Guest Editorial

imre@hit.bme.hu
sallai@tmit.bme.hu

This Special Issue is devoted to the 60th anniversary of the estalishment of our two predecessor departments, the Department of Wireline Communications and the Department of Wireless Communications. Since 1949 not only the organization structure and the name of departments in this area of our university has changed several times but we have been witnessing several dramatic changes, even paradigm shifts in communication technologies, telecommunications and broadcasting. It is sufficient to mention that there was no Internet at that time (another round figure this year: we are celebrating the 40th anniversary of the birth of the Internet as a packet switched wide area computer network). To follow these developments, the activities and profiles of the two departments underwent corresponding changes until today when the convergence of these areas has almost fully come true.

The objective of this special issue is to highlight a few topics by invited overview papers that are representative samples from the wide scale of research activities of the two present departments: the Dept. of Telecommunications (HIT) and the Dept. of Telecommunications and Media Informatics (TMIT).

The first paper titled „On the security of communication network: now and tomorrow" by *Boldizsár Bencsáth, Levente Buttyán and István Vajda* of Crysys Lab *(Laboratory of Cryptography and Systems Security, HIT)* discusses some security issues in the Internet and sketches future research directions in this field. In particular, the authors discuss the security issues in wireless networked embedded systems through three examples: sensor networks, vehicular communications, and RFID systems. Finally a brief introduction is given to the field of network coding, which is a new and promising research area in networking.

Speech technology has been an area of intensive research worldwide – including Hungary – for several decades and the *Laboratory of Speech Technology of TMIT* has been in the forefront of this research. The paper by *Géza Németh, Gábor Olaszy, Klára Vicsi and Tibor Fegyó* „Talking machines?! Present and future of speech technology in Hungary" gives an overview of the challenges and results of the domain and the vision of the development and the application of the technology will also be introduced.

The next paper „Congestion control and network management in Future Internet" by *Márton Csernai, András Gulyás, Zalán Heszberger, Sándor Molnár and Balázs Sonkoly* addresses two key issues in the Future Internet research where the general objective is to encourage clean slate designs and thus overcome the barriers of the previously prevailing incremental development, proposing new visions, architectures and paradigms for the coming 10-20 years. The first area is congestion control where recent results show that networks operating without explicit congestion control (like TCP) may survive without congestion collapse if appropriately designed in network resources and if end systems apply appropriate erasure coding schemes. The second topic is related to the exponential growth of the Internet which makes it hardly impossible to manage the network with traditional centralized approaches (like manager-agent); hence research results of complex networks are expected to spread over the Internet with its autonomic behaviors. The authors are pursuing research in these fields within the framework of the *High Speed Networks Laboratory of TMIT.*

The last paper is titled "Media communications over IP networks – An error correction scheme for IPTV environment" by *László Lois, Ákos Sebestyén, Laboratory of Multimedia Networks of HIT*. Video transmission over IP networks has been gaining more and more popularity recently. One of the crucial problems of video transmission over IP networks through unreliable links is the susceptibility to errors in the transmission path. Packets lost or discarded must be somehow regenerated which can be accomplished by requesting a retransmission or by recalculating the packet provided that some redundanc y is introduced in the transmitter side. In addition to a general overview of issues around the media transmission over IP networks, the paper describes a novel method that can be used for forward error correction in IPTV applications.

Prof. Dr. Imre Sándor
Head, Dept. of Telecommunications

Prof. Dr. Sallai Gyula
Head, Dept. of Telecommunications and Media Informatics

60 years of
Department of Telecommunications and Department of Telecommunications and Media Informatics

szabo@hit.bme.hu

A short summary based on the article by Profs. Géza Gordos and László Pap, published in Hungarian in the Special Issue of "Híradástechnika" devoted to the anniversary of the two departments.

In the 40's, Hungary was considered as one of the most developed countries in the world in the field of telecommunications and broadcasting. The telephone line density was one of the highest in Europe by the beginning of WW2. Radio broadcasting started already in 1923, based on the program structure and studio of the so-called telephone news service, invented by Tivadar Puskás in 1893. Our radio broadcasting industry produced one-third of the radio receivers in the middle of the 30's!

To meet the increasing need in electrical engineers specialized in these new areas, 60 years ago the Budapest University of Technology and Economics (then Technical University) decided to establish its Faculty (~School) of Electrical Engineering. It had a new branch called "weak current electrical engineering" in addition to the already existing – within the Faculty of Mechanical Engineering – specialization in "heavy current electrical engineering" (~electric power engineering).

This new branch of education started in the academic year of 1949/1950 and two new departments were founded the same time, the Dept. of Wireline Communications and the Dept. of Wireless Communications. In 1971, these two departments were combined in a single unit called Institute of Communication Electronics which existed for 20 years when in 1991 – as a result of a natural polarization – it was split again into two units: the Dept. of Telecommunications and the Dept. of Telecommunications and Telematics. The latter was re-named to Dept. of Telecommunications and Media Informatics in 2003. These days the two departments cooperate closely and their scopes mostly complement each other in several areas, and some healthy competition also exists between them.

The Dept. of Telecommunications was established and lead until 2008 by Prof. László Pap, Member of the Hungarian Academy of Sciences and is currently lead by Prof. Sándor Imre. It consists of laboratories which address the key research areas of the department and are responsible for the educational activities in these areas:

- Data and network security
- Signal processing and networking algorithms
- Computing technologies
- Multimedia networks
- Acoustics
- Networking technologies and network modeling
- Network design
- Mobile communications and computing

The Dept. of Telecommunications and Telematics was established and lead until 2002 by Prof. Géza Gordos. Since then Prof. Gyula Sallai has been the Head of Department. The department focusing on the convergent information, communication and media technologies is organized in the following laboratories representing the key research and educational areas:

- Infocommunication Networks, Services and Applications
- Infocommunication Management, Strategy and Regulation
- Content Management and Multimedia Systems
- Speech and Multimodal Information Systems
- Speech Acoustics
- Intelligent and Cognitive Media Informatics

In addition, there are two educational laboratories and an accredited and notified Telecommunications Test Laboratory also operates within the department.

For detailed and up-to-date information about the two departments we suggest to visit their websites:
www.hit.bme.hu and *www.tmit.bme.hu*

On the security of communication network: now and tomorrow

BOLDIZSÁR BENCSÁTH, LEVENTE BUTTYÁN, ISTVÁN VAJDA

Budapest University of Technology and Economic, Department of Telecommunications
Laboratory of Cryptography and System Security (CrySyS)

{bencsath, buttyan, vajda}@crysys.hu

Keywords: information security, privacy, Internet, wireless sensor networks, vehicular communications, RFID systems, network coding

In this paper, we first discuss some security issues in the Internet, and we sketch some future research directions in this field. Then, we discuss the security issues in wireless networked embedded systems through three examples: sensor networks, vehicular communications, and RFID systems. Finally, we give a brief introduction to the field of network coding, which is a new, promising research area in networking.

1. Introduction

In this paper, we give a brief overview of Internet security issues. First, we discuss the security problems of the current Internet as we know and use it today. Then, we introduce some future research directions in the field of Internet security. We continue by considering a broader interpretation of the Internet than it is usually meant today, where the network is not limited to PCs and servers, but it also includes various embedded computers. This broader interpretation is often referred to as the *Internet of Things*.

We describe the related security and privacy issues through three examples: wireless sensor networks, vehicle communication, and RFID systems. Finally, in the last part of the paper, we give a short overview on the security of network coding, which is a new promising research area that may have impact on the design of future networks.

Obviously, the general field of security and privacy in communication systems is a large area that cannot be fully covered in the context of this paper. Our selection of the topics discussed in this paper have been biased by our research activities in the Laboratory of Cryptography and Systems Security (CrySyS) of the Department of Telecommunications at the Budapest University of Technology and Economics. More information on our research and other activites is available on the web site of the laboratory at *www.crysys.hu*.

2. Security of the Internet

The security of the internet has evolved a lot in the last decades. Today, nobody can imagine the Internet without security mechanisms such as TLS, SSH, PGP, IPsec, or without advanced authentication techniques (smart cards, captchas, two factor authentication tools, etc.) However, lot of the security problems of the Internet remained unsolved.

Malware, badware, viruses and *worms* have been known for decades, however, the problem is getting worse and worse instead of finding proper solution for the problem. Malware infected hosts are not individual problematic points anymore, *botnets* have been created. These botnets might aggregate the resources of millions of computers. Even the estimation of the size of the botnets is a hard problem.

Using botnets, millions of *spam* e-mail messages can be sent out easily and rapidly. Although solutions are continuously proposed against spam, spam is still one of the biggest problems of the Internet. In June 2009, 88.9% of the full Internet e-mail traffic was spam.

Most of the Internet servers and services are prone to some kind of *Denial-of-Service problems (DoS)*. The basic architecture of the Internet was not designed to be resilient against such attacks, and therefore the number vulnerable services, servers, protocols is unknown and might very high, therefore, a significant increase might happen in both the number and the severity of the DoS attacks at any time.

Another unsolved problem on the Internet is *cracking and defacing web pages and servers*. Today, a number of tools and possibilities are available to make servers secure: automatic updates, intrusion detection systems, secure authentication, vulnerability scanners, etc., but still, the web sites are not secure and cracked frequently. There is multiple reasons behind that. Although tools are available, they are not used, for different reasons to make systems secure. Special web content cannot be updated automatically at all the time. Administrator do not have right to fulfill necessary operation without the owners' consent.

Grid-computing, cloud-computing also raise new security and reliability concerns. In such cases the legal situation is even more complicated. The cloud-based service might be free or very cheap, but there is no guarantee on the reliability, availability and security properties, therefore the end users might have no options to solve security problems.

2.1 Research directions

Not just the problems and solutions have been significantly changed during the last years, but even the methodology, the point of view on the problems changed a lot.

Future internet: A number of new research projects started to design the next generation of the Internet, including Global Environment for Network Innovations (GENI), a Future Internet Design (FIND), funded by the U.S. National Science Foundation (NFS), or projects in the Seventh Framework Programme of the European Union (FP7). The main goal of these projects is to redefine the basic architecture of the Internet. Traditionally the Internet provides best-effort services, it is a multi-layer unreliable network. Most of the services are based on TCP/IP where the main task of the routers just to forward simple packets. The reason behind many security problems is this approach: because of the architecture and protocols of the Internet, it is simply impossible or almost impossible to solve some security problems. By redesigning the basic blocks and the architecture, new solutions can be made for the security problems.

The next generation Internet should give more than best-effort forwarding of packets. People need services instead of an unreliable transport network, the need quality of service (QoS) agreements. Today, Internet-wide secure authentication is not solved, especially for simple network protocols, and therefore, it is nearly impossible to track back the origin of an attack and punish attackers. This list of problems and new tools, features might be continued, but the most important message is that there is a need and also intention to modify even the foundations of the Internet.

Intelligent intrusion-detection: The intrusion detection (IDS, IPS, honeypot, etc.) tools have evolved a lot in the last years. From the basic event detection we arrived at complex systems which intelligently distinguish attacks from normal traffic and automatically carry out countermeasures. This change goes forward to make global intrusion detection systems (possibly with honeypot systems), to use modern network technologies, like P2P techniques in intrusion detection, or more intelligent reactions to security problems. To this last goal, to provide more dependable systems by automatic reconfiguration, an EU FP7 project, DESEREC *(www.deserec.eu)* has just been finished with the participation of BME's Laboratory of Cryptography and Systems Security (CrySyS).

In the area of trust, reputation and authentication, new methods give new tools to solve security problems and therefore there is a very intense research activity in this field. The work includes:

- Defining attack and defense incentives and providing new tools based on these properties.
- Using game theoretical methods to make such situations, where there is no use to attack, thus avoiding attacks.
- Providing new ways of authentication possibilities while retaining anonymity and civil rights.
- Dealing with trust on local level and on large scale.

Secure clients and secure, trusted platforms: Lot of the problems originate from the fact that the software elements and thus the whole client computer cannot be trusted. Trusted computing could provide a situation, where there are no malware programs, or, the can be easily removed at large scale. However, the concept of trusted platforms contradicts with the current philosophy of the Internet, e.g. the need of the users to install pirated software, downloading music illegally, etc. The typical research areas: secure software engineering; formal analysis and proving of protocols, rule sets, or even formally proven APIs; secure authenticaion in untrusted environment, etc.

This short introduction cannot provide a full picture of all the work that is currently in place to provide new solutions to security problem, we just tried to show the most interesting research areas in this field, that could have a great impact to the future of the Internet.

3. Security and privacy in wireless networked embedded systems

3.1 Security in wireless sensor networks

In the last decade, a considerable amount of research on wireless sensor networks has been carried out all over the world. This new wireless networking technology allows for a number of new and useful applications the monitoring of the parameters of our physical environment (such as temperature, pressure, humidity, vibration, acustic noise, etc.), and the automated collection and processing of all these monitored data. The potential applications include optimization of agricultural processes, making ecological observations on large scale or in environments that are difficult to access physically, forecasting natural disasters such as earthquakes, reducing cost in industrial process automation and control, prevention of accidents on the roads, remote monitoring of elderly or chronically ill people, and military tactical applications, just to mention a few.

Many of these applications have security requirements related to the protection of wireless communications on the one hand, and to the increased resistance of the networking mechanisms againts malicious attacks on the other hand. Although, security is a problem and has been addressed both in wired and in traditional wireless networks (e.g., in cellular and Wi-Fi networks), there are new security challanges in wireless sensor networks, and the solutions proposed for wired and traditional wireless networks can be used only with limited success, if at all. Such a challenge, for instance, is that the nodes in wireless sensor networks have severe resource constraints: the nodes are small, battery powered embedded computers designed for low energy consumption, and consequently, they have reduced computing, storage, and communication capabilities. Therefore, one needs new security mechanisms in wireless sensor networks that are computationally not so expensive, have small code size, and minimize energy consumption.

These requirements are usually not satisfied by the security mechanisms used in traditional networks.

Another distinct security problem in sensor networks is that the nodes are usually physically accessible and they are not tamper resistant. Thus, they can be relatively easily compromised, meaning that an attacker can obtain secrets stored in the node and install rogue software on the node such that it continues to behave arbitrarily. Node compromise may happen in traditional networks too, however, as the nodes of traditional networks are usually located in locked rooms, the attacker is essentially restricted to remote logical attacks. In wireless sensor networks, node compromise is easier to carry out due to the easy physical access to the nodes, and we must always assume that some nodes may have indeed been compromised.

In our CrySyS laboratory, we have been working on the problem of securing wireless sensor networks in the context of two projects, UbiSec&Sens *(www.ist-ubisec-sens.org)* and WSAN4CIP *(www.wsan4cip.eu)*, both funded by the European Commission. In the former project, we participated in the development of a security toolbox for sensor networks inlcuding a random number generator, new encryption algorithms, new key establishment protocols, and security enhancements for routing, clustering, data aggregation, and distributed data storage schemes.

In particular, we developed the Secure-TinyLUNAR secure routing protocol, the RANBAR and CORA resilient data aggregation algorithms, and the PANEL robust aggregator node election protocol. In the latter project, we are currently working on dependable networking mechanisms for wireless sensors in the context of critical infrastructure protection applications.

3.2 Security and privacy in vehicle communication systems

Unfortunately, more than 40.000 people die in road accidents in Europe, and the statistics in the US are similar. In addition, another problem is the ever increasing amount of road traffic in large cities that leads to traffic jams and vaste of tremendous amount of time and fuel. In both cases, the situation could be improved if the right information would be available at the right place at the right time (i.e., if drivers would be informed about hazardous situations and receive up-to-date information about the traffic conditions). This could be achieved by letting vehicles to communicate with each other and with roadside equipments. Such communications must obviosly be wireless due to the nature of the application.

Most of the large car manufacturers around the world are seriously considering the idea of vehicle-to-vehicle and vehicle-to-infrastructure communications, and they investigate the related technical problems in national and international projects (e.g., NoW, CVIS, Safespot and Coopers projects). Among those technical problems, they are also looking at security issues. Indeed, it must be clear that vehicle communications can only be adopted if it cannot be easily misused, or its operation cannot be easliy disabled. One thing to absolutely avoid is that someone sitting at the roadside injects fabricated messages in the system, and in this way, provokes accidents. A security hole in the system can easily translate into fatalities in this case. Therefore, it is indispensable that messages are aunthenticated and their contents are validated, for instance, through consistency checking and correlation with other similar messages.

Most safety applications require that the vehicles continuously inform nearby vehicles about their current location, direction of movement, and speed. For this reason, the vehicles send so called heart beat messages, with a rate of several hundred messages per second, which contain these data. This, however, raises privacy problems, as it becomes rather easy to track the whereabouts of the vehicles by sniffing the wireless channel. Note that, although tracking vehicles is possible by means of video surveillance technologies, tracking by eavesdropping is less expensive, more precise, and can be carried out at larger scale. We, and fortunately many others, believe that, in modern societies, new technologies should be introduced only if they are designed in such a way that they do not make privacy violations easier than they are today. Therefore, the protection of location privacy in vehicle communication systems is an important design requirement.

In our CrySyS laboratory, we have been working on the problem of securing vehicle communications and location privacy enhancing techniques in the context of the SeVeCom project *(www.sevecom.org)* that received funding from the European Commission. In particular, we have investigated various pseudonym schemes and their effectiveness in preventing location tracking in vehicle networks.

3.3 RFID privacy

Similarly to the problem of location tracking of vehicles, individuals can be tracked by sniffing the wireless transmissions generated by various devices that they carry with or on them. In particular, it is expected that in the future many objects will be tagged with RFID tags that emit unique identifiers each time they are queried by a nearby reader. Hence individuals could be identified and tracked by observing the identifiers of their RFID tags. Unfortunately, RFID tags are even more constrained in terms of resources than the previously described sensor nodes; hence, it is very challenging to design protocols for them that would prevent this kind of tracking.

In our CrySyS laboratory, we have been working on efficient private identification schemes for low-cost RFID tags that ensure that external eavesdroppers cannot obtain the identifier of the tag from the transactions between the tag and a valid readers.

4. Network Coding

The idea of network coding was born at the beginning of our new millenium, an important development in the theory and practice of infocommunication. 60 years ago

Shannon in his famous publication gave the information theoretical foundation for point-to-point communication (channel capacity, existence of optimal channel codes). From the beginning of the 1960's till the end of 1980's capacity calculations were extended to small/special networks and channels (e.g. broadcast channels, multiple-access channels, feedback channels).

For practical applications many important results were found in the field of random access channels and code division multiple access channels (CDMA). Then more than a decade spent without new ideas in the field of information theory of networks when, finally, in the year of 2000 the network coding was innovated [7].

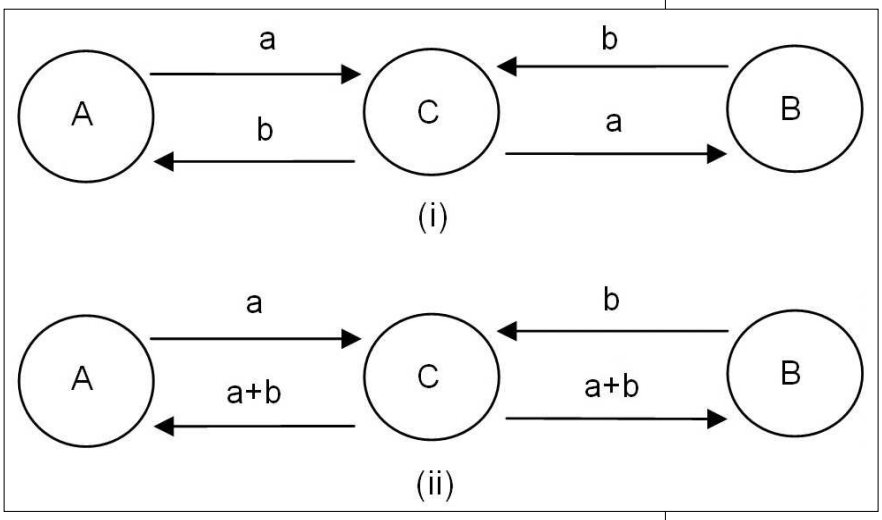


Figure 1. An illustration of the idea of network coding

One the simplest example for the strengh of network coding is given in *Fig.1*. Wireless communication nodes *A* and *B* want to exhange their messages *a* and *b*, respectively Obviously, they can solve this task in four steps of communication, as shown by version (i). However, they can also do it by sending only three messages in total: node *A* and *B* send their messages to node *C*, which linearly combines the messages by XOR addition *(a+b)* and broadcasts the sum in one step. It easy to see that network coding, i.e. allowing linear combination of messages by communication nodes, improves the communication efficiency compared to the classical store-and-forward solution: the store-and-forward solution is the trivial network coding, when no combination is done at nodes.

The principle of network coding can be extended to non-binary cases by straighforward generalization. In retrospect, the network coding in the case of the scenario of Fig.1 is trivial. But, how can we find appropriate combinations for large, general networks? Fortunately, it turns out that nodes can independently choose random linear combinations to achieve, essentially, optimal coding.

If a source (or a set of sources) wants to transmit a time flow of messages to destination nodes through the network, then the flow is partitioned into units of fixed number of messages (called generations) and network coding combines messages within the same generation. A destination node is able to decode messages of a generation, when it has collected a set of linearly independent combinations (so called innovative combinations) with set size equals to the size of a generation.

Network coding has important advantages for a diversity of practical application: for example, improvement of network throughput, improved robustness of communication in case of link/node failure, decrease of the number of communication steps in case of energy critical application (battery powered nodes).

Now, we shortly summarize the possible positioning of network coding in the protocol stack. Network coding can be placed in each layer, resulting in different potential applications. If we implement it in the application layer, it has the advantage that routing and MAC layers remain intact, and extension is needed only in the software of the source and destination nodes. A typical application is the case of overlay communication topologies, e.g. P2P file distribution systems.

When the network coding is implemented in the transport layer, a destination node does not send back an ACK for a received packet, but the node sends back information about possible combinations which were innovative for it. A source node according to received needs for innovative information selects combination which is innovative for largest possible set of destination nodes.

In the network layer during the discovery of network topology – which typically some kind of flooding technique – application of network coding is very appropriate. The so called opportunistic network coding in wireless networks is an excellent example for the usage of network coding in the data link layer. Network nodes are set into promiscuous mode and overhear the wireless communication in their neighborhood, according to which they can optimize the next network coding step they do. Network coding can also be implemented even in the physical layer, which is called analogic network coding.

A serious disadvantage of the network coding is its sensitivity to the so called pollution attack. This means that network coding does not check the integrity of the received packets, and if a – illegally – modified packet (polluted packet) is used in a combination it also becomes modified and will be combined into several further combinations accross the network, resulting in spreading of the pollution, which finally deteriorates the network communication. This attack can be stopped by detecting and dropping polluted packets.

For detection cryptographic techniques (MAC, digital signature or hash techniques in case of available authentic channels) can be used. This cryptography must fit

to network coding by having special algebraic property (homomorphic mapping). However, in case of resource constrained environments (e.g. sensor networks) cryptographic techniques with high computational complexity are excluded.

We can derive the conclusion that the theory and practice of network coding is developing fast in recent years. There are many important, potential applications, however, it is an open question yet, when will network coding be an ubiquitous option for networking protocols.

5. Conclusions

Communication systems play a fundamental role in today's society, and therefore, their dependability is crucial. Dependability includes also security, which means resistance against intentional attacks. In this paper, we reviewed some of the security issues pertaining in the current Internet, and those expected in future emerging wireless networks and communication systems.

As we could see, there are important challanges ahead of us that must be properly addressed by researchers and designers of future communication systems, such that we can live in a safer cyber world than we do today.

Authors

BOLDIZSÁR BENCSÁTH received his MSc diploma in Computer Science in 2000 from the Budapest University of Technology and Economics (BME), and his MSc diploma in Economics in 2001 the Corvinus University of Budapest. He has been working in BME's Laboratory of Cryptography and System Security since 2000, first as a PhD student and then as a researcher. His research interests are in practical Internet security, protection against spam and distributed denial-of-service attacks.

LEVENTE BUTTYÁN received the M.Sc. degree in Computer Science from the Budapest University of Technology and Economics (BME) in 1995, and earned the Ph.D. degree from the Swiss Federal Institute of Technology – Lausanne (EPFL) in 2002. In 2003, he joined the Department of Telecommunications at BME, where he currently holds a position as an Associate Professor and works in the Laboratory of Cryptography and Systems Security (CrySyS). His research interests are in the design and analysis of security protocols and privacy enhancing mechanisms for wireless networked embedded systems (including wireless sensor networks, mesh networks, vehicular communications, and RFID systems), and the application of formal methods in security engineering.

ISTVÁN VAJDA obtained his MSc degree in Electrical Engineering in 1977, and a diploma in Telecommunications Engineering in 1979 at the Budapest University of Technology and Economics (BME). He received his CSc degree (equivalent to PhD) in 1984, and his Doctor of Sciences (DSc) degree in 1997. Currently, he is a full professor at the Budapest University of Technology and Economics, and he is the head of the Laboratory of Cryptography and System Security at the Department of Telecommunications. His research interests include coding theory and cryptography.

References

[1] MessageLabs Intelligence Reports, Symantec, July 2009. http://www.messagelabs.com/mlireport/MLIReport_2009.07_July_FINAL.pdf

[2] Rajab, M.A., Zarfoss, J., Monrose, F., Terzis, A., "My botnet is bigger than yours (maybe, better than yours): Why size estimates remain challenging," Proceedings of 1st Workshop on Hot Topics in Understanding Botnets (HotBots'07), 2007.

[3] Global Environment for Network Innovations (GENI); http://www.geni.net/

[4] NFS NeTS FIND Initiative, http://www.nets-find.net/

[5] Seventh Framework Program, http://cordis-europa.eu/fp7/ http://www.future-internet.eu/activities/fp7-projects.html

[6] The Honeynet project, http://www.honeynet.org/

[7] Ahlswede, R. et al, "Network information flow." IEEE Transactions on Information Theory, July 2000, Vol. 46, No. 4, pp.1204–1216.

Talking machines?! – Present and future of speech technology in Hungary

GÉZA NÉMETH, GÁBOR OLASZY, KLÁRA VICSI, TIBOR FEGYÓ

Budapest University of Technology and Economics,
Department of Telecommunications and Media Informatics

{nemeth, olaszy, vicsi, fegyo}@tmit.bme.hu

Keywords: speech technology, speech synthesis, speech recognition, dialogue systems

Speech technology has been an area of intensive research worldwide – including Hungary – for several decades. This paper will give a short overview of the challenges and results of the domain and the vision of the development and the application of the technology will also be introduced.

1. Challenges of speech technology

Speech has been the most natural and most frequently used means of human communication. Speech usually fulfills the information transmission role between biological systems *(Fig. 1)*.

The science of speech technology has emerged a few decades ago. Its' results are used in replacing certain elements of the natural speech communication chain by artificial solutions (speech recognition, speech synthesis, human-machine dialogue, diagnosis by speech, speech training, speech-to-speech translation, etc.).

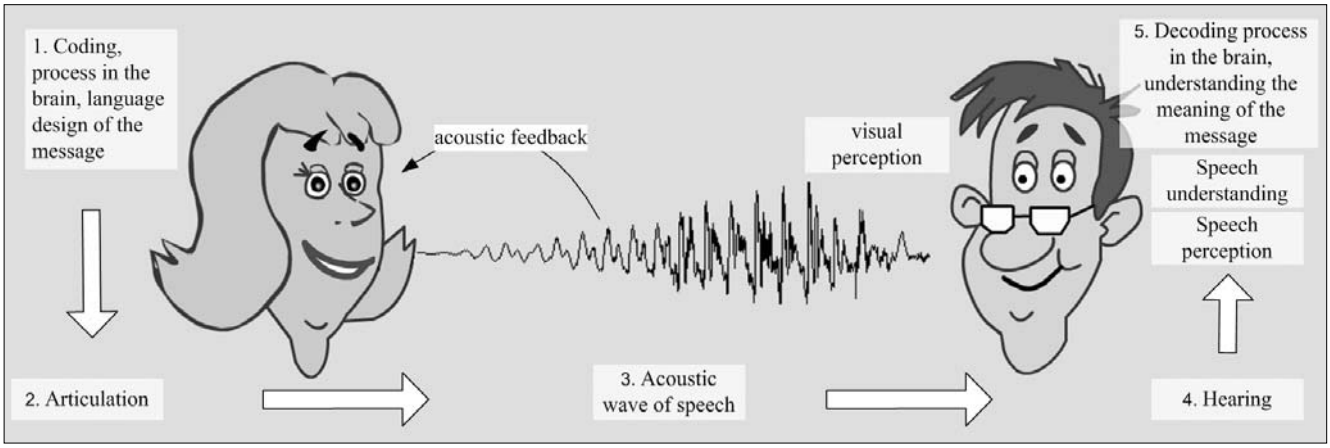
Out of the elements of the natural speech communication chain most practical engineering applications rely on the acoustic signal so in the following we shall also concentrate on this aspect. It should be noted, however, that the language is always behind the acoustic form of speech. The linguistic information determines several acoustic components of the spoken message. In order to create successful solutions of language and speech technology it is not only the processing of the acoustic signal that should be solved. Deep linguistic knowledge should also be coupled with it in order to achieve an artificial system comparable to natural communication.

The movie industry has given good visions for the practical applications of speech technology. One of the key "actors" is the HAL 9000 speaking computer in 2001: A Space Odyssey that was presented first in 1968 [26]. In 1977 in the first episode of Star Wars [1] robots perceive, store and present in many ways the multitude of information collected and transmitted through speech communication.

These visions of art created the impression for several people that all these technological breakthroughs can be reached in a short time. In practice just as interstellar spaceships, speaking and thinking robots are still to come. Because of the gap between huge expectations and significant but relatively slower technological advancements plus short time market success requirements there is a certain cyclic nature in the development of speech technology.

It is illustrated in *Fig. 2* along the dimensions of (technological) maturity and (media) visibility. The figure was created by combining Gartner's 2002 and 2006 key ICT (Information and Communication Technology) and HCI (Human Computer Interaction) forecasts in speech technology related areas. For example *natural language search* was expected to be mature for the market in 2-5 years by analysts in 2002 (i.e. between 2004-2009). The forecast

Figure 1. The process of speech communication



range changed to 5-10 years in 2006 (i.e. 2011-16). In the evaluation of *speech recognition on the desktop* similar trends can be observed. It was only *speech recognition in call centers* that moved from the 2-5 years category in 2002 into the less than two years expectation by 2006. *Text-To-Speech* (TTS) played the role of emerging technology both in 2002 and 2006 (less than two years until market penetration). It is important to note that these forecasts were created for the most developed, English speaking USA market where automation is a frequent business target (e.g. in telephone-based call centers). The real market situation varies greatly around the world, and evaluation is a continuous challenge both in Europe as a whole and in our homeland (Hungary) in particular.

In the next section of the paper an overview will be given about the results of speech technology in an international and a domestic setting with particular emphasis on existing and possible applications. In the 3rd section a short introduction will be given to the research and application vision of the area.

2. Domestic and international results of speech technology

It is worth looking at both the starting point and the current situation of R&D in domestic and international settings. It should be noted again that successful speech technology developments require advances in at least two areas: linguistic analysis and acoustic signal processing.

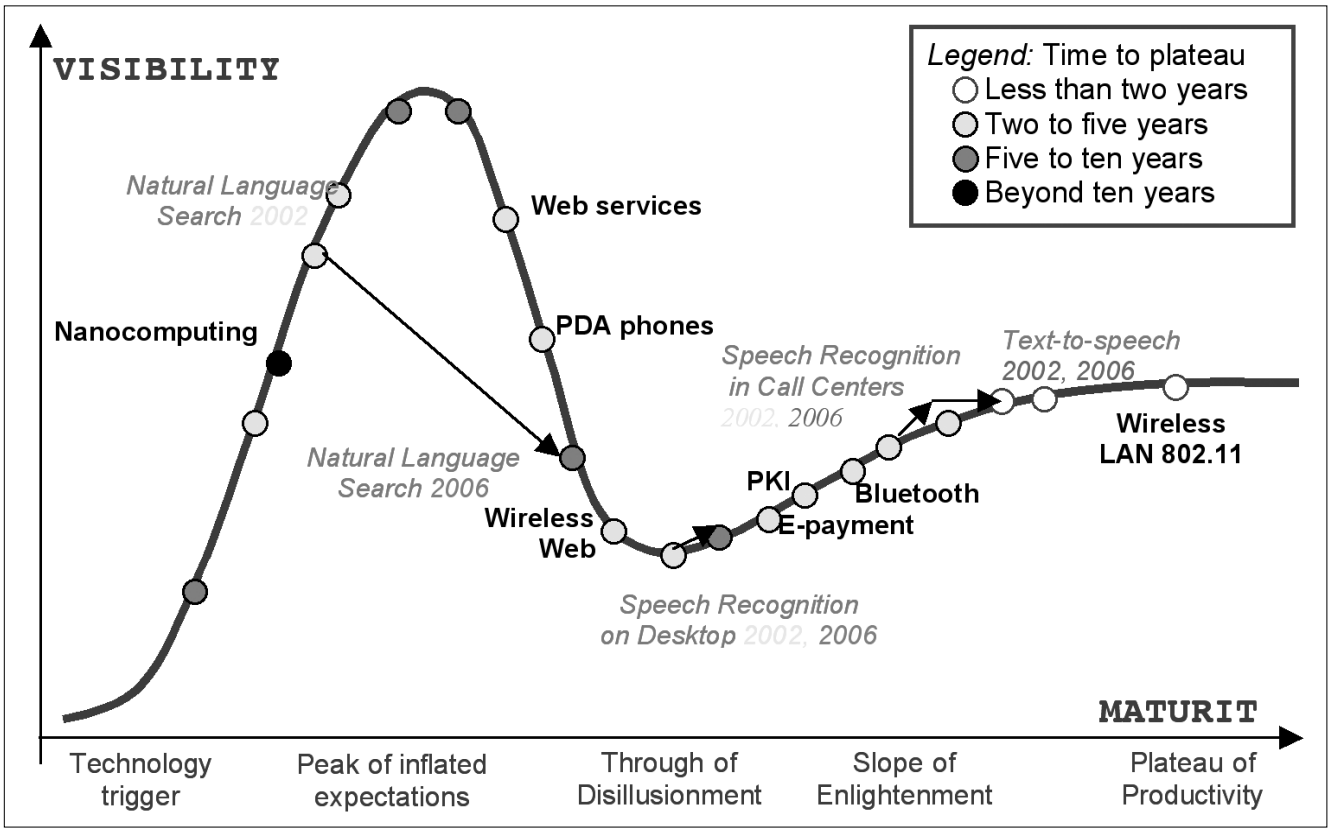
Automatic speech generation

In the area of automatic speech generation (often called speech synthesis) developers achieved as early as in 1984 that an English TTS system became part of the operating system of Apple computers [5]. This step was taken in the English versions of Microsoft Windows systems after 2000.

Hungarian research was already in the forefront of international research at the beginning of the 80s when the first general purpose, Hungarian TTS system – Hungarovox – was born in the Institute of Linguistics of the Hungarian Academy of Sciences [4]. Since then both linguistic analysis and acoustic signal processing algorithms have substantially improved. In the latter area the fourth generation is under study. The first systems modeled the human articulatory process by a time-varying filter bank and a simple excitation signal. This so-called formant synthesis solution allowed a coded acoustic database as small as a few kilobytes. The Hungarovox system was based on this technology, too. The system had a strongly robotic voice, with slow speech without rhythm and accent but with some level of intelligibility. At the predecessor of the Department of Telecommunications and Media Informatics (BME TMIT) the similar but improved MultiVox system was implemented in 12 languages [12]. The German version of MultiVox was licensed by an Austrian and a German company.

In co-operation with the developers of the Recognita optical character recognition (OCR) system we could demonstrate a Hungarian book-reading system in 1987. The first, commercially availably speech synthesizer for

Figure 2. Extended version of the Gartner Hype Cycle [Gartner Hype Cycle 2002, 2006]



the Commodore 64 computer was also developed at the same BME department [14 – p.269].

In the solutions of the second generation (from the beginning of the 90s) waveform segments were cut out from human voice, containing parts of two or three sounds (diphones and triphones, respectively). The acoustic database was compiled from these elements. The synthesized waveform was concatenated from these units (c.f. concatenative synthesis). In the following step digital signal processing algorithms were applied based on a prosodic model (pitch, timing and intensity). With this solution a speech quality resembling the given speaker could be achieved with a database on 1 to 10.000 units, and with storage space in the order of megabytes. In our research the ProfiVox system belongs to this category [13].

It was the basis in 1999 for the so-called MailMondó (MailReader) service, which read e-mail messages over the telephone for subscribers [6]. The same technology is applied in the SMS reading system for wireline telephony subscribers and in the SMSMondó (SMSReader) application for Symbian smartphones [10]. ProfiVox has also been integrated in the most widely used screen reader program for visually impaired people in Hungary (Hungarian version of Jaws).

The development of the third (corpus-based) technology started in the middle of the 90s. In this case there is no (or maybe some minor) prosodic modification by signal processing. Several hours of (usually read) speech from a speaker (or so-called voice actor/actress) are stored. This database is the acoustic database for synthesis. In case of good design there is a high probability that all sound units are available in several prosodic forms. The storage requirements of these solutions fall in the gigabyte range.

This technology is applied in the Hungarian name and address reading solution of BME TMIT that has already allowed the automation of the reverse directory service (reading out the name and the address of the subscriber based on the input phone number) of two mobile operators. In limited domains this technology can approach human quality. BME TMIT has prepared solutions for several domains. The weather forecast reader is publicly available *(www.metnet.hu)*, the automatic generation of auditory version of the price list of devices and services is applied in the IVR system of a mobile company [11]. The latest demonstration system is a railway timetable information system that "speaks" at the railway station of Sárospatak.

The fourth generation of TTS is based on Hidden Markov Models (HMMs). The basic principles are quite different from earlier TTS generations. One may say that it grew out from speech recognition experience. The acoustic basis in this case is recordings of several hours from one or more speakers (storage requirements may be in the terabyte range). These databases are used for training by statistical methods the control parameters of parametric speech coders. It is important that although a large database is required for training the resulting parameter database is typically much smaller (even just a few megabytes) which opens up several interesting applications. The quality of the latest HMM systems approach that of the third generation corpus-based systems. There are experiments with so-called hybrid systems which provide the data-driven, flexible features of HMM while maintaining the high speech quality of corpus-based systems. Our researchers conduct promising experiments in the HMM field, too [16].

It should be remembered, however, that although there are always new solutions the viability of older ones does not necessarily cease. They all have certain advantages that may be critical in certain applications. For example it is quite easy to generate whispering voice or speed up/slow down the synthetic speech with formant (or other parametric) technology which is quite difficult for corpus-based or HMM approaches.

Automatic speech recognition

The ASR has been a field of intensive research worldwide since the middle of the 20th century. From the initial sample-based systems with a vocabulary of a few words [15] the technology has advanced to large vocabulary, continuous, speaker independent technologies. The first Apple operating system containing speech recognition was announced in 1993 [5]. The latest ASR systems of industrial applications are typically based on HMMs. The basis of the technology was laid down in the 70s by the researchers of IBM. Nearly forty years have passed since then but we still cannot meet "omniscient machines" that perfectly comprehend our speech. In several narrower domains (e.g. medical dictation) though, there have been applications of regular practical use.

Current ASR systems – beyond standard software elements – have basically two language and application dependent components. Both the acoustic and the language model have to be trained according to the given application environment.

The acoustic model usually represents the speech sounds as derived from sound samples taken from several speakers. Even relatively small research databases contain at least 10 hours of speech of at least 100 speakers but there are training databases of up to several thousand hours. These samples have to be recorded in an environment that is identical (or at least similar) to the end-user application. For example there are different acoustic models for office (wideband) and telephony (narrowband) situations. The general acoustic model can be adapted to the voice of a particular speaker from a relatively smaller set of training data. The output of pattern matching based on just the acoustic models is not accurate enough. That's why the language model of a higher level is required. It is not so surprising if we remember that human speech perception and understanding have several layers, too.

Language models help the recognizer in matching the output of the acoustic model (sound sequence) to the probable linguistic content. In fact, the individual

speech sounds are connected to a complex network according to the given application environment. In a simple case, for example command word recognition, the language model is just a simple vocabulary where the possible commands are listed. In case of the more complex continuous ASR task the linguistic probability of words following each other also have to be taken into account. In practice statistical language models trained on large text corpora are applied.

In case of agglutinative languages – such as Hungarian – because of the large number of possible word forms morpheme-based approaches have also gained momentum besides traditional word-based ones. Language models are always domain specific, the narrower the domain the higher recognition accuracy can be expected. In case of isolated command words over 95% accuracy is not rare while in case of the recognition of spontaneous conversations a result over 60% is regarded as quite good on an international scale.

Researchers of BME TMIT have participated in co-operation with industrial partners and with significant state funding in the creation of several practical applications and in the composition of related indispensable databases. Their detailed introduction is beyond the scope of the paper so only a list of major results is given below.

- *MKBF 1.0 –*
  ASR engine and development environment:
  The ASR engine is HMM-based and provides real-time processing in case of moderate size vocabularies (1000-20.000 words). The toolset supports the training of both acoustic and language models and allows N-gram models and speaker adaptation as well.
- *Medical report generator:*
  The system allows the direct speech to medical diagnosis transcription [24].
- *Classification of prosody and segmentation of speech flow:*
  Prosodic-acoustic processing speech has turned from the interest of speech synthesis research to ASR focus as well. An application – based on accent and intonation contour based classification – was developed for word and phrase boundary detection for Hungarian and Finnish. A clause segmentation and a modality detection module was also implemented [21].
- *Automatic speech-based emotion recognition* [2,17].
- *Speech databases* [22]*:*
  A large amount of labeled and annotated sound material is required for the training of ASR systems. During the preparation of databases statistical language analysis, linguistic and phonetic modeling, corpus design, database qualification and validation tasks have been completed. Diagnostic (oto-rhino-laryngology, radiology, etc.), news (Broadcast News), and audio-visual databases have also been created.
- *Multilingual speech corrector (SPECO):*
  In the framework of an EU Copernicus project a system under the fantasy name of “Magic Box” was developed. This system provides help for speech training and speech therapy audio-visually for speech- and hearing-impaired persons in Hungarian, English, German, Slovenian and Swedish. This application will be extended with a prosodic module according to our latest research results [23,20,18].
- *Keyword recognition system:*
  Recognition of a keyword without recognizing previous and following speech segments.
  Due to integrating both co-articulation and higher level pronunciation into the pattern matching process the recognition accuracy may be quite high. Because of the lack of the linguistic level this approach is not suitable for detecting short keywords. Only one keyword may be found in an announcement. The solution is definitely recommended for recognition of proper names.
- *Large vocabulary, speaker independent, real-time application optimized for the transcription of broadcast news:*
  This solution takes into account the morphology of the Hungarian language extensively. Consequently the recognition error was nearly halved compared to traditional word-based technologies. In case of speaker adaptation the word error rate was reduced below 20% on a one hour test corpus which is at state-of-the-art level compared to similar languages [7].

Although automatic speech generation and recognition technologies have improved a lot, “talking machines” (so-called *dialogue systems*) can be used only among strongly constrained situations. The reason for this is that modeling such basic phenomena of human communication as linguistic, environmental and background knowledge is still at its infancy. In case of natural dialogues we know where and to whom we are talking to and based on our earlier experience we can guess the topic of the communication, the speaking style of the speaker, etc. Most current commercial recognizers do not convey such basic information as the sex and the speaking rate of the speaker. Speech synthesizers are typically not able to change speaking styles, to express emotions and to adapt to the partner.

Speech based dialogue systems

Taking into account the above mentioned limitations there are already speech based dialogue systems operating in Hungary. Such systems currently can be successful only if the domain of the conversation is sufficiently narrow and if we inform the human user that the other partner is a machine. Such an example is the DrugLine (in Hungarian: Gyógyszervonal; *www.gyogyszervonal.hu*) [9] information system that provides web, wap and telephone based interfaces. It ensures the availability of the Patient Information Leaflets of drugs that are approved by the Hungarian National Institute of Pharmacy through three different channels. The speech based dialogue was implemented in the telephony version (adapted speech recognition and speech recognition subsystems are integrated. The phone number of the system is +36-1 8869490.

In the USA there is a widespread technology that allows the connection of an operator or an appropriate department just by pronouncing the name without keying in an extension number. A similar technology is available in Hungary as well [13], but is used by smaller organizations yet – such as some local governments – although the technology would exhibit its real advantages in case of large entities (banks, insurance companies, ministries, etc.).

3. Vision for R&D and applications of speech technology

In the field of *automatic speech generation* one of the focus areas is applying the data-driven, easier to automate HMM technology while preserving the quality of the corpus-based approach. Increasing the naturalness, social appropriateness of the generated speech is of growing importance. As a consequence, besides the general-purpose systems there are increasing numbers of outstanding quality systems in limited domains.

Besides the above mentioned solutions an important area is voice telephony access to timetables and ticket ordering of public transport systems (railways, local and long-distance buses) and providing quick access over the telephone to information of banking, insurance, state and local government systems quickly, efficiently and 7 days/24 hours.

In the area of *automatic speech recognition* efficient applications can be implemented in several domains with currently available technologies. It cannot be regarded, however, a market of out-of-the-box solutions. On the contrary, each application needs thorough preparation and pre-processing work. In order to achieve a breakthrough for more widespread applications there should be advancements in some areas.

- Noise is the hardest limit on recognition accuracy. Noise robust models and noise resistant pre-processing algorithms need even greater attention. This task is language-independent to a large extent so results for a given language may be generalized.
- Another large research area is the recognition of spontaneous conversational speech. If we look at the advancement of past decades we can recognize that technology has proceeded from well defined read speech of both acoustic and linguistic viewpoint through designed and spontaneous speech to conversational one. The last one is just as "loose" from both acoustic and linguistic viewpoint as the language of Internet forums, for example. Intensive research in this area is of great importance in order to understand natural language communication.
- In case of Hungarian and similar agglutinative languages because of the variation of word forms the size of traditional language models can easily exceed the limits of several computing platforms. More efficient modeling techniques should be defined in order to find efficient solutions for these languages (e.g. Hungarian, Finnish, Turkish, Arabic, etc.). In Hungarian the relatively free word order is another dimension of future research.

Speech technology serving public information access

Nearly half of the Hungarian population is not an Internet user. Consequently interactive information services for all the citizens can only be solved by voice-based telephony. Speech technology is the key to provide automated, cost efficient solutions to this problem. This is the only way to bridge the widening "information gap". The idea of "digital public utility" may be worth to be extended to "information public utility" (the access channel to information important for the public).

Further information can be obtained from the authors and from the Hungarian Language and Speech Technology Platform *(www.hlt-platform.hu).*

Ackowledgements

The authors acknowledge the contribution of the following speech researchers of BME TMIT – Mátyás Bartalis, András Béres, Tamás Bőhm, Tamás Csapó, Géza Gordos, Krisztián Juhász, Géza Kiss, Laczkó Klára, Péter Mihajlik, György Szaszák, Bálint Tóth, Zoltán Tüske, Ákos Viktóriusz and Csaba Zainkó – to the results presented in the paper. Research presented in the paper has been supported by among others GVOP, NKFP, Jedlik and NTP programs of the Hungarian Government.

Authors

GÉZA NÉMETH (1959) obtained his MSc in Electrical Engineering, major in Telecommunications, at the Faculty of Electrical Engineering of BME in 1983, his Dr. Univ. degree in 1987 and the PhD degree in 1997. Dr. Nemeth is the Head of the Speech Technology Laboratory of BME TMIT. His research areas include speech technology, service automation, multilingual speech and multimodal information systems, mobile user interfaces and applications.

GÁBOR OLASZY (1943) is an electrical engineer and phonetician. He graduated from the BME, Faculty of Electrical Engineering, Branch of Telecommunications, in 1967, obtained his Dr. Univ. degree – BME (1985), Ph.D. in linguistics-phonetics (1988), DSc in phonetics (2003). Research areas include acoustics of speech, development of text-to-speech systems for different languages, embedding speech synthesis into applications, research for tools towards high quality synthesised speech, combination of stored speech method with synthesised items, modelling of prosody and sound durations for TTS.

TIBOR FEGYÓ (1973) obtained his MSc in Technical Informatics at the Faculty of Electrical Engineering and Informatics of in 1997. Research areas include automatic speech recognition, acoustic and language modeling, design and development of speech information systems, speech quality measurements of telecommunications channels.

KLÁRA VICSI (1948) is a speech acoustic expert. She obtained her M.Sc. degree at Faculty of Science of Eötvös Loránd University in 1971, the Dr.Univ. degree in 1982, her Ph.D. degree in 1992 and a DSc degree from the Hungarian Academy of Sciences in 2005, She obtained a Dr. habil. title from the Budapest Univ. of Technology and Economics in 2007. Research areas inlcude speech acoustics, computer speech recognition, preparation of speech databases, psycho-acoustics, project leader in phonetics and Hungarian speech databases, providing a basis for speech recognition tasks, she was the leader of an international project of the elaboration of multi-modal speech training and development processes. She was the organizer of numerous international conferences and summer schools.

References

[1] http://en.wikipedia.org/wiki/Star_Wars_Episode_IV:_A_New_Hope

[2] European COST Action 2102 (Cross-Modal Analysis of Verbal and Nonverbal Communication), http://www.cost2102.eu/joomla/

[3] Fegyó, T., Mihajlik, P., Szarvas, M., Tatai, P., Tatai, G., "Voxenter™ – Intelligent Voice Enabled Call Center for Hungarian". In: EUROSPEECH – INTERSPEECH 2003, 8th European Conf. on Speech Communication and Technology, Geneva, Switzerland, ISCA, pp.1905–1908.

[4] Kiss Gábor, Olaszy Gábor, "Hungarovox – a Hungarian language real-time dialogue speech synthetizer system". Információ Elektronika, 2. (in Hung.), Budapest, 1984, pp.98–111.

[5] MacinTalk, http://en.wikipedia.org/wiki/PlainTalk#The_original_MacInTalk

[6] Németh G., Zainkó Cs., Fekete L., "Statistical analysis for designing and developing an e-mail reader", Híradástechnika, Vol. LVI., 2001/1, pp.23–30. (in Hung.)

[7] Mihajlik, P., Tarján, B., Tüske, Z., Fegyó, T., Investigation of Morph-based Speech Recognition Improvements across Speech Genres, Proc. of Interspeech 2009, Brighton, U.K.

[8] Németh, G., Zainkó, Cs., Kiss, G., Olaszy, G., Fekete, L., Tóth, D., Replacing a Human Agent by an Automatic Reverse Directory Service; Proc. of 15th Int. Conference on Information System Development, Budapest, Hungary, Springer LNCS, 2006, pp.323–331.

[9] Németh, G., Olaszy, G., Bartalis, M., Kiss, G., Zainkó, Cs., Mihajlik, P., Speech based Drug Information System for Aged and Visually Impaired Persons, Proc. of Interspeech 2007, Antwerp, Belgium, pp.2533–2536.

[10] Németh, G., Kiss, G., Zainkó Cs., Olaszy G., Tóth, B., Speech Generation in Mobile Phones, In: D. Gardner-Bonneau and H. Blanchard (Eds.), Human factors and interactive voice response systems, 2n edition, Springer, 2008, pp.163–191.

[11] Németh, G., Zainkó, Cs., Bartalis, M., Olaszy, G., Kiss, G., "Human Voice or Prompt Generation? Can they Co-exist in an Application?", Interspeech 2009, Brighton, UK.

[12] G. Olaszy, G. Gordos, G. Németh, The MULTIVOX multi-lingual text-to-speech converter, In: G. Bailly, C. Benoit and T. Sawallis (Eds.): Talking machines: Theories, Models and Applications, Elsevier, 1992, pp.385–411.

[13] Olaszy, G., Németh G., Olaszi, P., Kiss, G., Gordos, G., "PROFIVOX – A Hungarian Professional TTS System for Telecommunications Applications", Int. Journal of Speech Technology, Vol. 3, Nr. 3/4. Kluwer Acedemic Publ., December 2000, pp.201–216.

[14] Olaszy Gábor, Electronic speech synthesis. Műszaki Kiadó, 1989, (in Hung.).

[15] Shoebox http://www-03.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_7.html

[16] Tóth, B., Németh, G., "Hidden markov chain-based artificial speech generation in Hungarian", Híradástechnika, 2008, pp.2–6. (in Hung.).

[17] Tóth, Sz.L., Sztahó, D., Vicsi, K., Speech Emotion Perception by Human and Machine. Proc. of COST Action 2102 International Conference, Patras, Greece, 9-31 October 2007. Revised Papers in Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction, Springer, 2007, pp.213–224.

[18] Vicsi, K., Roach, P., Öster, A., Kacic, Z, Barczikay, P., Tantos, A., Csatári, F., Bakcsi, Zs, Sfakianaki, A., A multimedia, multilingual teaching and training system for children with speech disorders. International Journal of Speech Technology, Vol. 3, Kluwer Academic Publisher, 2000, pp.289–300.

[19] Vicsi K, Velkei Sz., Development of a continuous speech recognition system, 3rd Hungarian Computer Linguistics Conference, Szeged, 2005, pp.348–359., (in Hung.).

[20] Vicsi, K., Computer Assisted Pronunciation Teaching and Training Methods Based on the Dynamic Spectro-Temporal Characteristics of Speech. In: Pierre Divenyi and Georg Meyer (Eds.): "Dynamics of speech production and perception", IOS Press, Amsterdam, 2006, pp.283–307.

[21] Vicsi K, Szaszák Gy., Using Prosody for the Improvement of ASR: Sentence Modality Recognition, In: Interspeech 2008, Brisbane, Australia, 2008.

[22] http://alpha.tmit.bme.hu/speech/databases.php

[23] http://rcs.hu/sc.htm

[24] http://alpha.tmit.bme.hu/speech/research.php

[25] http://alpha.tmit.bme.hu/speech/ikta_gastro.php

[26] http://en.wikipedia.org/wiki/2001:_A_Space_Odyssey

Congestion control and network management in Future Internet

MÁRTON CSERNAI, ANDRÁS GULYÁS, ZALÁN HESZBERGER,
SÁNDOR MOLNÁR, BALÁZS SONKOLY

Budapest University of Technology and Economics,
Department of Telecommunications and Media Informatics

{csernai, gulyas, heszi, molnar, sonkoly}@tmit.bme.hu

Keywords: Future Internet, congestion control, complex networks, socio-economics

Future Internet research programs try to ignore and overcome the barriers of incremental development and encourage clean slate designs, propose new visions, architectures and paradigms for the coming 10-20 years. Recent results in congestion control research has shown that networks operating without explicit congestion control (like TCP) may survive without congestion collapse if appropriately designed in network resources and if end systems apply appropriate erasure coding schemes. The exponential growth of the Internet makes it virtually impossible to manage the network with traditional centralized approaches (like the manager-agent one); hence research results of complex networks are expected to spread over the Internet with its autonomic behaviors. In this article we give overview and outlook of some interesting research areas connected to the future Internet research.

1. Introduction

The fast changes in the Internet usage and in its technology in the previous years have resulted in a growing expectation of possible paradigm changes in several mechanisms of the Internet. In this paper we address two challenging areas of these hot topics, namely, the solution for congestion problems and the management of future Internet.

The congestion is managed in the Internet by a congestion control mechanism called the Transmission Control Protocol (TCP), which is a complex transport protocol that has gone through several evolution steps since the beginning of the Internet. This evolution was driven by the ever-changing user and application requirements and also by the current technological limitations. As a result, a number of different TCP versions have been developed. TCP was originally designed as a reliable connection-oriented end-to-end transport protocol for the fixed wired Internet. However, the situation today is rather different concerning the pervasiveness of wireless technologies and also the increasing number of very high capacity links. It seems that the research community has arrived to a conclusion that it is very unlikely that an optimal TCP solution could ever be developed.

On the other hand, a new idea has arisen in the previous years, which basically says that we should design the solution for congestion problems from scratch. The idea challenges the researchers to think over the issue of congestion from a different point of view. What if we do not implement any congestion control in the network? How can the congestion be avoided in that case? Would it be possible to develop an Internet where the packet losses due to congestion events can efficiently be corrected by erasure coding? The TCP concept with its evolution and these exciting questions are discussed in the first part of the paper.

It is a well-accepted fact by current research concerning networks, that the future Internet will be characterized by the tons (in the order of trillions) of participating communicational entities and the complex and heterogeneous connections between them. Nevertheless the functional requirements of the next-generation networks will be highly diversified, and managing these networks with human interactions will be hardly solvable, thus the need for high-level automatic management is inevitable. Designing such complicated, large-scale systems is a complex task, and we still lack the adequate methodology for that. The means for describing large-scale networks developed a lot in the past years. We can see applications based on these findings, such as the search in complex networks, which will be discussed later in more detail.

Considering the future Internet, as in all complex systems, the analysis of the ongoing processes will require modern tools and methods. The conventional methodology, by which the global behavior of the system is treated as the collaborative functioning of different parts, lapses due to the complexity of the system. To handle these problems, we must come up with a self-organizing system model, where the centralized process control is replaced by distributed functioning and decentralized decision making. The monitoring of the network will take place on a macro level by analyzing the emergent features of the system, and not by independently observing the parts of it. In the second part of the paper, we discuss research topics in the area of large scale systems and consider self-organizing communication networks.

2. Congestion Control in Future Internet

Congestion control is a resource and traffic management mechanism to avoid and/or prevent excessive situations (buffer overflow, insufficient bandwidth) that can cause the network to collapse. It should not be confused with flow control, which prevents the sender from overwhelming the receiver. Congestion control has been accomplished by the *Transmission Control Protocol* (TCP) from the very beginning of computer networks and played an important role in the success of Internet.

The original protocol providing a reliable, connection-oriented service on top of IP networks dates back to 1981 (RFC 793). In the mid 1980s, serious incidents were experienced in the Internet when the network performance fell down by several orders of magnitudes. This phenomenon, called *congestion collapse,* raised the urgent need of some more sophisticated control mechanism in the transport layer. The original solution for the congestion collapse was provided in [1] by Van Jacobson. An essential part was added to TCP including the congestion control mechanisms. The congestion management of TCP is composed of two important algorithms. The *Slow-Start* and *Congestion Avoidance* algorithms allow the protocol to increase the sending data rate of sources without overwhelming the network and help to avoid congestion collapse. The protocol updates a variable called *congestion window* (`cwnd`, *w*) that directly affects the sending rate by means of limiting the number of unacknowledged packets in the network based on a sliding window mechanism which involves a *self-clocking* control. The congestion window variable is adjusted according to various algorithms in different phases of the connection. The basic mechanism was incrementally developed and tuned introducing new additional algorithms, e.g., RTO calculation and delayed ACK in 1989 (RFC 1122), SACK in 1996 (RFC 2018) and NewReno in 2004 (RFC 3782) just to mention a few. A standard TCP *(TCP Reno)* source starts sending according to the Slow-Start mechanism applying a multiplicative increase algorithm. More specifically, the congestion window is increased by a constant value for each acknowledgement received. This yields an exponential growth of the congestion window.

In Congestion Avoidance phase, the congestion window is adjusted by an AIMD (Additive Increase Multiplicative Decrease) mechanism which results in the classical "sawtooth" trajectory. When no packet loss is experienced, then the window is increased by $1/w$ per acknowledgements (AI) while it is halved as a response to packet loss (MD) as it is shown in the first row of *Table 2*. The main goal of the TCP's congestion control mechanism is to provide good network utilization, to avoid congestion collapse and to share the resources (now the link capacities) among end-users in a fair way. This is achieved by a distributed, closed-loop feedback mechanism. The last requirement regarding the *fairness* properties of the protocol is an important part of the next-generation transport protocol design.

TCP congestion control had managed successfully the stability of the Internet in the past decades but it has reached its limitations in "challenging" network environments. The new challenges of next-generation networks (e.g., high speed communication or the communication over different media) generated an urgent need to further develop the congestion control of the current Internet. In recent years, several new proposals and modifications of the standard congestion control mechanism have been developed by different research groups all over the world. These new mechanisms and TCP versions address different aspects of future networks and applications and improve the performance of

Table 1. High speed transport protocols

Protocol	Type	Proposed by	Main properties
HighSpeed TCP	loss-based	S. Floyd, International Computer Science Institute (ICSI), Berkeley University of California, 2003.	AIMD
Scalable TCP	loss-based	T. Kelly, CERN & University of Cambridge, 2003.	MIMD
BIC TCP / CUBIC	loss-based	I. Rhee et al., Networking Research Lab, North Carolina State University, 2004/2005.	good utilization, stability, linear RTT-fairness
FAST TCP	delay-based	S. Low et al., Netlab, California Institute of Technology, 2004. (now: FastSoft Inc.)	promising fairness properties
TCP Westwood	measurement-based	M.Y. Sanadidi, M. Gerla et al., High Performance Internet Lab, Network Research Lab, University of California, Los Angeles (UCLA), 2001–2005	several versions, different estimation methods
Compound TCP	hybrid	K. Tan et al., Microsoft Research, 2005.	AIMD + delay-based component
XCP	explicit	D. Katabi et al., Massachusetts Institute of Technology (MIT), 2002.	modification of the routers is necessary

regular TCP. For example, standard TCP (Reno version) cannot provide acceptable performance in wireless or mobile environments where the propagation delay and the available bandwidth can suddenly change (e.g., during inter-system handover) which can result in multiple back-offs or in extreme cases in disconnection. In order to remedy this problem, new TCP versions have been dedicated to this environment. The drawbacks of standard TCP Reno can be experienced in high speed wide area networks, as well. These networks can be characterized by *high bandwidth-delay product* (BDP) and TCP cannot efficiently utilize them due to its conservative congestion control scheme. As a response to this problem, the research community has proposed several new transport protocols recently referred as *high speed TCPs* or *high speed transport protocols*.

The huge number of new ideas has resulted in different new TCP versions implemented in several environments. In order to select the "optimal" transport protocol, extensive performance analysis is necessary in a wide range of network environments and applications. In the recent years, many papers were published deepening our understanding of these new protocols regarding performance characteristics, co-existence issues, and other important properties affecting the possibilities of their deployment. In the rest of the paper, a brief overview is given on some promising TCP versions. The main properties of the most important variants are presented in *Table 1* while a more detailed overview can be found in [2].

In TCP Reno, the congestion event is indicated by packet losses and the sending rate is reduced when losses occur. Protocols considering this type of congestion measure are generally referred to as *loss-based* protocols. This one-bit congestion indication does not allow sophisticated congestion control mechanisms. In addition, the permanent oscillation which is an intrinsic property of this mechanism raises stability issues. Therefore, fundamentally different approaches have also been emerged. In the case of *delay-based* algorithms, the round-trip time (RTT) is regularly calculated during the connection and the sending rate is adjusted according to the current value of the average delay estimation. In other words, a "multi-bit" congestion measure (delay) is considered in the control decision. The most recent TCP versions combine both principles and apply *hybrid* or *combined delay/loss-based* mechanisms. Other solutions suggest explicit congestion feedback from the network routers. These mechanisms using *explicit congestion indication* require the modification of the routers, as well.

HighSpeed TCP (HSTCP) [3] is a modification to TCP's congestion control mechanism for use with TCP connections with large congestion windows. It changes the TCP response function to achieve better performance on high capacity links. HSTCP is based on an AIMD mechanism where the increase and decrease parameters ($a(w)$ and $b(w)$) are functions of the current value of the congestion window (see the corresponding row of Table 2) yielding an adaptive and more or less scalable algorithm. HSTCP introduces a new relation between the average congestion window and the steady-state packet drop (or marking) rate. It is designed to have the standard TCP response in environments with mild to heavy congestion (packet loss rates of at most 10^{-3}) and to have a different, more aggressive response in environments of very low congestion event rate.

Ideas to introduce MIMD mechanisms for TCP have also been considered. Scalable TCP (STCP) [4] is a good example which has been suggested as an efficient transport protocol for high speed networks. Here, the multiplicative increase and multiplicative decrease algorithm guarantees the scalability of the protocol. The congestion window is increased by a constant parameter (a) as a response to a received acknowledgement, while it is reduced in a multiplicative manner (by bw) in case of packet losses (see Table 2). A proposed setting for the constants are $a=0.01$ and $b=0.125$ [4].

Table 2. Details of some TCP versions

Protocol	Window adjustment	When	Reaction to loss
TCP Reno	$w \leftarrow w + \frac{1}{w}$	per-ACK	$w \leftarrow 0.5w$
HSTCP	$w \leftarrow w + \frac{a(w)}{w}$	per-ACK	$w \leftarrow w - b(w)w$
STCP	$w \leftarrow w + a$	per-ACK	$w \leftarrow w - bw$
BIC TCP	$w \leftarrow w + \frac{a}{w}, \quad a \in \left\{ S_{\min}, \frac{W_{\max} - w}{B}, \frac{w - W_{\max}}{B-1}, S_{\max} \right\}$	per-ACK	$w \leftarrow \beta w$
FAST TCP	$w \leftarrow \min\left\{ 2w, (1-\gamma)w + \gamma\left(\frac{baseRTT}{RTT} w + \alpha \right) \right\}$	periodically	$w \leftarrow 0.5w$

In order to solve the TCP's severe RTT (round-trip time) unfairness problems, BIC TCP has been developed [5]. BIC TCP combines two schemes called additive increase and binary search. When the BIC TCP source gets a packet loss event, the congestion window is reduced by a multiplicative factor (β); and the maximum window parameter (W_{max}) is set to the value of the congestion window just before the reduction while the minimum window parameter (W_{min}) is set to the current value. Then the protocol performs a binary search between these parameters by jumping to the "midpoint" between the bounds. (More exactly, this jump is based on the B parameter of the protocol.) If packet loss does not occur at the updated window size, that window size becomes the new minimum; if packet loss occurs, that window size becomes the new maximum.

An important restriction is also introduced, the growth cannot be more aggressive than a linear one with a constant parameter (S_{max}). This process continues until the window increment is less than a small constant (S_{min}), when the window is settles down around W_{max} (increasing slowly on a "plateau"). This mechanism yields an "AIMD-like" behavior where the growing function is most likely composed of a linear phase (additive increase) and a logarithmic one (binary search). When the updated window size exceeds the current maximum, then a new equilibrium state has to be found and BIC TCP enters into the max probing state. During this phase, the growing function is the inverse of the previous ones, more exactly, the window is increased exponentially first (which is very slow at the beginning) and then linearly.

This complex mechanism is also summarized in the corresponding row of Table 2. The good performance of the protocol, including good utilization, linear RTT fairness (RTT unfairness is proportional to the RTT ratio as in AIMD), good scalability, and TCP-friendliness, comes from the slow increase around W_{max} and the aggressive linear increase of additive increase and max probing phases. Further research with BIC TCP has been resulted in CUBIC. CUBIC [6] is an enhanced version of BIC TCP. It simplifies the BIC window control and improves its TCP-friendliness and RTT-fairness. It is worth noting that the default TCP protocol of Linux kernels from version 2.6.8 was BIC TCP. From kernel version 2.6.19, the default protocol is CUBIC.

The research on the delay-based ideas has resulted in FAST TCP [7]. FAST TCP has the same equilibrium properties as TCP Vegas but it can also achieve weighted proportional fairness. FAST TCP seeks to restrict the number of its packets queued through the network path between an upper (β) and a lower (α) bound, however, the behavior is usually controlled by a single parameter (α) that can be considered as the targeted backlog (packets in the buffers) along the flow's path [7]. Under normal network conditions, FAST TCP periodically updates its congestion window based on the comparison between the measured average RTT and the estimated round-trip propagation delay (when there is no queueing). More exactly, the window is adjusted according to the formula presented in Table 2, where γ is the step size affecting the responsiveness of the protocol, and `baseRTT` is the minimum RTT observed so far which is an estimation of the round-trip propagation delay. The parameter α controls the equilibrium behaviour, therefore the appropriate setting of this parameter is crucial. FAST TCP also reacts to packet losses by halving its congestion window.

The delay-based control also appears in other proposals like TCP-Africa [8]. TCP-Africa is a hybrid protocol that uses a delay metric to determine whether the bottleneck link is congested or not. In the absence of congestion it uses an aggressive, scalable congestion avoidance rule but in the presence of congestion it switches to the more conservative Reno congestion avoidance rule. The combination of the delay-based and the loss-based approaches also appears in TCP-Illinois [9]. TCP-Illinois uses loss as a primary congestion signal and delay as a secondary one. The protocol uses an AIMD mechanism but adjusts the increase and decrease parameters based on experienced queueing delay.

Compound TCP [10] is another important example where a synergy of delay-based and loss-based approach has been implemented. It uses a scalable delay-based component in the standard TCP Reno congestion avoidance algorithm. Compound TCP has been developed in the Microsoft Research and it is the default TCP protocol of Windows Vista and Windows Server 2008. Moreover, it can be installed for other Windows versions by downloadable hotfixes and in addition, the Linux implementation is also available.

The idea of incorporating accurate bandwidth estimations into the TCP congestion control has also opened a new path in TCP research. TCP-Westwood [11] is a prominent example where eligible rate estimation methods to intelligently set the congestion window and slow-start threshold have been introduced.

Another important group of congestion control protocols is based on explicit congestion notification instead of the implicit congestion signals such as packet loss or delay. These congestion control schemes require the assistance of network routers by this means the modification of the routers is also necessary. This is a serious disadvantage from the aspect of deployment feasibility. One of the main representatives of this group is the eXplicit Control Protocol (XCP) [12] which generalizes the Explicit Congestion Notification (ECN) proposal. Instead of the one bit congestion indication used by ECN, XCP capable routers inform the senders about the degree of the congestion at the bottleneck. In addition, XCP decouples the utilization control from fairness control.

The history of the research of congestion control protocols revealed that it is difficult to find an optimal protocol that meets all the challenges of the evolving Internet. It is very likely that the task to find a universal and optimal congestion control protocol is impossible. This

view is supported by the fact that the developments of new applications show that they use their own congestion control mechanisms. These mechanisms in most of the cases are not *TCP friendly* so they cannot work together with TCP efficiently.

An interesting research is proposed in the framework of *GENI (Global Environment for Network Innovation)* advocating a *future internet without congestion control*. The basic idea is that flows do not attempt to relieve the network of congestion but rather send as fast as they can whenever they have data to send. Of course, if all flows are sending at maximal rates, then the packet loss rate within the network is probably high. To overcome this problem, flows can use efficient erasure coding.

This solution has several advantages but also raises some unsolved problems too. One of the biggest advantage is that we can achieve maximum resource utilization. It is because end hosts send packets as fast as possible and all available network resources between source and destination are utilized as much as possible. Links are constantly overdriven so any additional capacity is immediately consumed. This solution is *the most efficient regarding network resource utilization*. Another advantage is that this proposal can use *simple router architectures*. Routers no longer need to buffer packets to avoid packet loss so no need for expensive and power-hungry line-card memory. This can also result in significant decrease of the end-to-end packet delays for *supporting delay sensitive applications*. Moreover, this solution perfectly fits to a *network with all-optical cross-connects*.

With all these advantages we can also face a number of problems to solve. The most crucial question is that what *performance* can be achieved by implementing erasure coding techniques. The promising fact is that there is a number of new coding techniques proposed in the last decade with robust characteristics and high performance like fountain codes [13]. Another problem to solve is how to provide *fairness*. A mechanism is needed in the switches to perform selective packet dropping. As an example the *Approximate Fair Dropping* (AFD) [14] is a promising candidate to do this task.

Research is in the early phase but researchers can report some surprising results from their study. It seems that the congestion collapse is not as usual in networks without congestion control as it was believed [15]. Early results show that efficiency remains higher than 90% for most network topologies as long as maximum source rates are less than the link capacity by one or two orders of magnitude. It is also possible that a simple fair drop policy enforcing fair sharing at flow level is sufficient to guarantee 100% efficiency in all cases. Of course, there are several questions unanswered and new challenges to meet because present studies use some assumptions which are not fulfilled in practice.

An intensive research is needed to answer the exciting question whether we can build the future internet without congestion control and forget about TCP and its all problems.

3. Paradigm shift in managing networks

3.1 Large scale networks

One of the most important processes regarding the Internet today is the migration from the 20-year-old IPv4 protocol to the IPv6. The most demanding issue behind the process is that we are running out of the available IPv4 network addresses. The 32 byte address space of the IPv4 (~4.3 billion possible addresses) was created, when the computer (mainframe) to people ratio was 1:200. Nowadays this ratio is reaching 1:1 with 1.2 million users, which number can easily double in the near future due to the new users of the developing countries (according to estimates the number of users grow by 150 million every year). If we look at the spread of mobile devices, we can easily calculate that in the near future one person might even possess 200 network devices. According to assumptions, by 2010 the number of mobile devices will reach the number of PCs connected to the Internet. The typical tendency is that the PDA devices turn into a communicator device, but the real breakthrough will be the introduction of communications implants. In the world of sensors/actuators and intelligent materials we will see micro devices integrated into the human body, which are capable of wireless connectivity and will be used as life supporting or human-computer interaction devices. The RFID (Radio Frequency IDentification) is a really simple sensor network, where the active or passive devices can identify themselves, and it is already widely used.

The complexity of the Internet is growing not only by the addition of new network devices, but the rise of new available online services demands logically connected networks apart from the underlying physical one. This tendency is also reflected by the virtual ISPs (Internet Service Providers), mobile service providers and the development of virtual private networks. We can even mention the widespread use of peer-to-peer communicational methods, which all are built on the logically connected overlay networks. This significantly complicates the effective manageability of these networks.

For the effective analysis of large-scale complex networks first we need to develop such theory, which sets aside from the individual characteristics of the nodes, and concentrates on the structure of the connections and the character of the network. Moreover, the findings of this research field can be useful not only in the information technology. Many real world networks can be described with complex network models, for instance an organization, which is a network of people connected to each other. Also such networks are food webs, the global economical system or the connections between words in a language. We can also mention the diseases, which spread on the human social network (i.e. STDs). In general the research on complex networks concentrates on the various characteristics and the dynamic behavior of the networks.

Since the ‘50s the complex networks were described by the Erdős-Rényi [16] model, which was the only

reasonable and adequately precise approach at that time. Still researchers presumed that real world networks are neither completely regular, nor completely random. The widespread use of computers and the Internet generated large databases, and these databases are easily accessible.

By analyzing these topological data, researchers made two important discoveries in the last two decades, one of them is the Watt and Strogatz "small world" effect; the other is the Barabási-Albert scale-free network model. The small world effect describes the same phenomena, which was presented in Milgram's famous experiment in the '60s [17]. He found that even if there are many billions of people in the world, the shortest route consists of only several hops between two randomly chosen individuals in the social network. The scale-free network model highlights another interesting feature of complex networks, namely the complex networks have scale-free degree distributions, and not Poisson distribution, which is the characteristic of random networks. The scale-free distribution means that it is highly probable for high degree nodes (hubs) to evolve in large-scale network *(Fig. 1)*.

There are three fundamental characteristic features of the complex network models, which are worth to emphasize: *average path length, clustering coefficient and degree distribution* [18]. Average path length is the average of the shortest distances of any two random nodes in the network. This distance represents the effective size of the network. It was an interesting discovery that most of the real world complex networks have relatively short average path length, which feature led to the name "small world". The examination of the basic parameters of complex networks was an important step in this scientific field. Based on these parameters, we can intuitively build up different mathematical models, which result in networks with similar statistical characteristics.

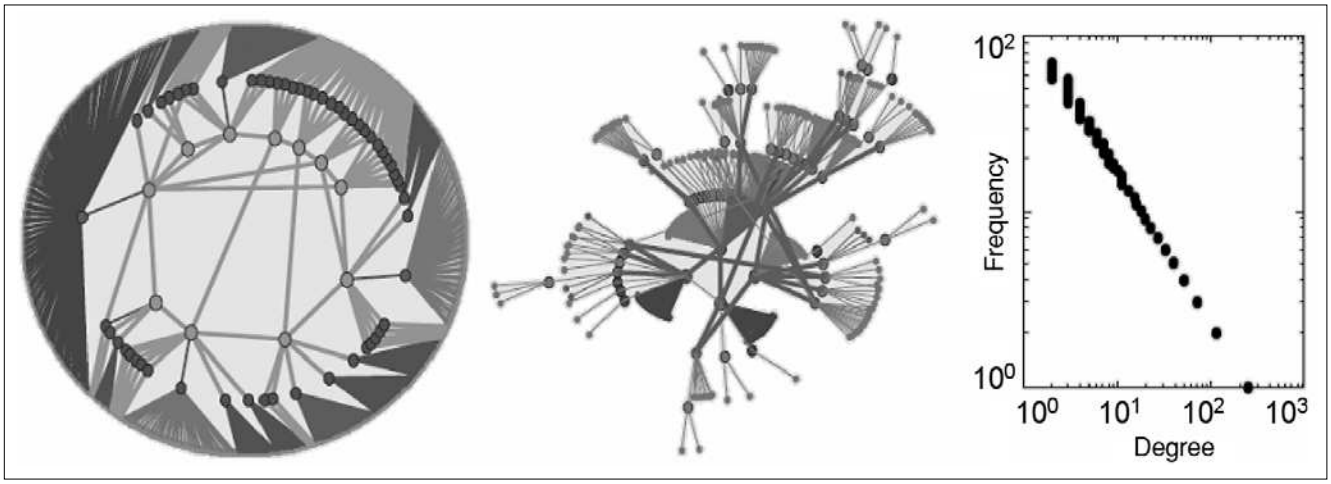
Another interesting topic in the field of complex networks is the problem raised by dynamic systems. Interesting observations were made about the synchronization of routing messages going over the Internet. Although the topology of the network was not specifically built for this purpose, the routers still synchronize their message exchanges easily, and when we break the synchronicity in one part of the network by introducing some randomness (altering a deterministic protocol), another part of the network will get in sync [19]. We can find more detailed treatment about these kinds of phenomena in the field of self-organizing systems.

3.2 Evolution and self-organization in communicational networks

Today telecommunication systems apply the widely used global network management paradigm. In this approach an external regulating unit controls the system. This unit constantly monitors the state of the system and the environment. When a problem occurs in the system, or the environment changes, the regulator calculates the appropriate response, and drives the system into the respective state *(Fig. 2)*. This solution is only feasible as long as we can find the appropriate response much faster, than the system changes its states. This criterion obviously delimits the permissible complexity and dynamism of the system, because the controller needs to be much more complicated than the system itself. For instance we can mention the link-state protocols, which cannot be used on large scale and complex topology.

A natural way to deal with complexity is self-organization, by utilizing the complexity of the system in the management plane. In a self-organizing system a large number of intricately connected devices achieve a global function by following simple local rules. It is shown in Fig. 2 that this way the control loop is integrated inside the system itself, thus the system can organize itself. Such a system evolves on its own obeying its given limitations. However, we don't know certainly what happens at a given point of the system, we can observe a precisely definable global behavior on the system level. Self-organization is not a feature of the system, but a paradigm, which can be helpful in understanding

Figure 1. Router level models of the Internet: engineering model (left), scale-free model (middle) and the degree distribution (right)



and designing certain real world systems (i.e.: complex telecommunication networks).

Algorithms showing classical signs of self-organization have played an important role from the beginning in the evolution and success of the Internet. We can mention the TCP protocol, which was previously discussed in the article. This protocol also uses a self-organizing technique while it deals with network congestion. It can control traffic flow parameters of a link in a decentralized manner, achieving an emergent result, such as fair resource distribution or high link efficiency. During the process every node makes strictly local decisions based on local information to reach a global goal.

Another example is the CSMA/CD algorithm, which is used in the Ethernet protocol. The rules of CSMA/CD guarantee that the communicating parties are able to detect the simultaneous transmissions and the consequent collisions on the common channel, and also prevent these collisions without a centralized control mechanism. If more nodes try to send packets on the same CSMA/CD channel at the same time, then the parties independently stop transmitting for a random time interval, hoping that next time they try to send packets, their packets won't collide with each other. If there is still collision on the channel, they increase the wait time interval, thus decreasing the probability of another collision. It is easily deductible that following these simple local rules the system realizes the fair and efficient resource distribution of the common transmission channel between the communicating parties.

The self-organizing systems belong to a highly active interdisciplinary research field, and they play an important role in many disciplines (biology, physics, social sciences). The systems utilizing these principles have many advantages, however we can only find few applications in the engineering fields. This can be explained by our lack of complete understanding concerning the mechanisms of self-organization. Designing such systems require an essentially new approach and design methods compared to our traditional ones.

3.3 Search in large-scale networks

Many large-scale networks, which are present in nature (human social networks, protein networks, neural networks, etc.), have good searchability as their important feature. Milgram's experiment [20] showed in 1961 that in human social networks not only short routes exist, but people are able to find them very efficiently depending on only the knowledge of a small local part of the whole network.

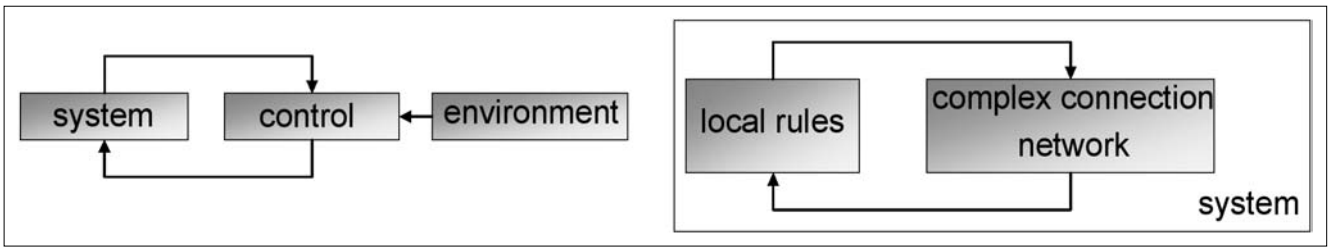
To achieve good searchability, which evolves in a self-organizing manner in nature, is a difficult task in artificial networks. An important example is any large-scale telecommunication network.

In the early development of the original concept of the Internet one of the most important elements was the design of a good routing protocol. The base of the framework that is still used consists of designated devices called routers, and the network emerging from the connections between these. This structure has the characteristic that some designated devices must have partial or even global knowledge of the whole topology to route messages. Since to gather the adequate information about the network topology takes time, the system must be quasi-static, and changes in the topology may happen with a limited speed. Considering that the next generation Internet will contain nodes by two or three times bigger order of magnitude than the present network, and these nodes will dynamically change the topology of the network, the original concept needs to be significantly adjusted.

The realization of efficient searchability is considered to be an important problem also in peer-to-peer networks. These networks utilize the Internet as a base structure, and they create an overlay network, which is responsible for the special P2P addressing and content searching functions. To cope with the strong network dynamism, and the high user activity as they connect and disconnect to the network, the system either uses a deterministic, but less scalable "network-flooding" technique, or apply a non-deterministic, more scalable method. These applications still highly depend on the underlying Internet infrastructure. Our research initiatives aim at developing an overlay network infrastructure (and the corresponding network protocols), which is able to handle numerous active (connecting, disconnecting or failing) users without a notable performance decline. The novel idea behind the research is the use of complex network structures instead of the ring structures present in P2P overlay networks.

The ongoing research about future networks more and more deals with "clean slate" designs. The recommended technologies show many similarities with the infrastructure-free ad-hoc networks. One of these concepts is the geographic position based addressing or geometry addressing and the search algorithms working with them. The communicating terminals can be mark-

Figure 2. The global management (left) and the structure of a self-organizing system (right)



ed by their geographical coordinates, and the routing is based on a simple greedy algorithm: if *X* is looking for *Y, X* first looks for its neighbor *Z*, which is closest to *Y*. In order for this algorithm to work, the network topology must meet some given conditions, for instance the Unit Disk Graph (UDG) is an applicable structure. In such cases, when the required connectivity conditions are not met, and for example it is true that *X* is not connected to *Y* and every neighbor of *X* is farther to *Y* thus there is no next step, some adjusted procedures need to be used in the algorithm.

The use of virtual coordinates can be the solution in this case. The task is to assign these virtual coordinates to the corresponding terminals, so that the greedy condition is realized, and the routing is always guaranteed [21]. The condition can be formalized in the following:

> For every pair of nodes *X* and *Y* ($X \neq Y$) there exists a node *Z* that $d(Z,Y) < d(X,Y)$,
> where $d(A,B)$ is the distance between *A* and *B*.

The virtual coordinates can be abstracted from the coordinates of the Euclidean plane or space, and they can be chosen from other abstract sets, after we defined a corresponding distance measure. The greedy routing capable virtual coordinate addressing assignment is called greedy embedding. If the greedy embedding is not possible due to the connection graph's special characteristics or other circumstances (i.e. dynamic topology change because of moving or failing nodes or links), we need to introduce complementary search procedures, for example face routing [22]. Our ongoing research deals with analyzing complex network search algorithms and topology management algorithms, which are efficient at given topology constraints (i.e. max node degree).

The routing techniques presented above have a common feature: they are easily scalable for large-scale networks, because each node needs to have information only about its neighbors. There is no need for large routing address tables, and the decision-making mechanism is really simple. This type of techniques is often called router-free routing. The procedures can be really effective, if they are supplemented with addressing algorithms, which can provide network addresses in a self-organizing manner based on local rules. Our research aims at realizing applications in accordance with our theoretical findings concerning these issues.

The management of large-scale, dynamic and structure-free networks is an active research field, and although we already have many basic results at will, the great challenges of technological applications are still ahead of us.

4. Conclusion

In this paper an overview was given about two research fields of the future Internet research where paradigm changes are expected. In the first part the issue of Internet congestion control was discussed. It was presented that Transmission Control Protocol (TCP) has always been serving as a solution for handling and avoiding congestion problems in the Internet. The mechanism of TCP was discussed and a short insight was given into the TCP versions that were developed during the history of the Internet.

A new idea is also discussed for solving the congestion problems in the future Internet. It is not based on control but rather on erasure coding. This solution makes the tempting promise that a future Internet could be developed without any congestion control. The possibility of this solution is a topic of current research. In the second part of the paper the manageability issues of large scale complex networks have been discussed. It is shown that in the future Internet network management methodologies featuring self-organization must play an important role. As a widely researched area, the special topic of searching in large networks has been presented in more detail.

Authors

MÁRTON CSERNAI is an undergratuate student on the Budapest University of Technology. He is about to finish his MSc degree in next-generation communication networks. He is participating in the ongoing research on future internet technologies at the Department of Telecommunications and Media Informatics. His main research fields are large scale complex networks and next-generation communication networks.

ANDRÁS GULYÁS received M.Sc. and Ph.D. degree in Informatics at Budapest University of Technology and Economics, Budapest, Hungary in 2002 and 2008 respectively. Currently he is a research fellow at the Department of Telecommunications and Media Informatics. His research interests are complex and self-organizing networks, network calculus and traffic management.

ZALÁN HESZBERGER recieved his M.Sc. and Ph.D. degree in electrical engineering at the Budapest University of Technology and Economics (BME), Budapest, Hungary in 1997 and 2007, respectively. Currently he is an assistant professor at the Dept. of Telecommunications and Media Informatics at BME. His main research interests are future internet technologies and complex networking.

SÁNDOR MOLNÁR received his M.Sc. and Ph.D. in electrical engineering from the Budapest University of Technology and Economics (BME), Budapest, Hungary, in 1991 and 1996, respectively. In 1995 he joined the Department of Telecommunications and Media Informatics, BME. He is now an Associate Professor and the principal investigator of the teletraffic research program of the High Speed Networks Laboratory. Dr. Molnár has been participating in several European COST and CELTIC research projects. He served on numerous technical program committees of IEEE, ITC and IFIP conferences. He is active as a guest editor of several international journals and is serving in the Editorial Board of the Springer Telecommunication Systems journal. Dr. Molnár has more than 130 publications in international journals and conferences. His main interests include teletraffic analysis and performance evaluation of modern communication networks.

BALÁZS SONKOLY received his MSc degree in software engineering from the Budapest University of Technology and Economics in 2002. Now he is a research engineer at the Department of Telecommunications and Media Informatics. His interests are in the field of traffic modeling and high speed transport protocols. He has authored international journal and conference papers in the are of high speed networks and the evaluation of TCP protocols.

References

[1] V. Jacobson,
Congestion avoidance and control,
In Proceedings of ACM SIGCOMM '88,
pp.314–329., Stanford, CA, USA, 16-18 August 1988.

[2] S. Molnár, B. Sonkoly, T.A. Trinh,
A Comprehensive TCP Fairness Analysis
in High Speed Networks,
Computer Communications, Elsevier,
Vol. 32, Issues 13-14, pp.1460–1484., August 2009.

[3] S. Floyd,
Highspeed TCP for large congestion window,
IETF RFC 3649, December 2003.

[4] T. Kelly,
Scalable TCP: Improving performance in high speed
wide area networks,
ACM SIGCOMM Computer Communication Review,
33(2):83–91, April 2003.

[5] L. Xu, K. Harfoush, I. Rhee,
Binary increase congestion control (BIC) for fast
long-distance networks,
In Proceedings of IEEE Infocom '04, Vol. 4,
pp.2514–2524., Hong Kong, China, 7-11 March 2004.

[6] I. Rhee, L. Xu,
CUBIC: a new TCP-friendly high-speed TCP variant,
In Proceedings of Third International Workshop on
Protocols for Fast Long-Distance Networks
(PFLDnet 2005), Lyon, France, 3-4 Februar 2005.

[7] D.X. Wei, C. Jin, S.H. Low, S. Hegde,
FAST TCP:
motivation, architecture, algorithms, performance,
IEEE/ACM Transactions on Networking (ToN),
14(6):1246–1259, 2006.

[8] R. King, R. Riedi, R. Baraniuk,
Evaluating and improving TCP-Africa: an adaptive
and fair rapid increase rule for scalable TCP,
In Proceedings of Third International Workshop on
Protocols for Fast Long-Distance Networks
(PFLDnet 2005), Lyon, France, 3-4 Februar 2005.

[9] S. Liu, T. Basar, R. Srikant,
TCP-Illinois: A loss and delay-based congestion
control algorithm for high-speed networks,
In Proceedings of First International Conference on
Performance Evaluation Methodologies and Tools
(VALUETOOLS), Pisa, Italy, 11-13 October 2006.

[10] K. Tan, J. Song, Q. Zhang, M. Sridharan,
A compound TCP approach for high-speed and
long distance networks,
In Proceedings of IEEE Infocom '06, Barcelona, Spain,
23-29 April 2006.

[11] R. Wang, K. Yamada, M.Y. Sanadidi, M. Gerla,
TCP with sender-side intelligence to handle
dynamic, large, leaky pipes,
IEEE Journal on Selected Areas in Communications
23(2):235–248, 2005.

[12] D. Katabi, M. Handley, C. Rohrs,
Congestion control for
high bandwidth-delay product networks,
In Proceedings of ACM SIGCOMM '02, Pittsburgh,
PA, USA, 19-23 August 2002.

[13] M. Luby,
LT-codes,
The 43rd Annual IEEE Symposium on
the Foundations of Computer Science,
pp.271–280., 2002.

[14] R. Pan, L. Breslau, B. Prabhakar, S. Shenker,
Approximate fairness through differential dropping,
ACM SIGCOMM Computer Communication Review,
Vol. 33, Issue 2, April 2003.

[15] T. Bonald, M. Feuillet, A. Proutière,
Is the "Law of the Jungle" sustainable for the Internet?,
IEEE INFOCOM '09, Rio de Janeiro, Brazil,
19-25 April 2009.

[16] P. Erdős, A. Rényi,
"On the evolution of random graphs", Publ.:
Math. Inst. Hung. Acad. Sci., Vol. 5, pp.17–60., 1959.

[17] S. Milgram,
"The small-world problem",
Psychology Today, Vol. 2, pp.60–67., 1967.

[18] Xiao Fan Wang, Guanrong Chen,
Circuits and Systems Magazine, IEEE,
Vol. 3, Issue 1, pp.6–20., 2003.

[19] S. Floyd, V. Jacobson,
"The synchronization of periodic routing messages,"
IEEE/ACM Trans. Networking,
Vol. 2, No. 2, pp.122–136., April 1994.

[20] Milgram, Stanley,
"Behavioral Study of Obedience",
Journal of Abnormal and Social Psychology,
67(1963):371–378.

[21] Cedric Westphal, Guanhong Pei,
Scalable Routing Via Greedy Embedding,
In Proceedings of IEEE INFOCOM'09 Mini-Conference,
Rio de Janeiro, Brazil, April 2009.

[22] J. Li, L. Gewali, H. Selvaraj, V. Muthukumar,
"Hybrid Greedy/Face Routing for
Ad-Hoc Sensor Network,"
Euromicro Symposium on Digital System Design
(DSD'04), pp.574–578., 2004.

Media communications over IP networks – An error correction scheme for IPTV environment

LÁSZLÓ LOIS, ÁKOS SEBESTYÉN

Budapest University of Technology and Economics, Department of Telecommunications
{lois, sebestyen}@hit.bme.hu

Keywords: IPTV, webTV, quality of service, error correction

Video transmission over IP networks has been gaining more and more popularity recently. One of the crucial problems of video transmission over IP networks through unreliable links is the susceptibility to errors in the transmission path.
Packets lost or discarded by the NIC due to CRC errors must be somehow regenerated. Regeneration can be done by requesting a retransmission, or the packet can be recalculated provided that some redundancy is introduced in the transmitter side.
After a general description of media transmission over IP links, the paper describes a method that can be used for forward error correction in IPTV applications.

1. Introduction

For seamless playback of digital audio and video contact frames must arrive in the decoder at the pace of the frame frequency. Moreover, as the misalignment of audio and video produces degradation of the perceived quality, synchronization between audio and video must be maintained. In an IP environment, however, audio and video frames are transmitted in packets. These packets travel independently within the network thus suffering from timing and alignment problems. It is the decoder that, through buffering operations, tries to resolve timing inaccuracies and presents accurately synchronised content to the viewer.

Another important issue is the error prone transmission of information. The IP network provides only very basic error protection, which, in some cases, is not adequate to keep up the quality of the service. Hence, for certain services more robust forward error correction and quality control schemes need to be introduced.

The paper is organised around the aforementioned two topics and is structured as follows: Section 2 gives an overview on the architecture of the different IP applications and reveals their timing requirements. The next Section deals with the most dynamically evolving application area, IPTV, and focuses mainly on the different error correction schemes that can be used in an IPTV environment. Section 4 presents an implemented approach based on Reed-Solomon encoding and erasure decoding [1] for reducing the number of retransmission requests together with some measurement data. Finally, Section 5 concludes our paper with some possible utilization of the implemented forward error correction scheme.

2. Media communication applications

2.1 Architecture [2]

The quality of the service offered by a media communication application is determined by the following factors:

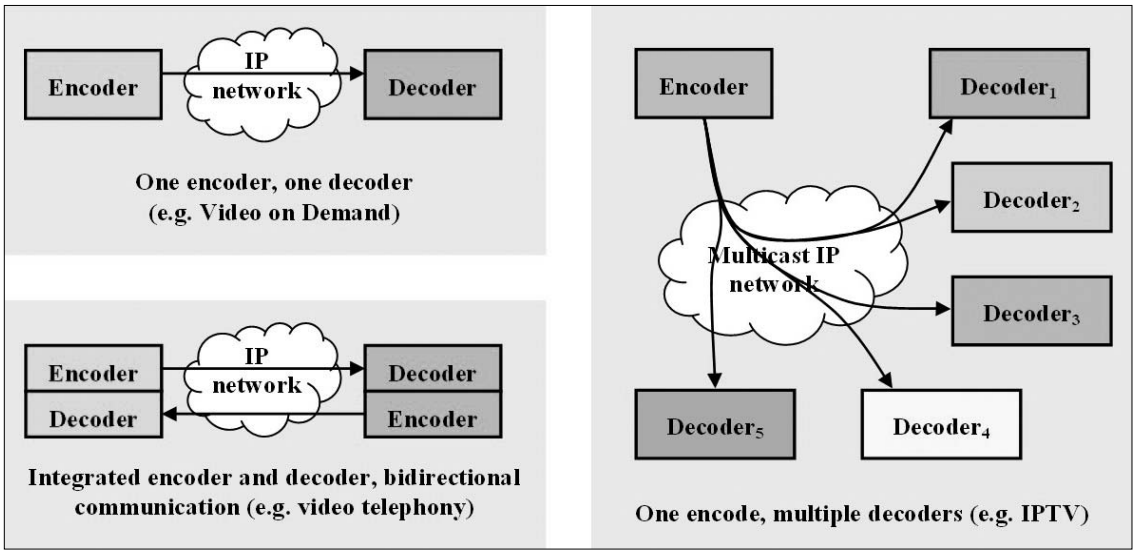


Figure 1. Basic layout of IP applications

- Quality requirements as set forth by the consumer and the service provider
- User terminal equipments
- Devices and media formats used by the service provider
- Quality of service parameters of the network or the interlinked heterogeneous networks

The basic building blocks of a media communication network are the encoder and the decoder. The encoder is responsible for converting the video, audio and data content to a format that can be transmitted through the network infrastructure, whereas the task of the decoder is to process such video, audio, and data content, restore their synchronism, and present them to the viewer.

Depending on the number and layout of the encoders and decoders, several basic network structures are possible *(Figure 1)*. The most simple media communication application consists of a single encoder (server) and a single decoder (client) which together make up a very basic peer-to-peer video on demand system (Fig. 1, top left). If both peers possess some encoding and decoding capabilities, they both can act simultaneously as a server and a client. Hence, encoded information can travel in both directions, and a video telephony system can be set up (Fig. 1, bottom left). The same architecture allows for conferencing services provided that more peers are allowed to join the telephony session.

A more sophisticated approach is when the IP infrastructure supports multicast transmission (Fig. 1, right hand side). In a multicast session, a single encoder supplies multiple decoders with a common output stream. The packets of this stream are duplicated (multiplied) and routed by special network components thus reducing the burden on the network.

The communication between the server and the client follows a hierarchical approach, in which each layer has its own task as depicted in *Figure 2*.

2.2 Timing accuracy in media communication

Media communications can also be characterised by the timing accuracy. The quality of timing accuracy is basically determined by the seamlessness of playback and the delay introduced by transmission and processing. Based on these quality factors three schemes can be distinguished. The properties of each are summarised in *Table 1*.

Regardless of the scheme used, the primary aim of transmission is to supply the user with as high an image and sound quality as can be ensured by the particular network in a reasonable time. A key point in that is choosing an appropriate buffering strategy which not only decreases jitter and ensures seamless playback, but also allows for either the possibility of retransmission or the correction of packet by FEC and interleaving.

An offline service is usually implemented by TCP/IP or HTTP protocols and is intended for downloading content to a temporary or final storage. In an offline service, the terminal equipment or a neighbouring network component has a storage capacity to store the whole content. The primary aim is safe delivery of content, transmission delay is only a secondary factor if relevant at all.

An online service such as a video telephony application tries to minimize the annoying delay of transmission and presentation. As any buffering operations increase the delay, the size of the transmitter and receiver buffer is usually limited to a couple of frames. As a result, there is no time to interleave the content in the transmitter side and perform FEC encoding (as that would also require buffering), or to request the resending of information in the receiver side. Because re-sending is not possible, unreliable protocols, such as UDP [3], are preferred.

Most of the current streaming services belong to the near-line scheme. Due to the lack of interactivity, basic streaming applications can tolerate longer delay. Buf-

Figure 2. Functional model of the media encoder and decoder

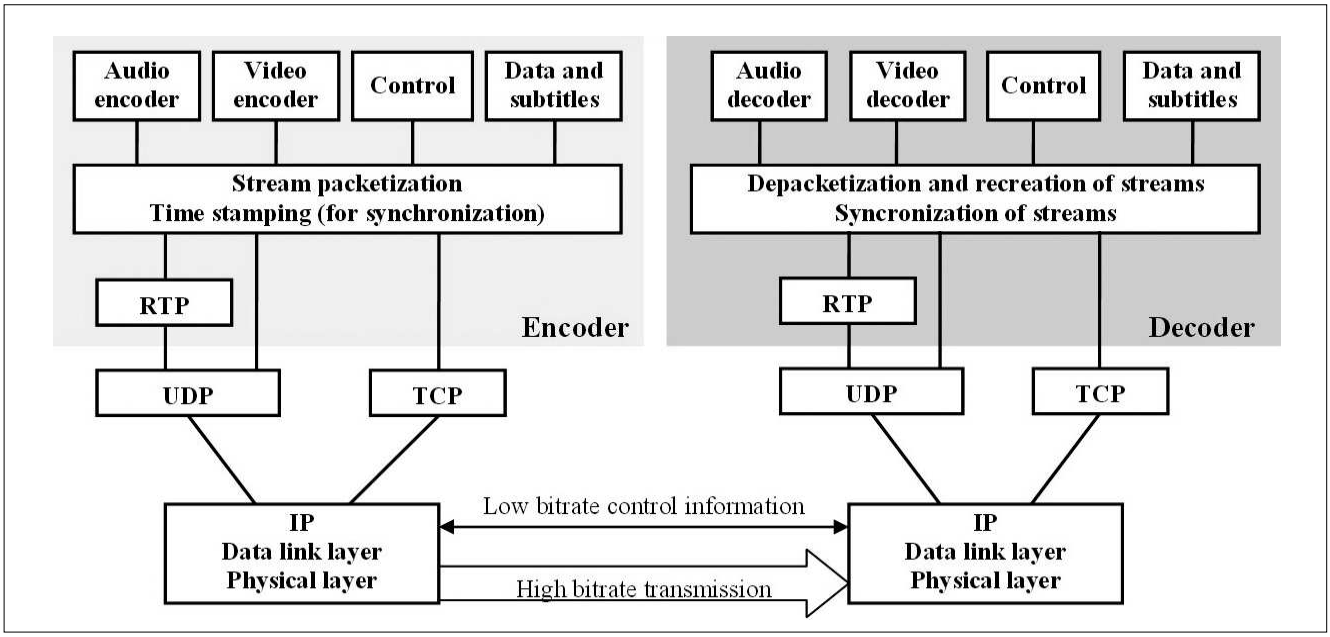


Table 1. Characterization of media applications by timing accuracy

	Offline scheme	Near-line scheme	Online scheme
Primary aim	Full and error free download	Seamless playback	Immediate playback with minimal delay
Seamlessness	Irrelevant	Yes	Yes
Presentation of each frame after reception	Irrelevant	After buffering delay	Immediately after reception
Buffering	Irrelevant	To ensure error correction and seamless playback	Minimal
Replacement of lost packets	Error correction or packet resending	Error correction or packet resending if allowed for by buffering	No resending of packets, only minor error correction
Typical application	Media file download	Media streaming	Video telephony, video conference

fering for a couple of seconds is common to streaming applications. The near-line scheme usually uses unreliable but more complex protocols to transmit information. In most cases a control channel is used through which control and link state information can be exchanged. A typical near-line service using real-time transport protocol (RTP [4]) is depicted in *Figure 3*.

The audio, video, or data sub-streams (depicted by a simple UDP block in Fig. 3) are either treated separately or combined to form a multiplexed stream. While in the former case separate UDP sessions and control channels are created for each individual sub-stream, in the later case the sub-streams are multiplexed and only one aggregate control channel is set up.

3. Streaming services, IPTV

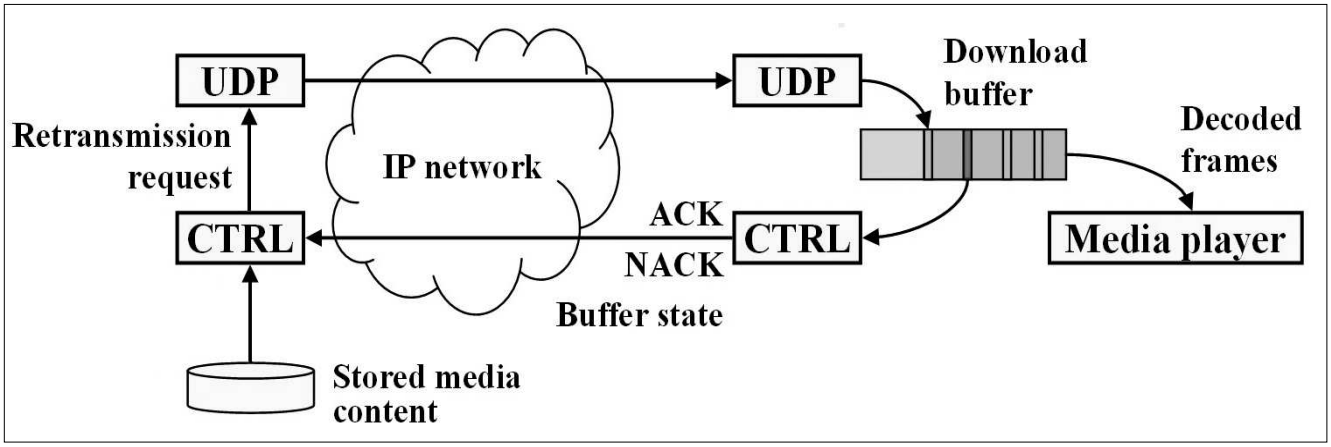
Streaming services usually use UDP/IP or RTP protocols. In a streaming environment not only error free transmission but also timing accuracy is a concern. A delay of some seconds is tolerable at the start of playback, but once started, playback must be uninterrupted and seamless. Although a short delay is allowed, it has to be minimized as much as possible. Since the delay introduced by the UDP protocol is minimal, it is considered to be a favourable choice.

Typical streaming services include applications like television through the Internet (webTV) and Internet Protocol Television (IPTV) services wich have been gaining more and more popularity recently. The difference between the two is while webTV is implemented on the public Internet, IPTV uses an IP infrastructure which is a private and closed system maintained by a service provider. It is this service provider who is responsible for providing the content, managing user access through access control mechanisms, and, in many cases, supplying the users with terminal equipments (so-called set-top-boxes). As the network is closed, the service provider has freedom to prioritise packets carrying television streams over other packets. (This is certainly not possible in webTV applications.)

IPTV is often bundled with other services like Voice over IP (VoIP) telephony, and supplementary Internet access. These three together form a service often referred to as TriplePlay.

The IPTV network is a switched digital video (SDV) architecture, in which only streams requested by the viewers are present. Programs that are not being viewed do not appear on the network or on the network segment. This approach, on the one hand, is very economic in the sense that it saves valuable bandwidth for the service provider which can be used to provide value-added ser-

Figure 3. A typical near-line streaming service based upon the RTP protocol



vices like Video on Demand (VoD). On the other hand, it allows for building an access network with only as much bandwidth as required by the subscription of each user. Therefore, the bandwidth of the access network for each user is determined by the number of programs that can be simultaneously accessed by them. If a user has access to only one standard definition (SD), MPEG4-AVC [5] encoded stream at a time, then the data rate can be as low as 2.4 Mbps. (Although 2.4 Mbps is considered to be a minimal data rate for standard definition AVC, due to technological reasons this is often reduced to 2 Mbps.)

The most typical consumer behaviour in an IPTV environment is watching live television streams. Since every program is watched by many users, the streams are transmitted in multicast mode. In a multicast session the network and the additional architecture take care of both packet multiplication and multicast group management.

If a program change is requested by the user, the user is moved from their previous multicast group corresponding to the program they were watching to a new multicast group corresponding to the newly requested program. Since multicast group switching is a slow process, and an additional buffering delay is introduced by the near-line scheme, a common method is to supply the user with packets of the newly requested stream through a VoD-like unicast connection of higher data rate. As soon as the multicast group switching is complete and the receiver buffer is full, the terminal equipment can switch to the normal multicast data stream and continue receiving that.

3.1 Error correction in IPTV networks

As far as the network and the data link layers are concerned, IP transmission features no forward error correction. The only error resiliency method used in a normal IP environment is the insertion of a CRC code into the packet that can be used for error detection. If a packet fails the CRC error check in the receiver, then it is discarded. How the system behaves in the case of packet loss is essential from the point of view of the service.

3.1.1 Retransmission of lost or erroneous packets

Packets can be lost due to congestion in the network, or they can be discarded by the network layer due to an invalid CRC code. Either way, missing packets of multimedia services, if not recovered, produce visual artefacts.

To reduce the quality degradation, missing packets can be recovered by requesting their retransmission. Retransmission makes sense only if the round trip time T_{RTT} that consists of the time needed to send the retransmission request plus the time the retransmitted packet arrives is less than the time T_{buf} the packet spends in the buffer till it is either presented or used as a reference to decode other frames:

$$T_{RTT} \leq T_{buf}$$

To satisfy the above condition the following approaches can be followed:

- RTT must be reduced by placing the server as close to the clients as possible. An IPTV network usually features one server at a predefined location and cannot be freely relocated. The problem of relocation, however, can be overcome by installing so-called secondary caching servers. The caching servers store a well defined portion of the streams that pass through them. Upon a retransmission request from a client, they are ready to resend the packet through a unicast connection.
- The buffering time T_{buf} must be increased. Since in an IPTV environment it means that the STB must be equipped with more memory, it is usually not a plausible approach.

3.1.2 Replacement of lost or erroneous packets

Missing packets can be replaced if enough redundancy is introduced in the system, and this redundancy can be used to regenerate the packets that have been lost. The redundancy information can either be inserted in the packets themselves, or it can be transmitted as a separate correction stream. As in the second approach the redundancy information can be discarded by receivers which do not need or do not support them, and there is no need to recalculate the CRC code of the original packets, this is a more plausible method.

The information can be recovered by utilising appropriate encoding (like systematic Reed-Solomon encoding) provided that the number of erroneous symbols remains below a well defined upper bound. The method for such encoding is the following:

(1) The transmitter appends N-K parity symbols to the K source symbols.
(2) The K source symbols and the appended N-K parity symbols together form the encoded word having a data length of N symbols.
(3) Out of the N encoded symbols E symbols are corrupted during transmission and N-E remain intact.
(4) The receiver receives the N encoded symbols, and recovers the E erroneous ones provided that:

$$2 \cdot E \leq N–K$$

if the error locations are unknown (normal error correction), or:

$$E \leq N–K$$

if the error locations are known (erasure error correction).

Both normal and erasure error correction means a trade-off between the channel capacity and error correction capability. To be able to correct E symbols out of the N received ones, $2E$ or E out of the N symbols must be parity information. A well-known systematic FEC encoding scheme that satisfies the above conditions and can be implemented quite easily is the Reed-Solomon encoding. In a test bed which is described in the next section, Reed-Solomon encoding and erasure decoding was used to provide a means to regenerate missing data.

4. Error correction as implemented in a real-world IPTV testbed

The method described in the previous section cannot be applied directly to the Ethernet packets, as that would imply an enormous calculation burden on the system. To reduce calculation complexity a so-called interleaving approach is followed, which is depicted in *Figure 4*.

4.1 Encoder

The encoder reserves two memory areas referred to as Network Data Table and RS Data Table. Each position within the reserved memory areas can hold one byte of information. The Network Data Table consists of *K* columns and 1346 rows, and is used to store incoming packets that need to be Reed-Solomon encoded. The RS Data Table has *N-K* columns and holds the parity information, a generated header and some additional information.

The Reed-Solomon encoder works above GF(8), therefore the upper bound for *N* is 255. K can be any arbitrary odd number between 1 and *N*. The difference *N-K* determines the error correction capability of the erasure Reed-Solomon decoder as seen in Section 3.1.2.

First, a copy of the Ethernet packets arriving at the encoder is buffered in the Network Data Table in a column-wise direction. If the packet length is less than the maximum packet length of 1340 bytes (corresponding to seven 188-byte transport stream packet plus the RTP header), the empty positions within the respective column are zero padded.

Once all *K* columns in the Network Data Table are completely filled, a 6-byte supplementary information consisting of the packet length and a CRC code is appended to the end of each column. Since this supplementary information can be regenerated in the receiver side provided that the packet it corresponds to does arrive, this is information is not transmitted.

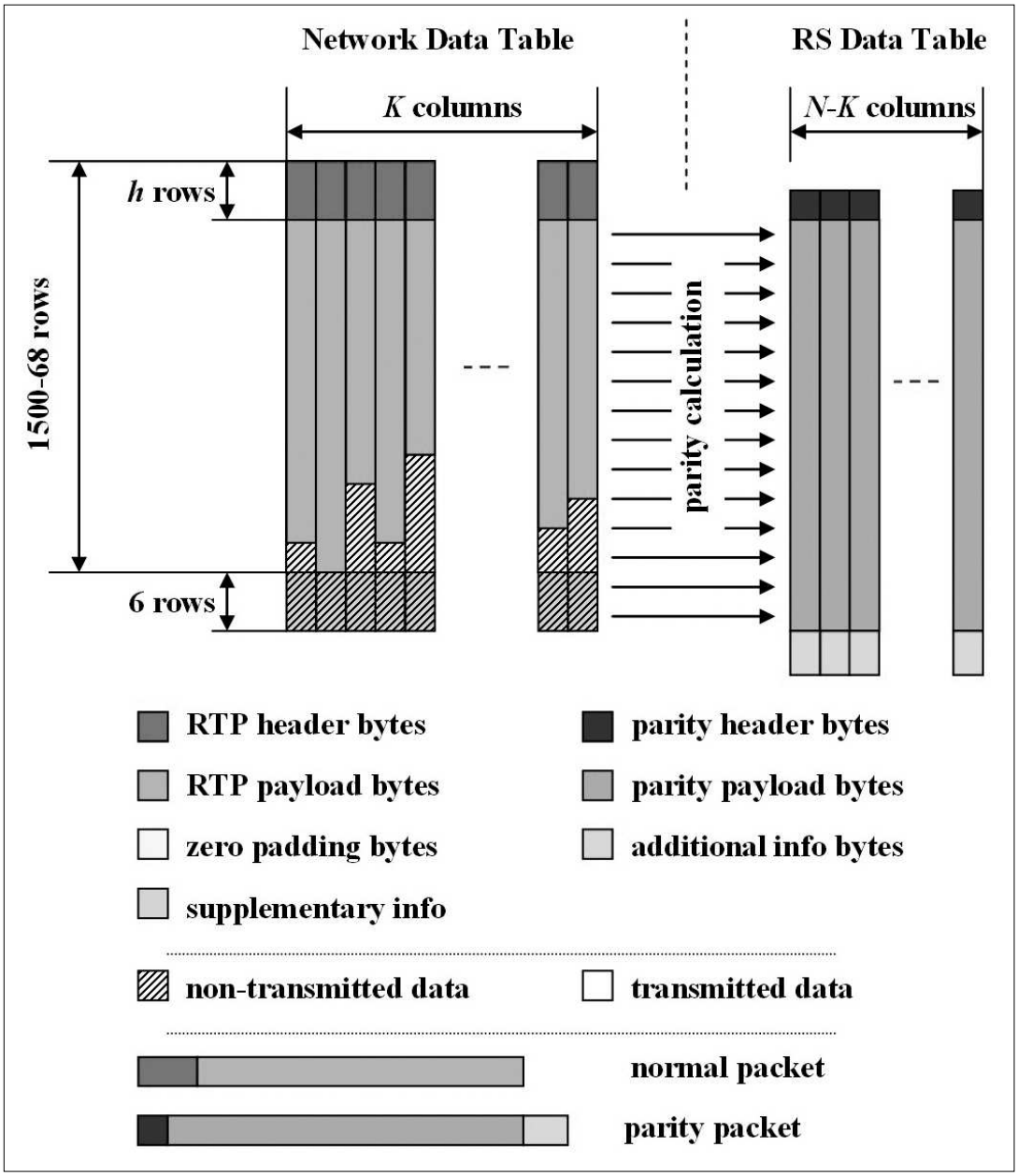


Figure 4. Network Data Table and RS Data Table used for interleaved encoding and decoding

In the next step, the useful payload of the packets (excluding the header), plus the zero paddings (if any) and the supplementary information are Reed-Solomon encoded row wise. The calculated parity information is stored in the appropriate row of the RS Data Table.

After the parity information is calculated for all rows, a parity header is generated for each column and some additional information is recorded in the RS Data Table. The header contains the source and destination addresses and ports as well as the position of the column within the RS Data Table. The additional information includes the sequence numbers of the RS encoded RTP packets, their positions within the Network Data Table plus parity information calculated from the length and CRC of the incoming RTP packets. These data are used in the receiver side for reconstructing the Network Data Table, spotting any missing packets, and for validating the data after reconstruction. Since the sequence numbers and the positions are crucial for the correct operation of the system, they are repeated in every column.

Finally the columns of the RS Data Table are transmitted as user packets.

4.2 Decoder

The decoder works in a similar fashion. It first restores the original order of both the normal and the parity packets, and saves the packets in the Network Data Table and RS Data Table respectively. If a packet, either normal or parity, is found to be missing, then all positions in the respective column are marked as erasures. Once all packets are accounted for (either inserted or marked as erased), the payload of the missing packets can be regenerated by performing erasure RS decoding row wise. After the payload is restored, the RTP header can easily be regenerated, thus all the missing information is regained.

4.3 System architecture and performance

The architecture of the testbed that uses interleaved RS encoding for IPTV streams is depicted in *Figure 5*. Both RS encoding and RS decoding of a predetermined stream were performed by computers with two NIC cards working transparently in a bridged configuration. The appropriate packets were filtered and passed to a user program which then performed RS encoding and RS decoding as described in Section 4. The packet loss in the network was modelled by a uniform distribution. The expected value of the ratio of lost packets and the parameters of the RS forward error correction could be freely chosen. The network traffic after RS decoding was monitored by a measurement device.

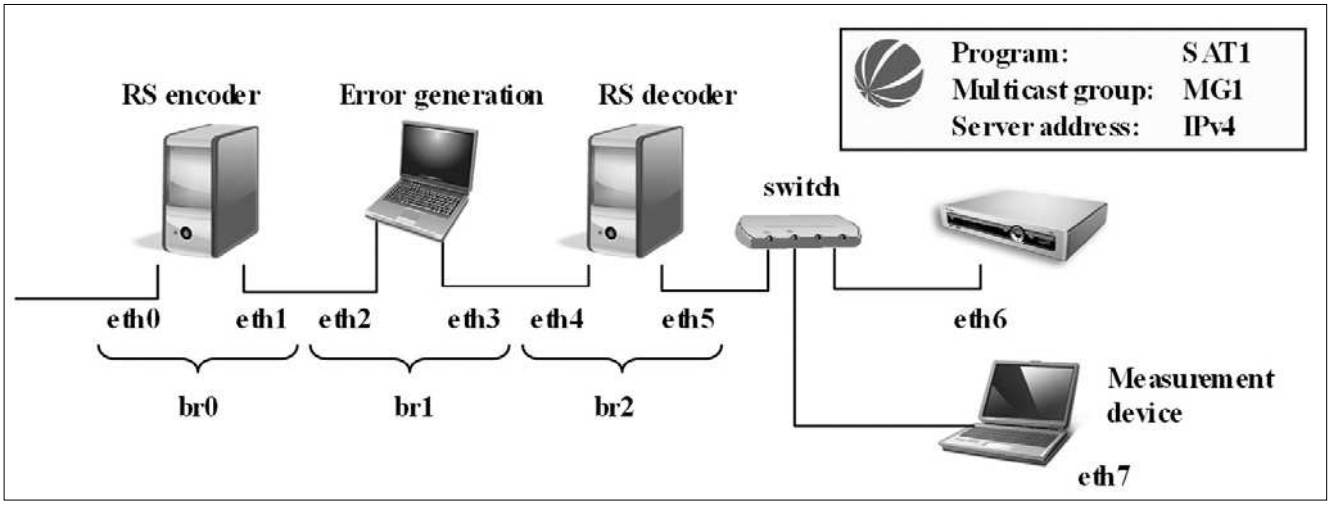
The measurement results for a packet loss ratio of 10% for different RS parameters are summarized in *Table 2*. (The packet loss ratio of 10% means that 10 percent of the packets after RS encoding is dropped by the error generator device.)

What is apparent from the figures (apart from the fact that packets with longer parity are more likely to be corrected) is that if the increase in bandwidth (i.e. the code rate) is kept constant, then the ratio of lost packets after RS decoding is decreased with the increase in message length. This agrees with our expectations as the longer the RS codeword, the less probable is that the number of errors per RS codeword will considerably exceed the expected value of the distribution, hence it is more likely that they can be corrected. In case of a codeword length of around 255 and uniform distribution, all packets could be corrected provided that the increase in bandwidth was twice the packet loss ration after RS encoding.

5. Conclusions and future work

In the paper we gave a general description of the two most severe problems of IPTV networks, and presented an approach that can be used for error correction in an IPTV environment. Although the method means an increase in the overall data rate for an IPTV stream (an increase inversely proportional to the code rate), it can be effectively used to decrease the unicast traffic thus making the caching servers obsolete.

Figure 5. The measurement setup



Message length (*K*)	Code word length (*N*)	Code rate	Increase in bandwidth	Ratio of lost packets after RS decoding
10	12	83.33%	20.0%	11.20%
10	14	71.43%	40.0%	0.90%
10	16	62.50%	60.0%	0.00%
20	22	90.91%	10.0%	37.70%
20	24	83.33%	20.0%	8.50%
20	26	76.92%	30.0%	1.20%
20	28	71.43%	40.0%	0.10%
20	30	66.67%	50.0%	0.00%
40	44	90.91%	10.0%	45.00%
40	46	86.96%	15.0%	17.20%
40	48	83.33%	20.0%	4.80%
40	50	80.00%	25.0%	1.00%
40	52	76.92%	30.0%	0.20%
40	54	74.07%	35.0%	0.00%
40	56	71.43%	40.0%	0.00%
40	58	68.97%	45.0%	0.00%

Table 2.
Measurement results for a packet loss ration of 10% after RS encoding for different coding parameters

As far as the future enhancements are concerned, the introduced error correction system can be expanded to use low density parity check codes instead of RS encoding.

Authors

LÁSZLÓ LOIS was born in 1971 in Tatabánya. He received his Masters and Ph.D. degree in software engineering from the Budapest University of Technology and Economics (BME) in 1995 and 2005, respectively. He is currently with the Department of Telecommunications at BME, and contributes to the research, development and educational activities. His research interests include source coding applications, media communications and IP television.

ÁKOS SEBESTYÉN was born in 1977 in Budapest. He received his Masters Degree in electrical engineering at the Budapest University of Technology and Economics, Department of Telecommunications in 2002. Now as a member of the staff of the Department, he contributes to the educational and research and development activities. His research areas include digital broadcasting (DVB-T/S/C/H/T2) and IP television.

References

[1] Didier F.,
"Efficient erasure decoding of Reed-Solomon codes"
http://arxiv.org/pdf/0901.1886v1
[2] ITU-T H.264:
Advanced video coding for generic audiovisual services,
Series H: Audiovisual and multimedia systems,
Infrastructure of audiovisual services –
Coding of moving video.
[3] RFC 768 – User Datagram Protocol.
[4] RFC1889 – RTP:
A Transport Protocol for Real-Time Applications.
[5] ISO/IEC 14496-10:2005:
Information technology – Coding of audio-visual objects
Part 10: Advanced Video Coding.

Mathematical algorithms of an indoor ultrasonic localisation system

ZSOLT PARISEK, ZOLTÁN RUZSA, GÉZA GORDOS

Bay Zoltán Foundation for Applied Research, Institute for Applied Telecommunication Technologies
{parisek, ruzsaz}@ikti.hu

Keywords: ultrasound, localisation, motion-tracking, trilateration, self-configuring

Our indoor ultrasonic localization and motion-tracking system (BATSY) is based on a mobile node capable of emitting radio and ultrasonic signals and a number of ultrasound sensors mounted at known positions in a room. The system uses the node's distance to the sensors (as derived from the arrival times of the ultrasound signal) for calculating its position.
In this paper we discuss the mathematical and measurement problems related to ultrasonic localization, we propose a possible solution algorithm, and we present a method for determining the sensors' position in an automated way.

1. Introduction

As part of the AAL (Ambient Assisted Living) programme, we have developed an indoor ultrasonic localisation system at BAY-IKTI (Institute for Applied Telecommunication Technologies of the Bay Zoltán Foundation).

Such a system can be used in many fields, e.g. for monitoring the daily routine of injured or elderly people [1], or for tracking the movement of customers in a supermarket in order to observe and analyze their shopping habits. A similar approach has been applied to the problem in a number of research projects worldwide, see "The Bat Ultrasonic Location System" developed at Cambridge University [2], or the "Ultrasonic Localisation System" developed at HomeLab, Lucerne University [3].

A common shortcoming of such localisation systems based on ultrasonic distance measurement is, however, that in order to achieve adequate accuracy (in the 3 to 15 centimetres range), one needs to map the exact position of the sensors with much higher precision (in the sub centimetre range). The process for this kind of high-precision positioning, which must be performed prior to the first use of the system, can be technically demanding (it is usually done either using a conventional tape measure or with a laser range finder), hence adding substantially to the installation time of the system.

In contrast with the above difficulties, the method we have developed allows the sensors' positions to be determined quasi-automatically (i.e. without any preliminary positioning, either manual or instrumental) through the accurate measurement capacity of our localisation system.

The rest of the paper is organized as follows. In Section 2 we introduce our BATSY system, in the next section we discuss how to determine the position of a point in space with its distance given from a number of known points, in Section 3 we review and correct some of the positioning errors resulting from possible distortions in the ultrasonic distance measurement process, and finally, in Section 4 we propose a method to obtain the sensors' positions in an automated way.

2. Description of the BATSY system

We have started to build our ultrasonic localisation system in 2006 to develop a localisation and motion tracking tool for our AAL (Ambient Assisted Living) laboratory. Since the system relies on distance measurement based on the speed of an ultrasound signal, it has been named BATSY (BAT SYstem). As a result of various improvements made to the calculation method, we have managed to reach an accuracy of 3 centimetres.

Figure 1. The Batsy mobile node

The BATSY ultrasonic localisation system consists of three main components:

1. A number (6-8) of sensor units mounted on walls, for receiving ultrasound signals.
2. A computer equipped with a radio module, for receiving radio signals and performing calculations.
3. A mobile node, capable of emitting radio and ultrasound signals simultaneously *(Figure 1)*.

The underlying concept of the system is based on the difference between the speed of sound and that of a radio signal. A radio sig-

nal emitted by the mobile node arrives to the computer almost instantly, while the 40 kHz ultrasound wave dispatched at the same time travels at the speed of sound, thus it reaches the sensors mounted on the walls significantly later. Therefore, through measuring the difference between the arrival times of the radio and the ultrasound signals to some given sensor, one can calculate the distance between the mobile node and the sensor. Assuming we know the exact location of the sensors as well as the distances between them and the mobile node, we can calculate the position of the mobile node.

3. Trilateration: calculating the position of a point in space with its distance given from three other points

Assume we have three known points in space with coordinates $S_1=(x_1,y_1,z_1)$, $S_2=(x_2,y_2,z_2)$, $S_3=(x_3,y_3,z_3)$ respectively, and we further know their distances, d_1,d_2,d_3 from some unknown point (x,y,z). This unknown point is located at the intersection of three spheres with S_1,S_2, S_3 as their centres and d_1,d_2,d_3 as their radii, respectively. This method is based on the measurement of three distances, so we call it *trilateration*. Three spheres intersect in two points generally, and these two points of intersection are symmetrical with respect to the plane through S_1,S_2,S_3. The two intersection points are obtained as the solution to the following system of equations:

$$(x_1-x)^2+(y_1-y)^2+(z_1-z)^2=d_1^2$$
$$(x_2-x)^2+(y_2-y)^2+(z_2-z)^2=d_2^2$$
$$(x_3-x)^2+(y_3-y)^2+(z_3-z)^2=d_3^2$$

In the course of the actual calculations, differences like $(x_1-x_2),(z_2-z_3),\ldots$, appear in the denominator, which causes a problem in case one of them is zero. In practice this happens quite often, since the sensors are usually mounted on walls and ceilings, thus some of their coordinates *x, y, z* are likely to coincide. The easiest way to avoid this problem is through adding small, independently chosen random numbers (in the range of a thousandth millimetre) to the coordinates. This practically does not decrease the accuracy of the calculations, while it avoids division by zero with a probability high enough.

In particular adaptations there is often an obvious opportunity for selecting the correct solution (i.e. the one corresponding to the actual location of the mobile node) out of the two candidates. For example, in case the three sensors are located on the ceiling of a room, the two solutions will fall on opposite sides of the ceiling, so it is straightforward to choose the one inside the room. When there is no chance for a solution of this kind, one can use the algorithm discussed in Section 4.1.

4. Correcting positioning errors resulting from distortions in distance measurement

In practical use, one has to apply different modifications to the theoretical algorithm described in Section 3.

4.1 Dealing with distortions in distance measurement

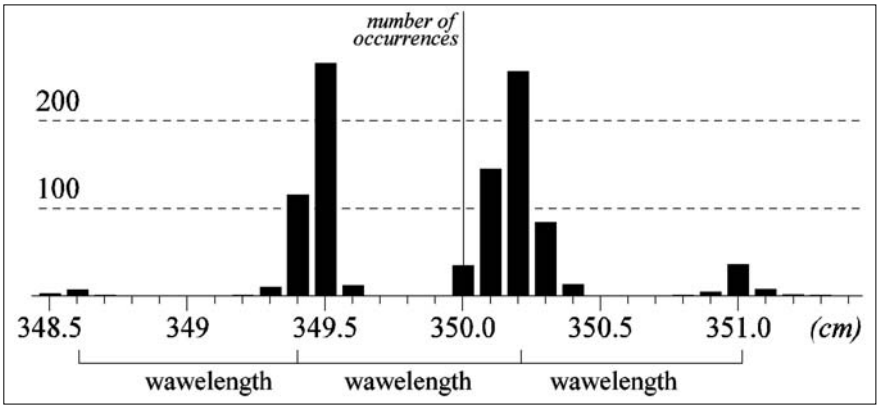
Sometimes, depending on the position of the mobile node, an obstacle along the straight line connecting the mobile node to a sensor may get in the way of the ultrasound wave. In that case, sound cannot reach the sensor along a direct path, hence distance measurement fails. This would not be of any problem in case no result was obtained from such occurrences of measurement; but what usually happens is that the sound wave reaches the sensor along a longer, indirect path. Thus the sensor detects a reflected signal, resulting in an estimated distance greater than the real one.

In order to avoid a possible localisation error caused by corrupted distance values, one should use a greater number (6-8) of sensors. This way, the estimated location of the mobile node would be obtained as the intersection of more than three spheres. However, there could still be errors made in the localisation of the centres and the measurement of the radii of the spheres (even additional to the previously mentioned ones), in which case more than three spheres may have no intersection at all.

Both problems can be dealt with simultaneously using an “Adaptive Fuzzy Clustering” algorithm [4].

The algorithm consists of two phases. The first phase involves calculating the intersection of all sphere-triplets derived from the measured distances, using the trilateration algorithm described in Section 3. If we have measured n distances, this will give $\binom{n}{3}$ possible locations as a set of points in space. Assuming the number of more or less accurate measurements is sufficiently high, the set of intersection points belonging to spheres with a correctly measured radius should be concentrated within a relatively small space segment, whereas the intersection points derived from one or more erroneously determined spheres will be dispersed in space essentially randomly.

Figure 2. Occurrences of different measured distances from 1000 measures at a distance of 350 cm



The second phase consists of finding the point with the maximum density within the above derived set of possible locations. This is done by calculating the vector average of the points, or the "centre of gravity" of the set. Next, we calculate the average distance of all points from this centre of gravity, and eliminate those whose distance is greater than average. In this way, we will have obtained a smaller set with which to repeat the second phase. We do the repetition until the diameter of the set falls below a required threshold value.

4.2 Theoretical and practical accuracy of distance measurement

The BATSY system uses 40 kHz ultrasound; its wavelength is about 8.6 mm in room temperature. The sensor we used detects the pressure peaks of air waves, so the arrival of an ultrasound packet will presumably be recorded at a pressure peak. These peaks lie 8.6 millimetres apart from each other, which adds an uncertainty factor of 8.6 mm to distance measurement.

The sensor's sampling frequency – in accordance with the clock pulse of its microcontroller, and taking as given the speed of sound – translates into a sub-millimetric measurement accuracy.

In *Figure 2* we can see the results from 1000 individual measurements performed at a given distance. One can clearly see the occurrence peaks situated 8.6 millimetres away from each other.

4.3 Distortions related to the directional angles

The sensors and emitters are directed in space. This means that the distance between the sensor and the mobile node is measured correctly as long as one is positioned facing the other, but when one or both of them are rotated, the measured distance increases with the angle of their axes. This results in a systematic bias related to the sensor's and the mobile node's directional angles *(Figure 3)*.

Figure 3. The sensor's (α) and node's (β) directional angles

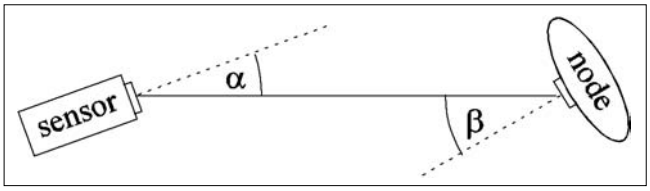
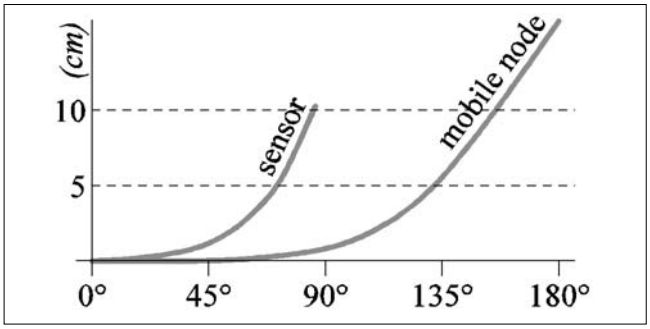


Figure 4.
Bias in distance measurement as a function of the sensor's and node's directional angle, at a distance of 350 cm



To offset the bias related to the sensor's directional angle, we have to measure it as a function of the angle and distance variables *(Figure 4)*.

In order to eliminate the distortion from the measured distance, one needs to know the approximate distance of the mobile node as well as the angle between the sensor-node line and the sensor's axis.

If the direction of the sensor's axis is known beforehand (i.e. it was recorded at the moment of the sensor being mounted), the bias related to the sensor's directional angle can be offset using the following two-step algorithm.

In the first step we determine the location of the mobile node using the method discussed in Section 3. This location will be inaccurate as yet, since it will contain a directional bias, but the deviation from the real position is not significant. So this inaccurate estimated location is suitable for calculating the angle between the sensor-node line and the sensor's axis.

In the second step we calculate the range correction value for the given (angle, distance) pair and subtract it from the previously measured distance. Do it for all the sensors, and recalculate the location of the mobile node using the method specified in Section 3. This new location will be free of any directional angle effect.

We have tested the method in our laboratory and found that the distance between the positions calculated with and without the sensor's directional angle correction is usually less than 20 millimetres. So if such accuracy is not required or the system is low on CPU performance, it can be omitted.

Correcting the bias related to the mobile node's directional angle is only feasible if the direction of the node's ultrasound emitter can be determined. In this case, one can use the same method as the one discussed above. In our specific application however, the mobile node's direction was not fixed, so we could only measure the localisation error related to the sensor's directional angle. Our chosen approach then was to place the mobile node to a pre-specified location and rotate it around while simultaneously calculating its indirectly estimated position through the above described algorithm. We have found the difference between the real and the calculated positions of the node to fall in the 0 to 45 mm range.

4.4 Distortions resulting from variations in the speed of sound

The speed of sound varies with temperature, humidity and air pressure, and so does the outcome of any distance measurement procedure relying on sound waves [5]. Whereas the effects of humidity and pressure are negligible from our point of view, a 1°C change in air temperature near the 20°C range causes a substantial, 0.176% change in speed. The consequence of this for our BATSY system is that a sensor will measure an erroneous distance d instead of the real distance d/q, where q is some quotient depending on temperature.

Assuming we have distances measured by four sensors, the value of q can in theory be obtained by solving the following system of equations:

$$(x_1-x)^2+(y_1-y)^2+(z_1-z)^2=q^2d_1^2$$
$$(x_2-x)^2+(y_2-y)^2+(z_2-z)^2=q^2d_2^2$$
$$(x_3-x)^2+(y_3-y)^2+(z_3-z)^2=q^2d_3^2$$
$$(x_4-x)^2+(y_4-y)^2+(z_4-z)^2=q^2d_4^2$$

However, we have found that imprecision in the sensors' coordinates and other inaccuracies prevent this idea from being put into practice. So if our goal is to set up a system operating in room temperature, and unless we have temperature data from other sources (e.g. from an internal thermometer), it is better not to use this kind of temperature correction method altogether.

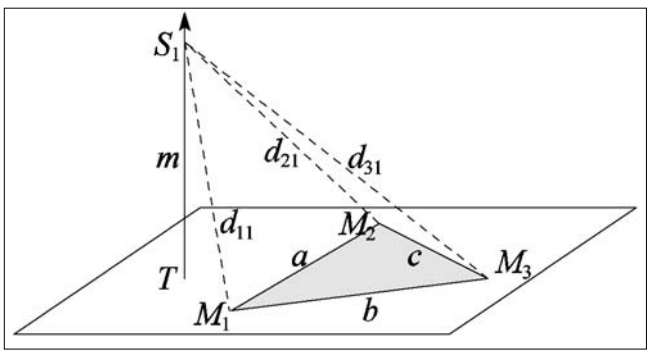
5. Determining the sensor's coordinates in an automated way

In order for the positioning system to yield an accurate result, the sensors' coordinates have to be measured with high precision. In a real-life deployment of the system, determining the sensors' exact position is the hardest and most time-consuming task. Using a traditional tape-measure or a laser rangefinder takes hours, and it leads to errors in the centimetres range. Is it possible to use the system itself for locating the sensors?

Obviously, it is impossible to have all the coordinates determined by the system, since the origin of the coordinate-system and the direction of the axes need to be chosen in some arbitrary way. (For convenience, we have chosen the vertical direction as the third axis of the coordinate system.) Thus our goal is to come to a self-configuring algorithm involving as little technical difficulty as possible and capable of determining the sensors' coordinates in some chosen Cartesian coordinate system. The algorithm we are about to discuss requires the following input data (these must be determined manually):

1. Coordinates of an arbitrary sensor S_1.
2. One of the non-vertical coordinates of some other sensor S_2.
3. The side lengths and the orientation of a triangle arbitrarily drawn on some horizontal plane.

Figure 5. The tetrahedron $M_1M_2M_3S_1$



5.1 Description of the algorithm

Denote the mounted sensors by $S_1,S_2,...S_n$, and their coordinates by (x_1,y_1,z_1), (x_2,y_2,z_2), ... (x_n,y_n,z_n), respectively. Let M_1,M_2,M_3 be the vertices of our horizontal triangle, and let a,b,c denote the distances between them *(Figure 5)*. According to our initial assumption, x_1,y_1,z_1,x_2 (or y_2), a,b,c are known, and our goal is to determine x_i,y_i,z_i ($i=1,...n$). Without loss of generality, one can impose $x_1=0$, $y_1=0$, $z_1=0$.

Step 1:
Measurement

Place the mobile unit at location M_1, and let the system measure its distance from the sensors. (Let $d_{11},d_{12},...d_{1n}$ denote these n distances.) Repeat the process with M_2 and M_3 so to obtain all the distances d_{ij} ($i=1,2,3$; $j=1,...n$).

Step 2:
Calculating the distance between S_1 and the plane $M_1M_2M_3$

First, calculate the volume of a tetrahedron with M_1, M_2,M_3,S_1 as its vertices. This can be done in two different ways. On one hand, according to Tartaglia's formula [6], we have

$$V^2=\frac{1}{288}\det\begin{bmatrix}0 & d_{11}^2 & d_{21}^2 & d_{32}^2 & 1\\ d_{11}^2 & 0 & a^2 & b^2 & 1\\ d_{21}^2 & a^2 & 0 & c^2 & 1\\ d_{31}^2 & b^2 & c^2 & 0 & 1\\ 1 & 1 & 1 & 1 & 0\end{bmatrix}$$

On the other hand (using the notation $s=(a+b+c)/2$), the area of the base triangle is, by Heron's formula [6], as follows $T=\sqrt{s(s-a)(s-b)(s-c)}$

Then, we have $V=Tm/3$, where m is the height of our tetrahedron as measured from the base triangle $M_1M_2M_3$. Making use of the equivalence of these two expressions for V, one can easily calculate m, which is precisely the distance between S_1 and the plane $M_1M_2M_3$.

Step 3:
Calculating the relative positions of M_1,M_2,M_3

Denote by T the foot of the tetrahedron's altitude line connecting S_1 to the base triangle $M_1M_2M_3$ (i.e. T is the orthogonal projection of S_1 to the plane $M_1M_2M_3$). Let r_1,r_2,r_3 be the distances of M_1,M_2,M_3 from T *(Figure 6 – on the next page)*. From the Pythagorean theorem, we have

$$r_1=\sqrt{d_{11}^2-m^2},\quad r_2=\sqrt{d_{21}^2-m^2},\quad r_3=\sqrt{d_{31}^2-m^2}.$$

Now fix a two-dimensional Cartesian coordinate system on the plane $M_1M_2M_3$, with T as its origin and one of its axes going through M_1. As r_1 denotes the distance $\overline{TM_1}$, point M_1 has coordinates $(0,r_1)$ in the afore mentioned system.

Similarly, r_2 denoting the distance $\overline{TM_2}$ and a denoting the distance $\overline{M_2M_1}$, coordinates (x,y) of point M_2 are obtained as the solution to the following system of equations

$$x^2+y^2=r_2^2$$
$$x^2+(y-r_1)^2=a^2$$

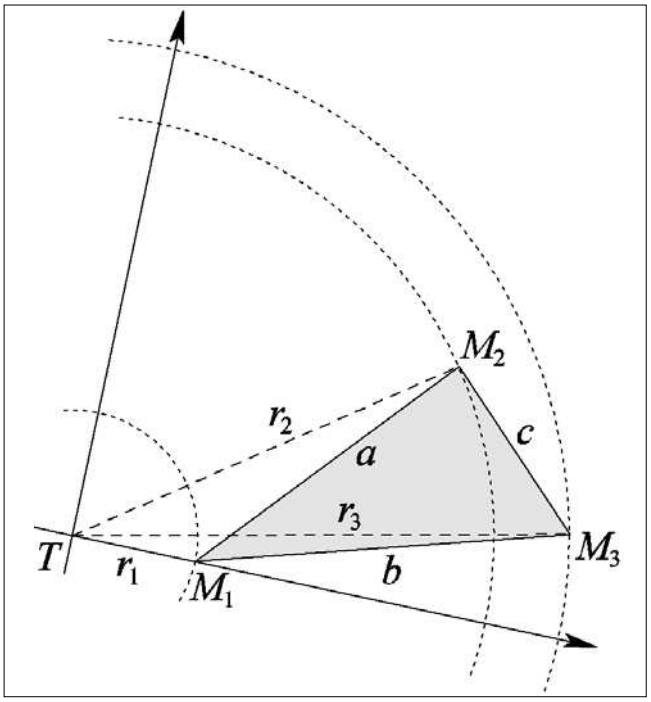


Figure 6. Orthogonal projection to the plane $M_1M_2M_3$

The above equation system has two solutions, among which the correct one is to be chosen in accordance with the orientation of triangle $M_1M_2M_3$.

The same method can be used for determining the coordinates of M_2. However, in this case the correct one out of the two solutions is to be selected imposing the further restriction that the distance $\overline{M_2M_3}$ must be equal with c.

The origin of this particular coordinate system is obtained through an orthogonal projection of the original system to plane $M_1M_2M_3$, yet the directions of their axes are different. Thus, we have so far determined the positions of M_1, M_2, M_3 relative to the new coordinate system, which we further need to rotate around the axis S_1T in order to get their coordinates in the original one.

Step 4:
Determining the angle of rotation around the axis S_1T

Let us return to the three-dimensional space. The two-dimensional relative coordinates of M_1, M_2, M_3 (as determined in Step 3) need to be complemented with a third one, which can be expressed as the negative distance between plane $M_1M_2M_3$ and point S_1, that is, $(-m)$.

Starting from these coordinates, and making use of the distances d_{12}, d_{22}, d_{32}, the coordinates of S_2 in the rotated system can be calculated through the trilateration procedure discussed in Section 2. Its two candidate solutions being symmetrical with respect to plane the $M_1M_2M_3$, we need to choose the one which lies on the same side of the plane as where the sensor is actually located.

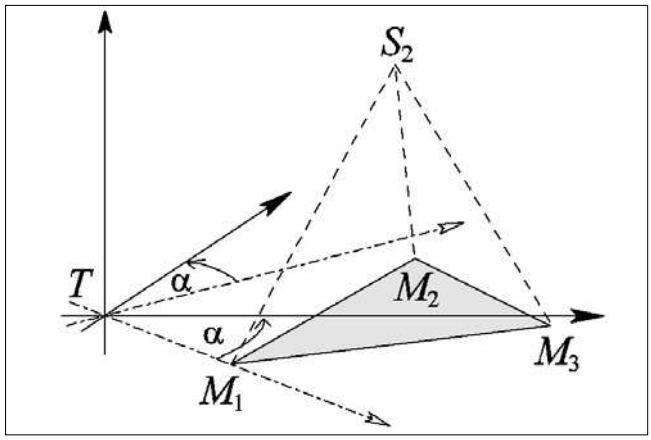
Denote by (x'_2, y'_2, z'_2) the coordinates of S_2 in the rotated system. Its coordinates in the original system are (x_2, y_2, z_2) where only x_2 is known for the present. Following from the properties of the rotated coordinate system, we have $z'_2 = z_2$.

If we transliterate the triplet (x'_2, y'_2, z'_2) to a different coordinate system derived from the original one through a rotation by α around the vertical axis *(Fig. 7)*, the new coordinates will be $(x'_2\cos\alpha - y'_2\sin\alpha,\ x'_2\sin\alpha + y'_2\cos\alpha,\ z'_2)$. From the equality of the first coordinates in the two systems, one easily comes to the trigonometrical equation

$$x_2 = x'_2\cos\alpha - y'_2\sin\alpha$$

This equation has two solutions for α, thus yielding two candidates for S_2, which lie symmetrically with respect to the plane parallel to axes (z,x) and containing S_1. Again, we have to select the correct α, the one corresponding to the particular candidate for S_2 which is located closer to its real position.

Figure 7.
Rotation from the temporary coordinate system to the original one



Step 5:
Calculating the coordinates of M_1, M_2, M_3 in the original coordinate system

In order to come to the absolute coordinates of M_1, M_2, M_3, their relative coordinates (as calculated in Step 3) need to be multiplied by the matrix

$$\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where α is the one determined in Step 4.

Step 6:
Obtaining the coordinates of sensors $S_2, \ldots S_n$

From Step 5 we know the positions of M_1, M_2, M_3, along with their distances from S_i: d_{1i}, d_{2i}, d_{3i}. Thus the coordinates of $S_2, \ldots S_n$ can be determined using the trilateration algorithm discussed in Section 3.

5.2 Practical considerations

Since the system's overall precision depends highly on the accuracy of the sensors' coordinates, it is essential to reduce any distortions relative to the process as much as possible.

1. The impact of temperature on the speed of sound can be offset using a reference measure taken at the beginning of the process (calibration). Place the mobile

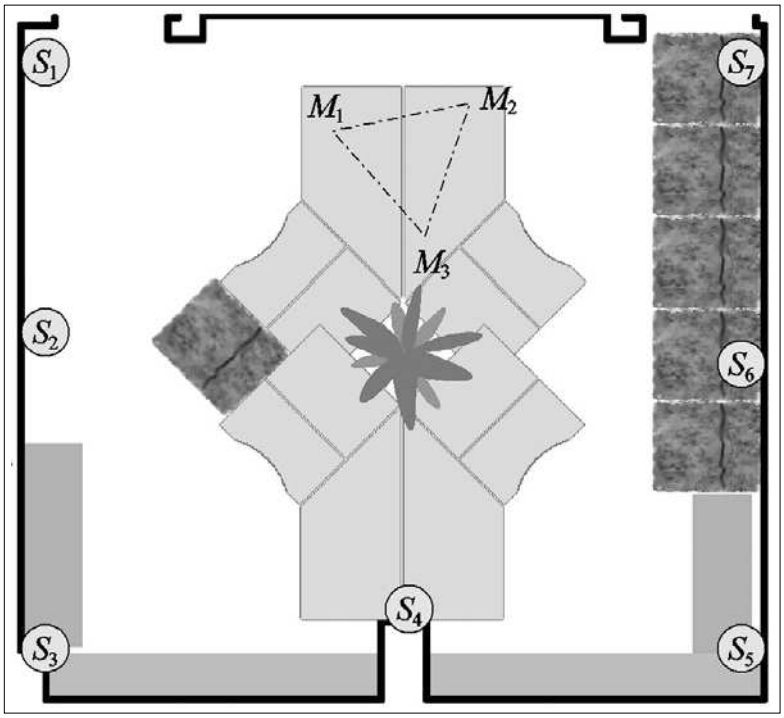


Figure 8. BATSY test configuration

node to a known benchmark distance from any particular sensor, and let the system measure its distance. Then, through multiplying all measured distances by the quotient of the afore measured and the real (benchmark) distances, the distorting effect of temperature is eliminated.

2. The inaccuracy related to the sensors' directional angle can be reduced through directing the sensors towards the triangle $M_1M_2M_3$.

3. The bias related to the mobile node's directional angle can be reduced using the method discussed in Section 4, assuming the ultrasound emitters are directed vertically while localising M_1, M_2, M_3.

4. The imprecision relative to the wavelength of the 40 kHz ultrasound (as discussed in Section 4.2) can practically be eliminated through taking the average of several measured distances.

5. There might be sensors whose positions cannot be determined by algorithm 4.1 (e.g. if points M_1, M_2, M_3 happen to be out of sight of a particular sensor). In this case, the already localised sensors can be used for determining the positions of some additional points M_4, M_5, M_6 and then for localising further sensors through Step 6 of the algorithm.

5.3 Test results

The algorithm described in section 5.1 and further amended in line with the considerations discussed in Section 5.2 was tested in a room 5x6 metres in area and 2.8 metres in height. Points M_1, M_2, M_3 were fixed on the surface of a 72 cm high desk, each of them situated 1 meter away from any other. We had put all three coordinates of sensor S_1 along with coordinate y of sensor S_7 into the system, and tried to determine the positions of sensors S_2,...S_7 *(Figure 8)*.

Sensor S_3 had no sight of view to points M_i so the algorithm couldn't estimate its position. A plant was obstructing the path between S_4 and points M_i, so we expected incorrect values for its coordinates. The results are shown in *Table 1*. One can see the accuracy of the method is good enough to determine the sensor's initial positions to use in the BATSY localisation system – if there are no obstacles along the ultrasound's path.

6. Conclusions

In this paper we presented the BATSY ultrasonic localisation system, discussed the positioning algorithm, observed the localisation errors resulting from possible distortions in distance measurement, and proposed ways of reducing them. We have provided an algorithm for configuring the system semi-automatically, and tested the results one can expect when operating the system in real-life conditions. The measurements confirmed that our configuring algorithm can be used to determine the sensors' position in ideal circumstances, however, when obstacles blocked the ultrasound's path, higher errors appeared.

Acknowledgement

We would like to express our gratitude to all those participated in the creation of the BATSY system: József Bánlaki, Gyula Bakonyi-Kiss, Martin Becker, Bence Csák, Pál Haraszti, Csaba Megyesi, Gábor Ruzsa.

Table 1. Test results

Sensor	Real coordinates *(cm)*			Calculated coordinates *(cm)*			Error *(cm)*
S_1	32.5	24.5	280.8	32.5	24.5	280.8	0
S_2	31.2	241.9	278.0	30.3	240.2	278.6	1.98
S_3	37.2	501.1	280.6	n.a.	n.a.	n.a.	n.a.
S_4	305.9	474.1	281.8	301.3	467.0	293.5	14.44
S_5	583.3	500.2	280.5	583.8	498.2	282.6	2.97
S_6	584.0	278.7	278.6	582.1	278.3	280.4	2.59
S_7	582.7	27.2	280.3	578.8	27.2	285.6	6.64

Authors

ZSOLT PARISEK is a researcher at the Bay Zoltán Foundation for Applied Research Institute for Applied Telecommunication Technologies (BAY-IKTI). He is doing research in the field of ambient intelligence. He has graduated at the University of Debrecen, Faculty of Informatics as a software engineer. His research interests include intelligent transportation systems, algorithms and data mining.

ZOLTÁN RUZSA received his M.Sc in Mathematics from the Eötvös Loránd University of Sciences in 2001. He was an assistant professor at the Budapest University of Technology and Economics to 2007. He is currently employed as a researcher at BAY-IKTI, the Institute for Applied Telecommunication Technologies. His research interests include optimisation, graph theory and localisation algorithms used in AAL (Ambient Assisted Living) and traffic analysis.

GÉZA GORDOS received his M.Sc, Ph.D and Dr. Habil. in Telecommunications from the Budapest University of Technology and Economics (BME) in 1960, 1966 and 1994, resp. He holds D.Sc. from the Hungarian Academy of Sciences. He is full-professor with the BME serving as head of two sections in the Department of Telecommunications and Media Informatics since 1976. His previous employments include 3 years at various universities abroad (UK, USA) and 7 years in industry, among others as Chairman of the Board of the Hungarian Telecommunications Co.(1992-93). His research activities have been focusing on the technology and management of telecommunication systems and services as well as on speech processing. In 2004 he accepted the invitation from the Hungarian equivalent of NSF to establish and run as Director the Institute for Applied Telecommunication Technologies (IKTI).

References

[1] Mitilineos A. Stelios, Argyreas D. Nick, Makri T. Effie, Kyriazanos M. Dimitris, Stelios C.A. Thomopoulos, "An indoor localization platform for ambient assisted living using UWB", International Conference on Mobile Computing and Multimedia, Proceedings of the 6th Int. Conf. on Advances in Mobile Computing and Multimedia, New York, ACM, 2008, pp.178–182.

[2] Andy Ward, Alan Jones, Andy Hopper, "A New Location Technique for the Active Office", IEEE Personal Communications, Vol. 4, No.5, October 1997, pp.42–47.

[3] S. Knauth, C. Jost, M. Fercu, A. Klapproth, "Design of an Ultrasonic Localisation System with Fall Detection", IET Assisted Living 2009 Conference, 24-25 March 2009, London, UK, http://www.ceesar.ch/fileadmin/Dateien/PDF/NewsEvents/IETAL2009_talk.pdf

[4] Jingbin Zhang, Ting Yan, John A. Stankovi, Sang H. Son, "Thunder: towards practical, zero cost acoustic localization for outdoor wireless sensor networks", ACM SIGMOBILE Mobile Computing and Communications Review, Vol. 11, No.1, 2007, pp.15–28.

[5] O. Cramer, "The Variation of the Specific Heat Ratio and the Speed of Sound in Air with Temperature, Pressure, Humidity, and Concentration", Journal of the Acoustical Society of America, Vol. 93, No.5, May 1993, pp.2510–2516.

[6] György Hajós, Bevezetés a Geometriába, Budapest, Tankönyvkiadó, 1960, pp.316–319.