BALTIC CONFERENCE

Learning in Networks

Kaunas, 20.-21.08.2010

Universität Rostock 2010

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

Erstellung der Druckvorlage: Jens Pankow

Entwurf des Umschlagbildes: Christine Bräuning

ISBN:

Bezugsmöglichkeiten:

Universität Rostock
Institut für Informatik
Frau Kerstin Krause
Albert-Einstein-Str. 21
18059 Rostock

Universität Rostock
Wissenschaftsverbund IuK
Frau Dr. Christine Bräuning
Albert-Einstein-Str. 21
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 sixth year, with the associated conference going into its fourth year.

It was in 2005, that the University of Bremen, the University of Luebeck, the ISNM - International School of New Media at the University of Luebeck, 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 (DAAD - Deutscher Akademischer Austausch Dienst), 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 program, 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 currently we are planning 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.

The present proceedings give proof of the research results submitted by the participants and lecturers of BaSoTI 6.

Clemens H. Cap Rostock, December 2010.

Programme Committee

Prof. Dr. Andreas Ahrens (University of Wismar)
Prof. Egidijus Kazanavicius (Kaunas University of Technology)
Prof. Sr. Stephan Lukosch (Delft University of Technology)
Prof. Dr. Alke Martens (University of Rostock)
Dr.-Ing. Peter Sobe (University of Lübeck)
Prof. Dr. Clemens Cap (University of Rostock)
Dr.-Ing. Andreas Könsgen (University of Bremen)
Dipl. Inf. Dennis Maciuszek (University ofRostock)
Dr.-Ing. Thomas Mundt (University of Rostock)

Ontology based configuration of virtual systems in a computer cloud

Andrejs Ābele

Department of Management Information Technologies
Riga Technical university, Faculty of Computer Science and Information Technology
1/3 Meža iela
Latvia, Riga LV-1048
andrejs@iti.rtu.lv

Abstract: Multiple combinations of hardware and network components can be selected to design an Information Technology (IT) infrastructure that satisfies requirements. This paper proposes depreciation a framework which helps to create an efficient virtual infrastructure and keep it up-to-date. The framework monitors changes in configuration of the virtual infrastructure and uses gathered data for generating infrastructure optimization recommendations. The ontology, which classifies virtual machines is the key part of the proposed framework. The structure of this ontology is outlined in the paper and an application example is presented.

1 Introduction

Investment in Information Technology (IT) infrastructure including hardware, networks and systems software are expensive and have high level of depreciation. Design and optimum sizing of IT systems is the result of balancing several conflicting requirements, including technical performance and cost, organization impact, and user acceptance [AF00][AF01]. Theoretical research often focuses on two design and sizing problems. The first problem is distribution of the overall computing load of a system onto multiple machines in order to minimize hardware costs. The second problem is where to locate machines that need to exchange information in order to minimize network costs [AFT02]. Virtualization [V07] enables resources sharing and scalability thus reducing the overall cost of IT infrastructure. The cloud increases scalability even further by enabling quick response to workload changes. It also eliminates dependency from the physical location of computing resources.

There are three types of services computer clouds provide. First one is Software as a service (SaaS). With SaaS, a single application is delivered to thousands of users from vendor's servers. Customers do not pay for owning the software rather for using it

[M05]. Second one, Platform as a Service (PaaS), is a complete platform, including application development, interface development, database development, storage, testing, and so on, delivered through a remotely hosted platform to subscribers. Based on the traditional time-sharing model, modern platform-as-a service providers provide the ability to create enterprise-class applications for use locally or on demand for a subscription price or for free [L04]. Third one, Infrastructure as a service (IaaS), delivers basic storage and compute capabilities as standardized services over the network. Servers, storage systems, switches, routers, and other systems are pooled and made available to handle workloads that range from application components to high-performance computing applications [R06].

A single cloud provides services to many users who deploy their virtual infrastructure in the cloud. Each user aims to optimize performance (e.g., speed, reliability) of her virtualized IT infrastructure and to utilize cloud's resources in the most efficient way (i.e., to reduce usage charges). Additionally, the virtual infrastructure in the cloud is updated frequently to take advantages of newly available technologies and releases. The cloud service provider has an opportunity to capture these data. The data can be provided back to users as suggestions for optimizing their virtual infrastructure.

The objective of this paper is to elaborate a framework for accumulating and systemizing clouds usage data and using these data in optimization of virtual infrastructure. The framework uses an ontology web language to classify virtual machines and appliances (similarly as products are classified in [YMWZ09]), an agent which collects system performance statistics and fuzzy logic to cluster collected statistics and update the ontology based on the results. The paper reports research in progress, and at its current stage the emphasis is on overall framework and outline of the ontology.

2 Framework

The main motivation behind the framework is that in a cloud virtual machines do not have a long life span. If a VM gets out-dated then the hole machine gets replaced with a newer version based on a predefined configuration pattern.[S08] Based on frequent VM changing frameworks optimizations method becomes meaningful. If VM would stay the same for long time, there wouldn't be possible to update ontology based on VM performance and because of that there wouldn't be created any new virtual machines combinations.

The framework (Figure 1) includes eight main parts: 1) Specifications of requirements, 2) VM Classifier, 3) VM Component repository, 4) OWL (Web Ontology Language) Register, 5) VM template and VM appliances template repository, 6) VM classifier based on system performances, 7) System performance statistics collecting agent 8) Cloud.

2.1 Specifications of requirements

To create virtual machine Templates or virtual appliance Templates, it is necessary to specify parameters of virtual machines. Specification of requirements consists of four possible choices:

- Chose VM template from repository – This means, that a template of a virtual machine is chosen from the repository. The main idea is to collect a large collection of virtual machine templates, so users could find best solution, without the need to create own template. It will be possible to modify existing templates, but the idea is, that user should be able to find a ready template which fulfils their needs.

- Chose VM appliance template from repository – As VM appliance here is meant combination of virtual machines that fulfils a certain task or combinations of tasks. You could say that VM appliance is a template of a systems infrastructure. As in previous point the idea is to collect so many different templates, as possible, so an infrastructure creation would be made as easy as possible.

- Combine VM manually – here, a possibility is given to combine a virtual machine, using components from VM component repository, which fulfils specific requirements. Of course there are some basic requirements that need to be fulfilled, so a VM could be created. For example, it must have an operating system, processor, RAM, etc.

- Combine VM appliance manually – Combine VM appliance can implement usage of ready VM templates and manually created. In this stage it is only possible to define what VM will be in an appliance and what will they contain, but configuring interaction between them will be possible only in later stages.

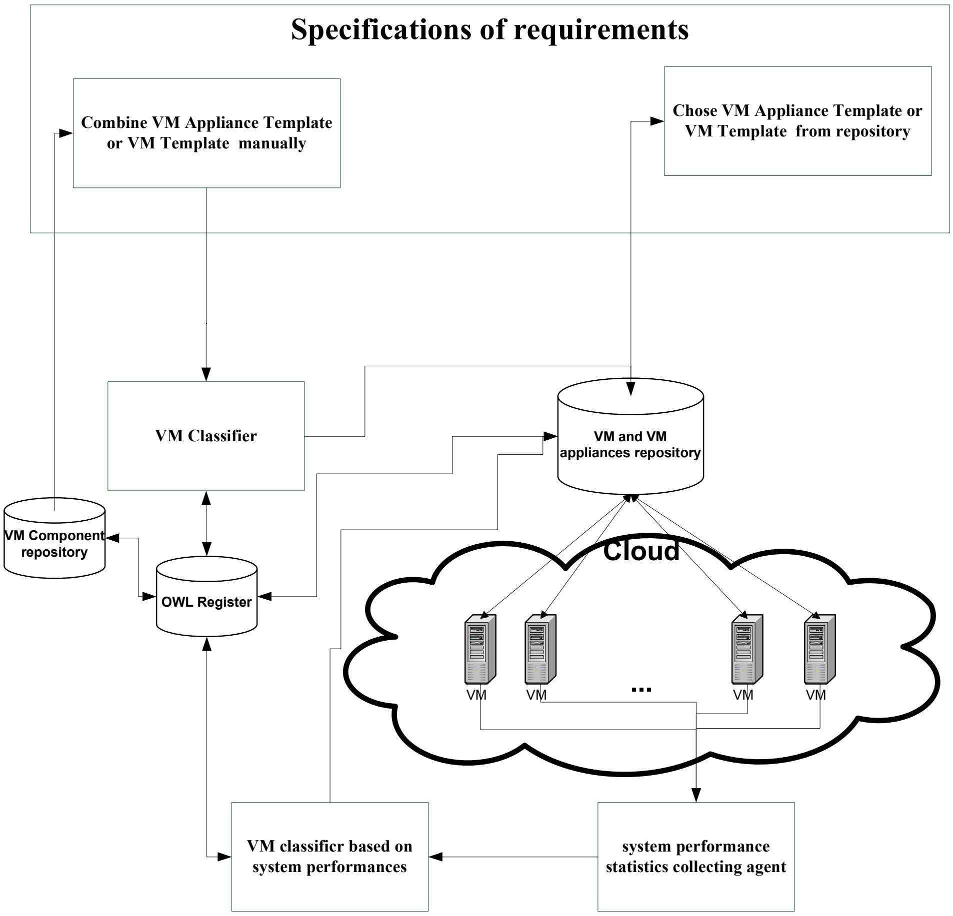

Figure 1: Framework

It is also possible to choose a ready template as a basis for manual configuration of VM, but that is different for VM appliances as VM in a VM appliance templates have been configured to work together, which is impossible in this stage.

2.2 VM Classifier

To compare performance of multiple VM, it is necessary to know which systems are similar and perform similar tasks. Therefore, it is necessary to classify newly created VM's and VM appliances. This classifier classifies VM's based on their architecture. The ontology is used for classification purposes. When VM is classified, its template is added to the VM repository. The ontology is described in the third section of this paper.

2.3 VM Component repository

This repository consists of components which can be installed automatically without human interaction. It contains components which represent hardware, as processor type, amount of RAM, HDD capacity. It also contains some software packages which can be installed automatically. The repository is connected to the OWL register, if the register wouldn't contain all the components, classifier wouldn't be able to classify the VM.

2.4 OWL Register

The OWL register contains the ontology which uses the OWL web ontology language. This ontology contains descriptions of every VM component, every VM machine an appliances. Purpose of this ontology is to help select and overview MV templates. This ontology is elaborated in the third section of this paper.

2.5 VM and VM appliances repository

The VM repository contains VM templates and VM appliances templates. It is interacting with the OWL register because all templates must be registered in the OWL register, otherwise it would be impossible to use those templates. Because the templates are just a description of VM configuration, so it needs actual VM components which it can find through OWL register. It is VM classifiers task to make sure that the VM repository and the OWL register are in sync.

2.6 System performance statistics collecting agent

This daemon collects information about VM performance. It is necessary, so it would be possible to determine which VM configuration is the most effective. Daemon monitors not only individual VM performance, but also VM appliances. From individual VM it collects – CPU data, file system disk usage, attached HDD usage, amount of swap space, processors and associated threads, system load, memory allocation. To determine appliances efficiency daemon collects – network interface activity, network latency, DB activity.

2.7 VM classifier based on system performances

This classifier classifies VM based on their performance on contrary to classifier which vas discus in Section 2.2 of this paper where VM's where classified based on their architecture. This classifier deals with QoS [S07][ZHP10]. As the range of data changes it is hard to maintain static limits for classification, so it is necessary to use Fuzzy logic [ZTJ11] [GGZ03].

2.8 Cloud

Cloud is where all the physical resources are located. It ensures efficient execution of VM by assigning necessary computing resources with limits and according to conditions specified in the hosting agreement. The proposed framework enables consumers to select and to configure their VM in order to maximize performance and to minimize hosting expenses by requesting only as much resources as they need.

3 Ontology

The Ontology is used to classify VM deployed in the cloud. This classification enables identification of VM providing similar services and accumulation of VM performance data. This ontology is developed using OWL DL. All VM components are defined as classes in the ontology (Figure 2). Newly created VM templates are defined as instances. If an instance, which does not belong to any ground level category, is created, then a new category is created. . Only an "is-a" relationship is visualized in the figure. There are also properties like "hasOS", "hasServices", "hasVMtipe", "worksWith". Below are written some logical conclusions which using an OWL reasoner can be deducted from this ontology.

1. Every VM is something
   that hasOSes exactly 1 Operating_system
   and
   that hasServiceses a Services
   and
   that hasVMtipes a VM_Tipe .
2. Every DB_SERVER is a VM
   that hasServiceses a Database
3. Every WEB_SERVER is a VM
   that hasServiceses a Web_server
4. Every LinuxWEBserver is a WEB_SERVER
   that hasOSes nothing but Linuxes .
   Every WEB_SERVER
   that hasOSes nothing but Linuxes is a LinuxWEBserver
5. Every Ubuntu is a Linux.

At the beginning reasoning is described that a VM must have one operating system, must have a service chosen (for now are available – web server, DB server, firewall) and virtualizations type. The second reasoning step describes a database server. A DB server is every VM that has a database service. The third reasoning step is similar to second, just with web server. The fourth reasoning step explains that Linux web server is a web server has operating system only Linux and that web servers that have Linux are Linux web servers. The fifth reasoning step shows that every Ubuntu is a Linux operating system.

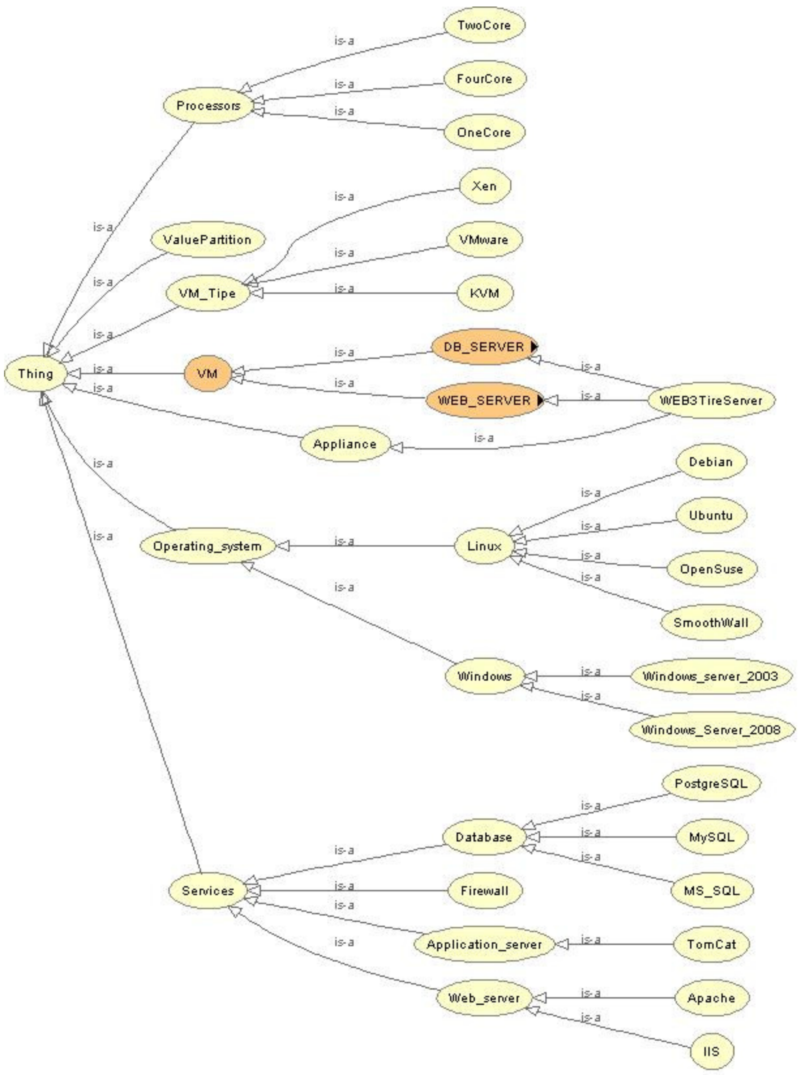

Figure 2: Ontology tree

4 Example

In this section, are demonstrated basic possibilities of the ontology. The example shows how manually configured VM are recognised and classified by the OWL reasoner.

A VM web server is defined (Figure 3. A). It has Ubuntu operating system and Apache web server installed. After activating the OWL reasoner this VM is recognised as LinuxWEBserver (Figure 3. B). Now we will try to create a VM which contains operating system Ubuntu and has VM type VMware, but reasoner does not categorises it (Figure 4. A), because it doesn't fulfil first condition, which was discussed in 3. section, it doesn't have a service. After a service MySQL is added, the reasoner recognises this VM as LinuxDBserver (Figure 4. B), which is a subcategory of VM.

Figure 3: Web server (A, B)

Figure 4: DB server (A, B)

5 Conclusions and Future Work

This paper presents the framework that will facilitate virtual IT system creations in a computer cloud. This framework allows using usage data to optimize performance of virtual infrastructure. Users are motivated to share data by getting the feedback and provider obtains a competitive advantage to other providers by providing an additional service to customers. Ontology is the key part of the framework; it allows recognizing the VM using reasoners. Because requirements frequently changes, ontology is subject to continuous refinement. Another key part is the performance classifier, which is the subject of further research. And main goal for the future is to create a working prototype where all parts work together.

References

[AF00] Ardabna,D; Francalanci, C:A Cost-Oriented methodology for the design of Web based IT architectures. SAC, Madrid, 2002

[AF01] Ardabna,D; Francalanci, C:A cost-oriented approach for the design of IT architectures. Journal of Information Technology, 2005

[AFT02] Ardabna,D; Francalanci, C; Trubian, M: Joint Optimization of Hardware and Network Costs for Distributed Computer Systems. IEEE Transactions on system, man, and cybernetics-part a: Systems and humans, vol 38, No. 2. 2008

[R06] Reese,G: Cloud Application Architectures: Building Applications and Infrastructure in the Cloud (Theory in Practice (O'Reilly)), O'Reilly Media, 2009

[YMWZ09]Yang, D; Miao, R.; Wu, H; Zhou, Y: Product configuration knowledge modeling using ontology web language. Expert Systems with Applications 36, 2009

[ZTJ11] Zhuang, L; Tao, H; JuanJuan, Z: Solving Multi-objective and fuzzy multi-attributive integrated technique for QoS-Aware Web Service Selection. Networking and Mobile Computing, pp.735–739, 2007

[ZHP10] Zhiyong C.; Haiyang W.; Peng P.: An Approach to Optimal Web Service Composition Based on QoS and User Preferences. International Joint Conference on Artificial Intelligence, Pasadena, 2009

[GGZ03] Guan Y.; Ghose A.; Lu Z.: Using constraint hierarchies to support QoS-guided service composition, International Conference on Web Services (ICWS '06), Chicago, 2006

[L04] Linthicum, D.: Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide. Addison-Wesley Professional. 2009

[S07] Stantchev V.: Performance Evaluation of Cloud Computing Offerings. 2009 Third International Conference on Advanced Engineering Computing and Applications in Sciences, Sliema, 2009

[M05] Miller, M.: Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online. Que, 2008

[S08] Sun Microsostems: IntroductIon to Cloud Computing ArChiteCture – Withe Pappre, Sun Microsystems, 2009

[V07] VMware :Understansing Full Virtualization, Paravirtualization, and Hardware Assist – White Papper, VMware, inc., 2007

Security practices of mobile phone student users in Šiauliai University (LT)

Iosif Androulidakis
Jožef Stefan International Postgraduate School
Jamova 39, Ljubljana SI-1000, Slovenia
sandro@noc.uoi.gr

Gorazd Kandus
Department of Communication Systems, Jožef Stefan Institute
Jamova 39, Ljubljana SI-1000, Slovenia
gorazd.kandus@ijs.si

Abstract: The present paper presents the results of a survey about users' security practices regarding mobile phone usage that took place in Lithuania's Šiauliai University in February 2010. Following the methodology of a previous similar survey in the University of Ioannina in Greece [An09b] we repeated the process in an extended pool of respondents reaching 759 answers. The main research hypothesis validated was that users are unaware of the necessary measures to avoid a possible unauthorized access and/or sensitive data retrieval from their phones and that they lack proper security education. Results further showed that users feel that mobile phone communication is not secure. It is unquestionable that since users fail to secure their phones, industry and academia should proceed to educating them and designing better user interfaces in order to mitigate the dangers.

Keywords: mobile phone security, security practices, user interface security, questionnaire survey, mobile phone usage.

1 Introduction

With the increasing momentum of wireless technologies, mobile phone penetration and the m-commerce growth [ITU09], it is evident that mobile devices are becoming a critical component of the digital economy, a style statement and useful communication device, a vital part of daily life for billions of people around the world [Wr10]. Used for personal entertainment or business purposes, the mobile phone has made a huge difference in how we do things nowadays. Modern mobile phones' enhanced capabilities allow them to be almost as versatile as a computer becoming a valuable business (mobile applications) and entertainment tool (mobile games, m-commerce).

At the same time users store more and process more and more data including sensitive information in their phones (e.g. private life photos shot by the phone's internal camera or credit card numbers and PINs). A few years ago the only concern of a mobile phone user would be his communication privacy. This is not the case anymore. Users have to be protected from unauthorized third party access to their data. It is logical that apart from the traditional security measures such as PIN usage and voice encryption, users have to take extra security measures and to follow new best practices. Unfortunately they fail to do so as the results of this paper clearly show. The main research hypothesis validated through the selection of specially crafted questions was that users are unaware of the necessary measures to avoid a possible unauthorized access and/or sensitive data retrieval from their phones and that they lack proper security education. We proceeded one step further, comparing the results to a previous survey in the University of Ioannina, Greece [An09b]. The University of Šiauliai was chosen as a typical Lithuanian University with many similarities to that of University of Ioannina, as to have the same distribution in the sample. However, the two cultures are obviously different and the important finding out of this comparison was that results are not "country" specific, but have the same more or less appearance among the two countries surveyed.

2 Related Work

Although there have been quite many theoretical studies concerning mobile services, a significant means for investigating and understanding users' preferences is asking their opinion via specific questioning techniques, as shown by several survey studies in this direction. The vast majority of these surveys indicate the growing importance of mobile phones in everyday life and the increased popularity of new features.

In any case, the security of mobile phones is proven not to be adequate in many research papers [RI02][An08][An09a]. There also exist several survey studies in this direction. Some of these surveys studies focus on mobile phone's security issues [AP08] while others focus on mobile phone services, touching also security issues [AA04][AA05a][VCS02][ABA07][ABA08].

A recent survey [AP08] published in November 2008 focused on mobile phones security issues and in which degree these issues concern the users. The conclusion was that a major part of the participants are extremely concerned about security and don't want any of their private data to be available to 3rd party unauthorized users. As mentioned earlier, following the same rationale and based on the results of a similar survey [An09b] that took place last year in a Greek University we used the technique in order to understand users' security practices and to possibly compare the results among the users of the two countries.

It is interesting to note that according to other surveys [ABA07][ABA08] a major part of the participants is interested in mobile services adoption only if the prices are low and the security framework is tight enough. This is why in the present paper we try to address users' security awareness and practices, as an enabler for greater mobile services market penetration.

3 Research Analysis and Results

3.1 Methodology

A very useful evaluation method for surveying user's practices is the use of multiple-choice questionnaires (i.e. in-person delivery or e-mail questionnaires) [Di99][PK01]. Our survey was conducted using the in-person delivery technique, with a total of 759 respondents participating in this survey. This method was selected instead of other alternatives because it is more accurate and has a bigger degree of participation from the respondents (e-mail questionnaires usually treated as spam mail from the respondents or they might misunderstand some questions). Data entry took place using custom software [AA05b].

The target group of the survey was university students from ages mostly 18-24 because these ages are more receptive to new technologies. They also understand better the technological evolution than older people who use mobile phones mostly for voice calls.

3.2 Survey Results

The questionnaire was divided in two parts. In the first part we asked the participants some demographic data including gender, age and field of studies as well as some economic data including mobile phone usage, connection type and budget spent monthly on phone service. In the second part we proceeded to our main contribution, the specific questions related with their practices and security perceptions regarding mobile phones' security issues.

3.2.1 Demographics

The participants were asked about their gender, age and field of studies. 56.5% of the participants were females and 43.5% were males while most of the respondents were aged 18-23 (73%). The main body of respondents was studying Engineering or Computer Science (43%) and one would expect these students to be more cautious regarding security, which was not the case as it will be shown later. Following in the sample there were equal parts of students of Mathematics or Natural Sciences (16%) and Economics and Business Administration (14.5%) with the rest of respondents in other fields.

Our fundamental research question was whether students are informed about how the options and the technical characteristics of their mobile phones affect the security of the latter and whether they are taking the necessary measures to mitigate the risks. The results that follow are totally in line with the initial response of students that only 13% believe they are high or very highy informed (Figure 1).

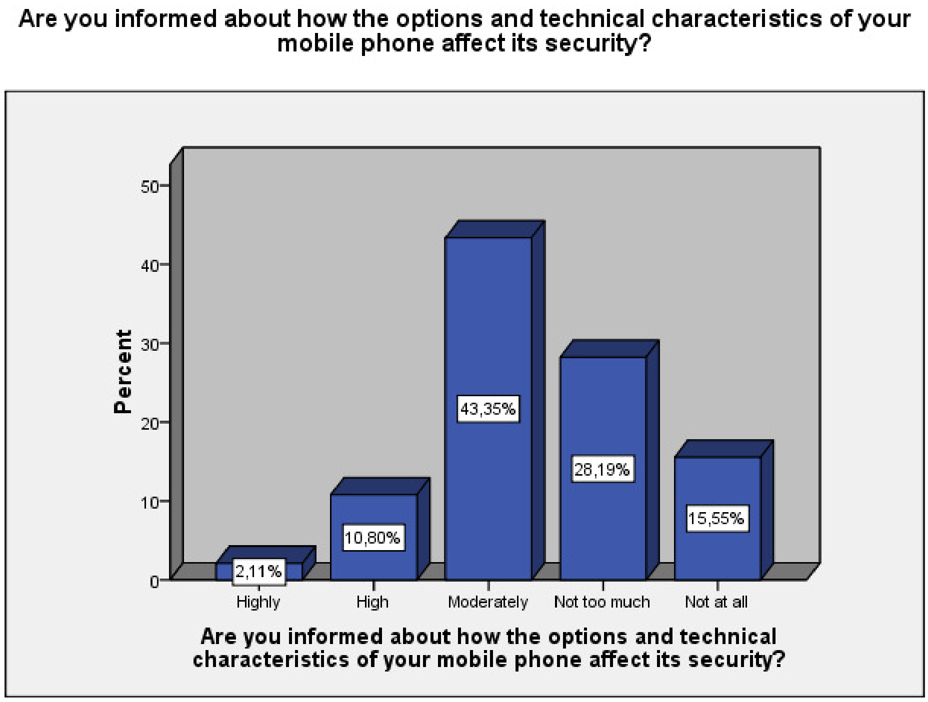

Figure 1: Knowledge of mobile phone security aspects.

Regarding mobile phone usage (Figure 2), almost 86% of them are using a single mobile phone daily, with some 8.5% using two phones regularly (compared to 34% of Greek students). Nokia is the favourite brand, reaching half of the students (51%) followed by Sony-Ericsson (22%) and Samsung (13%). Greek students preferred Sony Ericsson (46%) and then Nokia (26%) and Samsung (9.5%). Apple's iPhone seems to be scarce among students with just 4% of penetration (1.5% in Greece). It is immediately apparent that focusing on Nokia and Sony-Ericsson phones a security awareness campaign would immediately target almost 75% of users yielding a very high return. Such a campaign could be implemented in schools or universities, using social networks too. Of course the brand itself is not enough to categorize attack vectors and practices, since there is also the feature of the specific operating system running on each phone.

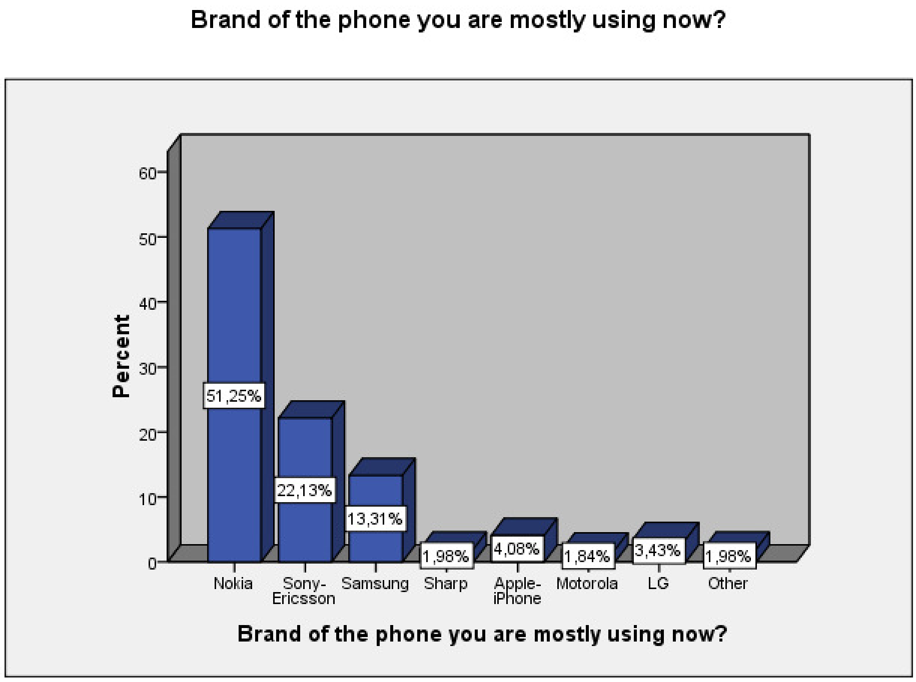

Figure 2: Favourite brands

3.2.2 Economics

Proceeding to economics, we asked participants whether they are using a pre-paid or post-paid (contract) mobile phone connection. We were surprised to find that 43% of students are using a contract based subscription, a rather high percentage. Usually students and young people go by pre-paid (card) subscriptions (i.e. 69% of Greek respondents had pre-paid connections).

Answering how much money they spent monthly, student mobile phone users have as it was expected limited budgets. More than 80% spend less than 20 Euros per month (currency converted) while most of them fall in the 11-20 Euros range. Greek students appear to spend slightly more, having 37% in the range of more than 20 Euros per month, compared to 19% of Lithuanians.

3.2.3 Security Specific Questions

The objective of this particular subsection and the main contribution of our research were to determine whether our participants acknowledge some security related features of their phone. The results are analyzed in the following paragraphs.

Starting with a general question about how "safe" mobile phone users feel, the majority (almost 60%) replied "moderately", with the rest of 40% equally distributed over and below that (Figure 3). Greek students on the other hand reach some 40% that are not too much or not at all sure that they are safe, appearing to be more mistrustful.

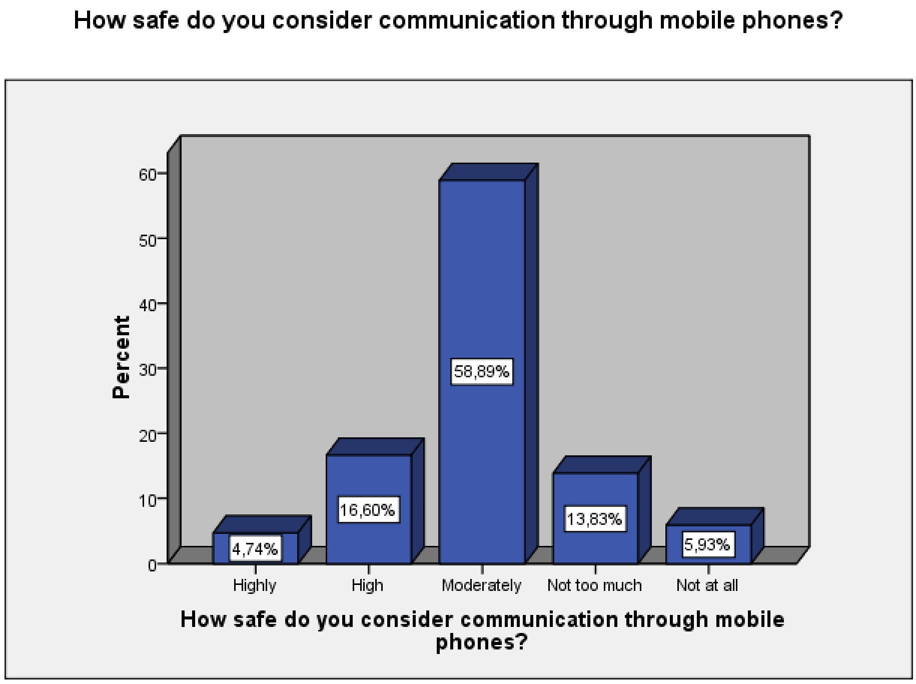

Figure 3: How safe do you consider communication through mobile phones?

A significant percentage of the participants (40%) doesn't know about the capabilities of his phone's operating system. Some 22% of students are using mobile phones with an advanced operating system, slightly more than Greeks (17%). In any case, the ignorance of the type of operating system renders users more vulnerable to hacker attacks with the use of exploits specifically targeted for their phones [An06a][An06b].

Similarly, only a very small percentage of the participants (less than 22%) knows his/her phone's IMEI (International Mobile Equipment Identifier – unique serial number of the mobile device) and has noted it somewhere. IMEI is very significant because if the phone is ever stolen, using this serial number the provider can block access to the stolen phone effectively mitigating stealing risks. Half of the students are completely unaware of its existence. Knowledge of this feature would help half of them (or 31% of Greeks) who unfortunately had their phone stolen once or more.

At the same time, just 17% (Figure 4) of users (in line with 15% of Greek students) are aware of the existence of the special icon that informs the user that his/her phone encryption has been disabled [An08]. Ignorance of this security icon leaves users vulnerable to man in the middle attacks since they can't recognize the attack taking place. In that scenario, a fraudster would impersonate the victim's network, using fake mobile base stations. So he would persuade the target's mobile phone to use this fake network, being thus able to intercept the calls [An08]. This was probably the most expected result as even professionals are not aware of this feature and another hint that user interfaces should help and not obscure security.

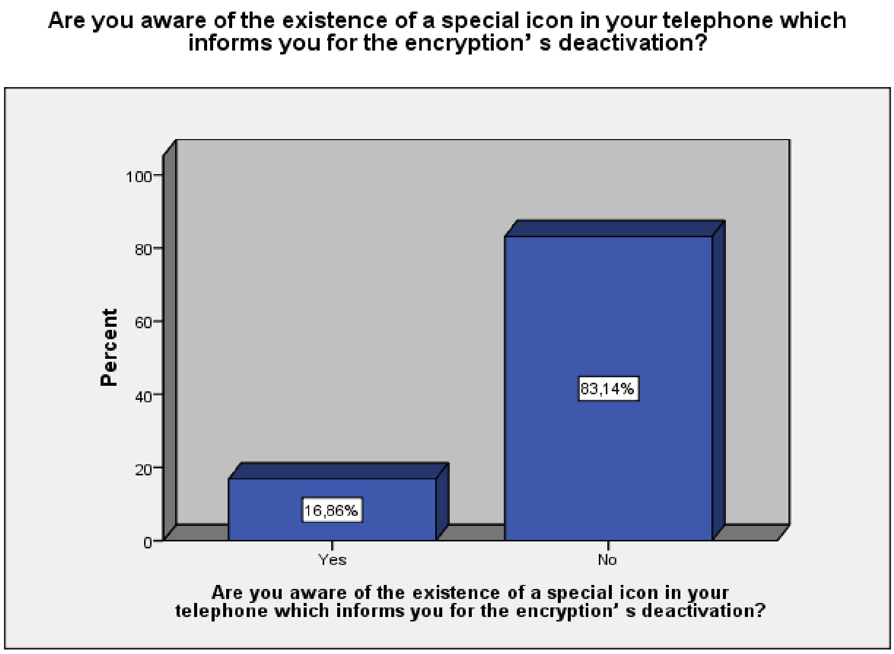

Figure 4: Encryption icon knowledge

Users, as expected, are actively (almost 80%) using SIM's PIN code. As a matter of fact, the same percentage of Greeks does exactly the same. The negative finding of the next question reveals that only a small percentage (16%) uses a screen-saver password while similar percentages do not know if their phone has such an option or are positive that their phone doesn't support the feature. That leaves 84% of users without a screen saver password, and their phones ready to be manipulated by "malicious" hands. An attack can take place in a few minutes by downloading specific software to the phone; this is why it is not enough to protect the phone only by PIN but also by a screen saver password.

A great attack vector of the past, Bluetooth, seems not to be the problem any more. Just one out of eight students has Bluetooth switched on and visible (leaving the phone vulnerable), while almost 70% of users have it switched off. It is not clear whether this is a security practice or a social practice that stemmed from the continuous harassements that messages over bluetooth caused upon users. Greek students are a little more tolerant since 20% of them are still waiting for Bluetooth messages (the rest 56% has it completely switched off).

In a question that touches upon issues of politeness and openness, 67% of students are lending their phones, but only while they are present. This is a major factor that compromises the phone's security even if the participant is present, because a single minute is needed for someone to install malicious software in the phone. In that respect 37% of Greeks refuses to lend their phone in any case, being better safe and "impolite" than sorry.

Following with a question of both security and economic importance, almost half of the participants don't download any software at all. There is also a 22% that actively downloads ringtones or logos, a 15% that tries applications and just 9% of "gamers". It is well interesting to note that the penetration is considerably higher than that of Greece's where 64.5% of users do not download at all (a somehow steady trend with results that haven't changed since a few years ago [AA04][AA05a]). In the antipode, getting familiar with downloading users are being more vulnerable to downloading and using unauthorised software that can harm their phone.

This is where a mobile phone Antivirus would help. It is sad to see that both Lithuanians and Greeks are still not using such products. In our case, 26% of users acknowledge there exists such a product but don't use it, while almost half of them (46.4%) do not know whether such a product exists. That leaves just a fraction of 4.5% using it. Compared with PC users where more or less users are educated and are using (at least) an antivirus shows a clear lack of security education and different mind-set.

Being young, 68% of university students (some 10% more than Greeks) keep sensitive information in their mobile phones. It seems that we consider our mobile phone to be a very personal device and we save equally important and sensitive information there. Such kind of information should be protected but again, the results from our survey show that users fail to do so. The consequences from a breach of data of this type could be devastating for the life of the victim.

In a rather positive finding, more than 65% of users are not saving important passwords in their phone. Some 21% are using some form of encryption (i.e. letter scrambling) while only 13.5% keep their passwords saved in plain. These results are almost identical to the ones we saw in Greece's survey. Since users follow the notion of encryption in these saved passwords, it is expected that they would be able to do the same with private information (i.e. photos) kept in the phone, should they be provided the necessary software. Once again, the issue of better designed user interfaces surfaces.

Closing our survey, we examined the issue of backup. As it was seen, a very large percentage of the participants surpassing 65% never performs a backup of their phone's data. One can argue that this was one of the most expected findings since even PC users don't actively back their data up.

Taking into consideration the answers given, we can deduce that users lack proper security education and are unaware of security measures and best practices, thus validating our research hypothesis.

4. Conclusions and future work

Conducting the survey in another country with a broader participating sample proved that the challenging findings of last year's survey were true. While the majority of the respondents care about security issues and are concerned about data interception and the fact that an intruder could gain unauthorized access to their devices, there is no culture of security and no advanced technical knowledge of their mobile phones. A very high percentage of users didn't know there is an icon that informs them about the phone encryption status. Most of them don't take backups at all while at the same time would lend their phone that contains sensitive data and passwords to somebody else. Contributing to the problem, badly designed interfaces are an additional factor of hindering the development of security culture.

In order to have comparative results, we have conducted a similar survey in more than 10 European countries reaching more than 7500 students and the results will soon be published. The preliminary findings however, show that users exhibit the same behaviour everywhere. Since students (who are young people and mostly receptive to technology and knowledge) do not actively follow most of security best practices, then academia, phone manufacturers and operators must team up informing users, raising awareness level and building more secure systems and user interfaces.

References

[AA04] Androulidakis N., Androulidakis I. m-Business: The base for creating competitive advantage. The case of Vodafone-Panafon, Wseas Transactions on Information Science and Applications, Issue 5, Vol 1, 1309-1313,2004

[AA05a] Androulidakis, N. and Androulidakis, I. Perspectives of Mobile Advertising in Greek Market, 2005 International Conference on Mobile Business (ICBM 2005), 2005

[Di99] Dillman, D. A. Mail and Internet Surveys: The Tailored Design Method, John Wiley & Sons, 2nd edition, November 1999

[NTT01] NTT DoCoMo. Current Trends in Mobile Phone Usage Among Adolescents, DoCoMo Report No. 10, March 2001

[PK01] Pfleeger, S. L. and Kitchenham, B. A. Principles of Survey Research Part 1: Turning Lemons into Lemonade ACM SIGSOFT Software Engineering Notes, vol. 26 (6), November 2001

[RI02] Rahman, M. & Imai, H., Security in Wireless Communication, Wireless Personal Communications [Online], vol. 22, issue, 2, pp.218-228,2002

[VCS02] Vrechopoulos, A.P., Constantiou, I.D. and Sideris, I. Strategic Marketing Planning for Mobile Commerce Diffusion and Consumer Adoption, in Proceedings of M-Business 2002, July 8-9, 2002

[AA05b] Androulidakis I., Androulidakis N. On a versatile and costless OMR system Wseas Transactions on Computers Issue 2, Vol 4, 160-165 (2005)

[An06a] I. Androulidakis, "Security Issues in Cell Phones", article-interview into Magazine "Defence and Diplomacy", Issue 187, pp 100-102, November 2006

[An06b] I. Androulidakis, "This is how hackers hack into our cell phones", article-interview into Sunday Newspaper "To proto thema", Issue 90, pp 40-41, November 12, 2006

[ABA07] I. Androulidakis, C. Basios, N. Androulidakis, Survey Findings towards Mobile Services Usage and M-Commerce Adoption, Proceedings of 18th European Regional ITS Conference, International Telecommunications Society, pp: CD-ROM, September 2007

[ABA08] I. Androulidakis, C. Basios, N. Androulidakis, Surveying Users' Opinions and Trends towards Mobile Payment Issues, Frontiers in Artificial Intelligence and Applications - Volume 169, pp. 9-19, 2008 (Techniques and Applications for Mobile Commerce - Proceedings of TAMoCo 2008)

[AP08] I. Androulidakis, D. Papapetros, Survey Findings towards Awareness of Mobile Phones' Security Issues, Recent Advances in Data Networks, Communications, Computers, Proceedings of 7th WSEAS International Conference on Data Networks, Communications, Computers (DNCOCO '08), pp 130-135, Nov. 2008

[An08] I. Androulidakis, Intercepting mobile phones and short messages, Computers and simulation in modern science, Selected papers from WSEAS conferences, Vol II, 2008, pp 320-321, Dec. 2008

[An09a] I. Androulidakis, Intercepting Mobile Phones, Article in «IT security professional» magazine, Issue 8, pp. 42-28, Jan-Feb 2009

[An09b] I. Androulidakis,V.Christou, N. Bardis, I. Stilios, Surveying users' practices regarding mobile phones' security features, Proceedings of 3rd International Conference on COMPUTATIONAL INTELLIGENCE (CI'09), pp 25-30, Jun. 2009

[Wr10] Richard Wray, In just 25 years, the mobile phone has transformed the way we communicate, guardian.co.uk, Friday 1 January 2010

[ITU09] Measuring the Information Society, The ICT Development Index, ITU 2009

APPENDIX: Questionnaire used

1) Male (A) or Female (B)?
2) Age? (A < 18, B 18-20 , C 21-23, D 24-26 , E >26)
3) Are you studying: (A: Humanities-Philology, B Medicine, C Law, D Engineering-Computer Science, E Maths-Natural Sciences, F Economics-Business Administration, G OTHER
4) How many mobile phones do you use (daily)?
A) 1B) 2 C) >2 D) None
5) Are you a contract subscriber or a prepaid subscriber?
A) Pre-paid (Card) B) Post-paid (Contract) C) Both
6) Your average monthly phone bill?
(A up to 10 Euros, B 11-20 Euros, C 21-30 Euros, D 31-40 Euros, E >40 Euros)
7) Brand of the phone you are mostly using now?
(A Nokia, B Sony-Ericsson, C Samsung, D Sharp, E Apple I-phone, F Motorola, G LG, H Other)
8) Does it have an advanced operational system (eg Symbian, Windows Mobile, Android)? (A I don't know, B yes, C no,)
9) Have you noted somewhere your mobile phone's IMEI?
(A, I don't know what it is, B yes, C no,)
10) Was your mobile phone ever lost or stolen? (A Never, B once, C more than once)
11) Are you aware of the existence of a special icon in your telephone which informs you for the encryption's deactivation? (A Yes, B No)
12) Do you have SIM card's PIN activated? (A Yes, B No)
13) Do you use password in your phone's Screen-Saver? (A I don't know if it has such a feature, B, doesn't have such feature, C, Yes, D No)
14) Do you have Bluetooth
(A Switched on and visible, B Switched on and invisible, C Switched off, D don't know the difference between visible and invisible, E My phone doesn't have Bluetooth,
15) Do you lend it to others? (A Never, B Only for a while and if I am present, C Yes)
16) Do you "download" software to your phone? (A I don't know if my mobile phone can download, B No, C mostly Ringtones/Logos, D mostly Games, E mostly Applications)
17) Do you use Antivirus software in your phone? (A Doesn't have the ability, B Don't know if there is such product for my phone, C I know there is but I don't use D Yes)
18) Do you store important passwords in your phone (eg Credit cards passwords, ATM passwords)? (A No, B Yes and "encrypted", C yes, without encryption)
19) How often do you create backup copies of your phone's data? (A Never, B >3 times per month, B 2-3 times per month, C Once per month, D Less often)
20) Do you keep sensitive personal data into your phone (photos/videos/discussion recordings)? (A Yes, B No)
21) How safe do you consider communication through mobile phones?
(A Highly, B High, C Moderately, D Not too much, E Not at all)
22) Are you informed about how the options and technical characteristics of your mobile phone affect its security? (A Highly, B High, C Moderately, D Not too much, E Not at all)

Performance Analysis of SVD-assisted Downlink Multiuser MIMO Systems

Sebastian Aust, Andreas Ahrens
Hochschule Wismar, University of Technology, Business and Design
email: andreas.ahrens@hs-wismar.de

César Benavente-Peces
Universidad Politécnica de Madrid
email: cesar.benavente@upm.es

Abstract: Multiuser multiple-input multiple-output (MIMO) downlink (DL) transmission schemes experience both multiuser interference as well as inter-antenna interference. Instead of treating all the users jointly as in zero-forcing (ZF) multiuser transmission techniques, the investigated singular value decomposition (SVD) assisted DL multiuser MIMO system takes the individual user's channel characteristics into account. This translates to a choice of modulation constellation and transmitter power and, in our proposed system, to a choice of number of activated user-specific MIMO layers. The performed joint optimization of the number of activated MIMO layers and the number of bits per symbol along with the appropriate allocation of the transmit power shows that not necessarily all user-specific MIMO layers has to be activated in both frequency-selective and non-frequency selective MIMO channels in order to minimize the overall BER under the constraint of a given fixed data throughput.

1 Introduction

Multiple-Input Multiple-Output (MIMO) systems are capable of increasing the achievable capacity and integrity of wireless systems and hence, they may be expected to form an integral part of next generation wireless systems [ZT03, ZVDL05]. However, single-user MIMO transmission schemes for both non-frequency and frequency selective MIMO channels have attracted a lot of attention and reached a state of maturity [KÖ6, AL08, ABP09b]. By contrast, MIMO-aided multiple-user systems require substantial further research where both multiuser as well as multi-antenna interferences has to be taken into account. Considering the entirety of the antennas of all mobile terminals at one end and the antennas of the base station at the other end of the communication link, state of the art interference cancellation is based on a central signal processing unit, e. g. a central unit at the base station, where joint detection can be applied in the uplink (UL) and joint transmission in the downlink (DL), respectively [MBW$^+$00, WMSL02, WSLW03, CM04, JUN05]. Widely used linear preprocessing techniques such as Minimum Mean Square Error (MMSE) or Zero Forcing (ZF) have attracted a lot of research and have reached a state of maturity [CM03]. In this work, a singular value decomposition (SVD) assisted downlink (DL) multiuser MIMO system is considered, which takes the individual user's channel characteristics into account rather than treating all users channels jointly as in ZF multiuser transmission techniques [LYH08]. Treating all user independently, adaptive modulation is a promising technique to increase the spectral efficiency of wireless transmission systems by adapting the signal parameters, such as modulation constellation or transmit power, dynamically to changing channel conditions. Therein, the most beneficial choice of the number of activated user-specific MIMO layers together with the number of bits per symbol and the appropriate allocation of the transmit power offer a certain degree of design freedom, which substantially affects the performance of MIMO systems in both frequency-selective and non-frequency selective MIMO links. Existing bit loading and transmit power allocation techniques are often optimized for maintaining both a fixed transmit power and a fixed target bit-error rate while attempting to maximize the overall data-rate [KRJ00, FH96, ZVDL05]. However for fixed-rate applications, such as video transmission schemes, it is desirable to design algorithms, which minimize the bit-error rate (BER) at a given fixed data rate.

Against this background, in this paper a SVD-assisted multiuser MIMO scheme is investigated, where both multiuser interferences as well as multi-antenna interferences are perfectly eliminated. The novel contribution of this paper is that we demonstrate the benefits of amalgamating a suitable choice of activated MIMO layers and number of bits per symbol along with the appropriate allocation of the transmit power under the constraint of a given fixed data throughput.

The remaining part of this paper is organized as follows: Section 2 introduces the system model. The considered quality criteria are briefly reviewed in section 3. The proposed solutions of bit and power allocation are discussed in section 4, while the associated performance results are presented and interpreted in section 5. Finally, section 6 provides some concluding remarks.

2 Flat Fading multiuser MIMO system model

The system model considered in this work consists of a single base station (BS) supporting K mobile stations (MSs). The BS is equipped with n_{T} transmit antennas, while the kth (with $k = 1, \ldots, K$) MS has $n_{\mathrm{R}\,k}$ receive antennas, i. e. the total number of receive antennas including all K MSs is given by $n_{\mathrm{R}} = \sum_{k=1}^{K} n_{\mathrm{R}\,k}$. The $(n_{\mathrm{R}\,k} \times 1)$ user specific symbol vector $\mathbf{c}_k$ to be transmitted by the BS is given by

$$\mathbf{c}_k = (c_{k,1}, c_{k,2}, \ldots, c_{k,n_{\mathrm{R}\,k}})^{\mathrm{T}} \quad . \tag{1}$$

The vector $\mathbf{c}_k$ is preprocessed before its transmission by multiplying it with the $(n_{\mathrm{T}} \times n_{\mathrm{R}\,k})$ DL preprocessing matrix $\mathbf{R}_k$ and results in the $(n_{\mathrm{T}} \times 1)$ user-specific transmit vector

$$\mathbf{s}_k = \mathbf{R}_k \, \mathbf{c}_k \quad . \tag{2}$$

After DL transmitter preprocessing, the n_T-component signal $\mathbf{s}$ transmitted by the BS to the K MSs results in

$$\mathbf{s} = \sum_{k=1}^{K} \mathbf{s}_k = \mathbf{R}\,\mathbf{c} \quad , \tag{3}$$

with the $(n_\mathrm{T} \times n_\mathrm{R})$ preprocessing matrix

$$\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K) \quad . \tag{4}$$

In (3), the overall $(n_\mathrm{R} \times 1)$ transmitted DL data vector $\mathbf{c}$ combines all K DL transmit vectors $\mathbf{c}_k$ (with $k = 1, 2, \ldots, K$) and is given by

$$\mathbf{c} = \left(\mathbf{c}_1^\mathrm{T}, \mathbf{c}_2^\mathrm{T} \ldots, \mathbf{c}_K^\mathrm{T}\right)^\mathrm{T} \quad . \tag{5}$$

At the receiver side, the $(n_{\mathrm{R}\,k} \times 1)$ vector $\mathbf{u}_k$ of the kth MS results in

$$\mathbf{u}_k = \mathbf{H}_k\,\mathbf{s} + \mathbf{n}_k = \mathbf{H}_k\,\mathbf{R}\,\mathbf{c} + \mathbf{n}_k \quad . \tag{6}$$

and can be expressed by

$$\mathbf{u}_k = \mathbf{H}_k\,\mathbf{R}_k\,\mathbf{c}_k + \sum_{i=1, i\neq k}^{K} \mathbf{H}_k\,\mathbf{R}_i\,\mathbf{c}_i + \mathbf{n}_k \quad , \tag{7}$$

where the MSs received signals experience both multi-user and multi-antenna interferences. In (6), the $(n_{\mathrm{R}\,k} \times n_\mathrm{T})$ channel matrix $\mathbf{H}_k$ connects the n_T BS specific transmit antennas with the $n_{\mathrm{R}\,k}$ receive antennas of the kth MS. It is assumed that the coefficients of the $(n_{\mathrm{R}\,k} \times n_\mathrm{T})$ channel matrix $\mathbf{H}_k$ are independent and Rayleigh distributed with equal variance. The interference between the different antenna's data streams, which is introduced by the off-diagonal elements of the channel matrix $\mathbf{H}_k$, requires appropriate signal processing strategies. A popular technique is based on the SVD of the system matrix $\mathbf{H}_k$. Upon carrying out the SVD of $\mathbf{H}_k$ with $n_\mathrm{T} \geq n_\mathrm{R}$ and assuming that the rank of the matrix $\mathbf{H}_k$ equals $n_{\mathrm{R}\,k}$, i. e., $\mathsf{rank}(\mathbf{H}_k) = n_{\mathrm{R}\,k}$, we get

$$\mathbf{H}_k = \mathbf{U}_k \cdot \mathbf{V}_k \cdot \mathbf{D}_k^\mathrm{H} \quad , \tag{8}$$

with the $(n_{\mathrm{R}\,k} \times n_{\mathrm{R}\,k})$ unitary matrix $\mathbf{U}_k$ and the $(n_\mathrm{T} \times n_\mathrm{T})$ unitary matrix $\mathbf{D}_k^\mathrm{H}$, respectively[1]. The $(n_{\mathrm{R}\,k} \times n_\mathrm{T})$ diagonal matrix $\mathbf{V}_k$ can be decomposed into a $(n_{\mathrm{R}\,k} \times n_{\mathrm{R}\,k})$ matrix $\mathbf{V}_{k\,\mathrm{u}}$ containing the non-zero square roots of the eigenvalues of $\mathbf{H}_k^\mathrm{H}\,\mathbf{H}_k$, i. e.,

$$\mathbf{V}_{k\,\mathrm{u}} = \begin{bmatrix} \sqrt{\xi_{k,1}} & 0 & \cdots & 0 \\ 0 & \sqrt{\xi_{k,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\xi_{k,n_{\mathrm{R}\,k}}} \end{bmatrix} \quad , \tag{9}$$

[1]The transpose and conjugate transpose (Hermitian) of $\mathbf{D}_k$ are denoted by $\mathbf{D}_k^\mathrm{T}$ and $\mathbf{D}_k^\mathrm{H}$, respectively.

and a $(n_{\mathrm{R}\,k} \times (n_{\mathrm{T}} - n_{\mathrm{R}\,k}))$ zero-matrix $\mathbf{V}_{k\,\mathrm{n}}$ according to

$$\mathbf{V}_k = (\mathbf{V}_{k\,\mathrm{u}}\,\mathbf{V}_{k\,\mathrm{n}}) = (\mathbf{V}_{k\,\mathrm{u}}\,\mathbf{0}) \quad . \tag{10}$$

Additionally, the $(n_{\mathrm{T}} \times n_{\mathrm{T}})$ unitary matrix $\mathbf{D}_k$ can be decomposed into a $(n_{\mathrm{T}} \times n_{\mathrm{R}\,k})$ matrix $\mathbf{D}_{k\,\mathrm{u}}$ constituted by the eigenvectors corresponding to the non-zero eigenvalues of $\mathbf{H}_k^{\mathrm{H}}\,\mathbf{H}_k$ and a $(n_{\mathrm{T}} \times (n_{\mathrm{T}} - n_{\mathrm{R}\,k}))$ matrix $\mathbf{D}_{k\,\mathrm{n}}$ constituted by the eigenvectors corresponding to the zero eigenvalues of $\mathbf{H}_k^{\mathrm{H}}\,\mathbf{H}_k$. The decomposition of the matrix $\mathbf{D}_k^{\mathrm{H}}$ results in

$$\mathbf{D}_k^{\mathrm{H}} = \begin{pmatrix} \mathbf{D}_{k\,\mathrm{u}}^{\mathrm{H}} \\ \mathbf{D}_{k\,\mathrm{n}}^{\mathrm{H}} \end{pmatrix} \quad . \tag{11}$$

Finally, the received downlink signal $\mathbf{u}_k$ of the kth MS may be expressed as

$$\mathbf{u}_k = \mathbf{U}_k\,\mathbf{V}_{k\,\mathrm{u}}\,\mathbf{D}_{k\,\mathrm{u}}^{\mathrm{H}}\,\mathbf{R}\,\mathbf{c} + \mathbf{n}_k \quad . \tag{12}$$

Taking all MSs received DL signals $\mathbf{u}_k$ into account, the $(n_{\mathrm{R}} \times 1)$ receive vector results in

$$\mathbf{u} = \left(\mathbf{u}_1^{\mathrm{T}}, \mathbf{u}_2^{\mathrm{T}}, \ldots, \mathbf{u}_K^{\mathrm{T}}\right)^{\mathrm{T}} \quad . \tag{13}$$

Then, the overall DL signal vector $\mathbf{u}$ including the received signals of all K MSs can be expressed by

$$\mathbf{u} = \mathbf{U}\,\mathbf{V}_{\mathrm{u}}\,\mathbf{D}_{\mathrm{u}}^{\mathrm{H}}\,\mathbf{R}\,\mathbf{c} + \mathbf{n} \quad , \tag{14}$$

with the overall $(n_{\mathrm{R}} \times 1)$ noise vector

$$\mathbf{n} = \left(\mathbf{n}_1^{\mathrm{T}}, \mathbf{n}_2^{\mathrm{T}}, \ldots, \mathbf{n}_K^{\mathrm{T}}\right)^{\mathrm{T}} \quad , \tag{15}$$

the $(n_{\mathrm{R}} \times n_{\mathrm{R}})$ block diagonal matrix $\mathbf{U}$

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{U}_K \end{bmatrix} \quad , \tag{16}$$

the $(n_{\mathrm{R}} \times n_{\mathrm{R}})$ block diagonal matrix $\mathbf{V}_{\mathrm{u}}$

$$\mathbf{V}_{\mathrm{u}} = \begin{bmatrix} \mathbf{V}_{1\,\mathrm{u}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{2\,\mathrm{u}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_{K\,\mathrm{u}} \end{bmatrix} \quad , \tag{17}$$

and the $(n_{\mathrm{T}} \times n_{\mathrm{R}})$ matrix $\mathbf{D}_{\mathrm{u}}$ which is given by

$$\mathbf{D}_{\mathrm{u}} = (\mathbf{D}_{1\,\mathrm{u}}, \mathbf{D}_{2\,\mathrm{u}}, \ldots, \mathbf{D}_{K\,\mathrm{u}}) \quad . \tag{18}$$

In order to suppress the DL multi-user interferences (MUI) perfectly, the DL preprocessing matrix $\mathbf{R}$ has to be designed to satisfy the following condition

$$\mathbf{D}_{\mathrm{u}}^{\mathrm{H}}\,\mathbf{R} = \mathbf{P} \ , \tag{19}$$

with the real-valued $(n_{\mathrm{R}} \times n_{\mathrm{R}})$ diagonal matrix $\mathbf{P}$ taking the transmit-power constraint into account. In order to satisfy (19), $\mathbf{R}$ can be defined as follows

$$\mathbf{R} = \mathbf{D}_{\mathrm{u}}\,\left(\mathbf{D}_{\mathrm{u}}^{\mathrm{H}}\,\mathbf{D}_{\mathrm{u}}\right)^{-1}\,\mathbf{P} \ . \tag{20}$$

Taking the ZF design criterion for the DL preprocessing matrix into account, the matrix $\mathbf{P}$ simplifies to an $(n_{\mathrm{R}} \times n_{\mathrm{R}})$ diagonal matrix, i. e. $\mathbf{P} = \sqrt{\beta}\,\mathbf{I}_{n_{\mathrm{R}} \times n_{\mathrm{R}}}$, with the parameter $\sqrt{\beta}$ taking the transmit-power constraint into account. When taking the DL preprocessing matrix, defined in (20), into account, the overall received vector of all K MSs, defined in (14), can be simplified to

$$\mathbf{u} = \mathbf{U}\,\mathbf{V}_{\mathrm{u}}\,\mathbf{P}\,\mathbf{c} + \mathbf{n} \ . \tag{21}$$

Therein, the $(n_{\mathrm{R}} \times n_{\mathrm{R}})$ block diagonal matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_K \end{bmatrix} \ . \tag{22}$$

In (21), the user-specific $(n_{\mathrm{R}\,k} \times 1)$ vector $\mathbf{u}_k$ can be expressed as

$$\mathbf{u}_k = \mathbf{U}_k\,\mathbf{V}_{k\,\mathrm{u}}\,\mathbf{P}_k\,\mathbf{c}_k + \mathbf{n}_k \ , \tag{23}$$

with the user-specific $(n_{\mathrm{R}\,k} \times n_{\mathrm{R}\,k})$ power allocation matrix

$$\mathbf{P}_k = \begin{bmatrix} \sqrt{p_{k,1}} & 0 & \cdots & 0 \\ 0 & \sqrt{p_{k,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_{k,n_{\mathrm{R}\,k}}} \end{bmatrix} \ . \tag{24}$$

As long as the transmit power is uniformly distributed over the number of activated MIMO layers, the matrix $\mathbf{P}_k$ simplifies to $\mathbf{P}_k = \sqrt{\beta}\,\mathbf{I}_{n_{\mathrm{R}\,k} \times n_{\mathrm{R}\,k}}$. After postprocessing of the received signal vectors $\mathbf{u}_k$ with the corresponding unitary matrix $\mathbf{U}_k^{\mathrm{H}}$, the user-specific decision variables result with $\mathbf{U}_k^{\mathrm{H}}\,\mathbf{n}_k = \mathbf{w}_k$ in

$$\mathbf{y}_k = \mathbf{U}_k^{\mathrm{H}}\,\mathbf{u}_k = \mathbf{V}_{k\,\mathrm{u}}\,\mathbf{P}_k\,\mathbf{c}_k + \mathbf{w}_k \ , \tag{25}$$

or alternatively with $\mathbf{U}^{\mathrm{H}}\,\mathbf{n} = \mathbf{w}$ in

$$\mathbf{y} = \mathbf{U}^{\mathrm{H}}\,\mathbf{u} = \mathbf{V}_{\mathrm{u}}\,\mathbf{P}\,\mathbf{c} + \mathbf{w} \ , \tag{26}$$

where interferences between the different antenna data streams as well as MUI imposed by the other users are avoided. The resulting system model is depicted in Fig. 1

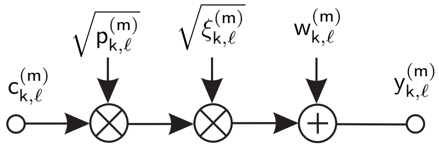

Figure 1: Resulting kth user-specific system model per MIMO layer ℓ (with $\ell = 1, 2, \ldots, n_{\mathrm{R}\,k}$) and per transmitted symbol block m

3 Theoretical Analysis of System Performance

In general, the user-specific quality of data transmission can be informally assessed by using the signal-to-noise ratio (SNR) at the detector's input defined by the half vertical eye opening and the noise power per quadrature component according to

$$\varrho = \frac{(\text{Half vertical eye opening})^2}{\text{Noise Power}} = \frac{(U_{\mathrm{A}})^2}{(U_{\mathrm{R}})^2} \quad , \tag{27}$$

which is often used as a quality parameter [AL08]. The relationship between the signal-to-noise ratio $\varrho = U_{\mathrm{A}}^2/U_{\mathrm{R}}^2$ and the bit-error probability evaluated for AWGN channels and M-ary Quadrature Amplitude Modulation (QAM) is given by [Pro00]

$$P_{\mathrm{BER}} = \frac{2}{\log_2(M)} \left(1 - \frac{1}{\sqrt{M}}\right) \operatorname{erfc}\left(\sqrt{\frac{\varrho}{2}}\right) \quad . \tag{28}$$

When applying the proposed system structure for the kth user, depicted in Fig. 1, the applied signal processing leads to different eye openings per activated MIMO layer ℓ (with $\ell = 1, 2, \ldots, L$ and $L \leq n_{\mathrm{R}\,k}$ describing the number of activated user-specific MIMO layers) and per transmitted symbol block m according to

$$U_{\mathrm{A}\,k}^{(\ell,m)} = \sqrt{p_{k,\ell}^{(m)}} \cdot \sqrt{\xi_{k,\ell}^{(m)}} \cdot U_{\mathrm{s}\,k}^{(\ell)} \quad , \tag{29}$$

where $U_{\mathrm{s}\,k}^{(\ell)}$ denotes the half-level transmit amplitude assuming M_ℓ-ary QAM, $\sqrt{\xi_{k,\ell}^{(m)}}$ represents the corresponding positive square roots of the eigenvalues of the matrix $\mathbf{H}_k^{\mathrm{H}}\,\mathbf{H}_k$ and $\sqrt{p_{k,\ell}^{(m)}}$ represents the corresponding power allocation weighting parameters (Fig. 1). Together with the noise power per quadrature component, introduced by the additive, white Gaussian noise (AWGN) vector $\mathbf{w}_k = \mathbf{U}_k^{\mathrm{H}}\,\mathbf{n}_k$ in (25), the kth user-specific SNR per MIMO layer ℓ at the time m results in

$$\varrho_k^{(\ell,m)} = \frac{\left(U_{\mathrm{A}\,k}^{(\ell,m)}\right)^2}{U_{\mathrm{R}}^2} \quad . \tag{30}$$

Using the parallel transmission over L MIMO layers, the overall mean transmit power becomes $P_{\mathrm{s}\,k} = \sum_{\ell=1}^{L} P_{\mathrm{s}\,k}^{(\ell)}$. Considering QAM constellations, the average user-specific

Table 1: Investigated user-specific QAM 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

transmit power $P_{\mathrm{s}\,k}^{(\ell)}$ per MIMO layer ℓ may be expressed as [Pro00]

$$P_{\mathrm{s}\,k}^{(\ell)} = \frac{2}{3}\left(U_{\mathrm{s}\,k}^{(\ell)}\right)^2 (M_{k\,\ell} - 1) \ . \tag{31}$$

Combining (30) and (31) together with (29), the layer-specific SNR at the time m results in

$$\varrho_k^{(\ell,m)} = p_{k,\ell}^{(m)}\,\xi_{k,\ell}^{(m)}\,\frac{3}{2\,(M_{k\,\ell} - 1)}\,\frac{P_{\mathrm{s}\,k}^{(\ell)}}{U_{\mathrm{R}}^2} \ . \tag{32}$$

Assuming that the transmit power is uniformly distributed over the number of activated MIMO layers, i. e., $P_{\mathrm{s}\,k}^{(\ell)} = P_{\mathrm{s}\,k}/L$, the layer-specific signal-to-noise ratio at the time m, defined in (32), results with the ratio of symbol energy to noise power spectral density $E_{\mathrm{s}}/N_0 = P_{\mathrm{s}\,k}/(2\,U_{\mathrm{R}}^2)$ in

$$\varrho_k^{(\ell,m)} = p_{k,\ell}^{(m)}\,\xi_{k,\ell}^{(m)}\,\frac{3}{L\,(M_{k\,\ell} - 1)}\,\frac{E_{\mathrm{s}}}{N_0} \ . \tag{33}$$

In order to transmit at a fixed data rate while maintaining the best possible integrity, i. e., bit-error rate, an appropriate number of user-specific MIMO layers has to be used, which depends on the specific transmission mode, as detailed in Table 1 for the exemplarily investigated two-user multiuser-system ($n_{\mathrm{R}\,k} = 4$ (with $k = 1, 2$), $K = 2, n_{\mathrm{R}} = n_{\mathrm{T}} = 8$). In general, the BER per spatial division multiplexing (SDM) MIMO data vector is dominated by the specific transmission modes and the characteristics of the singular values, resulting in different BERs for the different QAM configurations in Table 1. An optimized adaptive scheme would now use the particular transmission modes, e. g., by using bit auction procedures [WCLM99], that results in the lowest BER for each SDM MIMO data vector. This would lead to different transmission modes per SDM MIMO data vector and a high signaling overhead would result. However, in order to avoid any signalling overhead, fixed transmission modes are used in this contribution regardless of the channel quality. The kth user MIMO layer specific bit-error probability at the time m is given by

$$P_{\mathrm{e}\,k}^{(\ell,m)} = \frac{2\left(1 - \frac{1}{\sqrt{M_{k\,\ell}}}\right)}{\log_2(M_{k\,\ell})}\,\mathrm{erfc}\left(\sqrt{\frac{\varrho_k^{(\ell,m)}}{2}}\right) \ . \tag{34}$$

The resulting average kth user bit-error probability per transmitted symbol block m assuming different QAM constellation sizes per activated MIMO layer results in

$$P_{\mathrm{e}\,k}^{(m)} = \frac{1}{\sum_{\nu=1}^{L} \log_2(M_{k\,\nu})} \sum_{\ell=1}^{L} \log_2(M_{k\,\ell})\, P_{\mathrm{e}\,k}^{(\ell,m)} \quad . \tag{35}$$

When considering time-variant channel conditions, rather than an AWGN channel, the BER can be derived by considering the different transmission block SNRs.

4 Adaptive MIMO-layer Power Allocation

In systems, where channel state information is available at the transmitter side, the knowledge about how the symbols are attenuated by the channel can be used to adapt the transmit parameters. Power allocation (PA) can be used to balance the bit-error probabilities in the activated MIMO layers and has been widely investigated in the literature [KRJ00, AL07, JL03].

In order to suppress the DL MUI efficiently, the DL preprocessing matrix has to be designed according to equation (20).

However, the user-specific BER of the uncoded MIMO system is dominated by the specific layers having the lowest SNR's. As a remedy, a MIMO-layer transmit PA scheme is required for minimizing the overall BER under the constraint of a limited total MIMO transmit power. The proposed PA scheme scales the half-level transmit amplitude $U_{\mathrm{s}\,k}^{(\ell)}$ of the ℓth MIMO layer by the factor $\sqrt{\tilde{p}_{k,\ell}^{(m)}}$. This results in a MIMO layer-specific transmit amplitude of $U_{\mathrm{s}\,k}^{(\ell)}\,\sqrt{\tilde{p}_{k,\ell}^{(m)}}$ for the QAM symbol of the transmit data vector transmitted at the time m over the MIMO layer ℓ. Together with the DL preprocessing design, the layer-specific power allocation parameter at the time m results in:

$$\sqrt{p_{k,\ell}^{(m)}} = \sqrt{\beta^{(m)}}\,\sqrt{\tilde{p}_{k,\ell}^{(m)}} \quad . \tag{36}$$

Applying MIMO-layer PA, the information about how the symbols are attenuated by the channel, i. e., the singular-values, has to be sent via a feedback channel to the transmitter side and leads to a high signalling overhead that is contradictory to the fix transmission modes that require no signalling overhead. However, as shown in [AL09] a vector quantizer (VQ) can be used to keep the signalling overhead moderate. Here, a VQ for the power allocation parameters instead of the singular values guarantees a better adaption at a given codebook size, since the power level vectors has less or equal dimensions than the singular-value vectors [AL09]. Moreover, its elements are much smaller digits ranged from 0 to 1, rather than from 0 to $+\infty$ in the singular-value vector case. Hence, the entropy of the power level vectors is smaller, which benefits the quantization accuracy and the feedback overhead.

The aim of the forthcoming discussions is now the determination of the values $\sqrt{\tilde{p}_{k,\ell}^{(m)}}$ for the activated MIMO layers. Unfortunately, the Lagrange multiplier method often leads to excessive-complexity optimization problems. Therefore, suboptimal power allocation strategies having a lower complexity are of common interest [AL07]. A natural choice is to opt for a PA scheme, which results in an identical signal-to-noise ratio

$$\varrho_{\mathrm{PA}\,k}^{(\ell,m)} = \frac{\left(U_{\mathrm{PA}\,k}^{(\ell,m)}\right)^2}{U_{\mathrm{R}}^2} = \tilde{p}_{k,\ell}^{(m)} \, \frac{3\,\xi_{k,\ell}^{(m)}\,\beta^{(m)}}{L\,(M_{k\,\ell}-1)} \frac{E_{\mathrm{s}}}{N_0} \tag{37}$$

for all activated MIMO layers at the time m, i. e., in

$$\varrho_{\mathrm{PA}\,k}^{(\ell,m)} = \text{constant} \quad \ell = 1, 2, \cdots, L \ . \tag{38}$$

The power to be allocated to each activated MIMO layer at the time m can be shown to be calculated as follows [AL07]:

$$\tilde{p}_{k,\ell}^{(m)} = \frac{(M_{k\,\ell}-1)}{\xi_{k,\ell}^{(m)}} \cdot \frac{L}{\sum\limits_{\nu=1}^{L} \frac{(M_{k\,\nu}-1)}{\xi_{k,\nu}^{(m)}}} \ . \tag{39}$$

5 Results

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

5.1 Single-User System

Considering a non-frequency selective SDM (spatial division multiplexing) single-user MIMO link ($K = 1$) composed of $n_{\mathrm{T}} = 4$ transmit and $n_{\mathrm{R}} = 4$ receive antennas, the obtained BER curves are depicted in Fig. 2 for the different QAM constellation sizes and MIMO configurations of Tab. 1, when transmitting at a bandwidth efficiency of 8 bit/s/Hz. 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.

PA can be used to balance the bit-error probabilities in the different number of activated MIMO layers. As shown in Fig. 2, unequal PA is only effective in conjunction with the optimum number of MIMO layers and at high SNR. Using all MIMO layers, our PA scheme would assign much of the total transmit power to the specific symbol positions per data block having the smallest singular values and hence the overall performance would deteriorate.

Non-frequency selective MIMO links have attracted a lot of research. By contrast, frequency selective MIMO links require substantial further research, where spatio-temporal vector coding (STVC) introduced by RALEIGH seems to be an appropriate candidate for

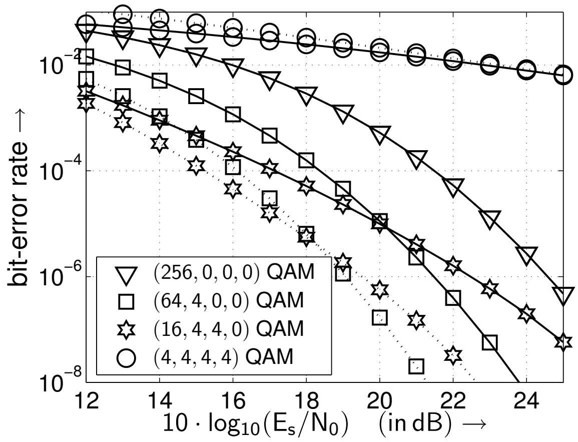

Figure 2: BER with PA (dotted line) and without PA (solid line) when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over non-frequency selective channels

broadband transmission channels, where multipath propagation is no longer a limiting factor [RC98, RJ99, Ges04].

When considering a frequency selective SDM MIMO link, composed of n_{T} transmit and n_{R} receive antennas, the block-oriented system has to take the $(L_{\mathrm{c}}+1)$ non-zero elements of the resulting symbol rate sampled overall channel impulse response between the given transmit and receive antenna combinations into account [ABP09a]. Throughout this paper, it is assumed that the $(L_{\mathrm{c}}+1)$ channel coefficients between the given transmit and receive antenna combinations have the same averaged power and undergo a Rayleigh distribution.

The obtained BER curves are depicted in Fig. 3 and 4 for the different QAM constellation sizes and MIMO configurations of Table 1, when transmitting at a bandwidth efficiency of 8 bit/s/Hz within a given bandwidth over a frequency-selective MIMO channel. Comparing the results, depicted in Fig. 3 for the two-path MIMO channel and in Fig. 4 for the five-path MIMO channel, it can still be seen that delay spread is highly beneficial for the overall performance.

5.2 Multi-User System

The parameters of the exemplarily studied two-users MIMO system are chosen as follows[2]: $P_{\mathrm{s}\,k} = 1\,\mathrm{V}^2$, $n_{\mathrm{R}\,k} = 4$ (with $k = 1, 2$), $K = 2, n_{\mathrm{R}} = n_{\mathrm{T}} = 8$. The obtained user-specific BER curves are depicted in Fig. 5 for the different QAM constellation sizes and MIMO configurations of Table 1 and confirm the obtained results for the single-user

[2] In this contribution a power with the dimension $(\mathrm{voltage})^2$ (in V^2) is used. At a real, constant resistor this value is proportional to the physical power (in W).

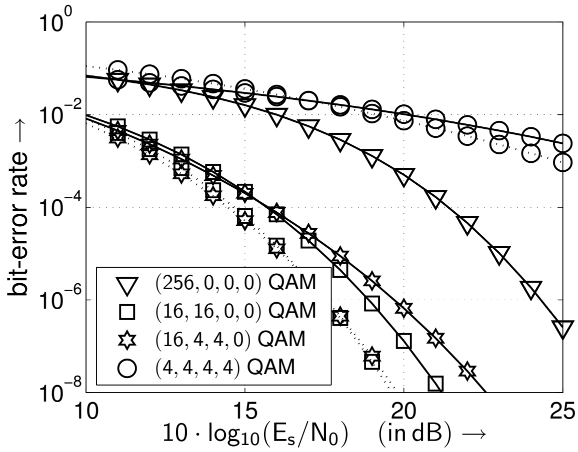

Figure 3: BER with PA (dotted line) and without PA (solid line) when using the transmission modes introduced in Table 1 and transmitting 8 bit/s/Hz over frequency selective channels (two-path channel model, $L_c = 1$)

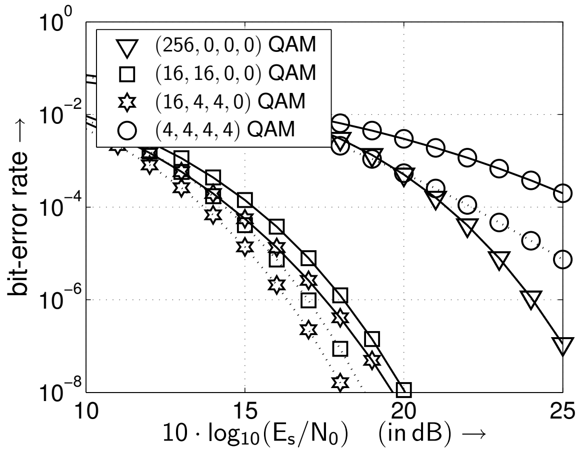

Figure 4: BER with PA (dotted line) and without PA (solid line) when using the transmission modes introduced in Table 1 and transmitting 8 bit/s/Hz over frequency selective channels (five-path channel model, $L_c = 4$)

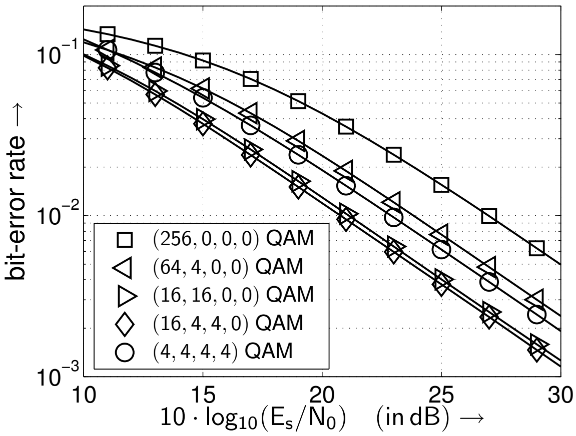

Figure 5: User-specific BERs without PA when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over non-frequency selective channels

system ($K = 1$). Assuming a uniform distribution of the transmit power over the number of activated MIMO layers, it still turns out that not all MIMO layers have to be activated in order to achieve the best BERs. However, the lowest BERs can only be achieved by using bit auction procedures leading to a high signalling overhead [WCLM99]. Analyzing the probability of choosing a specific transmission mode by using optimal bitloading, as depicted in Table 2, it turns out that only an appropriate number of MIMO layers has to be activated, e. g., the $(16, 4, 4, 0)$ QAM configuration.

Table 2: Probability of choosing specific transmission modes at a fixed data rate by using optimal bitloading ($10 \cdot \log_{10}(E_s/N_0) = 10$ dB)

mode	$(64,4,0,0)$	$(16,16,0,0)$	$(16,4,4,0)$	$(4,4,4,4)$
pdf	0	0.0102	0.9524	0.0374

Besides, as depicted in Fig. 5, the gap between the different transmission modes becomes smaller and the influence of PA, as depicted in Fig. 6, diminishes.

Analyzing frequency selective MIMO links instead of non-frequency selective ones, the obtained BER curves are depicted in Fig. 7 for the investigated frequency-selective MIMO channels and confirm the BER results depicted in Fig. 5. From Fig. 7 it can be seen that a high delay spread is beneficial for minimizing the overall BER.

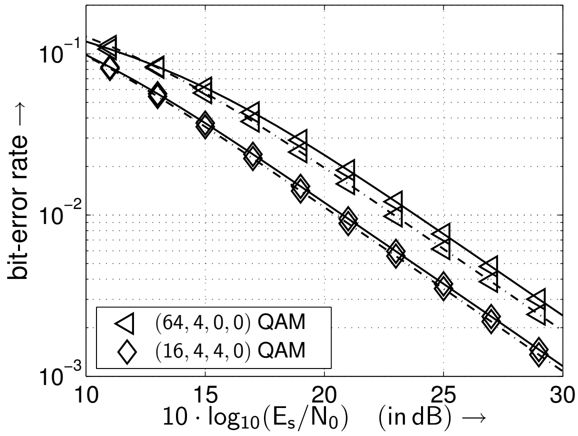

Figure 6: User-specific BER with PA (dashed line) and without PA (solid line) when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over non-frequency selective channels

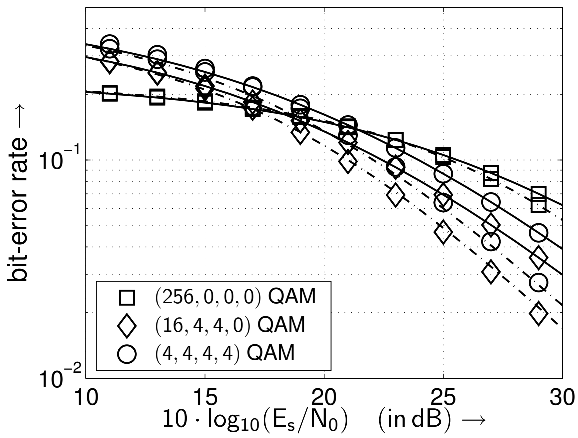

Figure 7: User-specific BERs without PA when using the transmission modes introduced in Tab. 1 and transmitting 8 bit/s/Hz over frequency selective channels (two-path channel model (solid line, $L_c = 1$), five-path channel model (dashed line, $L_c = 4$))

6 Conclusion

In this paper, the DL performance of multiuser MIMO system is investigated theoretically and by software simulation. Both frequency selective and non-frequency selective MIMO channels are considered and conditions to eliminate the multiuser and multi-antenna interferences are established using the SVD of individual user channel matrix. Furthermore, bit and power allocation in multiuser MIMO systems were investigated for constant throughput. Here, it turned out that the choice of the number of bits per symbol as well as the number of activated MIMO layer substantially affects the performance of a MIMO system, suggesting that not all user-specific MIMO layers have to be activated in order to achieve the best BERs. Additionally, unequal PA was found to be effective in conjunction with the optimum number of MIMO layers and delay spread was found to be beneficial for the overall performance.

References

[ABP09a] A. Ahrens and C. Benavente-Peces. Modulation-Mode and Power Assignment for Broadband MIMO-BICM Schemes. In *IEEE 20th Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, Tokio (Japan), 13.–16. September 2009.

[ABP09b] 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.

[AL07] A. Ahrens and C. Lange. Transmit Power Allocation in SVD Equalized Multicarrier Systems. *International Journal of Electronics and Communications (AEÜ)*, 61(1):51–61, 2007.

[AL08] A. Ahrens and C. Lange. Modulation-Mode and Power Assignment in SVD-equalized MIMO Systems. *Facta Universitatis (Series Electronics and Energetics)*, 21(2):167–181, August 2008.

[AL09] A. Ahrens and C. Lange. Modulation-Mode and Power Assignment in SVD-assisted MIMO Systems with limited Feedback. In *First Asian Conference on e-Business and Telecommunications (CeBT)*, Changhua City (Taiwan), 09.–10. February 2009.

[CM03] R. L. Choi and R. D. Murch. New Transmit Schemes and Simplified Receivers for MIMO Wireless Communication Systems. *IEEE Transactions on Wireless Communications*, 2(6):1217–1230, November 2003.

[CM04] R. L. Choi and R. D. Murch. A Transmit Preprocessing Technique for Multiuser MIMO Systems using a Decomposition Approach. *IEEE Transactions on Wireless Communications*, 3(1):20–24, January 2004.

[FH96] R. F. H. Fischer and J. B. Huber. A New Loading Algorithm for Discrete Multitone Modulation. In *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 724–728, London, 1996.

[Ges04] D. 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.

[JL03] J. Jang and K. B. Lee. Transmit Power Adaptation for Multiuser OFDM Systems. *IEEE Journal on Selected Areas in Communications*, 21(2):171–178, 2003.

[JUN05] M. Joham, W. Utschick, and J. A. Nossek. Linear Transmit Processing in MIMO Communications Systems. *IEEE Transactions on Signal Processing*, 53(8):2700–2712, August 2005.

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

[KRJ00] B. S. Krongold, K. Ramchandran, and D. L. Jones. Computationally Efficient Optimal Power Allocation Algorithms for Multicarrier Communications Systems. *IEEE Transactions on Communications*, 48(1):23–27, 2000.

[LYH08] W. Liu, L. L. Yang, and L. Hanzo. SVD Assisted Joint Transmitter and Receiver Design for the Downlink of MIMO Systems. In *IEEE 68th Vehicular Technology Conference (VTC)*, pages 1–5, Calgary, 2008.

[MBW+00] M. Meurer, P. W. Baier, T. Weber, Y. Lu, and A. Papathanassiou. Joint Transmission: An Advantageous Downlink Concept for CDMA Mobile Radio Systems using Time Division Duplexing. *Electronics Letters*, 36(10):900–901, May 2000.

[Pro00] J. G. Proakis. *Digital Communications.* McGraw-Hill, Boston, 2000.

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

[RJ99] G. 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.

[WCLM99] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch. Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation. *IEEE Journal on Selected Areas in Communications*, 17(10):1747–1758, October 1999.

[WMSL02] T. Weber, I. Maniatis, A. Sklavos, and Y. Liu. Joint Transmission and Detection Integrated Network (JOINT), a Generic Proposal for Beyond 3G Systems. In *International Conference on Telecommunications*, volume 3, pages 479–483, Beijing, June 2002.

[WSLW03] T. Weber, A. Sklavos, Y. Liu, and M. Weckerle. The Air Interface Concept JOINT for Beyond 3G Mobile Radio Networks. In *15th International Conference on Wireless Communications*, volume 1, pages 25–33, Calgary, Alberta, Canada, July 2003.

[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.

[ZVDL05] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li. MIMO Systems with Adaptive Modulation. *IEEE Transactions on Vehicular Technology*, 54(5):1073–1096, September 2005.

Online location data: Analysis of publicly available trajectory data sources

Martin Garbe
Research Group on Information and Communication Services
Department of Computer Science
University of Rostock

martin.garbe@uni-rostock.de

Abstract: With the availability of consumer GPS devices the amount of online location data publicly accessible are growing fast. Creating annotated photos or other resources with single location data is very common today. Moreover sharing of moving history data in the form of GPS tracks is becoming popular now. In this paper we analyze three leading services which provide GPS tracks for uploading, downloading and sharing. Besides comparing the growth of these services with other location databases like Flickr our results show that the data fits the requirements needed for doing location data recommendation.

1 Introduction

Web 2.0 raised users from consumers to producers and the amount of information was growing very fast. The key to this raising amount of information is an easy to use interface for users. The management for different kinds of resource is made easier by special programs or technologies. For creating texts blogs and wikis provide the right abstraction from the technique behind. Online photo communities make the uploading of photos easily possible. At the moment location data which is publicly online available are raising with the help of online photo databases like Flickr[1] or Panoramio[2] enabling easy annotation of images with locations. The timestamp when the image was taken can also be associated.

Since GPS receivers are inexpensive GPS Loggers and mobile devices with integrated GPS make history moving data easily accessible. This type of resource can be uploaded and managed with the help of online services like GPSies[3], bikemap[4] or OpenStreetMap(OSM)[5].

This paper presents work in progress of a PhD thesis in the context of location data recommendation. Here we answer the question how much publicly available trajectory data is uploaded to the leading services on the web. A second question whether this data can

[1] http://www.flickr.com
[2] http://www.panoramio.com
[3] http://www.gpsies.com
[4] http://www.bikemap.net
[5] http://www.openstreetmap.org

be used for recommendation algorithms will be answered, too. Creating suggestions with such recommendation algorithms in the form of GPS trajectories is the major aim of the PhD thesis. One key requirement to reach this aim is a large amount of track data per user. If a user has many GPS tracks one part can be used to describe the users profile. The algorithm will learn typical features of these tracks, e.g. user is always going jogging in the wood. The other part can be used to verify the suggestions. If the system made new suggestions for a user they will be checked against the second part of tracks. The recommendation algorithm here is primarily used to identify the typical user behaviour. The more tracks from the second part are suggested the better the algorithm can be rated. Therefore the quality of algorithms can be evaluated.

2 Categories of online location data

Location data can be used for different applications. Depending on the scenario the data is associated with resources. To narrow our problem description we classify the online location data into categories, see table 1. For a better understanding the categories are given with examples. A detailed description of each category follows.

	single spatial record	multiple spatial records
temporal	photo annotated with a position and a timestamp	record of a movement, e.g. jogging track
non-temporal	position information for describing a location	corners of a house

Table 1: Classification of online location data with examples

A single record of spatio-non-temporal data is a coordinate of any coordinate system. This data is used e.g. for giving the exact position of something.

A series of spatio-non-temporal data records describes spatial objects. The first coordinate of the series is connected with the second, the second with the third and so on. In a two dimensional space polygons are created this way. For example buildings can be represented on a map as two dimensional polygons using the longitude and latitude values of its corners. There is another connection between this data series and a single record. When no exact information is needed a series of spatio-non-temporal data can be represented by a single record. This means that geographic entities can approximately be addressed by a single record. For example people are describing the position of their homes with one single data record whereas their home is a building and can also be described as a series of data records.

A spatio-temporal record consists of coordinate and time information. This type of data is typically found in geotagged photos. While time information is immediately stored in the photos EXIF meta data after taking the picture the location information is mostly inserted later by hand. Due to maps as supporting tools the location can be associated without any knowledge of the GPS coordinate by the user only remembering the street or other landmarks.

Spatio-temporal data series are a set of coordinates with additional time information. While moving and logging the position of an object the result is a set of positions with time information of every single record. A GPS track recorded while going jogging is an example for this type of data.

3 Related work

The growing amount of online location data can be seen in many areas. One part are the geotagged photos with spatio-temporal records. Flickr represents one of the largest member of this category. Flickr was growing from 2 billion photos in November 2007 to 4 billion photos in October 2009 [Fli]. Besides new photos the amount of users, links between users and links from users to photos is growing. This was analyzed in [MKG+08]. The authors crawled Flickr over three month and recognized 950,143 new users and over 9.7 million new links.

In contrast to single location resources like geotagged photos there is little information about growing of multi location resources like GPS tracks. Publicly available GPS trajectories are found for example in the OpenStreetMap project. The project collects GPS tracks to defend against possible plagiarism allegations. A detailed statistic about new users, nodes and ways for OpenStreetMap can be found in [Opeb]. The authors show database statistics over time from the OpenStreetMap Project. The data show the exponential growth of the project. The number of new tracks per month or new tracks per user are missing.

Collecting data from websites can be done in different ways. One is to download all pages and extract the information with regular expressions or parsing the DOM tree, e.g. [YD09]. The problem here are mostly complex regular expressions, problems with DOM parsing because of invalid HTML and websites changing their website layout.

Another one is using APIs provided by the websites if available. This method is often used with online photo databases, see [LJ06], [BFL+08]. For this problem the authors of [HE08] implemented the framework IM2GPS for grabbing data of Flickr through the provided API. IM2GPS is limited to Flickr and while API changes are done users have to wait for a new updated version.

In contrast to these two methods the author of [Old08] suggests to use RSS feeds for collecting data and analyze growth, stability and convergence aspects. The data included in RSS sufficiently supports these analyses. The advantage here is standardised XML and a large amount of processing tools.

When analyzing GPS history data researchers always have the problem where to get source data. In [ZZXM09] the movement of 107 tourists were recorded and interesting locations extracted. Here the tourists have to be found and equipped with GPS loggers. Not taking real data but creating synthezised tracks is another choice. This was done in [SWR+04]. The authors took a small amount of real tracks, added noise to them, and ended up with a large amount of tracks. Tracks from public websites like GPSies & co are available but not used today.

4 Amount of online location data

The growth of single spatial data records temporal and non-temporal on the web, e.g. annotated photos, was already analyzed and shown in section 3. The growth of non-temporal multiple spatial records, e.g. map data of OpenStreetMap, was also shown in the previous section. For location movement data, which means temporal-spatial data series, no information about the amount and the growth are available.

We analyze this category of online location data here. There are no sources which state how much data these and other websites manage. So the problem is how to find the sites with the most data. Estimating that the most important and largest websites are found on Google's first searching result pages a list of possible services was created by searching for "gps", "gps track" and "gps route" with Google. The first three pages we examined and searched for further sites. As the second filtering step the amount of new items per day were estimated. This was done by looking at the RSS feeds and filtering out websites with an average of less than 500 new items a day.

As a result the following websites were chosen for further analyzes: GPSies, OpenStreetMap and bikemap. From these websites an RSS feed was regularly downloaded for a period of two months starting on May 1st, 2010 until June 30th, 2010. Alternatively crawling of websites would have been possible but due to several disadvantages compared to grabbing RSS feeds we decided against this method. RSS is a XML format where tools for processing are already available in contrast to website crawling where regular expressions or other techniques are necessary. The download interval was depending on the amount of new tracks. For example the RSS feed of GPSies contains 30 items. While inspecting different intervals for grabbing the RSS feeds a 10 minutes interval was chosen for GPSies and a 5 minutes interval for OpenStreetMap and Bikemap. The feed from the later two had to be fetched more often because there are less items in one feed than GPSies.

To support the idea of Open Science[Opea] and reproducible results the collected data will be freely available on `http://opsci.informatik.uni-rostock.de`.

The data presented here for OpenStreetMap and GPSies and the real number of uploaded tracks can differ. In OpenStreetMap there is a possibility to make "private" uploads which are not seen by the public and GPSies has the possibility to make anonymous uploads. This will be filtered later.

We provide an overview of the growth of new items in figure 1. Fastest growing is GPSies with 50,916 new tracks after two month followed by bikemap (40,737) and OpenStreetMap (38,560). The number of tracks on each of the websites is growing linearly in the two month period. Compared to Flickr with a growth of 1 billion pictures in 2009 the amount of new items is very small.

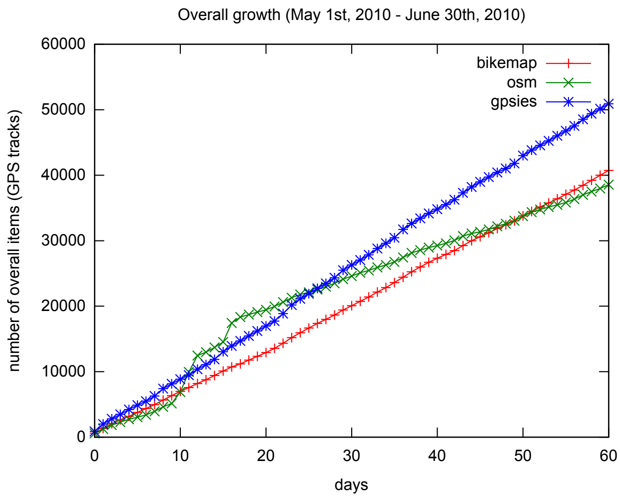

Figure 1: Growing number of uploaded tracks.

More detailed results are available in table 2. The number of users uploading differs between OpenStreetMap and bikemap very much. In OpenStreetMap nearly 4,000 users were active while bikemap had 10,533. Although the amount of OpenStreetMap users is the half of the amount of bikemap users the number of uploaded tracks is almost the same. There is no information for GPSies and "users uploaded data" in the table because the RSS feed of GPSies has no user information. Compared to Flickr the number of overall resources is very small. In Flickr there are 4 billion items in 2009 and the services analyzed here only have half a million as of June 2010.

	GPSies	OSM	bikemap
new tracks overall	50,916	38,560	40,737
ø(new tracks per day)	~835	~623	~668
users uploaded data	-	3,989	10,533
overall tracks	~430,000	~750,000	~560,000

Table 2: Statistics of collected data and overall tracks.

With the information above it is possible to make estimations about the amount of information saved in databases by the websites. The distribution of how much users are sharing which amount of information was not yet answered. But this distribution is important if the tracks should be used for creating recommendations. Recommendation algorithms need a first training set of moving information of the user to suggest new tracks. Creating suggestions for users with only one track is not feasible. With a few tracks per user the training could be done but rating the suggestions is a problem. Ideally there are many users associated with many tracks. In this situation the tracks can be divided into training set and test set. The profile of a user will be learned with the training set. Then the suggestions can be compared with the test set. If GPS tracks from the test set are suggested the recommendation algorithms can be rated positively. If only tracks not known by the user were suggested the user has to be consulted to rate these suggestions manually.

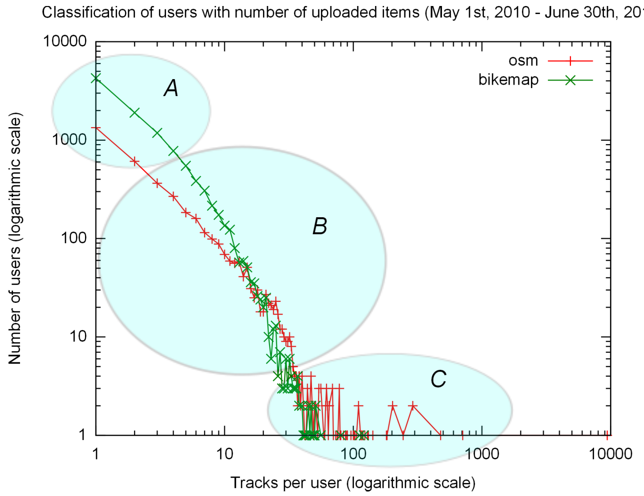

Figure 2: Classification of uploading users.

The results of analyzing the amount of tracks and classifying the users into groups are shown in figure 2. The RSS feed from the bikemap website has no user information so the assignment of tracks to users is not possible and is not included in this figure. The graph is plotted in logarithmic scale and reveals a distribution which follows nearly a power law. A power law graph can be described with a straight line in the logarithmic scaled coordinate system. Some outliers near the x-axis prevent a good approximation of our graphs with straight lines. All graphs in the figure are growing exponentially.

The graphs can be divided into three parts as shown in figure 2. In part A we find many users uploading less tracks. Part C is the opposite of part A. There we find a small number of users but many tracks per user. The rest of the users lies in B between A and C.

One the one hand both services shown in the graphic have many users with one single track associated. On the other hand there are less users with many tracks. On the bikemap website 10,533 users posted new tracks in the two month period. Nearly the half of them (4,258 users, 40%) uploaded only one track. But 1% of the users uploading much data created 10% of the new tracks. On OpenStreetMap 3,989 users posted new tracks. Here 34% of them (1,339 users) uploaded only one track. In contrast to bikemap 1% of the users here uploaded 39% of the new tracks. Which means some people are significantly more active than the rest.

This exponential behaviour is often seen in nature, e.g. earthquake magnitudes, word frequencies and city sizes. Another example is tagging and tag suggestion for bookmarking. Many users only use little tags. As a result a few tags are used by many users and on the other side many tags are only rarely used [GH05].

As result our research reveals that there are users with many tracks which can be used to test recommendation algorithms for GPS tracks. For the recommendation algorithms users with many tracks are needed. This is given in OpenStreetMap and bikemap. Depending on the needed amount of tracks per user the amount of available users vary. The less tracks are needed per user the more users are worth considering. If more users with many tracks are needed then collecting more data over a longer period should be done.

5 Conclusions and outlook

We presented analyzes of three leading public online GPS track databases. Our results from the overall collected data reveals that services managing spatio-temporal data series are not as popular as services with single temporal- or non-temporal spatial data (e.g. photo tagging services).

Our analyze of number of tracks per user revealed the possibility to use such online databases for testing recommendation algorithms in the area of moving behaviour. To conduct such tests users with many GPS tracks are needed. This requirement is fulfilled by at least two of the three observed services. Knowing this saves much time because no real GPS data has to be produced.

Further research will cover recommendation algorithms for GPS trajectories. On the basis of the presented research results users owning many tracks will be selected for further analysis. We will look for features which describe the tracks and the users. Using these features a profile can be build and suggestions of tracks can be rated.

References

[BFL+08] Michal Batko, Fabrizio Falchi, Claudio Lucchese, David Novak, Raffaele Perego, Fausto Rabitti, Jan Sedmidubský, and Pavel Zezula. Crawling, Indexing, and Similarity Searching Images on the Web. In *SEBD*, pages 382–389, 2008.

[Fli] Flickr 4 billion photos. `http://blog.flickr.net/2009/10/12/4000000000/`. Retrieved June 30th, 2010.

[GH05] Scott A. Golder and Bernardo A. Huberman. The Structure of Collaborative Tagging Systems. *CoRR*, abs/cs/0508082, 2005.

[HE08] James Hays and Alexei A. Efros. im2gps: estimating geographic information from a single image. In *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[LJ06] Kristina Lerman and Laurie Jones. Social Browsing on Flickr. *CoRR*, abs/cs/0612047, 2006.

[MKG+08] Alan Mislove, Hema Swetha Koppula, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Growth of the flickr social network. In *WOSP '08: Proceedings of the first workshop on Online social networks*, pages 25–30, New York, NY, USA, 2008. ACM.

[Old08] Steffen Oldenburg. Comparative Studies of Social Classification Systems Using RSS Feeds. In *WEBIST 2008: Proceedings of the Fourth International Conference on Web Information Systems and Technologies*, pages 394–403, Funchal, Madeira, Portugal, 2008. Springer Verlag.

[Opea] Open Science. `http://sciencecommons.org/resources/readingroom/principles-for-open-science/`. Retrieved June 30th, 2010.

[Opeb] OpenStreetMap Statistics. `http://wiki.openstreetmap.org/wiki/Stats`. Retrieved June 30th, 2010.

[SWR+04] Stefan Schroedl, Kiri Wagstaff, Seth Rogers, Pat Langley, and Christopher Wilson. Mining GPS Traces for Map Refinement. *Data Min. Knowl. Discov.*, 9(1):59–87, 2004.

[YD09] Erjia Yan and Ying Ding. Applying centrality measures to impact analysis: A coauthorship network analysis. *J. Am. Soc. Inf. Sci. Technol.*, 60(10):2107–2118, 2009.

[ZZXM09] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from GPS trajectories. In *WWW '09: Proceedings of the 18th international conference on World wide web*, pages 791–800, New York, NY, USA, 2009. ACM.

Implementation of an Interactive E-Learning Education Network in the Field of Electrical Engineering.

Sebastian Kula
Westpomeranian Business School University of Applied Sciences
71-210 Szczecin, Poland
sebastian_wsk@yahoo.com

Alexander Vasenev
LMN Laboratory, University 'Politehnica' Bucharest
060042 Bucharest, Romania
vasenev@lmn.pub.ro

Abstract: Modern digitisation and computerisation tendencies continuously raise demand for educated engineers and scientists with knowledge in the field of electrical engineering and electronics. Interactive e-learning networks could be effective solutions for knowledge transfer and professional education on different levels. This article describes an e-learning network implementation created within scientific project COMSON. Experience gained in this project was a basis for creating concept of an interactive e-learning education network dedicated to students pursuing to obtain the title of engineer. The concept presented here with proposal view to its implementation on studies conducted in the West Pomeranian Business School, University of Applied Science.

1 Introduction

Electrical engineering education is a complex and challenging mission for universities and industrial companies across Europe. Amount of related information increase tremendously, causing demand for collaborative actions to educate creative, flexible, high level engineers and scientists. Nowadays it is necessary to facilitate intensive technological progress and effective transfer of knowledge between academics and practitioners, university research centres and R&D departments of innovative companies. The e-learning networks can significantly help to achieve those goals [Al08].

Available visualization tools should be connected with simulation models to assist students to understand complex structures and processes, as well as to perform independent experiments and research tasks using just computer and Internet access. Those oriented to study phenomena and processes which in reality can appear in equipment laboratories only.

Successful example of an e-learning cooperation between industry and academic sector was created within COMSON project ("COupled Multiscale Simulation and Optimization in Nanoelectronics"), formed by R&D departments of three European semiconductor companies (Qimonda, NXP and ST Microelectronics) plus five academic partners from universities of Wuppertal, Bucharest, Calabria, Catania and Eindhoven. Being related to tasks of research institutes in academic and industrial centres, project was formed as research training network. Common goals and values in addition to communal relationship were presented from the very beginning of the project, forming key elements of exchange processes in learning network, valuable for creation and reuse of learning elements [Ko08]. One of the major achievements of the COMSON project was creation of an e-learning platform, able to facilitate create, revise and perform transfer of knowledge in connection with keeping track of code in repository, enabling re-experience by decreasing complexity and transactive group memory [He04].

This article presents a contribution and achievements of authors in the frame of the COMSON project with emphasis to e-learning and developed approach for creation of the e-learning network for students of the WPBS (Westpomeranian Business School) University of Applied Sciences.

2 E-learning network organization

The COMSON network was formed by individuals with varying level of knowledge and experience: early stage, experienced, very experienced researchers and experienced industrial engineers. Members of COMSON project worked at seven host nodes in four countries. Such complex structure required a binder, able to facilitate easy cooperation, exchange of information and knowledge. Success of the project was made possible through the e-learning platform and collaborative work, where roles were derived from experience and expertise field of participants. More experienced researchers were consulting and revising work and materials, developed by colleagues. Creation and code upload was iterative process, with accent on review and corresponding changes based on recommendation, stored in repository according to iterative design principles. Because of asynchronous type of interactions between students and supervisors in the middle and history of particular functional element changes, international level of interaction, even with people speaking different languages could be obtained. Although such option was available, this direction was not investigated as all participants were able to communicate effectively in English.

The e-learning platform in COMSON project consisted of following elements and tools:
- Internal forum of the project, where took place discussions, exchange of information between researchers and specialists.

- Repository (technically space on server dedicated for data storage) for course notes, slide presentations, articles, book chapters, etc.

- Simulation platform (technically each environment able to compile and run prepared code) that interfaces with a DP (Demonstrative Platform). The DP played central role in the project, contained repository and environment which was able to run, under DP all codes prepared on each singular simulation platform.

- Interactive courses, based on codes created for the DP, code tutorials and code documentation.

It is worth emphasising that authors were oriented to use and to rely on GPL (General Public License) during creating materials within all of the above items, as well as contribution made in the form of code, documentation, etc. Decision was based on EU requirement to use open source solutions for the project. Structure and organization of described element was the backbone for different type of materials and ways of interactions. Example of developed solutions could be an optimization course, oriented to wide audience and materials, developed for particular professionals, interested in study of interconnect analysis.

Creating of e-learning materials is a complex task with attention to deliver effective material as well as providing learning path to encourage future study by audience. According to requirements and students' background course should provide flexibility according to learners potential. During the COMSON project Bucharest node implemented e-learning optimization course, one of the important aspects of the project and within collaboration scheme between project's nodes. The course included materials with references to Bucharest, Catania (STMicroelectronics), Calabria and other nodes. Materials were presented in a form of a Moodle course and code was available as Octave and Scilab implementation, based on widely available open software.

Numerical optimization techniques are advanced effective modules for undergraduate and postgraduate students at the University Politehnica of Bucharest. The optimization course's purpose is to present the fundamental concepts and main numerical optimization methods used in scientific computing and the computer aided design of electromagnetic devices. The courseware is developed in two languages – local language of academic node and English. The English version is a translation of the original Romanian book [Ci02]. Taking into consideration the diversity of the optimization problem encountered algorithms and computing programs, it is difficult to initially find existing solvers that are optimized and efficient for a particular real problem. Usually to solve real problem appropriate baseline algorithm, as close as possible to the encountered problem, must first be selected. Offered course presents not only the theory, but also practical applications like solenoid optimization tasks and pseudocode, upon which a more refined solution can be developed.

With the deployment content on the DP, which can be referred as a distant available Linux-based machine, students have now an easy interactive access to the course. When used in conjunction with other materials, i.e. ROMI (Reduced Order Modelling of Interconnects) on the COMSON DP, they can personally exploit the algorithms and write the codes in order to gain a deeper understanding of the theory and methods, their advantages and drawbacks, work with test cases and consult supervisor accordingly. In combination with pseudocode, description of algorithms and history of materials changes student can develop better understanding of particular implementations. Exemplary codes can be archived for future use.

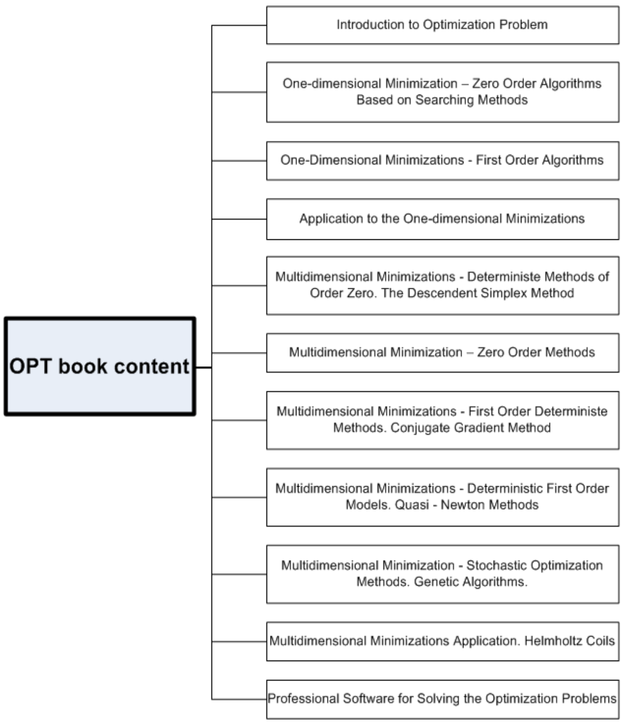

Figure 1: Numerical optimizations handbook.

Materials are arranged to develop higher-order thinking skills in students, referred to cognitive domain educational objectives outlined in Bloom's taxonomy of educational objectives [Bl56]. Presented pseudocode for a number of optimization routines and implementation in high level GPL language Octave encourages student to experiment with the code for better understanding, analyze, synthesis and evaluation of available solutions [Va09]. Different representation forms and exploration of the code interactively could provide better understanding and training results by giving opportunity to use discovery learning principles for students with appropriate background [JJ98]. Students can easy create hypothesis based on presented optimization routines, verify them by simulation and generate new hypothesis by code modification.

Example of professional-oriented materials developed ROMI code was mentioned earlier. The ROMI package was one such conduit developed within the COMSON framework to facilitate seamless knowledge transfer state of the art developments to various levels within the electromagnetic modelling community.

The ROMI was developed by Sebastian Kula at the Bucharest's LMN laboratory for the COMSON-DP and was dedicated to extraction of reduced order models of multi-conductor interconnects. Using the ROMI software one can calculate per unit length parameters of the interconnect elements. The extracted model is an equivalent (SPICE) circuit of lowest order which takes into consideration following critical effects of the EM field which manifests itself only at high frequencies (multi gigahertz) in the interconnect lines of nanoelectronic semiconductor devices: interconnect resistive losses, interconnect self and mutual inductive effects, interconnect transverse capacitive effects, skin depth effects in interconnects as a function of frequency, induced currents/losses in the semiconductor substrate. ROMI workflow was implemented in open source Octave language [Ku08]. All scripts and functions of the ROMI code are available at the COMSON Concurrent Versions System Repository webpage site. Each user can freely check the code. The ROMI code is well documented and contains tutorial and examples. Although the ROMI was orientated to slightly different target audience, scheme of usage and education stays the same as described earlier.

3 Visualisation in e-learning

Additional supportive materials were developed, including set of the e-learning courses, LayoutEditor tutorial, professional communication course and information about technology supported learning and training. Courses were oriented to students in purpose to provide additional knowledge and develop soft skills, necessary to work in collaborative environment and effective knowledge transfer.

In addition to professional training, number of software solutions with descriptive materials were developed. Special emphasis should be put on visualization aspect due to complexity of modern schemes and with orientation to e-learning. For better user's understanding and evaluation complex 3D forms software named LayoutEditor was utilized. Using imbedded script-language subset for visualization (fig.2) was created to represent form of the circuit in suitable way for interactive exploration. For further calculation was created solution [Va08] to visually identify fundamental loops on Graphic Data System layout for extraction of self and mutual reluctances using Finite Integral Technique between the hooks of Manhattan shapes (union of rectangles). Using corresponding commands it is possible to hide/show parts of the scheme, follow certain layer and export specified data for further calculations.

Also in the frame of the COMSON several scientific animations of current flow in a 24 GHz Low Noise Amplifier were created using Momentum software for enhanced visualization of electromagnetic effects in interconnects. The software helps to achieve high level of 3D visualization in motion with high potential for education purposes.

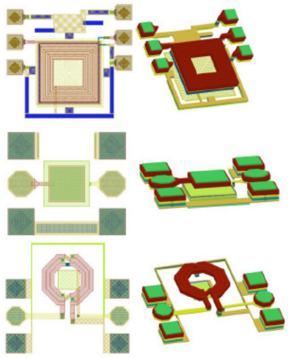

Figure 2: Visualization of the layouts.

4 Students e-learning network

The COMSON network was addressed to a group of scientists and engineers specialized in the field of electrical engineering. Described organization scheme will be used with slight modification in the future, oriented to SEN (Students E-learning Network) users with less electrical engineering knowledge than in COMSON project. It will be group of students who are intent to be engineers in the future. Despite such apparently fundamental difference between these networks, experience gained with e-learning organization and implemented in the COMSON project can also be used in the SEN.

The recipients of the SEN will be the WPBS University, namely its students of computer science faculty. The WPBS University conducts education in the field of electrical engineering just since one year, which means that the university has no real laboratory to conduct electrical engineering subjects. While being disadvantage on one hand, this could be considered as an opportunity on another, giving possibility to develop remote and unconventional teaching way of electrical engineering. Also, this is the opportunity for the WPBS University e-learning network, which at considerably lower cost is able to educate professionals.

The newly created WPBS University e-learning network will use majority of open source tools of the COMSON network [Va09]. The only difference will be replacement Octave program by Qucs program, integrated circuit simulator and electronic design automation tool, released under the GPL license. Pedagogical concept of usage LMS (Learning Management System) and SEN is described on fig. 3 and based on the Bloom's taxonomy. Student e-learning network will contain following elements:

- Forum, where students will be equivalent users on a par with the teacher.

- Repository of slide presentations, books, articles.

- Interactive courses made using LayoutEditor for typical electronic components.

- Simulation platform, based on the Qucs software, available through a local or remote access.

Bloom's Taxonomy	Level Verbs	Student network	Material Medium
Knowledge	arrange	e-materials	LMS
Compre-hension	translate classify describe	baseline device model selection	LMS + Qucs
Application	apply use	baseline device simulation	LMS +Qucs
Analysis	compare analyse test	circuit extraction and comparison	Qucs
Synthesis	arrange collect create	enhanced device model development	
Evaluation	choose compare rate	parameter tuning and validation with measurement	

Figure 3: Bloom's taxonomy for SEN.

Among these elements, one should pay more attention to Qucs. This program will be used not only to the usual Electronic Design Automation simulations but also to the implementation of elements of the virtual laboratory tools for the electrical engineering education, including function generator, multimeter, oscilloscope, etc. Implementation of these tools will be realized through the use of VHDL programming language, recognized by Qucs. From the standpoint of teaching it will give two results: students will not only have to truly learn the working rules of these electronic devices, but also will have to write programs in VHDL language. In this way it is possible to develop students' knowledge in the field of the electrical engineering and computer science simultaneously, which is highly desirable training effect within education of young engineers and computer scientists.

5 Conclusions

This paper presents recent work in area of the e-learning and visualization tools, aimed to improve interactive e-learning in electromagnetics, used by the COMSON academic partners in a form of e-learning network. Multiple roles were defined within the education network depending on experience and expertise background – creation, revision, supervision and evaluation of submitted materials. According to recommendations material evolved with respect to principles of iterative design. By providing supportive information like lectures, pseudocode, particular implementation and application to certain problems, opportunity for discovery learning was created.

Shown visualization techniques oriented to present complex objects by means of educational and scientific visualization, visual analytics, particularly valuable for e-learning. Visualization becomes an important part for educational purposes concerning complex electronic models and necessary for better understanding of processes and layouts.

In the paper was also described the effort of the authors in preparing e-learning materials in the frame of the network in the COMSON project. What is more, new concept plan for the WPBS University e-learning network was showed.

Acknowledgement

The authors would like to gratefully acknowledge support and insightful discussion with Prof. Daniel Ioan and Prof. Gabriela Ciuprina. The authors would also like to acknowledge the financial support offered by the European Commission under the Marie Curie Fellowship (FP6) programme.

References

[Al08] Ali, G.; Bilotta, E.; Gabriele, P.; Pantano, P. and Servidio, R.: An e-Learning Platform for Applications of Mathematics to Microelectronic Industry. Progress in Industrial Mathematics at ECMI 2006, Mathematics in Industry, 2008, Vol. 12, II, 736-740.

[Bl56] Bloom, S.: Taxonomy of Educational Objectives: The Classification of Educational Goals. Susan Fauer Company, Inc. 1956.

[Ci02] Ciuprina, G.; Ioan, D.; Munteanu, I.; Rebican, M.; Popa, R.: Numerical Optimization of Electromagnetic Devices. Printech, Bucharest, 2002.

[He04] Hemetsberger, A., & Christian, R. (2004). Sharing and creating knowledge in open-source communities: The case of KDE. Proceedings of the 5th European Conference on Organizational Knowledge, Learning and Capabilities (OKLC).

[JJ98] de Jong, Ton.; van Joolingen, W.R.: Scientific Discovery Learning with Computer Simulations of Conceptual Domains. Review of Educational Research, 1998, Vol. 68, pp. 179–201.

[Ko08] Koper, R., & Specht, M.: TEN-Competence: Life-Long Competence Development and Learning. In M-A. Cicilia (ed.), Competencies in Organizational e-learning: concepts and tools. Hershey: IGI-Global 2008; pp. 234-252.

[Ku08] Kula, S.: ROMI Code for Extraction of Reduced Interconnect Models. In Proc. 12th Int. Student Conf. on Electrical Engineering POSTER, Prague 2008.

[Va08] Vasenev, A.; Gim, S.; Stefanescu, A.; Kula, S. and Mihalache, D.: A novel graphical based tool for extraction of magnetic reluctances between onchip current loops. In Proc. Scientific Computing in Electrical Engineering SCEE, Helsinki, 2008, Mathematics in Industry, Editors J. Roos and R.J. Costa, Springer, Berlin Heidelberg 2010; pp.45-52.

[Va09] Vasenev, A.; Stefanescu, A.; Mihalache, D.; Kula, S.; Gim, S.: Interactive E-learning on Reduced Order Modeling in Electromagnetics Using COMSON Federated Repository. In Proc. 9th IEEE International Conference on Advanced Learning Technologies, Riga, 2009, Advanced Learning Technologies, IEEE International Conference, Washington, DC 2009; pp. 408-409.

Power Allocation for Adaptive Asymmetric distributed MIMO Networks

Fabian Monsees, Yidong Lang, Dirk Wübben and Armin Dekorsy
Department of Communication Engineering, University of Bremen, Germany
Email: f.monsees@uni-bremen.de

Abstract: Distributed MIMO network topology is one of the most promising technologies in modern communication networks. Compared to classical Single Input Single Output (SISO) links, where information is broadcasted from one sender over huge distances, the distance between two points in multi-hop networks is rather small. By means of such relay networks it becomes possible to guarantee a constant throughput while meeting a given end-to-end (e2e) outage constraint over extensive areas. In each hop several relay nodes form a virtual antenna array (VAA). By using this concept, capacity improving techniques like distributed space time codes can be implemented [Doh03]. The possibility of allocating recourses like transmit power is much more complex in distributed networks than in SISO links. The goal of power optimization in relay networks is to minimize the total transmit power while meeting a given e2e outage constraint. However, in practical scenarios, the structure of a relaying network does not have to follow certain geometrical regularities but can be arranged asymmetrically. As a consequence of that, the closed form expression of power allocation for asymmetric networks is intricate. In this paper the performance of power allocation for an adaptive asymmetric distributed relay networks is analyzed. The significant power saving compared to a non adaptive network is demonstrated by numerical calculations.

1 Introduction

By the concept of relaying networks a classical SISO link can be transformed into a virtual MIMO multi-hop system. A certain amount of relay nodes form a VAA. In each hop, the information is transmitted from one VAA to the next one until the destination is reached. In order to guarantee a give data rate and outage probability for each user, the transmit power in such a network has to be shared reasonable among the transmitting nodes. In the sequel the e2e outage probability is considered as a measure of a given Quality of Service (QoS) requirement that has to be fulfilled. Relays are due to their mobility highly energy restricted. Therefore we consider the transmit power as one of the most important factors that is required to meet a given QoS of a network. Since most of the wireless communication is done over slow fading channels, where the model of ergodic capacity is not applicable, the measure of the e2e outage probability is valid and will be considered in this paper. In order to reduce the overall transmit power in relay networks, several investigations have been made yet. In [LBW$^+$08] the authors derived a closed form expression for a non adaptive network. A symmetric multi-hop network was assumed, where the VAAs are distributed along a line in each hop. Therefore the channels in each hop could be modeled

to be equal with same pathlosses. Power optimization has been done in a non adaptive sense, means that if one relay fails, the whole communication fails. The proposed closed form solution was derived by utilizing common optimization techniques. In such a scenario the whole performance of the network is strongly dependent on the worst link. If one link fails, the whole e2e communication collapses. In order to overcome this drawback, a scheme for an adaptive network was introduced in [LWK09], where nodes that cannot decode the received signal correctly, just stop transmitting in order to save energy. The remaining nodes adapt to the new scenario by a new space time code scheme. The achieved gain by utilizing an adaptive network was for a three hop system with a distance between each VAA of 1km and three nodes per VAA about 15dB. As far as an optimal alignment of the nodes cannot always be presumed, asymmetric networks have been investigated in [LWK08]. A closed form power allocation scheme for asymmetric non adaptive networks was derived with the help of several approximations for the outage probability per hop. Since practical multi-hop networks can be modeled at best by an asymmetric structure, power allocation schemes for such structures have to be investigated. The contribution of this paper is to investigate the gain of power allocation for adaptive asymmetric distributed MIMO networks. The problem is formulated as a convex optimization problem. The performance is investigated later by numerical calculations.
The remainder of this paper is organized as follows. In section 2 we introduce the system model for asymmetric MIMO multi-hop networks as well as the mathematical description. The optimization problem for the overall transmit power is given in section 3. Finally the achieved performance is analyzed and conclusions are drawn from that in section 4 and 5 respectively.

2 System Model

The setup for the asymmetric MIMO multi-hop system is shown in Fig. 1.

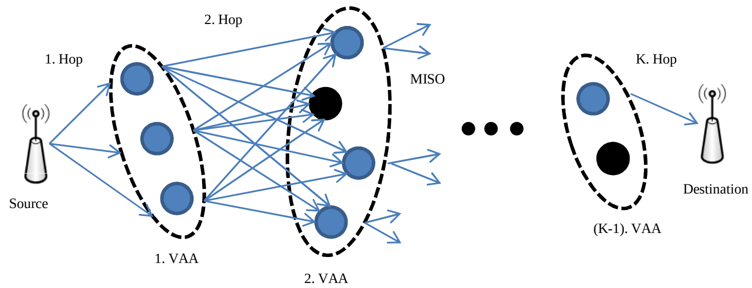

Figure 1: An asymmetric distributed MIMO multi-hop system

The source sends the information to the first VAA in the first time slot. Each node that has received the information successfully is considered to be active. The nodes that have not received or were not able to decode the information properly are considered to be in outage. In the next time slot the active nodes in a VAA adapt to the new setup of the VAA and transmit the data cooperatively according to a space time code, tailored for the number of active nodes. If all nodes within one VAA fail to decode the information, the VAA is considered to be in outage. In this case the whole e2e communication fails. This adaptive scheme is continued until the message reaches its destination. It is assumed that each relay node has only one antenna element, so that it communicates in half duplex mode. Decode and Forward (D&F) is presumed to be implemented in each node [LTW04]. Furthermore it is assumed that there is no exchange of data between the relays in a VAA. Since every node decodes the information separately and re-encodes it by using a fraction of a shared space time code, the communication in one VAA can be modeled as several multiple-input single-output (MISO) links. Data rate R and duration of transmission are assumed to be equal in the whole network. In all hops the total bandwidth denoted by W is used. In the sequel k indexes the current hop, t_k denotes the transmit nodes and r_k the receive nodes. The set of active transmit nodes in hop k is denoted by $\mathbb{T}'_k$ and the set of inactive nodes by $\tilde{\mathbb{T}}_k$ The pathloss between a sending node i and a receiving node j is denoted by $1/d^{\epsilon}_{k,j,i}$. Where $d_{k,j,i}$ denotes the distance between both nodes and ϵ denotes the pathloss exponent within a range of 2 to 5 as used for the most wireless channels. The Signal $\mathbf{S}_k \in \mathbb{C}^{t_k \times T_k}$ denotes the space time encoded signal of length T_k transmitted from the set of active nodes $\mathbb{T}'_k$ at hop k. The signal received by node j is defined by the following equation.

$$\mathbf{y}_{k,j} = \mathbf{h}_{k,j} \cdot \mathbf{\Lambda}_k \cdot \mathbf{S}_k + \mathbf{n}_{k,j} \ , \tag{1}$$

Where $\mathbf{\Lambda}_k$ corresponds to the diagonal matrix given by

$$\mathbf{\Lambda}_k = \mathrm{diag}\left\{ \sqrt{\frac{\mathcal{P}_{k,1}}{d^{\epsilon}_{k,1,j}}}, \cdots, \sqrt{\frac{\mathcal{P}_{k,t_k}}{d^{\epsilon}_{k,t_k,j}}} \right\} . \tag{2}$$

The gaussian noise vector with power spectral density N_0 is expressed by $\mathbf{n}_{k,j} \sim \mathcal{N}_C(0,\, N_0) \in \mathbb{C}^{1 \times T_k}$. The transmit power supplied by the ith node in the kth hop is expressed by $\mathcal{P}_{k,i}$. The channel from the t_k transmit nodes to the jth receive node at the kth hop is expressed as $\mathbf{h}_{k,j} \in \mathbb{C}^{1 \times t_k}$. Its elements $h_{k,i,j}$ obey the same uncorrelated Rayleigh fading statistics, i.e., they are complex zero-mean circular symmetric Gaussian distributed with variance 1. The capacity of an asymmetric MISO channel is given by

$$C_{k,j} = W \log\left(1 + \frac{1}{W N_0} \sum_{i \in \mathbb{T}'_k} g_{k,i,j} |h_{k,i,j}|^2\right) \ , \tag{3}$$

with the co-factors $g_{k,j,i} = \mathcal{P}_{k,i}/d^{\epsilon}_{k,j,i}$. Note that $g_{k,j,i}$ corresponds to the squared diagonal elements of $\mathbf{\Lambda}_k$ in Eq. 2 [Doh03]. The probability that the link from the set of $\mathbb{T}'_k$ sending nodes to the j_k receiving nodes in hop k is in outage can be interpreted as that this link cannot support the required data rate R.

This can be expressed by $p_{\text{out},k,j}(\mathbb{T}'_k) = \Pr(R > C_{k,j})$. The following closed form expression was derived in [Doh03, p.65]

$$
\begin{aligned}
p_{\text{out},k,j}(\mathbb{T}'_k) &= \Pr(R > C_{k,j}) \\
&= \Pr\left(\sum_{i \in \mathbb{T}'_k} g_{k,i,j} |h_{k,i,j}|^2 < \left(2^{\frac{R}{W}} - 1\right) W N_0 \right) \\
&= \sum_{i \in \mathbb{T}'_k} K_i \left(1 - e^{-g_{k,i,j}^{-1} Q_k}\right) \\
\text{where} \quad K_i &= \prod_{\substack{i' \in \mathbb{T}'_k \\ i' \neq i}} \left[\frac{g_{k,i,j}}{g_{k,i,j} - g_{k,i',j}} \right] .
\end{aligned} \tag{4}
$$

The parameters including data rate, bandwidth and noise are collected in the variable $Q_k = \left(2^{\frac{R}{W}} - 1\right) W N_0$.

3 Optimization Problem

The outage probability of node j in hop k depends on the outage probability $p_{out,k,j}(\mathbb{T}'_k)$, on the set of $\mathbb{T}'_k$ active nodes and on the set of $\tilde{\mathbb{T}}_k$ inactive nodes, which depends itself on the outage probability in hop $k-1$. Note that the intersection of $\mathbb{T}'_k$ and $\tilde{\mathbb{T}}_k$ is zero, so that

$$\mathbb{T}'_k \cap \tilde{\mathbb{T}}_k = 0, \quad \forall k \tag{5}$$

holds. Since the outage probabilities for the nodes in each VAA are not equal, no common distribution function can be applied here. For each VAA there exists 2^{t_k} different sets $\mathbb{T}'_k$ of transmitting nodes. The outage probability of a MISO system with a set of $\mathbb{T}'_k$ active nodes can be described by $\Pr(\mathbb{T}'_k) \cdot p_{\text{out},k,j}(\mathbb{T}'_k)$ that concludes that the outage probability $P_{\text{out},k,j}$ can be expressed by the sum over all possible outage scenarios, i.e. the sum over all possible sets of active nodes $\mathbb{T}'_k$.

$$P_{\text{out},k,j} = \sum_{i=1}^{2^{t_k}} \prod_{a_i \in \mathbb{T}'_{k,i}} \left[1 - P_{\text{out},k-1,a_i}\right] \prod_{b_i \in \tilde{\mathbb{T}}_{k,i}} \left[P_{\text{out},k-1,b_i}\right] \cdot p_{\text{out},k,j}(\mathbb{T}'_{k,i}) \tag{6}$$

An outage occurs in one hop, if all nodes in the active VAA cannot decode the message correctly. Taking into account that the single outage probabilities can be unequal, the outage probability per hop has to be defined as

$$P_{\text{out},k} = \prod_{j=1}^{r_k} P_{\text{out},k,j} . \tag{7}$$

The e2e outage probability for the adaptive relaying transmission can be expressed as

$$P_{\text{e2e}} = \sum_{k=1}^{K} P_{\text{out},k}. \tag{8}$$

For the required total power for this transmission scheme follows

$$\mathcal{P}_{\text{total}} = \sum_{k=1}^{K} \sum_{i=1}^{t_k} \mathcal{P}_{k,i}(1 - P_{\text{out},k-1,i}). \tag{9}$$

Eq. 9 considers the inactive nodes to stop transmission if they are in outage. To optimize the transmit power in the network, we formulate the following optimization problem

$$\begin{aligned} &\text{minimize } \mathcal{P}_{\text{total}} = \sum_{k=1}^{K} \sum_{i=1}^{t_k} \mathcal{P}_{k,i}(1 - P_{\text{out},k-1,i}) \\ &\text{subject to } P_{\text{e2e}} = \sum_{k=1}^{K} P_{\text{out},k} \leq e\ . \end{aligned} \tag{10}$$

Generally the problem stated in Eq. 10 is not convex. It can be shown that for a low outage probability constraint this problem can assumed to be convex [BV04]. We solved this optimization problem numerically with the help of common optimization toolboxes.

4 Performance Evaluation

In this work we investigated the performance of the adaptive transmission scheme for an arbitrary generated multi-hop network. It is assumed that the e2e outage probability should meet a maximum constraint of $1\,\%$. The total bandwidth is $W = 5\,\text{MHz}$, the pathloss exponent ϵ is assumed to be 3 and the noise power spectral density is assumed to be $N_0 = -174\,\text{dBm/Hz}$. We assume a $K = 4$ hop multi-hop system. The distance between source and destination is $4\,\text{km}$. Fig. 2 represents the network setup. Nodes are represented by circles, the solid lines show which nodes are mapped to one VAA. How the mapping of nodes to VAAs is archived is out of the scope of this paper. To show the improvement of the adaptive network compared to the non adaptive one, proposed by [LWK08], Fig. 3 shows the archived curves for the required total power versus data rate in Mbps. As a reference Fig. 3 shows also the required power for SISO links over the given distance of $4\,\text{km}$. It can be observed that the adaptive network with optimized transmit power offers a significant saving of total power compared to the non adaptive structure. This can be interpreted by the capability of the adaptive scheme to adapt to a new network setup if some nodes are in outage. The non adaptive scheme has to rely on the worst link in the network. Only if all nodes in on VAA fail, the adaptive scheme can be considered to be in outage.

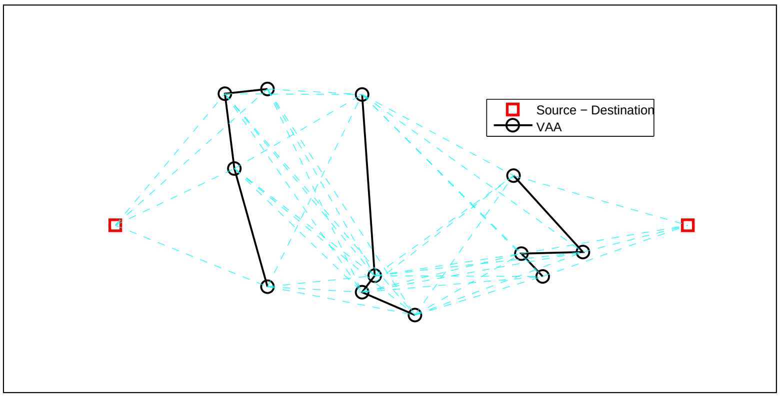

Figure 2: Setup of an asymmetric network with $K = 4$ hops and VAAs consisting of $t_k = 4$ nodes.

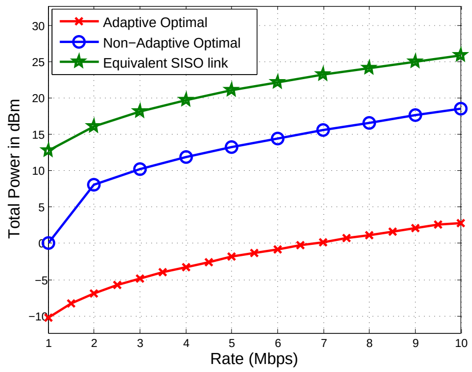

Figure 3: Total power consumption versus data rate for the asymmetric network with $K = 4$ hops and $t_k = 4$ nodes per hop.

5 Conclusions and Outlook

In this paper we introduced and investigated the performance of a power allocation scheme for adaptive asymmetric relay networks. It could be shown that power allocation in adaptive networks provides a huge potential to save overall transmit power. The solution for the optimal transmit power was obtained by common optimization tools and since there has no closed form been found yet, it is not possible to implement this scheme in practical setups. Since the performance of the proposed scheme is much higher than a non adaptive scheme, it is worth to spend further investigations in order to find a closed form solution that can be implemented in practical applications.

References

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambrigde University Press, 2004.

[Doh03] M. Dohler. *Virtual Antenna Arrays*. PhD thesis, King's College London, U.K., November 2003.

[LBW+08] Y. Lang, C. Bockelmann, D. Wübben, A. Dekorsy, and M. Söllner. A Novel Resource Allocation Strategy for Distributed MIMO Multi-hop Multi-commodity Communications. In *International ITG Workshop on Smart Antennas (WSA 08)*, Darmstadt, Germany, February 2008.

[LTW04] J.N. Laneman, D. Tse, and G.W. Wornell. Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior. *ITIT*, 50(12):3062–3080, February 2004.

[LWK08] Y. Lang, D. Wübben, , and K.-D. Kammeyer. Efficient Power Allocation for Outage Restricted Asymmetric Distributed MIMO Multi-Hop Networks. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Cannes, France, September 2008.

[LWK09] Y. Lang, D. Wübben, , and K.-D. Kammeyer. Power Allocations for Adaptive Distributed MIMO Multi-Hop Networks. In *IEEE International Conference on Communications (ICC09)*, Dresden, Germany, June 2009.

Channel Measurements and Performance Analysis of Optical MIMO Multimode Fiber Links

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

Abstract: Multiple-input multiple-output (MIMO) systems 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 optical transmission systems, multi-mode (MM) fibre offers the possibility to transmit different signals by different mode groups. In order to investigate the perspective of the MIMO philosophy within the field of optical transmission systems, channel measurements within a (2×2) MIMO system are carried out. 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 were taken into consideration in order to underline the potential of multi-mode fibres in optical high-data rate MIMO communication systems.

1 Introduction

The never ending desire for increasing the available transmission capacities has attracted a lot of research since Shannon's pioneering work in 1948. Bandwidth extension has long been seen as the only way to increase the channel capacity efficiently. However, as we know now, there are additional - and in particular cases - better ways: The capacity of a transmission channel can be utilized in a very efficient way, if multi-channel techniques treat several "traditional" channels as a whole and a "generalized" channel appears. 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]. Equipping the transmitter as well as the receiver with a number of antennas is nowadays common in the wireless transmission community to increase the channel capacity and has been integrated in a lot of applications (e. g. WLAN 802.11n). MIMO multicarrier techniques are also considered for the next generation of cellular mobile radio systems known as UMTS LTE (long term evolution). However, this MIMO technique isn't limited to wireless communication and a lot of scenarios can be described and outperformed by the MIMO technique.

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]. 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-input and multiple-output, also called MIMO systems. MIMO systems have attracted a lot of attention in the field of wireless communication, where the capacity increases linearly with the minimum number of antennas at both, the transmitter as well as the receiver side [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. Here, the spatial multiplexing 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 filers are implemented.

In general, different mode groups are activated at the transmitter side, which propagate along the fibre with different speed. Together, with the crosstalk between the different mode groups, the classical MIMO channel is formed.

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 general, the most beneficial choice of the number of activated MIMO layers offers a certain degree of design freedom, which substantially affects the performance of MIMO systems.

Against this background, the novel contribution of this paper is that we perform channel measurements within a (2×2) optical MIMO system. Based on channel measurements the perspective of the MIMO philosophy within the field of optical transmission systems is elaborated. The remaining part of this paper is organized as follows: Section 2 introduces the considered optical MIMO channel. The considered optical broadband MIMO channel model is shortly reviewed in section 3, while the associated performance results are presented and interpreted in section 4. Section 5 provides some concluding remarks.

2 MIMO Channel

In this work the potential of the MIMO philosophy is elaborated, based on channel measurements.

For launching light into the fibre, the transmitter side configuration is depicted in 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 an offset of $10\,\mu$m was chosen (off-center launch condition).

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 fibre whereas higher order modes, activated by light launched at an off-center position within the fibre, lead to power radiation pattern concentrated at the off-center of the fibre. Together, with the crosstalk between the different mode groups, the

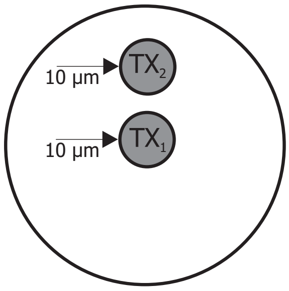

Figure 1: Light launch positions at the transmitter side

classical MIMO channel is formed. The corresponding electrical MIMO channel model can be characterized by the resulting impulse responses as depicted in Fig. 2.

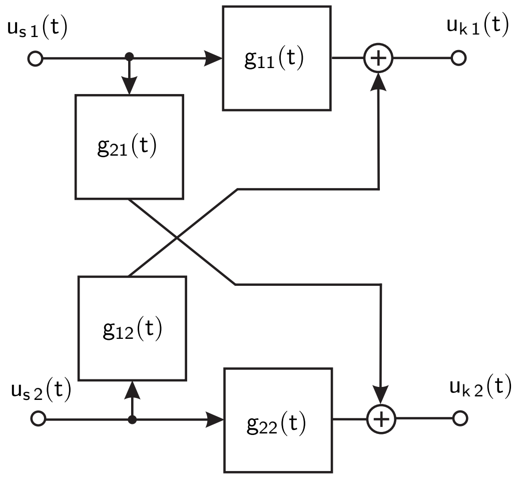

Figure 2: Electrical MIMO system model (example: $n = 2$)

In order to form the MIMO channel different spatial filter have to be implemented at the receiver side. On the one hand a centric filter was used for the low order modes and on the other hand an eccentric filter was chosen for higher order mode groups. Fig. 3 and 4 illustrate the investigated spatial configurations at the receiver side. The centric filter was designed as a centric $15\,\mu$m hole. For the higher order mode groups different eccentric filters such as the ring and the off-center filter were used. Thereby the off-center filter was designed similar to the centric filter but with an eccentric hole ($15\,\mu$m), whereas the ring filter offers a greater detection area compared with the off-center filter with only a $15\,\mu$m hole.

Together with the crosstalk between the different mode groups, the MIMO channel, as

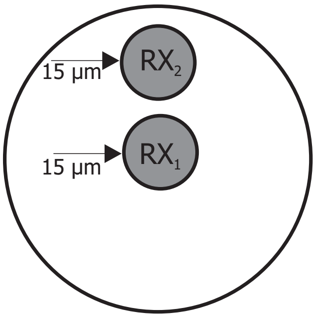

Figure 3: Investigated receiver-side off-center filter configuration

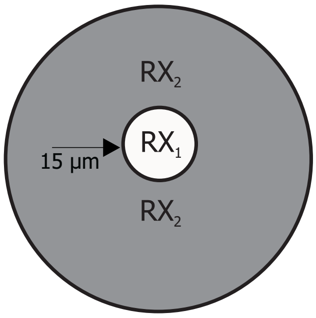

Figure 4: Investigated receiver-side ring filter configuration

highlighted in Fig. 2, will arise.

Fig. 5 shows the experimental setup for the investigated (2×2) MIMO system. The Picosecond Diode Laser System (PiLas) drives a $1,3\,\mu$m single mode diode laser which generates a short optical impulse of about 30 ps. Afterwards the splice equipment was used at the transmitter side to launch light at different radial offsets into a MMF. At one side of the splice equipment the $10\,\mu$m single Mode (SM) fiber was used and at the other side a $50\,\mu$m graded index (GI) MMF with a length of $1,4$ km was connected.

At the receiver side different spatial filters were used. After filtering, a metal-semiconductor-metal (MSM detector) converter was used for optical-electrical conversion. Finally, the MIMO channel transfer functions were obtained via de-convolution with the known input signal.

The novel contribution of this paper is to investigate the potential of a (2×2) optical MIMO system based on channel measurements which are carried out, without any loss of generality, separately. The obtained MIMO channel impulse responses, with respect to the pulse frequency $f_\mathrm{T} = 1/T_\mathrm{s} = 5,12$ GHz, are depicted in Fig. 6–9 when using the ring-filter configuration.

3 Optical Broadband MIMO System Model

When considering a frequency-selective MIMO link, composed of n_T optical inputs and n_R optical outputs, the resulting electrical discrete-time block-oriented system can be modelled by

$$\mathbf{u} = \mathbf{H} \cdot \mathbf{c} + \mathbf{w} \ . \tag{1}$$

In (1), $\mathbf{c}$ is the $(N_\mathrm{T} \times 1)$ transmitted signal vector containing the input symbols transmitted over n_T optical inputs in K consecutive time slots, i. e., $N_\mathrm{T} = K\,n_\mathrm{T}$. This vector can be decomposed into n_T transmitter-specific signal vectors $\mathbf{c}_\mu$ according to

$$\mathbf{c} = \left(\mathbf{c}_1^\mathrm{T}, \ldots, \mathbf{c}_\mu^\mathrm{T}, \ldots, \mathbf{c}_{n_\mathrm{T}}^\mathrm{T}\right)^\mathrm{T} \ . \tag{2}$$

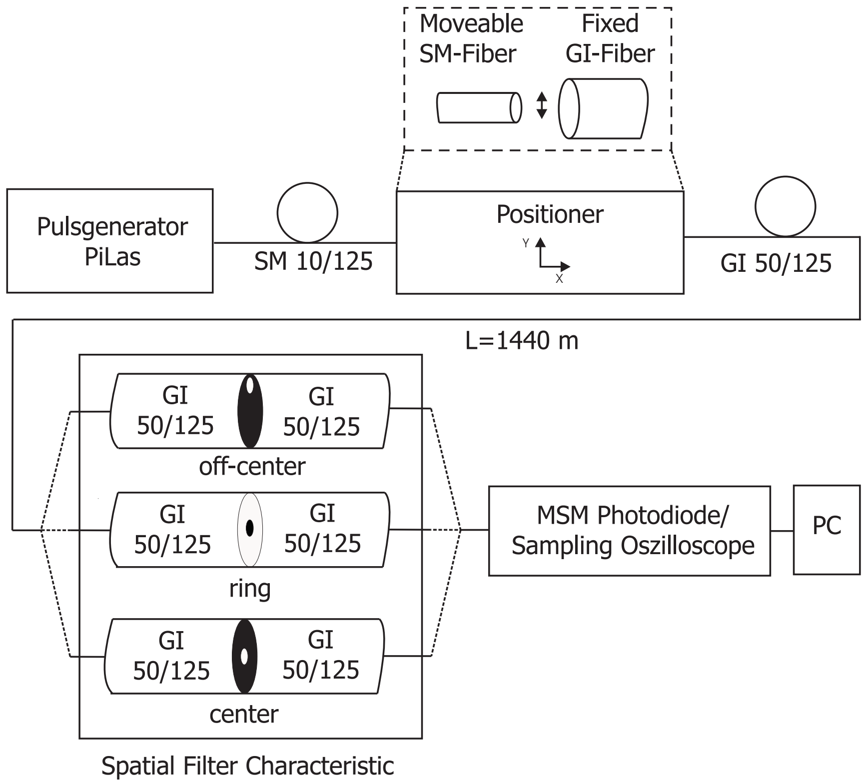

Figure 5: Measurement setup

In (2), the $(K \times 1)$ input-specific signal vector $\mathbf{c}_\mu$ transmitted by the optical input μ (with $\mu = 1, \dots, n_\mathrm{T}$) is modelled by

$$\mathbf{c}_\mu = (c_{1\,\mu}, \dots, c_{k\,\mu}, \dots, c_{K\,\mu})^\mathrm{T} \quad . \tag{3}$$

The $(N_\mathrm{R} \times 1)$ received signal vector $\mathbf{u}$, defined in (1), can again be decomposed into n_R output-specific signal vectors $\mathbf{u}_\nu$ (with $\nu = 1, \dots, n_\mathrm{R}$) of the length $K + L_\mathrm{c}$, i. e., $N_\mathrm{R} = (K + L_\mathrm{c})\, n_\mathrm{R}$, and results in

$$\mathbf{u} = \left(\mathbf{u}_1^\mathrm{T}, \dots, \mathbf{u}_\nu^\mathrm{T}, \dots, \mathbf{u}_{n_\mathrm{R}}^\mathrm{T}\right)^\mathrm{T} \quad . \tag{4}$$

By taking the $(L_\mathrm{c} + 1)$ non-zero elements of the resulting symbol rate sampled overall channel impulse response between the μth input and νth output into account, the output-specific received vector $\mathbf{u}_\nu$ has to be extended by L_c elements, compared to the transmitted input-specific signal vector $\mathbf{c}_\mu$ defined in (3). The $((K+L_\mathrm{c}) \times 1)$ signal vector $\mathbf{u}_\nu$ received by the optical output ν (with $\nu = 1, \dots, n_\mathrm{R}$) can be constructed, including the extension through the multipath propagation, as follows

$$\mathbf{u}_\nu = \left(u_{1\,\nu}, u_{2\,\nu}, \dots, u_{(K+L_\mathrm{c})\,\nu}\right)^\mathrm{T} \quad . \tag{5}$$

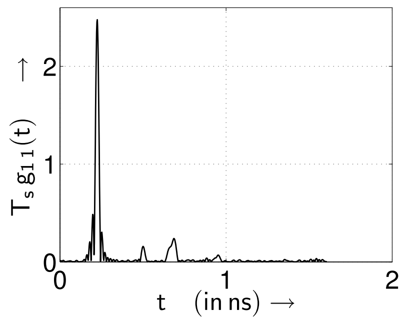

Figure 6: Measured electrical MIMO impulse response $g_{1\,1}(t)$ (center launch condition)

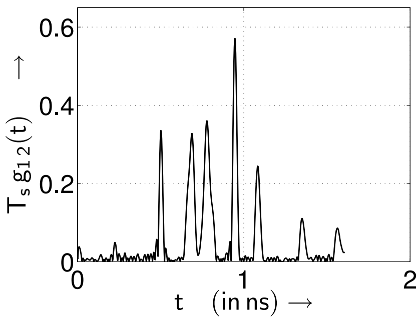

Figure 7: Measured electrical MIMO impulse response $g_{1\,2}(t)$ (offset launch condition, offset $10\,\mu$m)

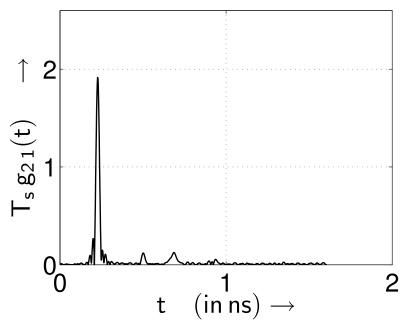

Figure 8: Measured electrical MIMO impulse response $g_{2\,1}(t)$ (center launch condition)

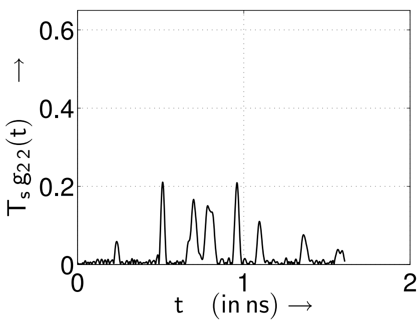

Figure 9: Measured electrical MIMO impulse response $g_{2\,2}(t)$ (offset launch condition, offset $10\,\mu$m)

Similarly, in (1) the $(N_\mathrm{R} \times 1)$ noise vector $\mathbf{w}$ results in

$$\mathbf{w} = \left(\mathbf{w}_1^\mathrm{T}, \ldots, \mathbf{w}_\nu^\mathrm{T}, \ldots, \mathbf{w}_{n_\mathrm{R}}^\mathrm{T}\right)^\mathrm{T} \quad . \tag{6}$$

The vector $\mathbf{w}$ of the additive, white Gaussian noise (AWGN) can still be decomposed into n_R transmitter-specific signal vectors $\mathbf{w}_\nu$ (with $\nu = 1, \ldots, n_\mathrm{R}$) according to

$$\mathbf{w}_\nu = \left(w_{1\,\nu}, w_{2\,\nu}, \ldots, w_{(K+L_\mathrm{c})\,\nu}\right)^\mathrm{T} \quad . \tag{7}$$

Finally, the $(N_\mathrm{R} \times N_\mathrm{T})$ system matrix $\mathbf{H}$ of the block-oriented system model, introduced in (1), results in

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{1\,1} & \cdots & \mathbf{H}_{1\,n_\mathrm{T}} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{n_\mathrm{R}\,1} & \cdots & \mathbf{H}_{n_\mathrm{R}\,n_\mathrm{T}} \end{bmatrix} \quad , \tag{8}$$

and consists of $n_\mathrm{R}\,n_\mathrm{T}$ single-input single-output (SISO) channel matrices $\mathbf{H}_{\nu\,\mu}$ (with $\nu = 1, \ldots, n_\mathrm{R}$ and $\mu = 1, \ldots, n_\mathrm{T}$). The system description, called spatio-temporal vector coding (STVC), was introduced by RALEIGH for wireless broadband MIMO channels. Every of these matrices $\mathbf{H}_{\nu\,\mu}$ with the dimension $((K + L_\mathrm{c}) \times K)$ describes the influence of the channel from transmitter μ to receiver ν including transmit and receive filtering. The channel convolution matrix $\mathbf{H}_{\nu\,\mu}$ between the μth input and νth output is obtained by taking the $(L_\mathrm{c} + 1)$ non-zero elements of resulting symbol rate sampled overall impulse response into account and results in:

$$\mathbf{H}_{\nu\,\mu} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & \vdots \\ h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & h_2 & h_1 & \cdots & h_0 \\ h_{L_\mathrm{c}} & \vdots & h_2 & \cdots & h_1 \\ 0 & h_{L_\mathrm{c}} & \vdots & \cdots & h_2 \\ 0 & 0 & h_{L_\mathrm{c}} & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & h_{L_\mathrm{c}} \end{bmatrix} . \qquad (9)$$

The interference, 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) [Hay91] of the system matrix $\mathbf{H}$, which can be written as $\mathbf{H} = \mathbf{S} \cdot \mathbf{V} \cdot \mathbf{D}^\mathrm{H}$, where $\mathbf{S}$ and $\mathbf{D}^\mathrm{H}$ are unitary matrices and $\mathbf{V}$ is a real-valued diagonal matrix of the positive square roots of the eigenvalues of the matrix $\mathbf{H}^\mathrm{H}\,\mathbf{H}$ sorted in descending order[1]. The SDM (spatial division multiplexing) MIMO data vector $\mathbf{c}$ is now multiplied by the matrix $\mathbf{D}$ before transmission. In turn, the receiver multiplies the received vector $\mathbf{u}$ by the matrix $\mathbf{S}^\mathrm{H}$. Thereby neither the transmit power nor the noise power is enhanced. The overall transmission relationship is defined as

$$\mathbf{y} = \mathbf{S}^\mathrm{H}\,(\mathbf{H} \cdot \mathbf{D} \cdot \mathbf{c} + \mathbf{w}) = \mathbf{V} \cdot \mathbf{c} + \tilde{\mathbf{w}} \,. \qquad (10)$$

As a consequence of the processing in (10), the channel matrix $\mathbf{H}$ is transformed into independent, non-interfering layers having unequal gains.

4 Performance Analysis

Exemplarily, in this work the efficiency of fixed transmission modes is investigated. 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 while maintaining the best possible integrity, i. e., bit-error rate, an appropriate number of MIMO layers has to be used, which depends on the specific transmission mode, as detailed in Table 1. The obtained bit-error rate (BER) curves are depicted in Fig. 10 and

[1]The transpose and conjugate transpose (Hermitian) of $\mathbf{D}$ are denoted by $\mathbf{D}^\mathrm{T}$ and $\mathbf{D}^\mathrm{H}$, respectively.

Table 1: Investigated transmission modes

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

11 for the different ASK (amplitude shift keying) constellation sizes of Table 1. For the investigations, a radial offset of $10\,\mu$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 layers have to be activated in order to achieve the best BERs. More explicitly, the main goal was to investigate the improvement by adopting MIMO principles known from wireless communication. Comparing the investigated spatial filter configurations, the advantage of the ring filter is clearly recognizable.

As elaborated by computer simulations, our first measurement results show the potential of the MIMO philosophy in optical communication systems clearly. However, there are still some unsolved problems in adapting well-known approaches from wireless communications and in this way the carried-out research has some limitations: the investigated modulation formats, combined with the SVD-based pre-processing, seems to be non-optimal for optical communication systems. This remains as an open point for further investigations.

5 Conclusion

In this contribution channel measurements within a (2×2) MIMO system are carried out and the perspective of the MIMO philosophy within the field of optical transmission systems is investigated. Our first results, obtained by channel measurements and computer simulations, show the potential of MIMO techniques in the field of optical transmission systems clearly. In combination with appropriate signal processing strategies, an improvement in the overall BER was obtained.

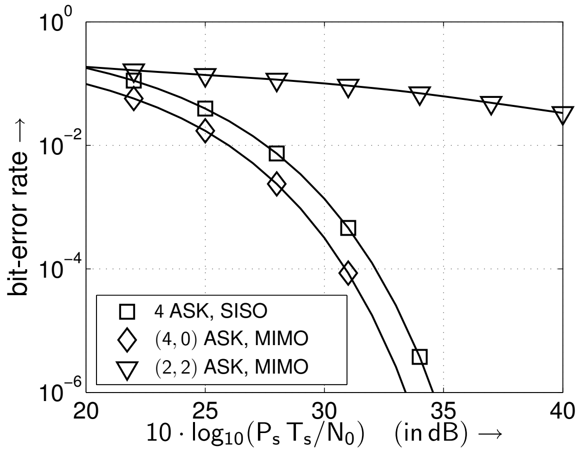

Figure 10: BER when using the transmission modes introduced in Table 1 and transmitting 2 bit/s/Hz over frequency selective optical channels using the off-filter configuration

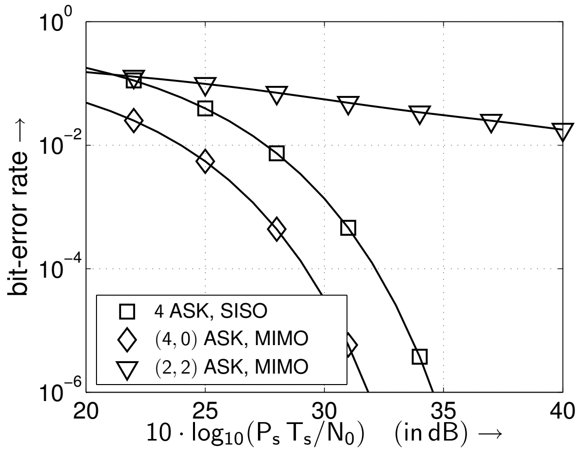

Figure 11: BER when using the transmission modes introduced in Table 1 and transmitting 2 bit/s/Hz over frequency selective optical channels using the ring-filter configuration

References

[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.

[Hay91] S. S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, N. J., 2. aufl edition, 1991.

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

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

[RJ99] G. 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.

Analysis of conference attendees mobility patterns

Tomasz Rybak
Faculty of Computer Science at Bialystok University of Technology

tomasz.rybak@post.pl

Abstract: Article presents analysis of data about movement of conference participants gathered from RFID tags. Active RFID tags were used during experiment to control range of transmission; custom tags were used instead of of-the-shelf ones to be able to control data send to the readers. Article shows that it is possible to determine whether tag wearers were conference organisers and, in some cases, details about personal identity of attendees can be inferred using available data. Article also shows ways of computing positions basing on RFID signals and computing activity of conference attendees using entropy. This can be used to infer further information about people wearing tags.
Keywords: RFID, privacy, identity, position estimation, entropy

1 Introduction

Sputnik is the RFID (Radio Frequency IDentifier) system build using OpenBeacon.org components. It was used during 23rd, 24th, and 25th Chaos Computer Congress, and during The Last HOPE conference to gather data describing movement and positions of volunteers participating in project.

This article does not describe internal details of the Sputnik system — they have been described in other presentations and publications:

- Milosch Meriac's presentations describing Sputnik and other RFID systems
  - OpenPCD/OpenPICC [Meriac(2006a)]
  - Project Sputnik [Meriac(2006b)]
  - Inside Sputnik and OpenBeacon — Smart Dust for the Masses [Meriac(2007a)]
  - Practical RFID attacks [Meriac(2007b)]
- Author's presentation from 24C3 describing data gathered during 23C3
  - Analysis of Sputnik Data from 23C3 [Rybak(2007)]

- Ciro Cattuto's presentation showing usage of the Sputnik system to find social connections of participants during 5C3
  - Mining social contacts with active RFID [Cattuto(2008)]

The main purpose of this article is to analysis of data send by unprotected RFID tags. Various methods of analysing data send by tags are presented, risks that those methods pose to privacy are analysed, and then those methods are used to attack privacy of conference attendees.

This article presents results of analysing data coming from only few of chosen tags. One could easily find more tags that are showing behaviour similar to ones shown here. Tags described in this article were chosen to show information that can be derived from public data, and to show privacy risks that arise while usage unprotected RFID tags.

Techniques similar to described in this article can also be used to analyse data gathered from devices using other technologies, like Bluetooth, WiFi or cellular networks. Such analysis was done by Vassili Kostakos (([Kostakos()]) for BlueTooth. Analysis of data gathered from cellular network was presented in Nature ([Gonzalez et al.(2008)Gonzalez, Hidalgo, and Barabasi]) and was done in PhD thesis by Nathan Eagle [Eagle(2005)].

2 Data gathered by Sputnik system

After every event during which Sputnik was deployed recorder data set was made available. Formats of released data sets are described in previous publication [Rybak(2007)]. The largest part of data was made available as text file with values separated by spaces. Each line was one packet received from single reader, consisting of tag id, time of receiving of packet, station ID, signal strength, internal counter value, and internal flags. Textual form of data set was chosen because it made it easy to import into different software packages. Data available as text files can be read by humans and it is possible to analyse it by using well-known Unix tools (grep, awk, perl).

During 24th Congress 215 tags were active. Additionally 23 strange packets were received; they were probably generated by WiFi and BlueTooth devices active during Congress. Because tags were sending readings more than once every second (and not once every few seconds like during previous Congress) 27.9 million packets has been read. It is 2.5 times more than during previous Congress in spite of having 4 times less active tags. Number of tags active during previous conference is only estimated based on number of sold units; exact number of active tags is unknown because of lost identities of devices.

Readers were placed to ensure that at least two of them were able to receive packets transmitted by tag (Figure 1). When packets transmitted by tag are received by more than one reader it is possible to estimate position by computing distance of tag from all stations that are able to communicate with tag.

More deployed stations and increased frequency of transmission increased probability that transmission was received and recorded by system. To increase number of received packets and ease calculations of positions even further, pattern of changes of power of signal was changed. The new tags cycle through all power levels so the highest power is active for 1/4th of the time. This change is visible in data set (Figure 2) and increased probability of receiving signal by more than one reader.

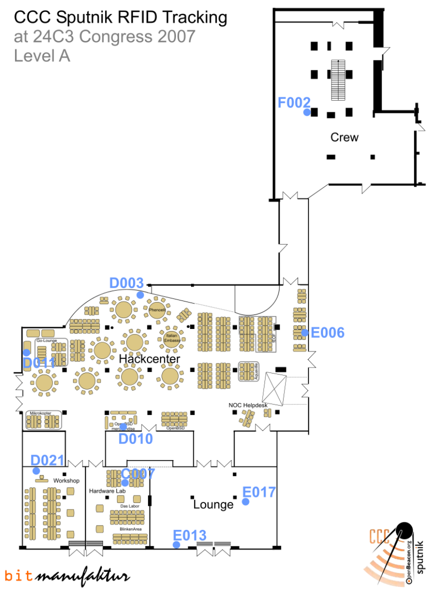

Figure 1: Readers during 24C3 on level A (created by Milosch Meriac, available at Sputnik web site)

Strength	count
0	8669189
1	7103390
2	6637093
3	5496778

Figure 2: Number of received packets according to their strengths

During The Last HOPE tags were included in conference badges. 1281 tags were active and packets send by them were recorded by system. Format of file containing data send by tags was the same as during 24C3. More than 176 million of packets were received. During this conference for the first time Attendee Meta-Data system was deployed; it allowed for tag wearers to record their nicks, emails, phone numbers and interests. AMD team could then use this data to remind people about interesting talks (basing on provided areas of interests) and to connect attendees with similar sets of interests.

The highest number of packets during Congress in Berlin were received by readers placed in biggest lecture room (Saal 1) and in the most frequently visited place — Canteen, where food was served and where people could socialise. Participants of conference in New York were mostly present in lecture rooms and in “vendor area” — place where people were able to experiment with hardware and software and work on various projects.

Figure 3 shows comparison of activities of tags during 23rd and 24th Congress. As common sense suggests, the most activity can be observed during afternoon; in the night most people were leaving conference venue (although few of the most dedicated stayed for all the night) to start returning in the morning. Comparison of charts shows that there was more data transmitted during 24th Congress, and that during 23C3 there were few minutes during which no packets were recorded. Reasons for lack of data is not known. Similarly Figure 4 shows comparison of activities during 24th and 25th Congress. During 25C3 there were 4 hours without any readings — there is no network data available for this time. Again, when comparing activity of previous and the next conference one can see increase in number of recorder packets. It suggests growing interest in Sputnik project among conference attendees.

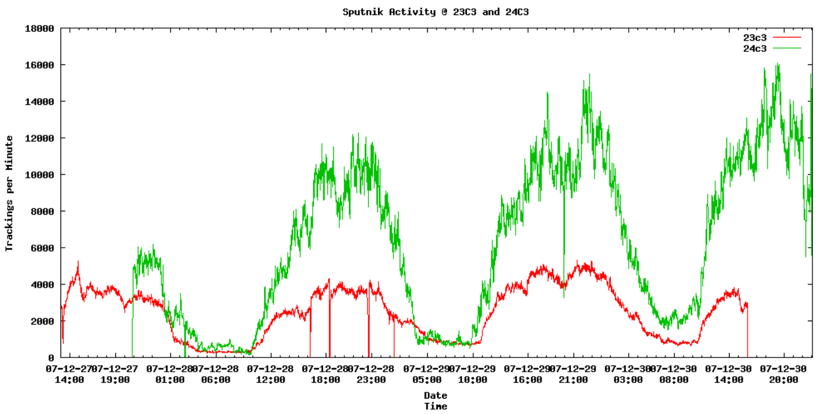

Figure 3: Number of packets read by minute during 23C3 and 24C3

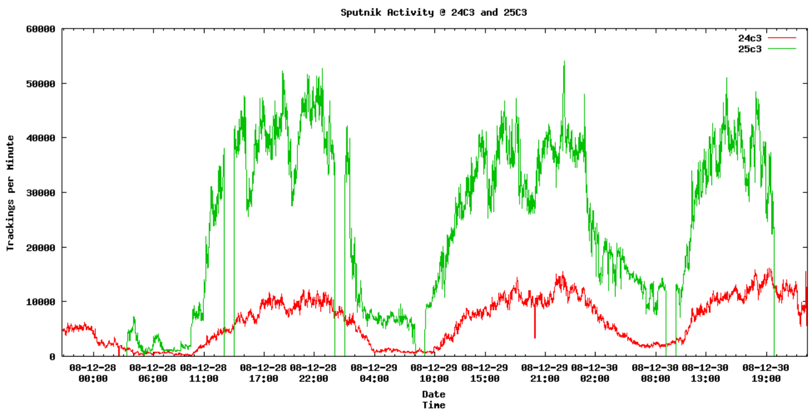

Figure 4: Number of packets read by minute 24C3 and 25C3

2.1 Attendee Meta Data

Attendee Meta Data (AMD) was project indented to increase value of data recorded by Sputnik system by gathering data about conference attendees. It was used during The Last HOPE and, in more limited manner, during 25C3. Tag wearers were able to log in into web page and to provide team with their nicks, phone numbers and email addresses. It was also possible for project participants to contact each other — system did not publish contact information but allowed for one person to send “ping” to another person. Usually it was used by people with similar sets of interests to find themselves in the crowd. System was also sending notifications (using either email of SMS) about interesting lectures, basing on data about their interests provided by participants.

Entire set of meta-data (excluding real names, phone numbers, and emails) were made available after The Last HOPE conference as set of CSV files containing data stored in database tables. Released data included countries of origin of participants, their interests, inter-person messages, list of attended talks, and snapshots of positions.

Although AMD team made effort to preserve privacy of participants, they involuntarily leaked some of the personal data. Emails and phone numbers were correctly removed and not released to the public. Nicks were removed from table “person” (each nick was changed into “userNNNN”), but they were duplicated in table “talks_presence”. This table still had correct nicks when database dump was released to the world. This error was the result of denormalisation of data, which probably was done to speed-up visualisation of positions of participants in conference venue. Although this was not the serious breach of privacy it suggests each data set containing personal data is potential risk for privacy.

2.2 Rebuilding of lost tag sequences

Because of a bug in aggregating software used during the first deployment of Sputnik system identities of all tags were lost from recorder data set. Description of attempts of recovering lost sequences can be found in author's previous article describing RFID systems [Rybak(2007)].

Attempts in reconstructing lost identities were moderately successful. Availability of the full data set from 24C3 made it clear why full recovery of lost sequences was not successful. Basing on analysis of firmware and observing behaviour of sample tag using USB reader I assumed that all sequences generated by tags increased linearly without large gaps. Some tags, however, had problems with sequence increments which caused large gaps in them. Frequent resets (probable causes of gaps in sequences) might be caused by loose connection to battery, shortcut in tag, shortcut caused by connecting clothes, broken tags, or wearer trying to do play with tag.

Regardless of reasons, non-linear sequences make it almost impossible to rebuild sequences in large quantities in automatic manner. Rebuilding sequences is necessary to recover lost sequences. During 24th Congress 39 of 215 tags presented non-linear behaviour — if similar ratio of tags with such behaviour was present during previous Congress it would mean that large portion of sequences could not be recovered. Such non-connected points make it also difficult for recovering values from "normal" tags. Such situation does not harm analysis of entire data set assuming that there is no other errors present in recorded packet.

Data recorder during 25C3 contains much less non-linear sequences, probably because of new hardware and new firmware. New software was based on old firmware, and was enhanced by Cattuto's team ([Cattuto(2008)]) to detect social interactions between participants. Although early version of tags still were sending broken sequences, final tags were generating much more linear sequences than during previous conferences.

3 Simple analysis

To start with analysis one needs only data send by tags and placement of readers that were recording packets. Such data set does not contain exact positions of tags. Even knowledge of regions of conference venue tag was present at given time can give insight into behaviour of person wearing tag. Those initial ideas can later be investigated on by using more sophisticated methods which may lead to either rejection of those ideas or to gaining deeper insight into behaviour of people.

Each of the figures shown below shows stations that were able to receive packets send by particular tags at given times. Knowing locations of RFID readers one can see where and for how long wearer of tag stayed during the conference. Each analysis must also take into consideration possibility of leaving tag in one place — each person can disable it or stop wearing them at any moment. Entropy, described in Section 5, allows for distinguishing between leaving tag unattended and staying in one place while still wearing tag.

Figure 5 shows Engel (volunteer helping during organisation of conference) which was responsible for checking tickets. In the beginning he or she was for some time in Heaven (station F002), and then for the rest of Congress was near entry. The Engels that were near entry were responsible for checking and selling tickets.

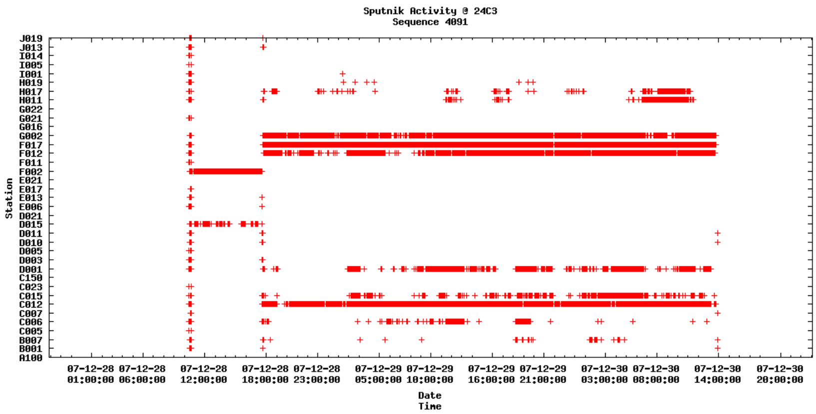

Figure 5: Engel responsible for tickets

We can assume that tag 4091 was worn by Engel because many packets send by it were received by station F002. This station was placed in restricted area (called Heaven) and was only able to read tags of people sitting there and being in corridor going to Heaven. This corridor also went to the toilets, but people going to bathrooms can be filtered out by requiring large number of consecutive readings from station F002. Further analysis of data shows that there is not many packets read by station F002 and other station. It supports hypothesis that packets received by station F002 can be used to identify volunteers and organisers of the conference.

Figure 6 shows someone who was attending many lectures. He or she was staying in one room for about one hour and then moving to another room. This person was leaving conference venue for the night, leaving between noon and 1:00, and coming in between 10:00 and 12:00 the next day.

Figure 7 shows packets send by tag 4095 most probably worn by Milosch Meriac — creator of Sputnik project. This tag was in the area where Sputnik project had its table (green readers) for almost entire conference. The only time when this person left Sputnik table was during lecture about Sputnik project (blue stations placed in Saal 2).

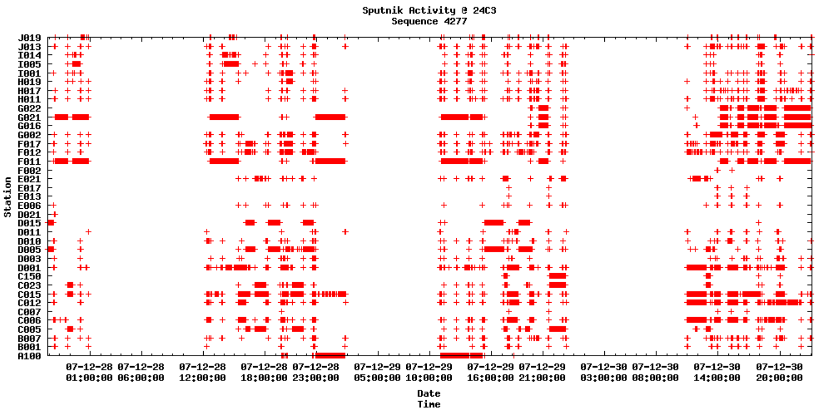

Figure 6: Someone who attended many lectures

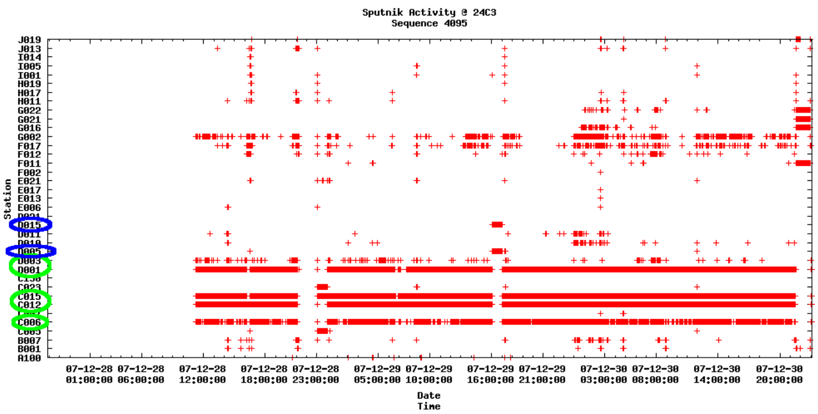

Figure 7: Milosch Meriac

Very little data is needed to deduct information about behaviour of the person. Additional knowledge (either about this person or about event, architecture, and so on) can be used to gain more knowledge, increasing probability of identifying this person or group of persons. Here, using very simple methods, I was able to distinguish organisers of conference and to deduce which tag was worn by creator of the Sputnik system.

4 Calculating positions of tags

Knowing exact positions of readers and power of signal's strength (transmitted inside each packet) it is possible to calculate position of tags. Weighted sum of positions of readers able to receive packets was used to compute position of tag.

Weights used in equations for calculating positions of tags depend on strengths of received packets. I used three different ways of calculating positions, using linear, quadratic, or exponential weights. All of them give very similar results. Exponential weights generate more "spread" positions of tags, while linear method generated more "smooth" plot of positions calculated for each of the tags. Differences of positions calculated by using three different methods mere not significant, though.

$$Pos_l(X) = \frac{\sum_{i=0}^{N} S_i(X) * (4 - strength_i)}{\sum_{i=0}^{N} (4 - strength_i)}$$

$$Pos_q(X) = \frac{\sum_{i=0}^{N} S_i(X) * (4 - strength_i)^2}{\sum_{i=0}^{N} (4 - strength_i)^2}$$

$$Pos_e(X) = \frac{\sum_{i=0}^{N} S_i(X) * 2^{4-strength_i}}{\sum_{i=0}^{N} 2^{4-strength_i}}$$

To calculate position in given period of time we need few packets. During 10-second interval system was seeing from 5 to 20 packets, sometimes up to 30 when subject was not moving (staying in one room). Minimal strength of packets received by particular station was used to compute weight for calculating position of tag.

Described method of computing position (i.e. usage of weighted sum of positions of known reference points) is similar to method used in GPS (Global Positioning System). This article uses differences in strengths of signals to compute position, not difference in time of arrival of signal that is used in GPS. It is not possible to use time of signal propagation in Sputnik system because of very small distance between tag and reading stations. It would require extremely expensive equipment to detect small periods of time that it takes for electromagnetic wave to travel through one room.

This method results in independence of all calculated positions; in other words computing position at time T does not require knowledge of any of previous or future positions. This method does not give very good results though. It is very sensitive to minimal visible strength of packet received by stations. It leads to very rapid changes in calculated positions. Signal emitted by Sputnik tags is weakened by human body (and any water molecules, e.g. wet walls) so one would need find other way of calculating positions if greater accuracy is needed. It is often enough to determine if tag is in particular room or not, even without knowing exact part of the room person is in. This allows for determining if someone attended lecture or not — so it is good starting point for analysing behaviour of person.

Figures 9, 10, and 8 show positions for the same tag calculated by using linear, quadratical and exponential methods. Calculated sequences of positions do not differ much.

All Sputnik tags offer only four levels of transmission strength. It results in very small number of possible combinations of stations and strengths of packets received by them. One way of increasing accuracy is to increase number of distinct signal strength levels, but this could require using different (more sophisticated or more expensive) hardware.

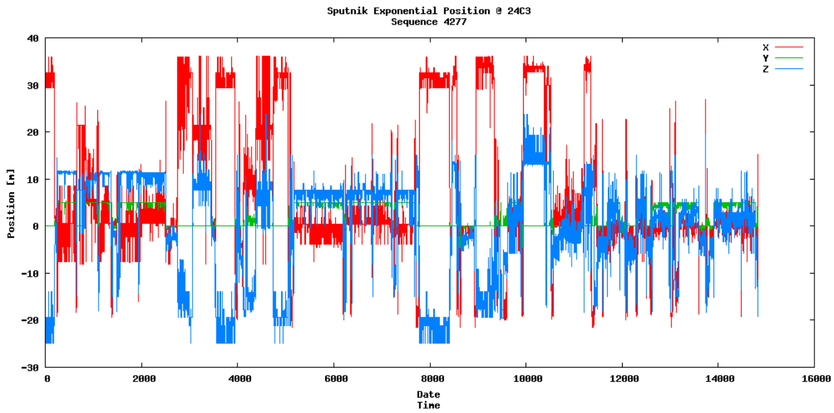

Figure 8: Generated sequence of exponential positions

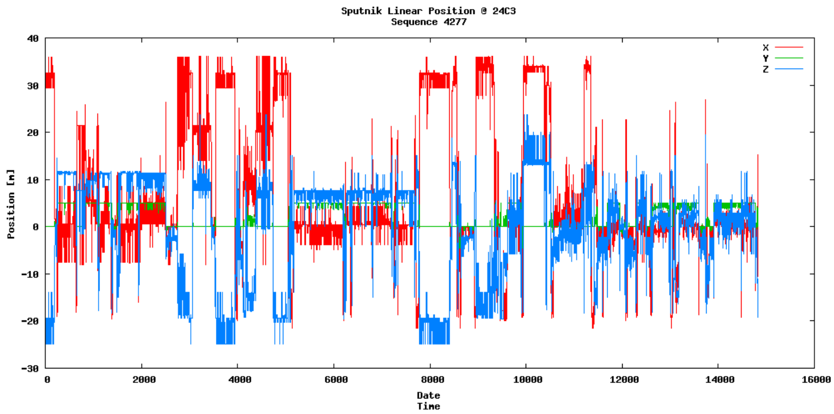

Figure 9: Generated sequence of linear positions

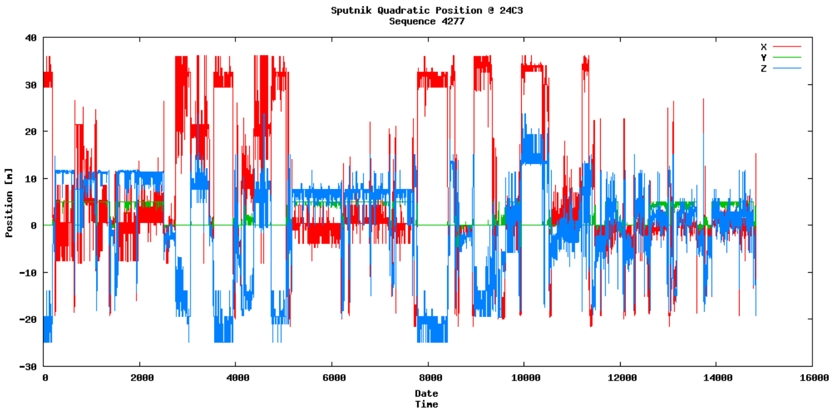

Figure 10: Generated sequence of quadratic positions

Another possible solution would be to increase number of reading stations and to decrease power level of transmission. First disadvantage of this solution is increased complexity of entire system and price of used hardware. Second disadvantage is increased risk of blocking weaker signals by human body or walls.

One can try to remove rapid changes in calculated positions by using interpolation of calculated values. This could, however, introduce some errors in final positions. It would also require knowing future positions. Such method of calculating positions can be used in off-line analysis but would be impossible in real-time analysis and visualisations.

Even solving problem with inaccurate positions does not gives answer for problem with missing data. There are many possibilities; one is to repeat the most recent position until new is available. It does not require any sophisticated analysis of signal. Another solution is interpolation of the last known position and new position, as soon as it is known. It requires knowledge of more than one position so this method is not appropriate for real-time analysis.

Sputnik system is to simple to calculate exact position of tags during the conference. As shown in previous section it is not always necessary to know exact position — one can come to interesting conclusions by having only rough positions of tag wearers. It would be useful to find better ways of calculating positions but it requires further research. Computing more detailed position inside building is interesting field of research but requires much experiments and researching through many different implementations.

5 Entropy

Entropy is used to determine how much information is present in data sequence. It can be used to distinguish between signals with different types of changes. According to Shannon it can be calculated using equation $E = \sum_{i=0}^{N} -p_i log(p_i)$. Entropy was used by Nathan

Eagle in his PhD thesis [Eagle(2005)] to describe differences between life styles of different social groups by analysing data coming from cellular phones. I defined probability of Sputnik signal as number of packets of given strength received by particular station divided by number of all packets recorded. This allows for direct usage of Shannon equation for computing entropy of Sputnik signal.

Entropy is calculated for short periods of time (usually few minutes) and can describe behaviour of person in this time. It can be used to determine whether someone was sitting in the one place (e.g. lecture room), which results in low values of entropy, or moving from one location to another, resulting in high entropy.

Figure 11 shows entropy of Engel responsible for tickets, previously shown in Figure 5. Entropy is small during the time spend in Heaven which suggests staying in one place; entropy is large while this person was on level B near the entrance, suggesting that person was not sitting but moving from place to place.

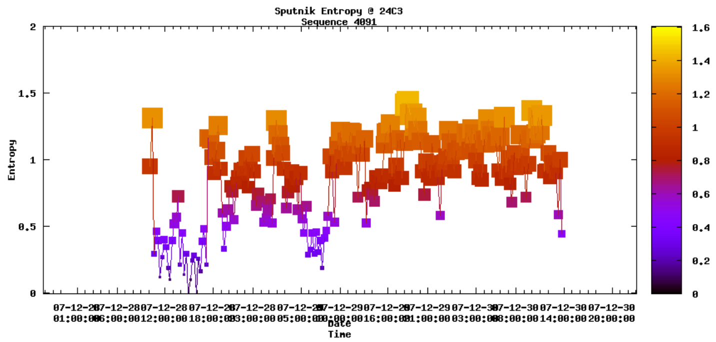

Figure 11: Entropy of Engel responsible for tickets

Figure 12 shows entropy for the person which attended many lectures shown in Figure 6. Entropy is small during periods when he or she was watching lectures. Entropy is much larger during breaks between lectures — more stations can receive signal of tag on the move. One can also notice that entropies in the last day are larger than on previous days; behaviour of this person has changed then, but we do not know why it has happened.

AMD team from New York released data set that contained more information than data from Berlin. There are two data sets containing positions of participants over time (one with packets read by server, like for 24C3, and another is the table “snapshot_summary”) for the New York conference. It allowed for calculating entropy in two ways. Values calculated in those two cases should not be compared with each other but they can be used to look for trends in data. One set contains larger entropy values because this data set contains more classes of values; in both cases entropy changes should give similar overview of behaviour of person.

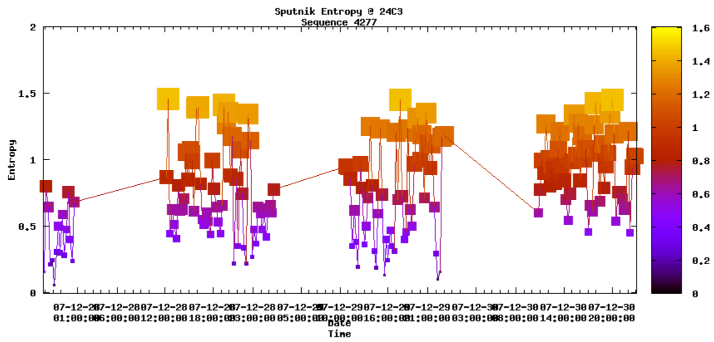

Figure 12: Entropy of person that attended many lectures

Figure 13 shows entropy calculated for tag by using aggregate data. Figure 14 shows entropy for the same tag calculated by using the raw data — the same method that was used to generate Figures 11 and 12.

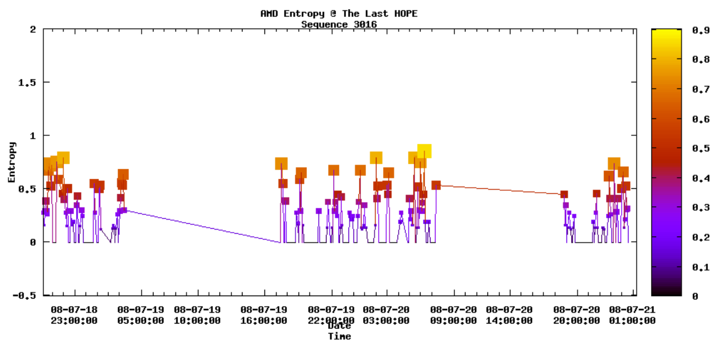

Figure 13: Entropy from The Last HOPE — aggregate data

Although entropy itself was not useful in breaking privacy of tag wearers, it can help with finding interesting candidates for further analysis. It shows whether person was active or not, whether it was sitting in one place or wandering in the building. More detailed analysis can then be used to determine whether person was sitting still during lecture (which could mean that this person was listening carefully) or whether he/she was constantly changing position, suggesting that lecture was not interesting to this person.

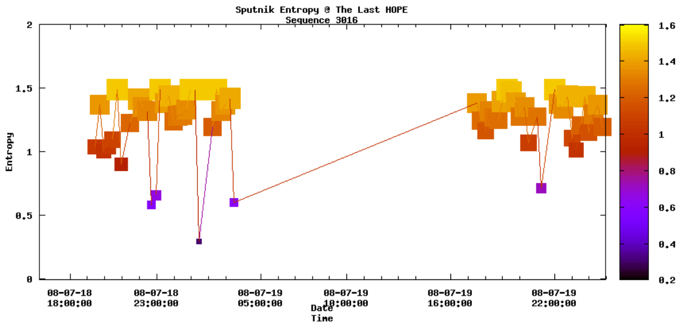

Figure 14: Entropy from The Last HOPE — raw data

6 Summary

Article shows privacy risks arising from using RFID tags. Although analysis was done using data gathered from custom RFID system (called Sputnik), similar techniques can be used to analyse data coming from any device that is worn by user and can be read from distance without his or her consent. The same techniques (entropy, position estimation) can be used to analyse data coming from cellular phone, BlueTooth device, laptop with WiFi transmitter, or even passport with RFID chip. For example if we see that someone is each day from 8.00 till 16.00 in the range of GSM cell tower covering city center, we can assume that this person works there.

Methods described in this article are simple; I am not sociologist, nor have any background in social sciences, so I have not used sophisticated analysis of social behaviour. But even using such simple techniques I was able to find the most active conference participants, determine who was involved in organisation of event, and to find which tag was worn by creator of Sputnik system. More advanced analysis could probably find connections between people or their interests.

We are buying more and more electronic devices, and those devices communicate with one another. Although it makes our lives easier, it also increases risks to our privacy. Living in always-connected world means that our behaviour can be analysed by aggregating data coming from different sources: communication and location from cellular phone, financial operations from internet banking, social ties from FaceBook or MySpace. Custom systems like Sputnik make analysis of privacy risks easier, because they allow for researchers to conduct controlled experiments. Only having large data sets freely available for analysis will allow for researching risks caused by spread of always connected devices.

Even such controlled experiments can lead to privacy violations — as was the case with the AMD data set which was not properly cleaned from all of the personal information. Further research is needed to guide law makers and researchers to properly protect privacy while not preventing development of new technologies.

References

[Cattuto(2008)] Ciro Cattuto. Mining social contacts with active rfid, 2008. URL http://events.ccc.de/congress/2008/Fahrplan/events/2899.en.html.

[Eagle(2005)] Nathan N. Eagle. *Machine Perception and Learning of Complex Social Systems*. PhD thesis, Massachusetts Institute of Technology, 2005.

[Gonzalez et al.(2008)Gonzalez, Hidalgo, and Barabasi] Marta C. Gonzalez, Cesar A. Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *Nature*, 453:779–782, jun 2008. ISSN 0028-0836.

[Kostakos()] Vassilis Kostakos. The privacy implications of bluetooth.

[Meriac(2006a)] Milosz Meriac. Openpcd/openpicc, 2006a. URL http://events.ccc.de/congress/2006/Fahrplan/events/1566.en.html.

[Meriac(2006b)] Milosz Meriac. Project sputnik, 2006b. URL http://events.ccc.de/congress/2006/Fahrplan/events/1736.en.html.

[Meriac(2007a)] Milosz Meriac. Inside sputnik and openbeacon — smart dust for the masses, 2007a. URL http://events.ccc.de/camp/2007/Fahrplan/events/1955.en.html.

[Meriac(2007b)] Milosz Meriac. Practical rfid attacks, 2007b. URL http://events.ccc.de/camp/2007/Fahrplan/events/1957.en.html.

[Rybak(2007)] Tomasz Rybak. Analysis of sputnik data from 23c3. In *24C3 Tagungsband Volldampf voraus!*, pages 19–51, Berlin, Germany, 2007. Chaos Computer Club. ISBN 978-3-934-63606-4.