Informatica

An International Journal of Computing and Informatics

Special Issue:
   Media in Information Society
Guest Editor:
   Jozsef Györkös

1977

Informatica

An International Journal of Computing and Informatics

Introduction:
Media in Information Society

The impact of contemporary technological/technical solutions to the contents of the message passed by the media is increasing. The Internet technology integrates different types of media and changes their traditional/basic form, impact and availability. Activities - yesterday treated as a "special effect" - are technically feasible - not only for experts but also for ordinary users - and can be an incorporated part of the message.

Main topics of the special issue discuss the media in information society from the view of technology, message itself and the user:

- Traditional media in information society

- Information sharing through the web – the availability and reliability questions

- Integrative internet technologies and their effect

- User/consumer incorporation – real interactivity

- Media competence – social resource to be developed

- Contents – the new challenge of media development

- Winners and losers – how to solve the knowledge gap

This special issue is based on the conference organized in the framework of the Multi-conference "Information Society" at the Slovene Festival of Science.

We hope that presented papers will serve as a reference to new ideas and solutions.

Jozsef Györkös
gyorkos@uni-mb.si

An Automatically Refereed Scholarly Electronic Journal: Formal Specifications

Stefano Mizzaro and Paolo Zandegiacomo Riziò
Department of Mathematics and Computer Science
University of Udine
Via delle Scienze, 206 — Loc. Rizzi, I 33100 Udine, Italy
e-mail: mizzaro@dimi.uniud.it, zandegia@tagliamento.sci.uniud.it
www: http://www.dimi.uniud.it/~mizzaro

Internet growth seems to amplify the critiques to peer review mechanism: many researchers maintain that Internet would allow a faster, more interactive, and more effective model of publishing. However, just removing peer review would lead to a lack of quality control in scholarly publications. We propose a new kind of electronic scholarly journal, in which the standard submission-review-publication process is replaced by a more democratic approach, based on judgments expressed by the readers. The new electronic scholarly journal is described in both intuitive and formal ways.

1 Introduction: scholarly journals and peer review

The communication mechanism that modern science still adopts nowadays arose in the 17th Century, with the publication of the first scientific journals reporting, in paper form, the ideas, discoveries, inventions, of researchers. Nowadays, since about 1930, the dissemination of scholarly information is based on *peer review*: the researcher that wants to disseminate her work writes a paper and submits it to a scholarly journal; the paper is not immediately published, but it is judged by some referees; if they judge it adequate, the paper is published.

The peer review mechanism ensures a reasonable *quality* of the published papers, and it is usually retained an adequate solution, though not the ideal one. Indeed, the peer review mechanism can be (and has been) criticized. Sometimes, the reviewing process takes too long, even one or two years, so that the published paper describes something old. Sometimes the reviewers do not do a good job, accepting a bad paper or not accepting a good one, that after two years cannot be resubmitted because too obsolete. Sometimes, referees introduce some bias in published papers: for instance, in medicine field, papers describing negative results seem more difficult to publish than paper describing positive ones. And one might go on.

Internet has changed, and is changing, this situation [1, 2, 7, 11]. A peer reviewed journal can be distributed by electronic means. The refereeing process too can take place completely electronically, drastically reducing time and money: see, *e.g.,* JHEP (http://jhep.sissa.it) or Earth Interactions [5] (http://EarthInteractions.org). Multimedi-

ality can lead to a more effective communication [5]. Of course, there are also some drawbacks of electronic journals (copyright problems, legal validity, accessibility, and so on), and they seem not to have a large impact by now [4], but the general feeling is that this is a temporary situation, and we just have to wait some years for overcoming these temporary problems.

Internet growth seems to amplify the critiques to peer review: many researchers maintain that Internet would allow a more fast, elastic, interactive, and effective model of publishing. Nadasdy [10] suggests substituting peer review with democracy: each submitted paper is immediately published and readers will judge it, selecting what they deem useful. Of course, the problem with this approach is that the readers may not be capable of correctly judging the paper: whereas the referees are chosen among the experts in the field, everybody can read and judge a paper published on Internet.

Nadasdy's proposal is not an abstract one. In a few years, this model of publishing might become a *de facto* standard, as witnessed by two examples, already existing today: the "do it yourself publishing" (authors publishing in a web site their ideas), and public repositories of scholarly papers (where authors can publish papers classified in some categories—see, *e.g.,* http://ArXiv.org). The threat that these (without-quality-control) mechanisms will replace the (with-quality-control) peer review journals is a real one.

A proposed solution is to replace peer review with *peer commentary*: readers will write in a public commentary their judgments on the read papers. It seems a viable solution, but Harnad, after some practical experience with this

solution, says that "peer commentary is a superb supplement to peer review, but it is certainly no substitute for it" [3].

We propose a more sophisticate mechanism. We describe a new kind of electronic scholarly journal, with the aim of changing the submission-review-publication process, making it more automatic, keeping anyway at a high level the quality of scientific papers, and providing also a way of measuring in an automatic and objective way the quality of *researchers*, extending and improving the well known *impact factor* mechanism (http://www.isinet.com/hot/essays/7.html). We try to make a step further on the road suggested by the not refereed journals just mentioned, and to present a mechanism that avoids some of the previously described problems.

This paper, that extends and revises previous work [8, 9], is structured as follows. In Section 2, the mechanism is described in an intuitive way. In Section 3, the behavior of the whole system is formally defined by means of some formulæ. Section 4 discusses some open problems and future developments.

2   General description

The basic idea is the following. Imagine a scholarly journal in which each paper is immediately published after its submission, without a refereeing process. Each paper has a score, measuring its quality. This score is initially zero, and is later dynamically updated on the basis of the readers' judgments. A subscriber of the journal is an author or a reader (or both). Each subscriber has a score too, initially zero and later updated on the basis of the activity of the subscriber (if the subscriber is both an author and a reader, she has two different scores, one as an author and one as a reader). Therefore, the scores of subscribers are dynamic too, and change accordingly to subscribers' behavior: if an author with a low score publishes a very good paper (*i.e.*, a paper judged very positively by the readers), her score increases; if a reader expresses an inadequate judgment on a paper, her score decreases accordingly, and so on.

Every object with a score (author, reader, paper) has also a *steadiness* value, that indicates how much steady the score is: for instance, old papers will have a high steadiness; new readers (authors) will have a low steadiness. Steadiness affects the score update: a low (high) steadiness allows quicker (slower) changes of the corresponding score. A steadiness value increases as the corresponding score changes.

While time goes on, readers read the papers, judgments are expressed, and the corresponding scores and steadinesses vary consequently. The score of a paper can be used for deciding to read or not to read that paper; the scores of authors and readers are a measure of their research productivity, then they will try to do their best for keeping their score at a high level, hopefully leading to a virtuous circle (publishing good papers and giving correct judgments to the read papers). A steadiness value is an estimate of how stable (and, therefore, reliable) the corresponding score is.

For understanding the details of the automatically refereed journal proposed here, let's follow the events happening when a paper is read and judged by a reader. The following scores and steadinesses change:

- **Paper.** The paper score is updated: if the judgment is lower (higher) than the actual paper score, the paper score decreases (increases). The score of the reader determines the weight of the judgment: judgments given by higher rated readers will be more important (will lead to higher changes) than judgments given by lower rated readers.

  The steadiness of the paper increases, since the score of the paper is now computed on the basis of one more judgment, and is therefore statistically more reliable.

- **Author.** The author's score is updated: when the score of a paper written by an author decreases (increases), the score of the author decreases (increases). Thus, authors' scores are linked to the scores of their papers.

  The steadiness of the author increases, since the score of the author is now obtained with one more judgment and is therefore statistically more reliable.

- **Reader.** The reader's score is updated: if one reader's judgment about a document is "wrong" (too far from the average), the reader's score has to decrease. Then, the reader's score is updated depending on the goodness of her judgment (how much adequate her judgment is, or how much it agrees with the current score of the paper).

  The steadiness of the reader increases, since her score, computed on the basis of the goodness of her judgments, is obtained on the basis of one more judgment.

- **Previous readers.** The scores of the readers that previously read the same paper are updated: if a judgment causes a change in a paper score, all the goodnesses of the previously expressed judgments on that paper have to be re-estimated. Therefore, a judgment on a certain paper leads to an updating of the scores of all the previous readers of that paper.

  The steadinesses of the previous readers increase since the goodnesses of the readers, that lead to their scores, are obtained on the basis of one more judgment.

The updating of the scores of the previous readers deserve further explanation. After the paper score has changed, it is possible to revise the goodness of the old readers' judgments, and to update the old readers' score consequently: for instance, if an old reader r expressed a judgment j that was "bad" (distant from the paper score) at that time, but after that the paper score changes and becomes more similar to j, then the score of r (s_r) has to increase. Let us take into account a simple concrete example (see Figures 1, 2, and 3):
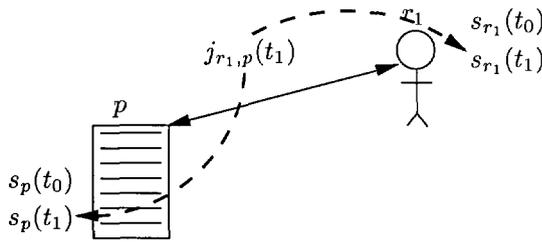
Figure 1: The updating of previous readers' scores: t_1.
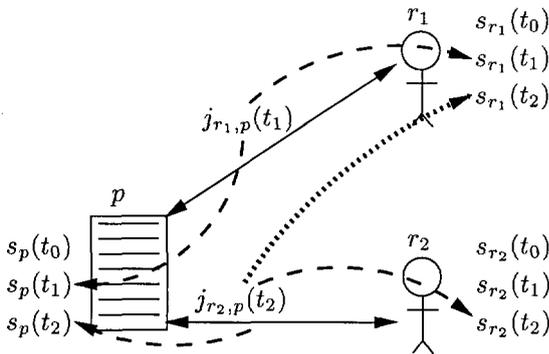
Figure 2: The updating of previous readers' scores: t_2.
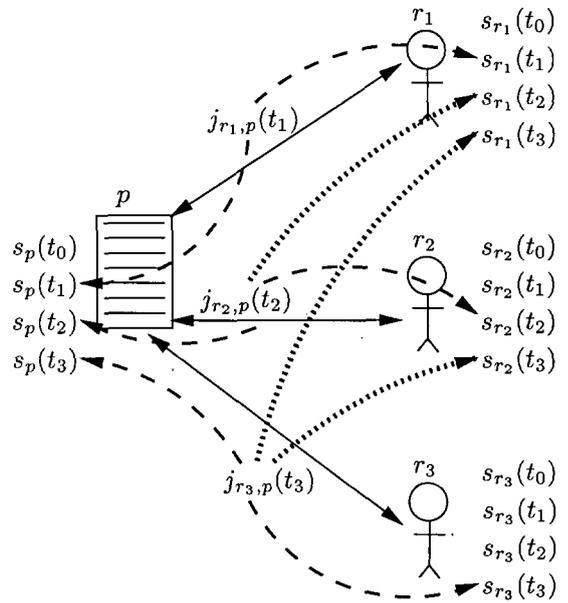
Figure 3: The updating of previous readers' scores: t_3.

- At time t_0, we have a paper p with score $s_p(t_0)$, three readers r_1, r_2, and r_3 with their scores $s_{r_1}(t_0)$, $s_{r_2}(t_0)$, and $s_{r_3}(t_0)$.

- At the following time instant $t_1 > t_0$ (Figure 1), reader r_1 reads paper p expressing the judgment $j_{r_1,p}(t_1)$ (continuous double arrow line in figure). This causes the updating of the scores of p and r_1 (dashed line in figure): we obtain $s_p(t_1)$ and $s_{r_1}(t_1)$.

- At time $t_2 > t_1$ (Figure 2), reader r_2 reads p expressing $j_{r_2,p}(t_2)$. The scores of p and r_2 are updated consequently, leading to $s_p(t_2)$ and $s_{r_2}(t_2)$. But also the score of r_1 has to be updated (dotted line in figure), since the goodness estimated at time t_1 for $j_{r_2,p}(t_2)$ with respect to $s_p(t_1)$ has to be re-estimated now that the score of p is $s_p(t_2)$.

- At time $t_3 > t_2$ (Figure 3), r_3 reads p expressing $j_{r_3,p}(t_3)$. This changes the score of p ($s_p(t_3)$), the score of r_3 ($s_{r_3}(t_3)$), and the scores of the previous two readers ($s_{r_2}(t_3)$ and $s_{r_1}(t_3)$).

In other words, the goodness of a reader's judgment is an approximation of the *ideal goodness*, defined as the difference between the reader's judgment and the final score of the paper (*i.e.*, the score obtained when the last judgment on that paper has been expressed). Since the final score is obviously not available when the judgment is expressed, it has to be guessed (updating of the reader), but this guess is revised and refined as time goes on and tends to $= +\infty$ (updating of previous readers).

3 Invariant properties

In this section, we will present some formulæ, in order to formally specify how to compute the values of the scores and steadiness of paper, author, reader, and previous readers, depending on the expressed judgments.

Let's start with some notation. We will denote with:

- t and t_i the discrete time instants. We assume that t_{i+1} immediately follows t_i, and that between t_i and t_{i+1} only the explicitly specified events will happen.

- $s_r(t)$, $s_p(t)$, $s_a(t)$ the score of a reader, a paper, and an author, respectively, at time t. We will sometimes omit the time indication when this does not rise ambiguity. The values for $s_p(t)$ and $s_a(t)$ are in the range $[0, 1]$ (0 is the minimum and 1 the maximum), whereas the values for $s_a(t)$ are in the range $]0, 1]$. This difference will be explained in Section 3.3.

- $\sigma_r(t)$, $\sigma_p(t)$, $\sigma_a(t)$ the steadiness of a reader, a paper, and an author, respectively, at time t. All the steadiness values are in the range $[0, +\infty[$.

- $j_{r,p}(t)$ the judgment expressed at time t by reader r on paper p.

- $t_{r,p}$ the time instant of the judgment expressed by r on p (we are implicitly assuming that each reader can judge each paper only once). We will write $j_{r,p}$ instead of $j_{r,p}(t_{r,p})$.

3.1 Paper

Given a paper p, its score is the weighted mean of the judgments previously expressed by readers on p. The weight of

each judgment is the score that the reader has when she expresses the judgment, to give more importance to the judgments given by better readers.

Definition 1 (Paper score s_p) Given a paper p, the set $R_p(t)$ of readers that expressed a judgment on p before time t, and the time instants of judgments expressions $t_{r,p}$, we have $\forall r \in R_p(t)$ the judgment $j_{r,p}$ expressed by r on paper p and the score $s_r(t_{r,p})$ of r at time $t_{r,p}$, i.e., the score that the reader has when she expresses the judgment. We can then define the score of paper p at time t as:

$$s_p(t) = \frac{\displaystyle\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p}}{\displaystyle\sum_{r \in R_p(t)} s_r(t_{r,p})}. \tag{1}$$

■

Remark 1 Consistently with Formula 1, the score of the paper before any judgment is expressed on it is zero.

Remark 2 The score $s_p(t)$ of a paper is modified only when a judgment on p is expressed.

Remark 3 In Formula 1, each judgment is weighted on the basis of the score that the reader has when she expresses the judgment $(s_r(t_{r,p}))$. The alternative of using the score that the reader has "now", i.e., when the mean is calculated $(s_r(t))$, seems less preferable since the reader's competence has probably changed in this period.

Let's now see how to measure the steadiness of a paper. The steadiness of a paper has to measure how stable its score is. To define it, a first attempt might be as the number of judgments expressed on that paper. However, it seems reasonable that a judgment expressed by a good reader should be more important, and give more steadiness to the paper, than a judgment expressed by a reader with a low score. Therefore, we define the steadiness of paper p at time t as the summation of the scores that readers have when they express their judgments on p.

Definition 2 (Paper steadiness σ_p) Given a paper p, the set $R_p(t)$ of readers that expressed a judgment on p before time t, and the time instants of judgment expression $t_{r,p}$, with $t_{r,p} \leq t$, the steadiness of p at time t is:

$$\sigma_p(t) = \sum_{r \in R_p(t)} s_r(t_{r,p}). \tag{2}$$

■

Remark 4 The steadiness value of a just published, and not yet judged, paper is zero.

Remark 5 The expression in the denominator in Formula 1 is exactly the steadiness of the paper. Therefore, we can rewrite Formula 1 as:

$$s_p(t) = \frac{\displaystyle\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p}}{\sigma_p(t)}. \tag{3}$$

3.2 Author

Given an author a, her score at time t can be defined in two equivalent ways:

- As the weighted mean of the scores of the papers previously published by a. The weight of each paper p is the steadiness of p, a value that sums up all the scores of the readers that expressed a judgment on p (see Formula 2).

- As the weighted mean of the judgments previously expressed by readers on the papers published by a. The weight of each judgment is, again, the score that the reader has when she expresses the judgment.

Let's define formally the first alternative, that uses the steadiness of a paper to weight the papers scores.

Definition 3 (Author score s_a) Given an author a and the set $P_a(t)$ of papers published by a before time t, we have $\forall p \in P_a(t)$ the score $s_p(t)$ of p at time t and the steadiness $\sigma_p(t)$ of p at time t. We can now define the score of author a at time t as:

$$s_a(t) = \frac{\displaystyle\sum_{p \in P_a(t)} \sigma_p(t) \cdot s_p(t)}{\displaystyle\sum_{p \in P_a(t)} \sigma_p(t)}. \tag{4}$$

■

Following the second alternative, we can define:

$$s_a(t) = \frac{\displaystyle\sum_{p \in P_a(t)} \left(\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p} \right)}{\displaystyle\sum_{p \in P_a(t)} \left(\sum_{r \in R_p(t)} s_r(t_{r,p}) \right)}, \tag{5}$$

where $P_a(t)$ is the set of papers published by a before time t, $R_p(t)$ is the set of readers that judged paper p before t, $s_r(t_{r,p})$ is the score of r at time $t_{r,p}$, $t_{r,p} \leq t$ are time instants of judgment expression, and $j_{r,p}$ is the judgment expressed by r on paper p.

Remark 6 Using Formulæ 2 and 3 to rewrite the summations in parentheses in Formula 5, it is easy to see that the score of an author computed with Formulæ 4 and 5 are equivalent.

Remark 7 Consistently with Formulæ 4 and 5, the score of the author before any judgment is expressed on her papers is zero.

Remark 8 $s_a(t)$ is modified only when the score of one of the papers published by a changes, *i.e.*, when a judgment on a paper published by a is expressed.

Remark 9 As discussed in Remark 3, in Formula 5 each judgment is weighted on the basis of the score that the reader has when expresses the judgment $(s_r(t_{r,p}))$.

The steadiness of an author has to measure how stable her score is. We can define it equivalently in two ways, as the summation of the steadiness of her papers and as the summation of the scores that the readers had when they expressed a judgment on a paper of the author. The first alternative leads to the following definition.

Definition 4 (Author steadiness σ_a) Given an author a, the set of papers $P_a(t)$ published by a, and the steadiness $\sigma_p(t)$ of each paper $p \in P_a(t)$ at time t, the steadiness of author a at time t is:

$$\sigma_a(t) = \sum_{p \in P_a(t)} \sigma_p(t). \tag{6}$$

■

Following the second alternative, we can define (with the usual notation):

$$\sigma_a(t) = \sum_{p \in P_a(t)} \sum_{r \in R_p(t)} s_r(t_{r,p}). \tag{7}$$

The equivalence of Formulæ 6 and 7 follows immediately from Formula 2.

Remark 10 Since the denominator in Formula 4 is exactly the steadiness of the author (Formula 6), we can rewrite Formula 4 as:

$$s_a(t) = \frac{\displaystyle\sum_{p \in P_a(t)} \sigma_p(t) \cdot s_p(t)}{\sigma_a(t)}. \tag{8}$$

3.3 Reader

First of all, we need to define a measure of how good a judgment is.

Definition 5 (Goodness) Given the score $s_p(t)$ of a paper p at time t and the judgment $j_{r,p}$ by reader r on p, the goodness, calculated at time t, of the judgment $j_{r,p}$ is defined as:

$$g_{j_{r,p}}(t) = 1 - |j_{r,p} - s_p(t)|. \tag{9}$$

■

Remark 11 The goodness depends on the time instant at which it is computed (t in Formula 9), since s_p can change also after the judgment expression (because of other judgments).

Given a reader r, her score is the weighted mean of the goodness of the judgments she has previously expressed. If we gave the same weight to all the goodnesses, we might define:

$$s_r(t) = \frac{\displaystyle\sum_{p \in P_r(t)} g_{j_{r,p}}(t)}{|P_r(t)|} \tag{10}$$

where $s_r(t)$ is the score of reader r at time t, $P_r(t)$ is the set of papers judged by r before t, $|P_r(t)|$ is the cardinality of $P_r(t)$, $j_{r,p}$ is the judgment expressed by r on paper p, and $g_{j_{r,p}}(t)$ is the goodness, calculated at time t, of the judgment $j_{r,p}$.

However, the weight of each goodness should be different on the basis of the steadiness of the paper being judged. For instance, a wrong judgment on a paper that has been previously judged by many readers should weight more than a wrong judgment on a paper whose score is calculated on the basis of very few judgments. The steadiness of the paper being judged should be taken into account, and the previous Formula 10 for $s_r(t)$ has to be rewritten as follows.

Definition 6 (Reader score s_r) Given a reader r and the set $P_r(t)$ of papers judged by r before time t, we have $\forall p \in P_r(t)$ the steadiness $\sigma_p(t)$ of p at time t, the judgment $j_{r,p}$ expressed by r on p, and the goodness $g_{j_{r,p}}(t)$ of the judgment $j_{r,p}$, calculated at time t. We can define the score of reader r at time t as:

$$s_r(t) = \frac{\displaystyle\sum_{p \in P_r(t)} \sigma_p(t) \cdot g_{j_{r,p}}(t)}{\displaystyle\sum_{p \in P_r(t)} \sigma_p(t)}. \tag{11}$$

■

Remark 12 The initial value of s_r should be, accordingly to Formula 11, zero. But this would lead to a division by zero in Formula 1 (if we compute the score of a paper that has not been judged yet). We therefore define this initial value as $s_r = \epsilon$, where ϵ is a small value that will be neglected after some judgments are expressed.

Remark 13 The score of a reader r changes when:

- r judges a paper. This adds one element to the set $P_r(t)$ and one new addendum to the summations in Formula 11.

- A reader $r' \neq r$ judges a paper that has previously been judged by r. This causes the score of the paper to change, and this leads to changing the goodness of the judgment previously expressed by r.

Remark 14 All the goodnesses are calculated at time t, to have the better estimate that is possible. We are implicitly assuming that the estimate improves with the number of judgments.

Remark 15 For the same reason, the steadinesses are calculated at time t too: in Formula 11, we have $\sigma_p(t)$, not $\sigma_p(t_{r,p})$.

The steadiness of a reader has to measure how stable the score of a reader is. A good measure might be the number of judgments expressed by a reader. However, since a judgment expressed on a paper with a high steadiness value should give more steadiness to the reader, we define the steadiness of a reader as the summation of the steadinesses of the papers judged by her.

Definition 7 (Reader steadiness σ_r) If $P_r(t)$ is the set of papers judged by a reader r before time t, the steadiness of r at time t is

$$\sigma_r(t) \;=\; \sum_{p \in P_r(t)} \sigma_p(t). \tag{12}$$

■

Remark 16 The denominator in Formula 11 is the steadiness of the reader (right hand side in Formula 12). Therefore, Formula 11 can be rewritten as:

$$s_r(t) \;=\; \frac{\displaystyle\sum_{p \in P_r(t)} \sigma_p(t) \cdot g_{j_{r,p}}(t)}{\sigma_r(t)}. \tag{13}$$

Remark 17 Although the steadiness of a reader is not used in the other formulæ, it is however useful to judge the reliability of a reader.

4   Conclusions and future developments

We have described, both intuitively (Section 2) and formally (Section 3), an electronic scholarly journal in which the standard peer review process is replaced by a more democratic approach based on judgments expressed by the readers.

The formulæ in Section 3 are different from what presented in [8, 9]: the formulæ presented here define some invariant properties of the system, that must hold at each time instant, whereas the formulæ in [8, 9] define how to update the values of score and steadiness of papers, authors, and readers as time goes on. The invariant properties are simpler than the updating formulæ, and thus it is straightforward to understand and justify them from an intuitive point of view; the correctness of the updating formulæ is more difficult to grasp at an intuitive level, but the equivalence of the two approaches can be formally proved.

Moreover, because of some improvements and simplifications, the formulæ presented here are not fully equivalent to those in [8, 9].

Generally speaking, this proposal can be seen as an improvement of the dissemination of scholarly information through on line journals. More specifically, it can be seen as an improvement of the well known impact factor mechanism, of the democratic journal proposed in [10], and of collaborative information retrieval and filtering [6], since it allows distinguishing among "good" and "bad" collaborators. Two other proposals that are similar, but less formalized, are those by Stodolsky [12] and by Varian [13].

This proposal is not free from problems. In general, one may wonder if democracy is a good approach to knowledge dissemination. Of course, it is difficult to have an objective opinion on that: it could be appropriate, or appropriate in some fields only, or not appropriate at all. We believe that only by further studies and experiments we can find an objective answer. However, it has to be emphasized that the mechanism proposed here is different from democracy, since different readers will have different importance. More specific problems are briefly discussed in the following.

A problem is *lazy readers*: a reader can simply confirm the previously expressed judgments, giving to each read paper a score equal to its actual score, with the aim of obtaining a high goodness. Two solutions seem suited here: give higher scores to fast readers (those that first read the papers), and do not show the paper score for a period after its publication (for instance, until when its steadiness reaches a certain value). Moreover, the strategy adopted by such a lazy reader might be a bad one, and lead to a low score for her, if the judgments by the following readers change the score of the paper.

Another problem might be the *lobbies*, *i.e.*, people that agree in mutually giving high scores. This might not be a problem at all, if the whole system can be modified to behave, by choosing appropriate formulæ and constants, in a way that discourages the lobbies; if this is not the case, the solution would probably be to implement some software able to detect such situations.

Of course there are technical difficulties too, *e.g.*, the identification of subscribers, or the huge amount of storage needed for recording the papers, the subscribers' data, and the history of expressed judgments. But these can be surely handled by database and cryptography technologies.

It is also easy to see some mandatory *improvements*:

- To deal with papers with more than one author. That should be easy: the judgments on the paper cause a modification of all the authors' scores, weighted on the basis of the importance of the contribution of each author to the paper.

- To have more scores, both for subscribers (authors and readers) and papers: comprehensibility, technical soundness, originality, and so on. In this way, a more

detailed evaluation is available. If just one single number is needed, a weighted mean of all the scores of a subscriber (or a paper) can be used.

- To have more than one journal, with different acceptance thresholds: an author must have a score larger than the threshold for publishing. Younger researcher will subscribe to lower rated journals, and "first class" journals will accept only well established researchers. The mechanism presented in this paper can also be used as a complement of, instead of a replacement to, peer review: the initial score of a paper can be given by a standard peer review, thus allowing an author with a low score to submit her paper to an higher rated journal.

- To allow the subscribers know why their score is decreasing, *i.e.*, which "wrong" judgment, or paper, causes that, and eventually let them revise their judgment or withdraw their paper.

- To introduce some sort of *rent* function, for decreasing the score of subscribers that are inactive for long periods of time.

- To allow the readers to express, besides the numerical judgments also a free text *commentary* on the paper. The commentary can then be considered as a paper itself, and judged by other readers, but it is linked to the paper it comments, and the score of the commentary can affect the score of the paper.

Finally, we sketch how we intend to proceed with this research. A software simulator of the electronic journal proposed here is currently under implementation. We also plan to use some mathematical models and techniques for formally studying the behavior of the whole system and for studying other similar approaches. More in detail, some typical cases could be simulated, *e.g.*, what happens to a reader's score when she expresses a judgment that is correct but far from the average; it would be possible to understand if and how the order of the judgments expressed by the readers affect their final score, or if the initial conditions— *i.e.*, the number of initial readers, their score, and so on— are a critical factor for having a stable system; we might use a bet-like approach, in which each reader has some money and bets on some papers; and so on.

These theoretical and experimental activities will allow to verify that the behavior of the system is correct and consistent and to choose in a more reliable way among the possible formulæ and parameters. After that, the software for the complete system will be implemented, tested, and evaluated. An ideal environment for these experiments is a repository of preprints, like ArXiv (http://ArXiv.org/). We plan to execute some laboratory experiments (with simulated papers, authors, and readers) and some real life experiments (involving real users), mostly within the TIPS project (http://tips.sissa.it), recently funded in the 5th Framework Program of the EU.

These theoretical and experimental activities will also allow to determine the relations between the values of some parameters and the real situation: indeed, it is likely that the values of parameters depend on figures like the number of subscribers, the rate of papers publishing, the rate of judgments expression, and so on. These dependencies must be singled out and, in case the above presented formulæ turned out to be inadequate, new ones need to be proposed and evaluated. Again, ArXiv seems the ideal environment for this issue.

Acknowledgements

References

[1] C. W. Bailey, Jr. Scholarly electronic publishing bibliography—Version 32: 8/1/00. Web document, visited 3/9/00. http://info.lib.uh.edu/sepb/sepb.html.

[2] C. W. Bailey, Jr. Scholarly electronic publishing on the internet, the NREN, and the NII: Charting possible futures. *Serials Review*, 20(3):7–16, 1994. http://info.lib.uh.edu/cwb/schpub.htm.

[3] S. Harnad. Learned inquiry and the net: The role of peer review, peer commentary and copyright. *Antiquity*, 71:1042–1048, 1997. ftp://ftp.princeton.edu/pub/harnad/Harnad/HTML/harnad97.antiquity.html.

[4] S. P. Harter. Scholarly communication and electronic journals: An impact study. *Journal of the American Society for Information Science*, 49(6):507–516, May 1998.

[5] J. Holoviak and K. L. Seitter. Earth Interactions— Trascending the limitations of the printed page. *JEP the Journal of Electronic Publishing*, 3(1), 1997. ISSN 1080-2711; http://www.press.umich.edu/jep/03-01/EI.html.

[6] M. Karamuftuoglu. Collaborative information retrieval: Toward a social informatics view of IR interaction. *Journal of the American Society for Information Science*, 49(12):1070–1080, 1998.

[7] M. Lesk. *Practical Digital Libraries: Books, Bytes, and Bucks*. Morgan Kaufmann, San Francisco, CA, 1997.

[8] S. Mizzaro. A concrete proposal for an automatically refereed scholarly electronic journal. Research Report UDMI/07/99, Department of Mathematics and Computer Science, University of Udine, Via delle

Scienze, 206 — Loc. Rizzi — Udine, Italy, April 1999. Presented at a Workshop at ECDL'99 Conference, Paris.

[9] S. Mizzaro. An automatically refereed scholarly electronic journal. In C. Bavec, M. Gams, V. Rajkovič, I. Rozman, and J. Györkös, editors, *Information Society 2000 — Proceedings of the Third International Conference*, pages 121–126, Ljubljana, Slovenia, 17–19 October 2000. ISBN: 961 6303-26-0.

[10] Z. Nadasdy. A truly all-electronic journal: Let democracy replace peer review. *JEP the Journal of Electronic Publishing*, 3(1), 1997.   ISSN 1080-2711; http://www.press. umich.edu/jep/03-01/EJCBS.html.

[11] V. H. Resh.   Science and communication: An author/editor/user's perspective on the transition from paper to electronic publishing.   *Issues in Science and Technology Librarianship*, 19, 1998.   ISSN 1092-1206;   http://www.library.ucsb.edu/istl/98-summer/article3.html.

[12] D. S. Stodolsky.   Consensus journals: Invitational journals based upon peer consensus. *Psycoloquy*, 1(15), 1990. Also appeared in: Datalogiske Skrifter (Writings on Computer Science), 29, 1990, Roskilde University Centre, Institute of Geography, Socioeconomic Analysis, and Computer Science (ISSN 0109-9779-29).

[13] Hal R. Varian.   The future of electronic journals.   *JEP the Journal of Electronic Publishing*, 4(1), 1998.   ISSN 1080-2711; http://www.press.umich.edu/jep/04-01/varian.html.

Digital Divide and the Gap Between North America and Europe in Internet Users per Capita: New Insight from Existing Data

Pavle Sicherl
SICENTER and Faculty of Law, University of Ljubljana
Brajnikova 19, 1000 Ljubljana, Slovenia
Phone: +386 1 5174 180, Fax: +386 1 5174 184
Pavle.Sicherl@sicenter.si

The emerging learning society requires a different development strategy, but also a conceptual and statistical framework that goes beyond the conventional approach and provides a broader dynamic framework for policy analysis. The novel time distance methodology is applied to the analysis of digital divide in the USA by income and educational level. It also shows that different statistical measures lead to diverse conclusions about the gap between North America and Europe in Internet users per capita. Percentage difference is decreasing, while time distance is increasing. Policy discussion relates to preparation of national strategy for information society in Slovenia.

1 Introduction and methodology

In the recent EU seminar on socio-economic research and European policy 'Towards a learning society: innovation, competence building with social cohesion in Europe', Lundvall [3] underlined two important characteristics of the emerging knowledge based economy: the major impact of the information technology revolution is that it speeds up the process of change in the economy, while the most important inherent contradiction of the learning economy has to deal with polarization and social exclusion. He concluded that under the new circumstances there is a need for reintegrating the strategies of firms, social partners and policy makers. We may add that now, more than ever, there is also a need that the conceptual and statistical framework employed to deal with the problems of interrelationship between growth and inequality goes beyond the conventional static approach and provides a broader dynamic framework for policy analysis and debate.

This paper delivers two messages. First, a novel broader dynamic conceptual and statistical framework can enhance the analysis and perception of the gap between two countries, firms or economic and social groups. By introducing time distance concept, and S-distance as a novel statistical measure, as an inherent dimension in the comparative analysis and benchmarking new insights from existing data can be provided. It will be shown that the degree of disparities may be very different in static terms and in time. The concept of time distance for a given level of the indicator can be usefully applied to many issues in economics, business, and statistics.

Second, the relation between the information society and government is dealt with at the broadest understanding of the term, i.e. the role of government in terms of establishing national strategy for development of the information society within the national consensus on economic and social development. It will be shown that in the USA as the leading country in the application of the information technology, as well as in the EU with the initiative entitled '*e*Europe An Information Society for All', policy decisions have been reached that the stakes for future development of society are so high that the development of information society cannot be left to spontaneous trends and has to be guided by conscious policy at the highest level in cooperation with all segments of the society.

A brief description of the time distance methodology will be presented first. Further details can be found in Sicherl [5, 6, 7]. The time perspective, which no doubt exists in human perception when comparing different situations, is systematically introduced both as a concept and as a quantifiable measure. Time distance in general means the difference in time when two events occurred. We define a special category of time distance, which is related to the level of the analysed indicator. The suggested statistical measure S-distance measures the distance (proximity) in time between the points in time when the two compared series reach a specified level of the indicator X. For a given level of X_L, $X_L = X_i(t_i) = X_j(t_j)$, and the S-distance, the time span separating unit (i) and unit (j) for the level X_L, will be written as

$$S_{ij}(X_L) = \Delta T(X_L) = t_i(X_L) - t_j(X_L)$$

where T is determined by X_L. S-distance has been generalized to complement conventional statistical measures in time series comparisons, regressions, models, forecasting and monitoring (Sicherl [5]). In this paper only the first application to comparisons between countries (regions) and socio-economic groups will be used. In practical applications it is important to distinguish backward looking (*ex post*) and forward looking (*ex ante*) time distances. They relate to different periods, past and future. Backward looking time distance belongs to the domain of statistical measures based on known facts, and there is no need to relate it to any static measure or growth rate. The second, the forward looking time distance, is important for describing the time distance outcomes of alternative assumptions about future developments or of alternative policy scenarios for the future.

2. Different statistical measures of the gap between North America and Europe in Internet users per capita show diverse conclusions

The indicator Internet users per capita is an example of an indicator with a much higher rate of growth than that of GDP per capita. As time distance is a negative function of the growth rate of the indicator, it is expected that the time distance for some of the information society indicators with very high growth rates will be substantially smaller than that for GDP per capita. This is so despite the fact that static disparities for some of these indicators are larger than that for GDP per capita. In this section the illustrative example is done for the indicator Internet users per 1000 people by region. Data from Computer Industry Almanac Inc. include also projections for this indicator until the year 2005. Despite the great importance of the information society indicators in shaping the tehnological change and the economy, the accuracy of international comparisons for these indicators leaves much to be desired. Our focus in this paper is in the first instance methodological, showing that different statistical measures lead to different perceptions about the magnitude and the direction of change in the gap between North America and Europe for Internet users per capita. Therefore we simply take the values and the geographical division from Computer Industry Almanac Inc. [14] for granted.

This data source shows that in 1998 the value for the indicator Internet users per capita was in North America (NAM) nearly three times higher than in Western Europe/Scandinavia (WES), and more than twenty times higher than in Eastern Europe (EE). This gap is in static terms much higher than that for GDP per capita. The time distance is roughly between 3 and 4 years between NAM and WES, as shown in Table 3. Table 2 shows the average growth rates for NAM, WES and EE for each of these three periods: the rate of growth of EE is higher than for WES, and that for WES is in turn higher than that for NAM. If one would be using only relative static measures of the gap the conclusion would be uncomplicated. Since the growth rate is higher for WES and EE than for NAM, the conclusion would be that the gap is decreasing and that convergence is well under way.

Table 1. Internet users per 1000 people by region

Time	North America	W.Europe /Scandinavia	Eastern Europe
1995	104.9	22.1	1.3
1998	311.2	105.8	13.0
2000	492.6	220.5	37.2
2005	720.6	529.9	157.7

Source: Computer Industry Almanac Inc. http://www.c-i-a.com/199908iu.htm

Table 2. Average percentage rate of increase per year

Period	North America	W.Europe /Scandinavia	Eastern Europe
1998/1995	44	69	117
2000/1998	26	44	69
2005/2000	8	19	33

The broader conceptual and statistical framework suggested here looks at the situation from more perspectives. Table 3 shows the numerical results for three views of the gap between the three regions, while Table 4 draws the main conclusions for the five statistical measures used in the discussion of the existing and projected gaps between NAM and WES for the Internet users per capita. In this case all three measures of the gap (absolute difference, percentage difference and S-distance) show diverse conclusions about the directions of the change of the gap between NAM and WES for the period 1998-2005.

Table 3. Three measures of differences in internet users per 1000 people between W.Europe/Scandinavia (WES=0) and North America and Eastern Europe

Time	Absolute difference		Percentage difference		Time distance in years	
	NAM	*EE*	*NAM*	*EE*	*NAM*	*EE*
1995	83	-21	375	-94		
1998	205	-93	194	-88	-3.0	
2000	272	-183	123	-83	-3.3	4.5
2005	191	-372	36	-70	-4.2	6.1

Table 4. Internet users per capita: comparing different views of the gap between North America and W. Europe/Scandinavia

Measure 1	Level of the indicator	*HIGHER IN NAM*	NAM > WES
Measure 2	Growth rate per year	*HIGHER IN WES*	NAM < WES
Measure 3	GAP: Absolute difference	*CONSTANT*	=
Measure 4	GAP: Relative difference	*DECREASING*	↓
Measure 5	GAP: Time distance	*INCREASING*	↑

Absolute difference is constant; at both compared points in time about 200 more people per 1000 people are Internet users in NAM as compared to WES. Percentage difference is decreasing substantially; in 2005 it is projected that the value of the indicator in NAM will be only 36 per cent higher than in WES. This degree of relative disparity would in 2005 be already lower than the disparity in GDP per capita in 1999. However, time distance would be increasing, in 1998 the time lead of NAM for Internet users per capita was according to this set of data 3 years, while for 2005 the projected time lead will be 4.2 years.

This is a surprising and counterintuitive conclusion that can be systematically obtained only within the outlined broader conceptual framework. The perception of and the conclusions about the degree of disparity that are based here on two-dimensional analysis of proximity (proximity in time and proximity in the indicator space) provide a better understanding of the situation. A new dimension is added to complement the conventional static measures of disparity, and all of the perspectives have to be studied simultaneously for a better perception of the reality.

The fact that growth rates for the same period are higher in Europe than in North America does not tell the whole story. When compared for given levels, at least until the projected level of 500 per 1000 people, the diffusion of Internet is shown to be always faster in North America than in WES. Such a conclusion is simply absent from the conventional measures. A country or a company that reduces the static percentage disparity by growing faster than its benchmark may erroneously believe that it is sufficiently improving its competitive position. However, in the present rapid changes in the economy it may be for a company much more important not to increase the lag in time behind its competition. In this numerical example the growth rate for Internet users per capita for Eastern Europe is much faster than for WES and NAM, but Eastern Europe is not growing fast enough to prevent the worsening of its position in two other views: absolute differences and time distances are increasing. It would be advisable that also the eEurope Initiative targets would be checked and that the implied reduction in various aspects of the gap with North America would be made explicit with the use of this broader framework. Such broader analysis could be useful also for market analysis of penetration rates for numerous products.

3. Digital divide in the USA by income and education level

U.S. Department of Commerce (1999) issued already the third report on the falling through the net, dealing with the digital divide by analyzing telecommunications and information gaps in America. In the nutshell, this report revealed that better educated Americans are more likely to be connected; the gap between high- and low-income Americans is increasing; whites are more likely to be connected than African-American or Hispanics; and rural areas are less likely to be connected than urban users. In this paper we shall touch only on one indicator, percentage rate of U.S. households with computers for selected years for the period 1984-1998, and observe what are the historical time distances for this indicator.

Table 5. Percent of U.S. Households with Computers By Education

Education	1984	1989	1994	1997	1998
Elementary	0.9	1.9	2.6	6.8	7.9
Some H.S.	2.3	4.5	6	10.9	15.7
H.S. Diploma	5.9	9.1	14.8	25.7	31.2
Some College	11.3	11.8	28.9	43.4	49.3
B.A. or more	16.4	30.6	48.4	63.2	68.7

Source: U.S. Department of Commerce (1999)

Table 6. Time Distance (in Years) for Digital Divide for Computer Penetration Rate from High School Diploma Level (- time lead, + time lag)

Education	1997	1998
Elementary	11.59	10.88
Some H.S.	6.42	3.75
H.S. Diploma	0	0
Some College	-3.94	-3.52
B.A. or more	-9.73	-8.83

Source: own calculation

This aspect of the digital divide will be illustrated for computer penetration rates in households by the level of education and by income. The 1998 level of households with computer for those with high school diploma corresponds with the 1989 level for the group with B.A. or more; the time lag is about 9 years. In turn, those with

Table 7. Percent of U.S. Households with Computers By Income

Income	1984	1989	1994	1997	1998
Under $5,000	1.6	5.8	8.4	16.5	15.9
5,000 - 9,999	1.7	3.7	6.1	9.9	12.3
10,000 - 14,999	3.3	4.5	8.2	12.9	15.9
15,000 - 19,999	5.3	8	11.7	17.4	21.2
20,000 - 24.999	8.1	9.6	15.2	23	25.7
25,000 - 34,999	11.7	14.6	19.8	31.7	35.8
35,000 - 49,999	17	22.5	33	45.6	50.2
50,000-74,999	22.4	31.6	46	60.6	66.3
75,000+	22.1	43.9	60.9	75.9	79.9

Source: U.S. Department of Commerce (1999)

only elementary education lag behind those with H.S. diploma for another 11 years. Roughly speaking, the time divide between those with B.A. or more and those with elementary education is about 20 years. This is a complementary view to the statements that the level in the highest group is eight times higher (relative difference) or that the absolute difference amounts to about 60 points in penetration rate.

Table 8. Time Distance (in Years) for Digital Divide for Computer Penetration Rate from the 20,000-24.999 level (- time lead, + time lag)

Income	1997	1998
Under $5,000	2.5	3.73
5,000 - 9,999	7.73	6.59
10,000 - 14,999	5.05	3.73
15,000 - 19,999	2.15	1.69
20,000 - 24.999	0	0
25,000 - 34,999	-2.19	-2.51
35,000 - 49,999	-7.76	-7.48
50,000-74,999	-12.67	-12.21
75,000+	-12.79	-13.17

Source: own calculation

The digital divide for computer penetration rate by income classes shows a similar time dimension. S-distance between the 75,000$+ class and the 20,000-24,999$ class is about 13 years, and for the highest and the lowest class between 17 and 20 years. Time distance of about 20 years shows, in addition to the analysis of static differences in the mentioned study, that the digital divide in the USA is substantial and that concerted action to reduce it is needed.

4. National strategy for information society

Both of the leading economic powers, USA and EU, have made a policy decision formulating an explicit strategy to enhance the conditions under which the benefits of the development of the information society

would come within the reach of all inhabitants. In the EU on December 8, 1999 the European Commission launched an initiative 'eEurope An Information Society for All'. The key objectives are:

1) Bringing every citizen, home and school, every business and administration into the digital age.

2) Creating a digitally literate Europe, supported by entrepreneurial culture to finance and develop new ideas.

3) Ensuring that the whole process is socially inclusive, builds consumer trust and strengthens social cohesion.

To achieve these objectives, the Commission proposes 10 priority areas for action with ambitious targets to be achieved through joint action of the Commission, Member States, industry and citizens of Europe (ISPO [15]). Further steps were taken under the Portuguese presidency.

In the USA President Clinton launched in April 2000 'National Call to Action' to close the digital divide. He challenged corporations and non-profit organizations to take concrete steps to meet two critical goals:

1) Provide 21st Century Learning Tools For Every Child in Every School.

2) Create Digital Opportunity For Every American and Community.

He announced that over 400 companies and non-profit organizations have signed a 'National Call To Action' to bring digital opportunity to youth, families and communities. The position is that access to computers and the Internet and the ability to effectively use this technology are becoming increasingly important for full participation in America's economic, political and social life. There is strong evidence of a 'digital divide', i.e. a gap between those individuals and communities that have access to these Information Age tools and those who do not (White House [12]). In December 1999 he directed five members of his Cabinet to take specific steps in their Departments to close the Digital Divide. He traveled around the country and participated in roundtable discussion with CEOs from the technology industry and leaders of the civil rights community and non-profit organizations, who pledged their contributions to this

action. Are there any lessons for the Slovenian government?

In Slovenia unfortunately some politicians and interest groups either because of lack of understanding of world trends or because of their particular interests entertain the view that even indicative planning of social and economic development is a characteristic of the former system and something which is not quite suitable for the market economy and political democracy. Consequently such state of mind in decisions on development policy underestimates the importance of professional know-how and objective elements and stretches of the role of political decision-making also to areas where this is not optimal. By doing this it introduces arbitrariness and inefficiency in the short run as well as in long-term decisions.

In professional circles there is substantial agreement that this situation is one of the most important obstacle for a more efficient development. Sicherl [4] states that the lack of coordination and social consensus on development startegy prevents that the synergy based on interrelationships in the economy and society could lead to better performance of Slovenia. In a study of development gaps between Slovenija and the EU Sicherl and Vahčič [8] put the achievement of social consensus for the strategy of development on the top of the list of factors needed to reduce the gap. Sočan [9] shows the importance of social consensus for successful development cases in international comparison. Banovec [1] emphasizes communication and dialog for the more demanding analytical use of information infrastructure. Caf [2] and Vehovar [11] also put emphasis on the lack of strategy and the need for a more proactive and consistent policy to promote further development of the information society.

The development of information technology competence in schools is a positive development, but otherwise there is an absence of active policy in the field of information technology, knowledge and its effective use as in many other areas. This field is a typical example of a necessary defensive investment: we have to invest in order to prevent falling behind the world; only later we will be able to discuss how much faster we should run to acquire a comparative advantage. The trend in the world is clearly reflected in the slogan at an Intel Conference:' every business= E-business, every home= E-home'. This will not be achieved tomorrow, but it is obviously one of great threats and opportunities of which the government and business are not fully aware of. From the point of view of the development strategy of Slovenia relying on the spontaneous development in this field cannot be an acceptable scenario.

However, all these opinions, together with open calls to the government by individual members of parliament, professional groups and researchers were not successful in the past to induce the government to take up its role in bringing together all interested groups and prepare an operational strategy for further development of information society in Slovenia. The Committee for Science and Technology of the Slovenian Parliament organized a public presentation of opinions on the information society as a challenge to Slovenia that was attended by researchers and businessmen, it published these opinions (Državni zbor [13]), but nothing more happened.

Two conclusions follow. The first conclusion is methodological. The concept of time lead or time lag for a given level of the indicator can be usefully applied to many issues in economics, politics, business and statistics. It offers a new insight to the problems, an additional statistical measure, and a presentation tool for policy analysis and debate expressed in time units, readily understood by policy makers, media and general public. Time distance is not a methodology oriented towards some specific substantive problem, it presents an additional view to many problems and applications. It can be used also for target setting, benchmarking and monitoring. By analogy, there is a wide-open possibility to apply this methodology to numerous business problems at the micro, corporate and sector levels. The shortest description of its benefits is new insight from existing data, as a new dimension is added while no earlier results are lost or replaced. In the empirical examples here very different conclusions were reached when the change in the gap in Internet users per capita was assessed by percentage difference and time distance. Both of them matter. Broader analytical framework means better understanding of the situation, which leads to better decisions.

The second conclusion is related to the need to prepare an operational strategy for further development of information society in Slovenia. The government and politicians did not realize that for the future of Slovenia building information highways to all people is more important than building the motorways across Slovenia, both for the international standing of Slovenia in the global economy and for the economic and social cohesion within Slovenia. It may very well be that Slovenia is in some aspects of information technology infrastructure still in a good position. However, this is only a necessary but not a sufficient condition for success. The crucial factor is diffusion of such opportunity and knowledge to all members of the society. It is a question of the value added in its use, a question of the modern learning economy and learning society. USA as the technology leader with a booming economy, at much higher level of development than Slovenia, showed that relying on the spontaneous development in this field is not an acceptable scenario. When will there be enough members of the government and parliament in Slovenia that will understand this message and prepare an explicit operational national strategy in full cooperation with the private and public sector and civil society? Will the present political divide prevent us for dealing with the digital divide, in relation to the leading countries and within Slovenia?

5. References

[1] Banovec (1999), 'Informacijska družba in izobraževanje', Razgledi – priloga Dosje, november

[2] Caf (1999), 'Informacijska družba v Sloveniji – v evropskem ogledalu', Razgledi – priloga Dosje, november

[3] Lundvall B-A. (2000), 'Europe and learning economy - on the need for reintegrating the strategies of firms, social partners and policy makers', Lisbon seminar on socio-economic research and European policy, May 28-30

[4] Sicherl P. (1994), Razvojne teme, Slovenska ekonomska revija, 1-2/1994

[5] Sicherl P. (1998), 'Time Distance in Economics and Statistics, Concept, Statistical Measure and Examples, in Ferligoj A. (ed.), Advances in Methodology, Data Analysis and Statistics, Metodoloski zvezki 14, FDV, Ljubljana

[6] Sicherl P. (1999a), 'The Time Dimension of Disparities in Our World', XIIth World Congress of International Economic Association, Buenos Aires, August

[7] Sicherl P. (1999b), 'A New Perspective in Comparative Analysis of Information Society Indicators', Informatica 23

[8] Sicherl P., Vahčič A. (1999), Model indikatorjev razvoja za podporo odločanju o razvojni politiki in za spremljanje izvajanja SGRS, Sicenter, Ljubljana

[9] Sočan L. (2000), 'Vzvodi za razvoj Slovenije v globalnem gospodarstvu', Razvojni izzivi pred Slovenijo, posvetovanje Združenja raziskovalcev Slovenije, maj, Ljubljana

[10] U.S. Department of Commerce (1999), Falling Through the Net: Defining the Digital Divide, A Report on the Telecommunication and Information Technology Gap in America, July

[11] Vehovar V. (1999), 'Informacijska družba v mrtvem teku', Razgledi – priloga Dosje, november

[12] White House (2000), 'The Clinton-Gore administration: A national call to action to close the digital divide', Office of the Press Secretary, April 4

[13] Državni zbor (2000), Informacijska družbe kot izziv Sloveniji, javna predstavitev mnenj, maj, Ljubljana

[14] Computer Industry Almanac Inc., http://www.c-i-a.com/199908iu.htm

[15] [15] http://europa.eu.int/ISPO/

Twelve Theses on the Information Age

Mario Radovan
Faculty of Philosophy, Dept. of Information Science,
University of Rijeka,
Omladinska 14, 51000 Rijeka, Croatia
Tel: +385 51 624 062 Fax: +385 51 345 207
mario.radovan@ri.tel.hr

The paper deals with the basic features of the contemporary life-space which has been created and shaped by information industry. On the basis of various analyses and positions put forward in published matters, and on the basis of the results of a survey we carried out, we devised twelve theses which address the most relevant features of the contemporary life-space, as well as the dominant attitudes and tendencies concerned with these features.

1 Introduction

The paper put forward the preliminary results of a research the aim of which has been to study the basic features of the life-space created by the contemporary information industry, as well as the dominant attitudes towards the opportunities and limitations that information technology offers and imposes. The first stage of the research comprised an extensive study of relevant published matter, as well as a survey by means of which we wanted to gain insight into public opinion about the issues the project deals with. The survey was carried out by means an interactive Web page (shaped as a questionnaire, realised in Java). The questionnaire started with a selection of pragmatic issues concerning life in the world pervaded by information technology; we addressed the problem of noise (in the contemporary communication space), the quality of public discourse, and new freedom as well as new limitations that information technology brings. The three traditional attitudes towards life - the aesthetic, the moral, and the religious - were considered in the context of the opportunities and limitations created by the contemporary information industry. The questionnaire also addressed the issues of *global homogenisation* of human activities and thoughts, as well as of the new *culture of the present*, created and promoted by information technology.

A hundred randomly chosen students from the departments of Information Science, Pedagogy, and Mathematics (at the University of Rijeka, Croatia) were invited (by e-mail) to visit the Web page and to take part in the survey. On the basis of the answers obtained, as well as on the basis of the insight into relevant published sources, we devised twelve theses about the contemporary life-space whose features are essentially determined by information technology. The theses we put forward do *not* express our position, nor are they

mutually coherent. They were shaped with an aim to address *the most relevant issues* concerned with the contemporary life-space, and to put forward *the dominant attitudes and tendencies* related to the basic features of current life-space. The theses we put forward are intended to be the subject of our further research, in which we intend to offer an extensive analysis and evaluation of the theses themselves, and of the prevalent position related to each of the theses. If the dominant tone in current presentation of the theses is a (too) critical one, that is primarily a sort of reaction to the excessive enthusiasm with which the information age is prone to speak about itself, completely neglecting the drawbacks of the present achievements.

2 The Twelve Theses

Thesis 1: *We live in the noise*
The concept *signal-to-noise ratio* designates the ratio between the magnitude of a useful signal and the magnitude of noise generated by the system which produces the signal. The development of information technology has been accompanied by constant *diminishing* of the signal-to-noise ratio in the global communication system, which means that our life-space is becoming more and more *noisy* (in the figurative and in the literal sense). The noise causes a stimulus overload, to which humans react to by blocking the reception, and by paying less and less attention to those signals that manage to get through their defensive barriers. To overpower the global noise, and to penetrate the defensive barriers of targeted subjects, *senders of messages* make their messages louder and louder, and more provocative. In other words, *aggressiveness and*

vulgarity increase alongside the noise, so that our life-space is becoming more and more crass (or "insensitive", "stupid").

Thesis 2: *We are becoming absent-minded*
As one of the specific problems of the information age, the increase of *Attention Deficit Disorder* (ADD) has been reported; ADD manifests itself as a lack/loss of the ability to keep concentration on any specific thing for more than a few moments. ADD has been called "the brain syndrome of the information age", and there are claims that we may be on the verge of an ADD epidemic (Shenk 1997: 36). For example, Kearney reports that since the appearance of multi-channel press-button television "less than 50 per cent of American children under the age of 15 have ever watched a single programme from start to finish" (Kearney 1994: 1). And the Internet's *hypertext style* of organising and approaching information seems to be the right thing to make the situation worse. Among the participants of our survey, 88 per cent notice the symptoms of ADD in their environment, and 16 per cent of them claim to feel this problem personally. Therefore, it seems that we have already crossed over the "verge" of the ADD epidemic, although we may well be too absent-minded to notice it.

Thesis 3: *Information industry does not promote understanding*
On the basis of the amount of *news* we are offered by the information industry, one could think that people today are well informed about the relevant local and global events. Indeed, more than half a century ago, it was announced that television would provide "truer perception of the meaning of current events, and a broader understanding of the needs and aspirations of our fellow human beings". However, television has been used mainly for "promoting voracious consumerism, political apathy, and social isolation" (Shenk 1997: 60). Furthermore, classical *news reporting* has been replaced by *news making*: attractiveness of an event became more important than its real relevance, and also more important than its objective presentation. The history of television offers a good example of the discrepancy between the opportunities created by the technology, and the ways these opportunities are being used. There are claims that the Internet (in many aspects) follows the way of television. Consequently, despite the huge advances of   information technology, the present "citizenry" is "no more interested or capable of supporting a healthy representative democracy than it was fifty years ago, and may well be less capable" (Shenk, 1997).

Thesis 4: *Privacy has been lost*
One of the outstanding problems brought about by information industry concerns the loss of *privacy*. There are specialised companies which collect all available data (primarily those connected with the use of credit cards)

about individuals. On the basis of these data they form a "consumer profile" for each individual, as well as for the specific "consumer groups". The Internet created new possibilities for data collection; for example, there are profile-making companies which copy and elaborate every message posted onto any of the *Usenet groups*, and on the basis of the content of these messages, they form consumer profiles of their authors. With such consumer profiles, marketing companies can directly (and often aggressively) "target" possible consumers (individuals and groups), with offers shaped in accordance with their consumer profiles. Such profiles are (allegedly) used only for marketing purposes, but the same or similar profiles can be used in political or personal fight of any sort against anybody. It means that by the unlimited data collecting, information industry threatens our privacy, and with this violates one of our *basic human rights*.

Thesis 5: *Humans reduced themselves to machine-friendly beings*
"Intelligent computers will design others, and they'll get to be smarter and smarter", says E. Fredkin from the MIT AI Laboratory. With a progressive increase of computer intelligence, it becomes difficult to imagine how "a machine that's millions of times smarter than the smartest person" could still be "our slave, doing what we want". Hence, the most we could hope for the future is that such machines "may condescend to talk to us", or perhaps that "they might keep us as pets" (Copeland 1993: 1). We consider Fredkin's prophecy a free play with words rather than a grounded prediction (or a plausible speculation) based on the present state of the art in Computer Science. However, the question of "slavery" does deserve attention; not because machines could become "millions of times smarter" than humans (and by this, our masters), but because they already are millions of times *faster and more productive* than humans. Hence, unless we find a way to limit the current absolute dominance of information industry, we could gradually slide away into a new slavery under the rather dull but omnipotent and omnipresent master we ourselves created.

Thesis 6: *We are entrapped in upgrading*
Our life-space is getting more and more *complex*, so that each of us is able to *operate* a lesser and lesser part of it, let alone to *understand* the ways it functions "under the surface". Ceaseless "upgrading" of software products overloads these products with "capacities" that hardly anybody needs, but which render the process of learning and using a product more demanding. (Upgrading proliferates because it is profitable for the information industry.) A large part of people do not feel able (or do not want) to cope with a highly technologized and permanently changing environment. There are social movements which advocate a *simpler* style of life (Elgin 1993). However, in the present technologized life-space, to be able to "function" at all, one is *compelled to be compatible* with the global system (*upgraded*, of course).

Hence, it is not possible to practise a "simple life" without excluding oneself from the production and social system. And the vast majority of people cannot afford it; therefore, we should consider ourselves *captives* of the ceaselessly "upgrading" technological world.

Thesis 7: *Technology is the most sublime fruit of human creative mind*
Technology emerged from human needs, as well as from authentic human *creativity*. On the other hand, technology itself (as a means) opens immense new opportunities for expressing human creativity. However, technology deserves to be considered not only a powerful means but also the *supreme end* itself, because technology embodies the highest level of *harmony and beauty* that humans ever created or encountered. If humans cannot live without a "commitment" to some *supreme entity/value* (a "god") then technology should be considered an entity worth of such a radical commitment and adoration.

Thesis 8: *By means of technology humans touch the eternity*
A human is only superficially a "reasoning animal"; but in essence, a human is desiring, suffering, death-conscious and hence, a *time-conscious* creature (Fraser 1981, 1987). Technology offers humans "more time", because it radically shortens the time needed to produce an object or to create a certain situation. However, what humans long for is not simply "more time" but a *way out* of time! People try to reach *timeless existence* by means of various "opiates" which create the *state of ecstasy* in which the feeling of time disappears. Some of the best known kinds of ecstasies are: *ecstasy of dance* (caused by the steady rhythmic music and motion), *ecstasy of the sacred* (caused by the feeling of the presence of the sacred), *ecstasy of the mushroom* (caused by means of drugs). The information industry brought about a new kind of ecstasy: the *ecstasy of the immediate*. Products of this industry (films, computer games, interactive interfaces of various kinds) keep us permanently engaged with the *timeless present*, and in that way they "raise us above time". Therefore, the products of information industry could be said to have a strong *opiatic dimension*.

Thesis 9: *Technology shall make us artists*
A human must become *an artist* who freely shapes his or her arbitrary existence in an absurd world into a *work of art*. To become such artists, humans must free themselves of all the traditional bounds, and follow only their authentic creative imagination. Values and truth are not something fixed and lasting; they are to be *created* ever and ever again in an infinite process of human active self-determination. Imagination is the only power by which humans can and must create their own world of values and truths, and in that way experience and express an endless multiplicity of their capacities. Technology is

only a *means* for the aesthetic existence; but it is an excellent means which opens to humans almost limitless opportunities to transform their lives and the world itself in an *aesthetic phenomenon*, and with that render them enjoyable, despite all the miseries which are inherent to life and to the world.

Thesis 10: *Moral sensibility must transcend the empty surface play*
We live in an empire of *mindless light and noise* in which the *present surface play* permanently occupies all our attention. Such *flattening and shrinking* of human experience threatens the very basis of civilisation which springs from the authentic human desire to understand and to communicate the *lasting* Truth, Beauty, and Goodness out of the superficial and transient phenomena. The present situation can be improved only by promoting the *moral sensibility*. We must (re)discover that beyond play of images, there are real human beings who suffer and struggle, hope and despair, live and die. The wakening of moral feelings, even if of no direct use to those towards whom they are directed, would be an essential step towards a new consciousness which would exceed the empty surface play and the arbitrary "commitments" to various "opiates".

Thesis 11: *Technology does not eliminate destructive inclinations; but it renders them more perilous*
Science and technology are not dogmatic. Every scientific theory is open to criticism and refutation, and every technological product is of a temporary value. Hence, it could have been expected that science and technology should have a radically different impact to the global human society than traditional religions and ideologies often had. However, the modesty of the spirit of science and technology has *not* made the life in modern age less violent and destructive than it was in the previous ages; on the contrary, the twentieth century has been considered the most violent period in human history. The reason for this lies in the fact that the inclination towards *creation* and the inclination towards *destruction* are equally authentic and *equally powerful* human features. Hence, science and technology cannot deliver humanity from the destructive behaviour (of various scales); on the other hand, they make the destructive human inclinations *more and more perilous*.

Thesis 12: *Information technology promotes a global human solidarity*
Intensive use of information technology in professional work, as well as in private life, makes people all over the world spend an increasing amount of time in essentially the *same environment*, thinking and acting in the similar way (determined by the specific tool they use), independently of the differences between the specific cultures to which they belong. In this way, information technology promotes a *homogenisation of the space of human thought and behaviour*. The homogenisation is

especially intensified by the appearance of the Internet, which introduced *direct links* and *direct communication* between people all over the world. With global homogenisation, specific traditions and ideologies (each usually hostile toward the others) gradually lose their dominance over the individuals that belong to these social traditions, so that people all around the world are getting ever more *similar* and *psychologically closer* to each other. Consequently, we can expect that the *range* of human solidarity - traditionally limited to the members of the *own social group* - will constantly grow and will embrace the *whole humanity*. Indeed, information technology raises and promotes a *new humanity*, freed from the hostilities and destructiveness that dominated human history.

3   Conclusion

Technological age opened an immense new space of possibilities for human self-understanding and self-realisation. In the age of technology humans have ceased to be merely powerless dreamers of the transcendental worlds, and became literally creators of their own (and only) world. Hence, it has become more important than ever to evaluate the good and bad sides of the world we create, as well as of the worlds we destroy. Modern information technology seems to be the most complex and the most impressive creation in human history. It delivers humans from many constraints, but at the same time it tends to enslave them into its own desensitised world of empty light and noise. Our future research will include an extensive public discussion about the theses we put forward in the present paper. However, it is difficult to promote a critical *and* coherent discussion about the life-space dominated and shaped by information industry, because to reach public attention in the present condition, one is compelled to use the same means and methods one is trying to criticise, which puts such a critic into a paradoxical position.

4   References

Let us mention some of the sources that had a relevant influence to the selection (and shapes) of the theses we put forward in the paper (although not all the sources were explicitly mentioned in the present text).

Borgmann, A. (1984) *Technology and the Character of Contemporary Life*, Univ. of Chicago Press.

Castells, M. (1996) *The Rise of the Network Society*, Blackwell.

Castells, M. (1997) *The Power of Identity*, Blackwell.

Castells, M. (1998) *End of Millennium*, Blackwell.

Copeland, J. (1993) *Artificial Intelligence: A Philosophical Introduction*, Blackwell.

Elgin, D. (1993) *Voluntary Simplicity*, Quill.

Fraser, J. T. (1987) *Time, the Familiar Stranger*, University of Massachusetts Press.

Fraser, J. T. (1981) *The Voice of Time*, University of Massachusetts Press.

Golomb, J. (1995) *In Search of Authenticity; From Kierkegaard to Camus*, Routledge.

Graham, G. (1999) *The Internet; A philosophical inquiry*, Routledge.

Johnson, S. (1997) *Interface Culture*, HarperEdge.

Kearney, K. (1994) *The Wake of Imagination*, Routledge.

Kearney, R. (1991) *Poetics of Imagining*, Routledge.

Shenk, D. (1997) *Data Smog: Surviving the Information Glut*, HarperCollins.

Simpson, C. L. (1995) *Technology, Time, and The Conversations of Modernity*, Routledge.

Taylor, C. M. and Saarinen, E. (1994) *Imagologies*, Routledge.

Winston, B. (1998) Media Technology and Society, Routledge.

Measuring New Media

Vasja Vehovar, Luka Kogovšek
University of Ljubljana, Faculty of Social Sciences, Kardeljeva pl.5, Ljubljana, Slovenia
Phone: +386 1 5805 297, Fax: +386 1 5805 101
info@ris.org

After some years of uncertainty, confusion and high expectations the Web is slowly becoming a real alternative to the conventional media. Within this framework, the measurement of the Web activities is an extremely important, although difficult task. However, the overall technological trends offer a radical solution (e.g. Pc-meter) to the problem of the global standardized meter. Even more, it seems that the globalisation of the content also follows the path of the globalized measurements.

1 Introduction

The Web emerged rapidly as a new media, in only few years. However, despite its extreme public attention, the actual Web media consumption is still relatively small. Average daily-spent amount of time on the Web (calculated on the total population) is still measured in minutes, while classic media (i.e. TV, radio) can be measured in hours.

As for Fall 2000, in Slovenia the average daily Internet media consumption (including the non-users) is only 3 minutes (RIS, 2000). On the other hand, the TV is watched around two hours, and radio is listened over three hours daily. Of course, the numbers for Slovenian Web consumption are relatively low, particularly because in year 2000 only one third of the time Slovenians spent on the Web was located at Slovenian sites. However even in the US the average daily consumption on the Web stands around 16 minutes (9 minutes for home usage and 7 for business usage). Here, again, we calculate this approximation as the average across the active population (Nielsen Net-Ratings, 2000).

Accordingly, the advertising expenditures already allocate few percents of their budgets to the Web. Of course, this is the case only in the most developed countries, whereas in small markets (i.e. Slovenia) this share is about ten times smaller. Nevertheless, the advertising budgets closely follow the increased Web usage, so huge growth in Web advertising is expected in next years. This is the case also for the countries such as Czech and Slovenia, where the Web advertising is rapidly growing, not to mention the US where the projections estimate that in few years the percentage of advertising budgets allocated to the Web will surpass 10% of all advertising expenditures (IAB, 2000).

Simultaneously, the need for specific Web measurements also constantly increases. It is an extremely convenient fact that the Web itself offers very effective way of measuring users' activities. As Internet usage is basically one-person-to-one-device activity it is in principle possible to automatically track very detailed activities. Today, we can thus observe trends of globalisation and centralization in Web activities' measurements. With so-called PC-meter or Net-Meter we are basically ending up with only few players (i.e. MediaMatrix, Nielsen NetRatings, NetValue) that can maintain global panels of the Internet users who accepted the installation of the measurement software into their PC. In year 2000, the largest players already claim that with samples of standardized national PC-panels they already cover more than 80% of Web users in the world.

Of course, the global PC-meter-based measurement has also many weaknesses. The most problematic is the low resolution for smaller sites, which can hardly get enough traffic on top level (introduction pages), while it is practically impossible to estimate the number of users on subsidiary pages within any level of rational costs. Another disadvantage stems from difficulties with business usage, due to number of restrictions companies impose to such measurement, what results in low coverage. In addition, the response rates tend to be small in these types of survey panels.

Despite high costs for software solutions and expensive methodological developments, further expansion of these measurements is expected. In the near future there will be thus a possibility to observe, moment by moment, the representative picture of Web activities of the total world population active on the Web. Of course, we can also expect the extension and unification of these

measurements to other devices than PC (e.g. TV, mobile).

Despite these global trends, the Web measurement in small audiences, either geographical limited local communities or language specific populations, is still problematic. It seems that, here, the measurements will remain local also for next few years, as the revenues from such measurement cannot cover the costs of maintaining the appropriate panels. The main problem in this aspect is the fact that, as opposed to TV, radio or printed media, Web is much more diversified with thousands of Web locations. Accordingly, we need much larger samples to capture these activities and consequently the corresponding costs are relatively large.

2 Measuring the Web in Slovenia

As for now, the global companies are not yet interested in performing measurements for small audiences, because there is a clear lack of the critical mass to enable profitability of these business activities. There are, of course, also other reasons for a relatively low priority Slovenia has in the future plans of MediaMetrix and Nielsen NetRatings, because Ireland and Singapore are already included, despite their relatively smallness.

However, besides global on-line PC-meters applied to the sample panels of Web users, the Web measurement can be performed also in other directions. As for now, in Slovenia we collect the data from the three sources presented in next subsections.

Recall survey measurements: user-centric approach

With user-centric approach we survey the Internet users as opposed to other approaches (log analysis) where units of measurement are basically the requests on servers. In the absence of the panel with the PC-installed programs (which also belongs to user-centric type of measurements), we can roughly estimate the monthly visits with recall techniques in telephone surveys. Such a data arise from a monthly RIS (Research on Internet in Slovenia) telephone survey where certain pages are asked to be evaluated by respondents. Of course, there do exist methodological problems with this methodology, however, the information is generally stable and consistent.

We should stress that the monthly reach of AltaVista and Yahoo still dominates among Slovenian users, with around 150,000 persons visiting them on a monthly basis. However the gap to Slovenian sites is rapidly narrowing each year.

Table 3 presents the list of top Web sites from the beginning of the year 2000. We should recall that there were around 280,000 users at that time.

Site	Reach
Matkurja.com	90,000
Tis.telekom.si	70,000
Arnes.si	65,000
Mobitel.si	50,000
Siol.net	50,000

Table 1: Monthly audience for top five domestic Web sites in Slovenia 2000/1

By the end of the year 2000 the above figures generally increased by 30-40%.

Log analysis - server analysis: Web-centric approach

Different results were obtained from the log analysis. The log file is a tracking that is recorded with every visit at certain Web site. The log files have been analysed within RIS project for 40 most visited sites in Slovenia at the beginning of the 2000. There does exist a complex methodological environment for understanding these results, as the standard results of Webtrends analysis gave different figures for a monthly number of user sessions.

We should keep in mind that each user could, of course, produce several sessions within a day or a month. A session is defined as a visit to a Web site by a single visitor, which is performed without a certain break (usually 30 minutes). Of course, this type of measurement is extremely vulnerable towards faking, specific technological settings at the server, and also to the patterns of visitation that is exhibited by the visitors. Certain sites have a lot of daily visitors, while other have majority of monthly visitors.

Site	Sessions
Siol	398,285
Slowwwenia	278,545
Mobitel	197,170
Arnes	167,623

Table 2: Top five sites according to monthly numbers of users session in Slovenia 2000/1

It is thus not surprising that another statistics, number of hosts, gave rather different results. The host is a device, usually the PC, with a specific IP number. The IP number thus tells from which computer the access was performed. The picture of top sites in this aspect is rather different:

Site	Distinct IP
CVI-sigov	28,794
Arnes	27,344
Tis.telekom.si	21,532
RCUM	21,098
Siol	13,243

Table 3: Top sites according to monthly number of distinct IPs in Slovenia 2000/1

Obviously, there exist certain ambiguities about these numbers too, however, it is very useful to expose them in a comparative manner.

The differences in measuring IP numbers as opposed to sessions well illustrates the problems of log analysis and the ambiguity when these data are used as a measure. On the other hand the gap compared to the telephone results reveal additional discrepancies, although the correlation was found to be high (R=0.8) between the IP measures and telephone results.

The difference between users centric and server centric statistics are common also with PC-meter technology approach. A well-known reconciliation study (MacEvoy, Kalyanam, 2000) reveals that the two types of data coincide only in the ranking of the first three sites, while with others considerable differences appear. Numerous methodological problems contribute to this gap and, of course, we will not address them here, but only repeat the general caution towards the usage of any type of this data.

We can thus conclude that, as for now, serious methodological problems exist in the field of measuring Web activities. However, the intensive work may soon remove the basic obstacles in this methodology, since it is only few years old and is thus lacking both, the experience and the research.

Ad-server analysis: ad-server centric approach

The measurement that is located at the server where the advertisements are lunched is technologically much more precise than the previous demonstrated standardized log analysis. Here, the organization that buy certain advertising space at the publisher Web sites pushes the advertisement to the users, so they have a centralized opportunity to measure the visits in an interactive and standardised way. However, this statistics still suffers from all deficiencies of log analysis. In addition, this approach is limited only to the pages that are included in a certain syndicated Web advertising pool.

3   Conclusion

At the beginning of the millennium we can observe the Web activities' measurements being in a sharp transition towards a global measurement. Of course, it will take few more years until the standards will be widely accepted. In addition to this global statistics, there will perhaps remain a certain space also for areas, where the global measurements cannot be performed. This is somehow similar to the TV audience measurement or radio measurement for small areas with few hundreds of thousands consumers, where the national-oriented surveys cannot provide representative data. Therefore, alternative and local surveys will be still preformed to provide the necessary data.

Today, the global (PC-meter based) measurements of Web activities already cover a large majority of World Web users and we can expect that the coverage will further increase. It is thus already possible almost instantly to measure certain significant global reaction to specific news, events or new sites that appeared on the Web. Such a high coverage of centralized global measurement is unique to the Web and cannot be compared to any other media.

However, we should add that it is not only the measurement that is becoming much more global with the Web than with any other media. It is also the content that is much more global with the Web than with other media too. Even a brief look at the top ranked sites (Nielsen Netratings, 2000) suffices to confirm that almost in all countries we find Yahoo, AltaVista, Excite, Microsoft, and MSN among the top five or top ten sites. Similarly, the Amazon.com is appearing among top shopping sites all over the world. It is not surprising that it is also the top-shopping site for Slovenian users. Such a globalisation is not at all the case with any other media (i.e. TV, radio, newspapers, magazines).

The Web content is thus becoming also extremely centralized. Of course, on the other hand, a massive stream of specific sites with narrow audience appears, too. However, the software, investment and promotion needed to establish and maintain a highly-visited Web site demand extreme resources, so we can expect that only the largest players will be able to keep the race of offering best information, technology, design and promotion. In a certain aspect we could even say that the global measurement itself also promote the globalisation of the content, as only the sites that appear in global top rankings can receive larger stream of advertisements what enables their further development and promotion.

In this aspect things are changing very rapidly. Only two years ago, when preparing a chapter for the monograph Cyber-imperialism (Ebo, 2000) we projected that the global English-language based sites will keep the lead also in small countries. At that time the visitation of

Slovenian Web sites was way behind the English-language well-known Web sites (e.g. Yahoo, Altavista). It thus seemed that Slovenian companies would have to locate and pay the advertisement at these foreign Web sites to target the Slovenian consumers. Eventually, that did not happen. There were no Slovenian advertisements on Yahoo or AltaVista yet. This was partially due to the fact that these sites were loosing their advantage as more and more Slovenian users without English knowledge appeared, but also due to the increased Slovenian content. The cheaper home-based advertising and the lack of competent comparison also strongly contributed to this fact.

Instead, however, the beginning of another much more radical trend can be observed. It seems that in near future the global companies will – directly or indirectly – perform not only the advertising but all Web-related operations: measuring the Web site visits, distributing the global and local (e.g. Slovenian) Web advertisement, and at the final stage they will also purchase all the worthy top Slovenian Web sites.

Thus, the trends in technology and the very nature of the Web strongly stimulate the globalisation of the industries linked to the Web, being it the advertising, the measurement or the content. It seems, however, that the measurement is the first activity that will become completely globalized. Of course, with globalisation, we mean both, the covering of the whole world Web population as well as a uniform methodology that provides an instant global insight. However, the globalisation also means that only few private but global companies will own this whole of this specific industry.

4 References

[1] Ebo, B (2000): Cyber-imperialism : Global Relations in the New Electronic Frontier, Northwest Publishing Group.

[2] Internet Advertising Bureau (2000). URL: www.iab.net.

[3] Kogovšek L: Analiza logov – jim lahko verjamemo, Marketing Magazin, 2000/junij.

[4] Nielsen Net Ratings, 2000. URL: www.nielsen-netratings.com

[5] MacEvoy, B, Kalyanam, K. (2000): Data reconciliation, reducing discrepancies in audience estimates from web survey and online panels. ESOMAR Online Media Measurement Conference and Exhibition, Paris, 2000.

[6] Research on Internet in Slovenia (2000). URL: www.ris.org.

[7] Vehovar V. (2000): Je slovenski trg za Internet premajhen? 3. marketinška konferenca. URL: http://www.ris.org/si/ris2000/novice/20000619.htm

Impact of Digital Radio-Television Development on Spatial Development of Slovenia

Miha Krišelj
RTV Slovenija, Kolodvorska 2, Ljubljana, Slovenia
Phone: +386 1 475 2681, Fax: +386 1 475 2680
miha.kriselj@rtvslo.si

Through the digitalisation of telecommunication (TK) systems the radio broadcasting network is also becoming a part of a unified - global telecommunication infrastructure. The new systems are related to new services and to specific technical characteristics of digital networks. The latter shall stem in the users needs of applications and services. The digital broadcasting system will make it possible to transmit not only the traditional radio-television but also multimedia contents, while the introduction of inter-activity will provide the access to the Internet. Consequently it follows from the above that a co-ordinated, economic and efficient construction and utilisation of public telecommunication systems shall be ensured (RTV, railways, electric power management, etc.), naturally considering potential spatial impacts.

1 Introduction

The digitalisation of broadcasting systems represents an important turning point not only in the development of what broadcasting has to offer, but also in the sense of a technological association with alternative telecommunication systems. In the area of production the digital radio-television will encounter major modifications, as in addition to classical audio-visual subjects others - so-called multimedia oriented services - will appear, which will also include PAD (Programme Associated Data). On the other hand the broadcasting radio-television systems are also becoming interesting for the distribution of different telecommunication services.

The last decade gave rise to several proposals of the telecommunication infrastructure development strategy. Thus in the HONET project, produced in the environment of the then ISKRA, it is possible to trace the networks which are supposed to form the future information telecommunication infrastructure. These are the following networks:

- Telekom
- Railways
- Electric Power Management of Slovenia
- RTV microwave links
- Cable television
- Special systems (fire service, emergency, taxi...)
- Ministry of Internal Affairs
- MORS (Ministry of Defence)
- Satellites
- GPS

In September 1992 the ruling government requested the co-ordination of HONET project with the "Public Administration" project, yet to no avail. As a result the Ministry of Traffic and Communications formed a committee to further the development of telecommunications project, named MITIS. The project dealt with 10 networks (then PTT, Railway, Electric Power Management, RTV, academic sphere, banks, tourism, state administration, and industry). Unfortunately MITIS was neither comprehensively dealt with nor adopted. It was commented on by the Chamber of Commerce of Slovenia and by the General Association of Electric Industry, and they proposed the establishment of a National Telecommunication Council, which was regrettably not appointed. In its documents MITIS also identified the third generation of telecommunication services, dealing with digital systems which will serve as a basis for the new telecommunications infrastructure.

Today it is only sensible to talk about the development of digital systems. In the above-mentioned studies on the telecommunications networks development there are quite a few so-called backbone networks. Here I have in mind the network of the Telecom, Slovene Railways Management, Slovene Electric Power Management, and microwave links of RTV Slovenia. Other networks provide communication to the final user (Telekom, Cable TV, RTV secondary network). Backbone networks have a common feature - they were built for the needs of the above-mentioned large companies. In the case of the Railways Management network we are dealing with

private telephone networks, a wide area network (WAN) and transfer of technical information (signals related to the status in railway traffic), with Electric Power Management - ELES we are mostly dealing with the development of the technical information system and a few other applications, while with RTVS links we are dealing with the primary distribution of RTV programmes and feeding the secondary network, which transmits the signal to the final user. With the development of technology all the above-mentioned networks were digitised. While in the case of Railway Management and Electric Power Management the utilisation of optical fibres, with RTVS the introduction of digital microwave links prevailed. All three may be attributed the following features:

1. Through digitalisation the networks became usable practically for the entire range of telecommunication services.
2. Through the introduction of optical fibres the capacity of the said networks enormously increased and - in spite of the increased needs to transfer information as a result of introduction of new applications - it offers a sufficient amount of free transfer capacities. A similar statement also applies to the introduction of digital link systems of RTVS, only that there are less free capacities (while it requires a rather high quality of digital audio-visual signals at the primary distribution level)

The networks may also be designed from bottom to top, i.e. in view of future telecommunication services. In such case it is sensible to contemplate the above-mentioned networks, with the access to the final user: Telekom (fixed and mobile networks), cable TV network, RTV (earth and satellite) network. In the research project of the Urban Planning Institute of the Republic of Slovenia, entitled Telecommunications System and Its Impact on Spatial Development (A. Gulic, S. Praper) one can detect the following statement: *"The designs of network operators shall stem from the users needs of applications and services, as the latter are the end product, the commodity which the users are willing to pay for. As with any product the quality and price are important here. Once the progress of technology conditioned the development of telecommunications systems and services, while nowadays technology is at hand. It will only be used if the user applications are clearly defined and the investments are economically based."*

The services and applications shall therefore serve as a basis of contemplation of new telecommunications systems. In the following we can divide the treated networks into:
- Physical (wire line and optical)
- Wireless

The Telekom network and cable TV network belong to the first category. Either the Telekom network or the cable TV network are no longer limited to the traditional telephone and television services. The digitalisation of these networks made also possible the transmitting of

wideband services through the Telekom system, among which television is the most characteristic representative, as well as the Internet services. Similar conditions prevail with cable distribution systems (200.000 connections in the Republic of Slovenia), which are becoming universal and two-way (interactive) through digitalisation.

In the wireless networks system the most important are the system of digital broadcasting and the system of third generation mobile telephony (UMTS - Universal Mobile Telecommunications Systems). The transmission systems of digital broadcasting can be divided into satellite DVB[1]-S, S-DAB[2] and terrestrial DVB-T and T-DAB. From the point of view of coverage the satellite systems are representatives of global telecommunications systems, while the earth systems may be divided into national and local - regional. The latter are closely related to the space. With data transmission through wireless systems we can not talk about routes - moreover - the density of the number of telecommunications connections <u>does not depend</u> on the telecommunications infrastructure, while it heavily depends on the quality of the wireless signal in space. The designing of wireless communications systems is therefore to a large extent **related to spatial planning** of the areas of coverage, considering potential spatial impacts.

Let us confine ourselves to the representatives of earth digital networks T-DAB, DVB-T and UMTS only. Through the development of digitalisation of broadcasting networks a tendency has arisen towards the transmission of additional services. One can trace the first essays of the type in the respective hybrid analogue-digital RDS (Radio Data System) system, in which the system capacity was too small (about 1 Kbit/s) for any major application. In the subsequent DAB system, which is digital and incompatible with the existing FM/RDS systems, it is possible to transmit much more data (2Mbit/s). With the arrival of the DAB system it became obvious that the traditional role of radio-television companies had changed. As we are dealing with systems which enable transmission of any telecommunication service, it has become clear that new rules of the game will have to be applied on this market. The need for a new regulation also originates in the fact that the future space in digital multiplexes will be shared by several mutually different radio-television programmes.

2 Digital Radio

Here we are talking about the issue of a balanced offer of various programme types (formats) of radio programmes within a single DAB multiplex, about the area of coverage with a digital radio signal, which will henceforth be equal for all multiplex members, and not last about the ratio of the bits intended for radio programmes and other bits for the so-called NON PAD

---

[1] DVB=Digital Video Broadcasting
[2] DAB=Digital Audio Broadcasting

(services not related to radio programmes). From the above one can draw a clear statement that a new deal has been established in the field of digital radio broadcasting, i.e. the distribution of roles to the so-called content providers, multiplex providers, and network operators.

We can say that the DAB systems development in Europe is coming to a crisis, as the initial interest of the providers of additional telecommunications services has substantially dropped along with the occurrence of digital television. In this way the tension and rivalry among radio companies and prospective new media subjects has decreased, and on the other hand has caused a slow-down implementation of digital radio networks. This is problematic, as the condition of establishment of a new digital network, considering the utilisation of new technology and related new services and users needs of those, has not been fulfilled. It has turned out that the users do not see additional value in a little better quality of radio sound and in improved conditions of reception in moving vehicles, while there is practically no new programme contents. With the problem of an insufficient available radio frequency space, which does not suffice for the distribution of all existing radio programmes, along with additional new programmes, and fresh telecommunications services, the story of the introduction of DAB systems gets in a blind alley. It is a well-known case that in Sweden they use DAB technology for the transmission of video pages in trains of Swedish railways.

For the time being in Slovenia the situation in the field of DAB network is at the designing stage. The analysis of the existing analogue network shows a rather unbalanced radio offer. This applies as much to the variety of programmes, as to the density of regional radio networks, which are difficult to spatially identify from the standpoint of statistic regions. An impression remains that the existing FM radio broadcasting networks create their own regional image, which is more or less the result of individual local conditions and ambitions of local radio stations.
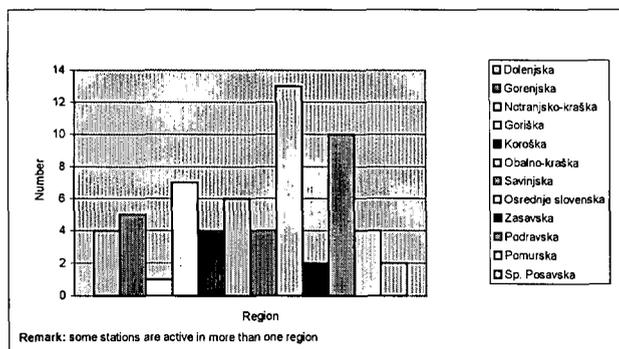
Chart: Spatial distribution of local FM radio stations.

In designing T-DAB networks we encounter some issues which relate to the problems of programmes, contents and coverage areas. A programme strategy should provide the answer to the question of what the offer of radio programmes will be like; whether in the new system there will be made up of all the currently operating radio stations, with the same programme which is on the FM network, or some capacity will also be dedicated to the new - multimedia radio programmes, which will provide fresh contents, enriched with added PAD (Programme Associated Data).

On the other hand it is suitable to ask what will be the relation between radio and other telecommunications services in DAB system. As an example let me quote English recommendations, which dedicate only 64 Kbit/s to the other telecommunications services.

Not of least importance is the spatial view of new networks. Considering that the same coverage area will henceforward be shared by at least 6 radio stations, which may currently have different areas of coverage by FM radio signal, it will be necessary to establish spatial strategy of T-DAB networks. This may either relate to the system of statistical regions or may create its own spatial pattern.

Slovenia has already made a few steps in the direction of T-DAB network development. A set of rules has been under preparation which is to solve the newly incurred conditions in the digital radio broadcasting market, and frequency blocks for the transmission of T-DAB signals have been co-ordinated. The spatial distribution and capacity of frequency allotments for T-DAB itself is not optimal, as it does not meet the current density of existing radio stations or spatial characteristics of the current distribution, nor any other potential strategy of spatial development of Slovenia.

The problem is even larger as a new mobile telephony system of the third generation - UMTS is on the horizon, which will enable rather high speeds of data transmission (384 Kbit/s in a moving vehicle and 2 Mbit/s when stationary). This system competes for the same frequency ranges as T-DAB (1.5 GHz), while in addition to telephone services it is also capable of providing other services.

3 Digital Television

As already mentioned DVB-T system offers even larger capacities. As applicable to all digital systems, DVB-T is also much more than television. In a single TV channel (8 MHz wide) it will be possible to transmit the following contents:

- 4 to 5 standard (4:3 format) television programmes or
- 2 to 3 improved (16:9 format) television programmes or
- 1 high definition television programme or
- combination of television, telecommunication, and data services

Naturally all the above contents are multiplexed in a single DVB-T channel. The real future of digital television is most probably in the so-called Multimedia Home Platform, which represents a giant leap in the development of telecommunication markets: from the present closed into the open - horizontal markets. MHP is the new, inter-active standard, which connects different services to the convergence of radio-broadcast, telecommunication and computer networks. MHP provides the universal Application Programming Interface, open and accessible in set top boxes, integrated digital TV sets, as well as in multimedia personal computers (MPC). MHP will enable the transmission of standard TV programmes, enhanced definition television programmes (EDTV), inter-active television and Internet contents, including the possibility of access to the world wide web, e-mail, e-commerce and the other Internet applications through the uniform API. The uniform standard will equalize the minor and major providers of services, lower equipment production costs (larger quantities and less varied articles) and consequently the prices, and therefore increase the selection of services and applications. In the beginning of 2000 the MHP standard was adopted by DVB, while the start of regular production of integrated digital TV sets has been scheduled for 2001.

The design of digital TV network DVB-T is similar to that described with digital radio T-DAB. Important here is the spatial aspect and the decision on the contents of the digital multiplex which - in relation to the digital radio network - are much larger. There are more promising possibilities of introducing new telecommunication services and applications, while it is also necessary to design the transition of the existing analogue TV programmes into the digital network. It is also possible to conceive the DVB-T system for the mobile reception of signals, which reflects in certain decrease of capacity of DVB-T multiplex.

4 Conclusions

It seems that - in spite of the increasing awareness of the convergence of the telecommunications market and the model of global information infrastructure (GII) - digital broadcasting still does not enjoy enough attention. In the planning documents of the Ministry of Environment and Space it is possible to trace the development charts of the Telekom network, ELES (Electric Power Management) network, and Railways network, while the broadcasting network is missing, although - from the point of view of the prospective telecommunication services and applications - it represents an equal part of the so-called Slovene Information Infrastructure. In the long-term plan of the Republic of Slovenia for the period from 1986-2000, amended in 1999, there are only the concepts of the analogue TV network for the minority programmes, TV networks (programmes 1 and 2 of TVSLO), and FM of radio network (programmes 1,2 and 3 of RA SLO) from 1984. The conditions in the field of broadcasting

have drastically changed in the last decade. The offer of programmes has broadened in the form of commercial radio and television stations, while the broadcasting networks are becoming a section of the Slovene telecommunication infrastructure through digitalisation. The telecommunications system undoubtedly influences the spatial development, while we can claim even more firmly that the telecommunications system - particularly the wireless part - is conditioned by space. In the studies on the telecommunications system and their impact on spatial development (6) it is possible to encounter emerging and important mode of access to the Internet and the world wide web - i.e. through TV sets. It is obviously more and more clear that the digital television and radio network will become increasingly important pillars of the future information infrastructure. URST (Administration of the Republic of Slovenia for Telecommunications) has outlined a plan of introduction of DVB-T in Slovenia by designing: a national SFN (Single Frequency Network) of DVB-T to cover at least 95% of population, three national networks, built on a combination of MFN (Multi Frequency Network) and SFN (Single Frequency Network), two MFN or SFN networks for regional programmes, and a few local transmitters. This principal orientation shall be followed by the principal strategy of the introduction of digital radio broadcasting in Slovenia. In its evaluations the Broadcasting Council, authorised for the treatment of principles of distribution of broadcasting frequency spectrum states that the law does not stipulate which body is competent for the preparation and the adoption of the principles of frequency channel assignment. As a result the Council has appointed a special expert project team, and - based on the proposal of this team, in co-operation with RTV organisations - has adopted the "Principles of the allotment of radio broadcasting frequency spectrum and assignment of radio broadcasting channels". In addition to the "local non-commercial programmes", defined by law, the Council also introduced regional (non)commercial programmes. But, considering the fact that the legislation does not identify the areas or regions, the Council took the number of population, covered by a signal as a criterion for regional programmes (10% of population of the Republic of Slovenia). This criterion can also be encountered in the draft Media Act, in the section dealing with local RTV programmes. In addition to the local programmes regional RTV programmes are also mentioned there, intended for the population of an area (region, town), with more than 10% and no more than 50% of the population of the Republic of Slovenia.

The lack of a transparent development strategy of the digital radio and television, not only in the sense of space but also in the sense of programme and services is probably a result of the fact that the development of radio and television has come to a turning point, as on one hand the role of RTV organisations is changing as a consequence of the new distribution on the digital radio broadcasting market (providers of contents, multiplexes and networks), and on the other hand due to the

indisputable fact that the national RTV has got competitive commercial RTV organisations, with which it will have to share the space in digital multiplexes in the future. Hereby one shall not neglect the fact that along with the above an interest also arises in the distribution of additional - non broadcasting services and applications.

5   References

[1]   Dolgoročni plan republike Slovenije za obdobje 1986-2000, dopolnjen 1999. In Slovenian.

[2]   Srednjeročni družbeni plan republike Slovenije za obdobje 1986-1990 dopolnjen 1999. In Slovenian.

[3]   Krišelj M. (2000) Vloga in pomen digitalne radiotelevizije v prostoru mobilnih multimedijskih telekomunikacij. *Proc.10th   Workshop   on Telecommunications VITEL*, Brdo 15.,16. May 2000. In Slovenian.

[4]   Krišelj M. (1999) Digitalno radiodifuzno omrežje in   prostor. *Proc.   9th   Workshop   on Telecommunications   VITEL,*   Brdo   22.,23. November 1999. In Slovenian.

[5]   Krišelj M. (1999) Perspectives of Spatial Development of Digital Audio Broadcasting in Slovenia. *A Graduation Thesis*. International Academy of Broadcasting, Montreux 1999.

[6]   Gulič   A.,   Praper   S.   (2000)   The Telecommunications System and Its Impact on Spatial Development - *Final Report*. Urban Planning Institute of the Republic of Slovenia. Ljubljana, February 2000.

[7]   Ravbar M. et al. (1998) Prostorski vplivi približevanja Slovenije Evropski zvezi - *zaključno poročilo*. Inštitut za geografijo. Ljubljana, december 1998. In Slovenian.

[8]   European   Spatial   Development   Perspective (ESDP). Achieving the Balanced and Sustainable Development of the Territory of the EU: *The contribution of the Spatial Development Policy*. Potsdam, May 1999.

[9]   Gulič A. et al. (2000) Zasnova regionalnega prostorskega razvoja Slovenije, *Raziskovalni projekt - zaključno poročilo*. Urbanistični inštitut Republike Slovenije. Ljubljana, avgust 2000.

[10]   Gulič A. et al. (1997) Vplivi sodobne informacijsko-komunikacijske infrastrukture na prostorkski razvoj Slovenije - *zaključno poročilo*. Urbanistični inštitut Republike Slovenije, Inštitut za geografijo, Ljubljana, november 1997.

Internet Based Art Installations

Franc Solina
University of Ljubljana, Faculty of Computer and Information Science
Computer Vision Laboratory
Tržaška cesta 25, SI-1000 Ljubljana, Slovenia
Phone: +386 1 4768 389, Fax: +386 1 4264 647
E-mail: franc.solina@fri.uni-lj.si

An overview of some internet-related computer applications which were used in several art presentations and art installations on the Internet is given. The technical solutions range from the design of typical hypertext contents combining text and images, creation of virtual environments, sending of life video images over the Internet and controlling remote robotic devices over the Internet. These technical solutions were successfully used for presentations of classical fine arts on the Internet as well as for creation of contemporary art installations.

1 Introduction

Already in the spring of 1995 we presented on the Internet the Slovenian Virtual Gallery which was a typical first generation web multimedia presentation of Slovenian fine art consisting of an interconnected set of texts, images, and video clips. Beside following the interconnecting links, an alternative way of exploring the pictorial information in this multimedia data set was by "walking" through a virtual gallery space.

This multimedia concept which we present in Section 2, in combination with our module for real-time video observation over the Internet (Section 3), was later used by the video artist Srečo Dragan for several of his art-Internet installations described in Section 4. By adding the possibility to get real-time video from any physical point, which can be connected to the Internet, one can effectively blend actual and virtual spaces.

2 Slovenian Virtual Gallery

The Slovenian Virtual Gallery (SVG) was developed in the first half of 1995 by senior students of computer and information science Andrej Lapajne, Bor Prihavec, Žiga Kranjec and Aleksandar Ruben as a project in the framework of a course taught by the author. The goal of the project was to present Slovenian fine-art on the internet [8, 13]. In cooperation with distinguished slovenian art-historians[1] we prepared an overview of Slovenian art from the gothic period up to the present day[2]. SVG consists of three main

parts (Fig. 1(a)):

1. Overviews of the main art historical periods which contain biographies of authors, each with an iconized index of their works (Fig. 1(b)). Icons can be blown up to the screen size (Fig. 1(c)).

2. Permanent collections and current exhibitions in selected Slovenian art galleries.

3. A 3D virtual gallery where the viewer can move through a virtual three-dimensional architecture and view the paintings hung on the walls (Fig. 1(d)). By clicking on the paintings the user switches to the works and their authors in the first part of the SVG.

SVG supports also search of authors or works of art by using different keys (names, years, art techniques).

SVG, judged by its implementation, was a typical first generation web site. Due to the lack of appropriate tools at the time of development we wrote our own data management tools and tools for automatic generation of HTML documents, all implemented in PERL. Data was stored in files which were directly manipulated. Since there were just a few typical types of documents in the SVG, we used patterns to generate HTML documents. The administrator of the system whose role was to add new content to the SVG didn't need to know the HTML syntax (Fig. 2). Additional features of the SVG system were a distributed database and remote management of the system. In the span of just a few years, however, the web related technology has experienced a tremendous growth. On the market is now available a range of relational and object-oriented data management systems which make such types of web applications much easier and faster to develop. We

---

[1]Dr. Samo Štefanac (gothic period), dr. Tomislav Vignjević (renessance), Matej Klemenčič (baroque), dr. Barbara Jaki (19th century), mag. Igor Zabel (20th century)

[2]Špela Zorčič helped with the graphical design of the SVG web site.
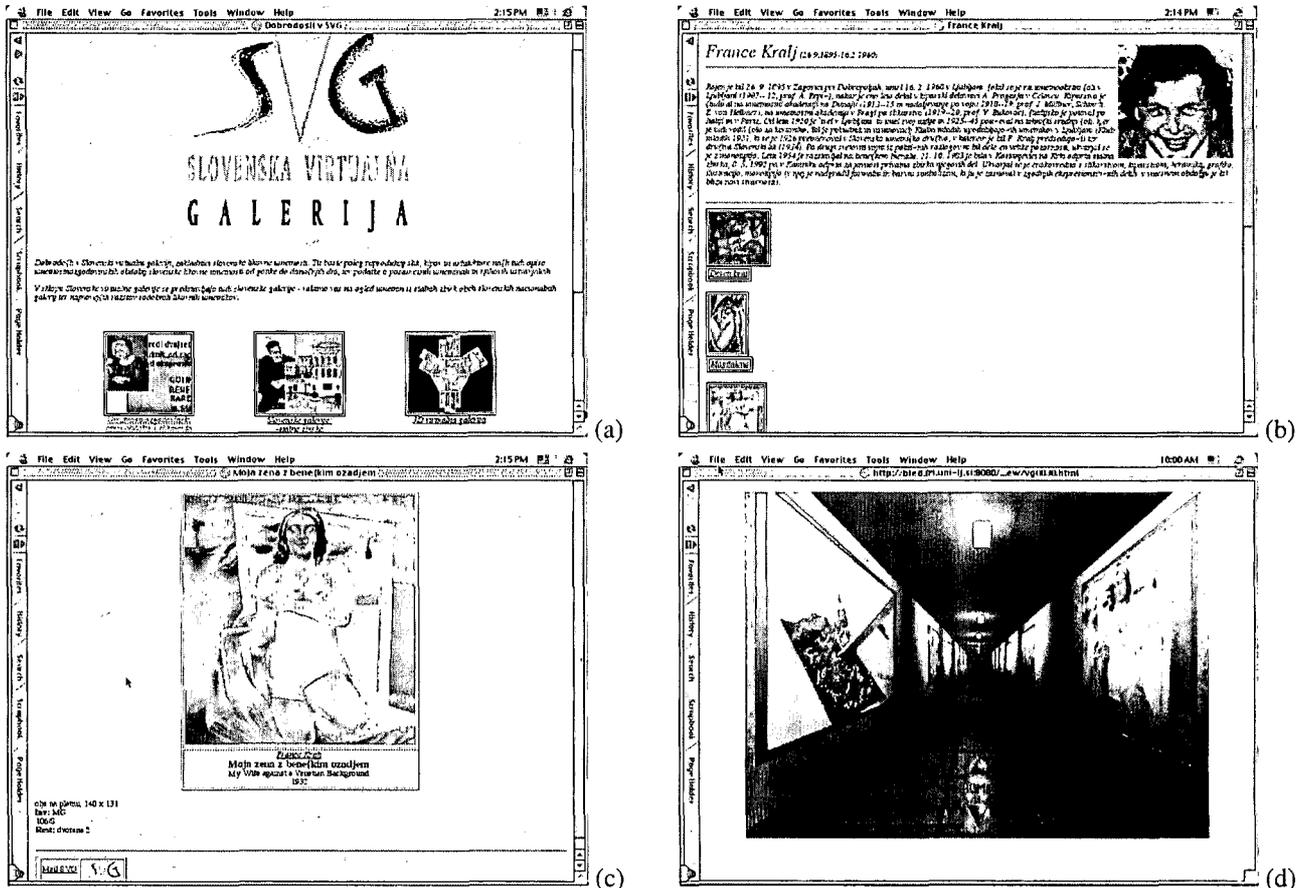
Figure 1: (a) Home page of the Slovenian Virtual Gallery showing its three main constituent parts, (b) Presentation of painter France Kralj, (c) "My wife against a Venetian background" by France Kralj (oil on canvas, 140×131 cm, 1932), (d) Inside the virtual gallery.
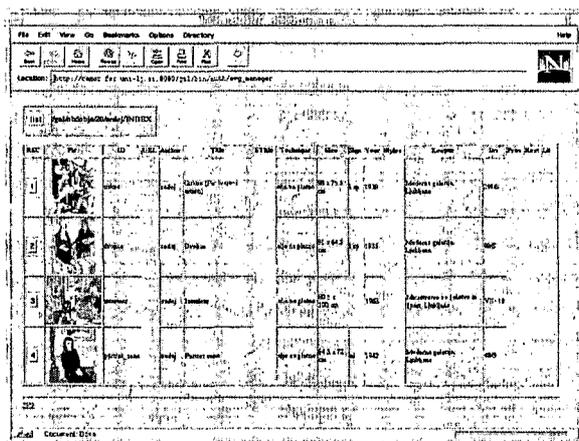
Figure 2: Management window for changing or entering new information into the SVG system

have in fact made a pilot re-implementation of SVG using a commercial object-relational data base [6]. The most critical and potentially time consuming step in such re-implementation is the conversion of existing data to the new data structure.

The virtual exhibition space of the SVG was implemented using a structure of inter-connected clickable-

maps. Each view of the 3D gallery space, which was initially constructed as a classical CAD model, was pre-rendered and converted to a clickable map by addition of carefully selected links to the next possible views. By clicking on the pre-selected areas of the clickable map the observer moves to the corresponding destination. Thus a

Figure 4: VRML model of Jakopič's death mask

selected sequence of such clickable maps forms a walk through the virtual gallery. If a visitor of the virtual gallery clicks on any of the paintings which are hung on the walls
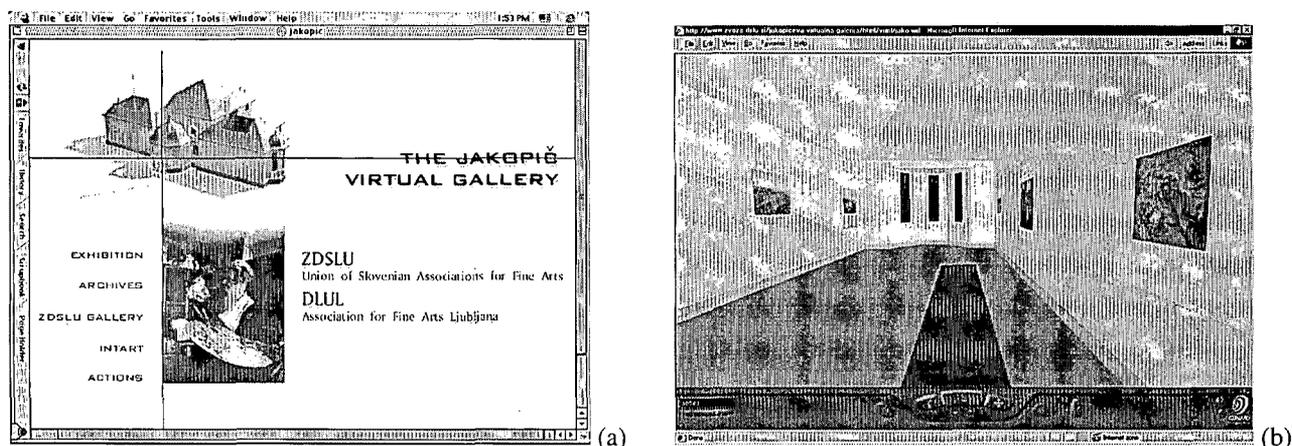
Figure 3: *(a) Home page of the Jakopič Virtual Gallery, (b) VRML model of one of the rooms*

he gets to the presentation of the paintings in the first part of the SVG. In this way all parts of SVG are interconnected. Although similar result can be obtained using a VRML model, our implementation was much faster and enabled greater flexibility in connecting to different parts of SVG since each step in a walk was just a link to another HTML document. Such predetermined paths through a virtual space are also easier to handle for a novice user who can get quite easily lost if a multitude of options are open such as in a typical VRML rendered virtual space.

SVG was warmly received in Slovenia [5, 7] and, as judged by a high number of visits, on the Internet in general. In 1996 the McKinley Group's online editorial team rated SVG as a "4-star" site excelling in "Depth of content", "Ease of Exploration", and "Net appeal". Unfortunately, no institution in Slovenia at that time showed any interest to support, maintain and upgrade the SVG system. The SVG system was a result of student work and was after the authors left the University no longer maintained. While the first part of SVG which contains the historical overview is fairly content stable, the second part was supposed to offer information on current exhibitions in several galleries in Ljubljana.

Recently, the Union of the Slovene Fine Artists Associations (ZDSLU) sponsored a project which was inspired by SVG. A VRML model of the Jakopič Pavilion, which was demolished in 1962, was built to serve as an environment for virtual exhibitions of Slovenian artists on the Internet and to celebrate the anniversary of Rihard Jakopič [3] (Fig. 3). Rihard Jakopič was the premier Slovenian impressionist painter who in 1908 actually financed the building of the pavilion in Ljubljana. In the virtual pavilion, which closely follows the original plans of the architect Maks Fabiani, is included also a 3D model of Jakopič's death mask (Fig. 4) which we rendered using our structured light range sensor [12].

3 Life video over the Internet

Life video transmission over the Internet is becoming more widespread as the capacity of the networks expands and the access speed of the end users increases. At this moment thousands of cameras, all across the earth, are sending images to web sites which can be used as our remote eyes. In 1996 we developed our own system for remote video observation over the Internet that we named Internet Video Server (IVS) [10, 11]. The IVS system consists of a camera mounted on a robot pan/tilt manipulator which makes possible to turn the camera in any direction.

The user of the IVS system observes the video image and controls the direction of the camera in a browser window shown in Fig. 5(a). This interface required the user to press the left/right and up/down buttons to move the camera. Due to buffering, slow, and uneven reaction times of the network these controls did not seem to be very predictable from the user's point of view. The reaction time of the system depended mostly on how the camera and the pan/tilt unit was connected to the Internet. Many types of connections were tested, ranging from direct computer network connections to GSM mobile phone networks. Due to the uneven network response the user could easily loose any sense of where the camera was pointing to, especially if he or she was not familiar with the location where the camera was placed.

These interface problems motivated us to design a better user interface for remote video observation [11]. Due to the precisely controlled position of the camera by means of the pan/tilt unit, individual images acquired by IVS can be assembled into panoramic 360^0 views of the surroundings (Fig. 6). These panoramic images are then used as a backdrop for live video images, to give the user the correct context for his observation. In the new "GlobalView" interface (Fig. 5(b)) one can simply drag the live video frame over the static panoramic image to define the new camera view directions.

This system for live video transmission over the Internet was used in June 1997 during the exhibition of the
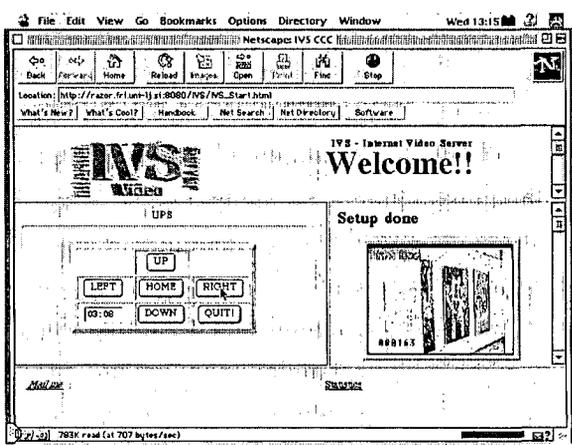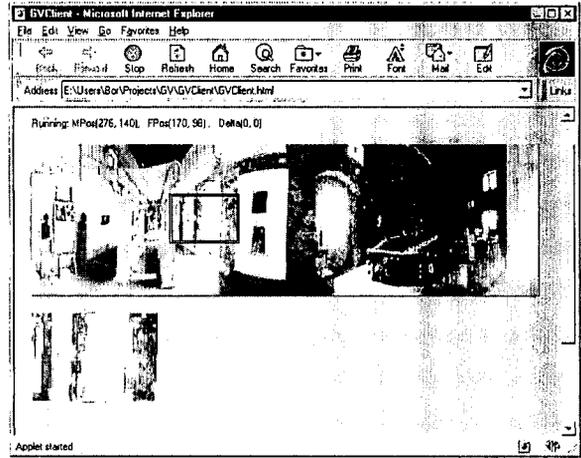
Figure 5: *(a) old Internet Video Server interface, (b) new "GlobalView" interface*

Figure 6: 360^0 *panoramic image taken in the ZDSLU gallery during the exhibition of Silvester Plotajs Sicoe in 1997.*

painter Silvester Plotajs Sicoe in the Gallery of Union of the Slovene Fine Artists Associations in Ljubljana. On the static panoramic images, taken in each room of the gallery (Fig. 6), one could click on paintings to get the corresponding pre-scanned images of these paintings (Fig. 7) and other information about the painter. From the current position of the camera platform, however, a web user could receive live video as well as control the camera to observe not only the sterile static exhibition, but also the visitors moving through the gallery.

4   Art-Internet projects

While the efforts of Computer Vision Laboratory in promoting Slovenian fine art over the Internet did not receive any institutional support, a very stimulating and fruitful collaboration started with the new-media artist Srečo Dragan. Dragan is one of the pioneers of video art and conceptual art in Slovenia. He was eager to explore and use any new technological solutions which related to his artistic interests. Our multimedia experience, in combination with our module for active Internet video observation, was used in several of Dragan's art-Internet projects and installations [14, 4, 15] (Fig. 8). These projects offered, in general, the visitor a blend of actual and virtual spaces which could be

Figure 7: *"Chair for van Gogh", Silvester Plotajs Sicoe (oil on jute, 100 × 180 cm)*

visited over the Internet. Visitors on the web could control the view direction of the camera to interactively observe
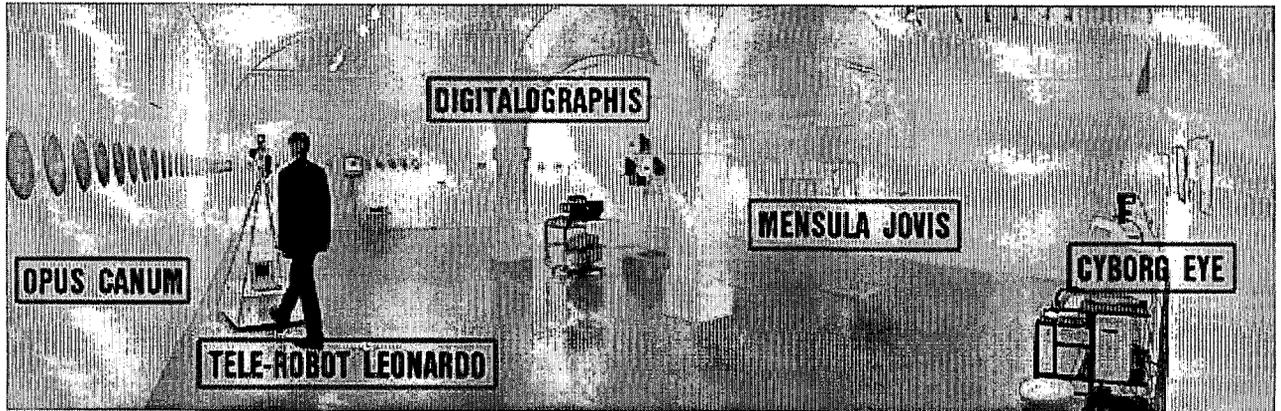
Figure 8: *Exhibition of Srečo Dragan's electronic art projects in gallery Equrna in 1997*
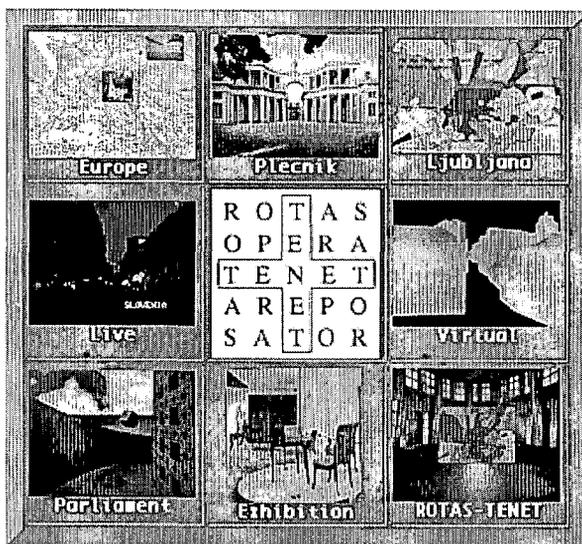
Figure 9: *Web project ROTAS–TENET*

actual physical locations which were again in an inventive hypertextural fashion connected to other virtual spaces or other visual or textural information.

The first joint interactive Internet installation ROTAS–TENET was entirely dedicated to the architect Jože Plečnik (1872–1957) and his exhibition "Architecture for the New Democracy" at the Hradčany castle in Prague (Fig. 9). To spiritually link Ljubljana and Prague by new technological means during the opening ceremony on Hradčany the IVS camera was set up on the Prešeren square next to the Three Bridges in Ljubljana which are one of the most famous demonstrations of Plečnik's mastery in urban development (Fig. 10). In the web site was included also a computer model of Plečnik's plan for a new Slovenian Parliament which was never realized. This event in May 1996 marked also the first occasion when life video from a public space in Slovenia was available on the Internet [10].
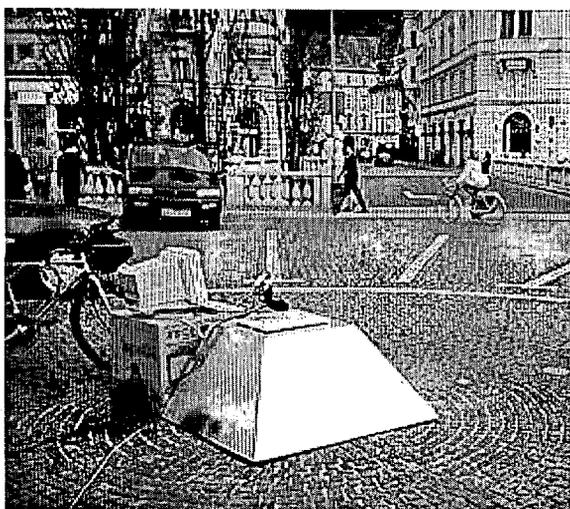
Figure 10: *The IVS camera mounted on the pan-tilt robot manipulator on Prešeren's square in Ljubljana on 23 May 1966*
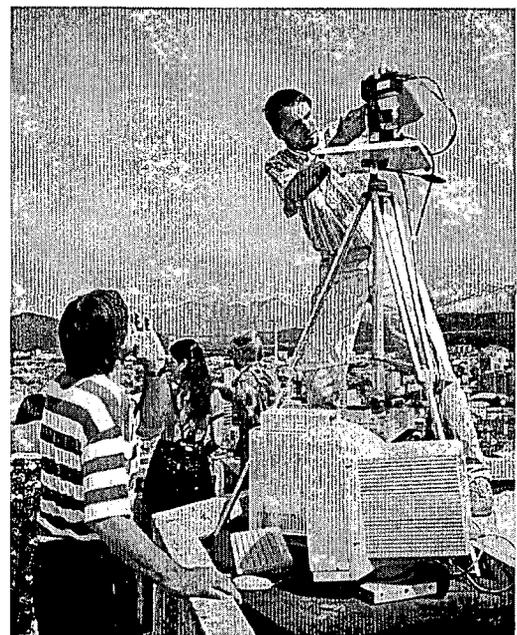
Figure 12: Scanning of panoramic images with the IVS system on top of Ljubljana castle.
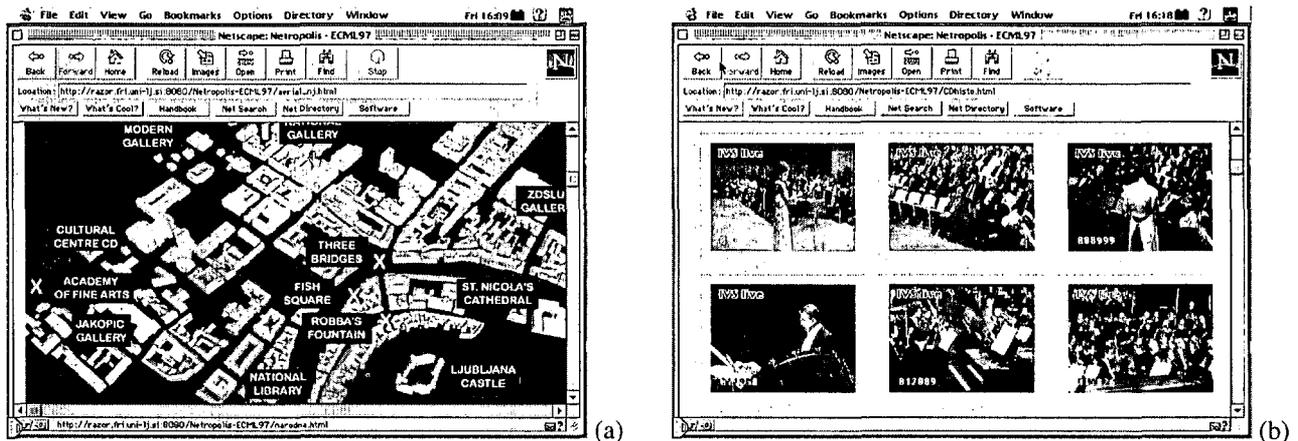
Figure 11: *Netropolis–Cyborg's Eye project: (a) 3D model of the center of Ljubljana with locations indicating where panoramic images can be interactively viewed, (b) video frames transmitted through IVS from the opening ceremony.*

The two most complex joint art-Internet projects were Netropolis–Cyborg's Eye and Netropolis–Clavis Urbis.

4.1 Netropolis–Cyborg's Eye

The Netropolis–Cyborg's Eye project was prepared for the European month of culture in Ljubljana (ECML) in the period of 15 May – 5 July 1997 [1, 2]. The project was dedicated to artistic research and experiment and was supposed to serve as a reference point for art at the entrance to the third millennium. The project was realized as an interactive installation of telepresence on the Internet which has at various locations and at different times connected Ljubljana with the world. The project was a combination of a navigable virtual architectural space, representing Ljubljana on the Internet—NETROPOLIS (Fig. 11 (a))—and of live video images, sent by means of a camera that could be controlled over the Internet from the selected points in the city—CYBORG'S EYE (Fig. 11 (b)). At these points one can cross from the (live or recorded) video images over into the digital city model and back again. At selected locations in the city 360° panoramic pictures were recorded which can be viewed interactively (Fig. 12 and 13). The conceptual plan of an interactive access from one space to another is only possible as a view from the utopian site, a non-existent, excluded point of view, which is the center of all actual and virtual plans of space and time. Additionally, images of other ECML events were digitized from video clips prepared by TV Slovenia and presented in the Chronicle section of the project web page. Unlimited possibilities of interactive dialogue formed a temporal image of the whole event which is still available on the Internet [14]. Clavis Urbis was a later "technical re-implementation" of the same concept but in the town of Slovenj Gradec.

4.2 Netropolis–Clavis Urbis

Netropolis–Clavis Urbis was the title of Srečo Dragan's participation at the U3, 2nd Triennale of Contemporary

Slovene Art in Modern Gallery Ljubljana, 14 November – 11 January 1998. The participants were selected by the curator Peter Weibel.

Netropolis-Clavis Urbis is based on a camera carrying robot telecontrolled via the Internet which was developed jointly with the Robotics group at the Jožef Stefan Institute in Ljubljana (Fig. 14). The robot [9] can be directed via the Internet to several locations in Ljubljana, represented in the gallery space by anamorfic images; at the same time, it sends a picture of the actual gallery space to the spectator–actor. But when the camera settles on the picture of the selected direction, the image on the screen transforms into a virtual, computerized picture of the city. The project thus connects the real gallery space with both the real and virtual spaces of the city of Ljubljana and with the dislocated word of computer networks [16].

5   Conclusions

We have built several art-related web sites, ranging from "classical" presentations of picture galleries to interactive art installations in cooperation with the media artist Srečo Dragan.

If a gallery or art museum web site serves only the general public, short information about collections, some pictures, opening hours, and directions for finding the museum generally suffice. Thousands of museum web sites from all over the world are now fulfilling this goal. If, on the other hand, a museum web site wants to serve also the art specialists, almost complete collections should be available in digital form, together with all other relevant information. This is of course a major, several years long undertaking, which must be carefully planed and is best organized in the framework of digitazing the collection also in high resolution for other purposes. The accompanying web design and technical solutions must be more conservative with the primary goal of finding the relevant information reliably and fast.

Web sites devoted to exhibitions of single authors or even
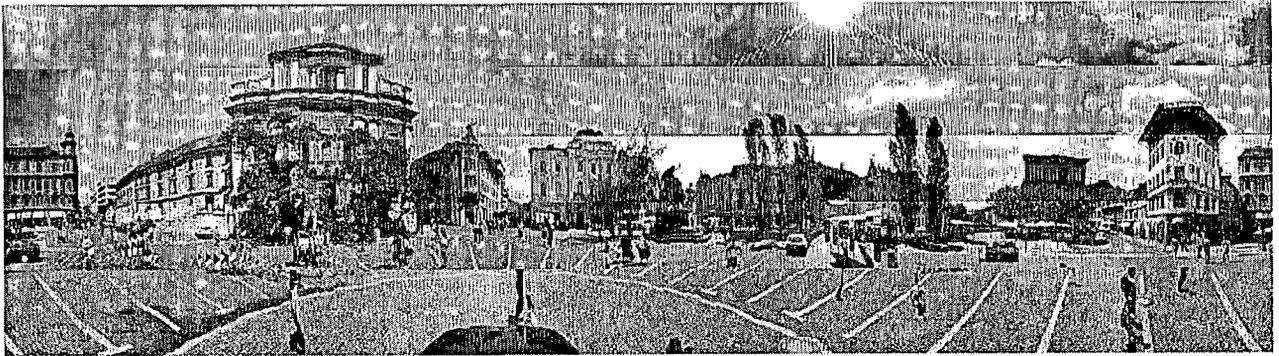
Figure 13: 360^0 view of the Prešeren square at the hearth of Ljubljana
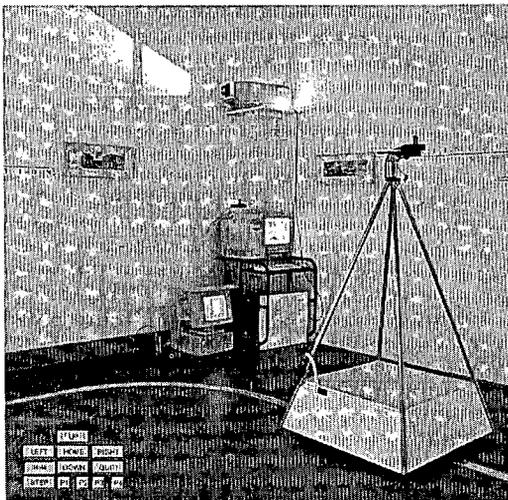
Figure 14: Srečo Dragan's exhibition at U3, Modern Gallery Ljubljana 1998

art-Internet installations are much less restrictive in design as well as in technological solutions since they are much smaller by content and usually short lived. Our art-Internet installations described in this paper served at the same time also as an excellent platform to test new and still unproven technical solutions.

From the human resources point of view, new educational programs are sorely needed to combine the knowledge of technology with design skills. Due to these art-Internet projects the collaboration between the Faculty of Computer and Information Science and the Academy of Fine Arts expanded. To fulfill part of their course requirements students from both institutions form mixed project teams to design and implement various web sites.

References

[1] S. Dragan, P. Grabnar, B. Prihavec, S. Rozman, F. Solina, and A. Vidmar. Netropolis — Cyborg's eye. In B. Zajc, editor, Proceedings 6th Electrotechnical and Computer Science Conference ERK'97, vol-

[2] European Cultural Month Ljubljana 1997, Programme. May 15 to July 5 1997. Ljubljana Municipality, 1997.

[3] The Jakopič Virtual Gallery. http://www.zveza-dslu.si/jakopiceva-virtualna-galerija/.

[4] J. Krivic Dragan, editor. Razstave/Exhibitions Srečo Dragan 1993/98. ZDSLU, Ljubljana, 1998.

[5] A. Lapajne. Slovene Virtual Gallery. M'ARS, Magazine of the Museum of Modern Art Ljubljana, VIII(3–4):90–92, 1996.

[6] A. Lapajne. Slovenska virtualna galerija — od datotečnega sistema k univerzalnemu strežniku. Diplomsko delo iz računalništva in informatike, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljàna, 1997.

[7] A. Lapajne, B. Prihavec, A. Ruben, and Ž. Kranjec. Slovenska virtualna galerija. Delo pripravljeno za prešernovo nagrado študentom, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, 1995.

[8] A. Lapajne, B. Prihavec, A. Ruben, Ž. Kranjec, and F. Solina. Slovenian Virtual Gallery. In F. Solina and B. Zajc, editors, Proceedings 4th Electrotechnical and Computer Science Conference ERK'95, volume A, pages 3–6, Portorož, Slovenija, September 1995. Slovenian Section IEEE.

[9] B. Nemec, B. Prihavec, and F. Solina. Leonardo — A mobile robot for gallery visit using Internet. In M. Ceccarelli, editor, Proceedings 6th International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD'97), pages 571–576, Cassino, Italy, June 1997. Studio 22 Edizzioni.

[10] B. Prihavec, A. Lapajne, and F. Solina. Aktivno video opazovanje preko interneta. In B. Zajc and F. Solina, editors, Zbornik pete Elektrotehniške in računalniške

konference ERK'96, volume B, pages 117–120, Portorož, Slovenia, September 1996. IEEE Region 8, Slovenska sekcija IEEE.

[11] B. Prihavec and F. Solina. User interface for video observation over the Internet. *Journal of Network and Computer Applications*, 21:219–237, 1998.

[12] D. Skočaj. Instructions to use range scanner. Technical report, Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, 1998.

[13] Slovenian Virtual Gallery. http://black.fri.uni-lj.si/SVG.

[14] Razstave/Exhibitions Srečo Dragan 1993/98. http://black.fri.uni-lj.si/Sreco.Dragan.

[15] J. Strehovec. Cybernetic art (Interactivity, game and holistic immersion). *Art Words*, 39–40:168–172, May 1997.

[16] P. Weibel and I. Zabel. Space is beyond geopolitics 1, Parallel institutional spaces, virtual and telematic spaces. *M'ARS, Magazine of the Museum of Modern Art Ljubljana*, IX(2):23–24, 1997.

Multivalued Optical Parallel Computation Using an Optical Programmable Logic Array

Abdul Ahad S. Awwal, Syed M. Munir, A. T. M. Shafiqul Khalid, Howard E. Michel and Oscar N. Garcia
Wright State University,
Computer Science & Engineering Department,
Dayton, Ohio 45435
Phone: 937-775-5111, Fax: 937-775-5133
email: aawwal@cs.wright.edu

A high speed parallel adder is designed which can perform carry-free addition of two modified signed digit quaternary numbers. Binary coded quaternary logic design is provided as one possible way of implementing the digital and/or optical system. Experimental results are also included for the proposed adder circuit.

1 Introduction

Optics provides a powerful computing paradigm (Karim and Awwal, 1992) for 2-D signal processing operations such as rotation/scale/translation invariant pattern recognition (Goodman, 1996) with possible application in machine vision, target tracking, optical memory subsystem and neural network associative memory. While the role of optics in these cases is mainly in the analog domain, digital processing using parallel optics has not reached comparable maturity. The interferenceless, free-space, parallel communication capabilities of optical signals makes it suitable for implementing digital algorithms which are inherently parallel. Thus, while a ripple carry addition is suitable for digital electronics, it is not necessarily the optimum scheme when using optical signals. It is difficult for optics to propagate the carry through different stages. On the other hand, a parallel adder will map elegantly to an optical architecture (Takagi, et al., 1986). An optimized utilization of the parallel processing capability of optics may take the avenue of multiple-valued-logic-system (MVL) which is more accurate than a purely analog system and is more powerful than simple binary based processors. MVL is gaining popularity in both digital circuits and fuzzy logic based intelligent sensors (Dao and Campbell, 1986, Hurst 1986 and Shanbag et al., 1990).

Binary signed-digit numbers are known to allow a limited-carry propagation through a complicated addition process (Takagi et al., 1985, Bocker et al., 1986, Cherri and Karim 1988, Avizienis 1961, Hwang and Louri, 1989, Awwal and Iftekharuddin 1999) requiring a very large circuit for implementation. A higher radix-based (Awwal et al., 1988, Alam et al. 1992, Hossain et al. 1998) representation of signed-digit numbers not only allows carry-free addition and borrow-free subtraction but also offers the im-portant advantages of simplicity in logic and higher storage density. When encoded using 2's complement binary, an modified signed-digit (MSD) binary number requires 2-bit/single bit whereas MSD quaternary numbers require only 1.5-bit/single bit of quaternary.

2 QSD Number

Any quaternary signed-digit number may be represented by

$$X = \sum_{i=0}^{n} x_i 4^i \qquad (1)$$

where the variable x_i can take any value from the set $\{\bar{3}, \bar{2}, \bar{1}, 0, 1, 2, 3\}$ and the symbol $\bar{x}$ represents the $-x$ value in QSD system. The modified signed-digit (MSD) numbers are also known as a redundant numbers since multiple representation exists for any number in this system. When two QSD numbers are added the resulting sum ranges from -6 to $+6$, and they may be represented in multiple ways as depicted below:

$$-6 = \bar{1}\bar{2} = \bar{2}2,$$
$$-5 = \bar{1}\bar{1} = \bar{2}3,$$
$$-4 = \bar{1}0,$$
$$-3 = \bar{1}1 = 0\bar{3},$$
$$-2 = 0\bar{2} = \bar{1}2,$$
$$-1 = 0\bar{1} = \bar{1}3$$
$$0 = 00,$$
$$+1 = 01 = 1\bar{3},$$
$$+2 = 02 = 1\bar{2},$$
$$+3 = 1\bar{1} = 03,$$
$$+4 = 10,$$
$$+5 = 11 = 2\bar{3},$$

$$+6 = 12 = 2\overline{2}.$$

bit.

The proposed addition scheme exploits the above redundancy to perform parallel addition of two quaternary numbers.

2.1   QSD Adder Design

When two QSD numbers are added $(a_i + b_i = t_i)$, a carry is generated when the absolute value of addition is greater than the radix of the number system, 3. The generation of the carry can be avoided by mapping the two digits into a pair of intermediate sum (w_i) and intermediate carry (x_i) such that the nth intermediate sum and the $(n - 1)$th intermediate carry never form any carry generating pair (3, 3), (3, 2), (3, 1), $(\overline{3}, \overline{3})$, $(\overline{3}, \overline{2},)$, $(\overline{3}, \overline{1})$. Note that if we restrict the representation such that the intermediate carry is limited to a maximum of 1, and the intermediate sum is restricted to be less than 2, then the final addition will become carry-free. Following are the selected representation of the sum of a pair of QSD numbers :

i) $-6 = \overline{1}\,\overline{2}$,
ii) $-5 = \overline{1}\,\overline{1}$,
iii) $-4 = \overline{1}0$,
iv) $-3 = \overline{1}1$,
v) $-2 = 0\overline{2}$,
vi) $-1 = 0\overline{1}$,
vii) $0 = 00$,
viii) $+1 = 01$,
ix) $+2 = 02$,
x) $+3 = 1\overline{1}$,
xi) $+4 = 10$,
xii) $+5 = 11$,
xiii) $+6 = 12$.

From the above one can show that the intermediate carry (x_i) can be expressed by the following equation

$$x_i \;=\; \begin{cases} 0, & |t_i| \;<\; 3 \\ 1, & t_i \;\geq\; 3 \\ \overline{1}, & t_i \;\leq\; -3 \end{cases} \qquad (2)$$

where $t_i = a_i + b_i$. The intermediate sum, w_i can be expressed as $t_i - 4x_i$.

Therefore, the final sum S_i of two quaternary numbers a_i and b_i can be expressed as

$$S_i \;=\; w_i + x_{i-1} = t_i - 4x_i + x_{i-1} \qquad (3)$$

Note here that x_{i-1}, which is the carry from the (i-1)-th stage is added to the intermediate sum, to get the final sum

Theorem: The addition process defined by Equation 3 is always carry-propagation-free.

It can be shown that the above equation results in a value of S_i which is less than 3. Note that when the magnitude of t_i is less than 3, x_i is 0. Again since the maximum value of x_i is 1, max S_i is less than 2+1, i.e., 3. Hence S does not exceed 3. When $t_i \geq 3$, $x_i = 1$, Therefore $(t_i - 4x_i)$ varies from -1 to $+2$. Thus $-2 \leq (S_i = t_i - 4x_i + x_{i-1}) \leq 3$. Using similar argument, it can be shown that when $t_i \leq -3$, $x_i = -1$, $(t_i - 4x_i)$ varies from -2 to $+1$. Therefore, for all possible values of t_i, $-3 \leq (S_i = t_i - 4x_i + x_{i-1}) \leq 3$.

The combinational circuit implementation of the QSD adder is possible using binary logic, when QSD numbers are represented by their equivalent 3-bit 2's complement binary numbers. Let $t'_i = a_i + b_i$. Then t'_i needs to be represented by 4 binary digits. The minterms for the correct sum, t_i are:

$$\begin{aligned} t_{i3} &= a_{i2} \oplus b_{i2} \oplus t'_{i3} \qquad (4)\\ t_{i2} &= t'_{i2} \\ t_{i1} &= t'_{i1} \\ t_{i0} &= t'_{i0} \end{aligned}$$

If each QSD input is represented using only 3 bits, then one needs to generate the correct sign bit t_{i3}. Otherwise, if 4-bit representations are chosen for the QSD inputs, then the sign bit is generated without any additional circuitry. The carry bit x is limited between -1 to $+1$. Therefore we need two bits for representing x_i. The equations for intermediate carry x_i are:

$$\begin{aligned} x_{i1} &= t_{i3}\overline{t_{i1}} + t_{i3}\overline{t_{i2}} \qquad (5)\\ x_{i0} &= x_{i1} + \overline{t_{i3}}t_{i2} + \overline{t_{i3}}t_{i1}t_{i0}. \end{aligned}$$

For the first stage, $x_{01} = x_{00} = 0$. Substituting Eqs. (4) and (5) into Eq. 3, we can get the final one bit sum.

Figure 1: Signal flow of a 3-bit adder

tronically. The smartness of each pixel could be derived from implementing Eq. 3 as shown in the circuit of Fig. 2. Second, an optical threshold logic could also be designed. In threshold logic (Cotofana et al. 2000), the inputs are summed and then compared against a threshold to determine the output. Note from Eq. 2 that the x_i bit is based on a threshold relationship with respect to the sum, t_i. The final equation could also be generated using a limited accuracy analog adder (Parhami, 1996) circuit. With the advancement of hybrid analog/digital circuits this approach may have considerable promise. In these cases, the quaternary input could be represented using a six pixeled input light (Parthasarathi and Jhunjhunwala, 1995). The position of the bright pixel will determine the value of the QSD number, and the zero will be represented by all dark pixels. The third possibility is to minimize the logic equations (mirsalehi and Gaylord, 1986, Awwal and Iftekharuddin, 1999b) for the output in terms of QSD inputs. Then one can choose an encoding for the QSD digits and derive the mask for a programmable logic array implementation (Ha and Li, 1994, Qian et al. 1999).

We demonstrate the operation the quaternary adder using an optical programmable logic array (PLA). This hardware setup and its error sources are documented in (Michel and Awwal, 1996). The details of the system, as it is used to implement the proposed efficient MSD quaternary mask, are briefly summarized below.

The PLA hardware consists of two stages of identical AND-OR logic, where the output of the first layer can be cascaded to the second one. Each logic stage consists of 16 light emitting diodes (LEDs), a liquid crystal display (LCD) and appropriate polarizing material functioning as a 16 pixel by 16 pixel spatial light modulator (SLM), 16 photodiode detectors, plus the associated control circuitry and software. The 16 LEDs, representing a 16 by 1 input vector, align with rows in the SLM array. The electro-optic setup, as shown in Fig. 3 performs a vector-matrix multiplication optically.

This multiplication generates negative logic minterms, which can be subsequently ORed to produce a sum of product expression. The SLM pixels can be individually programmed to alter their transmittances between opaque and transparent. The 16 photodiode detectors optically sum down columns in the SLM array, and represent output as a 1 by 16 vector. A personal computer serves as a user interface and controls the input LEDs, the SLM mask and displays the output of the system.

For input representation a dual-rail coding is used. In dual rail, two light sources are used to represent a binary value. The quaternary numbers are encoded in 3-bits 2's complement binary form. Therefore, 12 LEDs are required to introduce binary-encoded quaternary addend and augend.

For our experiment, we chose to generate an output for the input pair 11 that appears in the previous worked out example. The objective here is to produce a negative logic output from only this combination, while producing a zero

We show a typical example of using Eq. 3 to calculate the final sum of two QSD numbers.

a	2	3	1	$\bar{2}$
b	3	3	1	$\bar{2}$
t_i	5	6	2	-4
x_i	1	1	0	$\bar{1}$
$w_i = t_i - 4x_i$	1	2	2	0
S_i	1	2	2	1	0

The signal flow of a 3-bit adder is shown in Fig. 1. Using digital simulation software B2 Logic, Eqs. (4) and (5) were implemented in the computer as shown in Fig. 2. Simulation of the logic circuit was performed to verify the correctness of the designed circuit for QSD addition.

3 Optical Implementation

Programmable logic array (Arrathoon and Kozaitis, 1987), analog optical processors (Casasent and Woodford, 1994), content addressable memory (Ha and Li, 1994), polarization-encoding (Awwal and Karim, 1989, Zhou et al., 1995), interconnection network (Sun et al. 1996) and computer generated holograms (CGH) (Kawai and Kohga, 1992) have been used for optical implementation of MSD and other binary arithmetic processors. An optical processor for the proposed QSD adder could be developed in several ways. First, one may use a smart-pixel based architecture (Fey and Degenkolb, 2000). In these structures, both the input and the output of the 2-D processors are optical. Inside each pixel, the signal is converted to an electronic representation and then the information is processed elec-

Figure 2: Digital implementation of the QSD adder
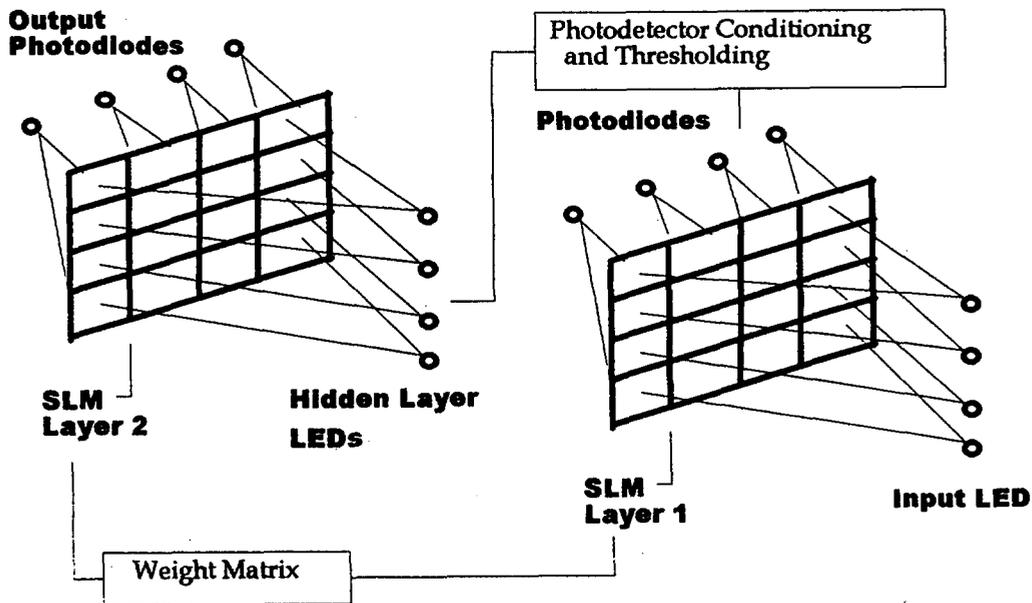
Figure 3: The PLA system

output for other combination such as 00,10, 20,30,21 and 31. These minterms are not an exhaustive list but is used to demonstrate the principle of operation.

The input LEDs representing the combination 11 are shown to the immediate left of the SLM depiction as shown in Fig. 4. Figure 4 shows both the actual output values and a superimposed graphic depicting the SLM mask and the input LEDs. The lowest LED (aligned with the lowest row) represents the a_3 term while the $\overline{a_3}$ is immediately above it. This continues up to the twelfth row from the bottom representing $\overline{b_1}$. The LEDs are colored white if they are illuminated and black if they are off. The input pattern, from the bottom to the top is 010110010110, representing:

$$a = \overline{a_3}\ \overline{a_2}\ a_1\ =\ 1 \tag{6}$$

and

$$b = \overline{b_3},\ \overline{b_2},\ b_1\ =\ 1 \tag{7}$$

The mask corresponding to $a_i b_i$ equal to 00, 10, 20, 30, 11, 21, and 31 are encoded in columns 7 through 13 respectively. Light squares represent transparent pixels, while dark ones represent opaque pixels.

The four output plots demonstrates the selection of a bias operating point and maximum noise in the system. These are chosen with one test run, along with several control runs. Since there is considerable variation in the amount of light transmitted along different optical paths, the control runs are necessary in order to calibrate the actual hardware.

Runs 0, 1 and 2 are the control runs. Run 0 plots the amplified photodiode values with all LEDs off and the SLM mask totally opaque. This line thus represents the noise floor of the detectors. Run 1 shows the basic lower operating limit for the test runs. It is obtained by making the top four rows of the SLM transparent by turning on the corresponding LEDs. This is done in order to bring the electro-optic components into a linear operating range. Run 2 has the same null input pattern as the previous run, however, the SLM mask is programmed with the selected mask. In an ideal system, run 1 and 2 would be identical. However, it can be seen that several of the photodiodes record a higher value in the later run. This is due to crosstalk in the optical components. Run 2 thus represents the actual lower operating limit for the system.

Run 3 represents the actual implementation of the selected mask. The difference between runs 2 and 3 thus represents the system's ability to perform the calculation. In the negative logic implementation, a low indicates a match. Here, clearly channel 11 represents a low, or true. All other channels are above the test floor established by run 2, although, one of them (channel 13) is barely above the test floor. The actual sum, S_i, is computed by ORing together all of the detector signals representing the entire mask.

4 Conclusion

Single step addition rules for quaternary signed-digit numbers are derived and the corresponding electronic implementation verified using digital simulation software. This
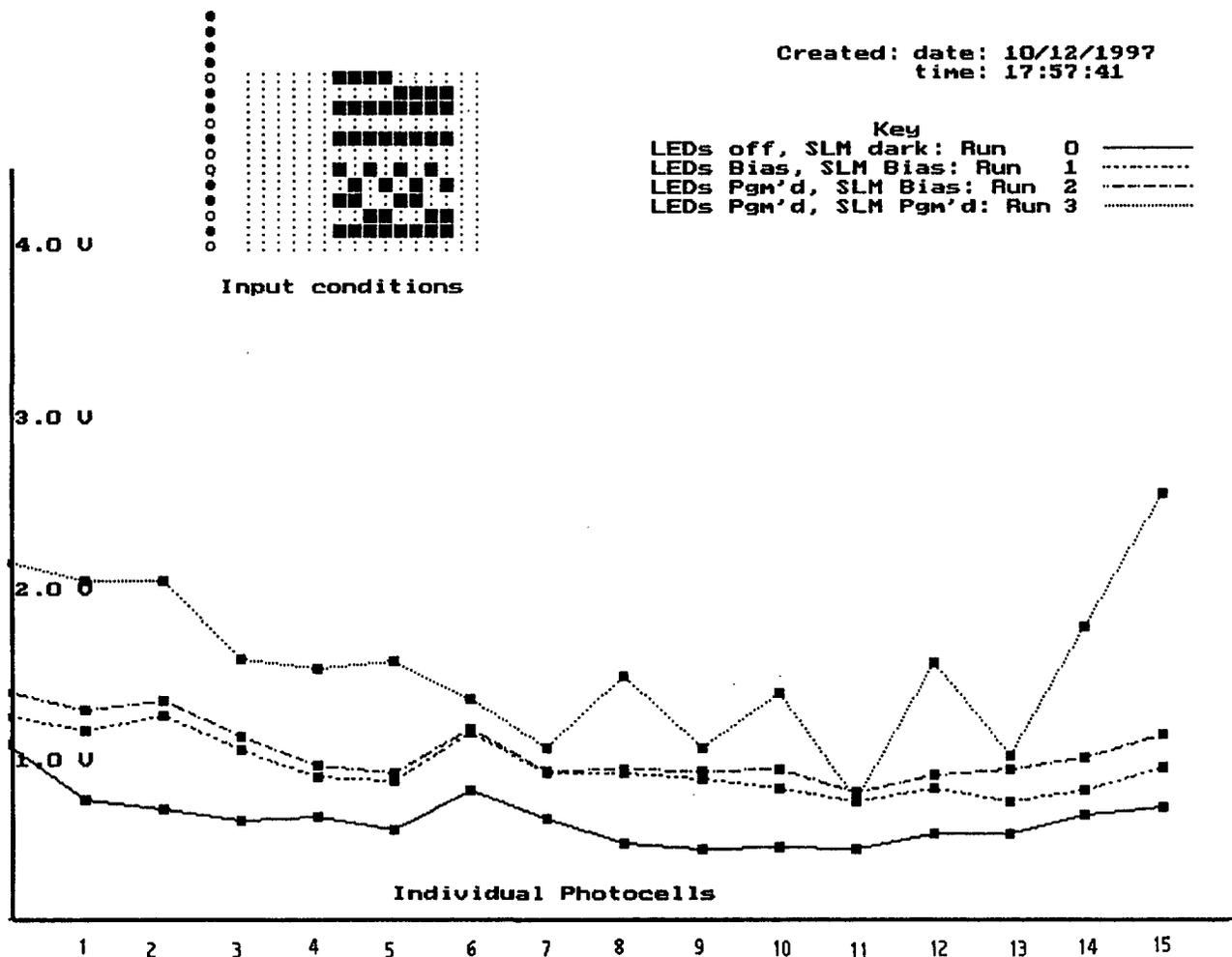
Figure 4: The optical output of the PLA

electronic implementation can be incorporated in the smart pixeled opto-electronic device. Such an implementation exploits the parallel optical communication capabilities for input/output and strength of digital electronic circuits for complex logic operations.

The optical implementation was done using an optical programmable logic array. The noise and cross talk was identified and its effect was compensated using an optical bias. The bias also demonstrates how an optical logic operation might be performed practically. Other efficient coding scheme as well as larger SLM will make implementation of the entire adder circuit possible. By controlling the mask in real time, reconfigurable computing can be achieved. The main advantage of the proposed QSD number system is the ease of conversion to and from binary.

References

[1] Avizienis A. (1961) Signed-Digit Number Representation for Fast Parallel Arithmetic. *IRE Trans. Electron. Comp.*, EC-10, p. 389-400.

[2] Alam M. S., Karim M. A., Awwal A. A. S. & Westerkamp J. J. (1992) Optical Processing Based on Higher-Order Trinary Modified-Signed-Digit Symbolic Substitution. *Applied Optics*, 31, p. 5614-5621.

[3] Alam M. S., Awwal A. A. S. & Karim M. A. (1992) Digital Optical Processing Based on Higher-Order Modified Signed-Digit Symbolic Substitution. *Applied Optics*, 31, 14, p. 2419-2425.

[4] Arrathoon R. & Kozaitis S. (1987) Architectural and Performance Considerations for a 10^7 - instructions/sec Opto-electronic Central Processing Unit. *Optics Letters*, 12, p. 956.

[5] Awwal A. A. S. & Iftekharuddin K. M. (1999) Special issue on Computer Arithmetic for Optical Computing, *Optical Engineering, March issue.*

[6] Awwal A. A. S. & Iftekharuddin K. M. (1999) A graphical approach for multiple valued logic minimization, *Optical Engineering*, 38, p.462-467.

[7] Awwal A. A. S. & Karim M. A. (1989) Polarization-encoded Optical Shadow-Casting: Direct Implementation of a Carry Free Adder. *Applied Optics*, 28, p. 785-790.

[8] Awwal A. A. S., Islam M. N. & Karim M. A. (1988) Modified Signed-Digit Trinary Arithmetic Using Optical Symbolic Substitution. *Applied Optics*, 31, p. 1687-1694.

[9] Berlin H. & Li Y. (1994) Parallel modified signed-digit arithmetic using an opto-electronic shared content-addressable-memory processor. *Applied Optics*, 33, p.3647-3662.

[10] Bocker R. P., Drake B. L., Lasher M. E., & Henderson T. B. (1986) Modified Signed-Digit Addition and Subtraction Using Optical Symbolic Substitution. *Applied Optics*, p. 2456-2457.

[11] Casasent D. & Woodford P. (1994) Symbolic substitution Modified Signed-digit Adder. *Applied Optics*, 33, p. 1498-1506.

[12] Cherri A. K. & Karim M. A. (1988) Modified Signed-Digit Arithmetic Using an Efficient Symbolic Substitution. *Applied Optics*, 27, p. 3824-3827.

[13] Cotofana S. & Vassiliadis S. (2000) Signed digit addition and related operations with threshold logic, *IEEE transactions on Computers*, 49, p. 193-207.

[14] Dao T. T. & Campbell D. M. (1986) Multiple Valued Logic: an Implementation. *Optical Engineering*, 25, p. 14-21.

[15] Fey D & Degenkolb M. (2000) Digit pipelined arithmetic for 3-D massive parallel optoelectronic circuits, *Journal of Supercomputing* 16, p. 177-196.

[16] Goodman J. W. (1996) *Introduction to Fourier Optics.* New York: McGraw-Hill.

[17] Hossain M. M., Ahmed J. U., Awwal A. A. S. & H. E. Michel (1998) Efficient Electronic Implementation of Modified Signed-Digit Trinary Carry Free Adder. *Optics and Laser Technology*, p. 49-55.

[18] Hurst S. L., (1986) Multiple Valued Logic: its Status and its Realization. *Optical Engineering*, 25, p. 44-55.

[19] Hwang K. & Louri A. (1989) Optical Multiplication and Division Using Modified Signed-Digit Symbolic Substitution. *Optical Engineering*, 28, p. 364-372.

[20] Karim M. A. & Awwal A. A. S. (1992) *Optical Computing: an Introduction.* New York: John Wiley.

[21] Kawai S. & Kohga Y. (1992) Modified signed-digit processors using computer-generated holograms, *Applied Optics*, 31, p.6193-6199.

[22] Michel H. E. & Awwal A. A. S. (1996) Noise and Crosstalk in an Optical Neural Network. *Proc. IEEE 1996 National Aerospace and Electronics Conference*, 2, p. 662 - 669.

[23] Mirsalehi M. M. & Gaylord T. K. (1986) Logical Minimization of multilevel coded function. *Applied Optics*, 25, p. 3078-3088.

[24] Parhami B. (1996) Comments on High-Speed Area-Efficient Multiplier Design Using Multiple valued current-Mode Circuits. *IEEE transaction on Computers*, 45, 5, p.637-639.

[25] Parthasarathi R. & Jhunjhunwala A., Symbolic substitution by means of one-of-many coding, *Optical Engineering*, 34, p. 1456-1463.

[26] Qian F., Li G. Q., Ruan H. & Liu L. R. (1999) Modified signed-digit addition by using binary logic operations and its optoelectronic implementation, *Optics and Laser Technology* 31, p. 403-410.

[27] Shanbag N. R., Nagchoudhury D., Siferd R. E. & Visweswaran G. S. (1990) Quaternary Logic Circuits in 2 micron CMOS Technology. *IEEE Journal of Solid State Circuits*, 25, 3, p. 790-799.

[28] Sun D. G., Wang N. X., He L. M., Wang D., & Chen R. T. (1996) Demonstration of an optoelectronic interconnect architecture for a parallel modified signed-Digit Adder and Subtracter. *Applied Optics*, 35, p. 1785-1793.

[29] Takagi N., H. Yasuura H., & Yajima S. (1985) High Speed VLSI Multiplication Algorithm with a Redundant Binary Addition Tree. *IEEE Trans. Comp.* C-34, p. 789-795.

[30] Zhou S., Campbell S., Wu W., Yeh P., & Liu H. K. (1995) Two Stage Modified Signed-Digit Optical Computing by Spatial Data Encoding and Polarization Multiplexing. *Applied Optics*, 34, p. 793-802.

MILENIO: A secure Java2-based mobile agent system with a comprehensive security.

Jesús Arturo Pérez Díaz, Darío Álvarez Gutiérrez and Sara Isabel García Barón.
University of Oviedo
Calvo Sotelo s/n, 33007 Oviedo, Spain.
Phone: +34-98-5103397, Fax: +34-98-5103354
{arturop,dario}@lsi.uniovi.es, uov01887@correo.uniovi.es

Current mobile agent systems provide very simple security models. There is a lack of implementations allowing the administrator to manage system resources appropriately while offering comprehensive security to the server and agents against mutual or third party attacks. SAHARA is a security model for Java-based agent systems that takes advantage of the benefits of the Java2 SDK v1.2.1. SAHARA's main goal is to create an integral security model that can be easily implemented by any system. This architecture offers the following features: specific assignment of privileges; agents' authorities authentication using digital signatures; allowances management to limit resource consumption by each agent and authority, and a dynamic security policy that allows to modify the security permissions at runtime. Secure agent transmission and server authentication is achieved by using the SSL protocol. Digital signatures are also used to protect agent's code and to assign responsibility on the agent's data. In order to verify the performance of the SAHARA security architecture we have created a mobile agent system called MILENIO with basic functionality, but with adequate security capabilities since it implements all SAHARA's features. MILENIO also has also a user access restriction to avoid undesired users. We managed to create a secure mobile agent system with a graphical interface that can be used in any kind of mobile applications since it assures secure transactions for most of the applications We also show the proper use of the Java2 security model in Java-based mobile agents systems.

1 Introduction

The absence of reliable security architectures is probably the key factor that constrains the wide spreading of agent technology. Security is crucial so that an agent system can be used in real world applications.

Regarding this important aspect in a mobile agent system, six different problems deriving from all possible inter-relationships between entities in an agent system (Hohl 1997) can be identified: interaction between agents, between agent systems and agents, between servers themselves, and between agent systems and unauthorized third parties. So, these are the six areas to take care of:

1. Protection of the machine or agent system against attacks from other agents. The machine must be capable of authenticating the agent's owner, as well of assigning limited resources upon this authentication. Violation of these limits must be prevented likewise, to ensure complete integrity.
2. Protection of the agent against other agents. No agent should be able to interfere with any other agent. Neither resource stealing nor any kind of forging should be possible. No agent should be capable to obtain private information from other agent by forging the identity of a third agent.
3. Protection of the agent against the agent system. A machine should not be able to manipulate the behavior of an agent or to extract sensible information without the agent's cooperation. This is a very difficult problem to tackle, as it seems that only hardware mechanisms (Wilhelm & Staaman 1998) fully solve this problem[1].
4. Protection of a group of machines against an agent. An agent could use an excessive amount of resources in a network even when only little resources were used in individual machines.
5. Protection of the machine or agent server against other machines. A machine could try to forge the identity of other machine to gain more privileges when asking for a service from a third server.

---

[1] There are some works that try to solve this problem such as Code Mess Up (Hohl 1997) and mobile cryptography (Sander & Tshudin 1998). However, their implementation is very difficult to achieve.

6. Protection of information transmission between agent servers against unauthorized third parties. Secure communications between two servers using an insecure network should be considered. Information spying as well as many known attacks detailed in (Stalling 1995) should be eliminated. A proven technique is the use of a trusted information transfer protocol.

These problems have been examined in the mobile agent literature. However, existing systems present only partial solutions to solve them. Some established agent systems suffered an evolution towards more complete security architectures . But security models for new Java-based agent systems only implement partial solutions to the problems mentioned above. Consequently, these systems do not offer enough warranties to be used in real applications.

In the second part of this paper we give an overview of the SAHARA security architecture, in the third part we show how we create our prototype implementing the SAHARA's features and in the last part we present the related work and our conclusions.

2   The SAHARA Security Model

The main goal of our research is to define a comprehensive security architecture for Java2-based agent systems that provides new security mechanisms and enhances the current security models. This architecture will guarantee an integral security for any agent system that implements the security model, so that it can be used in real applications.

The new features of the Java2 SDK v1.2.1 security infrastructure (SUN 1998) will be used as well, for greater versatility and ease of implementation of this security model.

Following is a list of objectives of the SAHARA security architecture:

- To authenticate the servers where agents come from and agents' authorities.
- To grant individual privileges to mobile agents that belong to each remote authority recognized by the system.
- To create a secure, faster way to send the agent and its classes.
- To specify the particular set of files to which access will be granted for the agents coming from each authority, as well as access privileges for each file or group of files.
- To impose allowances on each authority or remote system to control the consumption of resources in the local system by agents coming from these remote systems. The allowances are assigned not only to each remote agent but to each remote authority too.
- To allow agents to define privileges upon files generated by them at execution time.

- To protect the information transmission between servers.
- To protect the agent's code using digital signatures, and to use the signature of the authority of each server to sign the data state of an agent, so that each server signing a data state acquires a responsibility.
- To provide intelligence to the server so that it can modify the security policy when it recognize a lack of resources.
- Finally, to implement this architecture to test its feasibility and then to extend this architecture to current mobile agent systems.

2.1   Authorities, users and resources

The authorities considered in SAHARA´s security are the *users of agents*, and the *server machines*. The authorities will be responsible for signing agents when they are sent to other servers. Thus, authorities are the primary element to perform authentication and privilege granting.

The *user of the agent* is the person that initially sends the agent and is responsible for it. Authorization and accounting for the actions of an agent in a given server are usually based on the authority of this user. The *server machine's* authority is used by servers to authenticate themselves.

SAHARA protects the fundamental resources of the operating system such as files, disk space, main memory, number of net connections, etc.

Every user/authority of the agent system will have an account. Only the administrator will be able to create new users/authorities and to establish their privileges.

The administrator of the agent system will be able to grant privileges individually for every remote authority as well as to define the local user allowances.

2.2   Use of regions to create trusted networks

A region is a set of agent systems that usually have the same authority, but are placed in different locations. In practice, a common authority such as a company frequently administers these groups of servers. It is useful to treat a group of servers as a single entity from the viewpoint of security. For example, to avoid the costly cryptographic computations when the agents migrate between servers in the same group (figure 2.1)

2.3   Agent structure and agent migration

An agent is initially composed by a main class and a set of required classes (Myagent.class and other classes), agent properties (an object that includes name, authority,
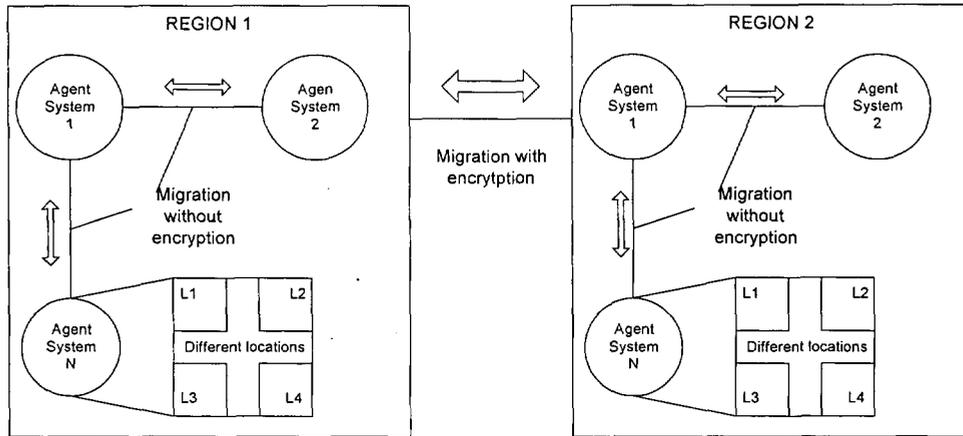
Figure 2.1 – Agents migration inside and outside regions

date of birth, etc) and its allowance (some of the privileges that are granted by the administrator). When an object is about to migrate, these elements are compressed into a Java jar file (Myagent.jar) using the JarFile API provided by Sun. The serialized state of the agent (Myagent.os) is also included in the file. All these elements are signed by the agent's authority. Signatures and the *manifest* file used for signature management are also included in the jar file (figure 2.2).

When the agent is in a remote server and its data state has changed in that server, a new Myagent.os file replaces the previous one before leaving the server. This new data state is signed with the authority of the remote server. The purpose of this signature is to assign some responsibility on the server in case the (possibly malicious) changes introduced by the server in the data state produce a misbehavior of the agent in the following server. The server's signature means "I have not maliciously altered anything in the data of this agent"[2].
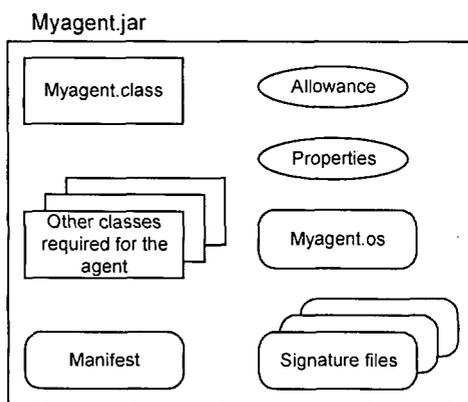
Figure 2.2 – Agents components

2.4   An Overview of the interactions scheme of SAHARA

The protection offered by this architecture is grouped into three sets:

------

[2] This process develops an idea suggested by Volker Roth.

- Protection of the machine or agent server against attacks.
- Protection of the information transmission against third parties.
- Protection of agents against malicious servers.

The overall scheme of interaction is depicted in figure 2.3, showing the different elements intervening to achieve an integral security that is concisely explained below. It will be explained in detail in the next sections.

A pair of keys is created by the agent system whenever a new user/authority is included in the system. The keys are sent to a certification authority in order certificate them. Once certified, the keys are stored in a key repository called *Keystore* and associated with their respective user/authority.

When an agent is created a unique universal identifier (in time and place) is generated with a message digests function and associated with it. Whenever the agent is about to migrate, the key belonging to its authority is retrieved from the *Keystore*, and the agent is signed with the attached allowance, which was associated from a preferences file (*DBPreferences*) when the agent was created. Once signed, the agent is sent so that it can be authenticated later.

When a remote agent is received in the system, it is authenticated using its digital signature, retrieving the associated public key for validation. Once its authority and origin are identified, the system looks out in the security policy files the granted permissions for this authority and a new protection domain is created (unless the domain were defined previously). This domain defines the privileges hold at run time by every agent originated from the same authority or server.

Once the privileges for an agent are determined, the agent is bound to its protection domain. In addition to the rights granted, the protection domain also has an allowances object called *DBResourcesFree*, which indicates the amount of resources still available for the authority owner of that protection domain. The amounts are updated whenever an agent enters the domain. When

the domain is created the *DBResourcesFree* object is constructed using the *DBResources* file that contains the upper resource limits for each remote authority in the system.

To transmit information in a secure way, the SSL (*Secure Sockets Layer*) protocol is used, as it is a standard for global networks. This protocol establishes an encrypted communications channel once the machines have been previously authenticated.

At run time, an agent can create new specific privileges upon the new files created by itself during its execution, in order to ease parallel and group work. In these privileges they can specify to which authorities the permission is granted and its access rights.

Finally, to protect agents against malicious servers the signature of the code of an agent is used as a primary measure. Signing the code guarantees that if it is altered, the next server hosting the possibly maliciously modified agent will refuse the agent, to avoid misbehavior of the agent, as it is presumably corrupted. Digital signatures of agents prevent any modification of an agent by any agent server[3].

3   The prototype MILENIO

We will describe how we take advantage of some security features of the Java2 SDK to implement the SAHARA architecture, and how we add some innovative concepts to provide comprehensive security to our mobile agent system. We just deeply describe the security implementation of the mobile agent system for space reasons.

3.1   Protection of the agent server against attacks

The protection of the agent server against attacks from mobile agents or even servers, involves two basic tasks:

- *Authentication*: It is the process of discovering which authority has made a given request.
- *Authorization and enforcement*: The purpose of authorization is to decide which privilege level should be assigned to an agent in order to perform its task. This privilege assignment is based on the authority to which the agent belongs. Enforcement monitors and assures that an agent does not violate or surpasses the limits and permission imposed by the agent server.
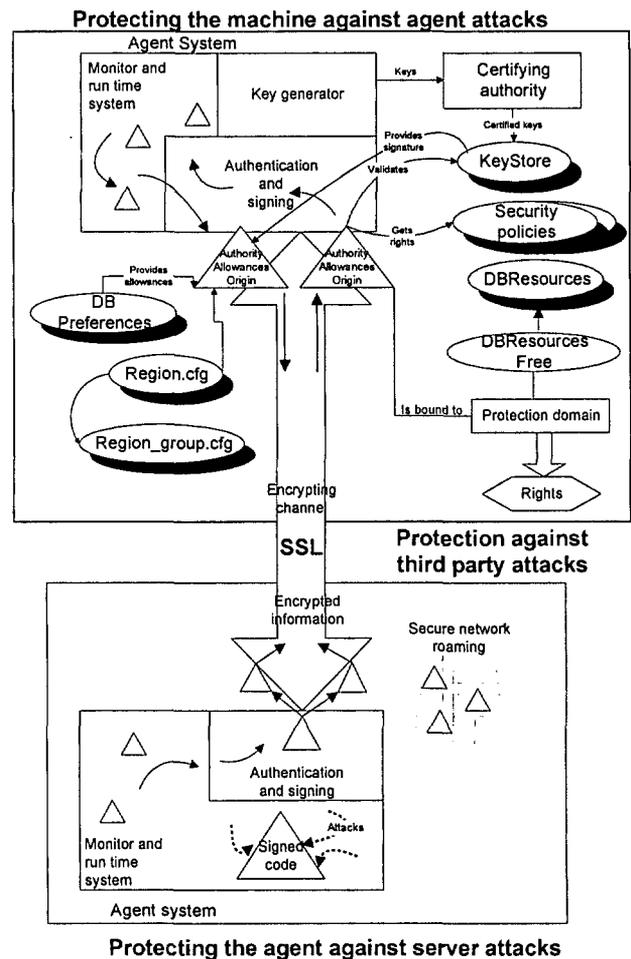
Figure 2.3 – Global scheme of SAHARA's security architecture interactions

3.1.1   Authentication

The authentication scheme is based on the digital signature attached to the files that are part of the agent. It is used to verify the signed code that represents the agent, which was originated in the remote agent server. SAHARA looks in the *Keystore* for the certificate of the corresponding authority to make the authentication and verification of the signature. Additionally, server authentication is also performed when an SSL connection is established.

Keys and certificates are stored for Each Authority. A database called *Keystore* is used to manage a repository of keys and certificates[4]. The actual format of the database depends on the implementation. The JKS format used by default in the Java2 SDK will be used initially for prototypes. The JKS format protects private keys with an individual password. The integrity of the whole database is also protected with another password.

The *Keystore* database manages two kinds of entries: key entries and certificate entries. *Key entries* store keys,

---

[3] Even though an agent system could freely manipulate the agent's code within the system and then send the unaltered original agent's code.

[4] A certificate is a digitally signed declaration of an entity stating that the public key of some authority has a particular value.

usually secret keys or public keys with the authentication certificate belonging to the matching public key, which are used for auto-authentication via digital signatures. *Certificate entries* contains public key certificates belonging to other subjects and trusted by the owner of the *Keystore*. These certificates are used to validate digital signatures of incoming agents.

Every entry is bound to an "alias" string. For a private key, the alias represents a user/ authority that is able to sign agents in the system to send the agents to roam the network. For a public key, the alias represents a remote user/authority that is authenticated in the local system.
As part of the agent authentication process, the system will retrieve from the *Keystore* database the necessary keys to sign the agents about to migrate, and to authenticate the incoming agents as well.

3.1.2 Authorization

Resources given to agents can be classified depending on the enforcement policy used to limit them:
- Privileges directly controlled by the kernel of the agent system.
- Privileges controlled by the security policy files of the agent system.

The authorization scheme is flexible enough to assign specific privileges to any remote authority recognized by the system, specifying the exact kind of privilege granted by the system.

3.1.2.1 Allowances

At agent creation time a unique identifier[5] is assigned to the agent. An **allowance**[6] object of class *ResourcesAvailable* is also attached. This object indicates the restrictions imposed on the agent whenever it migrates to a remote server. The number of child agents the agent will be able to fork and the number of Kilobytes of storage space that can be used in the remote server are among these constraints. The agent system kernel controls the **allowance** object, which has the following fields:

```
int disk        // Amount of storage KBs able to use
int childs      // Number of childs able to fork
int memory      // Amount of main memory KBs
boolean messages  // Ability (or not) to send messages
boolean duplicate // Ability (or not) to duplicate itself
......                      // others
```

The agent system administrator can bind a different *allowance* to every user/authority in the system so that

user's agents can have different privilege levels. The allowances defined for local users are stored in a system file called *DBPreferences* (figure 2.3).

To prevent an administrator from giving excessive allowances to his users when sending agents to other systems, every agent system has a database that stores the upper limit of main memory, storage space and maximum number of children for each remote or authority. This database is stored in the *DBResources* file. Whenever an incoming remote agent arrives, or a remote agent duplicates, current values of these amounts for the remote system/authority are updated. Agents coming from the same system or authority are thus able to work in other systems without trouble as long as the limit of resources imposed on these systems is not surpassed.

3.1.2.2 Kinds of privileges

There is a set of privileges that are granted according to the system security policy files. The Java2 SDK (SUN 1998) syntax is used to specify the privileges. The following list shows the permission that can be enforced:
1. **File System**: The format for these kind of rights can be specified in many ways detailed below (no blank spaces are allowed in file or directory names).

```
file
directory (same as directory/)
directory/file
directory/* (all files in the directory)
* (all files in the current directory)
directory/- (all the files in the file system of this
            directory)
-     (all files in the file system of the current
      directory)
"<<ALL FILES>>" (all files in the file system)
```

Actions that can be executed are read, write, delete, and execute. Some valid examples that create rights are shown below:

```
FilePermission p = new FilePermission
                   ("/arturo/mytmp", "read,delete");
FilePermission p = new FilePermission("/-",
                   "read,execute");
FilePermission p = new FilePermission("<<ALL
                   FILES>>", "read");
```

2. **Network access**: The format for this right is expressed as **hostname:port_range**, where the host name can be given in the following ways:

```
hostname (a particular server)
IP address (a particular server)
localhost (the local machine)
"" (same as "localhost")
hostname.domain (a particular server within a
                domain)
hostname.subdomain.domain
*.domain (all servers within a domain)
*.subdomain.domain
* (all servers)
```

---

[5] Which will be used to enter the agent in the system, and to find it in the agent references table.

[6] The concept of allowance is different from the Ara's concept. In our architecture an allowance not only impose the limits that an agent can burn up (as in the Ara system) but they also allow to impose the limits that the authority can consume.

That is, the server is given as a DNS name, an IP address, or as a *localhost*. The port range (*port_name*) uses the following syntax:

- N (a single port)
- N- (all ports starting with port number N and up)
- -N (all ports starting with port number N and up)
- N1-N2 (all ports within N1 and N2 inclusive)

Where N, N1, and N2 are integer numbers in the range 0..65535.

Accept, connect, listen, and *resolve* are the actions allowed upon sockets. Note that "resolve" is implicitly required when accepting, connecting, and listening. "listen" is applied on local ports, while "accept" can be used on both local or remote ports. Some example rights follow:

```
SocketPermission p = new SocketPermission("java
                       .uniovi.es","accept");
SocketPermission p = new SocketPermission
                       ("204.160.241.99", "accept");
SocketPermission p = new SocketPermission
       ("java.uniovi.es:8000-9000","connect,accept");
```

3. **Other privileges:** to allow the creation of windows and to restrict other security capabilities.

3.1.2.3 Security Policy Files

The security architecture defines one or more files where rights are granted to every specified authority. Every file can contain a *Keystore* and zero or more *grant* entries.

The *Keystore* stated in the configuration file is used to find the public keys of the signing parties in the *grant*[7] entries of the file. A *Keystore* must appear in the file in the case that one or more *grant* entries specify that agents belonging to the given *codesource* are signed. The syntax for this entry is:

```
        keystore "some_keystore_url", "keystore_type";
```

Where "*some_keystore_url*" specify the *keystore* URL and "*keystore_type*" states the type. The *keystore* type defines the data format and storing, and the algorithms used to protect private keys and the integrity of the database.

Each *grant* entry in the security policy file has a *codesource* and its corresponding rights. The syntax uses two reserved words. *Grant* marks the beginning of a new entry, while *permission* marks the beginning of a new right.

```
grant [SignedBy "signing parties "] [, CodeBase "URL"]{
    permission permission_class_name [ "target_name" ]
    permission ...
};
```

The order of the *codebase* (origin address) and *signedby* (signing parties) is not important. *Codebase* is optional, and will be omitted in some cases, as agents are usually sent signed along with the classes required, to be

recreated in remote servers. *Codebase* would be probably used for agents that have to be from one authority and coming from one specific site.

3.1.2.4 Privileges Authorization

The privilege authorization for agents coming from the same region, which is specified in the security policy file, could be granted depending on their origin. That is, the origin from a particular address will be verified. For example:

```
grant signedBy "156.35.31.66" {
    permission FilePermission "/tmp/*", "read";
};
```

For agents coming from outside the region, the privilege assignment will be based on the signing authority. Therefore, every agent signed by the same authority will have the same privileges. For example:

```
grant signedBy "Oviedo3_group" {
    permission FilePermission "/tmp/*", "write";
    permission SocketPermission "*", "accept";
};
```

It is worth to note that privileges can be granted depending not only on the authority, but also on the source machine. This can be specified adding the source IP address to the grant entry. For example:

```
grant signedBy "SAHARA", Codebase "156.35.31.66" {
    permission FilePermission "/tmp/*", "read";
    permission SocketPermission "*", "accept";
};
```

If an agent matches more than one entry, the permissions granted to the agent and associated with its protection domain will be the sum of all the permission granted by each grant entry.

3.1.2.5 Versatility in File Sharing Protection by Agents

The ability to perform parallel tasks is one of the advantages of mobile agents. This usually requires an agent to process information, and store results, which will be later accessed by other agents in order to continue the work. To facilitate this task in a secure way, the security policy can be updated at run time by the *share* method that was included in the agent abstract class so that the agents can protect access to their files when they created them.

The *share* method is overloaded with two forms to specify the arguments. The first option uses a variable that specifies the rights granted on the file besides the name of that file. This variable holds three unix-style octal numbers representing the rights for different groups of users on that file: file owner authority, source machine, and for all active agents in the system.

The octal numbers are an abbreviation of a three-bit binary number encoding rwx (read, write, execute) rights. When the bit is on it represents that the right is granted on that activity.

---

[7] Grant is the reserved word used to mark an entry specifying privileges for a given authority or server.

To include these privileges, the agent system must identify the authority and the origin of the agent invoking the method. The security policy file should then be edited. The system will find all the authorities and addresses involved and then it will assign the specific privileges. Finally, a refresh command on the security policy file should be issued for the new file to become active.

The second *share* option invokes the method passing additional arguments besides the file address; it passes the authority and the *codebase* to which privileges are granted. An octal number representing the rights assigned is passed to the agent system as well. Once the privileges have been edited in the policy file the refresh command is executed.

3.1.2.6 Privileges Enforcement

The unit of protection and privileges enforcement is the **protection domain**. A protection domain is a set of classes for which instantiated agents share the same set of privileges.

Every class/agent signed with the same keys and coming from the same machine is bound to the same domain, and therefore shares the same privileges. However, classes with the same signing entity but coming from different locations are bound to different domains.

The agent system administrator defines a security policy editing the policy security files. Every *grant* entry in these files specifies which new protection domains should be created and the set of permissions that will be granted to that domain.

A mapping between agents (code classes and instances) and their protection domains as well as between protection domains and their rights is created and updated by the execution environment of the agent system (figure 3.1).

Protection domains are created on demand. Whenever a new incoming agent arrives from a given location and authority, the system searches for an existing protection domain for that *codesource*. A new protection domain is not created always. If a corresponding existing domain is found, the agent is just bound to that domain instead, so that rights for that domain take effect for the agent.

3.1.2.7 Allowances Enforcement

As mentioned before, every agent system will use an allowances database called **DBResources** specifying maximum amounts for main memory, storage space, and number of children allowed for each remote system and authority.
At run time, a **DBResourcesFree** object will hold the available amount of these quantities for every authority or remote agent server that is included in *DBResources*.

This object will be bound to its existing protection domain.

Protection domains based on the address of the agent system (*codebase*) will also have a *DBResourcesFree* object showing the available resources for this remote system. Protection domains created for servers within the same region will be a usual case of having this kind of object bound.

Agents that have its protection domain based on the authority will again have a *DBResourcesFree* object bound to its domain. The only difference is that values stated in *DBResources* will be based on the authority, and not on the source server. It is shown in the next figure.

Agents, Domains/Allowances, and Rights
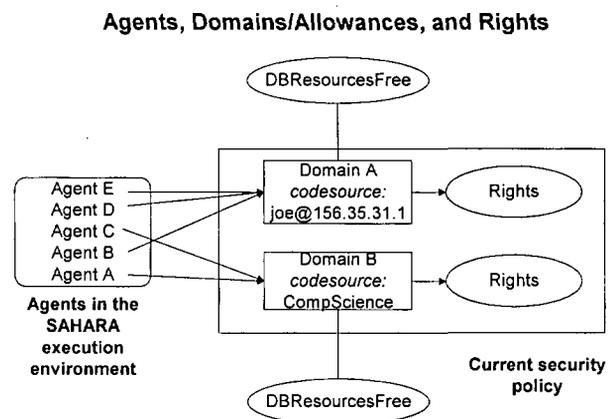
Figure 3.1 – Protection domains and allowances

Hence, whenever an agent arrives from a remote server or it duplicates itself, values in the *DBResourcesFree* object bound to the agent's protection domain are updated accordingly.

3.1.2.8 Server Self-protection

Our architecture also defines a server self-protection model where the server is monitoring the resources automatically since it is started up. When it detects a lack of one of them (i.e. RAM, H.D.) it modifies the security policy files and does not accept more mobile agents to avoid running out all the resource. In order to apply the changes made, the server executes the refresh method.

When the server detects that it has recovered enough amounts of the resources, it would then modify the policy again to allow the privileges that it had restricted before.

3.2 Protection of Information Transmission against Third Parties

Espionage, tampering, or forgery of agents (or of the information carried by agents) is another kind of attack against mobile agent systems. To avoid this attack, the MILENIO system uses the SSL protocol (Secure Sockets

Layer), currently a de-facto standard for information transmission on the Internet. In essence, this protocol provides integrity, authentication and confidentiality (privacy).

SSL connects to the socket layer of an application, and can conveniently be placed below other protocols such as HTTP and RMI, basically providing security to these protocols. The implementation is easier in comparison with other solutions.

Another advantage is that SSL is an application independent protocol, being transparent for higher-level protocols. Since SSL can authenticate the servers taking part in the connection, the possibility of attacks from other machines is avoided. The possibility of server forgery is also eliminated since the connection will not be established without the server authentication.

3.3 Protection of Agents against Attacks from Malicious Servers

The most difficult task when protecting mobile agent systems is to protect agents from attacks performed by a malicious server since the server usually has access to all parts of an agent like code and data, and has total control on it.

3.3.1 Protection of Code and Data

Fortunately, as it has been seen before, agents are digitally signed before dispatching. That is, the authority that the agent represents signs the code of the agent, not only for authentication purposes but for self-protection as well.

If the code of an agent is altered by a malicious server and then sent to other server, the new server will refuse the agent. The new server will check the signature of the agent's authority and if the code was previously altered, this signature will no longer be valid, and the agent will be rejected.

Since the alterations in the agent's code will be promptly detected the potential damages to the server caused by malicious alterations to an agent's code are avoided.

Digital signatures are also used "to protect" the data state as we describe in section 2.3. SAHARA includes partial result authentication codes (PRAC) described in (Yee 1997) in order to verify agent data integrity. However this feature has not been implemented in this first prototype.

3.4 Additional features of the Prototype

MILENIO's graphical user interface (Figure 3.2) allows creating users/authorities and agents. The main window

shows the active agents in the system. The tool bar has an icon for the most common operations. Through the information icon or using the context mouse menus is possible to see the available resources for both mobile agents and authorities. There are some options in the menus to visualize the history of the system, which specifies the creation, sent and deletion of any agent. Finally an import/export authorities' certificates tool is also available through the interface.

The lowest part of the graphic interface has two small windows. The output window shows the standard output of the system. The log window records the actions that have been carried in the core of the system.

MILENIO provides a help manual and a search utility where is possible to specify a string like "agent creation" and the system will show all pages of the system manual where it is described how to create an agent, with a specific percentage for each match. The help manual can also be viewed by contents.

The system also includes an access restriction feature that validates any user who attempts access to the system. When the user tries to log into the system the security manager validates its login and password in the keystore, where all the users allowed by the system administrator are found. When the login or password is not correct, access to the system is denied.

The prototype thus provides integral security since it implements the SAHARA architecture. MILENIO's security is comprehensive, versatile and easy to manage and enhances most of the current mobile agent systems' security (Peine 1998 and Lange 1996). The goals of the design of the security architecture are achieved too.
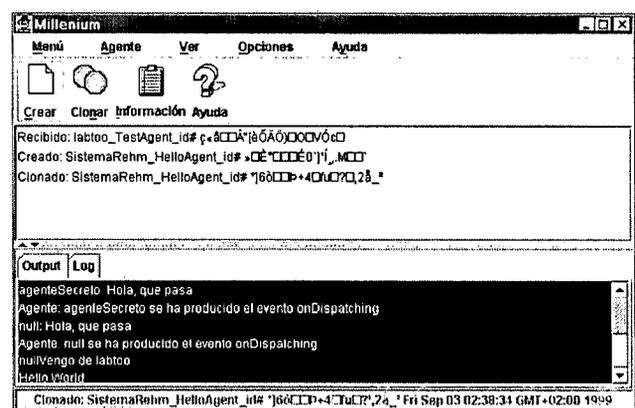
Figure 3.2 MILENIO's graphical user interface

4 Related work

There are many mobile agent systems that have implemented some elemental security features (i.e. D'Agents (Gray 1995) uses digital signatures and Ara(Peine 1997) uses allowances) but there is not any robust model implemented. The most relevant work related with our research is the security model for aglets

(Karjoth, Oshima & Lange 1997) but it has not been implemented. Nevertheless, the aglets' security model could take advantage of our architecture imposing limits to the system resources per authority and not only per agent. Also, their model could include the concept of *dynamic security policy* so that it could be more versatile and reliable according to the needs of the moment. Besides, it is possible that an agent includes a file (specifying its privileges) in the security policy during its execution protecting the work group.

Condordia's security (Walsh 1998) includes SSL and some other features like authentication, however this system as any other could take advantage of the same SAHARA's features as the aglet's model. Also any system could include the system access restrictions that SAHARA provides in order to avoid undesirable users in the mobile agent system.

Other important research on mobile agent security are the Fritz Hohl's work (Hohl 1997) and Thomas Sander's work (Sander & Tshudin 1998) that try to solve the problem of the malicious server. The Fritz's work basically consists on messing up the code of the agent before it leaves in order to avoid its readability or at least to ensure that a possible malicious host takes some time to understand and modify the code. The rest of the algorithm requires that the agent just has to spend a few seconds in the remote server to guarantee that the server will not be able to modify its code state.

The Sander and Tshudin's work uses homomorphic functions to demonstrate the possible execution of a encrypted mobile agents, in this way the malicious server will not be able to know and later modify the code of the agent.

The biggest drawback of the first method is that most of mobile agent's applications require that agents spend more than a few seconds inside each remote server they visit. So, the method is not useful for the most of applications. The disadvantage of the second method is that its implementation is not possible in practice, so it cannot be use it in real applications. One of the advantages of the SAHARA security model is that all the techniques that it uses can be implemented easily.

4.1 Comparison of security features with other systems

As expected the security architecture overloads the performance of the system, specially the digital signature of each agent. We made a comparison with the aglets system[8] that has not implemented any security mechanism for the agents. We found the system

considerably slower just when the agents are signed, however faster than we had expected. When the agents were not signed our system was faster. The table 4.1 shows the averages in milliseconds of the creation and migration time[9].

Agent System	Aglets	MILENIO
Creation Time First agent	140	420
Creation Time Next Agents	18	10
Migrating Time First agent	5398 not signed	15383 signed
Migrating Time Next Agents	4567 not signed	11627 signed
Migrating Time Next Agents NOT SIGNED	4567	2300

Table 4.1 – Performance comparison between MILENIO and Aglets system

The creation time for the first agent is slower in MILENIO because it opens many files in order to control all the security capabilities it has. We do not consider the sending time because this time mainly depends on the network traffic.

We also compare the resume time, which is the time that agent systems take since it receives the mobile agent and resumes its execution. The Aglets system took 2,1 seconds while MILENIO took 2,6 seconds. The difference of 0,6 seconds is worthless if we consider that MILENIO authenticates the authorities and performs privileges assignments and many other security actions.

4.2 Current work on Mobile Agents System Security

There are a few new mobile agent systems which the two main design objectives are security and interoperability. They are SOMA acronyms of Secure and Open Mobile Agents (Bellavista, Corradi & Stefanelli 1999) and SeMoA acronyms of Secure Mobile Agents.

SOMA basically offers security creating different places with different security policies. A place can accept or discard agents depending on authentication, which is perform while the agent is entering to the place basis on a series of protected information (i.e. name of the originating place, authority). Additionally SSL can be use to protect transmission among servers (but the protocol has not been implemented). To provide interoperability SOMA implements the CORBA's MASIF. However, SOMA does not use public key cryptography for authentication and does not have

---

[8] We executed the agents in an old network with Pentium MMX 200 Mhz and 64 Mb computers, in modern computers the differences between both systems are smaller. The results are in milliseconds.

[9] This is the time required by an agent to be ready to go, it implies the serialization, compression and signed process (if exists)

complete control of the amount of resources consumed. Also SAHARA is much more flexible and versatile.

SeMoA has a wider security than SOMA, its security model is very similar to ours. However, it does not allow modifying and updating the security policy during execution time when an agent shares a file or when the security manager detects the lack of one resource. SeMoA also goes deeper in interoperability. It defines interfaces that allows a SeMoA agent go to different platform and vice versa maintaining the uniqueness of each platform (e.g. by providing appropriate wrappers or deriving a common model). Unfortunately some parts of the mobile agent system are still under develop and this research work is still unpublished. The reader can find it in the Mobile Agent List (http://www.informatik.uni-stuttgart.de/ipvr/vs/projekte/mole/mal/preview/preview.html).

5   Conclusions

The SAHARA security architecture is designed to offer comprehensive security for any Java2 mobile agent system that implements it. To test the concept and to evaluate the performance of the architecture, a prototype implementation of a simple mobile agent system called MILENIO with this security architecture has been developed.

MILENIO offers agent and server authentication, fine-grained security since particular rights for every authority in the system can be defined down to the level of individual resources (i.e. files), and the use of allowances to restrict resource consumption by every agent and remote authority. Agent transmission between agents systems is protected by SSL. Agent code is also protected using digital signatures.

It is worth to note that the security policy can be updated at run time, so files generated dynamically by agents can be protected. The system can update the policy as well for self-protection reasons when a low resource level alarm is detected.

All of these capabilities provide a great versatility to implement a wide variety of applications from simple electronic commerce applications or digital libraries where access to one or more books of a specific subject is paid for, to complex applications that are required to modify the security policy in runtime such as banks transactions.

We managed to implement this prototype in one year and a half, with four persons. It achieves the primary goal of our research: to overcome the lack of a comprehensive security architecture for Java-based agent systems.

With this architecture and with its future enhancements we want to provide an efficient and easy way to implement a security model for agent systems. The model will contribute to the development of mobile agents since it removes many security concerns on mobile agents systems, which are possibly the biggest obstacle for the wide spreading of the mobile agent technology.

6   Future Work

Our primary research line will be to research how to verify the RAM and hard disk consumption by each agent since the JVM does not provide any interface to verify threads consumption. We will also try to find more efficient way to protect the data state of an agent during its itinerary.

In a second phase, we will try to improve this security architecture in order to avoid other attacks that could happen during messages and classes transmission. This involves the definition of a secure scheme for messages and classes management (including a cache manager that gets the classes that have been brought to the system by previous agents instead of going to the original server for them. The manager also must also consider versioning).

Finally, we will design patterns of interoperability with another agent systems in order to allow secure agents information exchange among mobile agents of different platforms (Aglets, Concordia, etc.).

7   References

[1] Bellavista P., Corradi A. & Stefanelli C. (1999) A Secure and Open Mobile Agent Programming Environment. *Proceedings of the Fourth International Symposium on Autonomous Decentralized Systems (ISADS '99), pages 238-245, IEEE Computer Society Press,* 1999

[2] Gong Li. (1999) Inside Java 2 Platform Security. *Addison Wesley Reading,* Massachusetts, June 1999.

[3] Gray Robert S. (1995) Agent Tcl: A transportable agent system. *Proceedings of the CIKM Workshop on Intelligent Information Agents.* Baltimore, Maryland, December 1995.

[4] Hohl Fritz (1997) An approach to solve the problem of malicious hosts in mobile agent systems. *Institute of parallel and distributed systems,* University of Stuttgart, Germany. 1997.

[5] Lange Danny & Chang Daniel T. (1996) IBM Aglets Workbench Programming Mobile Agents in Java A White Paper Draft. *IBM Corporation.* September 1996.

[6] Karjoth Guenter, Oshima Mitsuru & Lange Danny. (1997) A security Model for Aglets. *IBM Research division.* July-August 1997.

[7] Peine Holger & Stolpmann Torsten (1997) The Arquitecture of the Ara Platform for Mobile Agents.

Proceedings of the first International Worshop on Mobile Agents, MA'97. Berlin, Germany. April 1997

[8] Peiner Holger (1998) Security concepts and implementation in the Ara Mobile Agent system. *Proceedings of the WETICE '98.* Palo Alto California, USA, 1998.

[9] Sander Thomas & Tshudin Christian (1997). Protecting Mobile Agents against Malicious Host. *International workshop in Mobile Agent Security.* November de 1997.

[10] Stalling William (1995) Network and Internetwork Security. *IEEE Press,* 1995.

[11] Sun Microsystems (1998) JavaTM 2 SDK, *Standard Edition Security Documentation,* Version 1.2.1. 1998.

[12] Walsh Tom, Paciorek Noemi & Wong David (1998) Security and Reliability in Concordia. Mitsubishi Electric ITA, *Horizon System Laboratory.* USA, 1998.

[13] Wilhelm Uwe & Staamann Sebastian (1998) Protecting the Itinerary of Mobile Agents. *Proceedings of ECOOP Workshop on Distributed Object Security and 4th Workshop on Mobile Object Systems.* Belgium, 1998.

[14] Yee Bennet S.(1997) A Santuary for Mobile Agents. *USCD.* April 1997.

Heuristic Clustering of Reusable Software Libraries[*]

Anestis A. Toptsis
Dept. of Computer Science
York University, Toronto, Ontario, M3J IP3, Canada
Phone: (416) 736-2100 ext. 66675; Email: anestis@yorku.ca

In this paper we address the problem of software classification in the context of organizing libraries of reusable software components, and propose a method for organizing software libraries. The described method relies on an off-the-shelf heuristic that can automatically estimate the similarity between two software components. A domain analysis on the software collection at hand is required prior to starting to build the repository. This analysis needs be performed only once during the lifetime of the collection. Once such an analysis is completed, the repository is organized automatically. The software components are expressed in a standard knowledge representation language designed for information systems. The proposed method is tested through a prototype, on a small, but realistic, software collection. The experiments demonstrate that the method organizes a software repository in such a way that it facilitates high levels of retrieval quality as well as having the property that functionally similar components are clustered together. Also, the method is very robust and indifferent to the order of insertion of the software components into the repository. Certain drawbacks and pitfalls are identified and discussed.

1 Introduction and background

Software reuse is the process of developing software for a new system by using the software from other systems, thereby lessening or even avoiding the need for developing new software from scratch [17]. It is of growing importance for corporations to have effective reuse of software artifacts as they invest in developing and maintaining large software systems [28]. At the time of this writing, there are numerous large caliber software reuse activities and initiatives. Notable examples are the U.S. Dept. of Defense (DoD) *Global Command and Control System (GCCS)* initiative [4], and GTE's *Rapid* method [2]. The GCCS is the next generation of the current World Wide Military Command and Control System, and it is aimed to be built entirely with software reuse. According to [4], this will potentially save billions of dollars in software development costs. The key technical issues of software reuse are in the areas of domain analysis, classification of software components, interoperability of software repositories, adaptation of software components, reuse of systems designs and architectures, and software metrics that quantify eligibility for reuse [31]. Reportedly, no standards exist for any of these areas.

At the heart of the classification of software components issue is the organization of the software library -- also called the repository -- which is used to store software artifacts in various forms (or descriptions). Beginning with Prieto-Diaz's Ph.D. dissertation in 1985, a number of software repository organization methods have been reported [7, 19, 11, 3, 1, 25, 24, 6, 28, 14, 29, 32]. All software classification schemes contain two main ingredients: (1) a *software representation mechanism*, and (2) a *software repository organization* method. Both ingredients serve the purpose of providing the means for placing a collection of software components into a data bank which facilitates efficient and effective retrieval of software for reuse. The *software representation mechanism is* responsible for "translating" the software of the original collection into a uniform language, thus removing several incompatibilities such as source programming language and target operating system.

Typically, the most important operation of the software representation mechanism is to capture the functionality of a component at hand. Once all components have been expressed in such a uniform format, the *software repository organization* method is responsible for placing the components into a repository. The main requirement in this task is that the repository maintain the components in some order which facilitates systematic browsing and, upon request, retrieval of the most pertinent components for reuse. Table 1 shows the taxonomy of the existing

---

software classification methods in the light of the above two ingredients. The acronym NL in Table 1 and

Represent ation method	System	Organization and retrieval mechanism
Facet	Diaz [7]	conceptual graph
	CLIS [3]	hierarchical conceptual graph
	AIRS [25]	conceptual graph
	REBOOT [28]	conceptual graph
Natural language	GURU [19]	hierarchical clustering, NL processing
	CATALOG [11]	Boolean expressions, NL processing
	REUSE [1]	ad-hoc indexing, conceptual graph
	Nie [24]	ad-hoc indexing, NL processing
	LaSSIE [6]	ad-hoc indexing, NL processing
Frame	ROSA [14]	frame index, NL processing
	NLH/E [29]	frame index, NL processing
	WS [32]	frame index, NL processing

Table 1: Taxonomy of existing software classification schemes

In terms of *software representation* we observe that all the above systems use some representation paradigm directly borrowed from a discipline other that software engineering or information systems. The facet-based approach is borrowed from library science, where facets have been the standard classification method. The NL-based approaches have a clear artificial intelligence and text document retrieval flavor, although due to its nature, NL can be considered a bit more generic. The above choices may be due to historical reasons, that is, lack of a representation paradigm geared toward software systems at the time that the methods in Table 1 were designed. Recent advances in the development of such knowledge representation

hereafter stands for "Natural Language".

paradigms, allow (or even necessitate) the investigation of software repository organization methods in which software is expressed using a representation methodology especially designed for such purposes. A notable example of such a representation paradigm is the Telos language [21], especially designed for information systems.

In terms of *organization and retrieval*, it is apparent from Table 1 that the existing methods typically rely either on a conceptual graph or NL processing. The construction of a conceptual graph typically requires significant manual effort and expertise. Moreover, it does not effectively facilitate incremental changes, that is, in case that new software artifacts are inserted into the repository, the update of the current conceptual graph is a major task requiring additional amount of significant manual effort, expertise, and in-depth knowledge of the current structure of the graph. The NL approach is seemingly by far the other most popular organization and retrieval mechanism. Its advantages are based on the wide availability of NL processing algorithms that can be instantly employed and applied for software artifacts, provided that the latter are treated as generic text documents. Additional advantages come from the possibility of parsing those documents automatically, and then, based on recorded term frequencies, or aided by a knowledge base, classifying the software artifacts either automatically or with nominal manual effort. However, the whole concept of NL processing obviously relies on the availability of software in the form of text documents. To the best of our knowledge, this is unfortunately not the case for the vast majority of available software. Also, for many cases where textual descriptions are available for software, these documents provide very skim and/or cryptic descriptions, as, for example, is the case in legacy systems. This makes the foundation of NL approaches for software organization and retrieval rather unrealistic, at least under the present status of software availability.

In this paper we address the problem of software classification in the context of organizing libraries of reusable software components, using a heuristic (and non-hierarchical) clustering method. The method relies on an off-the-shelf heuristic that can automatically estimate the similarity between two software components. Given this estimator, the method is then able to create clusters of software components in such a way that functionally similar components belong to the same cluster. A domain analysis on the software collection at hand is required prior to starting to build the repository. This analysis needs be performed only once during the lifetime of the collection. Once such an analysis is completed, the repository is organized automatically. The software components are expressed in a standard knowledge representation language (subset of Telos) designed for information systems. The proposed method is tested through a prototype, on a small, but realistic software collection. The experiments demonstrate that the method organizes a software

repository in such a way that it facilitates good levels of retrieval quality. Certain drawbacks and pitfalls are identified and discussed.

The rest of this paper is organized as follows. Section 2 describes our method. Section 3 provides experimental results that illustrate the effectiveness of the method in constructing the software repository. In Section 4 we summarize our findings and discuss future research directions.

2 Software Repository Organization

In this section, we present a method for organizing the software library. The proposed method is an enhancement based on the *heuristic clustering* method of Salton for document classification [27], and a modification of a similar method used for heuristic search algorithms in [23].

2.1 Software Component Representation

The software components to be inserted into the repository are expressed in terms of their functionality, using a knowledge representation language based on Telos [21]. The importance of using a Telos-like language to represent knowledge about information systems is discussed in [21], as is the value of using a knowledge representation language. We use the representation technique and structure outlined in [10, 13] to represent the components. In that representation method, each component is expressed as a *functional description* (*FD*). A FD consists of one or more *features*. A feature is a tuple (*verb, noun, weight*), where the *verb* is the action or operation performed, the *noun* is the object upon which the operation is performed, and the *weight* is a number indicating the relative importance of the feature within its FD. In other words, the weight represents the degree of functionality that a particular *verb-noun* pair contributes to a software component. A sample FD with three features is shown below.

FD: < (open, file, H); (read, file, M); (close, file, L) >
The letters H, M, L represent weights for High, Medium, and Low respectively. The mapping between the letters and weight values is (Very High, High, Medium, Low, Very Low) = (1, 1/2, 1/4, 1/8, 1/16). In case of synonym verbs or nouns, a thesaurus is used to store those synonyms.

2.2 Software Component Similarity

Given any two software components, our repository organization scheme needs to compute their *similarity*. The similarity between two components is a number indicating how close the two components are, for the purpose of using them interchangeably during software development. Several similarity computation methods have been reported in various disciplines and contexts. Possibly the most well-known are the ones dealing with text retrieval [26, 27]. Since we deal with

software artifacts rather than plain text documents, we feel that the techniques described in [10, 13] are most suited for our purposes. The similarity computation method in [10, 13] takes as input two software components expressed as functional descriptions and returns a number between 0 and 1, which indicates how similar the two components are. The closer this number is to 1, the more similar the components are, and visa versa. For example, a similarity value of 1 between two components means that the one component can completely replace the other in the software development process. [10] gives an excellent step-by-step example of the similarity computation method. For convenience, we also provide such an example next. Following the terminology in [10], given two components expressed as FD's, one will be referred to as the *source* FD and the other as the *stored* FD. The computation involves several matrices as described below.

The EQ (equivalence) matrix expresses the degree of keyword compatibility between the i-th feature of the source FD and the j-th feature of the stored FD[1]. EQ is a *f-source* x *f-stored* matrix, where *f-source* is the number of features in the source FD and *f-stored* is the number of features in the stored FD.

The IMP (importance) matrix shows the degree of satisfaction that a source description is compatible (or can be replaced) with a stored description. Its entries show the importance between the j-th feature of a stored FD, and the i-th feature of a source FD. This importance is computed as $min(1, B/A)$ where B is the weight of a (verb-noun) pair of the source FD, and A is the weight of a (verb-noun) pair of the stored FD[2]. The values of A and B are members of the set {1, 1/2, 1/4, 1/8, 1/16}, which represent weight values corresponding to {VH (very high), H (high), M (medium), L (low), VL (very low)}. IMP is a *f-stored* x *f-source* matrix, where *f-source* is the number of features in the source FD, and *f-stored* is the number of features in the stored FD. For any entry of the EQ that EQ[i][j] = 0, in the IMP matrix, the entry IMP[j][i] is also zero. The value of the remaining entries of IMP is computed as min(1, B/A) as described above.

The W (weight) matrix holds the *normalized* values of the source FD. Each entry of matrix W is a number between 0 and 1, and represents the percentage of the corresponding weight within its FD. There is one weight per feature in any given FD, therefore, the size of matrix W is f x 1, where f is the number of features in the FD.

---

[1] A thesaurus is used to store synonyms (if any) and their degree of compatibility. In case of no synonyms, the compatibility degree is 1.

[2] We define the IMP matrix slightly different from [10]. Namely, in [10] it is stated that B corresponds to the stored FD while A corresponds to the source FD (i.e., the reverse of what we use). When carrying out tests using that definition we observed that some seemingly similar components gave erratic similarity value. By reversing the definition we got similarity values more consistent with the intuitive similarity of the components.

verb-noun pairs. The average number of features per FD is 1.9.

3.1.2 Construction of the Reference Components

Our next task was the construction of reference components. This is tightly related with the issue of domain analysis, since the reference components must define "adequately" the multidimensional space that serves as our global software repository. By "adequately" is meant that the reference components should be well spread out and not themselves clustered in a certain small region of the space, otherwise computed component similarities will be mostly identical. This implies that the reference components should capture the common characteristics of the software components of that domain, so that, in turn, "the essential functionality required in that domain is captured" [8]. Preliminary tests revealed that randomly generated reference components do not capture our domain. In general, domain analysis is a highly manual task, similar to the knowledge acquisition process in expert systems. According to [9], "domain analysis is a knowledge intensive activity for which no methodology or any kind of formalization is yet available". After thorough examination and closer look of the FD's at hand, we were able to construct a set of reference components that capture the essential functionality of the data structures domain. Our set consists of 27 reference components. By no means should this be considered a unique set. The reference components contain a total of 55 features, and an equal number of distinct verb-noun pairs.

3.1.3 Query Processing

Component retrieval is done as follows. A query component is first provided by the reuser, and takes the form of a functional description. A similarity computation is then performed between the query and the reference components. This transforms the query into a vector in the multidimensional space defined by the reference components. The distance of this vector from the centers of all clusters is then calculated in the same way that this is done when placing components into the clusters as described in Section 2. If no cluster has a center within distance T from the query vector, then no components are retrieved and the answer to the query is null. In the case that one or more clusters have their center within distance T from the query vector, then the closest such cluster is identified and all components of that cluster are retrieved as answer to the query. In either case, the *recall* and *precision* values of the retrieval operation are computed (if no qualifying cluster exists, the recall is obviously zero, and the precision is also set to zero). The recall and precision are typical information retrieval evaluation measures, defined as follows.

$$Recall = \frac{number\ of\ retrieved\ and\ relevant\ components}{total\ number\ of\ relevant\ components\ in\ the\ collection},$$

$$Precision = \frac{number\ of\ retrieved\ and\ relevant\ components}{number\ of\ retrieved\ components}.$$

3.2 Evaluation methodology

We have implemented a prototype tool that organizes the repository using our clustering scheme, accepts queries in the form of software components represented as FD's, and retrieves components under the guidelines set in subsection 3.1.3. In order to illuminate the behavior of the clustering algorithm of section 2 from different angles, numerous tests were conducted. We test the clustering algorithm in terms of *quality of response* to the user

Quality of response to the user: We examine how the clustered collection performs in terms of recall and precision for a variety of different queries. The major difficulty in designing this benchmark is to establish criteria for query construction. Obviously, given any software collection there is practically an infinite number of queries one can pose. Moreover, different queries may result to widely different recall and precision values. Two factors that clearly cause such fluctuations are the *type of match* between a query and certain FD's and the *query size*.

Type of match: Among the posed queries, some may have a *complete match* with certain FD's of the collections, some may have a *partial match*, and some may have *no match*. By *complete match* we mean the occasions where all verb-noun pairs of a query Q can be found within a component FD-i. By *partial match* between Q and FD-i, we mean that some of the verb-noun pairs of Q can be found within component FD-i, and some not. By *no match* we mean that no verb-noun pair of Q can be found in component FD-i. Note that in our definition of match, we require only the verb-noun pairs to match, i.e., we are not interested in the weights of the matching features. (If the weights match too, then all the better, but it is not required.) This is because we assume that identical verb-noun pairs represent identical portions of software artifacts and the weight only indicates the importance of these portions within the artifact. Therefore, when such an artifact is requested via a query, the importance of the individual pieces within the artifact does not alter the level of desirability for the artifact.

Query size: Our second consideration in designing the queries is the number of features in the query. Note, provided that a query's verb-noun pairs are in the collection, large-sized queries have little chance of being completely matched since this requires that a certain FD contains all verb-noun pairs appearing in the query. By the same token, small-sized queries have a better chance of being completely matched since they

only require the collection to have a FD that contains the small number of verb-noun pairs that appear in the query.

Based on the above query design considerations, we constructed 305 queries which are divided in the following four sets.

Query set A: 150 queries. In this set, each query is identical to a FD from the collection, i.e., all FD's that are inserted into the repository are also later used to query the repository.

Query set B: (small size queries). This set contains 55 queries. Each query consists of one feature. The 55 verb-noun pairs appearing in this query set are all distinct 55 verb-noun pairs appearing in our collection.

Query set C: (average size queries). This set contains 50 queries. Each query consists of two features with distinct verb-noun pairs. The two features of each query are randomly selected among the 285 features available in our collection.

Query set D: (large size queries). This set contains 50 queries. Each query consists of four features, with distinct verb-noun pairs. The four features of each query are randomly selected among the 285 features available in our collection. Since the average FD size in our collection is 1.9 features, a query with four features is considered quite large.

3.3 Experimental Results

The 150 FD's arrive in a particular OIC and they are clustered using threshold $T = 0.35$. The recall values per query fluctuate widely when all 305 queries of the sets A, B, C, and D are posed to the repository. The best performing queries are those in set B, followed by the queries in set A, followed -- with significant degradation in recall -- by the queries in set C. The recall values for the queries of set D are mostly zero.

Note, each query in query set B consists of only one feature. As such, the queries of set B are easy to match with FD's in the collection. Moreover, each such query potentially has multiple matches within the collection -- one match for each FD that contains the verb-noun pair appearing in the particular query. This is why query set B exhibits the highest recall values. The abrupt dips in the recall values are caused by the low threshold value of 0.35. At 0.35, the T value is too discriminating of which FD's qualify for inclusion with each cluster, that is, it allows into the same cluster only FD's which are highly similar. Therefore, any two FD's which are somewhat similar (but not highly similar) are placed into different clusters. On the other hand, since those FD's are similar, most (or all) of them are probably relevant to the same query. Since our query answering mechanism selects only *one* cluster for retrieval, only the relevant FD's placed into the selected cluster will be retrieved and the remaining relevant FD's would not be retrieved. This results in a low recall value. The higher recall values in Figure 1 mean that most/all relevant FD's are placed into the same cluster *and* this cluster is subsequently chosen for retrieval. In these cases, the

FD's which are relevant to the query are also very similar to each other, so that the low threshold (0.35) does not succeed in partitioning them into different clusters.

Query set A is the next high-match query collection, since each query in set A is guaranteed to have a complete match with one FD from the collection (this follows by the way the query set A was constructed, as described in subsection 3.2). However, since each of the queries in set A has usually more than one feature, the number of FD's that match these queries is smaller than the number of FD's that match the smaller queries of set B.

This trend continues for the recall values for query set C. The degradation in the recall for query set C is due to the phenomenon illustrated in Figure 1.

Since each of the queries in set C consists of two features *randomly* selected among all the features of the collection, it is quite unlikely that both features f1, f2 of a query Q belong to the same FD. As shown in Figure 1, given a query Q = {f1, f2}, feature f1 belongs to a FD with features f1, f3, f4, while feature f2 belongs to another FD with features f2, f5. Since, in general, nothing guarantees that the two FD's are similar, they may have been placed into different clusters. Since when posing query Q only the FD's of *one* cluster are retrieved (see subsection 3.1.3), the other FD's that happen to be similar to Q but were placed in another cluster would not be retrieved. This causes the recall value to drop.
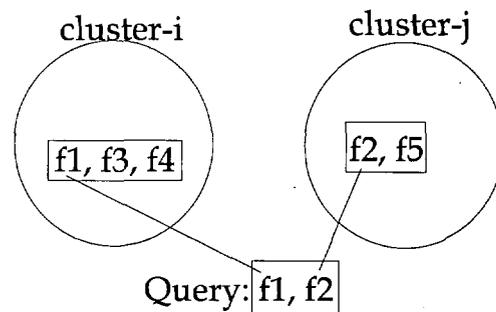
Figure 1: Clustering of query set

Since each query in set C consists of two features, it is expected that, on the average, half of the FD's that have one of the two features of Q are placed into one cluster and half are placed into another cluster. This means that the expected recall for the queries of set C is about half the expected recall in the case when both features of Q were in the same FD (note, this happens for the queries of set A). This expectation is confirmed by the average recall value for query set C (0.31 recall for set C vs. 0.59 recall for set A).

Besides the numerous cases of set D, there are seven cases in which the attained recall value is zero. This happens for queries 26, 160, 161, 191, 228, 247, and 253. In these cases no center of any cluster is of distance 0.35 or less from the query vector. As a result, following the guidelines for retrieval ("query processing" in subsection 3.1.3), no cluster is selected for retrieval and, therefore, no FD's are retrieved. This results in a

recorded recall value of zero. In conjunction with the results for the "OIC" and "cluster tightness" benchmarks (see later in this subsection), this means that the formed clusters are pretty tight, that is, the FD's of each cluster rally very close around the center of the cluster. In fact, this also explains the large number of zero recall values for query set D. Note, a query in set D consists of four features -- say f1, f2, f3, f4. Since those features have been randomly selected among all 285 features of the collection, they most likely belong to different FD's. Assume FD-1 contains f1, FD-2 contains f2, FD-3 contains f3, and FD-4 contains f4. During the process of inserting these FD's into the repository, the similarity value between the FD's and the reference components determines the location of each FD to be close to the reference component that shares a feature or a verb-noun pair with the FD. Consequently, the four FD's are inserted into four different clusters, with each cluster formed nearby the reference component that shares a feature with that FD. When query Q containing all four features f1, f2, f3, f4 is posed, its location in the multidimensional space is computed with respect to the reference components R1, R2, R3, R4. Since each of R1, R2, R3, R4 shares a feature (or verb-noun pair) with Q, the location of Q is estimated to be somewhere in the center of mass of R1, R2, R3, R4. This results to Q being located *outside* the influence of any of the four clusters that contain the FD's with features f1, f2, f3, f4. Consequently, none of those four clusters would be designated for retrieval. Although the above situation may seem to be an ironic and undesirable side-effect of the clustering scheme, it demonstrates that the center of each cluster is located in such a way that it acts as a true representative of its cluster's contents. Clearly, this is desirable. Considering the fact that none of the four FD's FD-1, FD-2, FD-3, FD-4 is actually a good match for query Q (since each FD only matches 25% of Q), we argue that the low recall values recorded for query set D actually constitute a desirable result.

The high precision values mean that most components retrieved as an answer to a query are relevant to that query. Since these components constitute a single cluster (by the way we perform the retrieval), we can infer that our clustering mechanism behaves well, in the sense that it places components that have high degree of similarity among each other into the same cluster. This is supported by the high precision values observed for query sets A, B, and C.

The average of the precision values were recorded for query sets A, B, C, and D. The extremely low precision values for set D are due to the reasons that we described for the explanation of the low recall values for the same query set, where no cluster was selected for retrieval for a query from set D. In these case, the precision value is set to zero.

The true merits of the clustering scheme were revealed. They show the variation of recall and precision against the threshold T. A single query from set A is used for the tests. We see that as T increases, the recall also increases, up to a certain T (0.6), and then, for greater values of T (0.8, 0.9) it falls rapidly. The observed increase in recall is due to the fact that larger T values allow more components to be placed into the same cluster. In conjunction with very high precision values, this means that the components that are placed into the same cluster are also similar.

For smaller values of T (e.g., 0.1), although similar components are placed into the same cluster, the small T value becomes too discriminatory, and this results in forbidding components which are relevant to the query but exhibit a smaller degree of similarity among each other, to be placed into the same cluster. Continuous increases in the value of T have desirable effects on both recall and precision. However, when the T value becomes too large, the clustering mechanism loses its ability to cluster the components effectively. This affects both the recall and precision, and it is due to the following anomalies.

(a) *Garbage-in-cluster syndrome*: When T is very large, the criteria for a component to qualify for membership into a certain cluster become very loose. As a result, dissimilar components are allowed to be placed into the same cluster. This results in having clusters which in addition to containing a core of similar components, contain a lot of components that are mostly unrelated. The undesirable result of this anomaly is illustrated for T = 0.8 and 0.9. The extremely low attained precision values mean that among all the retrieved components (which are all the components of a certain cluster), only few (or none) are relevant.

(b) *Center derailment syndrome*: As large T values allow many components to enter the same cluster, the center of that cluster is updated many times until the cluster formation process is complete. Since, due to the *garbage-in-cluster syndrome* many of these components are not mutually related, the k-dimensional points that represent those components are far away from each other. As a result, the center of the cluster under formation moves *widely* before it settles.   Since the movement of the center is caused by mutually unrelated components, the final position of the center within the k-dimensional space is most likely away from the center of mass of the small core of mutually related components in that cluster. In other words, for large values of T, when the clustering process is complete, the center of each cluster does not represent the mass center of the most mutually related components of that cluster, but it represents the mass center of a collection of mutually unrelated (or very little related) components. The latter may be far away from the core of mutually related components. As a result, the wrong cluster may be selected for retrieval when a query is posed.
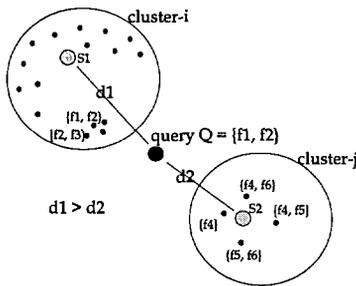
Figure 2: Center derailment syndrome

As illustrated in Figure 2, although cluster-i contains more components which are relevant to Q, due to the *center derailment syndrome*, the final positions of the centers S1 and S2 cause the illusion that cluster-j is the most relevant cluster. The undesirable effect of the center derailment syndrome is felt by the rapid fall in recall values for T = 0.8 and 0.9. Note, the high recall (100%) for T = 1.0 is because when T = 1.0, only *one* cluster is formed and, therefore, all relevant components are found in that cluster. The low precision for T = 1.0 shows, however, that this cluster (which is the entire collection) contains also many components which are not relevant to query Q.

4   Conclusion

We presented a heuristic clustering scheme for organizing software repositories for software reuse. We use a subset of a knowledge representation language (Telos) geared to information systems for describing the functionality of the software components, and an off-the-shelf heuristic that is able to compute similarity between two components. As such, our scheme is generic in at least two ways. First, the detail of description of each software component can be enhanced (or reduced) to any level deemed appropriate by the person who translates the actual software artifacts into Telos descriptions. Second, the heuristic that computes software similarity can be replaced by other more accurate estimators, should such are, or become, available.

The effectiveness of the proposed scheme was tested on a software collection in terms of quality of response to the user and robustness. Our tests indicate that the quality of response is very satisfactory on the average, however, it exhibits fluctuations on a query-by-query basis. We also determined that for large queries, the recall and precision suffer significantly. However, as we discuss in section 3, this is a phenomenon caused by the inflexibility of the query processing mechanism that we use as well as by the intrinsic nature of the query (i.e., the fact that there is not actually any good match for large queries within the collection), rather than from the clustering mechanism per se. In terms of robustness, our scheme behaves extremely well for reasonable threshold values (determined to be up to 0.7 in a scale from 0.1 to 1) used to perform the clustering of the software components. In the process of testing, we identified two intricacies in our method. First, for threshold values

above 0.7, the clustering mechanism degrades rapidly. Second, regardless of the threshold value used, the user should be aware of large clusters. As discussed in section 3, large clusters are prone to the center derailment syndrome, which is a condition causing clusters to accumulate dissimilar components. Fortunately, for small threshold values, the phenomenon of large clusters (and center derailment) is quite rare, and also, its effects are not as severe.

We currently investigate additional heuristic clustering approaches for organizing software libraries. Especially, we are experimenting with a method that does not rely on prior knowledge of the software domain's characteristics. Clearly, this has the advantage that the tedious domain analysis task can be skipped prior to organizing the software repository. On the other hand, we aim in automating the domain analysis task -- at least in the context of building reference components automatically. We are currently developing a tool that implements automatic generation of reference components based on knowledge acquired by observing users' querying patterns over a period of time.

Finally, we recognize that better query processing protocols are needed for querying a repository organized by our clustering scheme. An obvious proposal is to perform intra-cluster search, once a target-for-retrieval cluster has been identified. This will allow ranking the components within the target cluster, and thus produce better recall and precision. In cases where no cluster qualifies for retrieval under the currently used threshold (as it happens especially for large queries), a possible approach is to select all closest clusters and then perform an intra-cluster search in them for component screening. Note, besides addressing our query answering concerns, such a method may illuminate useful strategies for software synthesis, since the union of all components retrieved from different clusters would hopefully be a valid answer to the query.

5   Acknowledgments

6   References

[1]   S. Arnold and S. Stepoway, "The REUSE System: Cataloging and Retrieval of Reusable Software", *Proc. COMPCON '87 Conference*, 1987, pp. 376-379.

[2]   M. Bucken, "GTE Reengineers Development Method", *Software Magazine*, Sentry Publ. Co. Inc., Westborough, MA, USA, March 1995, pp. 42-43.

[3] W.G. Cho, Y.W. Kim, and J.H. Kim, "CLIS: A Software Reuse Library System with a Knowledge Based Information Retrieval Model", *Proc. Pacific Rim Conference on AI*, 1990, pp. 402-407.

[4] P. Constance, "New Joint Ops System will be Built with Reused Code", *Government Computer News*, June 5, 1995, pp. 68, 70.

[5] M. D'Alessandro, P. Iachini, and A. Martelli, "The Generic Reusable Component: an approach to reuse hierarchical OO Designs", *Proc. 2nd International Workshop on Software Reusability*, Lucca, Italy, March 1993, pp. 39-46.

[6] P. Devanbu, R. Brachman, P. Selfridge, and B. Ballard, "LaSSIE: A Knowledge-based Software Information System", *CACM*, Vol. 34, No. 5, May 1991, pp. 34-49.

[7] R. Prieto-Diaz and P. Freeman, "Classifying Software for Reusability", *IEEE Software*, Vol. 4, No. 1, January 1987, pp. 6-16.

[8] R. Prieto-Diaz, "Domain Analysis for Reusability", *Proc. COMPSAC'87 Conference*, Tokyo, Japan, 1987, pp. 23-29.

[9] R. Prieto-Diaz, "Domain Analysis for Reusability", *Domain analysis and software systems modeling*, R. Prieto-Diaz and G. Arango (eds), IEEE Computer Society Press, 1991.

[10] S. Faustle and M.G. Fugini, "Retrieval of Reusable Components using Functional Similarity", *Institution Universita di Pavia, Technical Report ITHACA.POLIMI.E.6.92*, 1992.

[11] W. Frakes and B. Nejmeh, 1988, "An Information System for Software Reuse", *IEEE Tutorial: Software Reuse: Emerging Technology*, W. Tracz (ed), IEEE Computer Society Press, 1988.

[12] W. B. Frakes and C. J. Fox, "Sixteen Questions About Software Reuse", *CACM*, June 1995, Vol. 38, No. 6, pp. 75-87, 112.

[13] M. Fugini and S. Faustle, "Retrieval of reusable Components in a development Information system", *Proc. 3rd International Workshop on Software Reusability (IWSR-3)*, 1993, pp. 89-98.

[14] M.R. Girardi and B. Ibrahim, "Automatic Indexing of Software Artifacts", *Proc. 3rd Inter. Conf. on Software Reuse*, W. Frakes (ed.), IEEE Computer Society Press, Rio de Janeiro, Brazil, November 1994, pp. 24-32.

[15] M. L. Griss, 1991, "Software Reuse at Hewlett-Packard", *Proc. 1st International Workshop on Software Reusability*, W. Frakes (ed.), Dortmund, July 1991.

[16] Capers Jones, "Software Challenges: Economics of Software Reuse", *IEEE Computer*, July 1994.

[17] Charles W. Krueger, "Software Reuse", *ACM Computing Surveys*, Vol. 24,. No. 2, June 1992.

[18] Robert L. Kruse, *Data structures and Program Design in C*, 2nd edition, Prentice-Hall, 1991.

[19] Yoelle Maarek and Daniel Berry and Gail Kaiser, "An Information Retrieval Approach for Automatically Constructing Software Libraries", *IEEE Trans. Software Engineering*, 17, 8, August, 1991, pp. 800-813.

[20] W. Myers, "Workshop explores large-grained reuse", *IEEE Software*, January 1994, pp. 108-109.

[21] J. Mylopoulos, A. Borgida, M. Jarke, and M. Koubarakis, "Telos: representing knowledge about information systems", *ACM Trans. Information Systems*, October 1990.

[22] Proceedings of the Sixteenth Annual NASA/Goddard Software Engineering Workshop: Experiments in Software Engineering Technology, Software Engineering Laboratory, December 1991.

[23] P. Nelson and A. Toptsis, "Search Space clustering in parallel bidirectional heuristic search", *Proceedings of 4th UNB AI Symposium*, 1991, pp. 563-573.

[24] J. Nie, F. Paradis, and J. Vaucher, "Using Information Retrieval for Software Reuse", *Proc. 5-th International Conf. on Computing and Information (ICCI'93)*, O. Abou-Rabia, C. Chang, W. Koczkodaj (eds.), IEEE Computer Society Press, Sudbury, Ontario, Canada, May 1993.

[25] E. Ostertag, J. Hendler, R. Prieto-Diaz, and C. Braun, "Computing similarity in a reuse library system: an AI-based approach", *ACM Trans. Software Engineering and Methodology*, July 1992, pp. 205-228.

The Impact of Visualisation on the Quality of Chemistry Knowledge

Margareta Vrtačnik, Vesna Ferk and Danica Dolničar
University of Ljubljana, Faculty of Science and Engineering, Department of Chemical Education and Informatics,
Vegova 4, Ljubljana, Slovenia
Phone: +386 61 214 326, Fax: +386 61 125 8684
margareta.vrtacnik@guest.arnes.si, vesna.ferk@uni-lj.si, danica.dolnicar@uni-lj.si
AND
Nataša Zupančič-Brouwer
University of Amsterdam, Department of Chemistry,
Nieuwe Achtergracht 166, NL - 1018 WV Amsterdam, The Netherlands
nbrouwer@chem.uva.nl
AND
Mateja Sajovec
Osnovna šola Simona Jenka Kranj
Ulica XXXI. Divizije 7A, Kranj, Slovenia
mateja.sajovec@guest.arnes.si

The most important result of the long-term project entitled Computer literacy is that the majority of Slovenian primary and secondary schools are now equipped with multimedia computers and with LCD projectors. However, these computerised classrooms should not be used only for teaching computer science and informatics; they should also serve for teaching and learning other subjects e.g. chemistry, physics, languages, etc. Internet offers chemistry teachers previously unavailable possibilities for bridging the gap between concrete and abstract chemical concepts and processes. Research on the spatial ability of students and the quality of knowledge show that well developed spatial abilities enable better results when solving complex chemical problems, especially when dealing with 2-D representation of 3-D chemical structures. In this article we discuss how chemistry teachers can use specialized Internet websites for visualising chemical structures and processes on the macro- and microscopic level, and correlate properties of molecules with their structure. We also present results which demonstrate the effects of different visualisation elements on the quality of chemical knowledge.

1 Introduction

The ancient Chinese knew that one picture can symbolise the meaning of ten thousands words, and modern neuro-psychological studies have proved this is possible. A human being can remember 20% of what has been read, 30% of what has been heard, 40% of what has been seen, 55% of what has been told, 60% of the content after personal involvement in a subject, and 90% if all the elements are combined [1].

The pioneer work in the field of visual methods has been done by Jan Amos Komensky in his Orbis sensualium pictus (1657). His basic idea is that teachers should present everything through pictures, thus trying to visualise the subject-matter presented at school. Modern technology offers numerous possibilities to achieve these goals, especially through websites which offer various possibilities for making chemistry classes more attractive, thus helping the student to bridge the gap between concrete and abstract notions of chemical concepts and processes.

The significance of the visualisation of abstract knowledge structures can be linked with Gibson's theses of ecological visual perception, which state that mental conditions depend on the interactions between the entities and the objects from the environment. The objects can be either concrete, direct observations of phenomena or processes at a macroscopic level, or visualised microscopic explanations of processes, or their symbol presentation, using chemistry knowledge. In all these cases of perception the underlining idea is that visualisation supports cognitive processes [2].

The research so far indicates that a well developed spatial ability can have a significant impact on solving chemical problems [3, 4, 5, 6]. The correlation between visualisation and understanding is particularly strong

when solving   problem-based tasks, or those which require memorisation, but less so for algorithmic tasks. In chemistry there are some difficult areas, e.g. visualising three dimensional objects from a two-dimensional presentation. Even more difficult is to visualise objects from another perspective, or visualising a picture of an object after rotation, reflection, inversions, etc. In such cases, visual presentation can

2   Visualisation elements and the Internet in view of chemistry curricula for primary and secondary schools

The need for more extensive use of visualisation tools has been highlighted in the new chemistry curriculum for Slovenian schools. The Internet, which can provide quick access to various useful visual elements, can play an important role. In order to help Slovene chemistry teachers to access the relevant websites more easily, a special website, KemInfo (Slovenian Chemical Information Network) was designed at our Department. It offers numerous links to other relevant websites; including a special module on the visualisation of chemical structures and processes, organised according to relevant curricular chemistry topics. These have been analysed for primary and secondary schools and after extensive Internet searches we found a number of visualisation elements, which could be used for supporting presentations of chemistry themes. The summary of this search   is listed below:

- images; visualisation at the macroscopic or microscopic level.
- films; visualisation at the macroscopic or microscopic level.
- process animations; visualisation at a macroscopic or microscopic level.
- molecular models; microscopic visualisation of structures.
- symbol presentation of molecules and reactions; symbol visualisation.
- schemes; reaction schemes for symbol visualisation; concept schemes for visualisation of the relations between concepts and concept groups.
- graphs, tables; visualisation of the relationships between data.
- knowledge test; visualisation at the macroscopic or microscopic or symbol levels.

The KemInfo website (http://www.keminfo.uni-lj.si) includes numerous links to other related web pages which offer either visual elements or complete teaching units.

The following symbols are used to denote links:
◈   symbol for molecules and reactions
▧   molecular models
▦   tables

▥   graphs
▦   teaching units
△   experiments
▨   images
▣   films, animation
✎   task items
品   schemes

For accessing some web pages we need additional web browser programs (e.g. CHIME, RASMOL, VRML, FLASH). These allow the simple visualisation of molecular structures in different formats, or show an animated image.

3 The impact of multimedia visualisation on student knowledge

In spite of the rich assortment of multimedia packages for chemistry, very little has been published regarding the impact of multimedia on the knowledge and motivation of students, let alone the cost-effectiveness of these products.

In the masters thesis by M. Sajovec [18], the impact of multimedia on cognitive, motivational and motoric development of students was studied. Based on the results of the study the evaluation criteria for assessing multimedia packages for chemistry teaching were set up. The research included 50 third year secondary-school students in 1996/97, who were divided into the experimental group (26 students: 15 males and 11 females) and the control group (24 students: 8 males and 16 females). Both groups were pre-tested. The pre-test consisted of six multiple choice questions and open-answer questions to check understanding of the following topics:

- matter and building blocks of matter,
- light,
- electromagnetic radiation spectrum,
- interactions between light and substances (absorption and in emission),
- photosynthesis, and
- radical halogenation of alkanes.

The students from the experimental group worked in pairs, using the CD ROM "Light and Chemical Change" [19], which was developed at our department and implemented in Toolbook computer based training (CBT) software package (Asymetrix). This teaching unit was designed with an interdisciplinary approach,

integrating chemical, biological and physical concepts which are related to the phenomena of photosynthesis, and describing interaction of light with matter on the macro- and microscopic levels. The topics were designed as four interrelated segments:

- Oxygen, the life-supporting gas
- Light phase of photosynthesis

Figure 1: Correct answers from the pre-test by both groups
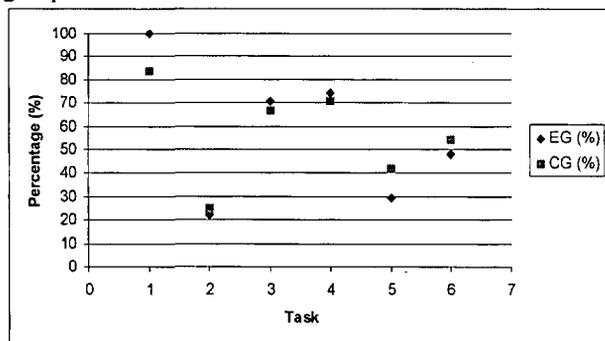
Legend: EG (=experimental group), CG(= control group)

course books available, including a 1st year biology course-book, a 4th year physics book, and a 3rd year chemistry course-book. The topics from these course-books were first presented by the teacher in three one hour periods. After a week, both groups took a common test. The test had a similar concept structure to the pre-test, the difference being that greater emphasis was placed on understanding the concepts at the microscopic level. The results of the pre-test are shown in Figure 1 and the final test results in Figure 2. The comparison of both graphs shows that the performance of the two groups prior to the test (i.e. before working with the CD ROM and the teacher presentation) was equal. Even the distribution of points achieved is similar, while the results of the final test show significant differences in favour of the experimental group which can be ascribed to the visualisation elements included in the CD ROM.

Figure 3: Visualisation elements – spectra, animation
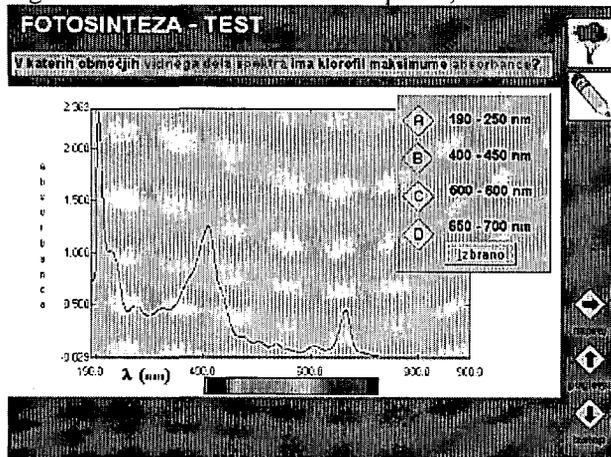
- Interaction of light with matter
- Example of a simple photochemical reaction - bromination of hydrocarbons

After studying with the CD ROM, the students had three hours available for solving the tasks. They did not use any additional literature for consolidating their knowledge. Students from the control group had several

Figure 2: Correct answers from the final test by both groups
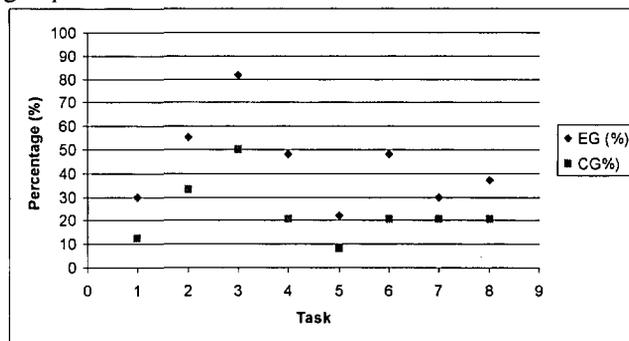
Legend: EG (=experimental group), CG(= control group)

To illustrate the impact of visualisation on understanding the concepts and the use of knowledge, the analysis of some final-test items is given below: Each item is presented with key visualisation elements from the CD ROM, textual description, and graphic presentation of student performance of both groups.

Task 4
Visualisation elements: spectra, animation of the absorption (Figure 3, Figure 4)

Question
The graph below presents the absorption spectrum of chlorophyl. In which wave-length range of the visible spectrum will chlorophyl absorb the light? Will chlorophyl absorb the light in the UV part of the spectrum? Does chlorophyl absorb green light?

Figure 4: Visualisation elements – spectra, animation
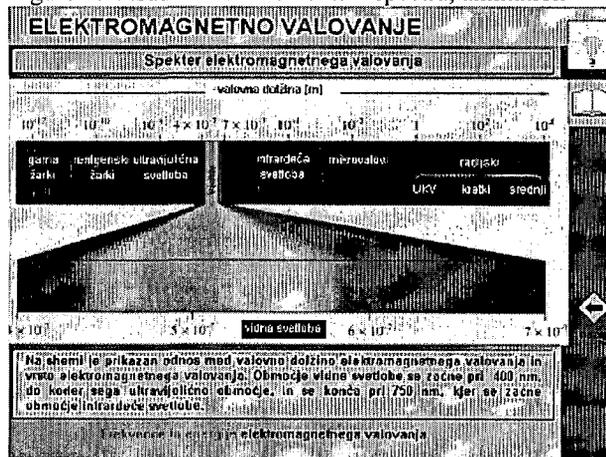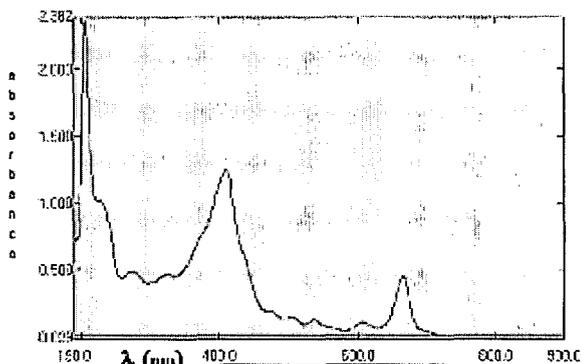
Results

The results presented in Figure 5 indicate that visual presentation of the chlorophyl spectrum and the written explanation about absorption on the CD ROM have contributed significantly to a better understanding of the phenomenon, and that students were able to apply their knowledge in solving the question about the absorption maximums of chlorophyl in the visible spectrum. As a result, 60% of students from the experimental group gave correct answers while only 28% from the control group answered correctly. There were only 10% wrong answers from the experimental group related to the question of the absorption maximum at 650 nm, while there were 30% wrong answers in the control group. In addition, 14% more students from the control group thought that chlorophyl would absorb light at 190 nm, (which is already in the UV range).

Figure 5: Final-test results for the first part of the task,

Legend: EG (=experimental group), CG(= control group)

Task 5
Visualisation elements – microscopic level
Absorption of light quantum, geometrical change of the molecule and explanation of the phenomenon (Figure 6, Figure 7, Figure 8)

Question
Select a combination of correct answers
a) *Upon interaction with light, the substance will be transformed into an electron-excited state, regardless of the wave-length or the energy of the light.*
b) *Light energy will be absorbed by outer electrons.*

c) *Molecules will remain in an excited state for an unlimited period of time.*
d) *After the transformation into the excited state the molecular geometry will change.*
e) *Compared with the initial state of the molecules, the reactivity of excited molecules will not change.*
f) *Interaction with the light can lead to homolytic cleavage between the atoms of the molecule.*
g) *Excited molecules of the substance can return into their initial state by emitting a photon (light quantum).*

Figure 6: Visualisation elements – microscopic level
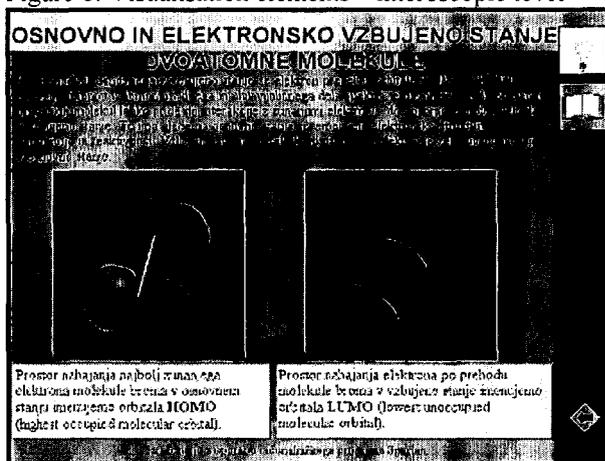
Figure 7: Visualisation elements – microscopic level
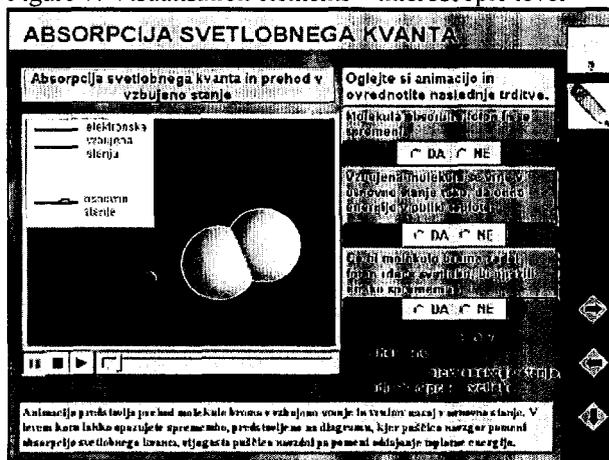
Figure 8: Visualisation elements – microscopic level

Figure 9: Percent distribution of the correct combination of answers (C) and incorrect answers (IC)   by EG and CG

Figure 10: Distribution of answers by EG and CG
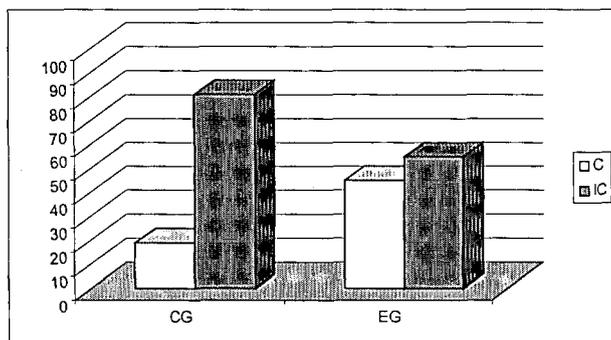
Results

The overall performance in solving Task 5, which was basically a rather difficult task, was rather poor for both groups.   However, there is a noticeable   difference between the two groups: 45% of students from the experimental group made a correct combination   of statements while the result of the control group was only 19% (Figure 9).

Correct combinations for Task 5 are: b, d, f and g (Figure 10). A more detailed analysis of the results shows that   the experimental group achieved better results since 20% more students selected (b) as a correct answer, while more than 35% of students from the control group   selected a wrong statement (a). This indicates that students who saw the simulation of light absorption using the CD ROM better understood the concept. The majority of students, regardless of the group, correctly answered that   molecules do not remain in the excited state for an unlimited period of time (statement c). The experimental group was again better in observing that the molecular geometry changes upon transformation into the   excited state   (statement d) which is indicated by a 35% better performance of the experimental group compared with the control group.
Similarly, the majority of students from both groups knew that the reactivity of molecules in the excited state is changed (statement e), however, there were 11% more students from the control group who knew that after the

interaction between light and matter a homolytic bond cleavage between the atoms may occur (statement f). It should be noted that in the multimedia teaching unit the concept of homolytic cleavage was not used. In answers related to the statement that   molecules from the excited state can return into their initial state by emitting a photon, the performance of both groups was   similar (statement g).

Task 6
Visualisation elements – animation of   the reaction mechanism in the bromation of cyclohexane (Figure 11, Figure 12)

Questions
Read the description and answer:
The bromation of cyclohexane proceeds according to the radical mechanism.
- *The first stage is the initiation of the reaction which runs due to the presence of light. Which of the molecules will absorb the light? Draw the reaction scheme for this stage.*
- *During the next stage the reaction expands. What processes will occur during this stage? Draw the reaction schemes.*
- *During the third stage the reaction is completed. Which processes run during this stage? Draw the reaction schemes.*

Figure 11: Visualisation elements – animation
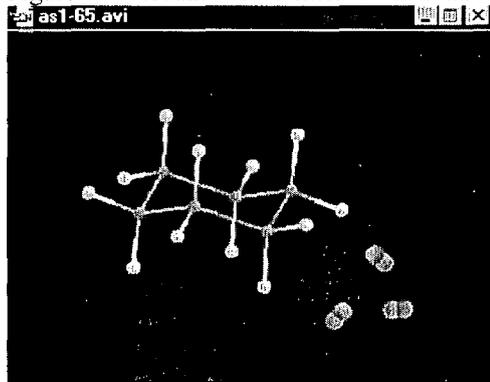
Figure 12: Visualisation elements – animation
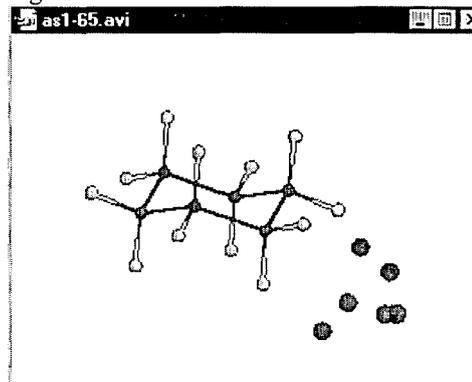
Results

Figure 13: Performance of students in drawing the reaction schemes
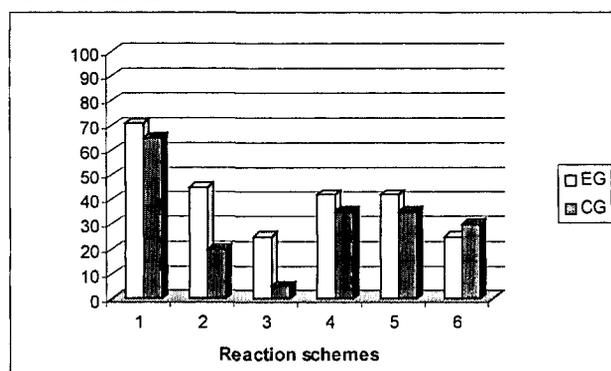
As can be seen from the graph in Figure 13, only 8% of the students from the experimental group were better at drawing the first reaction scheme. In drawing the second reaction scheme (expansion of the reaction) the performance of the experimental group was significantly better (25%) compared with the control group, while in the third part, i.e. drawing the reaction schemes for the completion of the reaction, the differences between the two groups become smaller again. However, in the control group there were 12% more students who did not even try to accomplish this task.

On average, the performance of the experimental group was 20.5% better than the control group. The experimental group was doing much better at solving both, difficult and easy tasks. It was also noted that the experimental group achieved much better results in the experimental test compared with the marks which the students achieved during regular chemistry courses.

We are aware that the results should not be over generalised, yet they do indicate that there is a positive impact on the motivation and quality of knowledge if various visual elements are integrated into chemistry teaching. Some differences should be noted, however: the evaluation of the task showed that some students found it difficult to use the CD ROM unit on their own, and expressed a desire that a teacher be present to direct them, particularly when dealing with the animation of microscopic processes.

4 Conclusions

Visualisation skills are very important for an easier understanding of abstract science concepts. The Internet offers manifold visualisation support and visual elements can be easily integrated into teaching, which can help students enormously when it comes to understanding difficult chemical concepts. The KemInfo website which has been used by chemistry teachers in Slovenia, has proved to be very supportive in preparing teaching units with integration of visualisation elements.

5 References

[1] Collin, R., Goll, L., 1993, Umetnost učenja, Tangram, Ljubljana.

[2] Vrtačnik, M., 1999, Vizualizacija v kemijskem izobraževanju, Kemija v šoli, 11,1,2-8.

[3] Baker, S.R., Talley, L.H., 1972, The relationship of visualization skills to achievement in freshman chemistry, Journal of Chemical Education, 49, 11, 775-777.

[4] Baker, S.R., Talley, L.H., 1974, Visualization skills as a component of aptitude for chemistry: a construction validation study, Journal of Research in Science Teaching, 11, 2, 95-97.

[5] Carter, C. S., La Russa, M. A., Bodner, G. M., 1987, A study of two measures of spatial ability as predictors of success in different levels of general chemistry, Journal of Research in Science Teaching, 24, 7, 645-657.

[6] Keen, E. N., Fredman, M., Rochford, K., 1988, Relationships between the academic attainments of medical students and their performance on a test requiring the visual synthesis of anatomical sections, South African Journal of Science, 84, 205-208.

[7] French, J. W., 1951, Description of aptitude and achievement in tests in terms of rotated factors, Psychometric Monographs, 5.

[8] Lohman, D. F., 1979, Spatial ability, A review and re-analysis of correlation literature, Technical Report 8, Stanford University, Stanford.

[9] McGree, M., 1979, Human spatial abilities, psychometric studies and environmental, genetic, hormonal, and neurological influences, Psychological Bulletin, 86, 889-918.

[10] Eliot, J., Hauptman, A., 1981, Different dimensions of spatial ability, Studies in Science Education, 8, 45-66.

[11] Seddon, G. M., Moore, R. G., 1986, An unexpected effect in use of models for teaching the visualization of rotation in molecular structures, European Journal of Science Education, 8, 1, 79-86.

[12] Tuckey, H. P., Selvaratnam, M., Bradley, J. D., 1991, Identification and rectification of student difficulties concerning three-dimensional structures, rotation and reflection, Journal of Chemical Education, 68, 6, 460-464.

[13] Saunderson, A., 1973, The effect of special training programme on spatial ability test performance, Journal of College Science Teaching, 5, 15-23.

[14] De Bono, E., 1976, Teaching thinking, Template Smith, London.

[15] Lord, T. R., 1985, Enhancing the visuo – spatial aptitude of students, Journal of Research in Science Teaching, 22, 5, 395-406.

[16] Rochford, K., 1987, Underachievement of spatially handicapped chemistry students in an academic support programme at U.C.T. in 1986, Paper presented to staff of the Department of Education, Cornell University, Ithaca.

[17] Tuckey, H. P., Selvaratnam, M., 1993, Studies involving three-dimensional visualization skills in chemistry: A review, Studies in Science Education, 21, 99-121.

[18] Sajovec, M., 1998, Evaluation of Application of Multimedia in Chemical Education, Master Thesis, University of Ljubljana, Faculty of Science and Engineering, Department of Chemical Education and Informatics, Ljubljana.

[19] Vrtačnik, M., Zupančič-Brouwer, N., Dolničar, D., Orel, M., 1997, Svetloba in kemijska sprememba (Light and Chemical Change): interactive multimedia teaching unit, Zavod R Slovenije za šolstvo, Ljubljana.

A Digital Watermarking Scheme Using Human Visual Effects

Chin-Chen Chang and Kuo-Feng Hwang
Department of Computer Science and Information Engineering,
National Chung Cheng University, Chaiyi, Taiwan 621, R. O. C.
Email: ccc, luda@cs.ccu.edu.tw
AND
Min-Shiang Hwang
Department of Information Management,
Chaoyang University of Technology, Wufeng, Taiwan, R.O.C.
Email: mshwang@mail.cyut.edu.tw

A watermarking technique for copyright protection of images is proposed in this paper. The idea is to use the human visual effects as the feature of images. The embedding strategy is to put the extracted image feature together with the watermark to generate the secret key for watermark retrieval. This strategy for watermarking system is wholly different from previous works. In addition, a voting approach is used to improve the correctness of the retrieved watermarks. The experimental results show that our method can against attacks by many image altering algorithms, such as filtering, lossy compression, rotation and so on. Furthermore, the proposed scheme is not only applicable to ordinary natural images but also to cartoon graphics.

1 Introduction

Digital watermarking is applicable in copyright protection. In recent years, many digital watermarking schemes have been proposed. To achieve copyright protection, digital watermarking must satisfy the following requirements.

1) Imperceptibility: the quality of the watermarked image must be very high.

2) Security: the security of watermarking must be hold whether possible attackers do or do not know how the watermark is embedded into the image.

3) Robustness: it must be possible to retrieve the watermark after various image processes, such as low-pass filtering, high-pass filtering, lossy compression, scaling, rotation and so on.

There is an issue on invisible digital watermarking. As pointed out by Craver et al. [6, 7], how to decide the rightful ownership of the invisible watermarking scheme is a problem. Craver et al. argued that a watermarked image could allow multiple claims of ownerships. To avoid such a thing from happening, a trusted third party should be introduced. Voatzis and Pitas [20] proposed a generic model for protecting copyrights. They have included a trusted registration authority in their model. Similarly, we use a time-stamping [2, 8, 9] mechanism to solve this problem in the proposed scheme.

Digital cartoon graphics have significant differences from ordinary natural images. Cartoon images to be without complicated colors and texture variations. This characteristic makes it difficult to embed watermarks into cartoon images. Moreover, cartoon images can be easily repainted using other colors without affecting the original intention. This is another challenge for robust digital watermarking techniques.

Conventionally, digital watermarking techniques can be classified into two categories. The first type includes those which embed the watermark into the spatial domain [13, 21, 19]. In general, this type of methods has a computing performance advantage. The second type embeds the watermark into the frequency domain [1, 5, 17]. This method transforms the original data into the frequency domain for watermark embedding. The Fourier transform, Discrete Cosine transform and Wavelet transform are applied usually. The proposed scheme embraces a different concept from those two categories [3, 4]. That is the original image remains not being altered even when the watermark is embedded. In Section 5, some advantages of our method will be discussed. For more information about digital watermarking, the interested readers may consult [14, 18, 22].

In this paper, we propose a new robust digital watermarking algorithm using the human visual effects. The human visual effect means the sensitivity of the human eye to a luminance under a background. We simply use the human visual effect to be the feature of images. This feature

is used to generate a secret key for the digital watermarking system.

The rest of this paper is organized as follows. Section 2 reviews the human visual model briefly. Section 3 demonstrates the details of our method. The experimental results of the proposed digital watermarking process are demonstrated in Section 4. Finally, Sections 5 and 6 present the discussions and the conclusions, respectively.

2 Human Visual Model

Kuo and Chen [11] proposed the human visual effect for DPCM (differential pulse code modulation). They considered the Weber's law [16] in their human visual model. In 1998 [12], they added some considerations and applied the human visual effect to the vector quantization (VQ) scheme. The human visual model constructs the contrast function in the gray-valued domain (from 0 to 255). The contrast function is used to evaluate the sensitivity of the human eye to a luminance under a background.

In [11], Kuo and Chen modelled the contrast function $C(x)$ from the combination of the bright background and dark background as follows. Note that the background β is the mean of the gray values in the background.

For the bright background ($\beta \geq 128$), $C(x)$ is defined as

$$C(x) = \begin{cases} ln\left(\frac{c_1(c_L - x)}{c_L(127.5 - (x - c_1))}\right), 0 \leq x \leq 128, \\ ln\left(\frac{(x - c_1)(x - c_H)}{c_1(255 - c_H)}\right), 128 \leq x \leq 255, \end{cases} \quad (1)$$

where c_1 is a constant and is equal to $127.5/2$, $c_L = \frac{128}{1 - e^{-k}}$, $c_H = \frac{128 - 255e^{-k}}{1 - e^{-k}}$. Here k is defined as $k = \frac{2.5}{1 + e^{(255 - \beta)/55}}$. Similarly, the contrast function $C(x)$ for the dark background ($\beta < 128$) is defined as

$$C(x) = \begin{cases} ln\left(\frac{c_1 \times c_L}{(127.5 - (x - c_1))(c_L - x)}\right), 0 \leq x \leq 128, \\ ln\left(\frac{(x - c_1)(255 - c_H)}{c_1(x - c_H)}\right), 128 \leq x \leq 255, \end{cases} \quad (2)$$

where c_1 is equal to $127.5/2$, $c_L = \frac{-128e^k}{1 - e^k}$, $c_H = \frac{255 - 128e^k}{1 - e^k}$. Here k is defined as $k = \frac{2.5}{1 + e^{\beta/25}}$.

In our proposed method, we use the contrast function to evaluate the sensitivity of a 5×5 sub-image. The sensitivity of the centric pixel x is measured via Equation (1) or Equation (2) according to β (mean of the 5×5 sub-image). The evaluated sensitivity stands for the feature of the 5×5 sub-image. We use this feature to generate a secret key for the digital watermarking system. Next section will describe the details of our method.

3 The Proposed Scheme

There are two primary characteristics of our scheme. First, the watermarked image is the same as the original image. Second, the watermarking secret key K is generated using the human visual effect, the watermark, and a seed of pseudo-random number generator. When the secret key K is generated, a signed time-stamp for K via a trusted third party (TTP) as K_s is required. The signed time-stamp K_s stands for the secret key K, which was generated at a certain time t. In other words, if K successfully retrieves the embedded watermark, then K_s represents the original image produced before or exactly at this time t. Consequently, K, t and K_s altogether make the evidence of identifying the rightful copyright owner of the watermarked image. We will depict the secret key generation algorithm in the first subsection. The second subsection describes the watermark detection algorithm. In addition, the voting strategy for improving the fidelity of the retrieved watermark will also be described.

3.1 Secret Key Generation

The original image O requires β bit(s) per pixel, where

$$O = \{o(i,j) \mid 0 \leq o(i,j) \leq 2^\beta - 1\}. \quad (3)$$

Here $0 \leq i \leq O_H - 1$, $0 \leq j \leq O_W - 1$. O_H and O_W are the height and width of the original image, respectively. Note that in our experiments, β equals 8. In other words, $o(i,j)$ ranges from 0 to 255. We define the watermark W as follows:

$$W = \{w(i,j) \mid w(i,j) = 0 \text{ or } 1\}. \quad (4)$$

Here $0 \leq i \leq W_H - 1$, $0 \leq j \leq W_W - 1$, and W_H is the height of the watermark and W_W is the width of the watermark.

Key generation algorithm:

Step 1: A set B, consisting of 5×5 sub-images, is constructed from O according to a pseudo-random number generator. The cardinality of B is equal to $W_W \times W_H$.

Step 2: The human visual effect set V of the centric pixel x of sub-image $b_{m,n}$, $b_{m,n} \in B$, is calculated via the contrast function $C(x)$.

Step 3: V is transformed by the transform function $Tran(\cdot)$ to a temporal key T. Here T is with the same data type (binary) as the watermark.

Step 4: The secret key $K = T \oplus W$, where $\oplus$ means the exclusive-OR (XOR) operator.

Next, we shall explain the steps of the above algorithm. Let the sub-image set $\mathcal{B}$ be defined as follows.

$$\mathcal{B} = \{b_{m,n} \mid m = 0, 1, \ldots, W_H - 1, \\ n = 0, 1, \ldots, W_W - 1\}, \tag{5}$$

where

$$b_{m,n} = \\ \begin{bmatrix} o(i,j) & o(i,j+1) & \cdots & o(i,j+4) \\ o(i+1,j) & o(i+1,j+1) & \cdots & o(i+1,j+4) \\ \vdots & \cdots & \ddots & \vdots \\ o(i+4,j) & \cdots & \cdots & o(i+4,j+4) \end{bmatrix}.$$

Here $0 \le i \le O_H - 5$ and $0 \le j \le O_W - 5$. Note that any sub-image $b_{m,n}$ allows mutual overlapping. Each sub-image $b_{m,n}$ is given from the coordinate (i,j) to the coordinate $(i + 4, j + 4)$ of the original image O. Coordinate (i,j) is determined by the pseudo-random number generator $Rand(\cdot)$ so that $(i,j) = Rand(s, m, n)$. Here s is the seed of the random number generator, and $m = 0, 1, \ldots, W_H - 1, n = 0, 1, \ldots, W_W - 1$. Since the seed s is a part of the secret key for watermark retrieval, a highly secret pseudo-random number generator is required. Hwang et al. [10] proposed a highly secure pseudo-random number generator $Rand(\cdot)$ using one-way hash functions. In our method, we use their pseudo-random number generator.

The human visual effect set **V** is defined as

$$\mathbf{V} = \{v_{m,n} \mid v_{m,n} \in \mathbf{R}\}, \tag{6}$$

where

$$v_{m,n} = C(x). \tag{7}$$

Here $m = 0, 1, \ldots, W_H - 1; n = 0, 1, \ldots, W_W - 1;$ W_H and W_W respectively are the height and width of the watermark; and x is the centric pixel of the sub-image $b_{m,n}$. The contrast function $C(x)$ has been defined in Equations (1) and (2). In Step 3, we transform $v_{m,n}$ to a temporary key T by a transform function $Tran(\cdot)$. T is defined as

$$T = \{t_{m,n} \mid t_{m,n} \text{ is 0 or 1}\}, \tag{8}$$

where

$$t_{m,n} = Tran(v_{m,n}). \tag{9}$$

Here $m = 0, 1, \ldots, W_H - 1$ and $n = 0, 1, \ldots, W_W - 1$. It is obvious that T has the same sizes and value domain as the watermark. Now, we define the transform function $Tran(\cdot)$ as follows.

$$Tran(v_{m,n}) = \begin{cases} 0, & \text{if } v_{m,n} \le A, \\ 1, & \text{if } v_{m,n} > A. \end{cases} \tag{10}$$

Here A is a threshold used to determine the output value of $Tran(\cdot)$. In our experiments, we set A to be the mean of all the visual effects $v_{m,n}$, for $m = 0, 1, \ldots, W_H - 1$ and $n = 0, 1, \ldots, W_W - 1$. Finally, in Step 4, the secret key K is constructed.

Signing time-stamp:
    When the secret key K is generated, the time-stamp K_s for K is signed by TTP using its private key as follows.

$$K_s = sig_{TTP}(h(O), t, K). \tag{11}$$

Here $h(\cdot)$ denotes a publicly known collision-free and one-way hash function [15], and t is the time while K_s is signing. The watermark embedding stage is finished when the time-stamp K_s is verified via the public key of TTP. For the later watermark detection stage, the image owner has to keep K, K_s and s secretly.

3.2 Watermark Detection Algorithm

The watermark detection algorithm is the same as the key generation algorithm except for Step 4. We can obtain the retrieved watermark W' from the temporary key T and the secret key K. As mentioned above, the pseudo-random number generator seed is a part of the secret key for watermark retrieval. When a right seed is introduced, the rightful temporary key T can be obtained. The retrieved watermark W' is obtained by

$$W' = T \oplus K. \tag{12}$$

The retrieved watermark W' is equal to W if the watermarked image is still the same old one. On the other hand, if the watermarked image has been modified, there may be a little difference between W' and W because some features of sub-images are altered.

When there is conflict between the right owner and someone else for the copyright, each one can verify the time-stamp K_s to examine if the secret key was indeed generated at certain time t via the public key of TTP. The real copyright owner is the one who registered to TTP earlier.

To improve the fidelity of the retrieved watermark, we introduce the voting strategy in the proposed scheme. For instance, a watermark is embedded α ($\alpha \ge 3$ and α is odd) times in one image. Every element of the retrieved watermark is decided by the majority. Let W'' denote the final retrieved watermark. The element $w''(m,n)$ of W'' is given by

$$w''(m,n) = \begin{cases} 1, \sum_{k=1}^{\alpha} w'_k(m,n) \ge \lceil \frac{\alpha}{2} \rceil, \\ 0, \sum_{k=1}^{\alpha} w'_k(m,n) < \lceil \frac{\alpha}{2} \rceil. \end{cases} \tag{13}$$

Here $m = 0, 1, \ldots, W_H - 1; n = 0, 1, \ldots, W_W - 1$; and $w'_k(m, n)$ denotes the kth retrieved element at the coordinate (m, n) by the detection algorithm mentioned above. The experimental results of the proposed algorithm will be shown in the next section.

4 Experimental Results

Figure 1(a) shows the original image of "Lena" (8 bits/pixel, 512×512), and the watermark in Figure 1(b) is the logo (64×64) of "National Chung Cheng University (CCU)". To evaluate the correctness of the retrieved watermark, we define an accuracy bit ratio R as below:

$$R = \frac{\sum_{i=0}^{W_H - 1} \sum_{j=0}^{W_W - 1} \overline{w(i,j) \oplus w'(i,j)}}{W_H \times W_W} 100\%. \qquad (14)$$

Here $w(i, j)$ is the element of original watermark, $w'(i, j)$ is the element of the retrieved watermark and $\oplus$ is the exclusive-OR operator.

First, the blurring algorithm and JPEG compression are applied to alter the watermarked image. The altered results are shown in Figures 2(a) and 2(c), respectively. The JPEG compression ratio is about 14:1 in the compressed image. Note that, following the voting approach, we embed the watermark three times ($\alpha = 3$) in the experiments. Figure 2(b) is the retrieved watermark of "CCU" from Figure 2(a). Figure 2(d) shows the retrieved watermark from Figure 2(c).

Next, the original image rotated one degree clockwise as shown in Figure 3(a). Besides, the image size is changed to 521×521. The other modified image using the sharpening algorithm is shown in Figure 3(c). Figures 3(b) and 3(d) are the retrieved watermarks from Figures 3(a) and 3(c), respectively. In particular, our method can directly retrieve the watermark from the rotated image. In other words, it is not necessary to turn back the size of the rotated image for watermark retrieval.

A cartoon graphic "Bunny" is used in the following experiment. Figure 4(a) shows the original "Bunny" graphic, and Figure 4(c) represents a repainted image of "Bunny". In Figure 4(c), both the face and background of "Bunny" are replaced with other similar gray levels. The retrieved watermark from the original image of "Bunny" is shown in Figure 4(b). Figure 4(d) shows the retrieved watermark from Figure 4(c).

Finally, the retrieved watermarks under different images are shown in Table 1. The experimental results show that the retrieved watermarks are recognizable under various attacks.

5 Discussions And Security Analyses

In the proposed scheme, the secret key length $|K|$ can vary according to the sizes of watermarks. However, the larger the watermark size is, the higher the security will be. Nevertheless, the memory space requirement also increases. Similarly, applying the voting strategy will also increase the memory space requirement for secret keys. The experimental results show that, if α equals 3, the retrieved watermark can be clearly recognizable. Consequently, the required storage space for the secret keys is practical. For example, let $\alpha = 3$ and the size of the watermark be 64×64, and then the extra space equals $64 \times 64 \times 3 = 12288$ bits (about 1.5 kb). In contrast to intellectual property loss, the cost of the extra space is worthwhile.

The main feature of the proposed algorithm is that the watermarked image remains the same as the original one. This characteristic makes the proposed scheme have no time limit for watermark embedding. In other words, the copyright can be traded and traced any number of times through the secret key K and its time-stamp K_s. Furthermore, the quality of the watermarked image is still as good as that of the original one.

The proposed scheme can avoid multiple claims of copyrights with the time-stamp. It is true that anyone can use our algorithm to construct the secret key for his/her images, even though the images are stolen from others. A pirate can also forge a time-stamp K'_s to make the signing time earlier than that of the original owner. After that, with the help of his/her secret key as well as K'_s, the pirate can steal the image successfully. Nevertheless, since the time-stamp is generated by a public-key cryptosystem, such as RSA, to forge an illegal time-stamp K'_s is as difficult as breaking a public-key cryptosystem. We assume that the rightful copyright owner always has the earliest signing time, which can be revealed in K_s. Consequently, using the time-stamp to identify the rightful copyright owner is very reasonable.

6 Conclusions

A new digital watermarking technique is proposed in this paper. In the proposed scheme, the human visual effect works as the feature of images for embedding invisible watermarks. The watermarked image is still the same as the original one. In addition, this method is robust against various attacks as shown in the experimental results. To avoid the multiple claims of copyrights, a time-stamping mechanism is introduced in this work. Furthermore, a voting approach is used to improve the fidelity of retrieved watermarks. The experimental results also show that the proposed scheme is not only applicable to ordinary natural images, but also to cartoon graphics.

References

[1] A. G. Bors and I. Pitas, "Image watermarking using DCT domain constraints," *Proceedings of 1996 IEEE International Conference on Image Processing (ICIP'96)*, vol. 3, pp. 231–234, 1996.

[2] A. Buldas, H. Lipmaa P. Laud, and J. Villemson, "Time-stamping with binary linking schemes," *Advances in Cryptology - CRYPTO'98*, pp. 486–501, 1998.

[3] C. C. Chang, K. F. Hwang, and M. S. Hwang, "A block based digital watermarks for copy protection of images," in *Proceedings of APCC/OECC'99*, pp. 977–980, Beijing, October 1999.

[4] C.C. Chang, T.S. Chen, and P.F. Chung, "A technique for copyright of digital images based upon (t, n)-threshold scheme," *to appear in Informatica*.

[5] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[6] S. Craver, N. Memon, B.-L. Yeo, and M. Yeung, "Can invisible watermarks resolve rightful ownership?," *Proc. SPIE Storage and Retrieval for Still Image and Video Databases V*, vol. SPIE 3022, pp. 310–321, 1997.

[7] S. Craver, N. Memon, B. L. Yeo, and M. M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 573–586, 1998.

[8] S. Haber and W. S. Stornetta, "How to time-stamping," *Journal of Cryptology*, vol. 3, no. 2, pp. 99–111, 1991.

[9] S. Haber and W. S. Stornetta, "Secure names for bit-strings," *Proceedings of 4th ACM Conference on Computer and Communications Security*, pp. 28–35, April 1997.

[10] M.S. Hwang, C.C. Chang, and K.F. Hwang, "A watermarking technique based on one-way hash functions," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 2, pp. 286–294, 1999.

[11] C. H. Kuo and C. F. Chen, "A prequantizer wiht the human visual effects for the DPCM," *Signal Processing: Image Communication*, vol. 8, pp. 433–442, July 1996.

[12] C. H. Kuo and C. F. Chen, "A vector quantization scheme using prequantizers of human visual effects," *Signal Processing: Image Communication*, vol. 12, pp. 13–21, March 1998.

[13] M. Kutter, F. Jordan, and F. Bossen, "Digital watermarking of color images using amplitude modulation," *Journal of Electronic Imaging*, vol. 7, no. 2, pp. 326–332, 1998.

[14] F.A.P. Petitcolas, R.J. Anderson, and M.G. Kuhn, "Information hiding - a survey," *Proceedings of the IEEE*, vol. 87, pp. 1062–1078, July 1999.

[15] B. Schneier, *Applied Cryptography*, WILEY, 2nd edition, 1996.

[16] Jr. T. G. Stockham, "Image processing in the context of a visual model," *Proceedings of the IEEE*, vol. 60, pp. 828–842, 1972.

[17] P.C. Su, C.C. J. Kuo, and H.-J. M. Wang, "Blind digital watermarking for cartoon and map images," *Proceedings of SPIE*, vol. 3657, January 1999.

[18] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of IEEE*, vol. 86, no. 6, pp. 1064–1087, 1998.

[19] G. Voyatzis and I. Pitas, "Embedding robust watermarks by chaotic mixing," *Proceedings of 13th International Conference on Digital Signal Processing(DSP'97)*, vol. 1, pp. 213–216, 1997.

[20] G. Voyatzis and I. Pitas, "Protecting digital-image copyrights: A framework," *IEEE Computer Graphics and Applications*, vol. 19, no. 1, pp. 18–24, 1999.

[21] G. Voyatzis and I. Pitas, "Chaotic mixing of digital images and applications to watermarking," *Proceedings of European Conference on Multimedia Applications, Services and Techniques (ECMAST '96)*, vol. 2, pp. 687–695, May 1996.

[22] Minerva M. Yeung, "Digital watermarking," *Communications of the ACM*, vol. 41, pp. 30–33, July 1998.

Table 1: The correcting ratio of the retrieved watermarks (CCU) from several attacks using various host images

Image		Blurring	JPEG	Sharpening	Rotating
Lena	PSNR(dB)	29.62	34.07	23.10	NA
	R(%)	99.40	99.73	99.73	95.12
Barbara	PSNR(dB)	23.41	31.29	17.15	NA
	R(%)	98.05	98.56	95.65	90.60
Airplane	PSNR(dB)	26.38	31.65	21.18	NA
	R(%)	98.32	99.02	97.31	90.45
Bunny	PSNR(dB)	16.37	28.67	21.35	NA
	R(%)	98.71	99.90	99.37	97.68

(a)   (b)
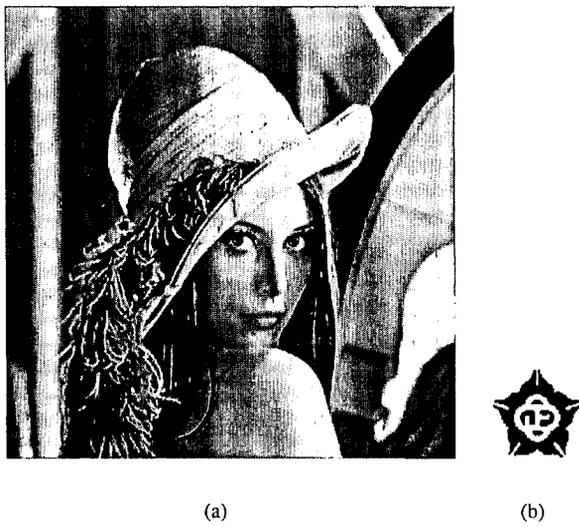
Figure 1: (a)Original image of "Lena" (512×512), (b)Watermark of "National Chung Cheng University (CCU)" (64 × 64)

(a) PSNR=29.62dB   (b) R=99.40%

(c) PSNR=34.07dB   (d) R=99.73%

Figure 2: (a)Blurred image of "Lena", (b)A retrieved watermark from Figure 2(a), (c)Reconstructed JPEG image, (d)A retrieved watermark from Figure 2(c)

(a) Resized to 521×521

(b) R=95.12%

(a) 256×256

(b) R=100.0%

(c) PSNR=23.10dB

(d) R=97.73%

(c) PSNR=18.68dB

(d) R=92.82%
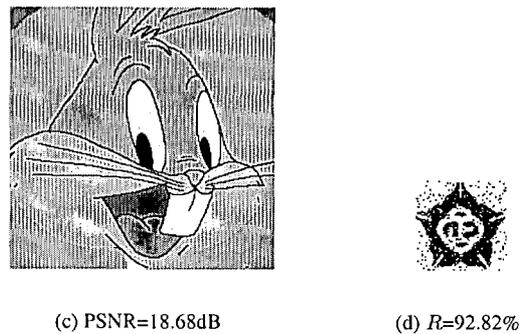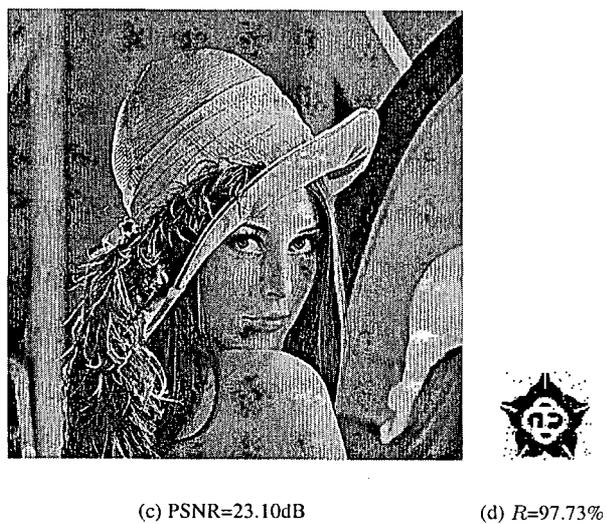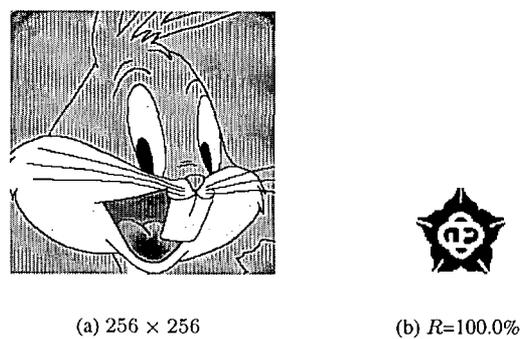
Figure 4: (a)Original graphic of "Bunny", (b)A retrieved watermark from Figure 4(a), (c)Repainted graphic of "Bunny", (d)A retrieved watermark from Figure 4(c)

Figure 3: (a)Rotated one degree image of "Lena", (b)A retrieved watermark from Figure 3(a), (c)Sharpened image of "Lena", (d)A retrieved watermark from Figure 3(c)

Business Process Reengineering and Information Systems Renovation Projects: Problems and Assessment

Andrej Kovačič
University of Ljubljana, Faculty of Economics
Kardeljeva pl. 17, SI-1000 Ljubljana
Tel: +386 61 1892 400
E-mail: andrej.kovacic@uni-lj.si

The following paper points out problems encountered in several business process reengineering (BPR) projects in Slovenia in the last five years and provides an assessment of BPR's strategic impact on the quality of information systems (IS). Based on his own experience and his knowledge of the opportunities offered by the BPR concept and the use of new information technology (IT), the author assesses of the strategic relevance of reengineered processes and consideres the impact of BPR projects on the quality of an organization's information systems management.

1 Introduction

Technical progress together with the opening of a global market are definitely among the primary factors playing roles in modern society. One of the consequences of recent development in the fields of information technology, communications, and transportation is an ongoing process of standardization in both the business and government arenas. In national economies, the global characteristics are becoming localized, and vice versa the local ones are becoming globalized. The modern global society is becoming more and more universal.

Only businesses that are able to adjust their mass production and generalized marketing approaches to more dynamically serve clear market niches will survive in the future. By individualizing and personalizing their products or services, businesses can adjust to the needs of their customers. Producing for a known customer and satisfying him as much as possible are the goals of each progressive organization. The best way to achive such a goal is thrugh reorganized business processes.

Organizations react very differently to projects or attempts of renovation of business processes, though the purpose is clear: reduction of costs, shortening the business cycle, and improvement of quality. The difficulties in the public sector are larger than those in the private sector. The increased employment in the public sector during the past few years has further entrenched bureaucracies; the problems of efficiency are then most often solved through purchasing computer hardware and software. Moreover, if managers feel the corporation they work for is successful at the present time, they usually reject the idea of renovating the business. Of course, when a company faces trouble,

there are never seem to be enough financial or human resources to start such a project.

Coping with these problems while working on BPR and IS renovation projects in the last few years, we have noted a strong correlation between improvement of overall corporate culture and strategies, and the quality of IS within an organization. This paper explores these relationships.

2 BPR and IS renovation

We cannot forget the role that information technology plays in business process renovation, as we must also keep in mind that an incorrect use of information technology can deliver partial solutions which do not consider the system a whole and are by all means unsatisfactory. Generally speaking, we should use a two-step approach by which the described problems can be eliminated. First, we should renovate existing processes with respect to the necessary characteristics of business; and second, we should provide a proper organization and informatics to support the new processes.

One of the main ways of connecting the difficulties of our enterprises is the implementation of the Business Process Reengineering (BPR) method. BPR is a new method of improving the operation and therefore the outputs of organizations (private and public). It means analyzing and altering the business processes of the organization as a whole. For a thorough and effective reengineering project, organizations should first meet certain conditions before starting such a project. First, the management should abandon all the obsolete rules and procedures that have been used up to that time. In addition they should abandon other inadequate organizational and production principles. At this point, the

design of a renovated and redesigned organization should start.

BPR was first introduced in a research program at MIT (Massachusetts Institute of Technology) in the early nineties. The term was used in the description of Davenport and Short's 1990 research project. They found out that the implementation of modern information technology in organizations means not only automation of managerial and production tasks but also has an enormous and direct effect on the means and quality of the work done. This point was further discussed and developed by Hammer in his "Reengineering Work: Don't Automate, Obliterate", Harvard Business Review, 1990.

In Slovenia, there have been several recent symposiums on the subject of business process redesign and IS renovation. Numerous papers can be found about this topic in publications that cover this area. In everyday practice the first attempts to implement techniques mentioned above usually employed modern CASE (Computer Aided Software Engineering) tools which unfortunately more or less merely automate existing processes. In the worst cases the primary objective of the renovation projects was to buy new equipment (mostly PCs). Slovenian experts follow the current trends and try to integrate them in our projects of informatization and business reorganization. However, most (approximately 90of current projects of informatization are still oriented to automation of existing business procedures and activities.

On the other hand, organizational projects are mostly oriented to ISO 9000 standardization (documentation), so there is no room for important simplification and organizational changes. There were also some attempts to move companies from rigid hierarchical structures to flat, dynamic structures that can rapidly regroup in response to changing requirements. In some cases traditional departments have been transformed into ad hoc teams or "centres of excellence".

3 Projects and open problems

The best way to understand the situation in the field of business process reengineering and informatization is to analyse open problems and situations in some typical projects. We empirically analyzed the most recent 5 projects on which we have worked. The analysis shows various results and implementation problems. We think that some of these problems are specific for the Slovenian business culture and the others are not. For further analysis and assessment we grouped the problems and possible solutions by strategic/business cultural aspect and from the strategic IS impact point of view.

3.1 Projects

The author of this paper has worked or participated as senior consultant or project manager on more than 25 consulting and engineering projects in different industry

branches and in the public sector. He has developed a methodology of business renovation and informatization. The methodology has been successfully implemented in different organizations. The projects selected for further analysis in this paper are listed below:

Project A

Initial goal: Shortening of business cycle
Type of business: Metalwork production
Duration: 1,5 years
Main results: Time and cost reductions, business process model, proposed IS renovation

Project B

Initial goal: BPR and IS renovation
Type of business: Paper industry
Duration: 2 years
Main results: Reengineering of logistics, new IS developed and implemented

Project C

Initial goal: IS auditing and assessment (starting project: Process assessment and renovation)
Type of business: Insurance
Duration: 2,5 years
Main results: Modeling and standardization of key business processes, strategic IS planning

Project D

Initial goal: IS development
Type of business: Energy supplier (electricity)
Duration: 4 years
Main results: Reengineering of key business processes, new IS being developed and implemented

Project E

Initial goal: Management consulting
Type of business: Construction & housing
Duration: 2 years (still in progress)
Main results: New business model, proposed organizational changes, new IS developed

We can see that particular starting points or goals of the listed projects most often lead to different results. But if we look at the results more carefully we can group them into BPR and IS oriented steps. In the BPR step strategic problems from the organizational culture point of view should be solved. Any unsolved problems in the BPR step have a direct (and negative) impact on the quality of results from the IS renovation or IS development project step.

3.2 BPR strategic problems

Organizational strategic problems come mostly from aspects related to the way the firms operate, the scope of BPR projects, and strategic and operational modeling and benchmarking.

The way the firms operate. Most Slovenian experts subscribe to the traditional organization theory beginning from Adam Smith and followers. They see organizational structures and work as a collection of tasks, each of which can be described by a precise procedure whose steps are particular motions and activities of a worker. Problems related to corporate costs, quality, and the business cycle are the result of a cross-functional dispute where nobody has total control of the whole process. These input-process-output models or models of individual specialization are no longer sufficient, however. They do not represent human commitments and the business processes view of an organization in which people request work and agree on what will be done, who will do it, and when it will be done (Denning and Medina-Mora, 1995).

We see the workflow or the business loop as a closed cycle in which the fulfiller (performer) completes action leading to the satisfaction of a request by the requester (customer). The business loop begins with preparation and formulation of a customer's request and culminates in a delivery (proposal) to the performer. The second step is negotiation, culminating in an agreement between customer and performer. The third step is performance, resulting in the requested product or service and the performer's declaration that the work is done. The final step is acceptance, culminating in the customer's declaration of satisfaction. Figure 1 shows the workflow loop adopted by Action Technologies and Lei Yu (Yu, 1996).

On our BPR projects we find various kinds of problems we must solve:

- The transition from the traditional, functional, and departmental way of work on organization focused on business processes is very difficult to achieve. We cannot literally say that "People who work in different functional areas hate each other" (Palermo's law, see: Watson, 1994 p. 60), but nonetheless we found a strong resistance to a cross-functional way of work.

- We see that the way the firms operate cannot be changed without some kind of redesign of its business processes. Such redesign can include rationalization, simplification, and standardization of the procedures, as well as the introduction of important organizational changes and modern information technology. We should also indicate that the information infrastructure enables the growth of the firm.

- Many BPR projects lack the fourth, satisfaction step of the main workflow loop (see Figure 1). This should in practice trigger a new set-up of missing workflow activities.

Scope of BPR projects. A "business process" is defined as a system of logically united executive and control procedures and activities that end in the desired result, product, or service. The efficiency of a business process is measured through the outputs we get from the transformation of the input resources. The resources are often measured by the time consumed or the costs of the process. Effectiveness of a business process can be defined as the compatibility of the given results with the results that we desired from the process.

BPR projects can differ depending on their scope. The scope of a project increases as more processes, people, and information are included in the project. On the basis of our experience we must take into account the following defining criteria of scope of the project: the questions of **process automation, business efficiency, and business effectiveness**. The meaning of the word "efficient" means "to do things in a correct way". It is an internal characteristic of the company's ability to do what it is supposed to do in a limited period of time. It is a function of availability, adaptability, and capacity. It is often measured with some kind of output/input ratio. The term "effectiveness" means "to do the right things". By its nature it is external to the firm, measured by the rate of satisfaction of all kind of needs.

Process automation projects are normally influenced by modern information technology and focus on one process, usually within the department. They do not radically change any existing procedures but merely automate existing procedures. In such cases (applying technology) the focus of the reengineering project is narrow; only the limited procedures of the particular process will be changed.

By contrast, **business process redesign** projects focus on the underlying business reasons why certain processes exist in their current form. Reengineering questions are related to efficiency and correctness of certain processes under consideration.

It can be stated without any doubt that the corporate-wide **business process reengineering** projects cannot be seen only as a problem of automation and implementation of modern information technology or efficiency improvement of existing processes. Such a project begins by asking the questions related to business effectiveness, "Why are we even doing this?" (Kubeck, 1995). The project focuses on the underlying vision of the corporation, drastically changing corporate culture and realigning corporate goals, changing organizational structure skills and processes.

The first one to consider this issue was Leavitt (Leavitt, 1965) who noticed that there is more to organizational changes than just the technological view. To ensure optimal results, we must define all correlated key business elements: **structures, people** and other resources that enable the business processes and help us meet objectives, the **business processes**, and finally the **technology**. The structural view is more or less the organization of the corporation. He presented a diagram (Leavitt's diamond) that exposes the need to recognise the connection of business process redesign with all other factors that form the socio-technical frame of the organization. If we add the factor of corporate culture, the diamond looks as presented on the following diagram (Figure 2) (Adopted from: Galliers, 1995, p. 124):

The culture is one of the key factors of business renova-
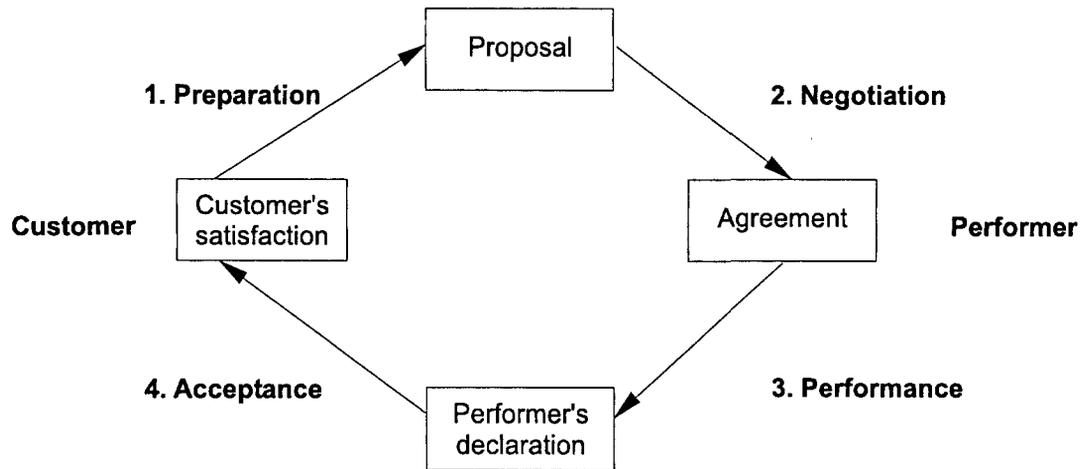
Figure 1: The Action Workflow loop
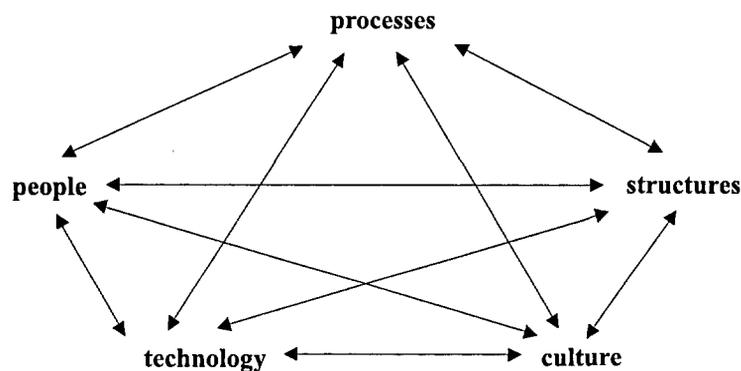
Figure 2: Amended Leavitt's diamond

tion. The appropriate business culture is required to enable the changes in the firm and in the whole society. Without the recognition of culture as an important factor, it is impossible to predict the chances of success of the project, as it is hard to form the objectives and strategy of the renovation project. Because of its complexity, the culture should be considered from the individual, corporate, and society point of view. From the corporate point of view, the scope of the project depends mostly on the particular organizational level involved in the project and its interrelationships. In general, the levels are Enterprise (strategic), Business (tactical), and Operations (operational) (Watson 1994, p. 47).

In our experience, the relatively slow process of privatization of the Slovenian economy has caused a weak "voice of ownership", and the impact of shareholders on changing the corporate culture toward business renovation has been limited. Most corporations that have existed for more than 10 years have retained hierarchical and "self management" features in their corporate culture. The enterprise and business levels are weak compared to the operations level. The results in terms of renovation have been automa-

tion and information technology projects that have embedded ineffective or inefficient work processes. The number of successful corporate-wide BPR projects has been low and the successful ones have been primarily in "private" corporations. In Slovenia and in other developing countries of Central Europe experts also have recognized that transforming opportunities into business success requires managers with insight, flexibility, and decisiveness (see: Janson and Wrycza, 1996).

We consider the change from a functional to a process type of organization to be of the utmost importance in the business renovation projects we have worked on. Our single unsuccessful project failed in the moment of implementation because agreed upon organizational changes were not carried through. We expect to overcome these types of problems through working with people, not only raising their availability, flexibility, or productivity but also improving their knowledge, managing their natural resistance to change, and helping to convert that resistance into commitment. It is also easier to deal with people that have better education and are prepared and skilled to use new information technology. Progressive organizations build on

the potential of their employees by ensuring that they know more, work better, and do more. So the employees should meet certain conditions to be able to perform: They should be skilled and well- informed, with the possibility of further education when needed, they should use right tools and technology, and they should be personally stimulated and motivated.

Strategic and operational modeling and benchmarking. Business modeling should be divided into strategic and tactical/operational levels (Krallmann and Derszteler, 1996). Strategic modeling includes the analysis of corporate strengths, weaknesses, and culture, the assessment of information systems in the organization, and organization and management competencies. It is the basis of all further actions and is carried out by corporate management. Corporate goals, strategies, and critical success factors form the basis for selecting and modeling core business processes on the global level of description. Such a model together with information on the organization's current state is fundamental for evaluating and benchmarking compared to other corporations. Benchmarking is a business practice that leads "reengineers" towards implementing strategic change initiatives in key business processes (Watson, author of strategic benchmarking). The benchmarking effort focuses on measuring other corporation processes to determine where excellence of performance exists, learning what those corporations did, and then producting excellence in those processes.

Detailed modeling of the processes or workflow structures takes place on the tactical/operational level. Workflows are refined and modeled at the level of particular interdependent business activities that are performed by actors (subjects) in an organization in order to achieve common goals. On this level, the more exact and certain information about workflow is, the better the modeling results will be. The problem lies in the conflict of aims between the need for accurate information and the difficulties of obtaining it due to the often obsolete documents describing flow structure, varying or even contradictory statements from employees, and time constraints (Krallman and Derszteler, 1996). On this level of detail, the benchmarks in the field of workflow modeling can be called reference process models. Reference process models are developed and evaluated on the basis of best practice for several selected types of business.

When working on our projects we discovered several problems related to business process modeling and benchmarking. Many Slovenian corporations are not willing to put their energy and time into business process modeling and to invest in benchmarks on the strategic level. They want quick and tangible results like many other companies from all over the world. Owing to a lack of national benchmarks, we also found that some international benchmarks would be suitable for problems of our corporations. Result: the number of corporate-wide BPR projects is trifling in Slovenia. We have recognised that business process modeling "from scratch", based on several decomposition

iterations from the strategic level to the appropriate level of detail of the process is a time consuming and risky job. We can say that our work relies mostly on our intuition and experience.

The situation on the tactical/operational level is not much better. Only a few Slovenian software houses produce application software solutions based on their own reference process models. On the operation level of modeling and benchmarking we have noticed a strong impact of ARIS designed EPC (Event-driven Process Chain) models (mostly produced for SAP R/3 application packages). But we recognized that many of the models are not appropriate for our corporations' way of work.

3.3   The actual role of IT

Information technology plays the key role in business process renovation. We should point out again that mere automation of procedures brings more or less negative results. Even if some of the achievements of such actions are positive, they prevent us from seeing all the opportunities that are offered by the informatization of a redesigned business process and an infrastructural role of informatics. That is why we first have to analyze the business process in order to find out if it is well-defined, adequate, and ready for the implementation of new information technology. Only if we do all this correctly can we expect an improvement of quality, lower costs, and shorter performance times of renovated business procedures and activities.

Our experience on BPR projects shows that the informatization of business processes should cause several changes and enhancements such as:

- data entry and concurrent control is performed only once in the process, at the beginning, so the users in later stages only update these data,

- data entry and control is performed by the user, who is responsible and able to have his/her part of the activities of the process under control,

- wherever possible, the workflow activities are triggered, performed, and controlled automatically,

- the focus of managerial accounting is no longer on ensuring the quality of information, but it moves towards its controlling function.

We can find the "right moment" or starting point for the implementation of modern information technology on BPR projects, if we borrow some experience from Total Quality Management (TQM) theory. Figure 3 (adopted from Watson, 1994) shows a model of a natural process improvement sequence that occurs as the corporation applies TQM to its work processes.

Watson sees this model as a road map for process improvement and application of basic quality tools and quality improvement processes before there is a need to automate work processes or seek IT solutions. We can add that

Level of Improvement
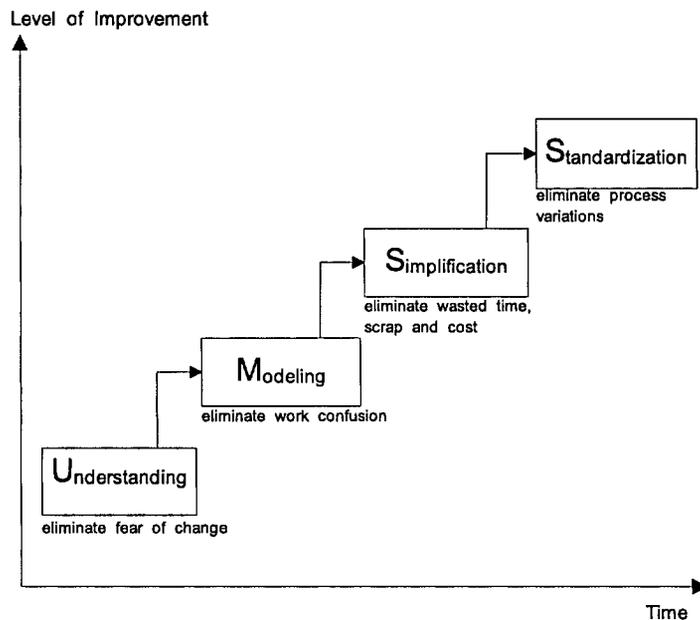
Figure 3: Process improvement sequence

IT has the strongest impact on standardization or elimination of process variations. We recognize that informatization cannot start before the first three levels of work process improvement have been successfully implemented.

4 BPR and IS project assessment

This section attempts to answer the question: What role and impact might BPR play in changing the quality of an organizational IS? In order to answer this question in our research we must first evaluate or assess the projects listed in the previous section from the following points of view: **strategic relevance and progress of the reengineered processes** and **improvement of the quality of the organization's IS management**. Although the terms evaluation and assessment are often used synonymously in regard to BPR and IS projects, there are slightly different connotations associated with each word. Evaluation is the act of placing an exact value on the object. In the context of BPR and quality of IS the ideal evaluation is not possible. Rather than evaluating, our aim in this paper is to derive a set of measures for assessing the strategic results of BPR projects and the quality of IS. For both parts of further analysis we will assign the following values to the answers given:

$$
\begin{aligned}
\text{none; ignore} &= N \\
\text{very low} &= 1 \\
\text{low} &= 2 \\
\text{moderate} &= 3 \\
\text{high} &= 4 \\
\text{very high} &= 5
\end{aligned}
$$

4.1 Strategic relevance and progress assessment

While assessing the impact and results of a BPR project in the light of organizational efficiency and effectiveness, the evaluator must address several critical questions (more about strategic assessment in: Revenaugh, 1995). The strategic relevance of the reengineered process should be assessed through:

1. Cost displacement or cost reduction (COSTS)

2. Development and offering of new or improved products or services (PRODUCTS)

3. Development of new administrative control and planning processes (CONTROL)

4. Offering significant tangible benefits (e.g., business cycle reduction, inventory reduction) (BENEFITS)

5. Offering new ways of competing and customer-supplier relations (COMPETITION)

6. Obtaining organizational changes (shift toward process organization) (ORGANIZATION)

7. Work quality improvement, focusing on the organizations key measures and possible outsourcing (QUALITY)

4.2 IS assessment

Assessing the information system function within organizations has been identified as one of the most critical issues of information systems management (Dickson et al.,
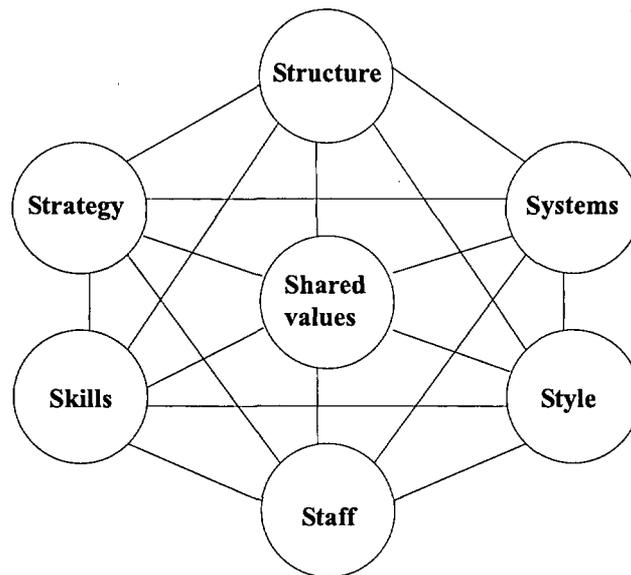
Figure 4: The McKinsey model

1986). We believe that a more comprehensive assessment can be achieved through a model provided by Galliers (1995). This model combines the wellknown Nolan's stages of growth with the so-called '7S' analysis pioneered by McKinsey and Co. (Morris and Haigh, 1998, p. 381) (4).

The premise behind the McKinsey model is that to function effectively organizations have to rely on the interdependence of the seven variables: **strategy** - the plan leading to the allocation of resources; **structure** - the organizational chart showing lines of authority and responsibility; **systems** - procedures, guidelines, and control mechanisms; **staff** - people employed; **style** - the management style of the organization; **skills** - the strengths and capabilities of employees; **shared values** - the goals of all employees (Morris and Haigh, 1998, p. 381).

Galliers model is illustrated in outline in Table 1. It is best used in a workshop environment where key stockholders can debate their different perceptions about the current state of affairs becouse there will almost certainly be different perceptions expressed. Certain elements of the 7S framework will almost always be at different stages of growth. The resultant profile will help to identify where particular urgent action is required (Galliers, 1995).

5 Results and concluding remarks

Comparing the current corporate state in terms of reengineered business processes and IS management to the state that existed when starting a particular project (listed in section 3.1) gives an indication or an assessment of the rate of progress that has been made.

a. Assessment of the strategic relevance of reengineered process. To recognize where the projects are on

the development continuum we employ transition analysis. The progress (or changes) that has been obtained on the analyzed projects is assessed in Table 2. The first value of each result provides the starting position or state of the assessed strategic element whereas, the second value presents the obtained status (values are presented in section 4). Wherever possible, the strategic relevance parameters of the previous and attained states are compared or benchmarked against an industry average and not against the most aggressive competitor. These assessments are particularly appropriate in the field of cost reduction (COSTS), business cycle or inventory reduction (BENEFITS), shifting toward a process organization (ORGANIZATION), and quality improvement (QUALITY). The previous state of the other parameters are in most cases estimated by the researcher, with improvements evaluated by comparing to the planned values.

b. Changes in the quality of corporate IS management (assessment has been made by the model proposed in the previous section). Similar to Table 2, results are presented with starting and final values, so the comparison and the progress can be seen. Table 3 summarizes these results.

Comparing the final results presented in Tables 2 and 3, we can find both expected and unexpected results and relationships:

- The analyzed BPR projects have no negative impact neither on the BPR strategic relevance nor on the IS quality improvement parameters.

- The analyzed BPR projects show a significant impact especially in the development of new organizational control and planning processes and the shift toward a process organization; some positive changes are also obtained in cost displacement or cost reduction and work quality improvement.

STAGE ELEMENT Values:	I N	II 1	III 2	IV 3	V 4	VI 5
Strategy	acquisition of IT	audit of IT provision	top-down analysis	integration coordination	strategy linkage	interactive planning collaboration
Structure	informal	finance controlled	centralized DP department	information centre(s)	departmental coalition	coordinated coalitions
Systems	ad hoc, operational accounting	gaps/ duplication: large backlog: heavy maintainance	uncontrolled end-user comp. vs. centralized systems	decentralized approach, some EIS	coordinated centralized and decentralized IS some strategic IS	inter-organizational systems: IS/IT-based products & services
Staff	programmers, contractors	system analysts, DP manager	IS planners, IS manager, DB specialists	business analysts, information resource manager	business & IS planners integrated	IS/IT director (board level)
Style	unaware	"don't bother me (I'm too busy)"	abrogation, delegation	partnership, benefits management	individualistic (product champion)	multidisciplinary teams (key themes)
Skills	individual: technical low level	system development methodology: cost-benefit analysis	IS awareness product management	IS business awareness	entrepreneurial marketing	lateral thinking (IS/IT potential)
Shared values	obfuscation	confusion	senior management concern, DP defense	cooperation	opportunistic	strategy making & implementation

Table 1: Amended Galliers' model: Assessing the quality of an organization's IS management

PROJECT:	A	B	C	D	E
Costs	2 ⇒ 3	2	1 ⇒ 2	1 ⇒ 3	2 ⇒ 3
Products	1 ⇒ 2	N	2 ⇒ 4	N	N
Control	2 ⇒ 3	1 ⇒ 3	3 ⇒ 4	2 ⇒ 3	1 ⇒ 4
Benefits	3 ⇒ 4	3	2 ⇒ 3	1 ⇒ 2	3 ⇒ 4
Competition	1 ⇒ 2	3	2 ⇒ 4	N	3
Organization	N ⇒ 1	1 ⇒ 3	1 ⇒ 3	1 ⇒ 2	2 ⇒ 5
Quality	2 ⇒ 3	1 ⇒ 2	1 ⇒ 3	2 ⇒ 3	2 ⇒ 4

Table 2: Results of the progress on key performances

- The impact on the quality of IS is quite similar across projects, though the projects had different objectives and duration and were performed in differend types of businesses.

- The progress of BPR project has (in average) a stronger impact on the quality of IS management than on the key BPR performance elements (Figure 5).

- IT enabled BPR projects to have a strong impact on the quality of IS management only from particular aspects. These aspects are: strategic position of IS management, organization of IS organizational unit (e.g., DP department, Information Centre) and systems concerned (dealing) with new planning and control procedures (Figure 5).

We discovered a strong two-way correlation between BPR and IS project activities, leaving aside the question 'what is cause and what are consequences'. Recent BPR research and papers state that information technology represents the key role in business process renovation. But in our research we also point out an impact in an opposite direction, toward the quality of IS management.

References

[1] Denning, P.J., and Medina-Mora, R., (1995), Case Study: George Mason University, In : Fisher, L. (ed.), The Workflow Paradigm, Future Strategies Inc., Florida

[2] Dickson, G.W. et al., (1986), Toward a Derived Set of Measures for Assessing IS Organizations, In: Information Systems Assessment, IFIP WG 8.2 Working Conference, Noordwijkerhout

[3] Galliers, R.D., (1995), IT and Organizational Change: Where does BPR Fit In?, In: Burke, G., and Peppard, J., Examining Business Process Re-engineering, Kogan Page Limited, London

[4] Hammer, M., (1990), "Re-engineering Work: Don't Automate, Obliterate", Harvard Business Review, 68(4), July-August, 104-112

[5] Janson, M.A., and Wrycza, S., (1996), IT as an Enabler of Business Process Designing During Macroeconomic Transformation, In: Scholz-Reiter, B., and Stickel, E., Business Process Modelling, Springer, Berlin

[6] Kovačič, A., (1996), Business Process Reengineering: Research Project Directions, International Con-

PROJECT:	A	B	C	D	E
Strategy	2 ⇒ 4	2 ⇒ 4	1 ⇒ 4	N ⇒ 3	1 ⇒ 4
Structure	2 ⇒ 3	2 ⇒ 3	2 ⇒ 4	2 ⇒ 4	2 ⇒ 3
Systems	1 ⇒ 3	1 ⇒ 3	2 ⇒ 4	2 ⇒ 4	2 ⇒ 4
Staff	1 ⇒ 2	1 ⇒ 2	1 ⇒ 3	1 ⇒ 2	2 ⇒ 4
Style	1 ⇒ 2	1 ⇒ 2	2	1 ⇒ 2	1 ⇒ 2
Skills	1	1 ⇒ 2	1 ⇒ 2	1	1
Shared values	1 ⇒ 2	1 ⇒ 2	1 ⇒ 2	1 ⇒ 2	1 ⇒ 3

Table 3: Progress on the quality of organization's IS management
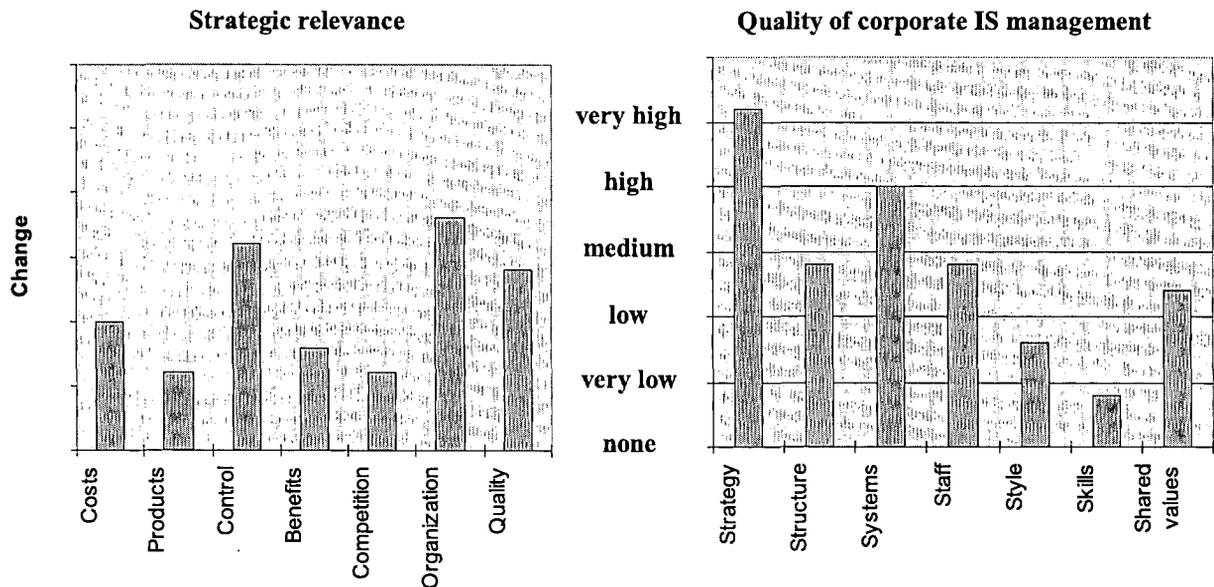
Figure 5: Impact of BPR projects on strategic relevance parameters and quality of IS management compared

ference, Summary of Abstracts, University of Ljubljana, Faculty of Economics

[7] Krallmann, H., and Derszteler, G., (1996), Workflow Management Cycle, In: Scholz-Reiter, B., and Stickel, E., Business Process Modelling, Springer, Berlin

[8] Kubeck, L.C., (1995), Techniques for Business Process Redesign, John Wiley & Sons, New York

[9] Leavitt, H.J., (1965), 'Applying Organizational Changes in Industry', In: March, J.G., (ed.), Handbook of Organizations, Rand McNally, Chicago

[10] Morris, D., and Haigh, B., (1998), From Case Study to Paradigm:From Information to Explanation, In: Montanheiro, L. at al., Public and Private sector Partnerships: Fostering Enterprise, Sheffield Hallam University Press

[11] Revenaugh, L.D., (1995), Crossing the Final Hurdle in Business Process Re-engineering, In: Burke,

G and Peppard, J, Examining Business Process Reengineering, Kogan Page Limited, London

[12] Yu, L., (1996), A Coordination-based Approach for Modelling Office Workflow, In: Scholz-Reiter, B., and Stickel, E., Business Process Modelling, Springer, Berlin

[13] Watson, G.H., (1994), Business Systems Engineering, John Wiley & Sons, New York

THE MINISTRY OF SCIENCE AND TECHNOLOGY OF THE REPUBLIC OF SLOVENIA

Address: Trg OF 13, 1000 Ljubljana,
Tel.: +386 61 178 46 00, Fax: +386 61 178 47 19.
http://www.mzt.si, e-mail: info@mzt.si
Minister: Lojze Marinček, Ph.D.

Slovenia realises that that its intellectual potential and all activities connected with its beautiful country are the basis for its future development. Therefore, the country has to give priority to the development of knowledge in all fields. The Slovenian government uses a variety of instruments to encourage scientific research and technological development and to transfer the results of research and development to the economy and other parts of society.

The Ministry of Science and Technology is responsible, in co-operation with other ministries, for most public programmes in the fields of science and technology. Within the Ministry of Science and Technology the following offices also operate:

Slovenian Intellectual Property Office (SIPO) is in charge of industrial property, including the protection of patents, industrial designs, trademarks, copyright and related rights, and the collective administration of authorship. The Office began operating in 1992 - after the Slovenian Law on Industrial Property was passed.

The Standards and Metrology Institute of the Republic of Slovenia (SMIS) By establishing and managing the systems of metrology, standardisation, conformity assessment, and the Slovenian Award for Business Excellence, SMIS ensures the basic quality elements enabling the Slovenian economy to become competitive on the global market, and Slovenian society to achieve international recognition, along with the protection of life, health and the environment.

Office of the Slovenian National Commission for UNESCO is responsible for affairs involving Slovenia's co-operation with UNESCO, the United Nations Educational, Scientific and Cultural Organisation, the implementation of UNESCO's goals in Slovenia, and co-operation with National commissions and bodies in other countries and with non- governmental organisations.

General Approaches – Science Policy
Educating top-quality researchers/experts and increasing their number, increasing the extent of research activity and achieving a balanced coverage of all the basic scientific disciplines necessary for:
- quality undergraduate and postgraduate education,
- the effective transfer and dissemination of knowledge from abroad,
- cultural, social and material development,
- promoting the application of science for national needs,
- promoting the transfer of R&D results into production and to the market,

- achieving stronger integration of research into the networks of international co-operation (resulting in the complete internationalisation of science and partly of higher education),
- broadening and deepening public understanding of science (long-term popularisation of science, particularly among the young).

General Approaches – Technology Policy
- promotion of R&D co-operation among enterprises, as well as between enterprises and the public sector,
- strengthening of the investment capacities of enterprises,
- strengthening of the innovation potential of enterprises,
- creation of an innovation-oriented legal and general societal framework,
- supporting the banking sector in financing innovation-orientated and export-orientated business
- development of bilateral and multilateral strategic alliances,
- establishment of ties between the Slovenian R&D sector and foreign industry,
- accelerated development of professional education and the education of adults,
- protection of industrial and intellectual property.

An increase of total invested assets in R&D to about 2.5% of GDP by the year 2000 is planned (of this, half is to be obtained from public sources, with the remainder to come from the private sector). Regarding the development of technology, Slovenia is one of the most technologically advanced in Central Europe and has a well-developed research infrastructure. This has led to a significant growth in the export of high-tech goods. There is also a continued emphasis on the development of R&D across a wide field which is leading to the foundation and construction of technology parks (high -tech business incubators), technology centres (technology-transfer units within public R&D institutions) and small private enterprise centres for research.

R&D Human Potential
There are about 750 R&D groups in the public and private sector, of which 102 research groups are at 17 government (national) research institutes, 340 research groups are at universities and 58 research groups are at medical institutions. The remaining R&D groups are located in business enterprises (175 R&D groups) or are run by about 55 public and private non-profit research organizatios.

According to the data of the Ministry of Science and Technology there are about 7000 researchers in Slovenia. The majority (43%) are lecturers working at the two universities, 15% of researchers are employed at government (national) research institutes, 22% at other institutions and 20% in research and development departments of business enterprises.

JOŽEF STEFAN INSTITUTE

Jožef Stefan (1835-1893) was one of the most prominent physicists of the 19th century. Born to Slovene parents, he obtained his Ph.D. at Vienna University, where he was later Director of the Physics Institute, Vice-President of the Vienna Academy of Sciences and a member of several scientific institutions in Europe. Stefan explored many areas in hydrodynamics, optics, acoustics, electricity, magnetism and the kinetic theory of gases. Among other things, he originated the law that the total radiation from a black body is proportional to the 4th power of its absolute temperature, known as the Stefan–Boltzmann law.

The Jožef Stefan Institute (JSI) is the leading independent scientific research institution in Slovenia, covering a broad spectrum of fundamental and applied research in the fields of physics, chemistry and biochemistry, electronics and information science, nuclear science technology, energy research and environmental science.

The Jožef Stefan Institute (JSI) is a research organisation for pure and applied research in the natural sciences and technology. Both are closely interconnected in research departments composed of different task teams. Emphasis in basic research is given to the development and education of young scientists, while applied research and development serve for the transfer of advanced knowledge, contributing to the development of the national economy and society in general.

At present the Institute, with a total of about 700 staff, has 500 researchers, about 250 of whom are postgraduates, over 200 of whom have doctorates (Ph.D.), and around 150 of whom have permanent professorships or temporary teaching assignments at the Universities.

In view of its activities and status, the JSI plays the role of a national institute, complementing the role of the universities and bridging the gap between basic science and applications.

Research at the JSI includes the following major fields: physics; chemistry; electronics, informatics and computer sciences; biochemistry; ecology; reactor technology; applied mathematics. Most of the activities are more or less closely connected to information sciences, in particular computer sciences, artificial intelligence, language and speech technologies, computer-aided design, computer architectures, biocybernetics and robotics, computer automation and control, professional electronics, digital communications and networks, and applied mathematics. ·

The Institute is located in Ljubljana, the capital of the independent state of Slovenia (or S♡nia). The capital today is considered a crossroad between East, West and Mediter-

ranean Europe, offering excellent productive capabilities and solid business opportunities, with strong international connections. Ljubljana is connected to important centers such as Prague, Budapest, Vienna, Zagreb, Milan, Rome, Monaco, Nice, Bern and Munich, all within a radius of 600 km.

In the last year on the site of the Jožef Stefan Institute, the Technology park "Ljubljana" has been proposed as part of the national strategy for technological development to foster synergies between research and industry, to promote joint ventures between university bodies, research institutes and innovative industry, to act as an incubator for high-tech initiatives and to accelerate the development cycle of innovative products.

At the present time, part of the Institute is being reorganized into several high-tech units supported by and connected within the Technology park at the Jožef Stefan Institute, established as the beginning of a regional Technology park "Ljubljana". The project is being developed at a particularly historical moment, characterized by the process of state reorganisation, privatisation and private initiative. The national Technology Park will take the form of a shareholding company and will host an independent venture-capital institution.

The promoters and operational entities of the project are the Republic of Slovenia, Ministry of Science and Technology and the Jožef Stefan Institute. The framework of the operation also includes the University of Ljubljana, the National Institute of Chemistry, the Institute for Electronics and Vacuum Technology and the Institute for Materials and Construction Research among others. In addition, the project is supported by the Ministry of Economic Relations and Development, the National Chamber of Economy and the City of Ljubljana.

Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel.:+386 1 4773 900, Fax.:+386 1 219 385
Tlx.:31 296 JOSTIN SI
WWW: http://www.ijs.si
E-mail: matjaz.gams@ijs.si
Contact person for the Park: Iztok Lesjak, M.Sc.
Public relations: Natalija Polenec

CONTENTS OF *Informatica* Volume 24 (2000) pp. 1–524

Papers

KUBAT, M. 2000. Recycling decision trees in numeric domains. Informatica 24:195–204.

LIANG, W. & M.E. ORLOWSKA. 2000. Computing multidimensional aggregates in parallel. Informatica 24:107–115.

LIU, C. & X. ZHOU. 2000. On incremental global update support in cooperative database systems. Informatica 24:27–31.

LUKOVSZKI, C., R. SZABÓ & T. HENK. 2000. Performance evaluation of a hybrid ATM switch architecture by parallel discrete event simulation. Informatica 24:259–267.

MARTZ, W.B. 2000. A discussion on process losses in GSS: exploring the consensus gap. Informatica 24:167–174.

MERNIK, M., M. LENIČ, E. AVDIČAUŠEVIĆ & V. ŽUMER. 2000. Multiple attribute grammar inheritance. Informatica 24:319–328.

MILAN-FRANCO, J.M., R. JIMÉNEZ-PERIS & M. PATIÑO-MARTÍNEZ. 2000. DD-Mod: a library for distributed programming. Informatica 24:379–389.

MIZZARO, S. & P.Z. RIZIÒ. 2000. An automatically refereed scholarly electronic journal: formal specifications. Informatica 24:431–438.

MOGHRABI, I.A.R. & F. RAIDAN. 2000. Fractal geometry for natural-looking tree generation. Informatica 24:117–123.

MOOR, DE, O., K. BACHHOUSE & S.D. SWIERSTRA. 2000. First-class attribute grammars. Informatica 24:329–341.

OUSSALAH, M. & A. SERIAI. 2000. A framework for query formulation aid in a multi-user environment. Informatica 24:57–62.

PAPA, G. & J. ŠILC. 2000. Linear algebra in one-domensional systolic array. Informatica 24:249–257.

RADOVAN, M. 2000. Twelve theses on the information age. Informatica 24:445–448.

SASAKI, T., T. HIRONAKA, S. FUJINO & T. TAKAYAMA. 2000. Performance improvement of Thakore's algorithm with speculative execution technique and dynamic task scheduling. Informatica 24:33–38.

SHEPHERD, M.M. 2000. An empirical study to measure the diffusion of GroupSystems in organizations. Informatica 24:187–194.

SICHERL, P. 2000. Digital divide and the gap between North America and Europe in Internet users per capita: new insight from existing data. Informatica 24:439–444.

SIROMONEY, A. & K. INOUE. 2000. Elementary sets and declarative biases in a restricted gRS–ILP model. Informatica 24:125–135.

SOBOL, M.G. & M.A. WINNIFORD. 2000. Factors affecting the use, adoption and satisfaction with groupware. Informatica 24:175–186.

SOLINA, F. 2000. Internet based art installations. Informatica 24:459–466.

SOUZA BAPTISTA, DE, C. & Z. KEMP. 2000. STepLib: a spatio temporal digital library. Informatica 24:69–73.

SULEIMAN, J., R. EVARISTO & G.G. KELLY. 2000. Facilitating and coordinating distributed joint application development. Informatica 24:159–166.

TOPTSIS, A.A. 2000. Heuristic clustering of reusable software libraries. Informatica 24:487–496.

VEHOVAR, V. & L. KOGOVŠEK. 2000. Measuring new media. Informatica 24:449–452.

VRTAČNIK, M., V. FERK, D. DOLNIČAR, N. ZUPANČIČ-BROUWER & M. SAJOVEC. 2000. The impact of visualisation on the quality of chemistry knowledge. Informatica 24:497–503.

YANG, Y. 2000. An architecture and the related mechanisms for web-based global cooperative teamwork support. Informatica 24:13–19.

Editorials

MERNIK, M, & D. PARIGOT. 2000. Introduction: attribute grammars and their application. Informatica 24:285–286.

KLEIN, G. & M.M. SHEPHERD. 2000. Informatica 24:147.

ZHANG, Y., V.A. FOMICHOV & A.P. ŽELEZNIKAR. 2000. Introduction: Special Issue on database, web and cooperative systems. Informatica 24:1–2.

GYÖRKÖS, J. 2000. Introduction: Media in Information Society. Informatica 24:429.

A Conference Report

DEY, S.K. 2000. Charleston, IL, Consciousness Conference, November 6–7,1998. Informatica 24:137–142.

Calls for Papers

Information Society—IS'2000. An international multi-conference. 2000. Informatica 24:143,282.

ITI 2000.   22-nd International conference on information technology interfaces. 2000. Informatica 24:144.

ERK'2000.   Electrotechnical and Computer Science Conference. 2000. Informatica 24:281.

Professional Societies

The Ministry of Science and Technology of the Republic of Slovenia. 2000. Informatica 24:145,283,427,523.

Jožef Stefan Institute. 2000. Informatica 24:146,284,428,524.

INFORMATICA

AN INTERNATIONAL JOURNAL OF COMPUTING AND INFORMATICS

INVITATION, COOPERATION

Submissions and Refereeing

Please submit three copies of the manuscript with good copies of the figures and photographs to one of the editors from the Editorial Board or to the Contact Person. At least two referees outside the author's country will examine it, and they are invited to make as many remarks as possible directly on the manuscript, from typing errors to global philosophical disagreements. The chosen editor will send the author copies with remarks. If the paper is accepted, the editor will also send copies to the Contact Person. The Executive Board will inform the author that the paper has been accepted, in which case it will be published within one year of receipt of e-mails with the text in Informatica LaTeX format and figures in .eps format. The original figures can also be sent on separate sheets. Style and examples of papers can be obtained by e-mail from the Contact Person or from FTP or WWW (see the last page of Informatica).

Opinions, news, calls for conferences, calls for papers, etc. should be sent directly to the Contact Person.

QUESTIONNAIRE

☐ Send Informatica free of charge

☐ Yes, we subscribe

Please, complete the order form and send it to Dr. Rudi Murn, Informatica, Institut Jožef Stefan, Jamova 39, 61111 Ljubljana, Slovenia.

Since 1977, Informatica has been a major Slovenian scientific journal of computing and informatics, including telecommunications, automation and other related areas. In its 16th year (more than five years ago) it became truly international, although it still remains connected to Central Europe. The basic aim of Informatica is to impose intellectual values (science, engineering) in a distributed organisation.

Informatica is a journal primarily covering the European computer science and informatics community - scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the Refereeing Board.

Informatica is free of charge for major scientific, educational and governmental institutions. Others should subscribe (see the last page of Informatica).

ORDER FORM – INFORMATICA

Name: ....

Title and Profession (optional): ....

....

Home Address and Telephone (optional): ....

....

Office Address and Telephone (optional): ....

....

E-mail Address (optional): ....

Signature and Date: ....

Informatica WWW:

Referees:

Witold Abramowicz, David Abramson, Adel Adi, Kenneth Aizawa, Suad Alagić, Mohamad Alam, Dia Ali, Alan Aliu, Richard Amoroso, John Anderson, Hans-Jurgen Appelrath, Vladimir Bajič, Grzegorz Bartoszewicz, Catriel Beeri, Daniel Beech, Fevzi Belli, Francesco Bergadano, Istvan Berkeley, Azer Bestavros, Andraž Bežek, Balaji Bharadwaj, Ralph Bisland, Jacek Blazewicz, Laszlo Boeszoermenyi, Damjan Bojadžijev, Jeff Bone, Ivan Bratko, Jerzy Brzezinski, Marian Bubak, Leslie Burkholder, Frada Burstein, Wojciech Buszkowski, Rajkumar Bvyya, Netiva Caftori, Jason Ceddia, Ryszard Choras, Wojciech Cellary, Wojciech Chybowski, Andrzej Ciepielewski, Vic Ciesielski, David Cliff, Maria Cobb, Travis Craig, Noel Craske, Matthew Crocker, Tadeusz Czachorski, Milan Češka, Honghua Dai, Deborah Dent, Andrej Dobnikar, Sait Dogru, Georg Dorfner, Ludoslaw Drelichowski, Matija Drobnič, Maciej Drozdowski, Marek Druzdzel, Jozo Dujmović, Pavol Ďuriš, Johann Eder, Hesham El-Rewini, Warren Fergusson, Pierre Flener, Wojciech Fliegner, Vladimir A. Fomichov, Terrence Forgarty, Hans Fraaije, Hugo de Garis, Eugeniusz Gatnar, James Geller, Michael Georgiopolus, Jan Goliński, Janusz Gorski, Georg Gottlob, David Green, Herbert Groiss, Inman Harvey, Elke Hochmueller, Jack Hodges, Rod Howell, Tomáš Hruška, Don Huch, Alexey Ippa, Ryszard Jakubowski, Piotr Jedrzejowicz, A. Milton Jenkins, Eric Johnson, Polina Jordanova, Djani Juričič, Sabhash Kak, Li-Shan Kang, Orlando Karam, Roland Kaschek, Jacek Kierzenka, Jan Kniat, Stavros Kokkotos, Kevin Korb, Gilad Koren, Henryk Krawczyk, Ben Kroese, Zbyszko Krolikowski, Benjamin Kuipers, Matjaž Kukar, Aarre Laakso, Phil Laplante, Bud Lawson, Ulrike Leopold-Wildburger, Joseph Y-T. Leung, Barry Levine, Xuefeng Li, Alexander Linkevich, Raymond Lister, Doug Locke, Peter Lockeman, Matija Lokar, Jason Lowder, Kim Teng Lua, Andrzej Malachowski, Bernardo Magnini, Peter Marcer, Andrzej Marciniak, Witold Marciszewski, Vladimir Marik, Jacek Martinek, Tomasz Maruszewski, Florian Matthes, Daniel Memmi, Timothy Menzies, Dieter Merkl, Zbigniew Michalewicz, Gautam Mitra, Roland Mittermeir, Madhav Moganti, Reinhard Moller, Tadeusz Morzy, Daniel Mossé, John Mueller, Hari Narayanan, Rance Necaise, Elzbieta Niedzielska, Marian Niedq'zwiedziński, Jaroslav Nieplocha, Jerzy Nogieć, Stefano Nolfi, Franc Novak, Antoni Nowakowski, Adam Nowicki, Tadeusz Nowicki, Hubert Österle, Wojciech Olejniczak, Jerzy Olszewski, Cherry Owen, Mieczyslaw Owoc, Tadeusz Pankowski, William C. Perkins, Warren Persons, Mitja Peruš, Stephen Pike, Niki Pissinou, Aleksander Pivk, Ullin Place, Gustav Pomberger, James Pomykalski, Dimithu Prasanna, Gary Preckshot, Dejan Rakovič, Cveta Razdevšek Pučko, Ke Qiu, Michael Quinn, Gerald Quirchmayer, Luc de Raedt, Ewaryst Rafajlowicz, Sita Ramakrishnan, Wolf Rauch, Peter Rechenberg, Felix Redmill, David Robertson, Marko Robnik, Ingrid Russel, A.S.M. Sajeev, Bo Sanden, Vivek Sarin, Iztok Savnik, Walter Schempp, Wolfgang Schreiner, Guenter Schmidt, Heinz Schmidt, Dennis Sewer, Zhongzhi Shi, William Spears, Hartmut Stadtler, Olivero Stock, Janusz Stoklosa, Przemysław Stpiczyński, Andrej Stritar, Maciej Stroinski, Tomasz Szmuc, Zdzislaw Szyjewski, Jure Šilc, Metod Škarja, Jiři Šlechta, Chew Lim Tan, Zahir Tari, Jurij Tasič, Piotr Teczynski, Stephanie Teufel, Ken Tindell, A Min Tjoa, Wieslaw Traczyk, Roman Trobec, Marek Tudruj, Andrej Ule, Amjad Umar, Andrzej Urbanski, Marko Uršič, Tadeusz Usowicz, Elisabeth Valentine, Kanonkluk Vanapipat, Alexander P. Vazhenin, Zygmunt Vetulani, Olivier de Vel, John Weckert, Gerhard Widmer, Stefan Wrobel, Stanislaw Wrycza, Janusz Zalewski, Damir Zazula, Yanchun Zhang, Zonling Zhou, Robert Zorc, Anton P. Železnikar

EDITORIAL BOARDS, PUBLISHING COUNCIL

Informatica

An International Journal of Computing and Informatics