Molecular Processes as a Basis for Autonomous Networking

Bettina Krüger and Falko Dressler

Abstract— ***Autonomous networking has become the buzzword for attempts of building high-scalable network architectures, which are self-organizing, self-maintaining and self-healing. Few of these approaches were successful and none has shown to provide all the promised functions. We try to study the processes in computer networks using molecular processes as the paradigm. This novel approach shows many similarities between computer networking and cellular mechanisms. In this paper, we focus on the area of network security as one research area with high demand for high-scalable mechanisms providing the needed functionality. After identifying similarities between nature and technology, we discuss potential research domains, which are high potentials for learning directly from molecular biology using the example of security threats in communication networks. We see the proposed mechanism as a generic approach for autonomous networking. The countermeasures against attacks in computer networks are only a special example to introduce the mechanisms.***

Index Terms— ***Autonomous Networking, Bio-Inspired Networking, Network Modeling, Next Generation Networks***

1. INTRODUCTION

WE investigated the possibilities to utilize the infinite experience of the nature to address questions in computer science, focused on problems in networking. Among other things, this approach was motivated by a book from Nobel Laureate Prof. Manfred Eigen [7]. Unlike most activities in bioinformatics, where computerized methods are employed to study natural processes in more detail, we discovered that at least some of our nowadays problems might be easier to understand and even to be solved if we directly learn from natural mechanisms. In this paper, we focus on mechanisms known from molecular biology, which we can adopt to improve internet technology.

During the last couple of years, great progress was made to make computer networks more stable, more efficient, and more secure. Nevertheless, we also experience that there are still many open issues, especially in terms of network security. Distributed denial of service (DDoS) attacks, worms, and viruses are getting more aggressive and much harder to prevent [9]. Autonomous networking should help to solve these problems. We address these issues by applying mechanisms learned from natural processes and show potential solutions.

In simple configurations of low-speed networks, well-known mechanisms can be employed to examine all the network traffic, to filter unknown or suspicious data packets, and to program firewall rules blocking typical attacks [3, 11]. Unfortunately, these mechanisms fail in high-speed backbone networks due to restrictions in CPU capacity and free memory for processes and queues. Therefore, it is not possible to run all kinds of required monitor processes at once directly on the network components.

Our idea is to limit the number of active processes to those which are really required in the current situation, e.g. if a particular worm is being distributed, only a prevention scheme for this event is needed. The solution is provided by achievements of molecular biology. In every single cell of an

Manuscript received August 10, 2004; revised September 7, 2004.

B. Krüger is with the Medical Clinic IV, Nephrology, University of Erlangen-Nuremberg, Germany.

F. Dressler (contact) is with the Department for Computer Sciences, University of Erlangen-Nuremberg, Germany (e-mail: bettina.krueger@fd42.de; dressler@informatik.uni-erlangen.de).

organism, the program for all reactions and mechanism is coded in the DNA and is activated exactly in the situation when it is needed. Every signal from the extracellular environment is recognized specifically and results in a specific signal transduction cascade. Signal transduction normally initiates the translation of the necessary genes which finally leads to a cellular answer, e.g. to maintain the cellular function, but also to react to pathological situations, e.g. any kind of inflammation, etc.

The same procedure seems to be adequate for reactions in computer networks. If we are able to describe solutions for typical problems in form of processes and their course of actions, we are able to build flexible network components with low resource requirements and a high efficiency in terms of network operations. Fortunately, we can employ methods from software engineering for such descriptions. Monitor components act as the receptors and try to identify the behavior of the network. Due to the availability of sampling methodologies and statistical methods, such monitoring can be done even in high-speed network environments [6, 13]. We already addressed similar scalability issues in other areas such as network measurements [5].

Based on the recognized behavior of the network, new processes can be created and activated using a description of the available mechanisms. Therefore, only these processes exist in memory and are activated which are required in the current situation.

We see the proposed mechanisms as a generic approach for autonomous networking. The detection of security threats and the activation of countermeasures are only special examples to introduce the mechanisms.

In this paper, we show the similarities between the signaling mechanisms in cellular systems and networking entities responsible for packet forwarding, intrusion detection, and firewalling.

The rest of the paper is organized as follows: in section 2, the cellular mechanisms are described followed by a description of a general networking architecture in section 3. Possible research issues resulting from an analysis of the similarities between both the cellular systems and computer networks are discussed in section 4. A section describing related work and some conclusions complete this paper.

2. *Cellular Mechanisms*

All organisms share one common information system which is mostly the DNA and which codes for the organization of the whole organism. This organization is a highly regulated process from the single cell up to complex organs of the body. The hierarchy in the organism is very high. Every process, e.g. movement, metabolism, communication, etc. is organized by interactions of several organs. Organs represent an assembly of one or more tissues, which fulfill a common function. One tissue is build by different cell types. One cell type consists of identical cells, which are associated and communicate with each other to fulfill a common function within the tissue. Single cells communicate with each other as well as with cells of other tissues by sending signals to which the target cells respond by specific gene expression. In this way, a signal can be carried out and influences the function of higher units such as the organs.

In this paper we want to focus on mechanisms how cells interact with their environment in general and, secondly, on intracellular processes which are initiated by extracellular signals and result in the specific cellular response out of the pool of information given by the DNA.

D. Cellular Signaling

The functionality of an eukaryotic cell relies on the complex network of biochemical processes. Within these processes, single reactions take place in a coordinated fashion. They can take place simultaneously and successively. Thus, these processes must be highly regulated and controlled. This also means that these mechanisms are very specific for the given result.

The main goals of cellular processes are to regulate the intracellular metabolism and to communicate with their environment.

Physical or chemical attractions from the environment are signals for the cell to change intracellular processes. Chemical attractions can be low-molecular metabolites, hormones, or ions. They can be sent by other cells of the same tissue or by cells from other tissues. Furthermore, physical parameters such as heat, pressure, or electrical signals can induce cellular reactions.

E. Intercellular Signaling

Communication between cells can occur by different processes. First, cells can release soluble molecules such as hormones etc. that are transported via the blood (long distance in the organism, e.g. hormones) to the target cell. Other soluble factors are released into the extracellular space to reach the neighboring cells in a short distance (Fig 1 A).

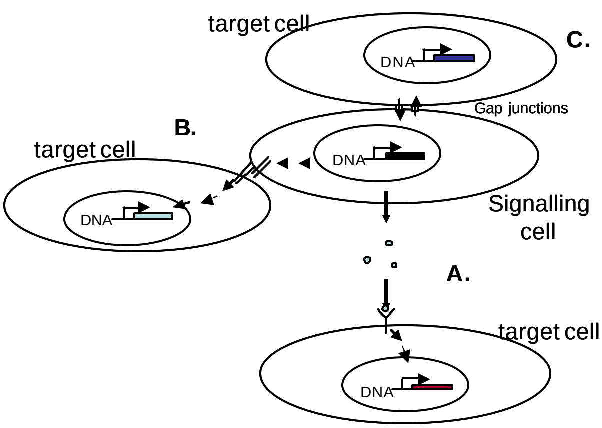

Fig 1. Intercellular communication. One cell can communicate with the neighboring cells via A. soluble factors, B. receptor interactions, C. direct contact, e.g. "gap junctions". The target cells react on the current stimulus by gene transcription.

These molecules are recognized by the target cells and induce a specific biochemical answer. By the release of soluble factors many cells can be activated simultaneously which results in a coordinated reaction of the organ or organism.

Secondly, cells can also communicate via cell surface molecules. In this process, a surface molecule of one cell or even a soluble molecule, which is released by one cell, directly binds to a specific receptor molecule on another cell (Fig 1 B). Thirdly, communication between cells occurs via direct connections between two neighboring cells, which allow a direct exchange of metabolites (Fig 1 C).

F. Intracellular Communication

In either case, the signal from the extracellular source is transferred through the cell membrane. Inside of the target cell, complex signaling cascades are involved in the information transfer (signal transduction), which finally result in gene expression or an alteration in enzyme activity and, therefore, define the cellular response.

Because of the great variety of signal transduction pathways, only one example for receptor-mediated signal transduction is presented here which might have a great relevance for computer networks: The MAPK signaling pathway is a major pathway in eukaryotic cells, which is activated by different types of receptors, e.g. receptor-tyrosine-kinases or G-protein-coupled receptors [15].

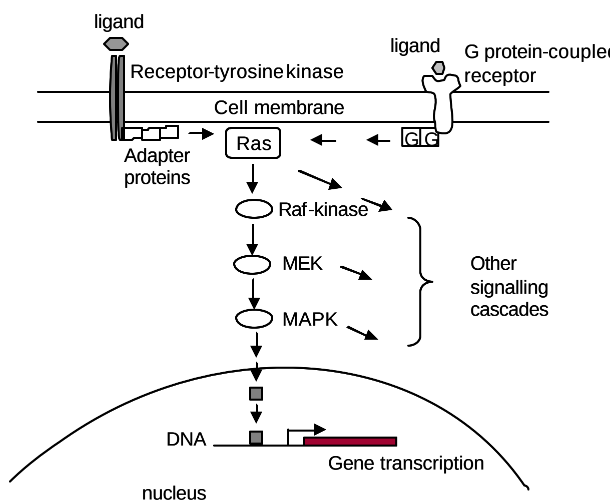

Fig 2. Intracellular signal transduction, shown on the example of MAPKinase pathway. Different ligands can activate different receptor types which result in MAP kinase activation.

Upon binding of the signaling factor (ligand), the receptor is activated. This activation is mostly attended by a phosphorylation or conformational change of the receptor, which make it possible for adapter proteins inside of the cell to bind to the receptor. Similarly, these adapter proteins are activated.

The signal is carried on to a signaling molecule called Ras. As shown in Fig 2, the activation of different types of receptors comes together in the activation of Ras. The signal is carried on by protein kinases. Protein kinases phosphorylate other proteins. Phosphorylated protein kinases are able to conver the phosphorylation to the next kinase. Finally, a transcription factor is activated by this signaling cascade which move into the cell nucleus and bind to the DNA.

The binding of a specific transcription factor to a specific binding site on the DNA result in gene transcription, which finally induce a specific cellular response.This example shows only one straight-forward signaling pathway. For example, the same protein kinases Ras, MEK and MAPK are also involved in other signaling cascades. Thus, signaling cascades are often highly networked, but at least result in a very specific gene transcription and, therefore, result in a very specific cellular response [12].

3. Network Architecture

We describe the relevant elements of general network architectures starting at a high abstraction level, the structure of the global Internet, and dig deeper until we reach the internal operation of a single networking node. The focus is on the characterization of the information and data paths in the network concerning packet forwarding, monitoring, and firewalling as well as management functions including intrusion detection mechanisms.

G. Internetworking Structure

From a high-level point of view, an Internet consists of a multitude of individual networks. Each of theses networks, which are called domains, hides its internal structure from the outside world. Such a scenario is shown in Fig 3. In reality, all these domains are managed from different service providers. Thus, the interaction between the domains is limited, due to unequal configurations and mistrust.

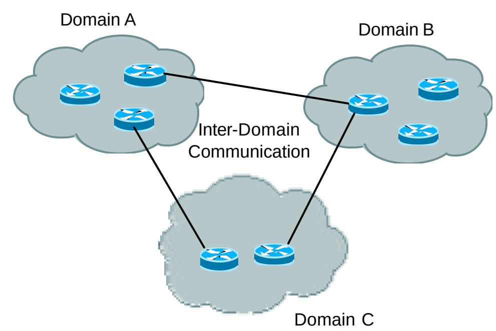

Fig 3. Internetworking Structure. Large internets consist of multiple network domains

Looking at security issues, this is a serious problem. Denial of service attacks can only be detected if knowledge about the network traffic in multiple domains is available. Inter-domain signaling and interaction is required, but there are no or, at least, few of such mechanisms developed or even deployed. The interaction between the domains is currently restricted to the exchange of routing information.

H. Intra-Domain Mechanisms

If we look a little closer and examine the components of single domains (still with the focus on network security), there are, among other entities, routers, monitors, firewalls, and intrusion detection systems. These entities and their data and signaling paths are shown in Fig 4.

It is important to understand that this is only a logical point of view. In reality, some or even all of these functions can be implemented in a single box.

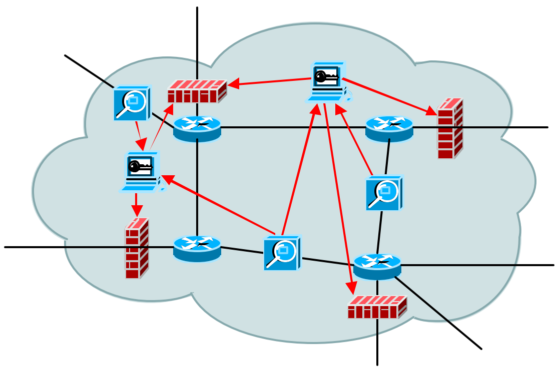

Fig 4. Communication in a single network. Typical entities are routers (packet forwarding), monitoring probes, firewalls, and controlling IDS systems

The routers are responsible for the raw packet forwarding. Sometimes, statistical information is sent to the management or IDS systems. Network monitors are employed to analyze the traffic in the network in order to detect suspicious data flows or unusual network behavior. Based on such information, examinations that are more precise can be initiated, e.g. based on attack signatures, and countermeasures become possible. The intrusion detection systems signal signatures of violent hosts to the firewall systems. At this place, rules are installed which prevent any further attacks from the identified intruders.

Today, there are some intrusion detection systems available, a few of them as open source software. Unfortunately, they do not work in highly distributed configurations and their interaction is limited. Research needs to be done to improve these mechanisms, especially the interoperation between different kinds of networking components.

I. Single Node

In the last step, we zoom into a single network node and analyze the primary components in order to compare them later with corresponding parts in cellular environments. Fig 5 provides a schematic overview. Shown are only those parts, which we think are directly improvable by studying cellular mechanisms.

First, each network node has interfaces connecting it to other network nodes. Logically, each interface has two sub-interfaces. One for inter-node communication and one for the raw data transfer. Internally, each interface consists of input and output queues with fixed, but configurable size and behavior, e.g. the selected queuing algorithm.

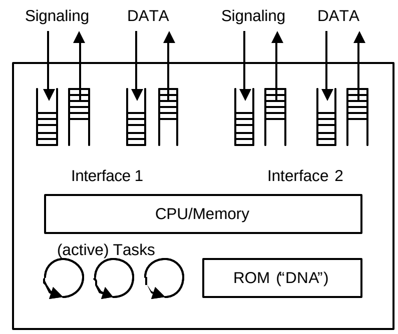

Fig 5. Schematic overview of the resources in a single network node those are required for processing and signaling.

Secondly, limited resources such as CPU capacity and memory are available for sharing among numerous tasks. Some handle the routing tables and algorithms, others coordinate the packet forwarding and still others perform security checks and inter-node communication. Tasks "exist" per default, i.e. they are started when the system is booted. Other tasks can be created on demand allowing a high flexibility.

In terms of network security, it would be desirable to have tasks searching for security threats while at the same time forwarding data packets at very high data rates. Typically, this is not possible because the main resources such as CPU capacity and memory are very limited.

Therefore, the primary question is which task to run at which time. Furthermore, not all tasks and algorithms can exist in memory waiting to get started. Other solutions are required when and how new tasks are created from saved high-level descriptions of their composition and behavior.

4. *Similarity Analysis and Research Issues*

We discuss potential research issues in this section by analyzing similarities in both worlds.

J. Composition of the Components

A first obvious result is the cognition that cellular systems consist of similar units (single cells), which react on changes in the environment to survive, and transfer information to similar units, respectively. They are directly comparable with network nodes.

Furthermore, the interaction of different cell types in a tissue is comparable to networking domains. If we zoom into both systems, the same similarities appear as tissues are made of different cells, each focused on its very specific purpose. Network nodes, on the other hand, are also very specific for their particular tasks such as packet forwarding or network monitoring.

Currently, we only assume the possibility that some of the cellular information transfers might be useful and can be copied to advance the specificity of operations in network technology.

In computer science, great effort was made in simulating mechanisms based on constructive building blocks. We intend to create object-oriented programming modules representing network nodes, their internal behavior, and their external communication mechanisms. This is the basis for representative simulations of the new mechanisms. The novel approach is to adopt and to transfer mechanisms how organs, tissues, and single cells communicate to computer networks to enhance their specificity and velocity of responses in every situation.

K. External Signaling Pathways

The primary difference in the external signaling pathways between communication networks and cellular systems is the presence of data traffic in computer networks sharing the same infrastructure that is used for inter-node communication. Besides this fact, the concepts of inter-node communication and cellular signaling are similar. A typical problem in communication networks is the scalability of mechanisms, e.g. routing or configuration tasks.

We believe that there are high potentials in examining the mechanisms for signal transduction in cellular systems and applying them to communication systems.

The most impressing issues are given by the autonomic behavior of individual cells by reacting to signals from the direct environment on the one hand, and refer the answer to other tissues or “other networks”. One example is given here: During inflammation, the cells of the affected tissue send signals to recruit cells of the immune system from the blood which means that a local problem in the body is referred to those cells which are able to initiate a response, namely the cells of the immune system. The signals that are sent by the cells of the inflamed tissue reach the target cells, the blood vessel cells (endothelial cells) outside of the cell. The signal which is mostly a small protein (in this specific example it is called IL-1ß), very specifically recognizes one receptor on the surface of the target cell. This recognition results in the gene transcription and translation of a specific protein, called E-selectin. Furthermore, this E-selectin protein is one of the first proteins that are involved in the recruitment of cells of the immune system in the blood via the endothelial cells which are found on the luminal site of all blood vessels and represent the connection between the blood and the surrounding tissue on their way to the center of the current inflammation. This example shows a two step process in which a specific signal is sent from the signaling cell (in the inflamed tissue) to a first target cell (the endothelial cells) in the environment. Secondly, these target cells recognize the problem and are able find an answer to recruit help from far away (the cells of the immune system).

Similar mechanisms help to solve problems in communication networks. An intrusion detection system is working all the time on analyzing network traffic. If suspicious traffic is detected, the system might, as usual today, contact its management system or an attached firewall to enforce countermeasures. This methodology has many drawbacks, e.g. the necessity of the possible and undisturbed communication between the IDS and the management/ firewall node and the dependency on the proper function and availability of the management system.

Looking at network security, one can

imagine a similar two step information transfer as described for the biological systems.

L. Internal Behavior of Individual Components

In this section, we focus on one single cell. This cell receives information from outside and processes this information inside and gene activation is the answer. But gene activation as a result of the intracellular signaling cascade is often influenced by the presence of intracellular inhibitor and effector molecules which are, in turn, regulated by other control mechanisms in the cell [2, 8]. These control mechanisms appropriate the specificity of the gene translation and, therefore, the cellular answer. This means that a signal transduction cascade is often not a straight-line cascade, but can be networked with other signaling cascades and all these processes succeed highly coordinated and regulated. A lot of relationships between signaling cascades are known already and it turned out to be the importance of timing and combination how they work together to reach a very specific answer.

The same autonomic behavior applied to network nodes would help us making communication networks more efficient. There are a number of topics to discuss. Much pressure is on high scalable and effective intrusion detection with automatically involved countermeasures. We are working on such mechanisms allowing monitoring, analyzing, and processing of high-speed traffic. Processing power of individual nodes is first allocated for the supervision of ongoing attacks and the effectiveness of taken countermeasures. Then, the remaining capacities are directed to the analysis of unknown traffic and the detection of new attack traffic. Considering this platform, a lot of similarities between the biological processes of signal transduction and communication networks can be expected. The knowledge on the specificity of receptor-ligand interactions as well as the specificity of intracellular signaling pathways and the cellular answer that results from the particular pathway might give impressions how the mechanisms in computer networks can be improved.

5. *Related Work*

The first approaches to identify mechanisms in nature to address technological problems, especially in computer science date back to the mid 1970s. Later, the human immune system was used for investigations on computer viruses and their detection [4]. This is still the best-known example of so-called bio-inspired computing [10].

The group of Prof. Suda is investigating an architecture, which they named the bio-networking architecture [16]. The basis of their project is a middleware platform [14], which aims to incorporate mechanisms known from swarms of bees and ant colonies in order to achieve a high level of autonomy and reliability.

Especially in the field of computer networking, we believe that there is still much work to be done employing bio-inspired mechanisms. The relevance is also demonstrated, for example, by the current research-funding programs of the NSF [1] and the European Union.

Other research areas in computer science and engineering do also profit from bio-inspired research. Examples are pattern recognition, robot control, and cognitive sciences.

6. *Conclusions and Further Work*

In conclusion it can be said that we were able to show many similarities in the signaling pathways known from molecular biology and from computer networks.

The attempt to adopt mechanisms from the basis of the organism, the single cell, looks promising for further studies because the cell has to maintain its own assembly and thus react specifically on changes of their direct environment without the necessity to know about all interactions in the organism. The efficient communication system of cells can be copied to computer networks. The next step will be to analyze processes in network entities which might react in a more “cellular” or “biological” manner to improve their effectiveness. Our primary goals are the development of self-configuring, self-maintaining, and self-healing network systems proving a better protection against security threats.

In conclusion, actual problems of computer networks have to be discussed on the model that each part in the network corresponds to an adequate structure in the organism. Looking at the defined adequate structure in the organism, the mechanisms of interaction with other parts of the system can be analyzed in detail and assigned for computer networks. We are going on in our research activities creating models representing the signaling pathways in communication networks and allowing us to incorporate the studied mechanisms from biology. From such a model, simulative analyses can be done to show the advantages of our new algorithms.

REFERENCES

[1] "National Science Foundation: Summary of FY 2003 Budget Request to Congress," NSF, 2003.
[2] N. Abdel Wahab and R. M. Mason, "Connective tissue growth factor and renal diseases: some answers, more questions," *Curr Opin Nephrol Hypertens*, vol. 13, pp. 53-8, Jan 2004.
[3] R. K. C. Chang, "Defending against Flooding-Based Distributed Denial-of-Service Attacks: A Tutorial," *IEEE Communications Magazine*, vol. 10, pp. 42-51, October 2002.
[4] P. D'haeseleer, S. Forrest, and P. Helman, "An Immunological Approach to Change Detection: Algorithms, Analysis and Implications," Proceedings of IEEE Symposium on Security and Privacy, Oakland, CA, USA, May 1996, pp. 110-119.
[5] F. Dressler, "A Scalable Environment for Quality of Service Measurements in the Internet," Proceedings of 2nd IASTED International Conference on Communications, Internet, and Information Technology (CIIT 2003), Scottsdale, AZ, USA, November 2003, pp. 307-312.
[6] N. Duffield, "A Framework for Packet Selection and Reporting," Internet-Draft, draft-ietf-psamp-framework-08.txt, September 2004.
[7] M. Eigen and R. Winkler, *Laws of the Game*, Princeton Univ Pr, 1993.
[8] A. Hahn, J. Heusinger-Ribeiro, T. Lanz, S. Zenkel, and M. Goppelt-Struebe, "Induction of connective tissue growth factor by activation of heptahelical receptors. Modulation by Rho proteins and the actin cytoskeleton," *J Biol Chem*, vol. 275, pp. 37429-35, Dec 1 2000.
[9] M. Handley, "Internet Denial of Service Considerations," draft-iab-dos-00.txt, January 2004.
[10] S. Hofmeyer, "Who says biology need be destiny?," in *CNET News.com*, April 13, 2004 2004.
[11] R. Kemmerer and G. Vigna, "Intrusion Detection: A Brief History and Overview," *IEEE Computer*, pp. 27-30, April 2002.
[12] T. Pawson, "Protein modules and signalling networks," *Nature*, vol. 373, pp. 573-80, Feb 16 1995.
[13] J. Quittek, T. Zseby, B. Claise, and S. Zander, "Requirements for IP Flow Information Export," draft-ietf-ipfix-reqs-16.txt, June 2004.
[14] J. Suzuki and T. Suda, "Design and Implementation of a Scalable Infrastructure for Autonomous Adaptive Agents," Proceedings of 15th IASTED International Conference on Parallel and Distributed Computing and Systems, Marina del Ray, CA, USA, November 2003.
[15] T. van Biesen, B. E. Hawes, D. K. Luttrell, K. M. Krueger, K. Touhara, E. Porfiri, M. Sakaue, L. M. Luttrell, and R. J. Lefkowitz, "Receptor-tyrosine-kinase- and G beta gamma-mediated MAP kinase activation by a common signalling pathway," *Nature*, vol. 376, pp. 781-4, Aug 31 1995.
[16] M. Wang and T. Suda, "The Bio-Networking Architecture: A Biologically Inspired Approach to the Design of Scalable, Adaptive, and Survivable/Avilable Network Applications," Proceedings of 1st IEEE Symposium on Applications and the Internet (SAINT), San Diego, CA, USA, January 2001.

Bettina Krüger (bettina.krueger@fd42.de) received her M.Sc. in biology from the University of Goettingen, Germany and her Ph.D. from the University of Erlangen, Germany in 1998 and 2003, respectively.
She is currently a PostDoc at the Medial Clinic IV, Nephrology at the University of Erlangen-Nuremberg, Germany.

Falko Dressler (dressler@informatik.uni-erlangen.de) received his M.Sc. and Ph.D. in computer science from the University of Erlangen in 1998 and 2003, respectively.
Currently, he is an assistant professor of computer science at the University of Erlangen-Nuremberg, Germany.