BALTIC CONFERENCE

Human – Computer Interaction

Riga, 23.-24.08.2011

Universität Rostock 2012

Herausgeber: Prof. Dr. Clemens Cap
Wissenschaftsverbund „Informations- und Kommunikationstechnologien“ (IuK)

Erstellung der Druckvorlage: Sebastian Aust

Entwurf des Umschlagbildes: Christine Bräuning

ISBN:

Bezugsmöglichkeiten: Universität Rostock
Institut für Informatik
Frau Doreen Schulze
Albert-Einstein-Str. 22, Raum 163
18059 Rostock

Universität Rostock
Wissenschaftsverbund IuK
Frau Dr. Christine Bräuning
Albert-Einstein-Str. 22, Raum 364
18059 Rostock

Druck: Universitätsdruckerei Rostock

Table of Contents

Preface

Due to the generous support of the German Academic Exchange Service (DAAD - Deutscher Akademischer Austausch Dienst), BaSoTI summer school is going in to its seventh year, with the associated conference going into its fifth year.

It was in 2005, that the University of Bremen, the University of Lübeck, the International School of New Media at the University of Luebeck (ISNM), and the University of Rostock joined forces for the first Baltic Summer School in Technical Informatics (BaSoTI). Supported by a sponsorship of the German Academic Exchange Service a series of lectures was offered between August 1 and August 14, 2005 at Gediminas Technical University at Vilnius, Lithuania. The goal of the Summer School was to intensify the educational and scientific collaboration of northern German and Baltic Universities at the upper Bachelor and lower Master level.

In continuation of the successful programme, BaSoTI 2 was again held at Vilnius in 2006, BaSoTI 3 took place in Riga, Latvia at the Information Systems Management Institute in 2007, BaSoTI 4 and BaSoTI 5 were held at the University of Tartu, BaSoTI 6 took place in Kaunas, Lithuania and BaSoTI 7 at the Technical University of Riga, Latvia.

Since BaSoTI 3, the Summer School lectures have been complemented by a one day scientific event. The goal is to give young, aspiring PhD candidates the possibility to learn to give and to survive an academic talk and the ensuing discussion, to get to know the flair and habits of academic publishing and to receive broad feedback from the reviewers and participants. Moreover, the Summer School students would have a chance to participate in what most likely would be their first academic research event.

In 2011, due to a close cooperation with the robotic lab of the Technical University of Riga, an emphasis was placed on the topic of human computer interaction and robotics. However, since the primary focus of the conference is its educational aspect, also some papers with other topics were accepted. In addition to the publications in the current volume, also several presentations were given by summer school attendants, who made use of the occasion to receive feedback on their work and their presentation skills, without aspiring to produce a formal paper.

Clemens H. Cap
Rostock, January 2012.

Programme Committee

Prof. Dr.-Ing. Andreas Ahrens (University of Technology, Business and Design, Wismar)
Prof. Dr. rer. nat. Clemens Cap (University of Rostock)
Prof. Dr.-Ing. Antje Düsterhöft (University of Technology, Business and Design, Wismar)
Prof. Dr. rer. nat. Frank Kirchner (University of Bremen)
Dr.-Ing. Andreas Könsgen (University of Bremen)
Dr.-Ing. Andreas Mundt (University of Rostock)
Dr.-Ing. sc. Agris Nikitenko (Riga Technical University)
Prof. Dr.-Ing. Ralf Salomon (University of Bremen)
Prof. Dr. rer. nat. Lars Schwabe (University of Rostock)
Prof. Dr.-Ing. Peter Sobe (University of Applied Sciences Dresden)
Prof. Dr.-Ing. Dirk Timmermann (University of Rostock)
Prof. Dr.-Ing. Bodo Urban (Fraunhofer IGD Rostock)

Adaptation in Speech Dialogues – Possibilities to Make Human-Computer-Interaction More Natural

Markus Berg, Antje Düsterhöft
University of Wismar
{markus.berg, antje.duesterhoeft}@hs-wismar.de

Bernhard Thalheim
University of Kiel
bernhard.thalheim@is.informatik.uni-kiel.de

Abstract: Adaptation is an important issue for the creation of pleasant user interfaces. In this paper we identify characteristics that can be adapted in speech dialogues. In order to realise adaptive speech dialogues we first have to develop a model that enables us to easily define dialogues and that supports adaptivity. Hence we propose the first step of a backend-oriented dialogue model.

1 Introduction

Human Computer Interaction is a successful, expanding and also hyped and controversially discussed area of research. Today nearly every device and even clothes already are or can be equipped with small microprocessors, resulting in ad hoc networks, sensor arrays and remotely controllable devices. But apart from the technical aspect, we also need interfaces to be able to interact with these devices and to benefit from the information they offer. We have permanent internet access and we use computers several times a day. As technology finds its way into our living rooms and also involves new user groups, conventional user interfaces become inconvenient. Usability suffers from inappropriate user interfaces and alternatives have to be found. Most prominently Touch- and Visualisation Technologies as well as Speech Processing Systems. In this paper we focus on the latter.

Text-based language interfaces already entered the market in the 1960s. With the turn of the millennium, speech recognition technology reached an acceptable quality and became self-evident in the field of telephony, resulting in so called Interactive Voice Response Systems. Although disliked by many customers, companies gladly used this new technology in order to save money. But not only in callcenters Speech Technology can help us to improve effectivity, convenience and usability. Think of a speech enabled living room or the possibility to search for holiday trips just by saying *“I’d like to go to Paris with my wife for one week, could you find me a cheap hotel near the city centre?”*. Besides, also in the field of eInclusion Speech Technology becomes an important factor.

The lack of today’s systems is not the underlying technology but a poor human-like behaviour and a deficient dialogue strategy. Users are often disappointed about the limited understanding today's systems offer and the missing feature of adaptivity. Like humans adapt to their dialogue partner, also the system should adapt to the user. In this paper we identify different possibilities to adapt a speech dialogue and point out how we can reach this aim.

2 Related Work

During the last years we have worked on different dialogue systems. In the *eOhr* (electronic Ear) and later *MAIKE*[1] [13] projects we developed a system to control rooms by voice. Unlike other systems we not only focussed on a command-and-control system but on a natural dialogue. In the *Travel Consult* project [12] we developed a text-based booking system. In contrast to current chatbots like *IKEA Anna* we realised a mixed-initiative dialogue that enables the user to freely choose what to say. As we received different opinions on the style of the user interface, we conducted a user study [14] to answer the question if people from different age groups and with different abilities use a different style of speaking. We could not infer a general rule set or a one-to-one relation between any of the regarded criteria. Instead we believe that the style of speaking is a matter of personal preference. Some people prefer systems which make use of full sentences and social elements (like greeting, thanking,...) and others like a telegraphic style claiming that one does not need to be polite when speaking to a machine.

Also Bell [11] finds that "speaking styles and dialogue strategies vary from one user to another". Edlund [8] explains this with the help of the point of view with regard to the system. From the point of the *interface metaphor* the dialogue system is perceived as a machine, from the *human metaphor* it is perceived as a human-like creature. As we have to support different user types with different preferences and skills, models for adaptive dialogue systems help to simplify the development of those systems and improve user satisfaction.

We don't want to hide the fact that some researchers don't support the idea of natural dialogues. They claim that reliability is more important than naturalness. Usability is defined by success rate, time and simplicity – not by the level of anthropomorphism. Unfortunately anthropomorphism can make it even worse as "users attribute the computer more intelligence than is warranted to it". This leads to "unrealistic expectations of the capabilities of the system" [7]. Brennan and Oheari [5], too, concluded that the anthropomorphic style is undesirable for dialogue systems because it encourages more complex user input which is harder to recognise and interpret [10]. This results in recognition errors and misunderstandings. In the worst case the user is not able to reach his aim, or at best needs much time and many attempts. Stent points out that "users choose not to interact with dialogue systems as they would with other humans" resp. that "humans adapt to the interaction style of their conversational partners" [1]. A reason why people adapt to systems is given

[1] Mobile Assistive Systems for Intelligent and Cooperating Rooms and Ensembles, in cooperation with the University of Rostock

by Brennan [4] who suggested that "users adopt system's terms to avoid errors, expecting the system to be inflexible" [10].

As you can see, the negative attitude is connected with inflexible systems, negative experience and systems which are not able to process natural dialogue. This strengthens our belief that we need to focus on the system behaviour. As there are different types of users with different demands, we need to support adaptive dialogues. The mere fact that man is able to adapt to a system does not guarantee that this form of communication is the most effective.

3 Adaptation in Speech Dialogues

Adaptation is the adjustment of software behaviour during runtime. We differentiate between *adaptability* which refers to a manual adaptation by the user and *adaptivity* which is an automatic, self-adjusting adaptation of the system. Besides users often have the possibility to manipulate settings before runtime. In this case we speak of *configurability*. Both terms are closely related as it is possible for most attributes to change them before and during runtime.

The function of a dialogue does not refer to its form. The following example leads to the same result, yet having a completely different realisation.

S: How can I help you?
U: I'd like to go to Paris with my wife.
S: Oh, Paris, what a nice city. And do you already have a date in mind?
U: Easter would be great.
S: Ok, um, let me see.
S: I have fou... yeah... I have found three wonderful hotels.

S: Where do you want to go to?
U: Paris
S: Did you say Paris?
U: Yes
S: Departure date?
U: April, 20th
S: Return date?
U: April, 24th
S: How many persons?
U: Two
S: I will now list three hotels.

Our user study [14] has shown that different users have different preferences. As most people are conditioned to graphical user interfaces, they also try to use speech in the same way. We call this a *menu-oriented mindset*. Like Brennan [4] we also believe that many people avoid the use of human-human-like language in order to avoid errors. Also we do believe that they are just not aware of the system's capabilities. Consequently most participants claimed that they would have used a more human-like language style if they had known of this possibility. Interestingly unease led to the use of informal language in full sentences, which indicates that people concentrate on not using this form of language when speaking to the system. Only in certain situations they fall back to their natural style of interaction. All these observations confirm the need of adaptive dialogue systems as they embody an efficient and easy usable interface. Apart from the manual configuration of the dialogue by the dialogue designer also a self-adjusting dialogue system would be

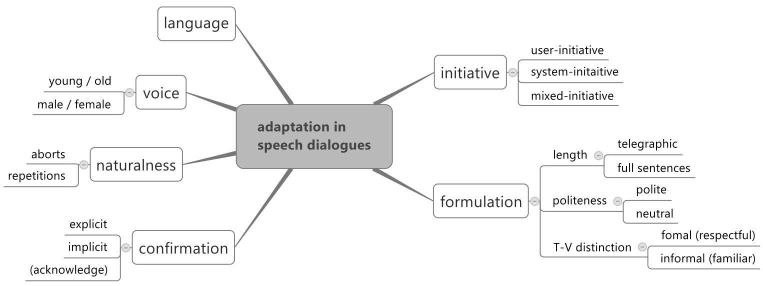


Figure 1: Adaptation in speech dialogues

beneficial. Every dialogue could start in a formal and mixed-initiative way and adapt to the user's style of speaking over runtime.

When speaking of speech dialogues, there are several characteristics to adapt: initiative, formulation, style, politeness, confirmation, naturalness, voice and language, as you also can see in figure 1. We can classify all attributes as *form-related* or *behaviour-related*.

Initiative

The initiative describes if the user is only passive and responds to questions or if he can actively influence the dialogue flow. An adaptation makes sense when errors occur. A dialogue could start with an open-ended question like *"How can I help you?"*. An inexperienced user may not know what to say or uses the wrong words. In most systems a recognition error leads to the repetition of the same question which of course is not of any help. Instead the system should create a direct question: *"Where do you want to go to?"*. Another example of an initiative-change can be seen in the following example:

S: When do you want to travel?
U: For two weeks starting next Monday.
S: I did not understand you. Please say the start date!
U: 12th of May.
S: Now specify the return date!
U: 26th of May.

The system expects answers in the form "from X to Y" and is thus not able to interpret the user answer. Consequently the system generates a more specific question. A change of the type of initiative also makes sense in the context of different user types, as "certain users are likely to voluntarily give a spoken dialogue system feedback throughout the dialogue, while others have to be explicitly asked to provide the same information" [11]. So the system has to adapt its style of asking questions to the user's way of responding to questions.

Confirmation

Confirmation is a method to ensure that the system correctly understood a user's answer. This is crucial for the usability of a system. Different from graphical user interfaces you don't automatically have feedback. At the same time feedback is much more important because natural language is fuzzy and harder to recognise than keyboard input and thus more error-prone. Despite its importance, confirmation in speech interfaces can easily be annoying. That's why we use different confirmation strategies based on the confidence of the recognition result. In the following example S1 represents explicit confirmation and S2 represents implicit confirmation. S3 and S4 use no confirmation at all whereas S3 produces the illusion that the system understood the user instead of only asking the next question. S3 can be seen as acknowledging the user's answer.

S: Where do you want to go?
U: Rome
S1: You want to go to Rome, correct?
S2: When do you want to come back from Rome?
S3: Ok, and when do you want to come back?
S4: When do you want to come back?

Formulation

The formulation of system prompts extremely contributes to the appearance, or *hear & feel*, of a voice interface. Some people prefer telegraphic sentences or even single words like *"Destination?"* and others expect full sentences with obeying politeness like *"Which city do you want to fly to?"* or *"Would you please inform me about your destination city?"*. In many languages you also have to obey T-V-distinction[2], which refers to respectful[3] or familiar[4] formulation. While common systems use a predefined formulation which is only synthesised by the system (TTS), we aim at a concept-to-speech-component (CTS) which automatically generates language.

Moreover the formulation style of the system has influence on the style of the user answers. Hence we can "influence users to behave in a certain way, for instance by implicitly encouraging a speaking style that improves speech recognition performance" [11]. Jokinen suggests an inclusion of context information in the case of low confidence levels. This refers to different confirmation styles depending on the confidence. In her example [9] the answer to the question *"When will the next bus leave for Miami?"* could be *"2.20pm"*, *"It will leave at 2.20pm"* or *"The next bus to Miami leaves at 2.20pm"*.

Further Attributes

Apart from these characteristics, also the language or voice can be adapted to the user's wishes. A language identification algorithm could automatically switch the language when

[2] lat. tu/vos
[3] german: "siezen"
[4] german: "duzen"

it recognises that the language of the user differs from the current interface language. Moreover we can adapt the gender or age of the synthesised voice. Another interesting approach is to increase naturalness or human-likeness by introducing typical speech phenomena like repetition or abortion. Moreover the use of coughing, hawking or filler words like “um” or “err” could improve acceptance within some user groups.

4 Dialogue Modelling

The task of the dialogue manager is the evaluation of the user goal and response generation [9], i.e. to find out *what* the user wants and *how* this goal can be realised with regard to different context information. In order to create an adaptive dialogue system, we first have to realise the interpretation of the user goal. Thus we need a dialogue model which allows an easy dialogue definition that also supports adaptivity. As we have mentioned, the dialogue logic is independent from the form of the dialogue. That’s why we plan to separate form and function with the help of language generation algorithms. Another important way of supporting this goal is dialogue acts. Dialogue acts base on speech acts which have been invented by Austin [2] and advanced by Searle [15]. They define what a user *does* by *saying* words. In other words they generalise utterances to their function or the user’s intention. By saying *“hello”* we greet, by saying *“Mind the gap”* we warn or by saying *“Do you know where James is?”* we ask. Speech Acts have been revised by different researchers (e.g. Harry Bunt [6]) resulting in Dialogue Acts. They have been used in many research projects like *Verbmobil* or *Trains* and embody an important tool in dialogue modelling.

In our research we focus on *task-oriented dialogues*. This includes *information-seeking* dialogues (e.g. travel booking systems), *question-answering dialogues* and *command dialogues* (e.g. smart room control). In these dialogue types we focus on the understanding of the question or command (we comprise this as *concern*) and the generation of an appropriate answer or the intended action (what we summarise as *reply*) [3]. We delimit from conversational or small talk systems and also from human-human-communication, i.e. we are only interested in information that helps us to support the user. We want to realise this as natural and human-like as possible but we explicitly do not focus on the exchange of opinions or the telling of stories.

When analysing existing dialogue act tagsets, we observe that they are very detailed and also include parts we excluded from our considerations, namely commitments, promises, oaths, nominations or threats. Also the recognition of turn-taking and stalling is not important to understand the intention of the user. In task-oriented dialogue systems the user has a specific goal. This goal is defined within the range of services that the system offers. Thus we limit the dialogue acts to those which directly refer to the backend. In [3] we identified three basic backend functions: *getInfo*, *setInfo* and *do* as you can see in figure 2. A *getInfo*-call puts a request to the backend, *setInfo* changes or sets the value of a variable in a frame and *do* executes a function (switch the light on, start a presentation software). The first two methods are information-related while the last one refers to all other actions. These backend-functions influence the selection of dialogue acts, so we propose the

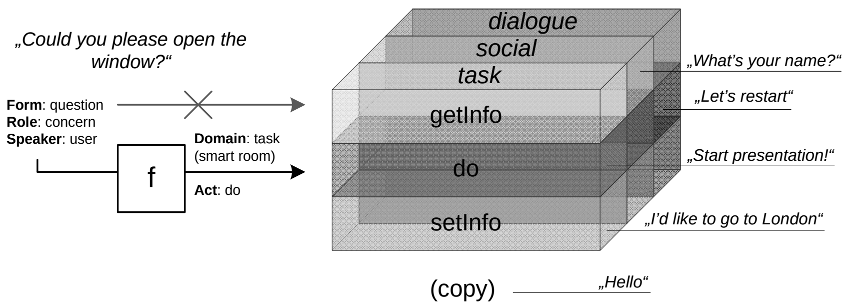


Figure 2: Backend-based dialogue acts in task-oriented dialogues

following top-level dialogue-acts: *information seeking*, *information providing* and *action requesting*. We must not mix up these categories. An instruction like *"Search for hotels in London that offer breakfast"* is not an action request but an information seeking act.

We now have the *role*-attribute (concern and reply) and the *act* or goal of the user (information seeking, information providing, action requesting). Apart from that, we can also define *domain*, *speaker* and *form*. Even a single-task system consists of different domains, i.e. task domain, social domain and dialogue domain. Whereas all task-related requests obviously belong to the task domain, requests like *"Could you repeat that?"* or *"Let's restart"* are dialogue-related (i.e. they influence the behaviour of the dialogue). Here we also distinguish between getInfo, setInfo and do. A dialogue-oriented do-request is especially interesting in the context of adaptability because it allows the user to influence dialogue attributes, e.g. *"Please use indirect confirmation"* or *"Set the voice to female"*. The social domain represents social obligations that need to store or retrieve information like *"Hello, my name is Markus"* (setInfo) resp. *"Do you remember my name?"* (getInfo). Of course most social obligations don't need any access to the backend. Most acts are symmetric adjacency pairs like openings and closings (*greet*, *say goodbye*). Hence we introduce a special dialogue act – *copy* – that doesn't need backend access. As already mentioned, another attribute is the *speaker*. Here we distinguish *user* and *system*. The *form* of an utterance is the most basic information. It can be seen as the foundation for any further analysis. But since form and function don't build a one-to-one relationship, the inference of the function based on form and context is an eminent task. Sometimes we regard form as *secondary illocution* and function as *primary illocution*.

The resulting dialogue model consists of role, speaker, act/function, form and domain, e.g. *"Could you please open the window?"* can be formalised as a quintuple: *(concern, user, action request, question, smart room)*. Of course this information only allows us to call the correct type of backend function, i.e. we identify the general user goal. The actual proposition has to be modelled in the next step.

5 Conclusion and Future Work

The adaptation of speech dialogues is important for the usability of the system. We have shown several possibilities how to adapt dialogues, i.e. when adaptation should take place and what characteristics can be adapted. Our examples include the adaptation of the initiative in case of errors or bad recognition confidence, the adaptation of the style or formulation, change of language and voice and the adaptation of the confirmation strategy.

In order to be able to create an adaptable dialogue system, we first need a dialogue model which allows us to define dialogue systems in an abstract way. As you can see in figure 3, in this paper we have realised a model for the *general user goal*. This is the first step for creating an adaptive dialogue system. In our future work we will address further modules, i.e. the inference of the proposition, the language generator and of course the dialogue manager that determines *how* communication takes place. All of these parts have to interact with the adaptation context for the system to be able to adapt properly.

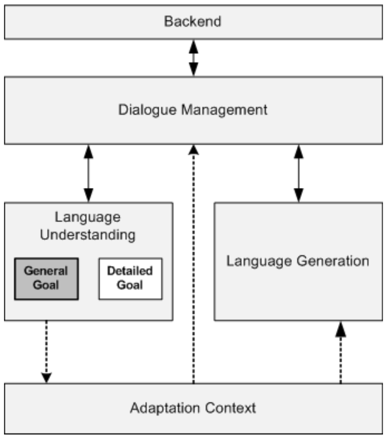


Figure 3: Interaction of the modules with the adaptation context

References

[1] Amanda J. Stent. *Dialogue Systems as Conversational Partners: Applying Conversation Acts Theory to Natural Language Generation for Task-Oriented Mixed-Initiative Spoken Dialogue*. PhD thesis, University of Rochester, 2001.

[2] J. L. Austin. How to do things with words. 1962.

[3] Markus Berg, Bernhard Thalheim, and Antje Düsterhöft. Dialog Acts from the Processing Perspective in Task Oriented Dialog Systems. In *Workshop on the Semantics and Pragmatics of Dialogue*, Los Angeles (USA), 2011.

[4] Susan E. Brennan. Lexical entrainment in spontaneous dialogue. In *Proceedings of the International Sympo- sium on Spoken Dialogue*, pages 41–44, 1996.

[5] Susan E. Brennan and Justina O. Ohaeri. Effects of message style on user's attribution toward agents. In *Proceedings of CHI'94 Conference Companion Human Factors in Computing Systems*, pages 281–282, 1994.

[6] Harry Bunt et al. Towards an ISO standard for Dialogue Act Annotation. In Nicoletta Calzolari et al., editors, *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC 2010)*, Valletta, Malta, 2010. European Language Resources Association (ELRA).

[7] Byron Long. Natural Language as an Interface Style. `http://www.dgp.toronto.edu/people/byron/papers/nli.html`, 1994.

[8] Jens Edlund, Mattias Heldner, and Joakim Gustafson. Two faces of spoken dialogue systems. In *Interspeech 2006 - ICSLP Satellite Workshop Dialogue on Dialogues: Multidisciplinary Evaluation of Advanced Speech-based Interactive Systems.*, 2006.

[9] Kristiina Jokinen and Graham Wilcock. Adaptivity and response generation in a spoken dialogue system, 2007.

[10] Ivana Kruijff Korbayova and Olga Kukina. The effect of dialogue system output style variation on users' evaluation judgments and input style. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, SIGdial '08, pages 190–197, Morristown, NJ, USA, 2008. Association for Computational Linguistics.

[11] Linda Bell. *Linguistic Adaptations in Spoken Human-Computer Dialogues*. PhD thesis, KTH Stockholm, 2003.

[12] Markus Berg and Antje Düsterhöft. Website Interaction with Text-based Natural Language Dialog Systems. In *7. Wismarer Wirtschaftsinformatiktage*, Wismar (Germany), 6 2010.

[13] Markus Berg, Nils Weber, Gernot Ruscher, and Sebastian Bader. Maike: Mobile Assistenzsysteme für intelligente kooperierende Räume und Ensembles. In *5. Kongress Multimediatechnik*, Wismar (Germany), 10 2010.

[14] Markus Berg, Petra Gröber, and Martina Weicht. User Study: Talking to Computers. In *3rd Workshop on inclusive eLearning*, London (UK), 9 2010.

[15] John R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, 1969.

16

Performance Analysis of an Optical MIMO Testbed in Comparison to Wireless MIMO Systems

Sebastian Aust, Andreas Ahrens, Jens Pankow and Steffen Lochmann
Hochschule Wismar, University of Applied Sciences: Technology, Business and Design
email: {andreas.ahrens, steffen.lochmann}@hs-wismar.de

Abstract: Wireless communication is nowadays one of the areas attracting a lot of research activity due to the strongly increasing demand in high-data rate transmission systems. The use of multiple antennas at both the transmitter and receiver side has stimulated one of the most important technical breakthroughs in recent communications allowing increasing the capacity and dropping the bit-error rate. However, multiple-input multiple-output (MIMO) systems are not limited to wireless MIMO systems and can be observed in a huge variety of transmission links and network parts and have attracted a lot of attention since the mid 90's. In the field of optical MIMO transmission systems, multi-mode (MM) fibre offers the possibility to transmit different signals by different mode groups. The perspective of the MIMO philosophy within the field of optical transmission systems is elaborated in this contribution based on channel measurements within a (2×2) MIMO system. For the channel measurements the second optical window and a fibre length of $1,4$ km was chosen. Computer simulations on an overall data rate of $10,24$ Gbps underline the potential of multi-mode fibres in optical high-data rate MIMO communication systems and show that in order to achieve the best bit-error rate, not necessarily all MIMO layers have to be activated.

1 Introduction

The increasing desire for communication and information interchange has attracted a lot of research since Shannon's pioneering work in 1948. A possible solution was presented by Teletar and Foschini in the mid 90's, which revived the MIMO (multiple-input multiple-output) transmission philosophy introduced by van Etten in the mid 70's [Tel99, Fos96, van75, van76].

Since the capacity of MIMO systems increases linearly with the minimum number of antennas at both the transmitter as well as the receiver side, wireless MIMO schemes have attracted substantial attention [MC05,MW02] and can be considered as an essential part of increasing both the achievable capacity and integrity of future generations of wireless systems [K0̈6,ZT03]. However, the MIMO technique isn't limited to wireless communication and a lot of scenarios can be described and outperformed by the MIMO technique.

In comparison to the wireless MIMO channel, the optical fibre is an important type of a fixed-line medium, which is used in several sections of telecommunication networks, where single- and multi-mode fibres are distinguished [SSB08, WG90].

Optimizing the transmission on high-data rate links is in particular of great practical interest for delivering voice or video services in mobile IP (Internet Protocol) based networks in the access domain. Unfortunately, the inherent modal dispersion limits the maximum data speed within the multimode fiber (MMF). In order to overcome this limitation, the well-known single-input single-output systems, also called SISO systems, should be transferred into systems with multiple-inputs and multiple-outputs, also called MIMO systems [HT06, SSB08]. Taking finally into account that delay-spread in wireless broadband MIMO transmission systems isn't any longer a limiting parameter, MMF links should be well suited for high-speed data transmission [RC98, RJ99].

Different research groups, e.g. [SR07, SSR08], have adapted the MIMO technique on optical communication channels. The experimental equalization of crosstalk within frequency non-selective optical MIMO systems has attracted a lot of research [SR07, SSR08]. By contrast, frequency selective MIMO links require substantial further research, where spatio-temporal vector coding (STVC) introduced by RALEIGH for wireless MIMO channels seems to be an appropriate candidate for optical transmission channels too [RC98, RJ99].

In this contribution the spatial multiplexing (SM) is implemented at the transmitter side via different sources launching light with different offsets into the fibre. At the receiver side, i.e. at the fibre end, various spatial filters are implemented [PALA11]. By launching light with different offsets into the fibre, different mode groups are activated, which propagate along the fibre with different speed and attenuation. Together, with the crosstalk between the different mode groups, the classical MIMO channel is formed, where the most beneficial choice of the number of activated MIMO layers and the number of bits per symbol offer a certain degree of design freedom, which substantially affects the performance [ABP09].

Against this background, the novel contribution of this paper is that based on channel measurements within a (2×2) optical MIMO system, the perspective of the MIMO philosophy within the field of optical transmission systems is elaborated. Our results show that even for relatively long (e. g. $1,4$ km) transmission lengths high data rates (e. g. $10,24$ Gbps) are feasible and that the choice of the number of bits per symbol and the number of activated MIMO layers substantially affects the performance of a MIMO system, suggesting that not all MIMO layers have to be activated in order to achieve the best BER.

The remaining part of this paper is organized as follows: Section 2 introduces the optical MIMO channel. The crosstalk impact within the optical MIMO channel is studied in section 3, while the associated performance results are presented and interpreted in section 4. Finally, section 5 provides some concluding remarks.

2 Optical MIMO Channel

In order to comply with the demand on increasing available data rates, systems with multiple inputs and multiple outputs, also called MIMO systems (multiple-input multiple-output), have become indispensable and can be considered as an essential part of increas-

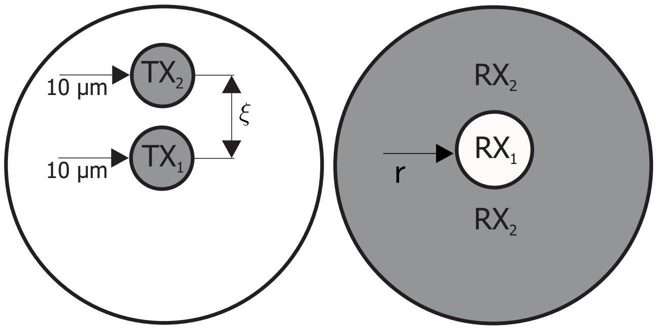


Figure 1: Forming the optical MIMO channel (left: light launch positions at the transmitter side with a given eccentricity ξ, right: spatial configuration at the receiver side as a function of the mask radius r)

ing both the achievable capacity and integrity of future generations of communication systems [KÖ6], [ZT03].

In this work the potential of the MIMO philosophy in optical channels is elaborated, based on channel measurements. Basically, light launched at different positions within the fibre activates different mode groups, which propagate along the fibre with different speed. Low order mode groups, activated by light launched into the center of the fibre, lead to a power radiation pattern concentrated at the center of the fiber end whereas higher order modes, activated by light launched at an off-center position, e. g., with a given eccentricity ξ, within the fibre, lead to power radiation pattern concentrated at the off-center of the fiber-end. Therefore, by launching light into the fibre with given eccentricities, as highlighted in Fig. 1, different spatially separated power distribution pattern can be obtained at the receiver side to form the optical MIMO channel.

According to Fig. 1, the optical input TX_1 was adjusted to launch light into the center of the core (center launch condition), whereas for the optical input TX_2 a given eccentricity ξ was chosen (off-center launch condition). The activated modes can be separated at the fibre end by the corresponding power distribution pattern. Fig. 2 illustrates the simulated power distribution pattern by activating low- and high-order modes. The simulations are in good agreement with the measured power radiation pattern as depicted in Fig. 3 for different parameters of the eccentricity ξ. Now, spatial ring filters at the end of the transmission line as depicted in Fig. 1 have been applied for channel separation. These spatial filters have been produced by depositing a metal layer at fiber end-faces and subsequent ion milling [PALA11], [PAL10].

Together, with the crosstalk between the different mode groups, the classical MIMO channel is formed (Fig. 4) [PALA11], [PAL10].

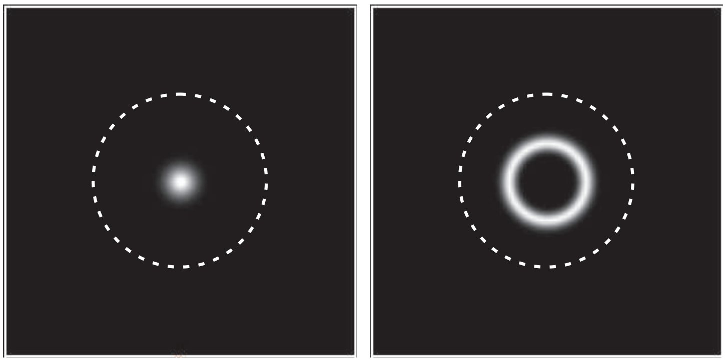


Figure 2: Simulated mean power distribution pattern (left: by the LP_{01} mode, right: by activating all solutions of LP_{81} modes); the dotted line represents the $50\,\mu$m core size.

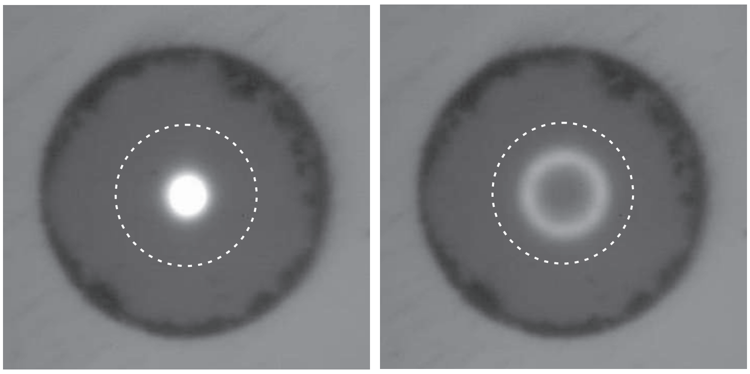


Figure 3: Measured mean power distribution pattern as a function of the light launch position (left: eccentricity $\xi = 0\,\mu$m, right: eccentricity $\xi = 18\,\mu$m); the dotted line represents the $50\,\mu$m core size.

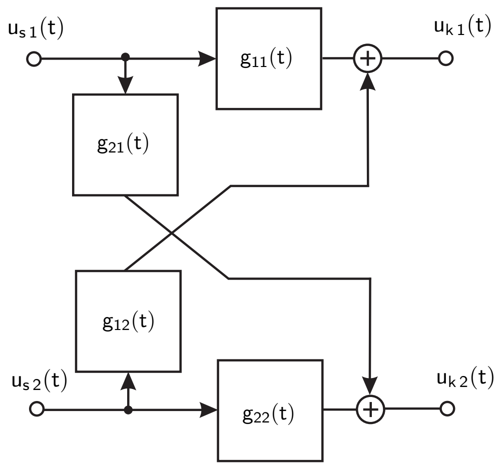


Figure 4: Electrical MIMO system model (example: $n = 2$)

3 Crosstalk Impact

In this section it is studied, how the crosstalk impact depends on the MIMO system configuration. Therefore at this point the eccentricity of the transmitter side MIMO configuration as well as the mask radius of the ring filter configuration at the receiver side are investigated in an exemplary system according to Fig. 5. It is assumed, that each MMF input

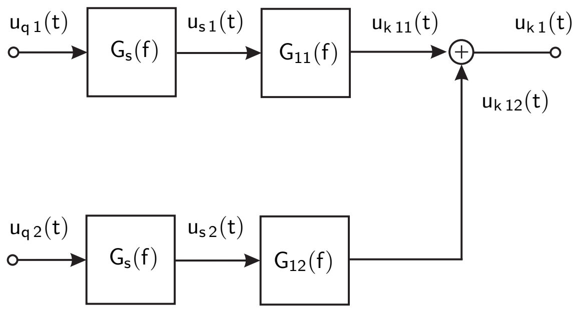


Figure 5: Electrical system model of transmitter and MMF with crosstalk (example: $n = 2$)

is fed by a system with identical mean properties with respect to transmit filtering, pulse frequency $f_\mathrm{T} = 1/T_\mathrm{s}$, the number of signalling levels and the mean transmit power P_s.

The source signals $u_{\mathrm{q}\,1}(t)$ and $u_{\mathrm{q}\,2}(t)$ traverse the transmit filters with the transfer function $G_\mathrm{s}(f)$. Then the wanted transmit signal $u_{\mathrm{s}\,1}(t)$ passes the MMF, modelled by the transfer function $G_{11}(f)$ and causes the signal $u_{\mathrm{k}\,11}(t)$ with power $P_{\mathrm{k}\,11}$ at the MMF output,

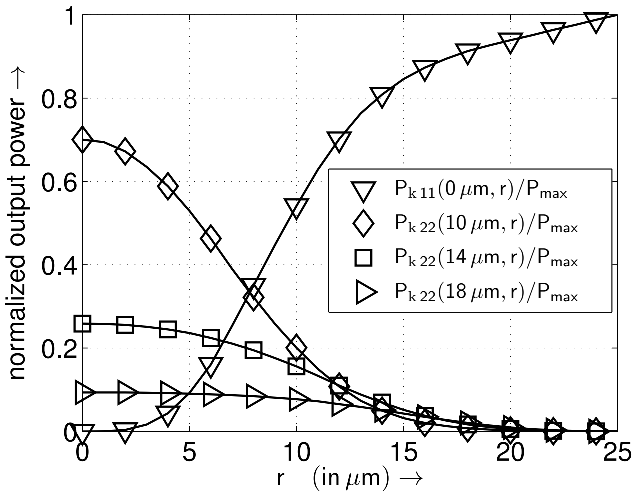


Figure 6: Measured electrical signal power $P_{\mathrm{k}\,\nu\,\mu}(\xi, r)$ at the MMF output as a function of the mask radius r for given parameters of the eccentricity ξ

whereas the crosstalk signal $u_{\mathrm{k}\,12}(t)$ (with power $P_{\mathrm{k}\,12}$), which is fed into the MMF with a given eccentricity ξ, originates at the MMF output, after the transmit signal $u_{\mathrm{s}\,2}(t)$ passed the filter with the transfer function $G_{12}(f)$, which models the coupling from optical input 2 to the output 1 (Fig. 5).

In general, the MIMO performance is affected by both the mask radius r and the eccentricity ξ. As highlighted in Fig. 6, the power of the wanted signal $u_{\mathrm{k}\,11}(t)$ at the MMF output, distributed in the inner ring, increases monotonically with rising mask radius r, whereas at the same time the power of the wanted signal $u_{\mathrm{k}\,22}(t)$, distributed in the outer ring, decreased with increasing mask radius and increased eccentricity. In addition a channel asymmetry can be observed which is caused by the larger differential mode attenuation of the higher order mode groups. From this point of view it can be concluded that a mask radius in the range of 5μm to 15μm should be chosen in order to have adequate power at both outputs.

From a practical point of view the power $P_{\mathrm{k}\,12}(\xi, r)$ of the crosstalk signal $u_{\mathrm{k}\,12}(t)$ at the MMF output is an interesting indicator for the strength of the crosstalk disturbance, which depends on the mask radius r and the eccentricity ξ. In order to assess the effect of crosstalk on the wanted signal not only the pure crosstalk signal power is of interest, but rather the behaviour of the powers of the wanted signal and the crosstalk signal to each other. This behaviour may be investigated by a signal-to-crosstalk-interference ratio (SCIR)

$$\varrho_{\mathrm{k}\,11}(\xi, r) = \frac{P_{\mathrm{k}\,11}(0, r)}{P_{\mathrm{k}\,12}(\xi, r)} \quad \text{and} \quad \varrho_{\mathrm{k}\,22}(\xi, r) = \frac{P_{\mathrm{k}\,22}(\xi, r)}{P_{\mathrm{k}\,21}(0, r)} \quad . \tag{1}$$

Since MIMO makes use of the interference for channel improvement the SCIR should not

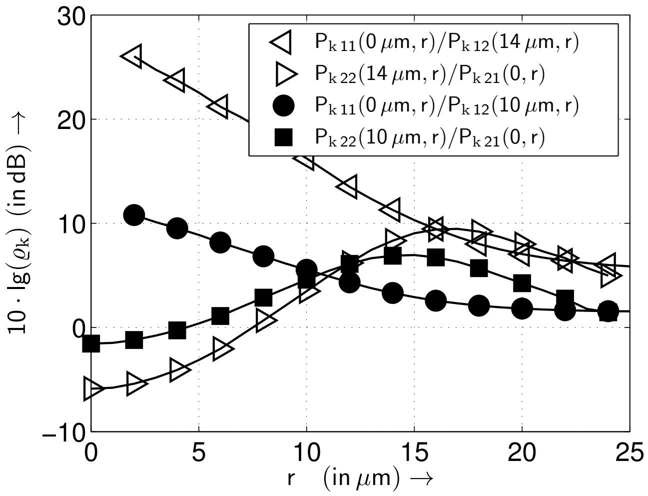


Figure 7: Electrical signal-to-crosstalk-interference ratio (SCIR) ϱ_{k} at the MMF outputs as a function of the mask radius r and the eccentricity ξ

be chosen as high as possible like in orthogonal transmission. Referring to Fig. 7 this means for lower order modes (output 1) a movement towards larger mask radiuses and vice versa for higher order mode groups (output 2).

Though a relationship between the spatial mode location and the channel's impulse response needs to be established for an exact BER (bit-error rate) trade-off it can already been concluded from Fig. 6 and Fig. 7 that mask radiuses in the range of about $5\,\mu$m to $15\,\mu$m should be used for further BER analyses.

4 Performance Analysis

For BER analysis, a frequency selective SDM (spatial division multiplexing) MIMO link, composed of n_{T} inputs and n_{R} outputs, is considered. The block-oriented system for frequency selective channels is modelled by

$$\mathbf{u} = \mathbf{H} \cdot \mathbf{c} + \mathbf{w} \ . \tag{2}$$

In (2), the transmitted signal vector $\mathbf{c}$ is mapped by the channel matrix $\mathbf{H}$ onto the received vector $\mathbf{u}$. Finally, the vector of the additive, white Gaussian noise (AWGN) is defined by $\mathbf{w}$ [PALA11, ABP09]. The interference between the different input's data streams, which is introduced by the off-diagonal elements of the channel matrix $\mathbf{H}$, requires appropriate signal processing strategies. A popular technique is based on the singular-value decomposition (SVD) [Hay02] of the system matrix $\mathbf{H}$, which transfers the whole system into independent, non-interfering layers having unequal gains [PALA11, ABP09].

In this contribution the efficiency of fixed transmission modes is studied regardless of the channel quality. Assuming predefined transmission modes, a fixed data rate can be guaranteed.

For comparison reason the performance of a wireless broadband MIMO system is studied. From wireless MIMO channels it is known that delay spread isn't any longer a limiting parameter [Ges04, ABP09]. In order to analyze the influence of delay-spread in wireless MIMO systems, the following broadband MIMO system model is considered: It is assumed that the number of transmit antennas equals the number of receive antennas, e. g., $n_{\mathrm{R}} = n_{\mathrm{T}} = 4$, and that the $(L_{\mathrm{c}} + 1)$ channel coefficients, between the μth transmit and νth receive antenna have the same averaged power and undergo a Rayleigh distribution [ABP09]. Finally, a block fading channel model is applied, i. e., the channel is assumed to be time invariant for the duration of one SDM MIMO data vector. The obtained

Table 1: Investigated wireless transmission modes

throughput	layer 1	layer 2	layer 3	layer 4
8 bit/s/Hz	256	0	0	0
8 bit/s/Hz	64	4	0	0
8 bit/s/Hz	**16**	**16**	**0**	**0**
8 bit/s/Hz	**16**	**4**	**4**	**0**
8 bit/s/Hz	4	4	4	4

BER curves are depicted in Fig. 8 and 9 for the different QAM constellation sizes and MIMO configurations of Tab. 1, when transmitting at a bandwidth efficiency of 8 bit/s/Hz within a given bandwidth. Assuming a uniform distribution of the transmit power over the number of activated MIMO layers, it turns out that not all MIMO layers have to be activated in order to achieve the best BERs. Comparing the results depicted in Fig. 8 and 9, it can be seen that a high delay spread is quite beneficial for a good overall performance.

Now, the results are applied to the optical MIMO channel. For numerical analysis it is assumed, that each optical input within the MMF is fed by a system with identical mean properties with respect to transmit filtering and pulse frequency $f_{\mathrm{T}} = 1/T_{\mathrm{s}}$. Within this work, the pulse frequency f_{T} is chosen to be $f_{\mathrm{T}} = 5,12$ GHz. The average transmit power is supposed to be $P_{\mathrm{s}} = 1\,\mathrm{V}^2$ and as an external disturbance a white Gaussian noise with power spectral density N_0 is assumed. In order to transmit at a fixed data rate, an appropriate number of MIMO layers has to be used, which depends on the specific transmission mode, as detailed in Tab. 2 for the investigated (2×2) optical MIMO system.

The obtained BER curves are depicted in Fig. 10 for the different ASK (amplitude shift keying) constellation sizes of Tab. 2. For the investigations, an eccentricity of $\xi = 10\,\mu$m and a mask radius of $r = 15\,\mu$m was assumed, which was found to be beneficial for minimizing the overall BER at a fixed data rate. Assuming a uniform distribution of the transmit power over the number of activated MIMO layers, it turns out that not all MIMO

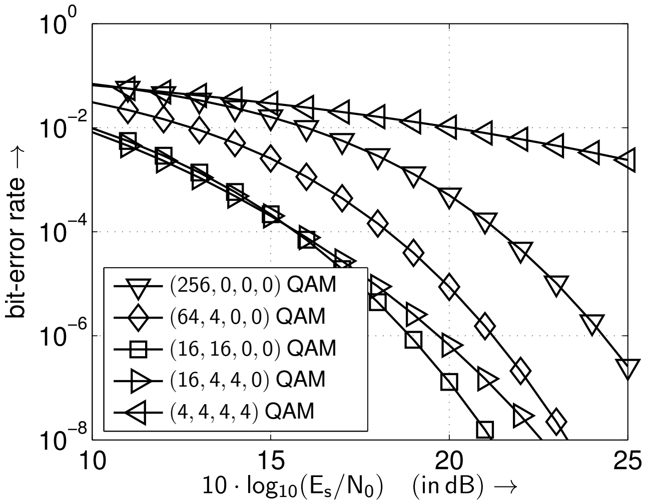


Figure 8: BER when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over frequency selective channels with $L_\mathrm{c} = 1$ (two-path channel model)

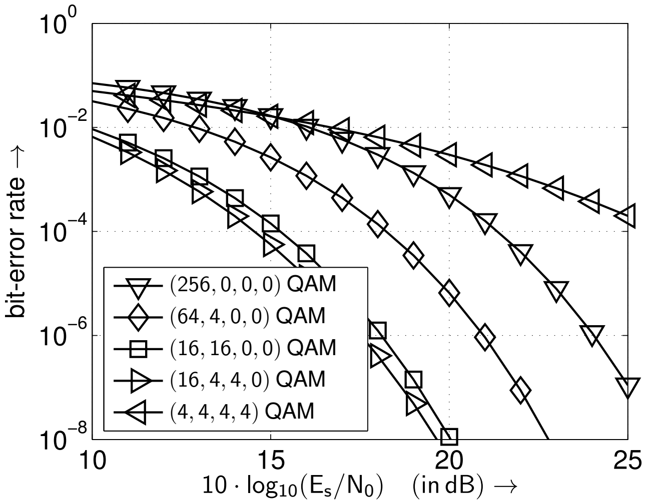


Figure 9: BER when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over frequency selective channels with $L_\mathrm{c} = 4$ (five-path channel model)

Table 2: Investigated optical transmission modes

throughput	layer 1	layer 2
2 bit/s/Hz	4	0
2 bit/s/Hz	2	2

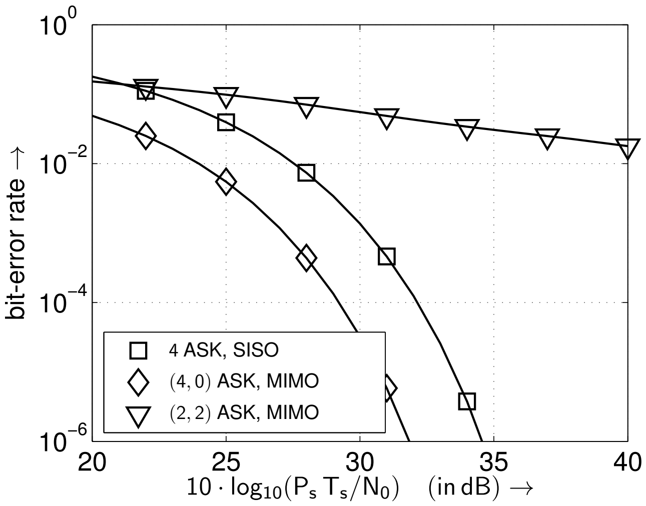


Figure 10: BER when using the transmission modes introduced in Tab. 1 and transmitting 2 bit/s/Hz over frequency selective optical MIMO channels

layers have to be activated in order to achieve the best BERs.

5 Conclusions

In this paper the perspective of the MIMO philosophy within the field of optical transmission systems is investigated. Our results, obtained by channel measurements and computer simulations, show the potential of MIMO techniques in the field of optical transmission systems. In combination with appropriate MIMO signal processing strategies, an improvement in the overall BER was obtained.

Acknowledgements

The authors wish to thank their co-worker, Mr. Ralph Bornitz, for supporting the measurement campaign.

References

[ABP09] A. Ahrens and C. Benavente-Peces. Modulation-Mode and Power Assignment in Broadband MIMO Systems. *Facta Universitatis (Series Electronics and Energetics)*, 22(3):313–327, December 2009.

[Fos96] G. J. Foschini. Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multiple Antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[Ges04] David Gesbert. Multipath: Curse or Blessing? A System Performance Analysis of MIMO Wireless Systems. In *Proceedings of International Zurich Seminar on Communications (IZS)*, pages 14–17, Zurich (Switzerland), 2004.

[Hay02] S. S. Haykin. *Adaptive Filter Theory.* Prentice Hall, New Jersey, 2002.

[HT06] R. Hsu and A. Tarighat. Capacity Enhancement in Coherent Optical MIMO (COMIMO) Multimode Fiber Links. *IEEE Communications Letters*, 10(3):195–197, 2006.

[KÖ6] V. Kühn. *Wireless Communications over MIMO Channels – Applications to CDMA and Multiple Antenna Systems.* Wiley, Chichester, 2006.

[MC05] M. R. McKay and I. B. Collings. Capacity and Performance of MIMO-BICM with Zero-Forcing Receivers. *IEEE Transactions on Communications*, 53(1):74– 83, January 2005.

[MW02] S. H. Mueller-Weinfurtner. Coding Approaches for Multiple Antenna Transmission in Fast Fading and OFDM. *IEEE Transactions on Signal Processing*, 50(10):2442–2450, October 2002.

[PAL10] J. Pankow, A. Ahrens, and S. Lochmann. Channel Measurements and Performance Analysis of Optical MIMO Multimode Fiber Links. In *4th Baltic Conference Learning in Networks*, pages 69–78, Kaunas (Lithuania), 20.-21. August 2010.

[PALA11] J. Pankow, S. Aust, S. Lochmann, and A. Ahrens. Modulation-Mode Assignment in SVD-assisted Optical MIMO Multimode Fiber Links. In *15th International Conference on Optical Network Design and Modeling (ONDM)*, Bologna (Italy), 8.-10. February 2011.

[RC98] Gregory G. Raleigh and John M. Cioffi. Spatio-Temporal Coding for Wireless Communication. *IEEE Transactions on Communications*, 46(3):357–366, March 1998.

[RJ99] Gregory G. Raleigh and V. K. Jones. Multivariate Modulation and Coding for Wireless Communication. *IEEE Journal on Selected Areas in Communications*, 17(5):851–866, May 1999.

[SR07] S. Schöllmann and W. Rosenkranz. Experimental Equalization of Crosstalk in a (2 x 2) MIMO System Based on Mode Group Diversity Multiplexing in MMF Systems @ 10.7 Gb/s. In *33rd European Conference and Exhibition on Optical Communication (ECOC)*, Berlin, September 2007.

[SSB08] A. C. Singer, N. R. Shanbhag, and H. M. Bae. Electronic Dispersion Compensation– An Overwiew of Optical Communications Systems. *IEEE Signal Processing Magazine*, 25(6):110 – 130, 2008.

[SSR08] S. Schöllmann, N. Schrammar, and W. Rosenkranz. Experimental Realisation of (3 x 3) MIMO System with Mode Group Diversity Multiplexing Limited by Modal Noise. In *Optical Fiber Communication Conference (OFC)*, San Diego, California, February 2008.

[Tel99] E. Telatar. Capacity of Multi-Antenna Gaussian Channels. *European Transactions on Telecommunications*, 10(6):585–595, November/Dezember 1999.

[van75] W. van Etten. An Optimum Linear Receiver for Multiple Channel Digital Transmission Systems. *IEEE Transactions on Communications*, 23(8):828–834, 1975.

[van76] W. van Etten. Maximum Likelihood Receiver for Multiple Channel Transmission Systems. *IEEE Transactions on Communications*, 24(2):276–283, 1976.

[WG90] J. Winters and R. Gitlin. Electrical Signal Processing Techniques in Long-Haul Fiber-Optic Systems. *IEEE Transactions on Communications*, 38(9):1439–1453, Sept. 1990.

[ZT03] L. Zheng and D. N. T. Tse. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096, May 2003.

Scene-Dependence of Saliency Maps of Natural Luminance and Depth Images

Lars Schwabe
Redwan Abdo A. Mohammed*

Universität Rostock
Dept. of Computer Science and Electrical Engineering
Adaptive and Regenerative Software Systems†

Abstract: Seamless human-computer-interactions (HCI) needs systems, which are designed in a human-centric way. This includes the use of user models, which take into account internal states of the user such as, for example, the focus of attention. Eye movements and the locations of fixations are an important indicator for the focus of visual attention. However, accurately predicting eye movements remains a challenging problem. So-called saliency maps are a promising tool to predict eye movements based on bottom-up "salient" image features. Here we report, for the first time, statistical properties of saliency in natural luminance *and* depth images. We find that the distribution of saliency in luminance images is unimodal whereas it is bimodal in depth images. As a consequence, low-saliency locations in luminance images can be highly salient in depth-images. This first characterization of joint luminance and depth saliency is an important first step towards developing models of eye movements, which operate well under natural conditions such as those encountered in HCI in ubiquitous computing settings.

1 Introduction

Truly ubiquitous computing and seamless human-computer-interaction (HCI) require systems, which are human-centric in the sense of taking into account a user's internal states such as intentions, current goals, or the focus of attention. The latter is widely acknowledged as an important factor to be considered in the design of user interfaces, and efforts to respect the human attentional system in the design of user interfaces have been made [HKPH03, WLLB04, Rod11].

Most user interface design principles, which respect the human attentional system, are essentially a collection of best practices, which respect theories of attention developed within, for example, cognitive psychology. Those theories, however, are often rather abstract with limited predictive power for particular HCI scenarios. Instead, we argue that

*{lars.schwabe,redwan.mohammed}@uni-rostock.de
†http://ars.informatik.uni-rostock.de

a satisfying theory of attention applicable to HCI shall make specific and testable predictions for interactions and even physiological signals based upon which the attentional state and focus of attention of a user can be inferred.

Eye movements and the visual field locations during fixation periods are often considered as an informative observable, but eye movements are at best an indirect measure of attention. It is known for a long time that task-demands affect the patterns of eye movements [Yar67]. Hence, properly predicting eye movements is still a challenging task, in particular for more natural scenarios [BKK10] such as those encountered in ubiquitous computing. The notion of a saliency map has been turned out helpful in visual attention research: Here, certain locations in the visual field are determined as "salient" if they are – in statistical terms – outliers relative to the surrounding visual field locations. Computational modeling of the visual system was quite successful in the sense of predicting saliency maps based on image properties, which closely match the experimentally measurable maps of eye movements and fixation periods [IK01].

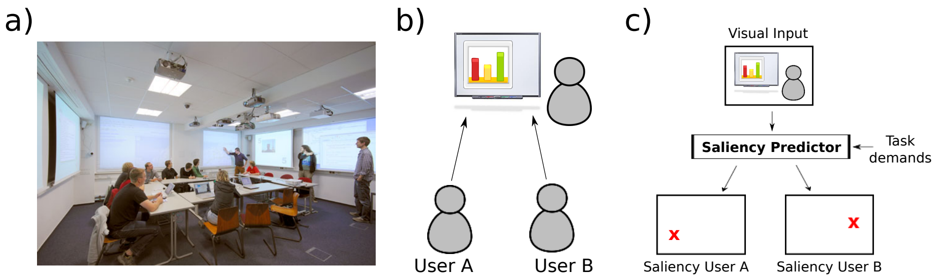


Figure 1: Application scenario for using saliency maps in smart environments. **a)** A typical scenario in the smart meeting room of the MuSAMA GRK (see musama.de). **b)** First abstraction with a speaker in front of a presentation screen and two users looking at that screen. **c)** Illustration of how to use a saliency predictor, which computes a saliency map, in such a setting: Users A and B have approximately the same visual input, but depending on their task demands, different locations in their visual field are rendered as most salient (red crosses). The saliency predictor shall combine bottom-up (the visual input) and top-down information. In this paper, however, we focus only on the bottom-up part.

Such saliency maps reflect bottom-up attentional processes, i. e. the attraction of attention by external cues. They do not consider task demands or the user's internal state. However, based on principles derivable from, for example, a notion of optimal information processing the responses in visual search experiments could be predicted [NI07]. Here we argue that a truly human-centric HCI shall be based on such theories, which can be used to make specific predictions for various HCI scenarios as compared to plain descriptive or even just heuristic design principles.

Figure 1 illustrates how saliency maps can be used within a so-called smart lab (Figure 1a): Various sensors may extract the gaze direction of users in a room. But even if the gaze direction is similiar, as for user A and B (Figure 1b), the saliency maps can still differ due to different task demands. For example, while user A aims at following the presentation, user B's task might be to spot spelling mistakes in slides, which shall make differnt visual field locations salient. Moreover, saliency maps yield richer information than just gaze direction, because they label the whole visual field of a user. This is valuable information for estimating the internal states of users such as in, for exmaple, intention recognition, to adapt visual interfaces, or to place important information.

In this paper, we report the results of a first experimental study to further improve the computation of saliency maps, i. e. to make them ultimately more predictive for eye movements. More specifically, we investigate a collection of natural scenes in terms of their saliency based on the two-dimensional (2D) pixel images *and* the corresponding depth images. The investigation of natural images in terms of their statistical properties is a prominent approach in vision research, because it informs us about the environment to which our visual system has been adapted during evolution and ontogenesis. However, here we go beyond the analysis of the 2D pixel images by incorporating depth images. Those cannot be sensed directly by our visual system but need to be inferred. The rational for investigating depth images is that they may reveal the "saliency that matters", because when interacting with the environment we evolved by interacting with objects in a three-dimensional (3D) world. Thus, we hypothesize that saliency maps respecting this will ultimately outperform saliency maps computed only on the basis of 2D pixel images in terms of predicting eye movements.

This paper is organized as follows: First, we describe the material and methods we used (2). Then, we present the results of our analysis, where we first compare the spatial correlations of the 2D and depth images for different scenes (Sec. 3.1) and then the estimated saliences for 2D and depth images (Sec. 3.2). We close with a summary and an outlook.

2 Materials and methods

2.1 Image collection

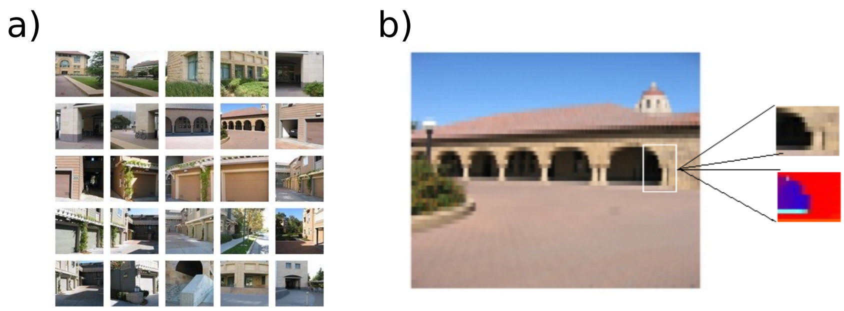


Figure 2: Examples from the image collection. **a)** Pixel images of city scenes. **b)** Resized pixel image (now with 55 rows) and a patch for the 2D image (top right) and the depth image (bottom right; red=closer, blue=more distant).

Our analysis is based on a collection of images obtained from Stanford University [AS09] (see Figure 2). The total number of images in this database is 400. The 2D color pixel images were recorded with a high resolution, but the depth images with a resolution of 305×55 pixel (width $\times$ height). The 2D pixel images were rescaled to the reduced resolution of the depth images and aligned. Details of this procedure are described in the analysis codes, which are provided together with the freely available image collection [AS09]. The luminance color images were transformed into a gray-scale images by taking the gray-value for each pixel in the 2D image as the mean of the RGB values in the color image.

All images were inspected manually and then labeled as either "forest scene", "city scene", or "landscape scene". Only 12 images were labeled as landscape scenes, and we did not include them in our analysis. 80 scenes were labeled as city scenes, and the remaining ones as forest scenes. Therefore, we compared only forest and city scenes. All analyses were repeated by taking random subsets of size 80 from the forest scenes (as we had 80+ forest scenes). The results reported here were not affected by this difference in the sample sizes.

2.2 Analysis methods

Spatial correlations

The correlation between pixels is probably the simplest characterization of images. It reveals how spatial dependencies in images fall off with distance. The luminance and depth values are each given by a single number. Based on these numbers we estimated the correlation coefficient as a function of the distance between any two pixels, i. e.

$$\text{corr}\,(d) := \text{corr}(X_1,\, X_2) = \frac{\text{cov}(X_1,\, X_2)}{\sqrt{\text{var}\,(X_1)\,\text{var}\,(X_2)}}, \tag{1}$$

where X_1 and X_2 are two random variables representing two gray-scale/depth values of two pixels separated by d pixel. Here,

$$\begin{aligned} \text{cov}\,[X_1,\, X_2] &= E\,[X_1 X_2] - E\,[X_1]\,E\,[X_2] && (2)\\ \text{var}\,[X] &= E\left[X^2\right] - E\,[X]^2 && (3) \end{aligned}$$

are the covariance and variance, respectively. $E\,[\cdot]$ denotes the expectation, which we estimated by the sample mean.

Local standard deviation as a feature

We separated each image, i. e. both the luminance and the depth image, into non-overlapping small patches of size 6×6 pixel. Given that the the resolution of these images is 305×55 pixel, a pixel-wise square patch corresponds to a vertically elongated rectangular patch in visual space. Therefore, we report all results for distances in pixel, which cannot be related directly to visual space, but still allows for a fair comparison between 2D and depth images, because the 2D images were resized and aligned to the depth images. Then, we computed the *standard deviation* of the 6×6 gray-scale values for each patch as the value of a local feature. The same was done for the depth image. This yielded, for each image i, the values $f^{2D}(\mathbf{x}; i)$ and $f^{depth}(\mathbf{x}; i)$ as the local feature for the visual field location $\mathbf{x}$.

The z-score to quantify saliency

Within an image i we consider a location $\mathbf{x}_1$ as more salient than another location $\mathbf{x}_2$ in terms of the 2D image feature when $f^{2D}(\mathbf{x}_1; i) > f^{2D}(\mathbf{x}_2; i)$; the same applies to the depth images. In order to abstract from the absolute values of these features, we consider the corresponding z-scores of these features and not the values of the features itself. The z-score at a location $\mathbf{x}$ in an image i is given by

$$z(\mathbf{x};\, i) = \frac{f(\mathbf{x};\, i) - \mu(i)}{\sigma(i)},$$

where $\mu(i)$ and $\sigma(i)$ are the mean and standard deviation of the feature *within* the i-th image, i. e. $f(\cdot)$ could be either $f^{2D}(\cdot)$ or $f^{depth}(\cdot)$. This way, saliency is defined relative to an image not in absolute terms.

3 Results

3.1 Spatial correlations

The luminance and depth images clearly differ in terms of their spatial correlations, which is illustrated in Figure 3a-c and summarized in Figure 3d. Consider the example luminance image and the corresponding depth image: While the gray-scale values of the pixels in the luminance image along a horizontal line (see arrow) is variable (Figure 3c, blue line) the corresponding depths are almost constant (Figure 3c, green line). This suggests that the 3D environment is spatially much more homogeneous than it appears from the 2D pixel images. Interestingly, the spatial correlation over many images reveals a scene-dependence: While the correlations in the luminance images is similar for forest and city scenes (Figure 3d, solid lines), the correlation in the depth images in the city scenes are more extended than in the corresponding luminance images (green dashed vs. green solid line). This is probably due to the presence of many spatially extended surfaces such as walls, streets, etc. On the other hand, the depth images in forest scenes seems to vary less smoothly than the luminance images (blue dashed vs. blue solid line). This could be due to the presence of many strong depth discontinuities in forest scenes such as in trees.

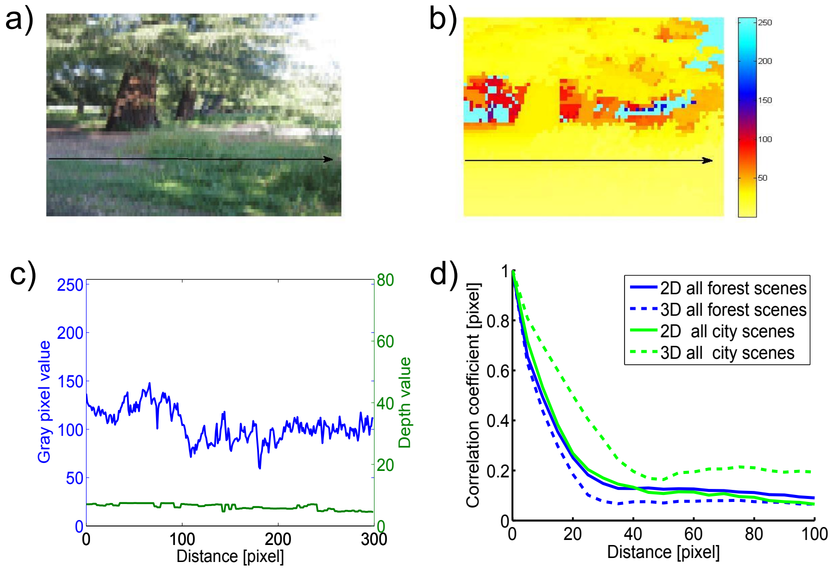


Figure 3: Example to compare the changes in gray-scale and depth values with distance. **a,b)** Color and depth image of an example scene. **c)** Gray-scale and depth values of the pixels along the black arrow in panels a,b. **d)** Spatial correlation as a function of distance measured in pixel. The pixel pairs selected for estimating this function were selected randomly from all possibly pixel pairs in an image.

3.2 Distribution of saliency in natural images

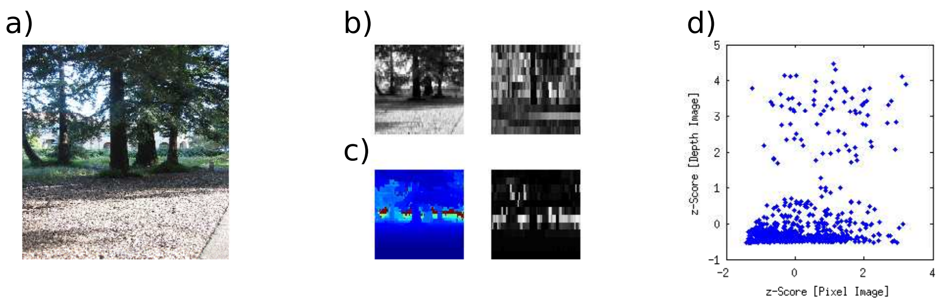


Figure 4: Example of saliency in a 2D and depth image. **a)** Color image of a scene. **b)** Resized image in gray-scale and corresponding saliencies based on the standard-deviation feature (see Section 2). Shown are the z-scores (white=high z-score, black=low z-score). **c)** Same as b) but for the depth image. **d)** Scatter plot of the z-scores in b,c with each point corresponding to an image patch. Saliencies in 2D appear to be unimodal, depth saliency is clearly bimodal.

To illustrate the computation of saliency maps we computed the saliency based on the standard-deviation feature for an example image shown in Figure 4a. Compare the z-scores shown Figure 4b,c (right panels). They are shown using the same color-scale. It is obvious that the luminance image has much more intermediate z-scores than the depth image, which has mainly small values with some high values at the location of the depth discontinuities at the tree trunks. This difference is also prominent in the scatter plot shown in Figure 4d, where the 2D saliencies appear to be unimodal but the saliencies for the depth image are bimodal, i. e. with either low or high values.

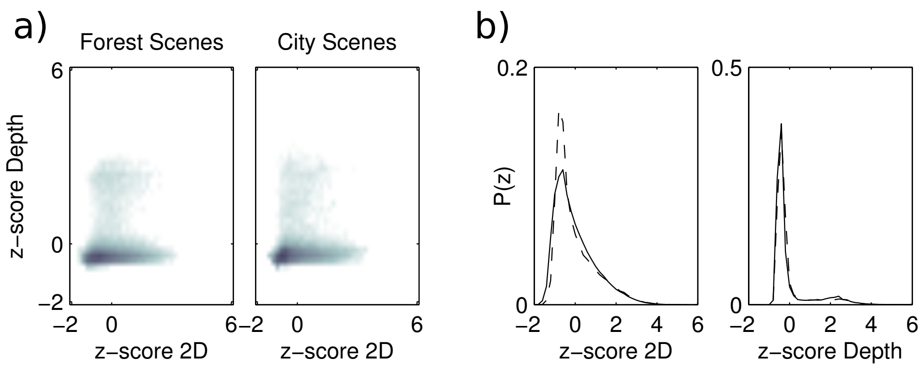


Figure 5: Joint and marginal distributions for the 2D and depth saliency. **a)** Joint probability distribution for the 2D and depth saliency computed for 80 forest and city scenes, estimated using a two-dimensional histogram. Both panels use the same (logarithmic) color-scale. Black corresponds to high probabilities. **b)** Marginal distributions for the forest (solid line) and city scenes (dashed lines).

We also performed this analysis for 80 forest and city scenes. The resulting joint distribution for the saliency in the 2D and depth images is shown in Figure 5a. Interestingly, the locations with high saliency in the depth image correspond to low salient locations in the 2D image. If the "saliency that matters" for our interaction with the 3D environment are the salient locations in the depth images, then they are not spotted by our (certainly very simplistic) 2D saliency algorithm. The marginal distributions are shown in 5b and reveal that 2D saliency is distributed unimodal, whereas depth saliency is bimodal. By visually comparing the joint distribution for forest with city scenes, it appears as if depth saliency is "more bimodal" in forest than in city scenes, but future studies need to explore this further.

4 Conclusion

In summary, we have analyzed the saliency in 2D pixel and depth images using a rather crude and highly simplistic feature: the local standard deviation of pixels. We find that saliency in depth images is bimodally distributed with highly salient locations corresponding to low salient 2D image locations. Given that most saliency algorithms work on the 2D images, this finding points towards including depth cues into the computation of saliency maps as a promising approach to improve their plausibility.

We also found differences between scenes in the spatial correlation functions and a tendency for saliency being more bimodal in forest than in city scenes. It is obvious that these two types of scenes differ in terms of their depth structure, but our first analysis could not make that distinction as clear as expected. We hypothesize that this is due to the (intentionally) rather simple feature we used and the lack of local “center surround interactions”, i. e. a cross-talk between neighboring image locations. Finally, we suggest that saliency maps shall become an integral part of any user model in HCI, in particular for HCI in ubiquitous computing settings, where users are experiencing a much richer visual environment than in the desktop-area of computing. Future work needs to i) include such saliency maps into user models for the sake of, for example, adapting user interfaces to the attentional focus, ii) use more and different features (both 2D and depth features/cues) for the computation of saliency maps, and iii) systematically compare the predictions with experimentally measured eye movements. Including a scene-dependence into applications of such saliency algorithms will further improve their utility.

References

[AS09] Andrew Y. Ng. Ashutosh Saxena, Min Sun. Make3D: Learning 3D Scene Structure from a Single Still Image. *IEEE Transactions of Pattern Analysis and Machine Intelligence (PAMI)*, vol. 30, no. 5:824–840, 2009.

[BKK10] Torsten Betz, Tim C Kietzmann, and Peter König. Investigating task-dependent top-down effects on overt visual attention. *Journal of Vision*, 10:1–14, 2010.

[HKPH03] E. Horvitz, C. Kadie, T. Paek, and D. Hovel. Models of Attention in Computing and Communication : From Principles to Applications. *Communications of the ACM*, 46(3):52–59, 2003.

[IK01] L Itti and C Koch. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3):194–203, March 2001.

[NI07] Vidhya Navalpakkam and Laurent Itti. Search goal tunes visual features optimally. *Neuron*, 53(4):605–17, February 2007.

[Rod11] Claudia Roda, editor. *Human Attention in Digital Environments*. Cambridge University Press, Cambridge, UK, 2011.

[WLLB04] Christopher D. Wickens, John D. Lee, Yili Liu, and Sallie E. Gordon Becker. *An Introduction to Human Factors Engineering*. Pearson Prentice Hall, Upper Saddle River, NJ, 2004.

[Yar67] A. L. Yarbus. *Eye movements and vision*. Plenum Press, New York, 1967.

The Conceptual Model for reliable Multi-Robot Map Merging

Ilze Andersone

Riga Technical university
Kalku 1, Riga, Latvia
ilze.andersone@rtu.lv

Abstract: This article addresses multi-robot map merging and the problem of acquiring reliable merging results. The conceptual model of the multi-robot map merging in the context of mapping is presented. The model can be used to dynamically propose and reject map merging hypotheses without losing the information acquired after the map merging. It consists out of three modules (Global map merging module, Local map merging module and Hypothesis verification module), and each module fulfills an important role to ensure that the map merging is reliable and easily reversible.

1 Introduction

One of the fundamental problems in mobile robotics is the environment mapping problem. Robots need to be able to construct a map of the environment and to use it for the proper behavior including motion planning and navigation. While the use of robot teams becomes more and more widely used, the issue of robot coordination becomes one of the central questions to be addressed. If multiple robots are used for the exploration of the environment, their collected information has to be fused into one general global map. The map fusion from multiple robots into one global map is called map merging [Ko03].

Map merging might be considered from two different viewpoints:

1. Map merging is considered as a physical fusion of two robot maps – the search process in transformation space leading to one or more transformations to be evaluated. In some cases also the preprocessing of the maps is performed. In this case research subjects address the search strategies, feature spaces that are used for map comparison and the evaluation metrics of the map merging result.
2. Map merging is considered in the context of the mapping. The main advantage of this approach is that the results of map merging can be verified. However, in this case the map merging reversion, the choice of the maps and other problems have to be addressed in addition to the research of physical map fusion.

There are two main sources of information used for robot map merging: relative positions of robots and heuristics [And10]. If the relative positions are known or might be determined during the mapping, it is possible to acquire the relative reference frames of the local maps and as a consequence it becomes feasible to merge the maps reliably.

Unfortunately in practice the positional information of the robots is usually unavailable. In this case the map merging hypothesis can be acquired with the help of some heuristic. The heuristic is a function that provides an estimate of solution cost [RN02]. In the map merging case function is a feature space that describes the maps. Unlike map merging that uses robot positional information, it is impossible to determine whether the map merging hypothesis is true, when only heuristics are available.

Most of the researches in the map merging area address solely the physical fusion of maps [BC06, Car08, LLW05, TEE10]. While these approaches evaluate the map merging by applying some similarity measure or metric, they do not guarantee reliable results, because only heuristics are used as the source of information.

Few researchers have addressed the problem of a reliable map merging [Kon03, HB05]. In [Kon03] the problem of the reliable map merging is addressed as a decision where both the absolute likelihood (the similarity of the two maps for a particular transformation) and the relative likelihood (the similarity of the two maps for a particular transformation compared with other transformations) are taken into account. This work, however, does not consider a particular case where the acquired result is wrong and has to be reversed. Instead, the emphasis is put on the avoidance of inaccurate results.

In [HB05] it is admitted that it is natural to make mistakes when merging maps. Therefore the merged maps need to be stored in a way that allows a simple discarding of incorrect hypothesis without losing the whole map or information acquired after map merging. It is proposed to store the maps of each robot in layers. Layer 0 stores the local map, Layer 1 – the maps created by merging Layer 0 map with other maps etc. The problem with this approach is that many maps have to be maintained simultaneously and it can prove to be computationally unfeasible.

When no positional information of the robots is available, it is not certain if the acquired map merging result is correct and therefore the map merging decision needs to be easily reversible. In this paper a conceptual map merging model is proposed to ensure the reliable and easily reversible merging of the robot maps. The proposed map merging model differs from the existing map merging approaches in two ways:

- It addresses not only the physical map merging but the whole process of map merging in the context of mapping. It allows to dynamically create new hypotheses and also offers a simple way to discard them and return to previous state without losing local information gathered by robots.
- It distinguishes the local map merging from the global map merging. Local map merging addresses only the search of transformation between two given maps. This is the point of view taken by most map merging researchers. Global map merging focuses on the creation of global map – how, when and which local maps should be merged to create the global map.

2 The representation of the map merging hypothesis

To ensure that the map merging hypothesis is easy to reverse it is important to choose the appropriate representation. If the robots merge their local maps and continue the mapping by using one common map, it is not an easy task to separate the maps without losing the information acquired after merging [Kon03].

An alternative – to maintain both local and merged maps – can become computationally inefficient, especially if the maps are large and/or several robots are involved in mapping. Therefore such merging should be executed only if the map merging is considered to be reliable, i.e. sometime during the mapping the relative positions of the robots are acquired. To avoid the potential problem of map separation, the maps after merging are still maintained locally.

In the Figure 1 the format of map merging hypothesis can be seen. Hypothesis contains the information about the two merged maps and the transformation between them. Transformation is the position of the second map against the first map – the translations along X and Y axis and the rotation. This representation of the hypothesis allows to use the local maps for the creation of the global map or to reject the hypothesis at any time without actually having to maintain multiple maps simultaneously.

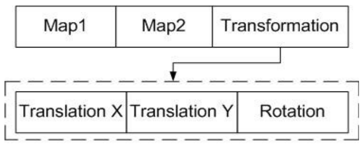


Figure 1: The representation of map merging hypothesis

One map merging hypothesis is obviously not enough when there are several robots that map the environment. Currently there is no map merging approach that would be able to merge more than two maps simultaneously without the positional information. Therefore some gradual merging is required to create the global map of multiple robots involved in mapping. In [AGG05] the pivot method is proposed for the merging of multiple maps. In this approach every two adjacent maps (maps are considered to be ordered sequentially) are merged and then the relative transformations with the rest of maps are computed. It means that for *n* robots *n-1* mergings and *n-2* relative transformation computations are necessary to acquire the global map.

However, it is possible to acquire the global map by using only *n-1* mergings without computing relative transformations. It is achieved by using each map for merging exactly once. Figure 2 depicts the resulting hypothesis hierarchy. In this case only the maps that have not been used in any other mergings are used in further mergings.

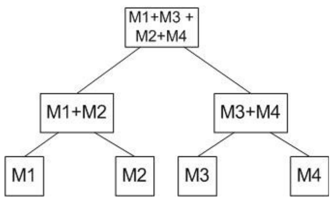


Figure 2: Hypothesis hierarchy for the minimal count of mergings

3 The conceptual model of map merging in the context of mapping

There is no single map merging approach that is suitable for all possible mapping cases. The conceptual model (see Figure 3) described in this paper makes the following assumptions:

- The maps used in mapping and map merging are metric maps. Metric maps are more detailed and therefore allow the robots to determine their position in the environment more precisely and to create optimal movement paths [Thr02].
- All information acquired by the individual robots during the exploration of the environment is accessible at any time and the computations are centralized.
- The initial relative positions of the robots are unknown at the beginning of the exploration. However, it is assumed to be possible that robots meet or sense each other during the exploration. It is very restrictive to assume that the robots start the exploration in a common reference frame. It would mean that either all robots start the exploration in the same exact location or the exact distance and angle between them is known.

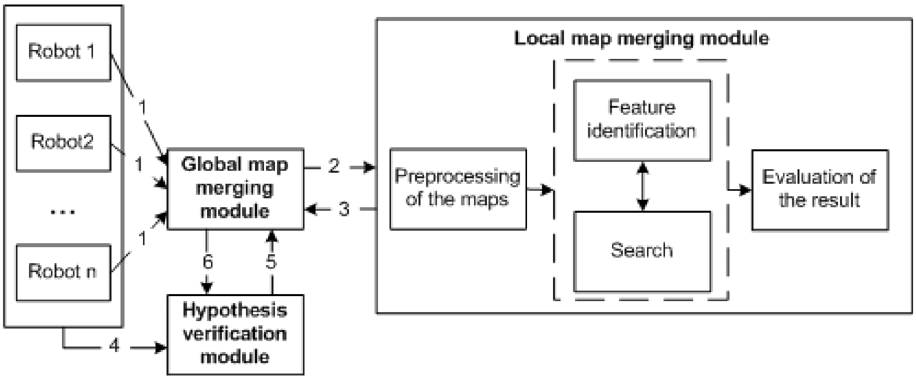


Figure 3: The conceptual map merging model

Information flows represented in the Figure 3: 1) local maps of the robots and information about the relative positions of the robots if the meeting occurs; 2) the maps to be merged and in some cases the relative positions of the robots 3) and 6) map merging hypothesis (transformations) and messages about map merging failure; 4) local maps of the robots; 5) updated hypothesis structure

There are three modules depicted in Figure 3: Global map merging module, Local map merging module and Hypothesis verification module. Each of these modules fulfills an important role to ensure that the map merging is reliable and easily reversible.

3.1 Global map merging module

The Global map merging module is responsible for the maintenance of the global hypotheses hierarchy to achieve the goal of creation of global map. It decides when the map mergings should be made and which maps should be merged.

Robots start the environment exploration by creating their local maps. When the robots have spent some time exploring (the exact timing might depend on the speed of robot exploration, size of local maps etc.), the merging of the maps can be initiated. For example, if four robots participate in mapping and create local maps *m1*, *m2*, *m3* and *m4*, then the goal is to acquire the global map *m1+m2+m3+m4*. The Global map merging module uses the local maps of the robots and rejected hypotheses from the Hypothesis verification module to determine the mergings necessary to acquire full global map.

It is possible that during the initial exploration some of the robots have met and determined their relative reference frames. In this case the local maps of these maps may be already merged before the global map merging process is initiated.

An important question in the creation of the global map is how to choose the order of maps to be merged. Some possibilities include choosing the maps that have similar map size or similar features count. To ensure more diverse map merging combinations the failed map mergings and rejected hypothesis can be taken into account to choose the next map mergings.

3.2 Local map merging module

There are two possible tasks that the Local map merging module can receive – map merging by using heuristics or map merging by using the positional information of the robots [And10]. As the information available is different, each of these cases should use a different map merging approach.

The map merging module in Figure 3 depicts the map merging process when heuristics are used. Map merging approaches that use heuristics consist of at least three steps [And11]: the identification of features in maps, the search and the evaluation of the result. The evaluation is an independent step that can be adapted to every map merging approach. On the other hand the feature identification and searching is often tightly interconnected within the bounds of one approach. For example, [Car08] describes the map merging approach that uses the Hough specters of the maps as features. These specters are then further used in the search of transformation. In approach proposed in [BC06] a stochastic search algorithm uses image similarity heuristic in each step. Besides the three mandatory steps map preprocessing might be used for map merging to reduce the time that is necessary for merging or to ensure the acquisition of a more precise result.

The map merging when robot positions are known is a much simpler task. In this case robot positions are used to determine the relative reference frames of robot maps. Even if the positional information is noisy, it is still useful for reducing the transformation search space.

In some cases an acceptable map merging transformation is not found. The common part of the maps may be too small or there might be no common part at all. In this case the map merging fails, and this information is passed back to the Global map merging module.

3.3 Hypothesis verification module

The Hypothesis verification module uses the computed map merging hypotheses and information from robots to determine if the map merging results are consistent or rollback should be applied. During the merging and further exploration typically three situations can happen:

1. The conflict arises.
2. The meeting of two robots.
3. Nothing happens.

Each of these events requires different reactions from the Hypothesis verification module.

1) The conflict. The information coming from different robots is conflicting indicating that one or more map merging hypotheses are incorrect.

Maps of higher levels are created by hierarchically merging the lower level maps. Periodically the global map (or maps, if there are several highest level maps) is updated. It is gradually recreated from the lower level maps and each step is verified. If a conflict arises in a hypothesis, then this hypothesis is discarded. All the higher level hypotheses that depend on it are also rejected.

Figure 4 represents an example of such situation. To create the global map *M1+M2+M3+M3* at first the sub-maps *M1+M2* and *M3+M4* have to be created. Let's assume that a conflict arises in sub-map *M1+M2*. It means that the map merging hypothesis of *M1+M2* and consequently the hypothesis of *M1+M2+M3+M3* are incorrect. These two hypotheses are then discarded leaving only one map merging hypothesis *M3+M4*. To restore the global map of all robots new mergings have to be performed.

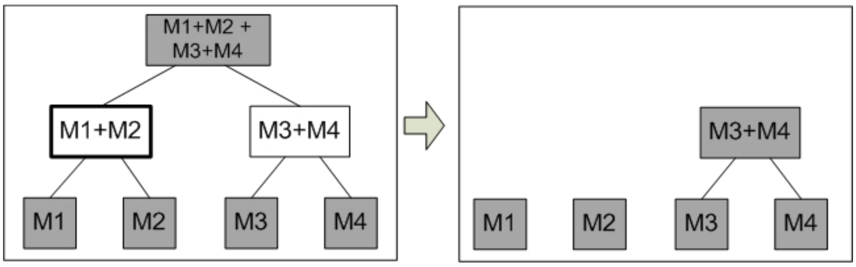


Figure 4: The example of hypothesis rejection

2) The meeting of two robots. The meeting means that robots reach each other's sensor range and consequentially the relative positions can be determined. If two robots meet, they acquire new information that can be used for map merging – the relative positions of the robots. If these two robots have a common map merging hypothesis then it is possible to prove or reject this hypothesis. In the case of rejection all other hypothesis that depend on it are also rejected.

It is more complicated to deal with the meeting, when the robots do not have a direct map merging hypothesis. Figure 5 depicts an example where two robots that have local maps *M1* and *M3* meets. The transformation between the two maps can be now determined. Currently maps *M1* and *M3* do not have a direct map merging hypothesis but they have a connection at a higher level – the map *M1+M2+M3+M4*. The newly acquired transformation must now be compared to the transformation of M1 and M3 that can be acquired from the existing hypothesis tree. If it is correct, then the tree remains unchanged with the exception that a new map merging hypothesis is added directly between *M1* and *M3*. On the other hand, if it is incorrect, then the changes must be made. The hypothesis *M1+M2+M3+M4* is discarded and a new hypothesis *M1+M3* added. There is no reason to discard hypothesis *M1+M2* and *M3+M4*, because they still might be correct. The map pairs *M1+M2&M1+M3* and *M1+M3&M3+M4* each have one common map and the relative transformations between these maps can be acquired. By computing transformations, maps *M1+M2+M3, M1+M3+M4* and consequently new global hypothesis *M1+M2+M3+M4* can be acquired.

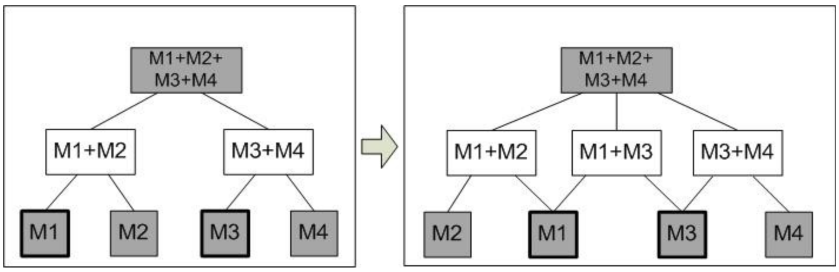


Figure 5: The example of possible hypothesis hierarchy changes when robots that meet do not have direct map merging hypothesis and the hypothesis *M1+M2+M3+M4* proves to be incorrect.

3) Nothing happens. Robots do not meet during the exploration and no conflicts arise. No map merging hypothesis are proven or rejected.

4 Conclusions

In this paper a conceptual model of the multi-robot map merging in the context of mapping is presented. The model can be used to dynamically propose and reject map merging hypotheses without losing the information acquired after the map merging. The representation of hypotheses as transformations in hierarchical format is used to reduce the computational and storage resources that are necessary for the map merging process. When using this representation, it is not necessary to maintain local and global maps simultaneously. Instead, it is possible to acquire the global maps any time by using the local maps of the robots and the hypothesis hierarchy.

The main goal of the proposed conceptual model is to enable reliable map merging and global map creation in the case where robots may not directly meet and acquire the estimations of each other’s position but the global map is still necessary for efficient task distribution. One such multi-robot system is currently under development by the researchers of Riga Technical university. The system consists of vacuum cleaner robots equipped with very simple sensors, which must first explore a room to efficiently distribute the cleaning areas. One central computer is responsible for all the map creation and merging computations so that the individual robots wouldn’t need powerful and consequently expensive processors. To avoid the placement of robots in predetermined positions, the initial locations of the robots are assumed to be unknown.

In future it is planned to implement and validate the proposed conceptual model in the multi-robot system described above. If the conceptual model will be successful, then it will enable the robot team to create a global map of the environment completely autonomously. Currently no mapping and map merging approaches known to the author offer the solution to the particular case. Before the implementation can begin, the main task is to elaborate the global map merging mechanism – the choice of local maps to be merged and the timing of mergings.

Bibliography

[AGG05] Amigoni F., Gasparini S., Gini M.: Merging Partial Maps Without Using Odometry. In Multi-Robot Systems. From Swarms to Intelligent Automata, III vol., 2005, pp. 133-144.

[And10] Andersone I.: The Characteristics of the Map Merging Methods: A Survey. In the Scientific Journal of Riga Technical university, 5th Series, Applied Computer Systems, volume 43, RTU, Riga, 2010, pp. 113-121.

[And11] Andersone I.: Multi-Robot Map Merging in the Context of Image Processing. In the Scientific Journal of Riga Technical university, 5th Series, Applied Computer Systems, RTU, Riga, 2011 (to be published)

[BC06] Birk A., Carpin S. Merging Occupancy Grid Maps From Multiple Robots. Proceedings of the IEEE. – IEEE, 2006, 94. vol., pp. 1384-1397.

[Car08] Carpin S.: Fast and accurate map merging for multi-robot systems. Autonomous Robots. Kluwer Academic Publishers, 2008, 25. vol., pp. 305-316.

[HB05] Huang W., Beevers K.: Topological Map Merging. In The International Journal on Robotics Research. Sage science press, 24. vol., No. 8, 2005, pp. 601-613.

[Ko03] Ko J. et. al.: A practical decision-theoretic approach to multi-robot mapping and exploration. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003, pp. 3232-3238.

[Kon03] Konolige K. et. al.: Map merging for distributed robot navigation. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003, pp. 212-217.

[LLW05] Lakaemper R., Latecki L. J., Wolter D.: Incremental multi-robot mapping. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005, pp. 3846-3851.

[RN02] Russel S., Norvig P. Artificial Intelligence: Modern Approach, 2nd edition. Prentice Hall, 2002, 1132 p.

[TEE10] Topal S., Erkmen E., Erkmen A. M.: A Novel Map Merging Methodology for Multi-Robot Systems. In Proceedings of the World Congress on Engineering and Computer Science, Vol I, 2010, pp. 383-387.

[Thr02] Thrun S.: Robotic Mapping: A Survey. In Exploring Artificial Intelligence in the New Millennium. – Morgan Kaufmann, 2002, pp. 1-35.

Neural Networks in Disease Diagnostics

Henadzi Vaitsekhovich

Department of Intelligent Information Technologies
Brest State Technical University
Belarus, Brest, Moskowskaja str. 267
224000
napster@tut.by

Abstract: In this paper the neural network model for transient ischemic attacks recognition have been addressed. The proposed approach is based on integration of the NPCA neural network and multilayer perceptron. The dataset from clinic have been used for experiments performing. Combining two different neural networks (NPCA and MLP) it is possible to produce efficient performance in terms of detection and recognition transient ischemic attacks. The main advantages of using neural network techniques are the ability to recognize "novel" TIA attack instances, quickness and ability to assist the doctor in making decision.

1 Introduction

Nowadays the use of Artificial Intelligence has become broadly applied in medicine. Every year in medical journals are issued over 500 academic publications concerning artificial neural networks in medical applications [1]. In accordance with published literature the artificial neural networks are powerful tool for automatic diagnostics of disease with potential to support clinical decision making.

There are different techniques to initial diagnostics of TIA: neuropsychological testing, statistical approach, artificial intelligence approach [3-8]. The main disadvantage of neuropsychological testing is correlation on a doctor qualification and low accuracy. The statistical approach demands large database. The artificial intelligence techniques use neural networks, genetic algorithm, fuzzy logic or combination of abovementioned approaches. Such a technique is characterized with high accuracy and demands not too big data set in comparison with the statistical approach. Therefore artificial intelligence techniques can be appropriate tools for TIA detection and recognition. In this paper we propose neural network model for TIA detection and recognition. Neural networks technique is used to reduce the diagnostic time and the number of misdiagnosis, as well as to assist a doctor in making decision. The database from the 5th city hospital of Minsk was used.

The efficiency of the proposed neural network model in detection and recognition of transient ischemic attacks is illustrated by the experimental results.

The paper is organized as follows. A brief review of related works is given in Section 2. The basic information about transient ischemic attacks is presented in Section 3. Section 4 describes patient's data. The proposed neural network model for TIA recognition is detailed in sections 5, 6 and 7. In Section 8 the experimental results are described. Finally, concluding remarks are made in the last section.

2 Related Researches

At present time there exist different approaches to preliminary diagnostics of TIA. Neuropsychological testing is used for initial evaluation of TIA very often [3]. However neuropsychological testing greatly depends on doctor qualification and therefore is very subjective.

The recognition tool for transient ischemic attack is described in [4]. The authors applied multivariate logistic regression using ROC (receiver operating characteristic curves) analysis to develop clinical scoring system. This system correctly identified 85% of patients with a cerebrovascular diagnosis and 54% with a non-cerebrovascular diagnosis. However such an approach demands the big database.

The neural networks for Ischemic Stroke are presented in [6]. The proposed models were developed for rapid classification into the following outputs: no event, TIA or stroke in left carotid, right carotid, verterobasilar.

In [7-8] the backpropagation neural networks for the prediction of thrombo-embolic stroke are described. The architecture of neural network consists of 20 input units, 10 hidden and 10 output units. The prediction accuracy is 78,52% using training set and 90,61% using testing data set.

The mentioned above neural networks approaches are differed each from other used input and output data, as well as database of patients. Therefore it is very difficult to compare different approaches.

3 Transient Ischemic Attacks

A transient ischemic attacks (TIA) is a transient episode of neurological dysfunction caused by focal brain, spinal-cord or retinal ischemia without acute infarction [3]. It is result of temporary reduction or cessation of cerebral blood flow in a specific neurovascular distribution due to low flow through a partially occluded vessel, an acute thromboembolic event, or stenosis of a small penetrating vessel. Transient ischemic attacks are named also thrombo-embolic stroke. After TIA the risk of early acute stroke is increased. Therefore the preliminary diagnosis of TIA is of great importance for prevention of acute stroke. However research has shown a high rate of misdiagnosis of TIA [2].

4 Patient Data

The data from the 5th city hospital of Minsk have been collected about 114 patients who have symptoms of different TIA diseases. 38 parameters have been selected for each patient, such as age, sex, residence, education, trade, conflicts on job, residence change for last 10 years, trade change for last 10 year, features of night dream, sleeplessness, heredity on pathology brain vessels, heredity on other diseases, arterial hypertensia, diastolic pressure, auscultation hearts, heart borders, changes on electrocardiogram, heart pain, cardiac arrhythmia, chronic bronchitis, chronic hepatocholecystitis, chronic gastritis, nephrolithiasis, osteochondrosis, meteodependence, the alcohol use, smoking (amount), smoking (age), working capacity, irritability, memory decline (degree), memory decline (occurrence time), vision acuity decrease (degree), vision acuity decrease (occurrence time), vision disorders, headache (nature), headache (occurrence time), dizziness. The TIA can be classified into 3 classes: TIA1, TIA2 and TIA3. The data set contain 28 patterns of TIA1, 25 patterns of TIA2, 27 patterns of TIA3 and 34 patterns of normal state without TIA.

5 Proposed Model

The proposed model is based on two different neural networks. The 38 features mentioned above are used as input vector, which contains information about patient. The output data of neural network model represent the 4-dimensional vector, where 4 are number of TIA classes plus normal state. The data processing consists of two stages. The first stage of data processing is feature selection. The important question concerning input data is the following: which input parameters are really useful and contribute significantly to the performance of neural networks? In this work nonlinear principal component analysis (NPCA neural network) for significant information from data extraction and dimensionality reduction is proposed. It transforms 38-dimensional input vectors into 12-dimensional target vectors.

The second stage of data processing is to detect and to recognize transient ischemic attacks. Compressed on the previous step data contain the useful information from input data and are used as inputs on the second stage of data processing. The multilayer perceptron (MLP) is applied for transient ischemic attacks recognition.

Thereby the neural network model consists of two neural networks: NPCA and MLP (Fig. 1).

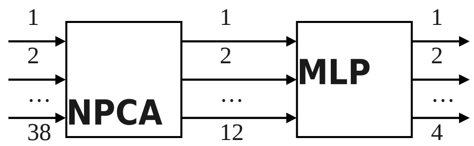


Figure 1: Architecture of the system

Before entering to NPCA the data should transform by the following way:

$$x_i^k = \frac{x_i^k - \mu(x_i)}{\sigma(x_i^k)} \;, \qquad (1)$$

$$\mu(x_i) = \frac{1}{L} \sum_{k=1}^{L} x_i^k \;, \qquad (2)$$

$$\sigma(x_i^k) = \frac{1}{L} \sum_{k=1}^{L} (x_i^k - \mu(x_i))^2 , \qquad (3)$$

where L is training data set dimension.

After training the neural network model have ability to transient ischemic attacks recognition.

6 NPCA Neural Network

Let’s consider an autoencoder, which is also called recirculation or replicator neural network as it is shown in Fig. 2. It is represented by multilayer perceptron, which performs the nonlinear compression of the dataset through a bottleneck in the hidden layer. As we can see the nodes are partitioned in three layers. The bottleneck layer performs the compression of the input dataset.

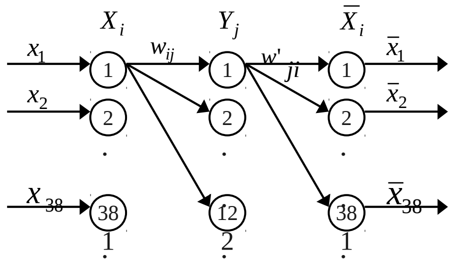


Figure 2: NPCA neural network.

After training the NPCA neural network can perform orthogonal compression of input data set.

7 Multilayer Neural Network

As it has been mentioned before the architecture of the neural network for TIA recognition used in this paper is the multilayered feed-forward network with 12 input units, 5 hidden units and 4 output units. The activation function for each unit of hidden and output layers is sigmoid function. The number of input neurons corresponds to dimension of compressed data and the number of output units corresponds to number of classes TIA and normal state. The number of hidden units was defined by experimental way. The backpropagation algorithm is used for training multilayer perceptron. Output value of a neural network is the number in a range from 0 up to 1 which characterizes probability of diagnostics for corresponding class of TIA.

8 Results and Discussion

For training and testing proposed neural network model the clinical observations of 114 patients with 38 parameters have been used. The clinical data set contain 28 patterns of TIA1, 25 patterns of TIA2, 27 patterns of TIA3 and 34 patterns of normal state without TIA. At the beginning the experiments with NPCA neural network have been performed using backpropagation algorithm together with the Gram-Schmidt procedure.

Let’s consider the mapping of input space data for normal state and TIA classes of attack on the plane of two and three principal components (Fig.3). As can be seen from the Fig. 3 the data, which belong different types of attacks are located in compact areas.

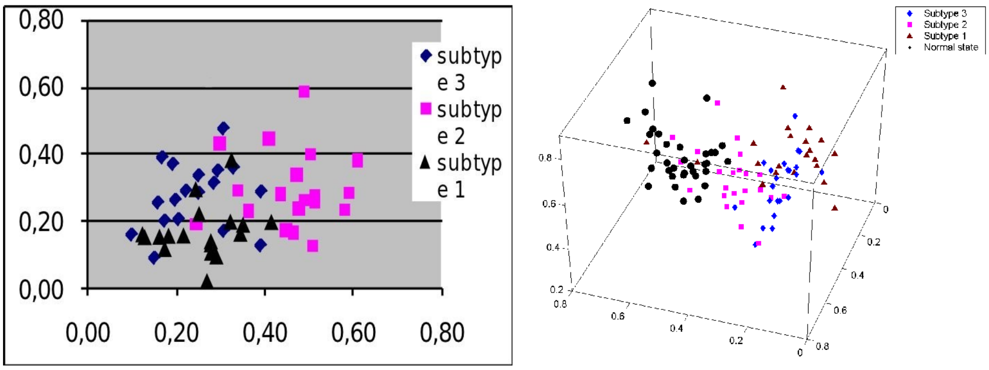
The all data set have been divided into 2 groups: the training data set and testing data set. The recognition accuracy is 100% using training data set and 78% using testing data set (Table 1).

9 Conclusion

In this paper the neural network model for transient ischemic attacks recognition have been addressed. The proposed approach is based on integration of the NPCA neural network and multilayer perceptron. The dataset from clinic have been used for experiments performing. Combining two different neural networks (NPCA and MLP) it is possible to produce efficient performance in terms of detection and recognition transient ischemic attacks. The main advantages of using neural network techniques are the ability to recognize “novel” TIA attack instances and quickness of work. Neural networks technique permits to reduce the diagnostic time and the number of misdiagnosis, as well as to assist the doctor in making decision.

Table 1: Recognition Accuracy

Number of patterns in training data set	**Number of patterns in testing data set**	**Recognition accuracy on training data set**	**Recognition accuracy on testing data set**
90	24	100%	78%



(a) (b)

Figure 3: Data Visualization on plane of two (a) and three (b) principal components.

References

[1] Paulo J. Lisba, Azzam Taktak. The use of artificial neural networks in decision support in cancer: A systematic review, *Neural networks* 19 (2006), p. 408-415.

[2] A. McNeill. How Accurate Are Primary Care Referral Letters For Presumed Acute Stroke? *Scottish Medical Journal*, V. 53, № 4, 2008. p. 11-12.

[3] J. Easton et all. Definition and evaluation of transient ischemic attack. *Stroke (Journal of the American Heart associations),* 2009.

[4] J. Dawson et all. A recognition tool for transient ischemic attack, *O J Med*, 192, p. 43-49, 2009.

[5] P. Lisba. A review of evidence of health benefit from artificial neural networks in medical intervention, *Neural networks* 15 (2002), p. 11-39.

[6]. R. Barnes, J. Toole, J. Nelson, V, Howard. Neural networks for ischemic stroke, *Journal of stroke and cerebrovascular diseases*, Vol.15, No.5, p. 223-227, 2006.

[7] D. Shanthi, Dr.G. Sahoo, Dr.N. Saravanan. Input feature selection using hybrid neuro-genetic approach in the diagnosis of stroke disease, *IJCSNS International Journal of Computer Science and Network Security*, Vol.8 No.12, December 2008, p.99-107.

[8] D. Shanthi, Dr.G. Sahoo, Dr.N. Saravanan. Designing an artificial neural network model for the prediction of thrombo-embolic stroke, *International journal of biometric and bioinformatics*, Vol.8, No.1,2008, p. 10-18.

An Artificial Immune System for Distributed Anomaly and Misuse Detection

Leanid Vaitsekhovich

Brest State Technical University,
Moskovskaja str. 267, 224017 Brest, Belarus
vspika@rambler.ru

Abstract: In this article the artificial immune system and neural network techniques for intrusion detection have been addressed. The AIS allows detecting unknown samples of computer attacks. The integration of AIS and neural networks as detectors permits to increase performance of the system security. The detector structure is based on the integration of the different neural networks namely RNN and MLP. The KDD dataset was used for experiments performing. The experimental results show that such intrusion detection system has possibilities for detection and recognition computer attacks.

1 Introduction

Different defense approaches exist to protect computer systems. All approaches can be divided into two main groups: organizational and technical. Technical approaches consist of network and hostbased techniques. In this article we will discuss network security tools namely intrusion detection systems.

The aim of *Intrusion Detection Systems (IDS)* is detecting inappropriate, incorrect or anomalous activity in computer systems or computer networks.

There are a lot of different means to protect computer networks: correct policy of security, gateway filters, anti-virus software etc. But as a rule IDS is assigned the role of a basic element of protection. IDS are used for early notification about network problems because generally they are allocated at a network level where suspicious actions can be found out earlier, then at higher levels. Besides IDS is able to gather necessary evidences of malicious activity as well as to reveal latent tendencies. This becomes possible due to analysis of plenty of the data.

The major problems of existing models are recognition of new attacks, low accuracy, detection time and system adaptability. The current anomaly detection systems are not adequate for real-time effective intrusion prevention. Therefore processing a large amount of audit data in real time is very important for practical implementation of IDS. It is difficult to eliminate stated disadvantages using only classical computer security methods. Therefore IDS have been closely studied recently. Researchers in this area have developed a variety of intrusion detection systems based on: statistical methods [1, 2], neural networks [3, 4], decision trees and SVMs [5], genetic algorithms and artificial immune systems [6, 7, 8, 9].

There exist two main intrusion detection methods: misuse detection and anomaly detection. *Misuse Detection* is based on the known signatures of intrusions or vulnerabilities. The main disadvantage of this approach is that it cannot detect novel or unknown attacks that were not previously defined. *Anomaly Detection* [10] defines normal behaviour and assumes that an intrusion is any unacceptable deviation from the normal behaviour. The main advantage of anomaly detection model is the ability to detect unknown attacks.

Researches of the natural mechanism of revealing of problems in the *Human Immune System (HIS)* can be used for building of an intrusion detection system owing to the fact that major principles of functioning are similar in both cases [11]. In the HIS the mechanisms of the nonspecific protection and the innate immunity realize the misuse detection function. The HIS consists of various immune cells, chemical signals, fibers etc. Their coordinated work allows to find out deviations in an organism of a person, to classify them and to start the mechanism of the immune answer. The properties of the distribution and self-organizing (adaptation to changing conditions), incorporated in the HIS, meet the basic requirements to systems of anomaly detection. Thus, modeling of the HIS includes development of algorithms of dynamic creation and updating of signatures, and also algorithms of anomaly detection by means of comparison to the current data.

In this work we propose our own solution of *Multi-agent Intrusion Detection System* that is a combination of *Artificial Immune System (AIS)* mechanisms and *Artificial Neural Networks (ANN)* to receive advantages from both approaches. We hope that such IDS will be able to perform dynamic anomaly detection and misuse detection in the real time mode.

This article is an extension of the previous work [12, 13, 14] associated with the development of intrusion detection system with the neural network classifier. Classification is the main problem in the intrusion detection domain.

The paper is organized as follows. The basic conception of an immune system functioning is given in Section 2. Section 3 deals with the neural network detector we use to build the multi-agent environment based on artificial immune system principles. Section 4 will briefly introduce the general ideas of the multi-agent modeling. In Section 5 the experimental results are described. Finally, concluding remarks are made in the last section.

2 Immune System

To begin with the Artificial Immune System for the Intrusion Detection domain, let's discuss how the Human Immune System works. This description will be simplified because the aim is to consider those basic elements that can be transferred to computer networks.

If one can say so, a major principle of the human immune system is a comparison of certain “patterns” with bodies located in a human organism. Thus we can reveal foreign bodes named antigens.

In real life lymphocytes carry out the role of the mentioned patterns. They are constantly generated by a spinal cord and thymus in view of the information contained in DNA (such information is accumulated, and this process is known as evolution of genic library). The lymphocytes spread in the organism through lymphatic nodes. Each type of the lymphocyte is responsible for detection of some limited number of antigens. There is an important stage during lymphocyte generation – *negative selection*. On this stage a special test on conformity with the native cells of the organism is executed. If such conformity takes place, the lymphocyte is killed. In fact otherwise it will struggle with the own cells of the organism. Due to the negative selection the “patterns” contain the information that is absent inside the organism. If any external body fit the given pattern than it is a foreign body.

In case of the lymphocytes detect an antigen on the ground of the corresponding pattern the new antibodies are produced and destroy the antigen. There is another mechanism that is known as *clone selection*. This mechanism is similar to the natural selection: only those antibodies survive that as much as possible correspond to the detected antigen. Thus the data on the generated antibodies get to the genic library mentioned above.

The most natural domain in which to begin applying the immune system mechanisms is computer security, where the analogy between protecting a body and protecting a normally operating computer is evident.

Experts working in the area of AIS mark out three fundamental properties of the human immune system:

- Firstly, it is distributed;
- Secondly, it is self-organizing;
- And thirdly, it is not especially demanding to computer resources.

In the opinion of many experts an efficient intrusion detection system should possess all of this properties.

3 Neural Network Detector

In our intrusion detection system artificial neural network detectors perform the function of lymphocytes in the HIS. *Neural Networks (NN)* have good generalization capabilities and can be efficiently used for approximation task, classification and processing of noisy data, what is especially important for intrusion detection.

We propose to use the integration of *PCA (Principal Component Analysis neural network)* and *MLP (Multilayer Perceptron)* as for basic element of the IDS (Fig.1).

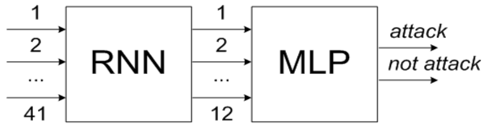


Figure 1: Detector for immune system construction

The 41 features from KDD dataset [15] are used for input data. These features contain the TCP-connection information. The PCA network, which is also called a *Recirculation Neural Network (RNN)*, transforms 41-dimensional input vectors into 12-dimensional output vectors. The MLP processes those given compressed data to recognize type of attacks or normal transactions.

Such a detector specializes in a certain type of attack. There are two output values "yes" (when the entrance pattern relates to the given type of attack) and "no" (when the entrance pattern is not attack of the considered type).

It is also possible to use detectors of another structure (Fig.2, for details see our previous works [12, 13, 14]) but in the article we will only refer to the detector shown in Fig.1.

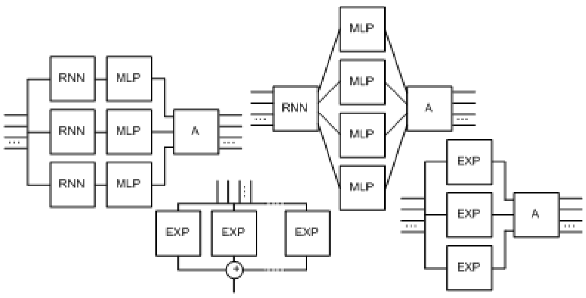


Figure 2: Other variants of detectors

After training of the neural networks they are ready to perform intrusion detection function.

4 Multi-agent System Based on Neural Network

The multi-agent system uses several neural network detectors that specialize different fields of knowledge.

Real immune systems are too complicated with a lot of complex protecting mechanisms. But we need not all of them. So constructing our own multi-agent system for intrusion detection we will use only the basic principles and mechanisms of the biological immune systems such as: generation and training of structurally diverse detectors, selection of appropriate detectors, ability of detectors to find out abnormal activity, cloning and mutation of detectors, forming of immune memory.

Let’s consider the generalized structure of the multi-agent IDS shown in Fig 3.

First of all, already known samples of normal network activity and attacks are placed into two databases of normal instances and attack instances respectively. Each sample is labeled either as an attack type or non-attack. These databases are used as a source for generation of training sets for the neural network detectors and for testing the performance of the IDS.

The next step includes creation of the detectors and the training procedure. Normal and attack samples are randomly selected from the mentioned above databases forming a training set for an immature detector.

The periodical testing phase is necessary to control current state of the IDS for revelation of the detectors failed to train (this detectors are immediately deleted from the system) and to calculate of efficiency parameters for each detector. We can use something like this expression shown in the simplified form as for the efficiency parameter: *EF = count_of_true_alarms – count_of_false_alarms.*

A collection of the immune detectors makes up a population that circulates in a computer system and performs recognition of network attacks. It is possible to generate hundreds and thousands of the detectors each of them is responsible for a definite attack type. Availability of various input instances and element of chance during the education stage provide large quantity of the structurally different detectors. During the network traffic scan each detector performs recognition of the input vector and an overall conclusion is reported to a human operator who decides whether there is a true anomaly.

Dynamic nature of the proposed intrusion detection system is provided by regular renovation of the detectors in the population. This is a result of continuous cloning and mutation procedures; updating the set of the detectors with new members and removal of inefficient or long-life detectors.

Samples selected for the training purposes greatly influence the results of training stage and neural network generalization abilities. So preparing different training sets we can change the detector behavior and its ability to recognize certain input instances. So we can use this property of neural networks for preparing detectors with various generalization capabilities.

The cloning procedure is equal to retraining the detector with the minimal efficiency rates on the training set of the detector with the maximum performance (both detectors have to specialize in the same attack type).

The mutation procedure is related with retraining of the randomly selected detectors from the population. Besides samples for the training are renewed. So the mutation introduces into the intrusion detection system an element of randomness.

Inefficient detectors and detectors with the completed lifetime parameter are removed from the system or replaced by new randomly created detectors.

However, if a detector achieves the highest values of the efficiency among the detectors specialized in the certain attack type, it enters the immune memory to reserve here its configuration parameters (in the case of neural network detectors – a vector of its weight coefficients). These memory detectors can be easily activated, for example, in the case when overall system performance will reduce greatly.

Each detector is represented by the artificial neural network consisted of the recirculation neural network and the multilayer perceptron, which functions were already discussed above.

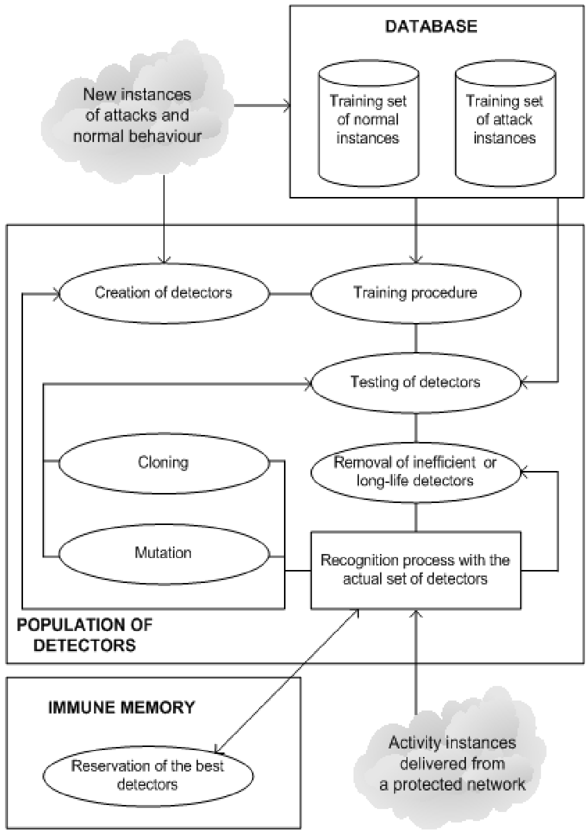


Figure 3: Simplified multi-agent IDS functioning

The detectors, which represent the same type of attack, are combined in groups from 3 to 10. Generally, all the detectors in the group give the diverse conclusions, which is the results of casual processes during the training. Theoretically, the number of detectors in the system is not limited and their number can be easily varied, but in real world problems with computational resources such as operative memory, speed etc…, arise.

Recognition process of an entrance pattern consists of the following sequence of steps:

1. Input pattern is transmitted to the multi-agent system.
2. Each detector gives a conclusion about entrance activity.
3. So-called *factor of reliability* on each group of the detectors is formed. This factor reflects percent of voices in the group, given for the type of attack the group is specialized in.

4. The analysis of factors of reliability, obtained from each group, is carried out. A decision of the group with the maximum value of the factor is considered to be the final decision.

After information about new attack have been received (from a network administrator or other sources) it is appended to the database and new group of the detectors specialized on this type of attack appears. Thus new data are involved into the system work.

The obvious advantage of such an approach is, (i) Training process is made comparatively easily; (ii) Detectors are trained on a smaller number of samples than models considered in the previous works; (iii) It allows to increase quality of their training and to considerably reduce time spent for preparation of the next detector.

5 EXPERIMENTAL RESULTS

In our work artificial immune system has been exploited for a development of multi-agent IDS.

Several important questions that strongly influence the efficiency of the model arise in the course of designing multi-agent structures: obtaining of the generalized decision on the basis of the set of detector opinions, selection of detectors, cloning and mutation, destruction of bad and/or irrelevant detectors.

There are a lot of random events during the multi-agent system functioning. So first of all it is necessary to be convinced of the IDS stable work. Look at Fig.4. Here we can see that the detection rate and the false positive rate for some testing set appear not to exceed certain boundaries during long time functioning.

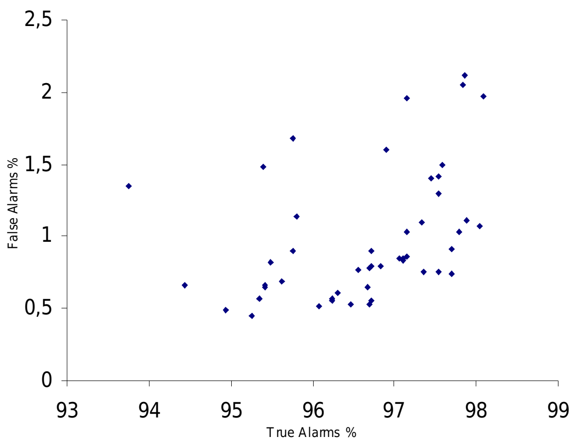


Figure 4: Demonstration of the multi-agent IDS stable work

Let's consider how such a multi-agent system works from an example of a population of detectors. The population consists of 110 detectors (5 detectors in a group for each attack type from the KDD dataset). The results (Table 1, 2) were prepared in the same way as for the models in our previous works [12, 13, 14] so that we can compare them easily.

Table 1: Training and testing sets

	DoS	U2R	R2L	Probe	Normal	Total count
training samples	3571	37	278	800	1500	6186
testing samples	391458	52	1126	4107	97277	494020

Table 2: Attack classification with the multi-agent IDS

class	count	detected	recognized
DoS	391458	386673 (98.78%)	368753 (94.20%)
U2R	52	47 (90.39%)	45 (86.54%)
R2L	1126	1097 (97.42%)	930 (82.59%)
Probe	4107	4066 (99.00%)	4016 (97.78%)
Normal	97277	---	82903 (85.22%)

The second experiment (Table 3) is related with the recognition of new attacks. For this purpose, we prepared a special set of samples for testing and training. The testing samples consist of network connection records that represent some of the most popular network services taken from the KDD dataset (http, ftp, ftp data, smtp, pop3, telnet). As dataset for training, we generated a considerably reducing number of samples for each attack type. Also what is necessary to draw attention is that the records of some scanty attack types were entirely excluded from the training set. Therefore, only 9 types of attacks have been selected here. Accordingly, 9 groups (5 detectors in each) have been generated. So, the quantity of the population has made up 45 detectors.

Table 3: Attack detection with the multi-agent IDS (Step 1)

type	count	detected
Normal	75952	75269 (99.10%)
Back	2203	2157 (97.91%)
Neptune	901	899 (99.78%)
Buffer_overflow	30	28 (93.33%)
Loadmodule	9	8 (88.89%)
Guess_passwd	53	52 (98.11%)
Warezclient	1015	966 (95.17%)
Ipsweep	9	9 (100.00%)
Portsweep	15	14 (93.33%)
Satan	10	8 (80.00%)

After several iterations had passed we added few instances of "warezmaster" attack to the database. As a result 5 additional detectors appeared in the population specialized in this attack type. The results taking after this manipulations are shown in Table 4.

Table 4: Attack detection with the multi-agent IDS (Step 2)

type	count	detected
Normal	75952	75169 (98.97%)
Back	2203	2174 (98.68%)
Neptune	901	900 (99.89%)
Buffer_overflow	30	26 (86.67%)
Loadmodule	9	8 (88.89%)
guess_passwd	53	53 (100.00%)
Warezclient	1015	947 (93.30%)
Warezmaster *	20	18 (90.00%)
Ipsweep	9	9 (100.00%)
Portsweep	15	14 (93.33%)
Satan	10	8 (80.00%)

* - *the attack that was added to the database*

The results shown in Table 5 show a lot of records corresponding to new attacks were detected and classified as an "attack". It means that multi-agent systems are capable of detecting new attacks and have high generalization capacity.

Table 5: Attack detection with the multi-agent IDS (Step 3)

type	count	detected
Normal	75952	74340 (97.88%)
Back	2203	2169 (98.46%)
Land*	1	1 (100.00%)
Neptune	901	900 (99.89%)
Buffer_overflow	30	26 (86.67%)
Loadmodule	9	9 (100.00%)
Perl*	3	0 (0.00%)
Rootkit*	7	3 (42.86%)
ftp_write*	6	5 (83.33%)
guess_passwd	53	53 (100.00%)
Multihop*	7	5 (71.43%)
Phf*	4	0 (0.00%)
Spy*	2	0 (0.00%)
Warezclient	1015	981 (96.65%)
Warezmaster	20	19 (95.00%)
Ipsweep	9	9 (100.00%)
Nmap*	2	2 (100.00%)
Portsweep	15	15 (100.00%)
Satan	10	8 (80.00%)

** - the attacks that were absent in the training set*

6 Conclusion

We have discussed only the prototype of multi-agent IDS that is based on artificial immune system and oriented to work on a single machine. Nevertheless, the results are promising due to the fact that many unknown records were detected. Extension of the proposed approach based on the multi-agent environment will allow us to build a real time intrusion detection system to protect local networks. Separate modules located on protected machines in different places of LAN will exchange data about their detectors what will make intrusion detection process more adaptable and quick reaction.

Bibliography

[1] D.J. Marchette. A statistical method for profiling network traffic // In Proceedings of the USENIX Workshop on Intrusion Detection and Network / 1999. – P.119–128.

[2] D.J. Marchette. Computer Intrusion Detection and Network Monitoring: A Statistical View-Point // Springer - 2001.

[3] J. Cannady. Artificial neural networks for misuse detection // In Proceedings of the 1998 National Information Systems Security Conference (NISSC'98) / October 5–8 - Arlington, VA, 1998. - P. 443-456.

[4] Beghdad R. Critical study of neural networks in detecting intrusions // Comput. Secur. – 2008. - doi:10.1016/j.cose.2008.06.001.

[5] Sandhya Peddabachigari, Ajith Abraham, Crina Grosan, Johnson Thomas. Modeling intrusion detection system using hybrid intelligent systems // Journal of Network and Computer Applications – 2007 – 30. – P.114–132.

[6] Morton Swimmer. Using the danger model of immune systems for distributed defense in modern data networks // Computer Networks – 2007. – 51. – P.1315–1333.

[7] Gerry Dozier, Douglas Brown, Haiyu Hou, John Hurley. Vulnerability analysis of immunity-based intrusion detection systems using genetic and evolutionary hackers // Applied Soft Computing – 2007. – 7. – P.547–553.

[8] Mohammad Saniee Abadeh, Jafar Habibi, Zeynab Barzegar, Muna Sergi. A parallel genetic local search algorithm for intrusion detection in computer networks // Engineering Applications of Artificial Intelligence – 2007. – 20. – P.1058–1069.

[9] M. Saniee Abadeha, J. Habibia, C. Lucasb. Intrusion detection using a fuzzy genetics-based learning algorithm // Journal of Network and Computer Applications – 2007. – 30. – P.414–428.

[10] Animesh Patcha, Jung-Min Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends // Computer Networks – 2007. – 51. – P.3448–3470.

[11] U Aickelin, P Bentley, S Cayzer, J Kim, J McLeod. Danger Theory: The Link between AIS and IDS? // In Proceedings of ICARIS-2003, 2nd International Conference on Artificial Immune Systems / P.147-155.

[12] V. Golovko and L. Vaitsekhovich. Neural Network Techniques for Intrusion Detection // In Proceedings of the International Conference on Neural Networks and Artificial Intelligence (ICNNAI-2006) / Brest State Technical University - Brest, 2006. - P. 65-69.

[13] V. Golovko, P. Kachurka and L. Vaitsekhovich. Neural Network Ensembles for Intrusion Detection // In Proceedings of the 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS-2007) / Research Institute of Intelligent Computer Systems, Ternopil National Economic University and University of Applied Sciences Fachhochschule Dortmund - Dortmund, Germany, 2007. - P. 578-583.

[14] V. Golovko, L. Vaitsekhovich, P. Kochurko and U. Rubanau. Dimensionality Reduction and Attack Recognition using Neural Network Approaches // Proceedings of the Joint Conference on Neural Networks (IJCNN 2007) / Orlando, FL, USA – IEEE Computer Society, Orlando, 2007. - P. 2734-2739.

[15] KDD Cup Competition. - Information on: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.