A Study of Self-Organization Mechanisms in Ad Hoc and Sensor Networks

Falko Dressler

Computer Network and Communication Systems, Dept. of Computer Sciences, University of Erlangen, Martensstr. 3, 91058 Erlangen, Germany
dressler@informatik.uni-erlangen.de

Abstract

Self-organization is a great concept for building scalable systems consisting of a huge number of subsystems. The primary objectives are improved scalability and dynamic adaptation to changing environmental conditions. Until now, many self-organization methods have been developed for communication networks in general and ad hoc networks in particular. Nevertheless, the term self-organization is still often misunderstood or misused. This paper contributes to the networking community by providing a better understanding of self-organization mechanisms focusing especially on the applicability in ad hoc and sensor networks. The main contributions of this paper are a clarification of the term self-organization and a categorization of self-organization methods. Additionally, well-known protocols in ad hoc and sensor networks are classified and selected case studies are provided. Primarily, solutions for the medium access control and the network layer are analyzed and discussed. Finally, open research issues with practical relevance are discussed.

Key words: self-organization, ad hoc networks, sensor networks, ad hoc routing, wireless communication, adaptive mechanisms

1 Introduction

Self-organization is one of the most fascinating concepts of many natural systems. Often, huge numbers of individuals participate on a common objective. This collaborative behavior can be observed in form of visible patterns emerging on a higher level. The primary objectives are improved scalability and dynamic adaptation to changing environmental conditions. From an academic point of view, the ideas of self-organization and the corresponding methods and techniques were first analyzed in biological systems [1,2].

Article published in Elsevier Computer Communications 31 (2008) 3018–3029

Since these days, a great number of (technical) solutions have been developed, which, either on purpose or unintentionally, inherently use the basic concepts of self-organization. The fundaments of self-organization are simple algorithms executed by autonomously acting systems. The decision process relies on sensitive feedback loops, interactions among these systems and with the environment, and probabilistic decision processes.

In communication networks, self-organization techniques are important building blocks. In this paper, we concentrate on issues in ad hoc and sensor networks, because of the spontaneous interaction of multiple heterogeneous components over multihop wireless radio connections [3,4] without human interaction. Eventually, self-organization is the only possible solution for many challenging problems in this area but it definitely is not the universal remedy.

Until now, many self-organization algorithms and protocols have been developed for communication networks in general and ad hoc and sensor networks in particular. Nevertheless, the term self-organization is still often misunderstood or misused. The objective of this paper is to introduce the basic concepts of self-organization and to provide a better understanding of the employed paradigms. Additionally, the basis methods are identified and analyzed. During these discussions, the main focus lies on technical systems with respect to ad hoc networking and wireless sensor networks. The main contributions of this paper are a clarification of the term self-organization and a categorization of self-organization methods. Additionally, well-known protocols in ad hoc and sensor networks are discussed based on selected case studies.

The rest of the paper is organized as follows. In section 2, the most challenging properties of ad hoc and sensor networks are briefly summarized that motivate the paradigms change towards self-organization. Section 3 introduces the basis methods of self-organization. Based on these information, a classification of self-organization techniques used in ad hoc and sensor networks is presented in section 4. Selected case studies are provided to enable a better understanding of the presented classification scheme. Finally, section 6 concludes the paper.

2 Properties of Ad Hoc and Sensor Networks

Ad hoc and sensor networks have become a major research domain [3,5]. This trend is mainly forced by advances in microelectronics and the development of new generations of wireless communication devices. Depending on the application scenario, different requirements have to be considered.

In the beginning, the communication in wireless networks was inspired by well-known approaches from wired networking standards such as the Internet. It quickly turned out that most of these algorithms are not appropriate for wireless communication.

In general, ad hoc and sensor networks have quite different characteristics and application domains. Nevertheless, many algorithms are used for both types of networks. Ad hoc networks are by definition operating without any pre-deployed network infrastructure. All the nodes are participating as end systems and provide network functions such as routing. Thus, the network topology is completely self-organized and all network functions are distributed to all the nodes. Sensor networks also incorporate these characteristics whereas different requirements in terms of mobility and resource constraints apply [6]. The most prominent challenges in ad hoc and sensor networks are outlined in the following with emphasis on the differences between both network types.

- *Unreliable wireless communication* – In many cases, especially with a growing number of nodes in the same communication range, wireless communication tends to become unreliable. The main reason is either the increasing number of collisions for CSMA access or the dramatically reduced bandwidth for TDMA variants. The collision probability is proportional to the network density, the path length, and the network load. Using non-deterministic techniques to form an ad hoc network and to exchange data can have strong implications on the performance [7].
- *Spatial and temporal mobility* – Mobility has a large impact on the behavior of ad hoc networks [8,9]. In general, spatial and temporal mobility need to be considered. Spatial mobility refers to geographical movements, i.e. changes of node locations over the time, which is the typical case for ad hoc networks. In contrast, temporal mobility that is common for sensor networks represents temporal changes of the network topology, e.g. due to employed duty cycles that enable the node to be switched off for an essential amount of time. Employed mechanisms for routing and data dissemination must be able to identify changes of the environment and to tolerate mobility [10,11].
- *Resource limitations* – In ad hoc networks, especially the battery capacity is a limiting factor. Sensor networks introduce special constraints that need to be addressed [5]. Resource limitations in terms of CPU power and storage can be handled much simpler than the strongly limited energy capacities. Optimization goals could be the *network lifetime* and the *sensor coverage* [12,13]. For example, the coverage must be improved while preventing any global state information at the same time.
- *Real-time requirements* – Real-time constraints need to be considered if network-controlled actuation needs to be coordinated [14]. In such applica-

tions, the coordination relies on real-time capabilities at application level – even if the same resource limitations and mobility issues have to be considered [15,16].

2.2 Operation and control: a paradigm shift

Management and control in massively distributed systems such as ad hoc and sensor networks needs to be completely distributed, i.e. each participating subsystem has its own control process as shown in Figure 1. Basically, the control paradigm of such systems can be compared to the centralized one. Each system operates on its own without the need for perfect coordination with all other participating systems. To achieve high scalability, a new control theory is required for managing massively distributed systems that finally ensures the cooperation of all subsystems according to a common objective. *Self-organization* is discussed as the ultimate solution to the mentioned problems. While pure self-organization comes along with inherent limitation that we will discuss later, complex systems become manageable and the scalability is even further increased.

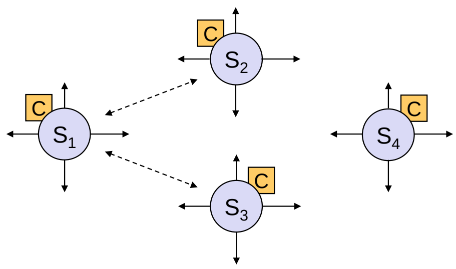

Fig. 1. Management and control in massively distributed systems: each system has its own control process without access to a centralized control or management process

In this paper, we focus on selected problem domains in the field of ad hoc and sensor networks to discuss the relevance and capabilities of self-organized management and control: medium access control (MAC), routing and forwarding, and coordination algorithms such as clustering techniques. In all these domains, the primary measures to evaluate a particular approach are protocol overhead vs. energy efficiency but also time constraints (e.g., latency) need to be considered.

Unreliable wireless communication, spatial and temporal mobility, resource limitations, system reliability, security, and real-time or QoS requirements demand for (sometimes conflicting) new solutions to control the network [4]. Self-organization is thought to be a solution for all the mentioned problems. Thus, we will discuss the basic meanings of self-organization in the following section followed by a classification of techniques available in the domain of ad hoc and sensor networks.

3 Self-Organization

Self-organization is not an invention nor it was developed by an engineer. The principles of self-organization have been evolved in nature and we finally managed to study and apply these ideas to technical systems. Especially Eigen [1] made the term self-organization popular in natural and engineering sciences.

The research domain of distributed systems is working on similar solutions. Novel approaches lead to control and collaboration paradigms that show the same behavior as have been described for self-organizing systems. For example, in the area of ad hoc and sensor networks, multiple, possibly different solutions have been elaborated. The common objective is to reduce global state information by achieving the needed effects based on local information or probabilistic approaches only. Most of these solutions are using (whether explicitly or implicitly) methodologies similar to biological systems [17,18]. Even if the area of bio-inspired networking is a novel research domain, first summaries and overviews to such approaches are available [19].

3.1 Understanding self-organization

The term *self-organization* refers to a specific control paradigm for complex systems. It covers a number of properties that can be regarded as the basic requirements motivating the use of self-organization. In the context of this paper, we use the following definition of self-organization: *Self-organization is a process in which structure and functionality (pattern) at the global level of a system emerge solely from numerous interactions among the lower-level components of a system without any external or centralized control. The systems components interact in a local context either by means of direct communication or environmental observations without reference to the global pattern.*

Comparing self-organized systems with centralized and distributed systems, the main difference is the lack of any globally valid state information. Centralized systems have by definition a central intelligence that maintains global state and derives operational behavior from it. In distributed systems, such state is also available whereas distributed among a number of nodes. Abstraction techniques are used to let the distributed system appear as a single coherent system to the user or application.

In contrast, the definition of self-organization focuses on the emergence of patterns. Similar definitions can be found in the literature concerning well-studied methodologies in biological systems [20]. The interaction of single components finally defines the behavior of the global system. Applied to ad hoc networks, self-organization can be seen as the interactions between nodes in the network leading to globally visible effects, e.g. the transport of messages from a source node to a sink node. Table 1 summarizes the main properties of self-organizing systems.

Property	Description
no central control	There is no global control system or global information available. Each subsystem must perform completely autonomously.
emerging structures	The global behavior or functioning of the system emerges in form of observable pattern or structures.
resulting complexity	Even if the individual subsystems can be simple and perform basic rules, the resulting overall system becomes complex and often unpredictable.
high scalability	There is no performance degradation if more subsystems are added to the system. The system should perform as requested regardless of the number of subsystems.

Table 1
Properties of self-organizing systems

Self-organization is often referred to as the multitude of algorithms and methods that organize the global behavior of a system based on inter-system communication. Most networking algorithms work like that. Therefore, self-organization in this context is not a new solution. Nevertheless, most of these algorithms are based on global state information, e.g. routing tables. In the networking community, it is commonly agreed that such global state is the primary source of scalability problems of the particular algorithms. Especially in the area of ad hoc and sensor networks, new solutions were discovered that show the properties of the new definition of self-organization. Most ad hoc routing algorithms as well as data centric data dissemination approaches are well-known examples [10,15].

3.2 Methods and techniques

Self-organizing systems completely rely on localized decision processes. Three basic methods can to be distinguished that enable the desired behavior. These are the building blocks for all approaches that should self-organize as depicted by our definition of self-organization:

- *Positive and negative feedback* – The local state of an individual system can

be adapted to changing environments based on adequate feedback on its own actions. Feedback can be divided into positive and negative feedback. Positive feedback provides amplification capabilities, to act efficiently and in real-time to particular changes. The amplifying nature may lead to snowballing effects and overloading amplifications. Negative feedback is used for keeping positive feedback under control. Such feedback can be created in form of rules that have been developed to keep the system state within a given parameter range. In other cases, such an inhibition arises automatically, often simply from physical constraints. Positive and negative feedback must be strictly coordinated. If these mechanisms are used uncontrolled, the system will either tend to do nothing, over-react, or to oscillate between multiple states. Different terms are used in the literature, for example, positive feedback is equal to reaction, amplification, and promotion, and negative feedback is often described as diffusion, suppression, and inhibition.
- *Interactions among individuals and with the environment* – In our discussion of self-organization and self-organizing system, we outlined as a main characteristic that the operation and control, i.e. the organization, arises entirely from multiple interactions among their components. We need to distinguish between two kinds of information transfer: Information transfer between individuals, i.e. direct communication between neighboring (in time and space) individuals via signals, and interactions with the environment, i.e. indirect information flows via cues arising from work in progress (stigmergy) [21].
- *Probabilistic techniques* – A third component for successfully building self-organizing systems is probability. Basically, all self-organizing systems include probabilistic techniques. Such mechanisms can be either used to entirely organize the local behavior of a single system or at least for parameter settings of other deterministic algorithms.

It is important to note that usually more than one of these methods is used to assemble the desired system behavior. For example, probabilistic methods can provide a good initial configuration and help to distribute the load over spacial and temporal boundaries. At the same time, feedback loops ensure a proper adaptation. Efficient communication is necessary to exchange state and control information. This can be done either direct or indirect via changes of the environment.

3.3 Limitations

Even though self-organization solves many problems that accompany the development of autonomously acting entities that are entitled to collaboratively work on a global goal, some problems are still unsolved or might even appear by introducing self-organization techniques. In the following, we discuss some of these open issues that can be seen as starting points to conduct further research on self-organization and distributed control.

- *Controllability* – As shown in Figure 2, the predictability of the behavior of a self-organizing system is rapidly decreasing while increasing its scalability. This problem is directly related to the controllability of the system. For example, classical network management solutions cannot be employed in self-organizing networks because the necessary state information cannot be retrieved. Therefore, the operability of the network can only be estimated or approximated. Even harder is the guarantee of quality of service parameters.
- *Cross-mechanism interference* – The composition of multiple self-organizing mechanisms can lead to unforeseen effects. For example, different energy-aware methods implemented at MAC and network layer may interfere and lead either to reduced throughput and reliability or to much higher energy consumption compared to the non-optimized behavior. Cross-layer design and cross-method validation techniques are needed to identify such interferences and to eliminate them.
- *System test* – The test of the system, its components, and the installed software becomes a complex task. It is not possible to create a lab environment showing exactly the properties of the desired deployment scenario. The same holds for field tests because it is not possible to predict future conditions influencing the system.

Figure 2 depicts the main problem ob self-organizing systems, the *reduced determinism*. The more scalable a system becomes by using self-organization techniques, the less control of individual entities is possible. The primary conclusion is that the predictability of the system behavior must be reduced for such a self-organizing system.

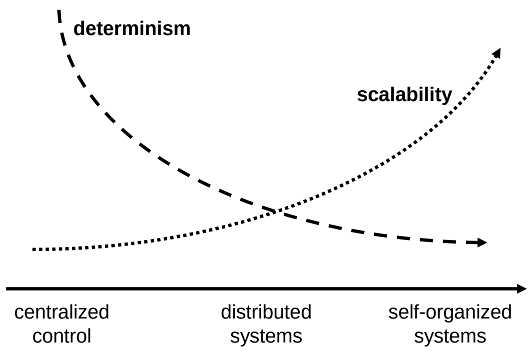

Fig. 2. Scalability vs. determinism in centralized controlled and self-organized systems

Similarly, one may argue that localized algorithms often may not result in a global optimum with respect to a certain property but only provide solutions close to the optimum. This sub-optimality, however, is not a real disadvantage in a dynamic system. Here, optimum configurations are subject to frequent changes and fast convergence to a stable system is more important.

4 Classification of Self-Organization Methods

For this classification, we start with a categorization of self-organization techniques in ad hoc and sensor networks and finally map well-known algorithms and protocols to the identified classes.

4.1 Categorization in two dimensions

The categorization of self-organization methods in ad hoc and sensor networks opens a multidimensional space. In general, these methods can be grouped *horizontally* by their use of state information and *vertically* by their function in the protocol stack.

Figure 3 depicts the horizontal dimension. Reading the figure from the left, necessary state information to perform the particular algorithm is decreasing. Also shown in this figure is a mapping of the basic self-organization methods as described in the previous section to the categorization in ad hoc and sensor networks. Protocols and algorithms developed for these networks need to avoid global state information in order to increase the scalability of the particular approach. The different categories are discussed in the following.

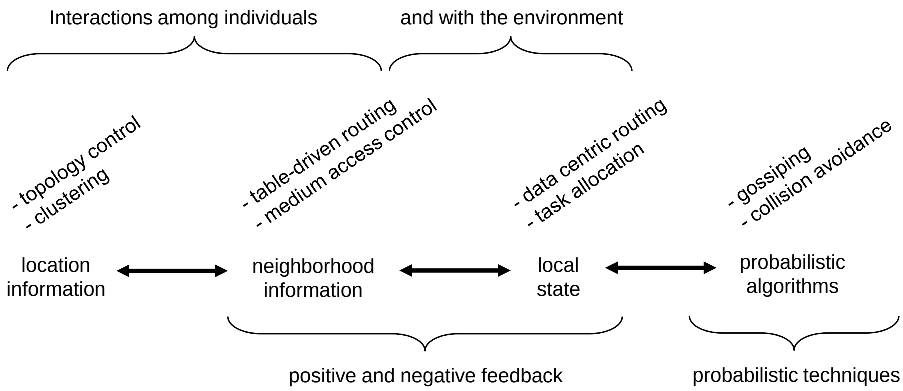

Fig. 3. Horizontal categorization of self-organization mechanisms in ad hoc networks

While the required state is reduced towards the probabilistic methods, the determinism or predictability of the algorithms is reduced as well. Therefore, the best solution for a particular application scenario must be chosen carefully by comparing all application requirements at once.

- *Location information* – Geographical positions, e.g. for Greedy Perimeter Stateless Routing (GPSR) [22], or affiliation to a group of surrounding nodes, e.g. clustering mechanisms [23], are used to reduce necessary state information to perform routing decisions or synchronizations. Usually, similar methods as known for global state operations can be employed in this context. Depending on the size of active clusters or the complexity to perform localization methods, such location-based mechanisms vary in communication and processing overhead. Regardless of the specific algorithms, absolute or relative location information must be collected and maintained.

- *Neighborhood information* – Further state reduction can be achieved by decreasing the size of previously mentioned clusters to a one-hop diameter. In this case, only neighborhood information is available to perform necessary decisions. Usually, so called `hello` messages are exchanged in regular time periods. This keeps the neighborhood information up-to-date and allows the exchange of performance measures such as the current load of a system.
- *Local state* – Local system state is always available. Especially in sensor networks, system parameters can be combined with environmental conditions as observed by attached sensors. Local state can be manipulated by different events. For example, previous message exchanges can influence the node behavior as well as the current time. Examples for local state techniques are data-centric routing and some specific task allocation schemes.
- *Probabilistic algorithms* – In some cases, it is useful to store no state information at all. For example, if messages are very infrequently exchanged or in case of high mobility. Then, pure probabilistic methods can lead to good results. Statistical measures can be used to describe the behavior of the overall system in terms of predicted load and performed operations. Obviously, no guarantee can be given that a desired goal will be reached. Additionally, probabilistic approaches help preventing global synchronization effects that influence the efficiency of the overall system.

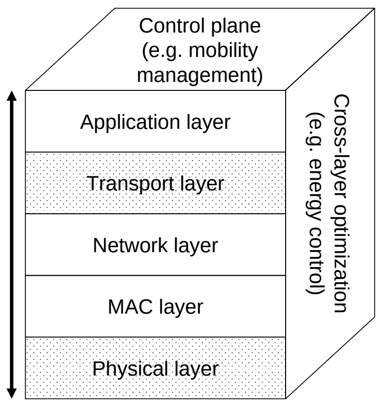

Fig. 4. Vertical categorization of self-organization mechanisms in ad hoc networks

In contrast, Figure 4 shows the layered system architecture. A common control plane coordinates and controls mobility questions and some additional cross-layer or cross-service issues have to be considered that are describing non-functional properties such as energy, security, end-to-end performance, and coverage. Based on the given application scenario, particular mechanisms from different layers might interact to achieve a common goal, e.g. to reduce the necessary amount of energy, but they might also interfere with one another, e.g. by defining different sleep cycles at different layers to reduce the energy consumption.

- *MAC layer* – Medium access control manages the access to the wireless radio link. Historically, contention-based mechanisms dominate this layer. Additionally, synchronization between neighboring nodes can be used to optimize the link sharing. While overhearing techniques enable the nodes to transmit messages at arbitrary times (obviously, the receiving node will be "always on"), low duty cycles can be used to reduce the energy consumption of the radio receiver. Self-organization mechanisms help to perform concurrent access, to synchronize nodes, and to maintain duty cycles in a distributed manner without the need of central management and pre-configuration of nodes and algorithms.
- *Network layer* – End-to-end forwarding of data packets is provided by the network layer. Two very different tasks must be solved at this layer: routing and data forwarding. We can distinguish between different techniques: proactive and on demand routing. Obviously, both mechanisms finally rely on state information must be synchronized between many (or even all) nodes in the network. On the other hand, forwarding refers to the delivery of messages to the next hop towards the final destination. Data-centric approaches have been proposed as promising solutions for sensor networks.
- *Application layer* – Besides the contents of the networked application itself, many coordination tasks must be organized such as the identification of master nodes are the allocation of tasks to one of the available nodes. Coordination algorithms such as clustering and distributed task allocation schemes have been studied in various kinds of networks that are based on self-organization methods.

4.2 Mapping of algorithms

In this section, methods developed for ad hoc and sensor networks are classified by employed self-organization mechanisms. We follow the horizontal classification and discuss the mechanisms in vertical order.

4.2.1 Global State

Using global state information, optimal solutions can either be directly calculated or at least approximated in the case of multi-objective optimizations. The primary purpose of global state algorithms in communication networks is twofold. First, this global state must be collected and maintained, and, secondly, the specific optimization algorithms must be performed. Depending on the size of the network in terms of participating nodes, both steps may require unsuitably high amounts of time and memory.

In MAC protocols, the optimal scheduling of sleep cycles can be calculated using such distributed state information. For example, mechanisms for delay efficient sleep scheduling [24] and energy efficient real-time medium access control [25] have been developed. Nevertheless, the maintenance of the perfect synchronization induces eminent scalability problems [26].

Similar to most Internet routing protocols, pro-active routing mechanisms have been developed for ad hoc networks first. All these protocols are based on periodic state exchange. Well-known examples are DSDV (Destination-Sequenced Distance-Vector Routing) [27] and hierarchical approaches such as HSR (Hierarchical State Routing) [28]. All these protocols are always able to find an optimal path through the network, whereas they differ in the convergence speed and the necessary maintenance overhead. Mobility management and energy control can be incorporated into global state algorithms, nevertheless the state maintenance is clearly too expensive [10].

4.2.2 Location information

Spatial restrictions of global optimization algorithms is a first approach to improve the efficiency of operation in ad hoc networks [29]. For example, the explicit calculation of a route towards a destination can be prevented if the positions of all the nodes are known [30]. Nevertheless, the retrieval of such geographic information can be expensive, especially in mobile networks [31] whereas first solutions for location-aided routing in mobile ad hoc networks are available [32].

On the other hand, clustering mechanisms have been studied to enhance the performance in ad hoc and sensor networks and to reduce the necessary amount of energy at the same time [33]. The primary idea is to group nodes around a so called clusterhead that is responsible for state maintenance and inter-cluster connectivity. Clustering is a crosscutting technology that can be used in nearly all layers of the protocol stack. Examples for efficient clustering algorithms are passive clustering [34], which reduces the necessary overhead for maintaining the structure of the clusters, and on-demand clustering [35], which mitigates the need for permanent maintenance of clusters by creating them on-demand. Routing algorithms make frequently use of efficient clustering mechanisms [36]. Typical examples of power-aware cluster-based communication solutions for ad hoc networks are LEACH (Low-Energy Adaptive Clustering Hierarchy) [37] and its competitor HEED (Hybrid, Energy-Efficient, Distributed Clustering Approach) [38]. These approaches have been extended to coordination on the application layer [39,16].

4.2.3 Neighborhood Information

Compared to cluster maintenance or even global state, neighborhood information can be gathered quite easily (usually, it is also used as a starting point for maintaining clusters or global state). The basic idea is to periodically exchange some `hello` or `sync` messages that include necessary information for the particular algorithm to take decisions based on its local state and the state of its neighbors. The overhead for maintenance is drastically reduced whereas globally optimal solutions are difficult to obtain, e.g. end-to-end communication paths and global allocation schemes.

At the MAC layer, `sync` messages can be employed to synchronize all neighbors to a common sleep cycle [24], to organize the message exchange using RTS/CTS, or to provide enhanced performance solutions such as adaptive listening [40]. The most prominent MAC protocol in the wireless ad hoc domain is IEEE 802.11 [41]. It features RTS/CTS-based solutions for the hidden/exposed terminal problems, adaptive sleep cycles, and energy-control using overhearing techniques. More specialized for wireless sensor networks is S-MAC (Sensor MAC) [42,40]. A techniques named adaptive listening has been integrated for enhanced performance and reduced energy consumption. PCM (Power Control MAC) [43] is an energy-aware extension to typical contention-based MAC protocols. It adapts the transmission power to the estimated distance between two nodes.

At the network layer, two different approaches rely on neighborhood information: reactive routing protocols and data-centric communication methods (also known as objectivity-driven). Reactive routing protocols do not keep global routing tables up-to-date. Instead, they only manage neighborhood relationships. When messages need to be transmitted, routing information is gathered on demand by flooding route requests through the network in order to find a suitable path towards the destination. Several optimizations in terms of adjustable caches for previously determined route information allow a fine-tuning of the algorithms depending on the application scenario. Best-known examples are AODV (Ad-Hoc On Demand Distance Vector Routing) [44,45] and its successor DYMO (Dynamic MANET On Demand) [46]. In contrast, data-centric communication methods prevent the calculation of routing paths by employing interest distributions. In the class of diffusion algorithms, multiple algorithms have been proposed. Of these, directed diffusion [15,47] is still the best known approach. Other variants try to optimize particular aspects such as the minimization of the energy consumption [48] or the inclusion of geographical information in GEAR (Geographical end Energy-Aware Routing) [49].

4.2.4 Local state

In most circumstances, the local system state can be maintained very easily. In this context, we consider the update of primary system parameters, the reception of messages, the update of sensor readings to examine environmental conditions, or internal events such as timers, which may be used to modify the system behavior over the time. In summary, any algorithm relying on the value of one or many local system variables only falls into this category.

Many MAC protocols employ local state information for various purposes. One example are message retransmissions after a timeout using a binary backoff. On the other hand, PCM [43] is an energy-aware extension to typical contention-based MAC protocols. It adapts the transmission power to the current needs in the local network.

At the network layer, a number of data centric routing algorithms have been proposed in the last years. The main idea is to process received messages according to their internal "meaning" [47,50]. The best known example for data centric message processing are various aggregation algorithms [51]. Additionally, complete message processing including message modification, data aggregation, actuation control, and routing is possible using rule-based approaches [52].

4.2.5 Probabilistic techniques

The category of probabilistic algorithms intents to keep no state information at all. Therefore, it shows the best behavior if very few messages per time have to be transmitted because the overhead due to state maintenance is negligible. The overhead for actual transmitting messages can be much higher. There cannot be an optimal path from a source towards a sink; probabilistic algorithms are used instead. In MAC protocols and congestion-aware communication mechanisms, stochastic distributions and random delays are employed to prevent the global synchronization effect. For routing and data dissemination in ad hoc networks, probabilistic algorithms are often used to prevent pure flooding of messages through the whole network. A comparison of data dissemination protocols in ad hoc networks is, for example, provided in [53,54].

Without routing tables, information exchange in communication networks can be organized by flooding the messages through the entire network. Optimized flooding strategies [34] try to prevent the forwarding of duplicates of the packet by using a maximum time-to-live or sequence numbers, i.e. limited local state information. The probability that a message will arrive at a destination is very high even in case of mobility and error-prone wireless channels. On the other hand, the overhead due to message transmissions into unessential parts of the network increases with the network size. Gossiping [55] and rumor routing [56] as alternatives to flooding have been developed to cope with this problem. Probabilistic parametric routing [57] and WPDD (Weighted Probabilistic Data Dissemination) [58] further improve the behavior of these algorithms. The optimization goal is the overhead due to unnecessary messages compared to the probability of reaching the final destination. This group of algorithms can be extended to probabilistic lightweight group communication [59] and task allocation schemes as well.

4.3 Evaluation of algorithms

The evaluation of self-organizing systems can be performed in multiple dimensions and according to a broad number of criteria. Besides classical performance measures, especially two parameters are of interest. The first one is *scalability*, i.e. the supported number of interacting systems or the maximum size of a system. Scalability must be regarded as an overall measure for how well the different protocols and techniques cooperate with each other. This also affects the *reliability* of the communication as connections might be influenced by the methodology itself. Depending on the application scenario, scalability is often the limiting factor in sensor networks as typically networks with huge amounts of nodes are considered.

The primary application constraint is the *network lifetime*, i.e. the ability of the network to fulfill all application requirements for a given (limited) time or to maximize network lifetime in order to enable further operation independent of any external maintenance. Besides pure energy concerns, most characteristics describing the quality of an application, e.g. sensor coverage or transmission delay, can be reduced to lifetime discussions. Additionally, it also describes the *availability* of the network, i.e. the question whether the sensor network is always available or if are there special protocol-inherent maintenance periods.

5 Selected Case Studies

Three case studies are provided in this section as examples for the previously depicted and classified self-organization mechanisms. The selection of these examples does not quality one of these approaches to be one of the best in its category. The discussed case studies were chosen only for their straightforward employed self-organization algorithms.

5.1 MAC layer – PCM

Power control is an important issue in MAC protocols. Most schemes vary the transmit power to reduce the overall energy consumption. In addition to providing energy saving, power control can potentially be used to improve spatial reuse of the wireless channel. PCM (Power Control MAC) [43] is an extension to typical contention-based MAC protocols.

The principles of PCM are shown in detail in Figure 5. The transmission of RTC/CTS handshake messages is performed with the maximum available power p_{max}. The required transmission power $p_{desired}$ that is used for the subsequent DATA/ACK transfer is determined from the received power level p_r in combination with some well-known minimum received signal strength Rx_{thresh} that is necessary for correctly decoding the messages according to $p_{desired} = \frac{p_{max}}{p_r} \times Rx_{thresh} \times c$. Therefore, the calculation is based on locally available information only by observing the neighboring environment. However, it was shown that this scheme can degrade network throughput and can result in higher energy consumption than when using IEEE 802.11 without power control. PCM proposes some enhancements that do not degrade throughput and yield energy saving.

The problem is that there are nodes that can sense the signal of the RTS/CTS exchange but cannot decode it because the signal level is too weak (nodes '1' and '8' in Figure 5). During the DATA/ACK period, these nodes do not sense a signal any longer. Therefore, they may initiate their own RTS/CTS exchange which results in a collision with the still ongoing data transmission. The power control MAC protocol addresses this issue by varying the signal level of the data transfer by periodically increasing it to p_{max} allowing distant nodes to sense the signal of the ongoing transmission. The period can be adapted to the carrier sensing algorithm in order to optimize the behavior of the protocol.

In summary, it can be said that the PCM is a good example for achieving optimal throughput by reducing the necessary transmission energy to a minimum. It does so using neighborhood information only, i.e. the knowledge extracted from monitoring and analyzing the surrounding behavior and conditions. Therefore, node mobility is supported as well as changes of the network topology.

5.2 Network layer – directed diffusion

Address-based routing depends on globally unique addressing schemes as well as on pro-actively or re-actively created and maintained routing tables for path

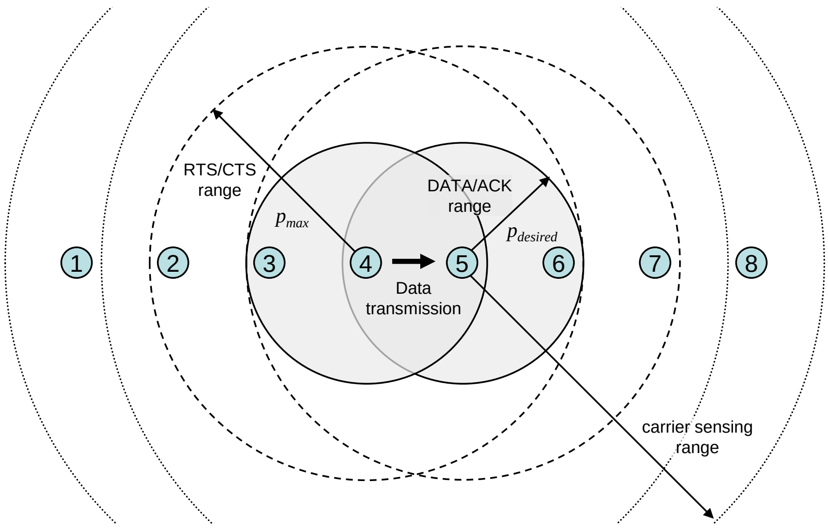

Fig. 5. Transmission ranges of the power control MAC protocol

calculations. Instead, directed diffusion [47] is a data-centric communication approach. The mechanism was developed for use in wireless sensor networks in which nodes can coordinate to perform distributed sensing of environmental phenomena. All nodes in a directed diffusion-based network are application-aware. This enables diffusion to achieve energy savings by selecting empirically good paths and by caching and processing data within the network.

Basically, directed diffusion consists of two different mechanisms: interest distribution and data propagation. Using naming methods, interest can be described in the form "I am looking for all measures of temperature higher than 20°C in the area [10,10,300,200]". Such interest is diffused through the network as shown in Figure 6 (a). Intuitively, this interest may be thought of as exploratory; it tries to determine if there are any sensor nodes that detect the requested measures. Such interest messages are renewed periodically to keep them up-to-date. Each node in the network that forwards this interest message sets up a gradient towards the source of the interest, i.e. the sink node (Figure 6 (b)). A special reinforcement process is employed to weight the gradients based on their qualities, e.g. loss ratio or hop count. If a sensor node receives the message that actually can provide the requested data, it will finally start to measure for example the temperature and to transmit the results along the chosen gradient towards the sink as depicted in Figure 6 (c).

Accordingly, directed diffusion operates on local requirements in form of interests, their diffuse distribution through the network, and temporary state maintenance in form of gradients. Depending on the number of nodes and the number of active interests, the utilization of the network is very low or comparable to other routing approaches. Directed diffusion allows to switch between different design choices for the implementation or even their runtime-change to adapt to changing environments. For example, the interest propagation can

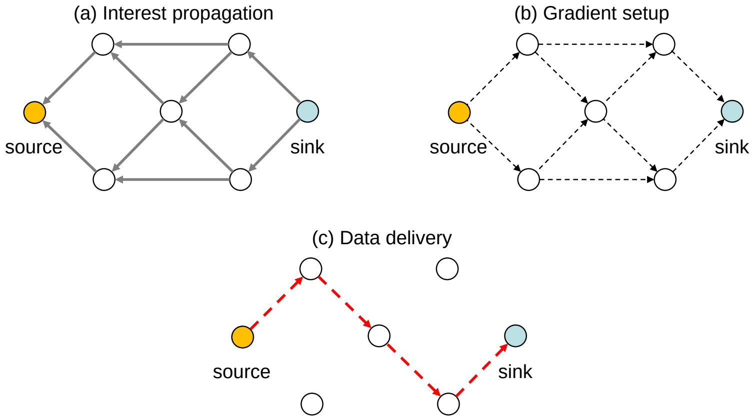

Fig. 6. Interest and data propagation usind in directed diffusion

employ network broadcasting, directed network broadcasting based on the location, or directional propagation using previously cached data. Data propagation can be implemented in form of single path delivery or probabilistic multipath forwarding. Additionally, data caching and aggregation algorithms can be employed for robustness and data reduction. For example, Intanagonwiwat et al. presented the benefits of suppressing received duplicate data [47]. This example shows again that multiple self-organization mechanisms can be successfully coupled to build a communication protocol that makes efficient use of the available resources. Because no addressing scheme is needed and the interest information is periodically updated, directed diffusion supports spatial and temporal mobility.

5.3 Topology control and clustering – LEACH

LEACH (Low-Energy Adaptive Clustering Hierarchy) [37] is a clustering-based protocol that utilizes randomized rotation of local cluster base stations (clusterheads) to evenly distribute the energy load among the sensors in the network. Energy load is defined in this context as the utilization in terms of awakness plus needed transmission energy. LEACH uses localized coordination to enable scalability and robustness for dynamic networks, and incorporates data fusion into the routing protocol to reduce the amount of information that must be transmitted to the base station. The primary goal is to equally distribute the energy load to all available nodes and to enhance the lifetime of the entire network. In [37], simulations were conducted that show that LEACH can achieve as much as a factor of eight reduction in energy dissipation compared with conventional routing protocols. In addition, LEACH is able to distribute energy dissipation evenly throughout the sensors, doubling the useful system lifetime for the networks that were simulated.

The operation of LEACH is broken up into rounds as depicted in Figure 7,

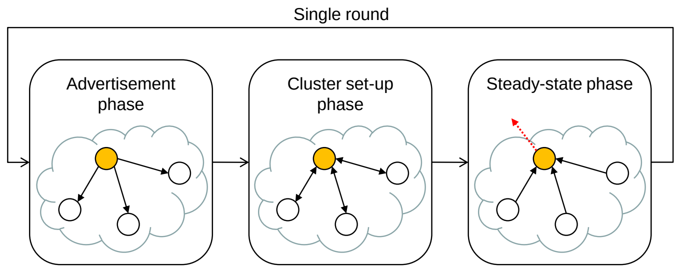

Fig. 7. Round based operation of LEACH

where each round begins with a set-up phase, when the clusters are organized, followed by a steady-state phase, when data transfer occurs. The election process works as follows. At the beginning of a round, each node decides whether or not to become a clusterhead. This decision process depends on only a single pre-defined value, the desired percentage P of cluster heads in the network, i.e. the number of clusters to be created. Each nodes elects itself to be a clusterhead with a certain probability. For this, each node choses a random number n between 0 and 1. If the number is less than a threshold $T(n)$, the node will become a clusterhead for the current round r. This threshold is calculated as follows (G denotes the set of nodes that have not been clusterhead in the last $1/P$ rounds):

$$T(n) = \begin{cases} \dfrac{P}{1 - P * (r mod 1/P)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Therefore, in round zero, each node has a probability P of becoming clusterhead. The probability of the remaining nodes must be increased because there are fewer nodes left for becoming clusterhead in round $r+1$. Finally, the cluster-head nodes broadcast their status to the other nodes in the network. Based on this information, each node determines to which cluster it wants to belong by choosing the clusterhead for which the communication energy can be minimized.

In summary, it can be said that LEACH operates on locally taken decisions that are broadcasted to all neighboring nodes. Based on the local decision and some (local) communications, the nodes organize themselves into a larger compound for energy-aware operations. Therefore, LEACH combines multiple mechanisms for self-organization: probabilistic algorithms (choice of becoming clusterhead) and neighborhood information (set-up of the clusters). This combination together with location information (the clusters are used for efficient data communication), LEACH provides an optimized behavior for communication in ad hoc networks based on self-organization methods. Obviously, in order to minimize the overhead, the steady state phase should be long compared to the set-up phase. Mobility is supported by LEACH, whereas new nodes have to be synchronized to the current round. Node failures may lead

positive and negative feedback	
MAC layer	feedback is used to control the transmission energy in PCM [43], or to synchronize nodes to a common schedule in S-MAC [40]
Routing	positive feedback for route discovery in table-driven protocols, negative feedback for error handling, e.g. in AODV [44]; in directed diffusion [47], received interest acts as a promoter whereas time represents negative feedback for suppressing further actions
Clustering	protocols such as LEACH [37], HEED [38], or Span [39] exploit the observed node density for the cluster election process
interactions among individuals and with the environment	
MAC layer	PCM [43] exploits the RTS/CTS handshake; TDMA schedules are negotiated in IEEE 802.15.4 [60]
Routing	topology maintenance is provides by local interactions; optimized gossiping relies on local interactions [55]; collected neighborhood information in GPSR [22]
Clustering	LEACH [37] and HEED [38] rely on local communication for achieving and maintaining cluster affiliations
probabilistic techniques	
MAC	reduced collision probability through randomized medium access
Routing	simple gossiping [55]; random walk strategies in rumor routing [56]
Clustering	randomized clusterhead selection in LEACH [37] and randomized back-off in Span [39]

Table 2
Summary of self-organization methods used by the discussed examples

to less clusterheads being elected than desired.

5.4 Mapping of self-organization techniques

Finally, Table 2 summarizes the employed self-organization mechanisms by mapping the discussed examples and other well-known algorithms to the basic self-organization methods.

6 Conclusion

This paper contributes to the networking community by providing a broad introduction and classification to the concepts and ideas of self-organization.

After outlining the basis methods of self-organization, we presented a general definition and classification of self-organization mechanisms in ad hoc and sensor networks. Additionally, we discussed the need for such techniques for operation and control in massively distributed systems. Based on the categorization and some clarifying case studies, we have shown that there are already a number of self-organization techniques used for communication and coordination in ad hoc and sensor networks.

Self-organization addresses the main requirements in such networks, i.e. scalability, reliability, and network lifetime and opens novel solution spaces. Nevertheless, it must be mentioned that the employment of self-organization techniques comes with non-negligible costs. The predictability of the communication methodologies is reduced by employing self-organization methods. Therefore, self-organization cannot be seen as a universal remedy. A proper analysis of the particular application and its requirements is necessary in order to choose adequate mechanisms.

In conclusion, we encourage other researchers to continue and intensify their studies on ad hoc network communication in general and self-organization mechanisms in particular. Possible interactions with interdisciplinary research domains should be carefully investigated in order to find and adapt well-studied solutions to the ad hoc community.

References

[1] M. Eigen, P. Schuster, The Hypercycle: A Principle of Natural Self Organization, Springer, 1979.

[2] C. Gerhenson, F. Heylighen, When Can we Call a System Self-organizing?, in: 7th European Conference on Advances in Artificial Life (ECAL 2003), Dortmund, Germany, 2003, pp. 606–614.

[3] C. S. R. Murthy, B. S. Manoj, Ad Hoc Wireless Networks, Prentice Hall PTR, 2004.

[4] J. Misic, V. B. Misic, Wireless sensor networks: Performance, reliability, security, and beyond, Elsevier Computer Communications 29 (13-14) (2006) 2447–2449.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Elsevier Computer Networks 38 (2002) 393–422.

[6] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, Y. F. Hu, Wireless Sensor Networks: a Survey on the State of the Art and the 802.15.4 and ZigBee Standards, Elsevier Computer Communications 30 (7) (2007) 1655–1695.

[7] S. Chakrabarti, A. Mishra, Quality of service challenges for wireless mobile ad hoc networks, Wireless Communications and Mobile Computing 4 (2) (2004) 129–153.

[8] M. Bhatt, R. Chokshi, S. Desai, S. Panichpapiboon, N. Wisitpongphan, O. K. Tonguz, Impact of Mobility on the Performance of Ad Hoc Wireless Networks, in: 58th IEEE Vehicular Technology Conference (VTC2003-Fall), Vol. 5, Orlando USA, 2003, pp. 3025–3029.

[9] F. Bai, A. Helmy, A Survey of Mobility Models in Wireless Adhoc Networks, in: Wireless Ad Hoc and Sensor Networks, Kluwer Academic Publishers, 2004.

[10] K. Akkaya, M. Younis, A Survey of Routing Protocols in Wireless Sensor Networks, Elsevier Ad Hoc Networks 3 (3) (2005) 325–349.

[11] Y. Han, R. J. La, H. Zhang, Path Selection in Mobile Ad-hoc Networks and Distribution of Path Duration, in: 25th IEEE Conference on Computer Communications (IEEE INFOCOM 2006), Barcelona, Spain, 2006, pp. 1–22.

[12] I. Dietrich, F. Dressler, On the Lifetime of Wireless Sensor Networks, Technical Report 04/06, University of Erlangen, Dept. of Computer Science 7 (December 2006).

[13] M. Cardei, J. Wu, Coverage in Wireless Sensor Networks, in: M. Ilyas (Ed.), Handbook of Sensor Networks, CRC Press, West Palm Beach, FL, USA, 2004.

[14] I. F. Akyildiz, I. H. Kasimoglu, Wireless Sensor and Actor Networks: Research Challenges, Elsevier Ad Hoc Networks 2 (2004) 351–367.

[15] D. Estrin, R. Govindan, J. Heidemann, S. Kumar, Next Century Challenges: Scalable Coordination in Sensor Networks, in: 5th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 1999), ACM, Seattle, Washington, USA, 1999, pp. 263–270.

[16] T. Melodia, D. Pompili, V. C. Gungor, I. F. Akyildiz, A Distributed Coordination Framework for Wireless Sensor and Actor Networks, in: 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM Mobihoc 2005), Urbana-Champaign, Il, USA, 2005, pp. 99–110.

[17] F. Dressler, Self-Organization in Ad Hoc Networks: Overview and Classification, Technical Report 02/06, University of Erlangen, Dept. of Computer Science 7 (March 2006).

[18] C. Prehofer, C. Bettstetter, Self-Organization in Communication Networks: Principles and Design Paradigms, IEEE Communications Magazine 43 (7) (2005) 78–85.

[19] F. Dressler, I. Carreras (Eds.), Advances in Biologically Inspired Information Systems - Models, Methods, and Tools, Vol. 69 of Studies in Computational Intelligence (SCI), Springer, 2007.

[20] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraula, E. Bonabeau, Self-Organization in Biological Systems, Princeton University Press, 2003.

[21] A. Abraham, C. Grosan, V. Ramos, Stigmergic Optimization, Vol. 31 of Studies in Computational Intelligence, Springer, 2006.

[22] B. Karp, H. T. Kung, GPSR: Greedy Perimeter Stateless Routing for Wireless Networks, in: 6th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2000), Boston, Massachusetts, USA, 2000, pp. 243–254.

[23] A. A. Abbasi, M. Younis, A survey on clustering algorithms for wireless sensor networks, Elsevier Computer Communications 30 (14-15) (2007) 2826–2841.

[24] G. Lu, N. Sadagopan, B. Krishnamachari, A. Goel, Delay Efficient Sleep Scheduling in Wireless Sensor Networks, in: 24th IEEE Conference on Computer Communications (IEEE INFOCOM 2005), Miami, FL, USA, 2005.

[25] M. Adamou, I. Lee, I. Shin, An energy efficient real-time medium access control protocol for wireless ad-hoc networks, in: 22nd IEEE Real-Time Systems Symposium (RTSS 2001), WIP-Session, London, UK, 2001.

[26] D. Dewasurendra, A. Mishra, Scalability of a scheduling scheme for energy aware sensor networks, Wireless Communications and Mobile Computing 4 (3) (2004) 289–303.

[27] C. E. Perkins, P. Bhagwat, Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers, Computer Communications Review (1994) 234–244.

[28] A. Iwata, C.-C. Chiang, G. Pei, M. Gerla, T.-W. Chen, Scalable Routing Strategies for Ad Hoc Wireless Networks, IEEE Journal on Selected Areas in Communications, Special Issue on Ad-Hoc Networks 17 (8) (1999) 1369–1379.

[29] D. Hollos, H. Karl, A. Wolisz, Regionalizing Global Optimization Algorithms to Improve the Operation of Large Ad Hoc Networks, in: IEEE Wireless Communications and Networking Conference, Atlanta, Georgia, USA, 2004.

[30] M. Mauve, J. Widmer, A Survey on Position-Based Routing in Mobile Ad-Hoc Networks, IEEE Network 15 (6) (2001) 30–39.

[31] R. C. Shah, S. Wiethölter, A. Wolisz, J. M. Rabaey, Modelling and analysis of opportunistic routing in low traffic scenarios, in: 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'05), Trento, Italy, 2005.

[32] Y.-B. Ko, N. H. Vaidya, Location-aided Routing (LAR) in Mobile Ad Hoc Networks, in: 4th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 1998), Dallas, Texas, United States, 1998, pp. 66–75.

[33] Y. P. Chen, A. L. Liestman, J. Liu, Clustering Algorithms for Ad Hoc Wireless Networks, in: Y. Xiao, Y. Pan (Eds.), Ad Hoc and Sensor Networks, Nova Science Publisher, 2004.

[34] T. J. Kwon, M. Gerla, Efficient Flooding with Passive Clustering (PC) in Ad Hoc Networks, ACM SIGCOMM Computer Communication Review.

[35] C. R. Lin, M. Gerla, Adaptive Clustering for Mobile Wireless Networks, Selected Areas in Communications 15 (7) (1997) 1265–1275.

[36] J. Ibriq, I. Mahgoub, Cluster-Based Routing in Wireless Sensor Networks: Issues and Challenges, in: International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS'04), San Jose, California, USA, 2004, pp. 759–766.

[37] W. Rabiner Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy-Efficient Communication Protocol for Wireless Microsensor Networks, in: 33rd Hawaii International Conference on System Sciences, 2000.

[38] O. Younis, S. Fahmy, HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks, IEEE Transactions on Mobile Computing 3 (4) (2004) 366–379.

[39] B. Chen, K. Jamieson, H. Balakrishnan, R. Morris, Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks, ACM Wireless Networks Journal 8 (5).

[40] W. Ye, J. Heidemann, D. Estrin, Medium Access Control with Coordinated Adaptive Sleeping for Wireless Sensor Networks, IEEE/ACM Transactions on Networking (TON) 12 (3) (2004) 493–506.

[41] IEEE, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Standard 802.11-1999 edition (1999).

[42] W. Ye, J. Heidemann, D. Estrin, An Energy-Efficient MAC Protocol for Wireless Sensor Networks, in: 21st IEEE Conference on Computer Communications (IEEE INFOCOM 2002), Vol. 3, New York, NY, USA, 2002, pp. 1567–1576.

[43] E.-S. Jung, N. Vaidya, A Power Control MAC Protocol for Ad Hoc Networks, in: 8th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2002), 2002.

[44] C. E. Perkins, E. M. Royer, Ad hoc On-Demand Distance Vector Routing, in: 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, 1999, pp. 90–100.

[45] C. E. Perkins, E. M. Belding-Royer, S. R. Das, Ad hoc On-Demand Distance Vector (AODV) Routing, RFC 3561 (July 2003).

[46] I. Chakeres, C. Perkins, Dynamic MANET On-Demand (DYMO) Routing, Internet-Draft (work in progress) draft-ietf-manet-dymo-10.txt (July 2007).

[47] C. Intanagonwiwat, R. Govindan, D. Estrin, Directed diffusion: A scalable and robust communication paradigm for sensor networks, in: 6th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2000), Boston, MA, USA, 2000, pp. 56–67.

[48] M. Durvy, P. Thiran, Reaction-Diffusion Based Transmission Patterns for Ad Hoc Networks, in: 24th IEEE Conference on Computer Communications (IEEE INFOCOM 2005), Miami, FL, USA, 2005.

[49] Y. Yu, R. Govindan, D. Estrin, Geographical and Energy Aware Routing: a recursive data dissemination protocol for wireless sensor networks, Technical Report UCLA/CSD-TR-01-0023, UCLA Computer Science Department (2001).

[50] F. Dressler, I. Dietrich, R. German, B. Krüger, Efficient Operation in Sensor and Actor Networks Inspired by Cellular Signaling Cascades, in: 1st ICST/ACM International Conference on Autonomic Computing and Communication Systems (Autonomics 2007), ACM, Rome, Italy, 2007.

[51] B. Krishnamachari, D. Estrin, S. Wicker, The Impact of Data Aggregation in Wireless Sensor Networks, in: International Workshop Distributed Event Based System (DEBS), 2002.

[52] F. Dressler, Network-centric Actuation Control in Sensor/Actuator Networks based on Bio-inspired Technologies, in: 3rd IEEE International Conference on Mobile Ad Hoc and Sensor Systems (IEEE MASS 2006): 2nd International Workshop on Localized Communication and Topology Protocols for Ad hoc Networks (LOCAN 2006), IEEE, Vancouver, Canada, 2006, pp. 680–684.

[53] T. Bokareva, N. Bulusu, S. Jha, A Performance Comparison of Data Dissemination Protocols for Wireless Sensor Networks, in: IEEE Globecom Wireless Ad Hoc and Sensor Networks Workshop, 2004.

[54] A. Boukerche, S. Nikoletseas, Protocols for Data Propagation in Wireless Sensor Networks: A Survey, in: M. Guizani (Ed.), Wireless Communications Systems and Networks, Kluwer Academic Publishers, 2004.

[55] Z. J. Haas, J. Y. Halpern, L. Li, Gossip-Based Ad Hoc Routing, in: 21st IEEE Conference on Computer Communications (IEEE INFOCOM 2002), 2002, pp. 1707–1716.

[56] D. Braginsky, D. Estrin, Rumor Routing Algorithm For Sensor Networks, in: First Workshop on Sensor Networks and Applications (WSNA), Atlanta, Georgia, USA, 2002.

[57] C. L. Barrett, S. J. Eidenbenz, L. Kroc, Parametric Probabilistic Sensor Network Routing, in: 9th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2003), San Diego, CA, USA, 2003.

[58] F. Dressler, R. German, B. Krüger, Adaptive Data Dissemination in Sensor Networks using WPDD, in: Frontiers in the Convergence of Bioscience and Information Technologies (FBIT 2007), IEEE, Jeju, Korea, 2007, pp. 827–832.

[59] J. Luo, P. T. Eugster, J.-P. Hubaux, Pilot: Probabilistic Lightweight Group Communication System for Ad Hoc Networks, IEEE Transactions on Mobile Computing 3 (2) (2004) 164–179.

[60] IEEE, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs), IEEE Sandard 802.15.4-2006 (2006).