministic parametric model of optimal selection of multi-direction activity drugs allows adequately and systematically complete process of selecting of effective treatment of all symptoms of patient.

The result of using the proposed method is optimally selected list of drugs most effective for treatment of patient. This method enables the physician to improve quality of the treatment of heart diseases and reduce time for optimal solutions, and also reduce impact of side effects of used drugs.

Bibliography

[MINAMI 1992] K. Minami, H. Nakajima, T. Toyoshima, Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network, IEEE Trans. Biomed. Eng. 46 (2) (1999), pp. 179—185.

[Laguna 1994] P. Laguna, R. Jané, P. Caminal Automatic detection of wave boundaries in multilead ECG signals: validation with the CSE database, In Proc. To Comput Biomed Res. Vol. 1 (1994), pp. 45-60.

[Strik 1988] H. Strik, L. Boves Averaging physiological signals with the use of a DTW algorithm. In Proceedings of SPEECH'88, 7th FASE Symposium, Edinburgh, Book 3 (1988), pp. 883-890.

[Al-Fahoum 1999] A.S. Al-Fahoum, I. Howitt, Combined wavelet transformation and radial basis neural networks for classifying life threatening cardiac arrhythmias, Med. Biol. Eng. Comput. 37 (1999), pp. 566—573.

[Linderberg 1994] T. Linderberg Scale Space Theory in Computer Vision, Kluwer Academic Publishers. (1994), 440p.

[Mokhtarian 1996] F. Mokhtarian, S. Abbasi, J. Kittler Efficient and Robust Retrieval by Shape Content through Curvature Scale Space, Int. Workshop on Image DataBases and Multimedia Search (1996), pp. 35-42.

[Chesnokov 2006] Y.C. Chesnokov, D. Nerukh D., R.C. Glen Individually adaptable automatic QT detector. In Proc. Of Computers in Cardiology (2006) pp. 337-340.

[Belous 2008] N. Belous, G.Kobzar Automatic ECG analysis for Preliminary and Detailed diagnostics based on scale-space representation. International Journal "Information Technologies and Knowledge" Vol.2 (2008) pp.53-60.

[Каркищенко 1998] Каркищенко А.Н., Лепский А.Е., Безуглов А.В. Об одном способе векторного и аналитического представления контура изображения // Сборник трудов Всероссийской научно-технической конференции с участием зарубежных представителей "Интеллектуальные САПР–97". – 1998. – С. 107-112.

[Пиотровский, 1987] В.К. Пиотровский. Модельные и модельно-независимые методы описания фармакокинетики: преимущества, недостатки и взаимосвязь. Антибиотики и медицинская биотехнология. №7, 1987.

[Wanger, 2004] Peter Wanger. Algorithms for optimizing drug therapy. In: BMC Medical Informatics and Decision Making. Vol. 4. # 10, 2004.

[Борзенков, 1987] Б.И. Борзенков, А.И. Довнарь. Матричный алгоритм метода хорд для решения трансцендентных систем нелинейных уравнений. АСУ и приборы автоматики. Вып. 59., 1987.

[Порван, 2003] А.П. Порван, Е.В. Высоцкая, А.И. Довнарь. Оптимизация рецептурного пространства лекарственно-растительной информации. Вісник МСУ. Серія «Технічні науки». Т.VI, № 2., 2003.

[Высоцкая, 2004] Е.В. Высоцкая, А.И. Довнарь, А.П. Порван. Минимизация рецептурного пространства растительных лекарственных средств. Клиническая информатика и телемедицина. № 1, 2004.

[Yoneyama, 2000] T. Yoneyama, M. Ceatano. Numerical Optimization Applied to the drug treatment. 4 th IFAC Symp. Modeling and Control. Greifswald. Germany, 2000.

[Загуровский, 2005] М.З. Загуровский, Н.Д. Панкратова. Системный анализ: проблемы, методология, приложения. К.: Наукова думка, 2005.

[Токмачев, 2003] М.С. Токмачев. Цепи Маркова в прогнозировании медико-социальных показателей. Обозрение прикладной и промышленной математики. Т.10. Вып. 2., 2003.

[Довнарь, 2000] А.И. Довнарь, И.В. Прасол, Е.Г. Кунник, В.В. Семенец. Решение компромиссных задач схемотехнического проектирования методом варьирования весовыми коэффициентами в многокритериальных свертках. Радиоэлектроника и информатика. Вып. 4 (13), 2000.

Authors' Information

***Nataliya Bilous** – Ph.D., head of ITLCVS laboratory, professor of Kharkov National University of Radio Electronics, Kharkov, Ukraine; e-mail: belous@kture.kharkov.ua*

***Elena Visotskaja** – PhD, lecturer of Biomedical Electronic Devices and Systems Dept. of KNURE P.O .Box: Lenin Ave. 14, Kharkov-166, Ukraine; e-mail: diagnost@kture.kharkov.ua*

Major Fields of Scientific Research: медицинская кибернетика

***Olga Kozina** – PhD, lecturer of Computers and Programing Department of National Technical University 'KPI'.,*

Major Fields of Scientific Research:

***Andrey Porvan** –Vice-president, Institute of Information Theories and Applications FOI ITHEA, P.O. Box: Lenin Ave. 14, Kharkov-166, Ukraine; e-mail: diagnost@kture.kharkov.ua*

Major Fields of Scientific Research:

***Gleb Kobzar** – Researcher of ITLCVS laboratory; Kharkov National University of Radio Electronics, Kharkov, Ukraine; e-mail: glebkobzar@yahoo.com*

***Alexey Krasov -** Researcher of ITLCVS laboratory; Kharkov National University of Radio Electronics, Kharkov, Ukraine;*

A USE CASE SCENARIO FOR TECHNOLOGY-ENHANCED LEARNING THROUGH SEMANTIC WEB SERVICES

Lilia Pavlova-Draganova, Desislava Paneva-Marinova, Lubomil Draganov

***Abstract**: A use case scenario dealing with the formal presentation of semantic Web services for technology-enhanced learning has been developed in the frames of the SINUS project "Semantic Technologies for Web Services and Technology Enhanced Learning"[1]. The main functionality and design of the SINUS platform, as well as its usage, are demonstrated on the basis of this scenario.*

***Keywords**: Use Case Scenario, Technology-enhanced Learning, Semantic Web Services, Multimedia Digital Libraries.*

***ACM Classification Keywords**: H.3.5 Online Information Services – Web-based services, K.3.1 Computer Uses in Education – Distance learning, H.3.7 Digital Libraries – Collection, Dissemination, System issues.*

Semantic Web Services in a Technology-Enhanced Learning Environment Presented Formally by Use Case Scenario

The project "Integration of Semantic Technologies for Service-Oriented Computing and Technology Enhanced Learning" (SINUS) is an interdisciplinary research project aiming at advancing two of the fastest evolving information technologies – Service Oriented Computing and Technology-Enhanced Learning by applying the Semantic Web Service (SWS)[2] methodology. The main project vision [Dochev and Pavlov, 2009] [Agre and Dochev, 2008] is to provide a dynamic adaptation of learning content to the context and the learner's needs during the learning process through:

- Developing new application-oriented methods and end-user oriented tools for Semantic Web services description, discovery and dynamic composition;
- Developing a new Semantic service-oriented architecture-based framework oriented to eLearning applications facilitating reusability and repurposing of learning objects.
- Developing new methods oriented to Technology-Enhanced Learning for Semantic Web services dynamic composition.
- Developing new methods and tools for creation and semantic annotation of learning objects compatible with SWS methodology.

---

[1] Research project No № Д-002-189 with the National Science Foundation of the Ministry of Education and Science. Project executors: consortium of two science organizations – the Institute of information Technologies at the Bulgarian Academy of Sciences and the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences – and one high technology software company –"Active solutions" Ltd.

[2] Web services, which are self-contained, self-describing, semantically marked-up software resources that can be published, discovered, composed and executed across the Web in task driven way [Arroyo, 2006].

During the search for technologically executable decisions for implementation of technology-enhanced learning in the project a use case scenario has been developed. This scenario covers the development of learning and semantic resources for presenting Bulgarian iconographical art with the potential to satisfy learning needs in various disciplines. The scenario aims at covering a broad set of possible learning situations. It formally presents a large spectrum of activities and functionalities required by the user to be implemented through Semantic Web services. The scenario illustrates the description, discovery and dynamic composition of these services (for this research aspect of SWS technology no general solutions have been found yet). The research on SWS description is directed to the graphical and ontology-based approach of SWS design. The research on discovery and dynamic composition of SWS is based on the data-driven SMS composition approach [Agre and Dochev, 2008].

The use case scenario also describes methods oriented to facilitate the reusability of learning objects and new options in the automation of learning objects discovery, selection and composition within a distributed services architecture seamlessly integrated through ontologies. Along these lines the research is for developing new semantic schemas for SMS, oriented to technology-enhanced learning application, that is based on further development and refinement of the INFRAWEBS Framework[1] (project INFRAWEBS, [Agre and Dochev, 2008]).

The scenario outlines the framework of the activity schema and the supporting information models for dynamic creation and adaptation of learning object (multimedia objects, annotated with content and context-oriented metadata), facilitating their reusability. This aims at developing methods and tools for creation and semantic annotation of learning objects compatible with SWS methodology. A wide range of activities expected by the authors and learners is included. The scenario emphasizes on the active authoring as a major learning activity, *i.e.* on supporting advanced learners in their work to find, collect, integrate and create digital objects with learning purposes. The information models and learning methods content organization described in the scenario are oriented to the mass authors of learning materials without need of specific knowledge and skills concerning information models and basics of ontology engineering. This permits the implementation of a user-centered learning content description tool for creating and editing descriptive, content- and context-oriented metadata. The learning description tool has to take into consideration the end-user profiles adaptable to characteristics, preferences and requirements of different user groups.

The project methodology aims at adapting and developing at larger scale the original methods and software components, developed by project member under the 6FP project INFRAWEBS (www.infrawebs.eu) and LOGOS (www.logosproject.com, [LADL 2007], [Arapi et al., 2007]).

Stages of the Use Case Scenario Development

"The learning scenarios in system design describe typical or important way of use of the system. They are designed to give all the partners in the project (both technical partners and content providers) a shared understanding on the purpose of the system and the ways it will be of use in practice." [LOGOS Deliverable D3, 2007].

The development of the present scenario went through the following stages [Paneva-Marinova et al., 2009]:

Stage 1 Determining the needed functionality of the use case scenario and all ontological and gnoseological assumptions, the methodological approach, basic requirements, detailisation phases of development, etc.

[1] A framework for semantic services engineering that covers the whole SWS life-cycle and allows creation of complex semantically-enable applications [Agre et al., 2008].

Stage 2: Description of a real learning situation in which the SINUS platform would be used according to the present knowledge resources (media objects, descriptions, glossaries) of the chosen knowledge repositories (Multimedia Digital library “Virtual Encyclopaedia of the East-Christian Art” [Pavlova-Draganova et al., 2007a] [Pavlov and Paneva, 2007]), learning resources needed to be developed, learning situations and context of use. Special attention is paid to the learning content in the areas of cultural heritage and in particular in the area of Bulgarian iconographic culture and art. The chosen learning domains have the potential to satisfy various learning needs in disciplines such as humanitarian studies, arts, history, social anthropology, cultural studies, theology, etc. The scenario aims at covering a wide range of possible learning situations – from the processes of development of learning content to the implementation of various types of learning (formal learning, professional qualification, self-training etc.) according to the needs of the target groups of users [Paneva-Marinova et al., 2008] [Paneva-Marinova et al., 2009].

Stage 3: Presentation of the general formulation of the scenario with clear definition of the objectives, basic type of resources, user groups, activities, requirements, preferences, motivation, needed and/or required information services along the learning process according to the various users, other required functionalities and information processes to be achieved.

Stage 4: Formalization of the scenario and presentation of the functionality required by the users via a combination of activities that can be carried out by the SINUS platform. For each activity the following information is defined: input and output data, user group performing the activity, steps leading to the accomplishment of the activity, information structures and tools needed for the development, and maintenance of resources supporting the activity.

Stage 5: Development of the scenario and definition of a minimal and an extended variant.

Stage 6: Determining of the basic requirements to the multimedia, semantic and learning resources supporting the use case scenario.

General formulation of the use case scenario of the SINUS platform and its versions

Overall objective of the scenario

The overall objective of this scenario is to describe the basic ways of exploitation (called *activities*) of the platform for technology-enhanced learning by the various types of users. For each activity the following components are defined: input and output data, users, the steps needed to accomplish the activity, the activity’s information structures and tools.

Basic resources

The basic types of resources managed or provided by the SINUS platform are:

- Primary annotated digital resources – informational content and multimedia objects of the target learning domain. A main source of these resources is the *Virtual Encyclopedia of East-Christian Art*. For each resource a primary annotation is made in a specially developed scheme in XML format that includes Dublin Core descriptors and an indicator for the position of the media file in the digital library repository.
- Semantically annotated digital resources – primary annotated digital resources with additional semantic annotation according to the domain ontology of the selected learning area. They are kept in a repository for digital resources.
- Learning resources – combinations of indicators for one or more semantic digital resources accompanied by text commentaries and annotation according to the LOM standard. According to the

level of elaboration of the learning resources the author defines the corresponding appropriate user types. These resources are kept in a repository for learning resources.

- Semantic resources (ontologies) – domain ontologies developed in the frames of the project that describe the learning domain, the users and the learning methods used.
- Profiles of users – these are semantic metadata describing the knowledge about every user of learner type and thus building his profile. These metadata are determined by the domain ontology of user profiles. The profiles are kept in the repository for user profiles.

Users

The main user groups of the SINUS platform are the developers of various resources by means of the SINUS platform and the consumers of those learning resources.

Developers of various resources by means of the SINUS platform:

- Authors of semantic resources – develop and maintain the domain ontologies in the frames of the project: learning domain ontology[1], user profile ontology [2], ontology of the learning methods[3]. They are experts in the domains covered by the corresponding ontologies. The authors formally present the semantic resources by means of the ontological language of the project using a tool for creating ontologies provided by the SINUS platform. These users define a knowledge level[4], a symbol level[5] and an execution level[6] [Schreiber et al., 2000] of the ontologies created.
- Authors of semantic digital resources (annotators) – annotate and semantically index the primary annotated digital resources of the target learning domain by means of specific tool in the SINUS platform so as to make *semantic digital resources*. The procedure is the following: the annotators find a primary digital resource to be semantically annotated in the multimedia digital library *Virtual Encyclopedia of the East-Christian culture and art*; they select an appropriate domain ontology, annotate and keep the annotations in a specific *semantic annotation repository* for digital resources (repository of semantic digital resources). The annotators use a specific software tool for creating semantic annotations.
- *Authors of learning resources* – create reusable learning resources by means of specific tool in the SINUS platform. Each learning resource is presented as a combination of indicators for one or more semantic digital resources accompanied by text commentaries and annotation according to the LOM

---

[1] *The learning domain ontology* is a domain ontology used for annotation and semantic indexing of digital resources. These resources are being taken from the multimedia digital library “Virtual encyclopedia of the East-Christian art” [Pavlova-Draganova et al., 2007b] [Staykova et al., 2007].

[2] *The user profile ontology* is utilized to create semantic metadata describing the level of knowledge and developing the personal profile of each user of the type *consumer of resources* [Paneva, 2006].

[3] *The ontology of the learning methods* is a domain ontology used to define the learning methods utilized in the process of technology enhanced learning in the frames of the project.

[4] *The knowledge level* determines the goal, range, use, level of formalization of the ontology and collection of data using different methods of extraction.

[5] *The symbol level* corresponds to the specification of categories, characteristics, rules, restrictions, individuals and their synonyms and facts in the ontology. The possible integrations with other ontologies could be examined on this level.

[6] *The execution level* – on this level the ontology is formally implemented on an ontological level and one could estimate its completeness, level of harmonization and rudimentarity.

standard. According to the level of elaboration of the learning resources the author defines the corresponding user types. These resources are kept in a repository for learning resources.

Consumers of learning resources in the SINUS platform:

- Academic users: students of various courses in the target learning domain – these users should be normally of medium or high level of knowledge in the target learning domain or should plan to use the SINUS platform to reach a high level in this domain. They actively search and use the learning resources being found to accomplish their learning goals – development of thematic projects, term projects, graduation works, preparation of analysis and analytic searches of various problems in the area, performing formal TEL education etc.
- Researchers in the target learning domain – the SINUS platform enable these users to search informational materials on specific themes related to their scientific work.
- Non-academic users – these users are not professionally involved in the target learning domain but have interest in it and want to introduce some basic parts of it. They have basic knowledge on search of resources specific for the Internet based environments and can use the SINUS platform to enrich their knowledge in the domain.

Motivation of the users

The basic motivation for the use of SINUS consists in:

- Achievement of the following main learning demands, summarized as follows:
- The academic users need additional specialized learning content with appropriate visual presentation. This content could be thematically organized or not, directly or indirectly related to the university subjects.
- The academic users need access to repositories with learning resources of high quality, which they could use in their projects, analysis, thematic discussions etc.
- The academic users need possibilities to make multicriteria search learning objects from the target learning domain, search with grouping of the results according to various values of the chosen criteria for search; consecutive filtering of the results etc.
- The researchers in the field need new resources of information materials and systematized results from serious studies, made by a specialist in the field.
- All the users need an environment providing them qualitative, well structured and adapted learning content according to their profile.
- The users want to have information about the actual state in the domain, to examine domain resources that cannot be reached in the standard way.
- The users want to be part of a community of interests and to communicate with each other.

Needed/wanted information services in the learning process according to the different user groups

When using the SINUS platform the users – developers of several types of resources are expecting the following:

- Authors of semantic resources: the SINUS platform to provide a software tool/editor for creation of ontologies. This tool should ensure the needed functionality for specification of the categories, characteristics, rules, restrictions, individuals and facts in the ontology; development of annotation patterns for the different types of informational resources, such as: iconographical objects, object presenting an iconographic technique, object presenting an iconographic school and object presenting an author. To facilitate the exploitation of the software editor a self instructor or a usage manual should be provided.

- Authors of semantic resources (annotators): the SINUS platform to provide software tool/editor for creation of semantic annotations which should ensure:
    - Access to the library with the primary digital resources so that the annotators could find and choose an appropriate resource for annotation;
    - Access to available domain ontologies, possibility to select and preview the content of the selected ontology, its categories, characteristics, rules, facts and annotation patterns for the different types of objects, as well as the possibility to choose an appropriate annotation pattern;
    - Annotation and semantic indexing of the selected resource using an appropriate annotation pattern from the chosen ontology;
    - Possibility to save the created annotation in a semantic annotation repository;
    - Possibility for multiple reusability of the existing semantic annotations – description of similar objects preserved in the semantic annotation repository;
    - To facilitate the exploitation of the software editor a self instructor or a usage manual should be provided.
- Authors of learning resources: the SINUS platform to provide software tool/editor for creation of learning resources which should ensure:
    - Access to the semantic annotation repository and the possibility to search appropriate resources so as to be grouped for the creation of learning resources;
    - Possibility for arrangement of the selected digital resources and inclusion of explanatory text commentaries (notes – introductory, intermediate, concluding; analysis, etc.);
    - Possibility to indicate which descriptors of each digital resource in the platform should be visualized, e.g. author, technique used, present location of the artifact;
    - Possibility to indicate the level of appropriateness of a digital resource for the different types of users according to the degree of detail of the digital resource description;
    - Annotation according to the LOM standard of the resource created;
    - Possibility for multiple use of the existing LOM annotations – description of similar resources preserved in the learning resource repository;
    - Testing of the visualization feature in a real-use mode;
    - Possibility to choose a visualization pattern for the learning resource created;
    - Possibility to edit a selected visualization pattern;
    - Publishing of the selected learning resource in the learning resource repository;
    - Possibility to edit a selected learning resource, its content and LOM annotations;
    - Search and preview of the existing learning resources;
    - Identification of the status of the learning resource – finalized or in edit mode.

The users of learning resources expect:

- The SINUS platform to enable their interaction with the learning content: semantic search of web resources, preview and selection of learning resources, creation of collections of learning resources chosen by the user himself; inclusion of commentaries and notes to selected learning resources, sharing of selected learning resources with other users, etc.

Other requested functionalities – information processes to be achieved:

- Registration and profiling of the users
- Administration of the users, rights, etc;
- Attractive visualization of the resources;
- Navigation of the resources;
- Communication between the users, etc.;

A Real Situation to Demonstrate the Scenario for Exploitation

Title of the learning situation: Development of the project The Iconography of Christ in the Historical territories of Bulgaria.

Learning domain: East-Christian culture and art

Primary source of digital objects for the learning domain: the multimedia digital library "Virtual Encyclopedia of the East-Christian art" '

Users: developers of various resources by means of the SINUS platform and users of the learning resources.

General scenario situation: Professor Ivanov is a lecturer at the National Art School. He is delivering a course on Iconography for students of different classes of the Wall painting and art history departments and students of the Faculty of theology of Sofia University. To all his course members he has given the task to prepare a project on the *Iconography of Christ in the Historical territories of Bulgaria.* The learners are expected to make:

- analysis of the theological meaning of the Iconography of Christ;
- art critic analysis of the development of Christ's image in the different iconographical schools in Bulgaria;
- study of the main iconographic techniques used in the historical territories of Bulgaria
- depiction of an icon of Christ or of part of wall painting of one of the Lord's feasts.

This project supposes division of the students in several teams according to their interests:

- theological team
- art critic's team
- art technique's team
- artistic team

Preliminary preparation: Professor Ivanov and his team should prepare and provide to all of his students access to some additional learning resources of the East-Christian art domain for the implementation of this project. To enable a complex semantic-based search of these objects the team of professor Ivanov should provide their semantic descriptions according to the specifics of the approved domain terminology.

Version 1 of the learning situation:

Users: Developers of various resources by means of the SINUS platform and users of the learning resources

Use case 1: Creation of learning resources

Users: Authors of semantic digital resources, authors of learning resources

Description of use case 1: The team of Prof. Ivanov has access to the *Virtual Encyclopedia of East-Christian Art* through the semantic services of the SINUS platform that provides a wide range of primary annotated digital resources – digital copies of icons, iconographic objects of various iconographic schools, authors, periods of creation, iconographic technique etc. The selection of multimedia objects (images, text etc.) is made by means of the search tools of the digital library. The next step of the developers team is the semantic description of the selected objects by means of the domain ontology developed (*Ontology of the East-Christian art*) and its annotation patterns. The semantic annotations of the digital resources are kept in a specific repository for semantic annotations. The authors of learning objects have access to this repository and use the services for searching, visualization, adding, deleting and editing of the digital objects. The authors of learning resources are annotating each selected digital object according to the LOM standard, describing the learning content and the learning situation in the context of the SINUS project.

Use case 2: Creation of semantic resources

Users: Authors of semantic resources

Description of use case 2: The semantic resource to be developed by Prof. Ivanov and his team by means of the SINUS platform is the domain ontology – *Ontology of the East-Christian Art*. Prof. Ivanov's team is responsible only for its formal description in ontological language. The conceptual level of the ontology is developed by a wide team of domain specialists.

Use case 3: Finding of appropriate learning resources

Users: Users of learning resources

Description of use case 3: In the frames of the proposed project Prof. Ivanov sets to the different workgroups the following concrete tasks:

1. to make an analysis of the theological meaning of the iconography of Christ (the theological team);
2. to make an art critic's analysis of the chronological development of the iconography of the iconography of Christ in the different iconographic schools in Bulgaria (art critic's team);
3. to examine the main iconographic techniques used in the best Bulgarian examples of iconography of Christ (the art technique team);
4. to make an icon of Christ or a part of a mural painting depicting one of the Christ's feasts (art team);

The members of the different teams are expected to make a number of semantic based searches in order to prepare their own analysis. For example:

- *General task:* Find all the iconographic artifacts in the SINUS platform containing the image of Christ.

  Main goals:
  - Studying the main iconographic scenes containing the image of Christ, so that it is needed its theological meaning to be analyzed and/or
  - Studying the theological meaning of the various symbols and signs in the different iconographic scenes.

- *General task:* Find iconographical artifacts, containing the image of Christ, from different periods.

  Main goal:

  - Defining the specific periodization in the depiction of the image of Christ (having in mind that it is independent and does not coincide with the periodization by centuries)
  - This goal may be achieved by repeated search and selection of objects.

    Step 1: One should start with *search* and *preview* of objects by/from century/centuries;

    Step 2: *Search* and *grouping* according to the iconographic schools should be made in the resulting set;

    Step 3: Search and grouping of the objects according to *part of century* they belong to, follows; Afterwards one should select iconographic objects similar in style.

    Step 4: Presentation of the search results and grouping of the objects according to the new periodization.

- *General task:* Find all the iconographic artifacts in the SINUS platform containing the image of Christ and compare their specifics from technological point of view.

  Main goals:

  - Comparison of different iconographic techniques, study on the periods of prevalent use of the different techniques, estimation of the performance quality, etc.

    This goal may be achieved by the finding of artifacts made using a single technique, others – with combined techniques. One should then perform searching and grouping in the resulting sets of objects according to the time frame and the iconographic school. One could also group the objects according to the different technological characteristics, such as: base, couch, varnish.

Version 2 of the learning situation:

Users: developers of various resources by means of the SINUS platform and the consumers of those learning resources

Version 2 of the learning situation includes *Use case 1* and *Use case 3* of Version 1. The following extension is added to *Use case* 2:

Use case 2.1.: Creating a semantic resource for profiling of the learning resources consumer

Users: authors of semantic resources

Description of use case 2.1: Professor Ivanov and his team develop an ontology describing learners' educational level in order to enable the automatic setting of tasks to the learners according to their educational level, experience and interests.

Use case 2.2.: Determination and integration in the platform of rules for automatic comparison of certain concrete tasks for comparative analysis of learning resources consumers with a certain profile.

Users: authors of learning resources

Description of use case 2.2: Professor Ivanov and his team include in the platform rules for comparison of concrete tasks with a relative user profile. These rules and concrete tasks are determined on the bases of the team's professional experience in the learning domain and more specifically in the work with students with different capacity, knowledge, learning goals, etc.

<u>Use case 2.3.:</u> Selection and inclusion in the platform of concrete tasks in the frames of the project *the Iconography of Christ in the Historical territories of Bulgaria.*

Users: authors of learning resources

Description of use case 2.3: Professor Ivanov and his team include in the platform a number of concrete tasks in the frames of the proposed project (see the examples in Use case 3 of Version 1 of the learning situation)

<u>Use case 2.4.:</u> Profiling of learning resources users and providing concrete tasks in the frames of the project *the Iconography of Christ in the Historical territories of Bulgaria*

Users: users of learning resources

Description of use case 2.4: The users register themselves in the platform and fill up a questionnaire, so that their profile to be specified according to the *User profiles ontology.* Each user receives as a result a concrete task from the platform corresponding to a greater extent to his profile

Version 3 of the learning situation

Users: developers of various resources by means of the SINUS platform and the consumers of learning resources

Version 3 of the learning situation includes *Use case 1* and *Use case 2* of Version 1

<u>Use case 3:</u> Learning course development

Users: authors and users of learning resources

Description of use case 3: The analysis prepared by the theological, art critic's and art technique teams should be provided to the artistic team as instructions for the examination of the final task in the project - depiction of an icon of Christ or of a part of wall painting of one of the Lord's feast. The students, helped by professor Ivanov and his team, prepare a complete learning module consisting mainly of the thematic teams' analysis developed within the platform.

Conclusion

The appearing new generation of information technologies is gradually alienated from the software as a basic term and starts to consider all the information resources as services in service-oriented architectures. What is meaningful for the user in the world of informational services is the service itself and not the software and hardware components behind it. The development of a use case scenario an determination of the specific requirements for the different components of the SINUS platform are supporting considerably the design of such a service-oriented architecture, as well as the determination of the functional specifications of the services provided by the platform, the needed informational services, methods and operations on the different levels of exploitation, the users and their activities, etc. One could plan on the basis of the use case scenario the testing and the evaluation of the platform and its components, as well as the functionality and the multiple reusability of the selected project approach.

Acknowledgements

This work is partly funded by Bulgarian NSF under the project D-002-189 SINUS "Semantic Technologies for Web Services and Technology Enhanced Learning".

This work is partially supported by Project BG051PO001/07/3.3-02/7 as a part of the grant scheme "Support for the Development of PhD Students, Post-doctoral Students, Post-graduate Students and Young Scientists" under the Operational programme "Human Resource Development" of the European Social Fund and the Bulgarian Ministry of Education and Science.

Bibliography

[Agre and Dochev, 2008] Agre G., D. Dochev - An Approach to Technology Enhanced Learning by Application of Semantic Web Services, International Journal "Cybernetics and Information Technologies", Vol. 8 (2008), № 3, pp. 60-72.

[Arapi et al., 2007] Arapi P., N. Moumoutzis, M. Mylonakis, G. Stylianakis, G. Theodorakis. Supporting Personalized Learning Experiences within the LOGOS Cross-Media Learning Platform, In the Proceedings of the Open Workshop "Cross-Media and Personalized Learning Applications on top of Digital Libraries" (LADL 2007) in conj. with the 11th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2007), 16-21 September, 2007, Budapest, Hungary, pp. 5-18.

[Arroyo, 2006] Arroyo, S. A Semantic Service-Based Micro-Learning Framework, In: T. Hug, M., Lindner, P. A. Bruck (Eds.) Micromedia&e-Learning 2.0: Gaining the Big Picture. Proceedings of Microlearning Conference 2006, Innsbruck University Press, pp. 260-272.

[Dochev and Pavlov, 2009] Learning Content for Technology Enhanced Learning – Experiments and Solutions, International Conference on e-Learning and the Knowledge Society - e-Learning'09, Berlin, Germany, September 2009.

[LADL 2007] Christodoulakis S. (Ed.). In the proceedings of the Workshop on Cross-Media and Personalized Learning Applications on top of Digital Libraries (LADL2007), Budapest, Hungary, September 2007, 110 p., Available at: http://ladl2007.cc.bas.bg/present/LADLWorkshopNotes_Final.pdf

[LOGOS Deliverable D3, 2007] Deliverable D3 "LOGOS project platform", project LOGOS "Knowledge-on-demand for Ubiquitous Learning", Available at: http://logosproject.com

[Paneva-Marinova et al., 2009] Paneva-Marinova D., L. Pavlova-Draganova, L. Draganov, R. Pavlov, M. Sendova. Development of a Courseware on Bulgarian Iconography for Ubiquitous On-demand Study. In: Szucs A. (Ed.) Proceedings of Open Conference "New Technology Platforms for Learning – Revisited". Budapest, Hungary, January 2009, pp. 37-46.

[Paneva-Marinova et al., 2008] Paneva-Marinova, D., L. Pavlova-Draganova, R. Pavlov, M. Sendova. Cross-media and Ubiquitous Learning Applications on Top of Iconographic Digital Library. In Proceedings of the 14th International Conference on Virtual Systems and Multimedia, Limassol, Cyprus, 20-25 October 2008, pp. 367-371.

[Paneva, 2006] Paneva D., Ontology-based Student Modeling, In the Proceedings of the Open Workshop "Ubiquitous Learning Challenges: Design, Experiments and Context Aware Ubiquitous Learning", 20 - 21 September, 2006, Turin, Italy, pp. 17-21

[Pavlov and Paneva, 2007] Pavlov R., D. Paneva. Toward Ubiquitous Learning Application of Digital Libraries with Multimedia Content, International Journal "Cybernetics and Information Technologies", vol. 6 (2007), № 3, pp. 51-62.

[Pavlova-Draganova et al., 2007a] Pavlova-Draganova L., V. Georgiev, L. Draganov. Virtual Encyclopaedia of Bulgarian Iconography, Information Technologies and Knowledge, vol.1 (2007), №3, pp. 267-271.

[Pavlova-Draganova et al., 2007b] Pavlova-Draganova, L., D. Paneva, L. Draganov. Knowledge Technologies for Description of the Semantics of the Bulgarian Iconographical Artefacts, In the Proceedings of the Open Workshop "Knowledge Technologies and Applications", Kosice, Slovakia, 31 May - 1 June, 2007, pp. 41-46.

[Schreiber et al., 2000] Schreiber, G., Akkermans, H., Anjewierden, A., De Hoog, R., Shadbolt, N., Van De Velde, W. & Wielinga, B. Knowledge Engineering and Management - The CommonKADS Methodology, MIT Press, 2000.

[Staykova et al., 2007] Staykova, K., D. Dochev, D. Paneva, L. Pavlova-Draganova, V. Saraydarova, Development of Domain Ontology, Targeted at the Creation of Learning Materials From Digital Archives, In the Proceedings of the Open Workshop "Cross-Media and Personalized Learning Applications on top of Digital Libraries" (LADL 2007) in conj. with the 11-th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2007), Budapest, Hungary, 16-21 September, 2007, pp. 91-100.

Information on the Authors

Lilia Pavlova-Draganova *– Assistant Professor, PhD Student, Laboratory of Telematics, BAS, Acad. G. Bonthev Str., bl. 8, Sofia-1113, Bulgaria; e-mail: lilia@cc.bas.bg*

Major Fields of Scientific Research: Multimedia Digital Libraries, East Christian Iconography, 2D and 3D Computers Graphics and Animation, Design and Modelling Cultural Heritage Objects, Preservation and Restoration of Digitalized Artifacts, eLearning.

Desislava Paneva-Marinova *– PhD in Informatics, Assistant Professor, Institute of Mathematics and Informatics, BAS, Acad. G. Bonthev Str., bl. 8, Sofia-1113, Bulgaria; e-mail: dessi@cc.bas.bg*

Major Fields of Scientific Research: Multimedia Digital Libraries, eLearning Systems and Standards, Knowledge Technologies and Applications, Personalization and Content Adaptivity in Multimedia Digital Libraries.

Lubomil Draganov *– Assistant Professor, PhD Student, Institute of Mathematics and Informatics, BAS, Acad. G. Bonchev Str.., bl. 8, Sofia 1113, Bulgaria; e-mail: lubomil@gmail.com*

Major Fields of Scientific Research: Multimedia Digital Libraries, East Christian Iconography, 2D and 3D Computers Graphics and Animation, Design and Modelling Cultural Heritage Objects, Preservation and Restoration of Digitalized Artifacts.

MODERN BUSINESS MODELING APPROACHES AND TOOLS FOR MANAGEMENT

Elena Serova

***Abstract**: Improving management systems is getting more and more important and topical, as the organizational structure of modern society consists of many features, peculiarities, relations, and is continuously getting more intricate. Although the society has been complex for a long time, it is only now that we are starting to comprehend its real complexity. It has become obvious that changing a characteristic can easily cause or require change to other parts of the system. Attempting to achieve an enterprise's highest efficiency, each modern executive should use computer modeling from time to time, as it is an efficient tool for management system research that weighs qualitatively and quantitatively characteristics of system functioning. A prerequisite of modern management is to know the principles and capabilities of modeling, be able to create research and use such models. This report considers the role of business modeling for modern management tasks. Generalizing theory development in the area along with international practices and domestic experience, the author chooses the main directions of modeling for modern management and those fields of management where applying computer modeling is more efficient.. This paper weighs perspectives of modeling used to solve business problems; considers modern modeling approaches used to describe architecture, develop operational models and carry out reengineering; briefs on existing techniques and instrumental tools used in modern business modeling.*

***Keywords**: business process, business modeling, business process modeling techniques, simulation modeling, structure-function approach, discontinuous event-driven approach, agent modeling, modeling tools.*

***ACM Classification Keywords**: I.6.5 Model development – Modeling methodology*

Introduction

Business modeling techniques and tools have proven their usefulness. Development and implementation of ERP (Enterprise Resource Planning Systems), decision making support complexes, consultancy on describing organization architecture, changing business processes, auditing and certifying operations are just a few examples of how these tools can be used. To successfully complete such projects, modern business modeling approaches and tools are indispensable.

When solving business problems, modeling tools are primarily used to ensure mutual understanding at every organizational level, bridge the gap between strategic vision and its implementation. To do that, modern business modeling tools use special software, languages and systems that help develop models and diagrams to demonstrate how business processes are built and how staff interaction is organized, and what needs to be changed to optimize the whole architecture.

Computer modeling allows for considering processes that run in a system at any level of detail. Almost any algorithm of managerial activities or system behavior can be modeled. In addition, models that can be researched with analytic methods can be analyzed with simulation methods as well. These are the reasons why computer modeling methods are becoming a principal research method for complex management systems.

Companies that actively use cutting-edge information technologies consider modeling as a stage of executive decision making [Lychkina, 2007]. They provide their managers with systems that help make strategic executive decisions. In addition, computer modeling based tools use methods and advantages such as object-oriented programming, video and multimedia supporting real-time animation.

Computer modeling allows for describing complex nonlinear business interactions, e.g., modeling economic agents' behavior in a crisis situation, weighing effects of different scenarios or forecasting further stream of events. The essence of computer modeling in business is to get quantitative and qualitative results from the existing model. By receiving from the analysis of business processes (structure-function modeling), qualitative results allow for finding previously unknown features of a complex system (e.g., management one): structure, development trends, sustainability, integrity, etc. Most quantitative results help forecast certain future values of variables characterizing the real system that is researched or explains those from the past, and can be obtained with modern simulation modeling techniques described in this paper. Naturally, all the modeling methods used to solve modern business problems are not mutually exclusive and can be applied to management system research either simultaneously or in a combination.

Goals and Objectives of Computer Modeling for Management Task

Computer models of complex management systems should show all major factors and correlations characterizing real situations, criteria and limitations. Models should be universal enough to describe objects close in application, simple enough to allow research needed at reasonable cost, and allow achieving the following objectives:

- to abolish a series of functions and reduce the number of management levels, to disengage mid-level workers;
- to rationalize solving management problems by implementing mathematical methods of data processing, using simulation modeling and artificial intelligence systems;
- to create a modern, dynamic organizational structure, improve the enterprise's flexibility and manageability;
- to reduce administrative costs;
- to reduce time spent to plan activities and make decisions;
- to increase competitive advantages.

To make the role of computer simulation modeling in the modern management more clear, I have to mention application of the structure-function approach to solving business problems. The essence of computer modeling in business is to get quantitative and qualitative results from the existing model. Qualitative results allow for finding previously unknown features of a complex system (e.g., management one): structure, development trends, sustainability, integrity, etc. Most quantitative results help forecast certain future values of variables characterizing the system researched or explain those from the past.

It is an essential difference of the computer simulation modeling from the structure-function one that the former gives both qualitative and quantitative results.

There is another direction of computer modeling. It solves management problems with mathematics and logic and, as a rule, uses Excel spreadsheets. The problems solved are those of stock management as well as transport, industrial or marketing logistics [Gorshkov et al., 2004]. The same is done with problems of linear and multiple regression forecasting, resource utilization review, etc. Such tools are quite popular, although specific management software, both scientific and commercial, that uses structure-function and simulation approaches, is more perspective.

Naturally, all the modeling methods listed above: simulation, mathematical logic and structure-function—are not mutually exclusive and can be applied to management system research either simultaneously or in a combination. Modeling tool Bpwin [Maklakov, 2003] can export models into a most efficient simulation modeling tool—the Arena system developed by Rockweel Automation (http://www.arenasimulation.com), and allows for optimizing business processes with simulation modeling (Fig.1). Using such an approach, various processes can be simulated and optimized: industrial technological operations, inventory control, banking, restaurant services, etc.

This is an example of how two leading directions of computer modeling can be integrated to solve modern management problems, demonstrating how simulation modeling can be applied to get quantitative results when modeling business processes.

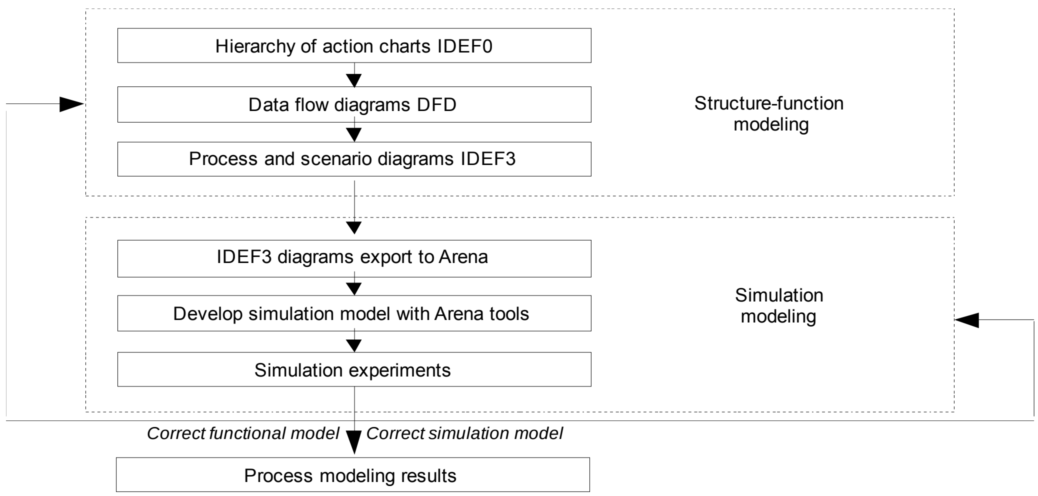


Fig. 1. Integrating simulation and structure-function modeling when solving modern business problems

Structure-Function Approach to Business Modeling

The most intuitive and quite popular example of the structure-function computer modeling in modern management is the business process modeling.

The market situation modern companies operate in is quite unstable which makes them respond to change quickly and accurately. Sooner or later, businesses have to restructure, and managers have to think how to change the existing business processes in order to improve the enterprise's operations. Thus, a manufacturer may wish to reconsider purchasing, ordering or delivery. Business process reengineering is tied together with altering the architecture of information systems. The key to success of a reorganization project is close cooperation of all the groups interested in solving the problem, primarily IT specialists and experts in the business area. It is achieved by building structure-function computer models that reflect business processes which are comprehensible for all participants. Such models should simultaneously help formalize and document the current state of affairs and find room for improvement. There are several computer technologies aimed at automating structure models—the CASE (Computer Aided Software Engineering) tools. The definition of CASE involves various tools used to analyze and model, and business process analysis tools are just a small fraction of the class.

Organization and structure changes in a company, especially when they involve an ERP implementation, bring serious risks. Implications of such changes should be carefully studied and analyzed before they start. Such ERP as SAP R/3, BAAN, ROSS iRenaissance, etc. use methods and tools proven by extensive experience and allow minimize risks and resolve issues arising from reorganization of business processes, including those linked with implementation of modern IT systems.

Today's approach to business process description suggests continuous improvement and modification, analysis and prognosis, as well as timely changes to the business model. The description should adequately reflect current state of affairs to underlie an integral comprehension of business development strategy and business automation. The following series of steps is best for business development or modification [Proshin, 2006]:

1. ***To define (correct) corporate strategy and major business goals*.**
2. ***To create a model of organizational architecture, incl. models of business processes (or to modify the existing)***
3. ***To weigh efficiency of the processes based on the business goals.***
4. ***To design information system architecture based on the business process model for the efficiency parameters of the company and its processes***
5. ***To form an implementation or modification plan for the information system***
6. ***To develop detailed requirements to the system based on the business process models***

Fig. 2. Business development (modification) steps

There are several techniques to describe and model business processes. The most popular are: Business Process Modeling, Work Flow Modeling and Data Flow Modeling [Repin, Yeliferov, 2008].

Suggested in the 1970s by Douglas Ross, the Structured Analysis and Design Technique (SADT) underlies the IDEF0 business process modeling standard. AllFusion Process Modeler 4.1 (BPwin 4.1) by Computer Associates (CA) is a modeling tool fully compliant with IDEF0 that allows analyzing, documenting and planning changes in complex business processes [Maklakov, 2003].

Another actively used process description methodology is the Work Flow Modeling—the IDEF3 standard to build process models as time sequences of jobs (functions, operations). The IRIS source environment by IDS Scheer AG that creates methodological and work instructions with eEPS (extend Event-driven Process Chain) models, is based on IDEF3 [Ilyin, 2006].

DFD (Data Flow Diagramming) notations allow reflecting job sequences within a process and information flows circulating between those jobs. The DFD methodology minimizes subjectivity of business process description and can be efficient when implementing process approach to organizational management.

The developing UML (Unified Modeling Language) methodology is also quite widely used. It considers a series of diagrams (e.g., the Activity Diagram) that can be used to describe business processes [Vendrov, 2000], although business modeling is not what UML is intended for.

Along with the techniques listed above, there are other ones offered by various software producers. Even such corporations as IBM and Oracle offer their own business process description and modeling tools. E.g., the Oracle Workflow technology used to automate job flows contains tools for process description and formalization. The most popular state-of-the-art business process management standard is BPEL (Business-Process Execution Language). Based on this product, an integral platform for all applications used can be created. Public and private institutions throughout the world are switching to BPEL. Certain pilot projects have been carried out in Russia as well, successfully solving IT infrastructure optimization problems [Proshin, 2006].

Simulation Modeling to Solve Business Problems

The structure-function method allows describing existing business processes, finding their drawbacks and building a model of the enterprise's operations. However, the difficulty is the optimization of particular processes, or the study of how various parameters influence a certain business process. To solve this problem, the structure-function model may be insufficient, and other modeling techniques and tools turn out to be more appropriate. An approach that solves such business problems and gives quantitative characteristics of business processes is simulation modeling. Simulation models can provide statistics of processes as if they were happening in reality. Normally, such models are built to find an optimum solution with limited resources, when other mathematical models are too complex. Owing to its simplicity, the idea of simulation modeling attracts both executives and system researchers. The simulation approach to business problems requires special software that is widely denoted with such terms as “simulation system” and “simulation modeling system”. The terms refer to an aggregate of a simulation model of a complex process, a set of simpler models of the same process, algorithms and relevant software associated with the models. Some examples of such systems applied to business modeling are the Arena simulation modeling system by Rockweel Automation (http://www.arenasimulation.com), AnyLogic by XJ Technologies (http://www.xjtek.com) or GPSS (General Purpose Simulation System) by Minuteman Software (http://www.Minutemansoftware.com). To create simulation models, one should know special algorithmic languages that can express concepts which modeling specialists use. Each language is specific in:

- how complex the concepts of simulation modeling are represented;
- language base;
- number of basic concepts.

An important factor to choose a simulation modeling language is if there is an efficient translator for the chosen computer hardware. A multifunctional user interface makes many language operators excessive. This is why a special simulation modeling language is ideal to build a simulation model for business problems.

The modern simulation modeling theory offers four major approaches [Borshchev, Filippov, 2006]:

- dynamic system modeling (simulation modeling systems MATLAB Simulink, VinSim, etc.),
- discontinuous event-driven modeling (GPSS, Arena, eMPlant, AutoMod, PROMODEL, Enterprise Dynamics, FlexSim, etc.) [Serova, 2007],
- system dynamics (VenSim, PowerSim, iThink, etc.), and

- agent modeling (AnyLogic [Karpov, 2005], Swarm, Repast, etc.).

Each direction develops its own tools, simulation modeling systems and languages.

System dynamics (SD) and dynamic systems are traditional, established approaches; whereas the agent modeling (AM) is comparatively new. SD and dynamic systems operate mostly with continuous processes, while the event-driven and agent modeling cover discontinuous ones.

The following two approaches are used most often to solve business problems as basic formalization and structuring conceptions in modern simulation modeling systems:

- process and transaction oriented modeling systems based on process description. In the modern IT market, they represent the discontinuous event-driven simulation modeling approach and are the most representative class of such systems. These are such systems as GPSS, Arena, Extend, AutoMod, ProModel, Witness, Taylor, eM-Plant, QUEST, SIMFACTORY II.5, SIMPLE++, etc. [Serova, 2007];
- agent modeling that uses models to study decentralized systems which dynamics and functioning is not defined by global rules, on the contrary, those rules are a result of the group members' individual activities. In Russia, such systems are represented by AnyLogic [Karpov, 2005];

Simulation modeling systems with discontinuous event-driven and agent approaches have proven most efficient in such areas of business modeling as business process and service modeling. The Arena simulation modeling system is integrated with a CASE tool, BPWin, whereas GPSS possesses tools needed to model processes relevant to such a dynamically developing area as the service-oriented economy [Serova, 2007].

Conclusion

Aiming at securing a stable economic position in a very competitive environment and attracting funding, the most forward-looking companies pay more and more attention to developing and implementing cutting-edge computer modeling systems. A principal tool to solve modern business problems related to cost cutting and restructuring, business modeling, service-oriented economy and decision-making procedures in management systems, is computer simulation modeling technologies. They include developed graphic interfaces for model construction, result presentation and output statistics filing. Moreover, simulation modeling widely uses methods and advantages of object-oriented programming, video and multimedia supporting real-time animation.

State-of-the-art simulation modeling technologies used at every level of enterprise management: strategic, tactic and operational—is a direction for development and a criterion for stable economic growth in the modern competitive environment.

Thus, an organization willing to use state-of-the-art business modeling tools can choose a methodology out of several standard options. The choice should be based on the clear understanding of the models' capabilities and drawbacks, as well as the purposes. Business modeling tools and instruments are evolving and tend to switch from a visual description of certain narrow business areas towards a holistic description of organizational architecture. The application of modeling is widening from information exchange within a small group of specialists to management of distributed organizations requiring comprehensive information of all the organization's operations. The opportunities of integration between different business modeling approaches that have appeared allow for fully implementing modeling and analysis tools into the organization's existing infrastructure. The most perspective direction seems to be the growing correlation of business modeling and analysis systems with management systems.

Bibliography

[]Ananiev, Serova, 2008] I. Ananiev, E. Serova. The areas of IDEFO notation effective application for tasks of business-processes description. Vestnik of St. Petersburg State University, 2008.

[Borshchev, Filippov, 2004] A. Borshchev, A. Filippov. AnyLogic — Multi-Paradigm Simulation for Business, Engineering and Research. The 6th IIE Annual Simulation Solutions Conference, Orlando, Florida, USA.March 15-16, 2004.

[Borshchev, Filippov, 2006] A. Borshchev, A. Filippov. From System Dynamics and Discrete Event to Practical Agent Based Modeling/ site by XJ Technologies., 2006.

[Gavrilova, 2007] T. Gavrilova, S. Puuronen. In search of a vision: ontological view on user modeling. KDS-2007, conferences' score.

[Gorshkov et al., 2004] A.F. Gorshkov, B. V. Yevteyev, V. A. Koshunov. Kompyuternoe modelirovanie menedzhmenta (Computer modeling in management). – Moscow: examen, 2004.

[Grabaurov, 2001] V.A. Grabaurov. Informatsionnye tekhnologii dlya menedzherov (IT for managers). – Moscow: Finansy i Statistika (FIS), 2001.

[Ilyin, 2006] V.V. Ilyin. Modelirovanie biznes-protsessov. Prakticheskiy opyt razrabotchika (Business process modeling. A developer's practical experience). – Moscow: Williams Publishing House, 2006.

[Karpov, 2005] Yu. Karpov. Imitatsionnoe modelirovanie sistem. Vvedenie v modelirovanie s AnyLogic (System simulation modeling. Introduction to modeling with AnyLogic). – SPb.: BHV-Petersburg, 2005.

[Lychkina, 2007] N.N. Lychkina. Imitatsionnye modeli v protsedurakh i sistemakh podderzhki prinyatiya strategicheskikh resheniy na predpriyatiyakh (Simulation models in enterprises' strategic decision making support procedures and systems). Biznes-Informatika, No. 1, 2007.

[Maklakov, 2003] S.V. Maklakov. Modelirovanie biznes-protsessov s AllFusion Process Modeler (Business processes modeling with AllFusion Process Modeler). – Moscow: Dialog MEPHI, 2003.

[Moshella, 2004] David Moshella. Customer-driven IT: how users are shaping technology industry growth. -- Boston, MA: Harvard Business Press, 2003.

[Proshin, 2006] F. Proshin. Biznes-modelirovanie: zadachi i instrumenty (Business modeling: problems and tools). IT News. 2006. http://www.olap.ru/home.asp?artId=295

[Repin, Yeliferov, 2008] V.V. Repin, V.G. Yeliferov. Protsessnyy podkhod k upravleniyu. Modelirovanie biznes-protsessov. Prakticheskiy menedzhment (Process approach to management. Business process modeling. Practical management).- 6th ed.– Moscow: Standarty i Kachestvo, 2008.

[Serova, 2007] E. Serova. Imitatsionnoe modelirovanie v sovremennom menedzhmente (Simulation modeling for modern management). SPb.: SIMMOD-2007.

[Turban, 2006] E.Turban, D. Leidner, E. McLean, and J. Wetherbe. Information Technology for Management. – WILEY, 2006.

[Vendrov, 2000] A.M. Vendrov. Proektirovanie programmnogo obespecheniya ekonomicheskikh informatsionnykh sistem (Designing software for economic information systems). Moscow: Finansy i Statistika (FIS), 2000.

GPSS World reference manual. Fourth Edition 2001. Copyright Minuteman Software. Holly Springs, NC, U.S.A. 2001; http://www.minutemansoftware.com/reference/reference_manual.htm

GPSS World Tutorial Manual. Copyright Minuteman Software. Holly Springs, NC, U.S.A. 2001. http://www.minutemansoftware.com/tutorial/tutorial_manual.htm

Authors' Information

Elena Serova *– PhD (Economics), Deputy Chair, Dept. of IT in Management, Graduate School of Management, St. Petersburg State University; Per. Dekabristov 16, St. Peterburg 199155, Russia; e-mail: serovah@som.pu.ru; serovah75@gmail.com*

A MODEL FOR THE UNIVERSITY COURSE TIMETABLE PROBLEM

Velin Kralev

***Abstract**: In this paper a model of the university course timetabling problem is presented. The basic data structures, parameters, vectors and matrices, which are used in the definition of the problem, are presented. The hard constraints, which include checking simultaneous involvement of lecturers, students or auditoriums, checking for exceeded capacity of auditoriums as well as verification of commitment of students and lecturers exceeding the predetermined number of hours, are considered. Objective functions for each of the soft constraints, which will be united in a common objective function, are defined. Thus the problem under investigation, will acquire multicriteria optimization nature. The future trends of work are presented.*

***Keywords:** timetable problem, multicriteria optimization.*

Introduction

In solving the university course timetabling problem, all hard constraints which are defined must necessarily be fulfilled. If in finding a solution, even one of these constraints is violated, then the whole solution is pronounced as unacceptable and needs searching for another solution. The solution of problem is limited to distribution of events in the week schedule. It is necessary to meet all hard constraints, while at the same time, be violated at least soft constraints. When defining the problem discussed in this paper, four hard and three soft constraints are defined. Although there are other examples of definitions of this problem (Burke at al., 2004; Carter and Laporte 1998, Rossi-Doria, at al., 2002, 2003; Socha at al., 2002), we propose a new model where events may be of varying lengths, in addition, each resource and its interactions will participate with a corresponding weight.

In our model a matrix of the preferences of lecturers and a common matrix of the preferences of students are provided. They combine some of the soft constraints of two data structures. If necessary it is also possible to introduce a similar matrix of the auditorium.

The advantages of our model, which has been discussed here are three:

- First, in the problem definition all the resources "lecturers", "students", "auditoriums" and "events" are included. Every element of the resource "events" represents the interaction of one or more elements from the three other resources that occur within a specified timeslot. A variant of the problem which is often considered is that in which some of the resources (eg "lecturers" and/or "auditoriums") are not taken into consideration. This is done in order to compile a model more easily and to create algorithms to solve the problem.
- Second, maintaining blocks of consecutive events that appear inseparable.
- Third, the soft constraints are combined into three groups - a) a minimum length of timetable; b) a minimum number of violations according to the lecturers' preferences and c) a minimum number of violated students' preferences. For each group an objective function with a certain weight is defined.

Basic Definitions

Basic data structures, parameters, vectors and matrices, which will be used to create the model will be presented.

Vector of events (EV) contains information about events (or activities). In our model, the events are marked with unique numbers from 1 to N:

(1) $EV = \{n_i\}_{i=1}^{N} = \{n_1, n_2, \dots, n_N\}$, where n_i is the *i*-th event.

Also a vector of the duration of events (EDV) and a vector of weights of events (WEV) are introduced, which contain additional information about each event:

(2) $EDV = \{n_i^d\}_{i=1}^{N} = \{n_1^d, n_2^d, \dots, n_N^d\}$, where n_i^d is the duration of the *i*-th event and

(3) $WEV = \{n_i^w\}_{i=1}^{N} = \{n_1^w, n_2^w, \dots, n_N^w\}$, where $n_i^w \in (0,1]$ is the weight of the *i*-th event.

The weight of each event is determined based on subjective factors and indicates the relative weight of an event to other events.

Students vector (SV), lecturers vector (LV) and auditoriums vector (AV) are with length respectively S, L and A. In our model, students, lecturers and auditoriums are identified by unique numbers, respectively, from 1 to S from 1 to L and from 1 to A:

(4) $SV = \{s_i\}_{i=1}^{S} = \{s_1, s_2, \dots, s_S\}$, where s_i is the *i*-th student.

(5) $LV = \{l_i\}_{i=1}^{L} = \{l_1, l_2, \dots, l_L\}$, where l_i is the *i*-th lecturer.

(6) $AV = \{a_i\}_{i=1}^{A} = \{a_1, a_2, \dots, a_A\}$, where a_i is the *i*-th auditorium.

It also introduces the weight vector of lecturers (WLV), the weight vector of auditoriums (WAV) and the vector capacity of auditoriums (ACV), which contain additional information of related resources:

(7) $WLV = \{l_i^w\}_{i=1}^{L} = \{l_1^w, l_2^w, \dots, l_L^w\}$, where $l_i^w \in (0,1]$ is the weight of the *i*-th lecturer.

(8) $WAV = \{a_i^w\}_{i=1}^{A} = \{a_1^w, a_2^w, \dots, a_L^w\}$, where $a_i^w \in (0,1]$ is the weight of the *i*-th auditorium.

(9) $ACV = \{a_i^c\}_{i=1}^{A} = \{a_1^c, a_2^c, \dots, a_A^c\}$, where a_i^c is the capacity of the *i*-th auditorium.

To have the matrix of weekly schedule (WPM) defined, it is necessary first to define: the days vector (DV), a vector of hours (HV) and the vector of absolute hours (AHV), which respectively have a length D, H and T:

(10) $DV = \{d_i\}_{i=1}^{D} = \{d_1, d_2, \ldots, d_D\}$, where d_i is the *i*-th day.

(11) $HV = \{h_i\}_{i=1}^{H} = \{h_1, h_2, \ldots, h_H\}$, where h_i is the *i*-th hour in each day.

(12) $AHV = \{t_i\}_{i=1}^{T} = \{t_1, t_2, \ldots, t_T\}$, where $T = D * H$, and t_i is the *i*-th absolute hour, which corresponds to a time interval (timeslot) of the weekly schedule.

Each timeslot is given by ordered pair $\langle d_i, h_j \rangle$, where $i = 1, 2, ...D;\ j = 1, 2, ...H$. A presentation of the matrix of weekly schedule (WPM) shown in (13):

(13) $WPM = (p_{ij})_{DxH}$, where p_{ij} is the timeslot in the *i*-th day and *j*-th hour.

The matrix of students and events (SEM) has a length of $S \times N$ where each element is given by "1" or "0". Value "1" indicates the *i*-th student takes *j*-th event where $i \in \{1..S\}$, $j \in \{1..N\}$ and "0" otherwise:

(14) $SEM = (p_{ij})_{SxN}$, where $p_{ij} = \begin{cases} 1, & if\ \ student\ \ i\ \ takes\ \ event\ \ j \\ 0, & otherwise \end{cases}$

Also the vector of the number of students at events (NSEV), the vector of the number of events by students (NESV) and the vector of students' hours of work (LHSV) are defined:

(15) $NSEV = \{n_i^s\}_{i=1}^{N} = \{n_1^s, n_2^s, \ldots, n_N^s\}$, where n_i^s is the number of students attending the *i*-th event and $n_i^s = \sum_{k=1}^{S} v_{ki}$, where $v_{ki} \in SEM$.

(16) $NESV = \{s_i^n\}_{i=1}^{S} = \{s_1^n, s_2^n, \ldots, s_S^n\}$, where s_i^n is the number of events, which must be attended by the *i*-th student and $s_i^n = \sum_{k=1}^{N} v_{ik}$, where $v_{ik} \in SEM$.

(17) $LHSV = \{s_i^t\}_{i=1}^{S} = \{s_1^t, s_2^t, \ldots, s_S^t\}$, where s_i^t is the total number of hours during which the *i*-th student is engaged and $s_i^t = \sum_{k=1}^{N} u_{ik} v_k$, where $u_{ik} \in SEM$, $v_k \in EDV$.

Because each event is held by only one lecturer, that information can be represented as a vector of events and lecturers (ELV) with length L:

(18) $ELV = \left\{n_i^l\right\}_{i=1}^{N} = \left\{n_1^l, n_2^l, \ldots, n_N^l\right\}$, where $n_i^l \in LV$ is the index of the lecturer who held the i-th event.

Also a vector of the number of events by lecturers (NELV), and a vector of lecturers' hours per week (LHLV) are defined:

(19) $NELV = \left\{l_i^n\right\}_{i=1}^{L} = \left\{l_1^n, l_2^n, \ldots, l_L^n\right\}$, where l_i^n is the number of events to be held by the *i*-th lecturer and $l_i^n = \sum_{k=1}^{N} \delta(k,i)$, where $\delta(k,i) = \begin{cases} 1, \; if \; v_k = u_i \\ 0 \; otherwise \end{cases}$, where $v_k \in ELV, \; u_i \in LV$, ie the function $\delta(k,i)$ will take value "1" if the i-th lecturer holds the k-th event.

(20) $LHLV = \left\{l_i^t\right\}_{i=1}^{L} = \left\{l_1^t, l_2^t, \ldots, l_L^t\right\}$, where l_i^t is the total number of hours during which the *i*-th lecturer is engaged and $l_i^t = \sum_{k=1}^{N} \delta(k,i)\, d_k$, where $d_k \in EDV$ is the length of the *k*-th event.

To be defined the matrix for appropriate auditoriums (SAEM) it is necessary that the vector of the kind of events (EKV) and the vector of the kind of auditoriums (AKV) to be defined before it, which have respectively length N and A:

(21) $EKV = \left\{n_i^k\right\}_{i=1}^{N} = \left\{n_1^k, n_2^k, \ldots, n_N^k\right\}$, where n_i^k is the kind of *i*-th event.

(22) $AKV = \left\{a_i^k\right\}_{i=1}^{A} = \left\{a_1^k, a_2^k, \ldots, a_A^k\right\}$, where a_i^k is the kind of *i*-th auditorium.

SAEM matrix is defined by the vectors ACV, NSEV, EKV and AKV. The dimensionality of this matrix is $A \times N$. Its elements are represented by "1" and "0" indicating that the auditorium a_i where $i \in \{1, 2, ..., A\}$, respectively, may or may not be used for the event n_j, where $j \in \{1, 2, ..., N\}$:

(23) $SAEM = (p_{ij})_{AxN}$, where $p_{ij} = \begin{cases} 1, \; if \; \left(a_i^c \geq n_j^s\right) \wedge \left(a_i^k = n_j^k\right) \\ 0 \; otherwise \end{cases}$, where $a_i^c \in ACV$ is the capacity of *i*-th auditorium, $n_j^s \in NSEV$ is the number of students who visit the *j*-th event, $a_i^k \in AKV$ is the kind of the *i*-th auditorium and $n_j^k \in EKV$ is the kind of the *j*-th event.

Similar to the lecturers, a vector of events and auditoriums (EAV), which has a length N is defined:

(24) $EAV = \left\{ n_i^a \right\}_{i=1}^{N} = \left\{ n_1^a, n_2^a, \ldots, n_N^a \right\}$, where $n_i^a \in AV$ is the index of auditorium in which the the *i*-th event is held.

Similar to the lecturers, the vector of the number of events by auditoriums (NEAV) and the vector of auditorium hours per week (LHAV) are defined:

(25) $NEAV = \left\{ a_i^n \right\}_{i=1}^{A} = \left\{ a_1^n, a_2^n, \ldots, a_A^n \right\}$, where a_i^n is the number of events to be held in the *i*-th auditorium.

(26) $LHAV = \left\{ a_i^t \right\}_{i=1}^{A} = \left\{ a_1^t, a_2^t, \ldots, a_A^t \right\}$, where a_i^t is the total number of hours during which the *i*-th auditorium will be occupied.

The matrix of conflicts of events arising from students (CESM) has dimension $N \times N$ and is generated by using the vector EV and the matrix SEM. Each element $p_{ij} \in CESM$ shows the number of the students who must attend the event i and event j as well, where $i, j \in \{1, 2, ..., N\}$. CESM is a symmetric matrix to its main diagonal and its definition is shown in (27):

(27) $CESM = (p_{ij})_{NxN}$, where $p_{ij} = \begin{cases} \sum_{k=1}^{S} v_{ki} v_{kj}, & i \neq j \\ 0, & otherwise \end{cases}$, where $v_{ki}, v_{kj} \in SEM$.

Similarly, the matrix of conflicts arising from lecturers and events (CELM) and the matrix of conflicts arising from auditoriums and events (CEAM) are defined and have dimension $N \times N$ and are generated using vectors EV, NSEV, ELV and EAV. If two events i and j are in conflict because of one lecturer or one auditorium, the element value p_{ij} is the number of students who must attend both events i and j, where $i, j \in \{1, 2, ..., N\}$. CELM and CEAM matrices are symmetrical to their main diagonal and are shown in (28) and (29):

(28) $CELM = (p_{ij})_{NxN}$, where $p_{ij} = \begin{cases} v_i + v_j, & if \;\; u_i = u_j \\ 0, & otherwise \end{cases}$, where $v_i, v_j \in NSEV$, $u_i, u_j \in ELV$.

(29) $CEAM = (p_{ij})_{NxN}$, where $p_{ij} = \begin{cases} v_i + v_j, & if \;\; u_i = u_j \\ 0, & otherwise \end{cases}$, where $v_i, v_j \in NSEV$, $u_i, u_j \in EAV$.

Also the vector of the number of conflicts arising from students and events (NCESV), the vector of the number of conflicts arising from lecturers and events (NCELV) and the vector of the number of conflicts arising from auditoriums and events (NCEAV) are defined, respectively:

(30) $NCESV = \left\{n_i^{cs}\right\}_{i=1}^{N} = \left\{n_1^{cs}, n_2^{cs}, \ldots, n_N^{cs}\right\}$, where $n_i^{cs} = \sum_{j=1}^{N} 1 \; if \; v_{ij} > 0$, $v_{ij} \in CESM$.

(31) $NCELV = \left\{n_i^{cl}\right\}_{i=1}^{N} = \left\{n_1^{cl}, n_2^{cl}, \ldots, n_N^{cl}\right\}$, where $n_i^{cl} = \sum_{j=1}^{N} 1 \; if \; u_{ij} > 0$, $u_{ij} \in CELM$.

(32) $NCEAV = \left\{n_i^{ca}\right\}_{i=1}^{N} = \left\{n_1^{ca}, n_2^{ca}, \ldots, n_N^{ca}\right\}$, where $n_i^{ca} = \sum_{j=1}^{N} 1 \; if \; x_{ij} > 0$, $x_{ij} \in CEAM$.

A summary matrix of conflicts arising from students, lecturers and auditoriums and events (GCEM) has dimensionality $N \times N$. It is generated using matrices CESM, CELM, CEAM and vectors WEV, WLV and WAV. The value of the element $p_{ij} \in GCEM$ showing the cost of the conflict among the events n_i and n_j, where $i, j \in \{1, 2, ..., N\}$. GCEM matrix is symmetrical and its main diagonal is shown in (33):

(33) $GCEM = (p_{ij})_{NxN}$, where $p_{ij} = \left\lceil \left(n_i^w + n_j^w\right)u_{ij} + f\left(b_i\right)v_{ij} + g\left(c_i\right)x_{ij} \right\rceil$, $u_{ij} \in CESM$, $v_{ij} \in CELM$, $x_{ij} \in CEAM$, $n_i^w \in WEV$, $n_j^w \in WEV$, $f(b_i) = l_k^w$, $l_k^w \in WLV$, $g(c_i) = a_q^w$, $a_q^w \in WAV$.

Also the vector of the number of conflicts arising from students, lecturers, auditorium and the events (NGCEV), the vector of the weights of conflicts arising from students, lecturers, auditorium and the events (GWCEV) and the vector for the total length by conflicting groups (GDCGV), which have a length N are defined:

(34) $NGCEV = \left\{n_i^{cg}\right\}_{i=1}^{N} = \left\{n_1^{cg}, n_2^{cg}, \ldots, n_N^{cg}\right\}$, where $n_i^{cg} = \sum_{j=1}^{N} 1, \; if \; v_{ij} > 0$, $v_{ij} \in GCEM$.

(35) $GWCEV = \left\{n_i^{cw}\right\}_{i=1}^{N} = \left\{n_1^{cw}, n_2^{cw}, \ldots, n_N^{cw}\right\}$, where $n_i^{cw} = \sum_{j=1}^{N} v_{ij}$, $v_{ij} \in GCEM$.

(36) $GDCGV = \left\{n_i^{cdg}\right\}_{i=1}^{N} = \left\{n_1^{cdg}, n_2^{cdg}, \ldots, n_N^{cdg}\right\}$, where $n_i^{cdg} = d_i + \sum_{j=1}^{N} b_j \delta\left(i, j\right)$, where

$\delta(i,j) = \begin{cases} 1, & if \ c_{ij} > 0 \\ 0, & otherwise \end{cases}$, where $d_i \in EDV$, $b_j \in EDV$, $c_{ij} \in GCEM$.

The matrix for the weekly lecturers' preferences (LWPM), can be represented by a matrix of dimensionality $L \times D \times H$, where L is the number of lecturers, D is the number of days in the week schedule and H respectively, the number of hours in one day. The elements in the matrix that have value "0", showing that these hours are the most preferred. The elements with value "1" indicate that this is the next favorite hours for the lecturer, etc. LWPM matrix is shown in (37):

(37) $LWPM = \{m_k\}_{k=1}^{L} = \{m_1, m_2, ..., m_L\}$, where m_k is the matrix of weekly preference of the k-th lecturer and $m_k = (p_{ij})_{DxH}$, where $p_{ij} \in \left[0, \left\lceil \frac{T}{l_k^t} \right\rceil - 1\right]$, $l_k^t \in LHLV$.

The common matrix for the preferences of students (CSWPM), can be represented by a matrix of dimensionality $D \times H$. Similarly, as in the matrix LWPM, the lower the value of the element, the more preferred the timeslot corresponding to that element is in the matrix CSWPM preferences belong to all students. The definition of the matrix CSWPM, is shown in (38):

(38) $CSWPM = (p_{ij})_{DxH}$, where $p_{ij} \in \left[0, \left\lceil \frac{TS}{\sum_{i=1}^{S} s_i^t} \right\rceil - 1\right]$, $s_i^t \in LHSV$.

Decision vector (RV) can be represented as a list of length N. Each element of the vector RV is an ordered pair of elements $\langle t_i^s, t_i^f \rangle$, respectively, for the beginning and end of the i-th event. The definition of the vector RV is shown in (39):

(39) $RV = \left\{n_i^r\right\}_{i=1}^{N} = \left\{n_1^r, n_2^r, \ldots, n_N^r\right\}$, where $n_i^r = \langle t_i^s, t_i^f \rangle$, $t_i^s \in AHV$, $t_i^f \in AHV$, $t_i^f = (t_i^s + v_i) - 1$, $v_i \in EDV$.

The vector of hours allocated by days (HDRV) and the vector of fixed events (FEV) is defined, respectively:

(40) $HDRV = \left\{n_i^{rd}\right\}_{i=1}^{N} = \left\{n_1^{rd}, n_2^{rd}, \ldots, n_N^{rd}\right\}$, where $n_i^{rd} = g(i)$ and

$$g(i)=\begin{cases}\dfrac{t_i^s}{H}, \; if \;\; t_i^s \bmod H=0\\ \left\lfloor\dfrac{t_i^s}{H}\right\rfloor+1, \; otherwise\end{cases}, \text{ where } t_i^s\in AHV.$$

(41) $FEV=\left\{n_i^{fx}\right\}_{i=1}^{N}=\left\{n_1^{fx},n_2^{fx},\ldots,n_N^{fx}\right\}$, where $n_i^{fx}=\begin{cases}-1, \; if \;\; user \;\; is \;\; fixed \;\; event \;\; i\\ 0, \; otherwise\end{cases}$.

A matrix for coefficients of events distance (CDEM) has dimension $N\times N$ and is generated using an EV vector, GCEM matrix, RV vector and HDRV vector. The value of each element in the matrix CDEM, shows the number of hours which put the events of a distance i and j (if they are in conflict), where $i\neq j$, $i,j\in\{1,2,...,N\}$. CDEM is a symmetric matrix to its main diagonal and its definition is shown in (42):

(42) $CDEM=(p_{ij})_{NxN}$, where $p_{ij}=\delta(i,j)*\lambda(t_i,t_j)$, where $\delta(i,j)=\begin{cases}1, \; c_{ij}>0\\ 0, \; c_{ij}=0\end{cases}$, $c_{ij}\in GCEM$,

$$\lambda(t_i,t_j)=\begin{cases}\left(t_j^s-t_i^f\right)-1, \;\; if \;\; \begin{pmatrix}\left(d_i=d_j\right)\wedge\left(t_i^s<t_j^s\right)\wedge\left(t_i^s<t_j^f\right)\wedge\\ \left(t_i^f<t_j^s\right)\wedge\left(t_i^f<t_j^f\right)\end{pmatrix}\\ \left(t_i^s-t_j^f\right)-1, \;\; if \;\; \begin{pmatrix}\left(d_i=d_j\right)\wedge\left(t_i^s>t_j^s\right)\wedge\left(t_i^s>t_j^f\right)\wedge\\ \left(t_i^f>t_j^s\right)\wedge\left(t_i^f>t_j^f\right)\end{pmatrix}\\ 0, \;\; if \;\; \left(d_i\neq d_j\right)\end{cases}$$

where $d_i,d_j\in HDRV$, $t_i,\,t_j\in RV$, $t_i=\left\langle t_i^s,t_i^f\right\rangle,\, t_j=\left\langle t_j^s,t_j^f\right\rangle$.

The matrix for the students' weekly schedule (SWLM) is a three-dimensional matrix with dimensionality $S\times D\times H$. This matrix is generated using a SEM matrix, RV vector, ELV vector and EAV vector. The definition of the SWLM matrix is shown in (43):

(43) $SWLM=\{m_k\}_{k=1}^{S}=\{m_1,m_2,...,m_S\}$, where $m_k=(p_{ij})_{DxH}$,

$$p_{ij}=\begin{cases}\langle n,l,a\rangle, \; where \; \exists!n, \; which \; \left(s_{kn}=1\right)\wedge\left(t_n^s\leq h\leq t_n^f\right), \; n=1,2,...,N\\ 0, \; otherwise\end{cases},$$

$s_{kn}\in SEM$, $t_n=\left\langle t_n^s,t_n^f\right\rangle\in RV$, $h=(i-1)H+j$, $l=u_n,\,u_n\in ELV$, $a=v_n,\,v_n\in EAV$.

Similarly, the matrices for lecturers and auditoriums weekly schedule, respectively (LWLM) and (AWLM) are presented. For them each element contains the value "0" or an ordered pair, respectively, $\langle n,a\rangle$ or $\langle n,l\rangle$, where n is the event which will be held by lecturer l in auditorium a.

Conditions for the Existence of a Solution

The following three conditions must be satisfied in order to have a possible solution:

1) For $\forall k,\ k=1,2,...,L$, to be met:

(44) $\left\lceil \frac{l_k^t}{t_m^l} \right\rceil \le D$, where $l_k^t \in LHLV$ and is the maximum number of hours a day, which may have the k-th lecturer.

2) For $\forall p,\ p=1,2,...,S$, to be met:

(45) $\left\lceil \frac{s_p^t}{t_m^s} \right\rceil \le D$, where $s_p^t \in LHSV$ and t_m^s is the maximum number of hours a day, in which the p-th student may be involved.

3) The total duration of the timetable must be less than or equal to the absolute number of hours in the weekly schedule.

To formulate this condition, we construct graph $G=(V,A)$ as follows: each event n_i will correspond to the vertex $v_i \in V$ with weight $w(v_i)$, equal to the length of the i-th event i.e. $w(v_i)=n_i^d$, $n_i^d \in EDV$ and the edge between vertex i and j exists if and only if events i and j are involved in a conflict of common resources (students, lecturers or auditoriums).

The graph $G=(V,A)$ is called *r-chromatic* if its set of vertices V can be broken to *r* subsets $V_1,V_2,...,V_r$, such that $\bigcap_{k=1}^{r} V_k=\varnothing$, $\bigcup_{k=1}^{r} V_k=V$ and which have two vertices of the same subset which are not joined with edge. Classes $V_1,V_2,...,V_r$ are called chromatic classes. The *chromatic number* $\gamma_v(G)$ of the graph *G* is called the minimum number *r*, for which the graph *G* is *r*-colorable [6].

To determine the minimum length of the timetable it is necessary for all chromatic classes to find vertices with maximum weights and to form their sum. It must be less than or equal to the absolute number of hours, i.e.:

(46) $T \ge \sum_{k=1}^{r} \max_{i\in\{1,...,|V_k|\}} \left(w(v_i)\right)$, where $r=\gamma_v(G)$ is the chromatic number of graph *G*, $|V_k|$ is the cardinal number of the *k*-th chromatic class and $w(v_i)$ is the weight of the i-th vertex belonging to the *k*-th chromatic class.

The problem of finding the chromatic number of graph is NP-hard and its solution in a large number of vertices (or edges) is impossible. It is possible to use heuristic algorithms, in which it is not always guaranteed that the found chromatic number is minimal.

It is necessary to note that if one of the three inequalities (44), (45) and (46) are not met, then the number of days D in a weekly schedule is not sufficient and should be increased.

Hard Constraints

For a solution to be acceptable the following four soft constraints must be fulfilled:

1) Any element of any resource cannot be in two events at the same time, i.e.:

(47) $\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} p_{ij} * \delta\left(t_i, t_j\right) = 0$, where

$$\delta\left(t_i,t_j\right)=\left\{\begin{array}{l} 0,\ if\left(d_i=d_j\right)\wedge \\ \qquad \left(\left(t_i^s<t_j^s\right)\wedge\left(t_i^s<t_j^f\right)\wedge\left(t_i^f<t_j^s\right)\wedge\left(t_i^f<t_j^f\right)\right)\vee \\ \qquad \left(\left(t_i^s>t_j^s\right)\wedge\left(t_i^s>t_j^f\right)\wedge\left(t_i^f>t_j^s\right)\wedge\left(t_i^f>t_j^f\right)\right) \\ 1,\ otherwise \end{array}\right\},$$

$p_{ij} \in GCEM$, $t_i \in RV$, $t_j \in RV$, $t_i = \left\langle t_i^s, t_i^f \right\rangle$, $t_j = \left\langle t_j^s, t_j^f \right\rangle$, $d_i \in HDRV$.

This condition requires the events that are in conflict to be allocated in different timeslots in the weekly schedule.

The number of students must be less than or equal to the capacity of the auditorium where the event takes place and the kind of auditorium, must match the kind of event, i.e. for $\forall i,\ i = 1,2,...,N$ must be met:

(48) $\left(n_i^s \le a_k^c\right)\wedge\left(a_k^k = n_i^k\right) \Leftrightarrow n_i^a = a_k,\ k = 1,2,...,A$, where $a_k^c \in ACV$, $a_k \in AV$, $n_i^a \in EAV$, $n_i^s \in NSEV$, $a_k^k \in AKV$, $n_i^k \in EKV$.

A lecturer cannot teach in more than t_m^l number of hours per day, i.e. for $\forall m_k, k = 1,2,...,L$, where $m_k \in LWLM$, to be met:

(49) $l_{k_d} \le t_m^l,\ d = 1,2,...,D$, where $l_{k_d} = \sum_{j=1}^{H} 1,\ if\ \ p_{ij} \ne 0$, $p_{ij} \in m_k$.

4) A student may not be involved in more than t_m^s number of hours per day, i.e. for $\forall m_k, k = 1,2,...,S$ where $m_k \in SWLM$, to be met:

(50) $s_{k_d} \le t_m^s,\ d = 1,2,...,D$, where $s_{k_d} = \sum_{j=1}^{H} 1,\ ако\ p_{ij} \ne 0,\ p_{ij} \in m_k$.

Soft Constraints

The soft constraints may be of different weights relative to each other. The quality of a solution is determined by the number of violations of those constraints. The higher this number is small the solution is better. We define the following three soft constraints:

1) The time span between the events to be the shortest possible.

To take into account the weights of the conflicts themselves, we will introduce the term "average cost of distance". For an event n_i, that cost is the sum of products between weights of conflicts and the coefficients of distance divided by the number of events.

We formulate this constraint as an objective function G_1:

(51) $G_1(i,j) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij} q_{ij}}{N^2} \to \min$, where $p_{ij} \in GCEM$, $q_{ij} \in CDEM$.

Because the matrix GCEM (respectively CDEM) is symmetrical to its main diagonal products of the elements above the main diagonal can be summed up and the resulting amount can be divided in half by the square of the number of events, i.e.:

$$G_1(i,j) = 2\frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} p_{ij} q_{ij}}{N^2} \to \min$$

A function $G_1(i,j)$ reaches a minimum at $\min(G_1(i,j)) = 0$.

2) Satisfaction of lecturers preferences to be maximized.

We formulate this constraint as an objective function G_2:

(52) $G_2(l,d,h) = \sum_{l=1}^{L} w_l \sum_{d=1}^{D}\sum_{h=1}^{H} p_{ldh}\delta(l,d,h) \to \min$, where $p_{ldh} \in LWPM$, $w_l \in WLV$ and

$\delta(l,d,h) = \begin{cases} 1, & x_{ldh} \ne 0 \\ 0, & otherwise \end{cases}$, where $x_{ldh} \in LWLM$.

A function $G_2(l,d,h)$ reaches a minimum at $\min(G_2(l,d,h)) = 0$.

3) Satisfaction of students preferences to be maximized.

Assuming that both resources "lecturers" and "students" have the same total weight, the weight of each student receives:

(53) $w_s = \frac{\sum_{i=1}^{L} l_i^w}{S}$, where $l_i^w \in WLV$.

Now we can formulate a third soft constraint as an objective function G_3 :

(54) $G_3(s,d,h) = \sum_{s=1}^{S} w_s \sum_{d=1}^{D} \sum_{h=1}^{H} p_{dh} \delta(s,d,h) \to \min$, where $p_{dh} \in CSWPM$ and

$\delta(s,d,h) = \begin{cases} 1, & x_{sdh} \neq 0 \\ 0, & otherwise \end{cases}$, where $x_{sdh} \in SWLM$.

Objective Function

The objective function of the problem can be formulated as a function that combines the three objectives G_1 , G_2 and G_3 , which were defined above.

It is necessary that a problem to be formulated mathematically as a linear multicriteria optimization.

The basic idea of multicriteria optimization is to convert output multicriteria problem in a problem with an objective function. The resulting problem leads to an efficient solution, which may not be optimal for all objectives of the original problem, because some of the objectives may conflict.

One possible approach for solving the multicriteria problems is to bring the relevant problem to problem with one criteria by replacing the set of objective functions $f_1,...,f_p$ with a single function $F(f_1,...,f_p)$. Defined in this way, the multicriteria problem is reduced to minimization (respectively, maximization) of total function $F(f_1,...,f_p)$, which is a function of the given criteria $f_1,...,f_p$.

One of the commonly used methods for the formation of the function $F(f_1,...,f_p)$ is the use of weight sums of functions $f_1,...,f_p$, presenting the criteria (objectives) in a given problem.

Let the multicriteria problem have *p* criteria (objectives) and *i*-th criteria be set as follows:

$\min(\max)\left\{f_{i,}\ i=1,...,p\right\}.$

Then the total objective function if used the weighing method is:

$\min(\max)F\left(f_i,...,f_p\right) = \sum_{i=1}^{p} w_i * f_i$, where $w_i \geq 0,\ i=1,...,p$ are weights that are set by the decision maker under the relative importance of each criteria.

The weighing method will be used for the task under investigation in this paper.

Definition. If X is convex set and $x_1,...,x_p \in X$, $w_i \geq 0$, $\sum_{i=1}^{p} w_i = 1$, then $x = \sum_{i=1}^{p} w_i x_i \in X$ is a convex combination of $x_1,...,x_p$ with coefficients $w_1,...,w_p$.

We determine the relative importance of each of the three criteria in such a way that the amount of weight equals to one.

Resources which interest us are lecturers L and students S because we have not defined soft constraints connected with auditoriums. The first objective G_1 is common for students and lecturers. The second objective

G_2 is connected only to the lecturers and the third objective G_3 is connected only to the students. The relative weight of the first objective is proportional to the sum of the number of lecturers and students i.e. $L+S$. The weighing of the second objective is proportional to the number of students S, and the weight of the third objective is proportional to the number of lecturers L. For for the three objectives we get a relative weight equal to $2(S+L)$. Since we have assumed that the sum of the weights of the objectives will be equal to one then for the weight coefficients we get the following values:

(55.1) $w_1 = \dfrac{L+S}{2(L+S)} = \dfrac{1}{2}$

(55.2) $w_2 = \dfrac{S}{2(L+S)} = \dfrac{1}{2}\dfrac{S}{L+S}$

(55.3) $w_3 = \dfrac{L}{2(L+S)} = \dfrac{1}{2}\dfrac{L}{L+S}$

For the common objective function G we get:

(56) $G = w_1G_1 + w_2G_2 + w_3G_3 \to \min$.

Because $w_1 \geq 0$, $w_2 \geq 0$, $w_3 \geq 0$ and $\sum_{i=1}^{3} w_i = 1$ the function G is a convex combination of G_1, G_2 and G_3, with coefficients w_1, w_2 and w_3.

By replacing the weights with their values and simplify (56), we obtain the objective function given in (57):

(57) $$G(i,j,l,d_1,h_1,s,d_2,h_2) = \frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} p_{ij}q_{ij}}{N^2} + $$
$$+\frac{1}{2}\frac{S}{L+S}\sum_{l=1}^{L} w_l \sum_{d_1=1}^{D}\sum_{h_1=1}^{H} u_{ld_1h_1}\delta_1(l,d_1,h_1) + \frac{1}{2}\frac{L}{L+S} w_s \sum_{s=1}^{S}\sum_{d_2=1}^{D}\sum_{h_2=1}^{H} v_{d_2h_2}\delta_2(s,d_2,h_2) \to \min$$

where N is the number of events S is the number of students, L is the number of lecturers, D is the number of days in the weekly schedule, H is the number of hours a day of the weekly schedule, $p_{ij} \in GCEM$, $q_{ij} \in CDEM$, $u_{ldh} \in LWPM$, $w_l \in WLV$,

$\delta_1(l,d_1,h_1) = \begin{cases} 1, & x_{ld_1h_1} \neq 0 \\ 0, & otherwise \end{cases}$,

$x_{ld_1h_1} \in LWLM$, $v_{d_2h_2} \in CSWPM$,

$w_s = \dfrac{\sum_{i=1}^{L} l_i^w}{S}$, $l_i^w \in WLV$,

$\delta_2(s,d_2,h_2) = \begin{cases} 1, & x_{sd_2h_2} \neq 0 \\ 0, & otherwise \end{cases}$,

$x_{sd_2h_2} \in SWLM$,

provided that the conditions for the existence of solution (44), (45), (46) and hard constraints (47), (48), (49) and (50) are satisfied.

The function G reaches a minimum at $\min\left(G\left(i, j, l, d_1, h_1, s, d_2, h_2\right)\right) = 0$.

Conclusion

In this paper a model for the university course timetable problem is presented. Fundamental data structures, parameters, vectors and matrices, which are used in defining the problem are presented. Hard constraints on the verification of concurrent involvement of lecturers, students and auditoriums, checking for excess capacity auditoriums and verification of commitment of students and lecturers for more hours per day than a predefined number are presented. Also, the soft constraints formulated as the objective functions are presented. The total objective function which is a convex combination of the others (with appropriate weighing) is also defined.

As future trends of work we may mention: the creation of algorithms for solving the problem which is based on the proposed model. Their integration into real information systems aims at demonstrating the usefulness of the developed model in solving combinational optimization problems.

Bibliography

1. Burke, E. K. and Landa Silva J. D. (2004) The design of memetic algorithms for scheduling and timetabling problems. Recent Advances in Memetic Algorithms, Studies in Fuzziness and Soft Computing, Springer, 166, pp 289-312.
2. Carter, M. W., Laporte G. (1998) Recent developments in practical course timetabling. Selected Papers from 2nd International Conference on the Practice and Theory of Automated Timetabling (PATAT II), Toronto, Canada, Springer-Verlag, pp 3-19.
3. Rossi-Doria, O., Blum C., Knowles J., Samples M., Socha K. and B. Paechter. (2002) A local search for the timetabling. The Proceedings of the 4th International Conference on Practice and Theory of Automated Timetabling (PATAT IV), Gent, Belgium, pp 124-127.
4. Rossi-Doria O., Samples M., Birattari M., Chiarandini M., Dorigo M., Gambardella L.M., Knowles J., Manfrin M., Mastrolilli M., Paechter B., Paquete L. and Stützle T.. (2003) A comparison of the performance of different meta-heuristics on the timetabling problem. The Practice and Theory of Automated Timetabling IV: Selected Papers from 4th International Conference on the Practice and Theory of Automated Timetabling (PATAT IV), Gent, Belgium, pp 329-354.
5. Socha K., Knowles J. and Samples M.. (2002) A max-min ant system for the university course timetabling problem. The Proceedings of the 3rd International Workshop on Ant Algorithms (ANTS 2002), Lecture Notes in Computer science 2463, Springer-Verlag, pp 1-13.
6. Мирчев, Ив. (2001) Графи. Оптимизационни алгоритми в мрежи, Унив. изд. "Неофит Рилски", Благоевград.

Authors's Information

***Velin Kralev** – Department of Informatics, Faculty of Mathematics and Natural Sciences, South-West University "Neofit Rilski"-Blagoevgrad, Bulgaria; e-mail: velin_kralev@abv.bg*

MOBILE SEARCH AND ADVERTISING

Vladimir Lovitskii, Colin McCaffery, Michael Thrasher, David Traynor, Peter Wright

Abstract: *Mobile advertising is a rapidly growing sector providing brands and marketing agencies the opportunity to connect with consumers beyond traditional and digital media and instead communicate directly on their mobile phones.* Mobile advertising will be intrinsically linked with mobile search, which *has transported from the internet to the mobile and is identified as an area of potential growth. The result of mobile searching show that as a general rule such search result exceed 160 characters; the dialog is required to deliver the relevant portion of a response to the mobile user. In this paper we focus initially on mobile search and mobile advert creation, and later the mechanism of interaction between the user's request, the result of searching, advertising and dialog.*

Keywords: *mobile text messages, mobile search, mobile advertising, question-answering system*

ACM Classification Keywords: *I.2 Artificial intelligence: I.2.7 Natural Language Processing: Text analysis.*

Conference*: The paper is selected from Second International Conference "Intelligent Information and Engineering Systems" INFOS 2009, Varna, Bulgaria, June-July 2009*

Introduction

This paper considers the results of our recent research in the areas of text data mining and the natural language processing [1-7] when restricted by mobile phone text-based SMS messaging. In our previous papers [6,7] the Question-Answering Mobile ENgine (QAMEN) has been discussed. During internet searching QAMEN converts web pages to a simplified format that is compatible with handheld devices. Moreover, QAMEN frees users to have an expensive mobile phone with a web browser. Internet connections from mobile devices continue to remain expensive and there is little prospect to an immediate decrease in pricing structure. QAMEN is based on industry-standard SMS messaging technology and thus works with any mobile in any GSM network. Hence QAMEN is useful for people on the move and probably unable to access a PC.

Worldwide Mobile Search Advertising Spending, by Region, 2007-2012 (millions)

	2007	2008	2009	2010	2011	2012
US	$34.5	$107.4	$241.8	$530.5	$910.2	$1,484.2
Asia-Pacific	$26.0	$72.0	$189.9	$372.8	$732.4	$1,160.0
Western Europe	$18.4	$52.0	$140.5	$339.7	$614.1	$968.2
Rest of World	$4.4	$12.4	$24.4	$47.0	$88.6	$160.9
Worldwide	**$83.3**	**$243.7**	**$596.6**	**$1,289.9**	**$2,345.2**	**$3,773.2**

The growth of mobile-search-related advertising is strong across many regions [8] (see the Table). Mobile search will eventually move away from the generic style of searches seen on the fixed internet to more personal services. In these circumstances local search will become increasingly important. Through these local searches location-based services and advertising will be able to gain vital revenues as mobile subscribers using search will reveal both their geographical location and the items being searched for.

Mobile search and mobile advertising should grow significantly because mobile phones have several advantages over PCs when accessing the Internet [9]:

- Mobile phones are always on, always available, and always "connected". Subscribers can gain access to information anywhere; at home, in the office, at a restaurant, or from the car.

- There are currently 3 billion mobile phones worldwide compared with less than 1 billion PCs.
- Ability to immediately connect people to phone numbers, since they already have a device in hand.
- When conducting mobile search, users are in an atmosphere, situation and environment more likely to result in a purchase. Most of the time they are out of the home.

But mobile phones have some limitations when compared to PCs:

- PCs have relatively large screens that can show more detailed information and numerous search results. By contrast, mobile phones have smaller screens with limited real estate.
- PCs have large, comfortable keyboards with easy-to-use pointing tools including mouse, trackball, or touchpad. Mobile phones have compact number pads, commonly with arrow keys limiting user-navigation to up, down, left and right.
- PCs and the Internet are relatively homogenous, using consistent colour displays, screen sizes, browsers, and open programming standards. Mobile phones are heterogeneous, employing varying input methods, display screens, browsers, operating systems, and user interfaces. Mobile devices do not conform to any standards.
- Mobile search usability issues include smaller screens, typing limitations of phone keypads and the cost of spending time scrolling through mobile search results.
- Mobile devices are currently less likely to be used for general browsing but more for retrieving specific information.

These limitations describe the problems that need to be taken into account when developing an acceptable mobile search and advertising procedure. In addition to these immediate differences, future mobile search applications may be able to capitalize on user specific information [7]. Mobile phones are increasingly associated with each individual's personal style, representative of their owner's personality, with specific demographics, behaviour patterns, and personal interests etc. This information offers the opportunity for more relevant search results to be determined. Mobile phones may also be able to leverage location as an additional search parameter, allowing for greater specificity for search results.

Reading this paper will tell you the following:

- The difference between mobile web search and PCs web search.
- Why mobile dialog is needed?
- What is mobile advertising?
- How mobile search and mobile advertising will work together.

QAMEN versus Mobile Web Search

The mobile web consists of web pages that are designed specifically for display on mobile devices. Due to their limited capabilities (relative to standard computers) mobile devices access and render web content using the specialized Wireless Application Protocol (WAP). WAP is separate and distinct from the Hyper-Text Transfer Protocol (HTTP) that computers use to access HTML pages. As such, WAP browsers can only access pages that are written in xHTML (eXtensible Hyper-Text Markup Language) or WML (Wireless Markup Language). WAP 1.x browsers access only WML pages, while WAP 2.0 browsers access both xHTML pages and WML pages. The mobile web is determined by the universe of content that is written in WML or xHTML.

Searching the mobile web returns only WAP pages. The world is short of quality WAP sites in general. Most WAP sites are poorly designed, under-used, unstable and limited in content and service. There are WAP search engines that index the mobile web but many do not work well due to the volatile nature of the current mobile web – there are many outdated and redundant links. Furthermore, most WAP site crawling technologies are underdeveloped [10].

QAMEN is an HTML compatible search engine. It works almost the same as computer browsers and can access HTML search pages from the huge internet webpage databases in the same way that computer browsers do. In providing the result of internet searching QAMEN converts web pages to a simplified format compatible with handheld devices. Search technologies of QAMEN are evolving to provide users with appropriate results despite the often unstructured web content.

QAMEN's Basic Commands

By default any request is considered by QAMEN as a request for searching in the local knowledge base, and/or in the Internet. Mobile question answering differs from standard information retrieval methods in some important respect. Firstly, it needs to retrieve specific fact information rather than whole documents. Secondly, it should select among the found facts the shortest and appropriate fact to meet the requirement of 160 characters. In short what a user normally requires is a precise answer to a question. But the concept "precise answer" is very fuzzy because of its reflects potential ambiguities in a user's request. For example, QAMEN even theoretically cannot define what is the precise answer to a user's request: ***"I'm looking for Hilton address in London"*** (see Figure 1). On the one hand, the most important, the result must be relevant to the user;s search, on the other hand, there is no guarantee that the displayed response: ***"The Hilton Trafalgar +44 20 7870 2900"*** exactly meets the user's expectation. Enhancing the user profile, of course, could significantly improve the selection of appropriate answer. Even an unexpected advert along with the correct response might improve user satisfaction.

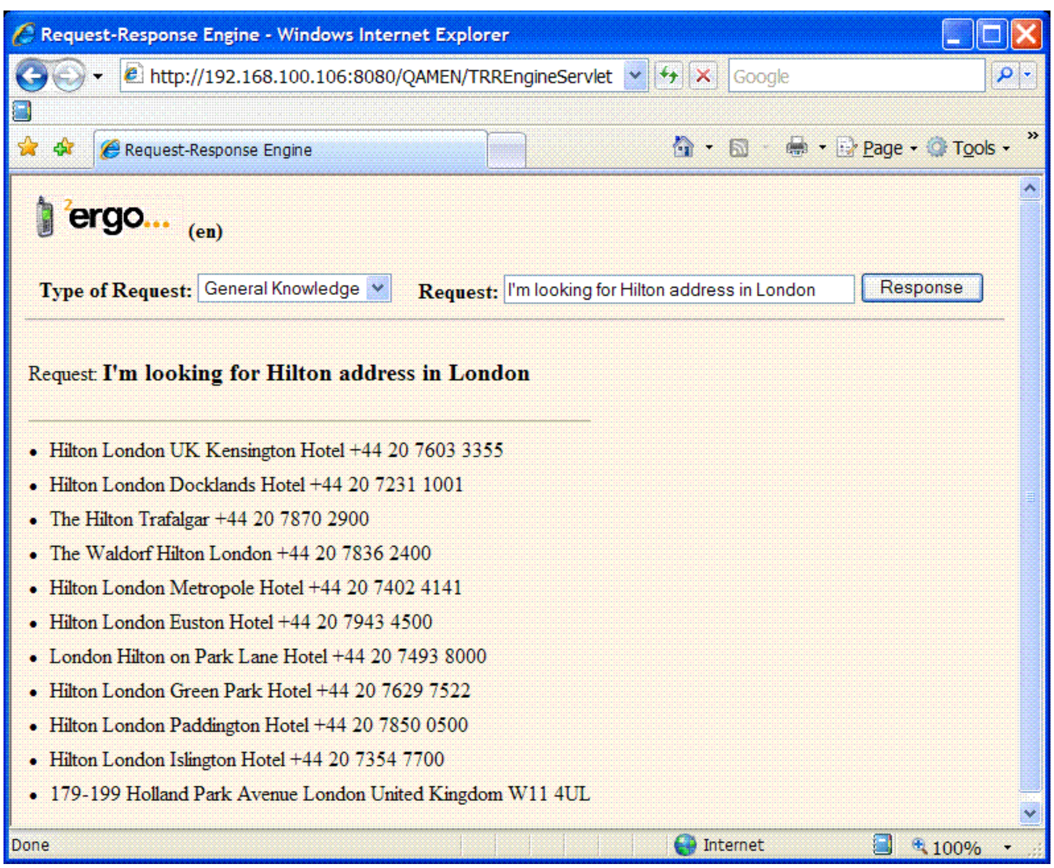


Figure 1. List of locations and phone numbers of Hilton hotels in London

The mobile search industry is still in its infancy and the primary barriers, cited by subscribers as affecting potential uptake, are not knowing how to use the search engine (23%) and not thinking about using mobile search on their phone (19%). Just 20% of UK subscribers actually search content on the mobile internet, despite an industry perception that 89% do [11]. This gap between reality and perception suggests that the industry needs to promote the benefits of mobile search more actively as well as educating consumers on how mobile search may be used to find relevant content.

Taking into account such a situation QAMEN provides users with the maximum flexibility for selection of the desirable combination of options. Mobile users seek quick and convenient access to information and services. The following basic commands, for example, allow a user to select any combination of options online:

- Type in ***Advert ON*** to activate advertising (or ***Advert OFF*** to quit).
- Type in ***Dialog ON*** to activate dialog (or ***Dialog OFF*** to quit).
- Type in ***Person ON*** to activate personalization (or ***Person OFF*** to quit).
- Type in ***Language BG*** to allow user to use Bulgarian language (or ***Language EN*** to back to English).
- Type in ***Election ON*** to select an application domain with data regarding the UK General Election 1997, 2001 and 2005 (or ***Election OFF*** to quit).

For example, in the result of Dialog activation when QAMEN provides interactions between itself and the user the response to the same request: ***"I'm looking for Hilton address in London"*** will be represented differently (see Figure 2). Space on the mobile phone screen is at a premium, and users have limited input mechanisms, so any result of mobile searching needs to be easy to navigate using only the mobile phone's own keypad.

Each user can have multiple active mobile dialogs. QAMEN holds continually the personal profile of each user as well as the current state of dialogue for each search activity.

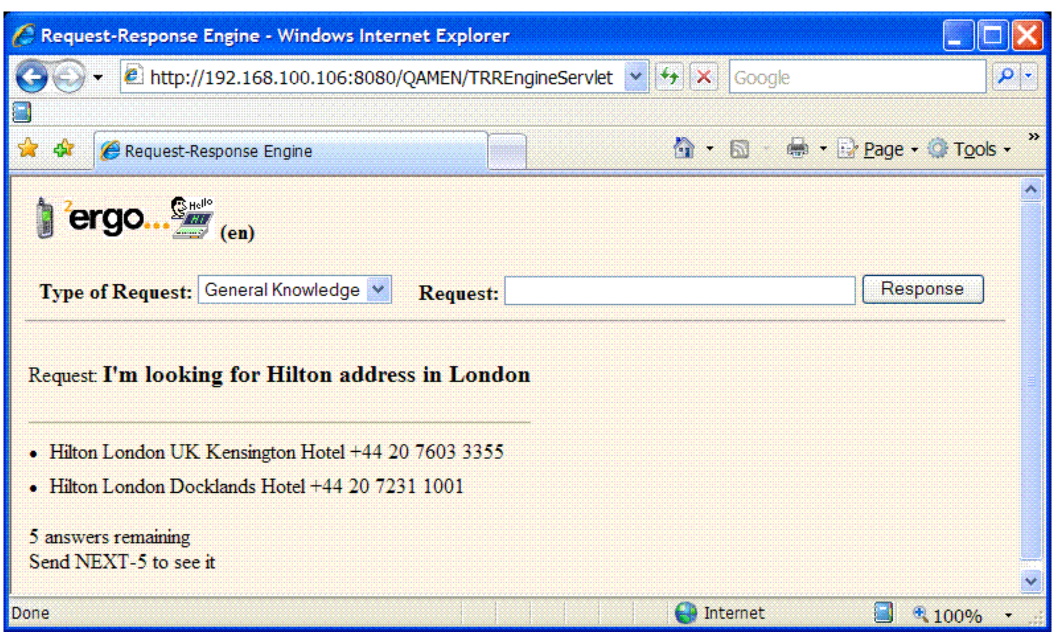


Figure 2. List of locations and phone numbers of Hilton hotels in London with dialog

Let us distinguish Answers Quantity (AQ) and SMS-Responses Quantity (RQ). For considered request AQ=11 (see Figure 1). RQ represents number of SMS, which need to be sent to cover AQ. As a rule the length of one response less the size of a SMS. That is why one SMS may include several answers i.e. RQ $\leq$ AQ. QAMEN shows how many RQ left (see Figure 2).

Mobile Advertising

Mobile advertising can basically be defined as the business that using the wireless channel as a medium to deliver the advertisement message or slogan encourages people to buy products and services. Mobile advertising is not simply linking users to websites by clicking on banners or pop-up windows through their mobile phone browser. It offers the ability to go beyond the mainstream of electronic advertising, which mainly targets the mass market predominantly over the Internet, to allow businesses to deliver their particular messages to specific individuals at specific locations.

QAMEN is an ideal way for advertisers to reach target markets and establish a one-to-one relationship with the consumer, a significant objective for all advertisers. For example, weather application is defined as personalized, localized weather prediction according to user location and personal profile. Weather related advertisement system knows how to match the right advertisement to the right weather in the most effective manner. For example, implementing a decision to start a soft drink campaign when the temperature approaches, 32°C / 90°F according to user location (if the user is close to the beach and experiencing higher levels of effective temp he will experience a different ad in different temperatures). Such approaches would help the advertiser to optimise its advertising campaign.

The following key characteristics drive the success of messaging as a source of mobile advertising inventory [12]:

- Ubiquitous SMS access – Virtually all mobile phones can receive SMS, and the majority of users use SMS on a regular basis. Today, SMS is the most widely used mobile phone service after voice.
- High attention level – Users almost never delete mobile message without opening them and reading at least part of content.
- Engaging – Once displayed, an effective advert can engage users directly in various ways, such as interactive message reply, click to transfer to the Web, or click to call.
- Compatibility – Messaging usually works between different networks and between different countries.
- Response collection is easier – It also may be achieved immediately. Brands may have access to real time response information and may modify the campaign according to the results, long before a campaign terminates.
- Direct and personal way of communication – Customer has the sense of feeling that the advert addresses only him/her.

Mobile Advertisement Creation

In this paper we consider adverts that are text only. Mobile messaging represents an opportunity for advertising placement. Advertisement can be inserted in SMS content that subscribers request and receive by using the free (non-used) space, up to the message size limit. The main task of advert creation is to establish an advert structure. There are, as yet, no standards for representing these, because there is no general agreement on what an advert should contain. That is why we were free to offer our vision of advert structure. First, let us describe the general requirements, restrictions and conditions to advert creation:

- A poor mobile advertising conveys a bad impression of any company, potentially turning away customers.
- An advert must be related to the original message content i.e. text adverts have to be shown in response to keywords entered on user queries.
- Mobile advertising is moving more towards dialogue type of advertising. Limiting user input to numeric or short sequence of text due to the limitation of the keypad on most mobile phones.
- Adverts should provide click-to-call links that allows the users to make a phone call directly from the displayed advert.
- Analysis of advert "activities" is a crucial feature in their future development in order to automate an improved process of advertising. The correctly specified metrics for an advert should include:
  - Total adverts sent.
  - Total adverts delivered.
  - Unique deliveries (distinct mobile phones).
  - Content of advert served.
  - Type of advert (independent or request dependent).
  - Location of served advert.
  - Date and Time when advert has been displayed.
  - Season of advert displaying.
- SMS advertising is appended to the bottom of the content message to be sent to the user. The available space for each advert is therefore dependent on how much space is left after subtracting the characters used in the main content body of the message.
- To optimize an advert placement:
  - The advertiser should develop several versions of adverts to be used, with the optimum length directly dependent on the length of sender's message.
  - Use punctuation only if required for clarity or emphasis.
  - Try to avoid a carriage return because it may count as two characters.

Let us distinguish between two types of advert: 1) independent if there is no any link with areas, season or location, and 2) dependent otherwise. For both types of adverts link with user request is required. Advert can be easily created and updated via the web-based interface (see Figure 3).

The advert structure consists of the following items:

- Areas: *game, music, video, product, weather, sport, health, finance, holiday, transport, omit*. For advert several areas might be selected e.g. *sport* and *weather*.
- Seasons: *all seasons, winter, spring, summer, autumn, omit*. If *spring* is selected the months *March*, *April* and *May* will be added to advert pattern automatically.
- Location: *country, county, city, omit*.
- Slogan is a short (< 100 characters) memorable text advert.
- Advert's Key Words (AKW) include not only synonyms of slogan's keywords but also describes the conditions for use of the advertised product. For example, if "ice cream" is advertised the words: *hot* and *heat* should be added to keywords.
- Advertiser and its contacts: *country, city, address, web site, email, telephone*.

- Priority depends on payment and its value lie within 1-10 range.
- Quantity shows how many times advertiser wants advert to be displayed and (in brackets) how many times advert has been displayed.

Before storing the advert description in the Adverts DataBase (ADB) which is the main part of advert – (the Advert Pattern (AP)) – is created. The general form of AP is the following:

AP=<SKW+AKW><Areas><Seasons><Location><Priority><Quantity>, where:

- SKW means Slogan's Key Words. SKW are extracted from the Slogan and together with AKW are saved in the advert dictionary;
- Meaningful selection (i.e. when the selected value is no equal *omit*) of Areas, Seasons and/or Location are saved in the advert dictionary as well. The selection of *All Seasons* means that all twelve months and four season names will be included in the dictionary.

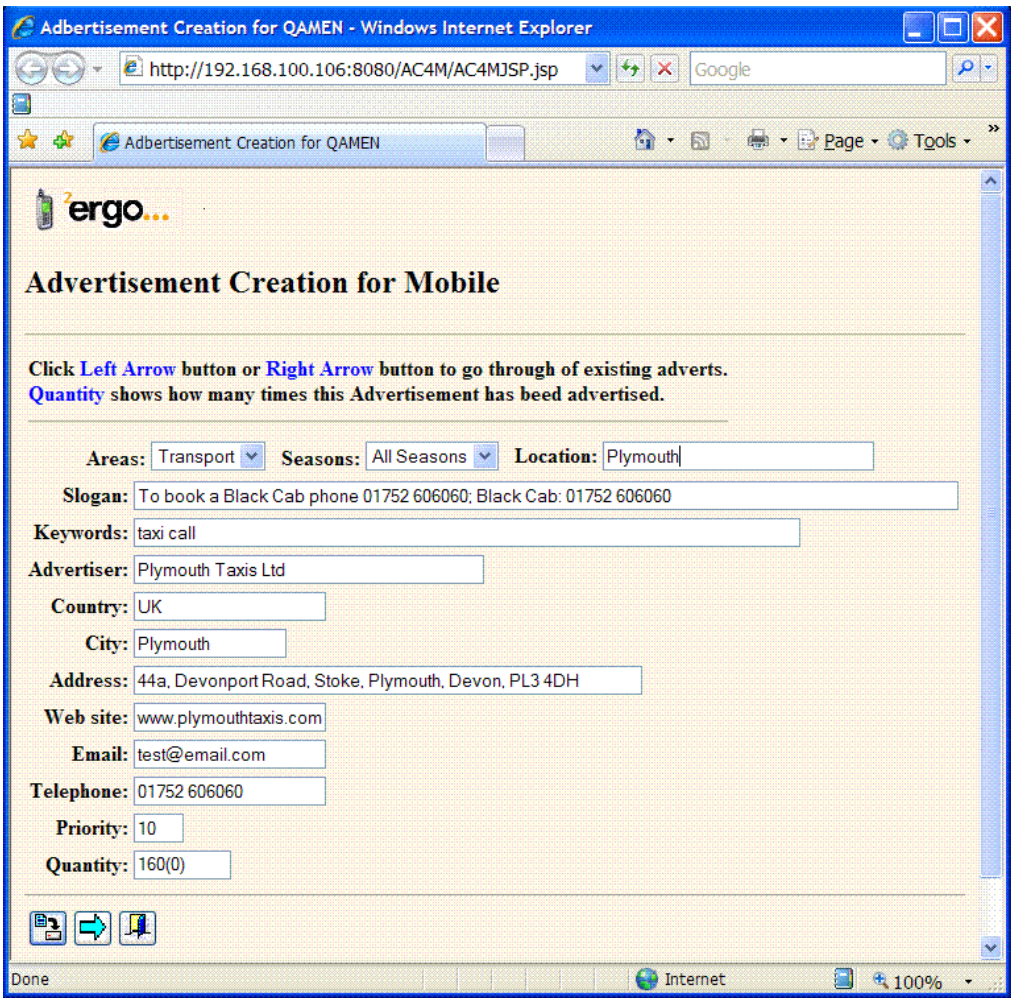


Figure 3. Web-based interface for Advert creation and editing

It is important to emphasise that any words are inserted in the advert dictionary just once but that the link chain related to any word will be extended, and will consist of the information of all adverts from ADB where this word has been involved. For example, word *May* has a link to the set of adverts *Seasons* that have been described as

All Seasons, or *Spring*. An example of advert, which is extracted from the ADB in accordance with user's request is shown in Figure 4.

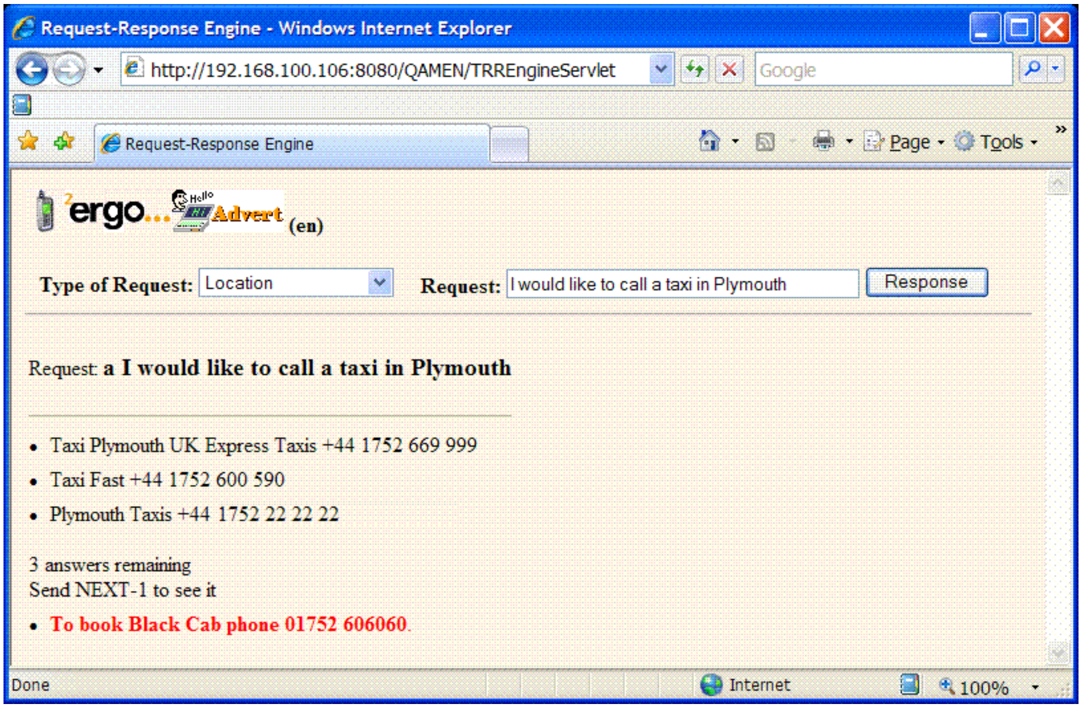


Figure 4. Example of Advert

Algorithm of Advert Selection

Let us consider the algorithm of advert selection:

1. QAMEN should be ready for advert i.e. command ***Advert ON*** should be sent from the mobile phone.
2. User type in and send the request e.g. *I would like to call a taxi in Plymouth* (see Figure 4).
3. QAMEN extracts KW from the request (RKW) i.e. *call*, *taxi* and *Plymouth*.
4. Using advert dictionary QAMEN extracts from ADB the list of advert patterns and calculates their "weights" i.e. APW. For APW calculation it is just enough to find intersection RKW with AP from ADB. If RKW belongs to SKW then 2 points is added to APW. In all other cases i.e. when RKW represents ***Areas***, ***Seasons***, ***Location*** explicitly, or belong to AKW just 1 point is added to APW. For the request in the example APW=3 but for request: "*I'm looking for phone No to book a taxi in Plymouth*" APW=6.
5. QAMEN makes descending sort of AP in accordance with their APW. If two AP have the same APW the ***Priority*** of AP defines the order. If ***Priority*** is equal the ***Quantity*** of AP and ***Date*** and ***Time*** when adverts have been displayed the last time define their order.
6. QAMEN calculates the free space for advert and in the result of calculation select an appropriate version of advert (see Figure 4).
7. If Web searching produces no response then just the advert can be displayed (see Figure 5).

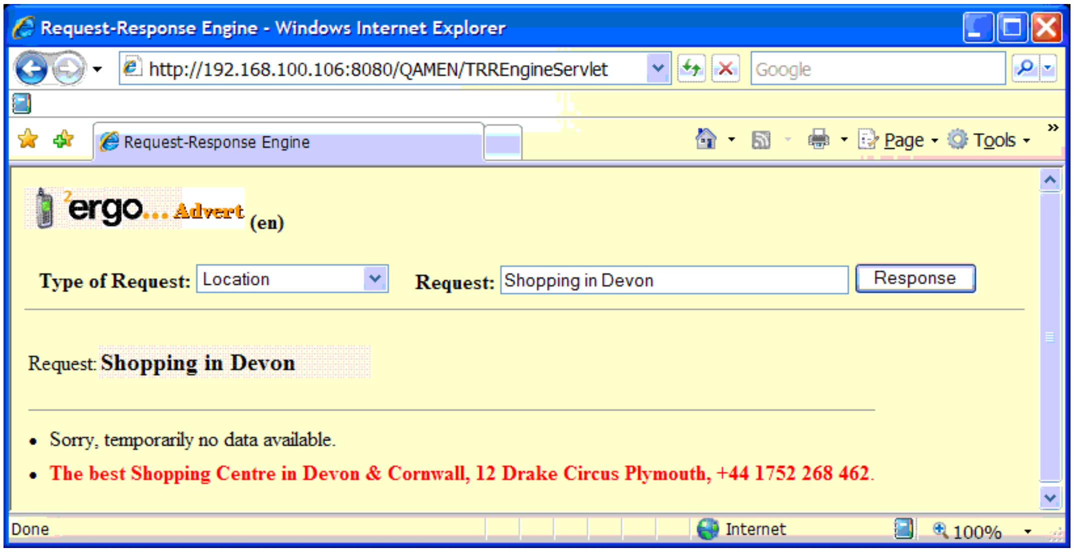


Figure 5. Web searching was not a success

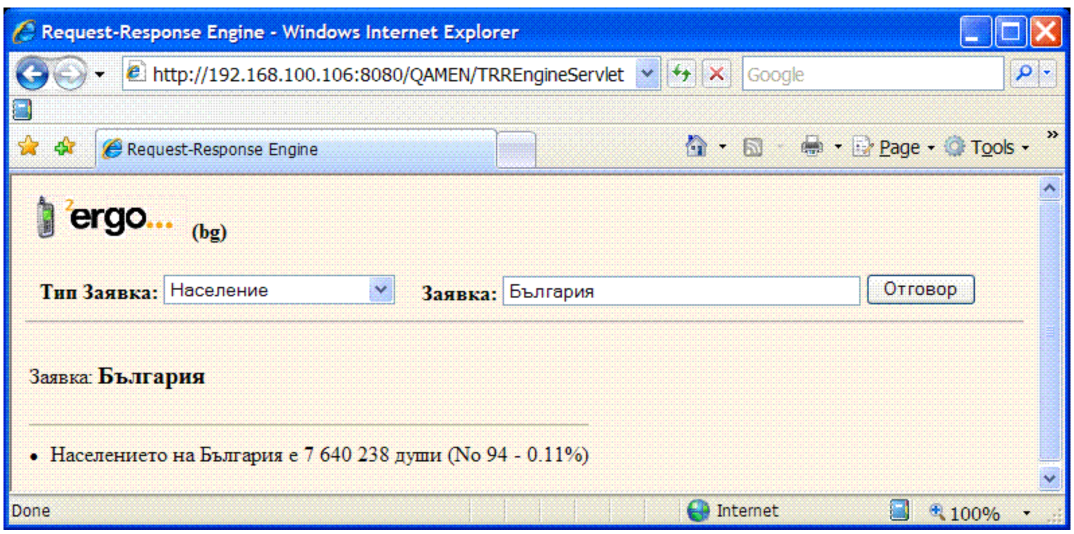


Figure 6. Bulgarian version of QAMEN

Conclusion

We believe that mobile search and mobile advertising offers a significant opportunity for many players in the wireless industry. While consumers across the world use Internet search applications as their primary access point to information and web sites, mobile users are just starting to experiment with the potential of mobile search and advertising applications. In this paper we describe our vision of mobile search and advertising to improve the efficiency of SMS. We turned our particular attention towards mobile advertising and its possible application in a mobile environment. The object of our research is to improve query response by adding adverts. It is important to offer and realize some ideas (not necessarily the best) when there are as yet no standards for representing mobile adverts. Of course, the ultimate criterion of "good" mobile search and advertising is that a user should be satisfied with search results and subscribe to the content of adverts. Future studies into this aspect of mobile phone use could be conducted as a series of further case studies in other countries in order to facilitate comparison: we have already started to create the Cyrillic version of QAMEN. The first step has been made in the creation of a Bulgarian version of QAMEN (see Figure 6).

Bibliography

[1] G.Coles, T.Coles, V.A.Lovitskii, "Natural Interface Language", *Proc. of the VIII-th International Conference on Knowledge-Dialogue-Solution: KDS-99*, Kacivelli (Ukraine), 104 -109, 1999.

[2] T.Coles, V.A.Lovitskii, "Text Searching and Mining", *Journal of Artificial Intelligence, National Academy of Sciences of Ukraine, Vol 3,* 488-496, 2000.

[3] D.Burns, R.Fallon, P.Lewis, V.Lovitskii, S.Owen, "Verbal Dialogue Versus Written Dialogue*", International Journal "Information Theories & Applications", Vol 12(4)*, 369-377, 2005.

[4] Ken Braithwaite, Mark Lishman, Vladimir Lovitskii, David Traynor, "Distinctive Features of Mobile Messages Processing",*International Journal "Information Theories & Applications", Vol 14(2)*, 154-160, 2007.

[5] Guy Francis, Mark Lishman, Vladimir Lovitskii, Michael Thrasher, David Traynor, "Instantaneous Database Access", *International Journal "Information Theories & Applications", Vol 14(2)*, 161-168, 2007.

[6] Vladimir Lovitskii, Michael Thrasher, David Traynor, "Automated Response To Query System", *International Journal "Information Theories & Applications", Vol 15(2)*, 143-152, 2008.

[7] Lee Johnston, Vladimir Lovitskii, Ian Price, Michael Thrasher, David Traynor, "Personalized Question-Answering Mobile System", International Book Series: "*Information Science and Computing",* book 2: "Advanced Research in Artificial Intelligence", 123-132, 2008.

[8] Loren Baker, "Mobile Search Advertising", *Search Engine Journal*, February 2008 http://www.searchenginejournal.com/mobile-search-advertising-to-hit-38-billion-by-2012-14-billion-in-us/6382/

[9] http://www.mmaglobal.com/uploads/MMAMobileSearchIntro.pdf.

[10] "Mobile Search and its Implications for Search Engine Marketing" provided by http://www.oneupweb.com

[11] Eden Zoller, Ovum, http://www.ovum.com/go/content/c,432,35105

[12] http://www.iabmexico.com/downLoads/mobileadoverview.pdf.

Authors information

***Vladimir Lovitskii** – 2 Ergo Limited, 4th Floor, Digital World Centre, 1 Lowry Plaza, The Quays, Salford, Manchester, M50 3UB, UK, e-mail: vladimir.lovitskii@fsmail.net*

***Colin McCaffery** – 2 Ergo Limited, 4th Floor, Digital World Centre, 1 Lowry Plaza, The Quays, Salford, Manchester, M50 3UB, UK, e-mail: colin.mccaffery@2ergo.com*

***Michael Thrasher** – University of Plymouth, Plymouth, Devon, PL4 6DX, UK, e-mail: mthrasher@plymouth.ac.uk*

***David Traynor** – 2 Ergo Limited, 4th Floor, Digital World Centre, 1 Lowry Plaza, The Quays, Salford, Manchester, M50 3UB, UK, e-mail: david.traynor@2ergo.com*

***Peter Wright** – 2 Ergo Limited, 4th Floor, Digital World Centre, 1 Lowry Plaza, The Quays, Salford, Manchester, M50 3UB, UK, e-mail: peter.wright@2ergo.com*

TABLE OF CONTENTS OF VOLUME 3, NUMBER 3