Bayerische Julius-Maximilians-Universität Würzburg
Institut für Informatik
Lehrstuhl für Verteilte Systeme
Prof. Dr. P. Tran-Gia

Performance Issues of MAC and Routing Protocols in Wireless Sensor Networks

Alexander Klein

Würzburger Beiträge zur
Leistungsbewertung Verteilter Systeme

Bericht 3/10

Würzburger Beiträge zur Leistungsbewertung Verteilter Systeme

Herausgeber

Prof. Dr. P. Tran-Gia
Universität Würzburg
Institut für Informatik
Lehrstuhl für Verteilte Systeme
Am Hubland
D-97074 Würzburg

Tel.: +49-931-31-86630
Fax.: +49-931-31-86632
email: trangia@informatik.uni-wuerzburg.de

Satz

Reproduktionsfähige Vorlage vom Autor.
Gesetzt in LaTeX Computer Modern 9pt.

ISSN 1432-8801

Performance Issues of MAC and Routing Protocols in Wireless Sensor Networks

Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
der Bayerischen Julius–Maximilians–Universität Würzburg

vorgelegt von

Alexander Klein

aus

Würzburg

Würzburg 2010

Eingereicht am: 25.05.2010
bei der Fakultät für Mathematik und Informatik
1. Gutachter: Prof. Dr.-Ing. P. Tran-Gia
2. Gutachter: Prof. Dr.-Ing. R. Kolla
Tag der mündlichen Prüfung: 16.11.2010

Danksagung

Zunächst möchte ich allen danken, die mich im Rahmen dieser Promotion begleitet, motiviert und unterstützt haben.

An erster Stelle möchte ich hiermit meinen besonderen Dank meinem Doktorvater Prof. Dr.-Ing. Phuoc Tran-Gia aussprechen, der sich zusammen mit mir auf das Abenteuer einer Promotion im industriellen Umfeld eingelassen hat. Seine uneingeschränkte Unterstützung hat es mir selbst in diesem oft marktgetriebenen Umfeld ermöglicht, die aus wissenschaftlicher Sicht interessantesten Themen zu bearbeiten. Ohne seine Hilfe und seinen fachlichen Rat wäre eine derart detaillierte Bearbeitung der Themen nicht möglich gewesen.

Des Weiteren möchte ich Herrn Prof. Dr. Reiner Kolla für das Erstellen des Zweitgutachtens, seine Funktion als Prüfungsvorsitzender und nicht zuletzt für die ansprechenden Diskussionen und Anmerkungen bezüglich der Kommunikationsprotokolle danken. Darüber hinaus gilt mein Dank Prof. Dr. Klaus Schilling, der in seinem stets gut gefüllten Terminkalender zeitnah Platz für meine Disputation gefunden hat.

Für die Hilfe und fachliche Unterstützung möchte ich mich bei Dr. Dirk Stähle und Dr. Rastin Pries bedanken. Ihr ehrliches Feedback und die sich daraus entwickelnden Diskussionen haben wesentlich zu dieser Arbeit beigetragen. In meiner Zeit als Doktorand bei EADS konnte ich viel über die Arbeitsabläufe im industriellen Umfeld lernen. Neben dem fachlichen Know-how konnte ich dort auch zahlreiche positive menschliche Erfahrungen sammeln. Dabei möchte ich meinen ehemaligen Kollegen bei EADS einen besonderen Dank aussprechen: Martin Kluge, Manuel Saez, Jirka Klaue, Sebastian Voss, Johannes Schels, Thilo Fath, Pascal Vermeer, Christoph Aulitzky, Christoph Heller, Dr. Bernhard Schön-

linner, Dr. Sergio Bovelli, Stefan Schneele und Josef Schalk.

Mein herzlichster Dank gebührt Frau Helga Bayer, die mir mit ihrer schnellen, unbürokratischen Unterstützung insbesondere bei zeitkritischen organisatorischen Angelegenheiten das Leben bei EADS sehr erleichtert hat. Das letzte halbe Jahr meiner Promotionszeit durfte ich an der Technischen Universität München am Lehrstuhl für Netzarchitekturen und Netzdienste verbringen. Mein Dank gilt an dieser Stelle Prof. Dr. Georg Carle, der mir insbesondere in der Endphase der Promotion den nötigen Freiraum gegeben hat, die Arbeit zeitnah fertigzustellen.

Neben dem fachlichen Umfeld trägt vor allem das private Umfeld zum erfolgreichen Abschluss einer Promotion bei. Deshalb möchte ich meinen aufrichtigen Dank meinen Eltern Brigitte und Erich Klein sowie meiner Schwester Annette Klein widmen, die mich immer ungefragt unterstützt und motiviert haben. Abschließend möchte ich meinen tiefsten Dank meiner Partnerin Madeleine von Bodisco aussprechen, die mit großem Verständnis Einschränkungen in der gemeinsamen Freizeit akzeptiert hat und meine wichtigste private Stütze in dieser entbehrungsreichen Zeit war.

Contents

1 Introduction

The improved capabilities of wireless sensors have raised the interest in Wireless Sensor Network (WSN) solutions. The higher demand results in lower hardware prices which further increases the number of applications that become economically feasible. Sensor applications, like home automation, have the potential to change our everyday life while others will not be recognized directly. Habitat and agricultural monitoring were one of the first applications which were investigated during the early stages of WSNs. Nowadays, sensor networks for environmental observation and forecasting have become more and more important since recent natural disasters, e.g. earthquakes, tsunamis and floods, have shown that these networks could help to mitigate the impact of these disasters since they provide a sophisticated solution to forewarn the population. Structural health monitoring is the most popular application for sensor networks in the industry. Sensor nodes are used in this field of application to monitor, e.g. pressure, temperature or stress and strain, in order to estimate the remaining lifetime of the monitored structure.

In recent years, sensor nodes are equipped with high data rate wireless interfaces which enable the nodes to transmit multimedia content. A network which is made up of these high speed nodes, is often referred to as Wireless Multimedia Sensor Network (WMSN). WMSNs close the gap between typical WSNs with their limited hardware resources and ad hoc networks. As a consequence of the highly varying hardware limitations and application requirements, communication protocols which are designed for WSNs have to be very flexible and adaptive to provide a high performance in different application scenarios.

1.1 Contribution

The focus of this work lies on the communication issues of Medium Access Control (MAC) and routing protocols in the context of WSNs. The communication challenges in these networks mainly result from high node density, low bandwidth, low energy constraints and the hardware limitations in terms of memory, computational power and sensing capabilities of low-power transceivers. For this reason, the structure of WSNs is always kept as simple as possible to minimize the impact of communication issues. Thus, the majority of WSNs apply a simple one hop star topology since multi-hop communication has high demands on the routing protocol since it increases the bandwidth requirements of the network. Moreover, medium access becomes a challenging problem due to the fact that low-power transceivers are very limited in their sensing capabilities.

The first contribution is represented by the Backoff Preamble-based MAC Protocol with Sequential Contention Resolution (BPS-MAC) which is designed to overcome the limitations of low-power transceivers. Two communication issues, namely the Clear Channel Assessment (CCA) delay and the turnaround time, are directly addressed by the protocol. The CCA delay represents the period of time which is required by the transceiver to detect a busy radio channel while the turnaround time specifies the period of time which is required to switch between receive and transmit mode.

Standard Carrier Sense Multiple Access (CSMA) protocols do not achieve high performance in terms of packet loss if the traffic is highly correlated due to the fact that the transceiver is not able to sense the medium during the switching phase. Therefore, a node may start to transmit data while another node is already transmitting since it has sensed an idle medium right before it started to switch its transceiver from receive to transmit mode.

The BPS-MAC protocol uses a new sequential preamble-based medium access strategy which can be adapted to the hardware capabilities of the transceivers. The protocol achieves a very low packet loss rate even in wireless networks with high node density and event-driven traffic without the need of synchronization. This makes the protocol attractive to applications such as structural health monitoring, where event suppression is not an option. Moreover, acknowledgments or complex retransmission strategies become almost unnecessary since the sequential preamble-based contention resolution mechanism minimizes the collision probability. However, packets can still be lost as a consequence of interference or other issues which affect signal propagation.

The second contribution consists of a new routing protocol which is able to quickly detect topology changes without generating a large amount of overhead. The key characteristics of the Statistic-Based Routing (SBR) protocol are high end-to-end reliability (in fixed and mobile networks), load balancing capabilities, a smooth continuous routing metric, quick adaptation to changing network conditions, low processing and memory requirements, low overhead, support of unidirectional links and simplicity. The protocol can establish routes in a hybrid or a proactive mode and uses an adaptive continuous routing metric which makes it very flexible in terms of scalability while maintaining stable routes. The hybrid mode is optimized for low-power WSNs since routes are only established on demand. The difference of the hybrid mode to reactive routing strategies is that routing messages are periodically transmitted to maintain already established routes. However, the protocol stops the transmission of routing messages if no data packets are transmitted for a certain time period in order to minimize the routing overhead and the energy consumption. The proactive mode is designed for high data rate networks which have less energy constraints. In this mode, the protocol periodically transmits routing messages to establish routes in a proactive way even in the absence of data traffic. Thus, nodes in the network can immediately transmit data since the route to the destination is already established in advance.

In addition, a new delay-based routing message forwarding strategy is introduced. The forwarding strategy is part of SBR but can also be applied to many routing protocols in order to modify the established topology. The strategy can be used, e.g. in mobile networks, to decrease the packet loss by deferring routing messages with respect to the neighbor change rate. Thus, nodes with a stable neighborhood forward messages faster than nodes within a fast changing neighborhood. As a result, routes are established through nodes with correlated movement which results in fewer topology changes due to higher link durations.

1.2 Outline

The structure of this thesis and the relations between different sections are illustrated in Fig. 1.1. The diagram points out that the focus of the thesis lies on three topics in order to describe and to analyze the most relevant communication issues of MAC and routing protocols for WSNs. At the beginning of each chapter, background information is given and related work is discussed. Moreover, the communication challenges and their corresponding solutions in the context of WSNs are emphasized.The identified challenges and solutions are represented by the building blocks in the left column. The building blocks in the center column refer to methodologies and derived mechanisms which are either based-on or inspired by existing solutions. Communication protocols, mechanisms and tools which were developed as part of the thesis are represented by the building blocks in the right column. Thus, these building blocks reflect the scientific contribution of this thesis. Arrows between two building blocks indicate that the described content is closely related. The numbers in parentheses represent the section number of the building block.

The remainder of this thesis is organized as follows. In Chapter 2, different wireless communication issues are discussed from the perspective of WSNs which have specific characteristics similar to other wireless networks.

MAC protocols apply different kinds of strategies to access the medium in an efficient way. Therefore, taxonomy of MAC protocols with respect to the used medium access strategy is given. In addition, the protocols make use of a large number of different algorithms for contention resolution to minimize the collision probability. For this reason, a selection of popular MAC protocols which apply different access and contention strategies is introduced. The efficiency of MAC protocols for WSNs is not only affected by low bandwidth. The turnaround time and the CCA delay of low-power transceivers are identified as performance limitation factors. Both factors are mainly responsible for the low efficiency of the wireless communication in sensor networks in terms of packet loss due to the fact that they limit the capabilities of a node to sense the medium which leads to collisions especially in dense networks with event-driven traffic. These communication issues are addressed by the BPS-MAC protocol which we developed to minimize the packet loss in asynchronous WSNs for Structural Health Monitoring applications. The protocol takes advantage of a new sequential contention resolution algorithm which is based on a slotted preamble transmission. The preamble-based medium access strategy of the protocol and its sequential contention resolution are analyzed and described.

Chapter 3 addresses the topic of routing in WSNs. Routing protocols which are designed for WSNs are optimized for a large number of sensor network specific issues, like high node density and limited hardware resources. The protocols use different mechanisms to establish new routes and to detect topology changes. These mechanisms are discussed in more detail since they are often used to classify the protocols. In addition, a survey of a selection of popular routing protocols, which covers almost the whole spectrum of the presented routing taxonomy, is given. Furthermore, the most essential routing tasks, like forwarding, processing, maintaining a valid topology and dissemination of information, are discussed.

The network topology is one of the key performance issues in WSNs since a non-optimized topology can lead to a shorter lifetime due to a partitioning of the network as a result of unbalanced energy consumption. This kind of performance issue was neglected in the early stages of digital communication where routes were only optimized according to the number of hops. Nowadays, routing metrics often consider more complex link characteristics, like available bandwidth, interference, energy consumption, delay, packet loss or other characteristics of interest. For this reason, the characteristics of popular routing metrics are compared and discussed since they affect the topology in a network.

Some routing protocols apply very interesting mechanisms to establish and maintain routes in a flexible manner, even in the presence of a fast changing topology as a consequence of mobility or unreliable data links. These mechanisms inspired us to develop the SBR protocol which combines several mechanisms in order to achieve high performance in a large number of different scenarios. The protocol and a new delay-based routing message forwarding strategy are introduced and evaluated.

Chapter 4 focuses on the simulated performance evaluation of routing protocols in WSNs. The implementation of routing protocols is time consuming since many protocols come with a huge set of functions which increases their complexity. In addition, the performance of the protocols is influenced by a large number of factors, like the node density, the number of nodes, the signal propagation, the underlying MAC protocol, the data rate of the wireless interface, the sensing capabilities of the transceiver and the mobility of the sensor nodes. Moreover, the utilization of the medium and the characteristics of the traffic pattern have to be considered due to the fact that both factors influence the route establishment process as well as the dissemination of routing information. It is clear that the re-implementation of models for every simulation is not a viable option.

A modular framework was developed to evaluate and compare the performance of the BPS-MAC protocol and the SBR protocol with other communication protocols. The modular framework can be easily extended by other simulations in order to create a co-simulation. Co-simulations are usually built of several simulation tools which are synchronized with each other. Each simulation tool can be executed on a different computer which speeds up the simulation since it may profit from higher computational power. Furthermore, it is possible to connect the simulated virtual network with a real network. This kind of simulation is called hardware-in-the-loop simulation. The idea is to connect applications from the real network through a virtual network which delays and drops packets according to the simulated network conditions. Thus, hardware-in-the-loop simulations enable the user of an application to get direct feedback on the perceived quality of the virtual network. However, hardware-in-the-loop simulation can only be applied to simulated networks which can be run in real-time. Otherwise, the system is not able to delay packets exactly as calculated by the simulator. The combination of the extended modular framework and the video evaluation tool EvalVid allows us to optimize the performance of routing protocols in terms of Quality of Experience (QoE) which reflects the perceived quality of an application.

Finally, Chapter 5 concludes this thesis.

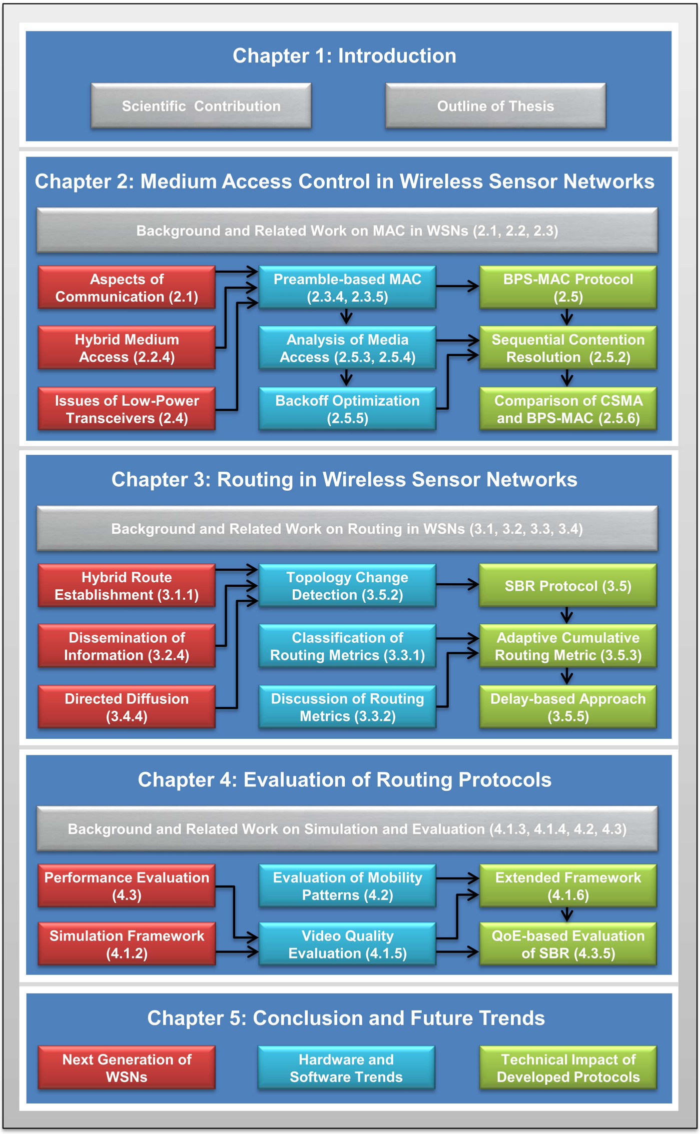


Figure 1.1: *Organization and Contribution of this Thesis*

2 Medium Access Control in Wireless Sensor Networks

MAC protocols for WSNs often have to deal with a large number of specific challenges. First of all, the protocols have to be optimized in respect to computational power and energy consumption. Furthermore, the limited data rate and the long power-up times of low-power transceivers have to be taken into account. Besides the hardware limitations, the characteristics of WSNs have to be considered by the protocol. Communication in WSNs is often unreliable [13–16] due to the low transmission power and the design of the chip antennas [17, 18] which also limits the transmission range in a significant way. Moreover, the traffic in sensor networks is mainly event-driven and thus highly correlated in nature [19]. The combination of highly correlated traffic and limited sensing capabilities of the low-power transceivers makes medium access a challenging task. Sensor nodes are usually built of low cost hardware to minimize the costs of WSNs which are often expensive due to the high number of nodes. The high node density intensifies the problem of medium access in WSNs even further. A solution to this problem is represented by protocols which try to synchronize the sensor nodes in order to schedule the medium access of the nodes. Nonetheless, the synchronization of a large number of nodes is a performance critical task, too. As a result of the high clock drift of the micro controllers, frequent resynchronization is required which consumes a lot of computational power and energy.

Sensor nodes are typically battery powered which makes energy consumption the primary goal of MAC protocols that are designed for WSNs. Therefore, the transceiver of a sensor node is switched off as often as possible to prolong its lifetime. Note that the transceiver is usually the part of a node which consumes far most of the energy. Furthermore, the latest generation of low-power transceivers, e.g. the CC2420 [20] and the CC2520 [21], consumes even more energy in the receive mode than in the transmit mode.

The impact of sleep scheduling becomes clear by taking a look at the lifetime of a node that is not allowed to switch off its radio. In this case, a running state-of-the-art sensor node, like the TmoteSky [22] from Moteiv Corporation that uses the MSP430 [23] micro controller and the CC2420 [20] transceiver from Texas Instruments, will drain a pair of two AA batteries (3000mAh) in approximately 100 hours [24].

Such a short lifetime is only acceptable in a very small number of scenarios. Due to the fact that sensor nodes are mainly used for long term monitoring such as surveillance applications and structural health monitoring, a long lifetime of the sensor nodes is inevitable. In addition, sensor nodes are often placed in areas which are hardly accessible. Thus, the frequent change of batteries is not an option. The discussed example shows that power management is a must for WSNs. However, sleep scheduling always represents a trade-off between latency and throughput. For that reason, the scheduling has to be tuned such that the requirements of the target application are still achieved.

2.1 Aspects of Communication

Besides the introduced aspects of energy consumption six aspects of wireless communication can be specified which are mainly responsible for the limitation of the lifetime and the performance of the sensor nodes. In the following, these communication aspects are discussed in detail. Furthermore, a short description of state-of-art solutions and their impact on the network is given.

Idle Listening

Idle listening represents the most obvious way to waste energy. The term idle listening or idle listening overhead is used for describing the time that a node listens unnecessarily to the medium. It represents the major waste of energy in WSNs since the majority of low-power transceivers consume most of their energy resources while listening to the medium. Typical traffic patterns in WSNs are mainly event-driven [19]. This means that sensor nodes wait for a certain event, e.g. pressure loss, stress, strain, or intruder detection, until they start to generate and transmit data packets. Almost no communication is required during the time between the events. On the one hand, the overall traffic load is very low which makes permanent listening to the medium not necessary. On the other hand, the events are usually restricted to a certain area. As a consequence, only a small number of nodes will recognize a local event. Moreover, the events occur at some indefinite future points in time which makes plan ahead scheduling almost impossible. Thus, all nodes in the network have to sense the medium from time to time in order to be able to forward the event-driven data traffic. For that reason, intelligent mechanisms and strategies are required to minimize the idle listening as much as possible.

Overhearing

Overhearing describes the reception of an incoming packet which is not dedicated and not of use for this particular sensor node. Overhearing may become a serious problem in wireless networks with a high node density [25]. Here, the event-driven traffic plays again a major role. In general, a large number of nodes recognizes and responds to the same event due to the redundancy caused by the high node density in WSNs. Overhearing often occurs when a large number of nodes transmit their sensed values at the same time.

There are three strategies which are mainly followed by MAC protocols to reduce the energy waste caused by overhearing. One possibility is to (locally) aggregate the data traffic [26–28] in order to prevent the forwarding of redundant information. However, this solution requires more intelligent sensor nodes which are able to evaluate and aggregate incoming data traffic. Another way to reduce the impact of overhearing is represented by probability-based event response and forwarding mechanisms [29, 30]. Thus, nodes only transmit their sensed values with a predefined probability. These kind of mechanisms can be used to reduce the number of packets containing redundant information if the number of nodes that will likely sense and respond to a certain event is known in advance. The third strategy minimizes the overhearing by using an event response and data forwarding metric. The metrics used are often based on the reception of redundant data, e.g. a node only responds to an event if less than a predefined number of nodes have responded to the same event [31].

Collisions

Collisions represent the worst-case for communication in WSNs since packets may have to be retransmitted over several hops which increases the energy consumption and the utilization of the medium. In some cases it is possible to recover a collided packet. There is a big chance to recover one of the collided packets if the transmissions were only partially overlapping and the difference of the signal strength is higher than 3 dBm [32]. Different strategies can be followed in order to avoid or to minimize the collision probability.

The most well-known mechanism is CSMA which is used in current Ethernet [33]. Nodes sense the medium before they start their data transmission. Nonetheless, CSMA neglects the time that low-power transceivers require to switch between sensing and transmission. The transceiver is not able to recognize a busy channel during the switching phase. As a consequence, two or more packets will collide if their corresponding nodes have started their data transmission within an interval which is shorter than the switching time. Another disadvantage of CSMA is that it is very energy consuming which makes it only a practical solution in combination with optimized sleep scheduling.

In addition, the hidden-node problem, which is discussed in detail in Subsection 2.3.1, plays a major role in WSNs due to the high node density and the asymmetric transmission area of typical chip antennas [17, 18, 34]. A solution to the hidden-node problem is given by the Ready-To-Send (RTS) / Clear-To-Send (CTS) handshake mechanism which is used e.g. by the IEEE 802.11 standard [35]. The basic idea can be described as follows. The sender and the receiver send out a message in order to inform their neighborhood about their intention to exchange data in the near future. Thus, nodes that receive either the RTS or the CTS message do not transmit any data until the communication has finished. However, the limited sensing capabilities of state-of-the-art low-power transceivers also limit the efficiency of the RTS/CTS handshake mechanism since neighbor nodes can miss one of the handshake messages.

Many protocols try to retransmit packets if they are not successfully received within a certain interval. In this case, the protocols assume a packet loss and thus request a retransmission. From the perspective of the used protocol it is not clear whether the packet was lost due to a collision or simply delayed as a consequence of temporary congestion. Note that a high number of collisions and a high delay are typical signs in WSNs for temporary congestion. Thus, the retransmission of packets may actually degrade the overall network performance since the traffic load is even further increased. In worst-case scenarios, the retransmission of packets will increase the delay and the packet loss such that more retransmissions are triggered. This behavior can lead to a total collapse of the network [24] due to the increased traffic load.

Clustering and synchronization are further strategies to minimize the collision probability. Clustering can be used to scale large sensor networks and to aggregate data which might reduce the overall traffic load. Synchronization provides the basis to allow collision-free medium access scheduling. Furthermore, the synchronization may also reduce the energy consumption if the nodes support sleep scheduling. Nevertheless, synchronization is a challenging task since micro controllers are very limited in their computational power and have a high clock drift. Therefore, resynchronization has to be done frequently which increases in turn the traffic load and the energy consumption. Due to these facts, the synchronization and medium access scheduling are often applied in heterogeneous networks where more powerful and less energy-constraint nodes cover these tasks. Another constraint is given by the complexity of the tasks which makes pre-planned scheduling necessary. For that reason, clustering and synchronization can only be implemented in an efficient way if the network characteristics, e.g. the traffic pattern, the node density, and the energy-constraints, are known in advance.

Traffic Fluctuation

Traffic Fluctuation results from the event-based communication pattern in WSNs. Peak loads are the consequence of the highly correlated traffic which drives the network into temporary congestion. The high number of competing nodes may lead to packet loss or very high delay depending on the used contention resolution mechanism. The usage of a long contention window only disseminates the traffic load to some extend due to the fact that the delay of packets is increased. As a consequence of the additional delay, the MAC protocol may trigger retransmissions which further increase the traffic load and the delay such that even more retransmissions are triggered. For this reason, contention resolution mechanisms for WSNs have to be configured very carefully.

TDMA-based MAC protocols are able to compensate traffic fluctuation as long as the peak traffic load does not exceed the available bandwidth. Thus, the MAC protocol can perform overprovisioning in order to deal with peaks of the traffic load. Nevertheless, the disadvantage of overprovisioning becomes obvious in dense wireless networks with a high traffic load where nodes have to spend most of their time in the idle listening mode which results in a higher energy consumption. Moreover, bandwidth has to be reallocated if the topology of network changes which makes overprovisioning only applicable in wireless networks with a low node density and a stable topology.

Synchronization and priority-based medium access represent feasible solutions to deal with the problem of high traffic load variation. However, both solutions have high demands on the hardware (micro controller and transceiver) and are thus not an option for every sensor platform.

Protocol Overhead

Protocol Overhead should not be neglected in WSNs since it greatly affects the scalability and the performance. The amount of protocol overhead is often larger in WSNs than the application data, due to the small payload that is transmitted by the sensor nodes. This is in contrast to other types of wireless networks like mesh and ad-hoc networks where the data rate and the payload are much higher. A popular strategy to minimize the protocol overhead is to (locally) aggregate the data [26, 27, 36] before a packet is transmitted in order to reduce the number of required MAC headers. Data aggregation often leads to an increase of the delay since packets are accumulated and buffered in a waiting queue before they are transmitted. Nevertheless, data aggregation may also decrease the delay which is an effect often neglected by protocol designers. Due to the fact that nodes have to compete less often for the medium access the protocol overhead is reduced which results in a lower delay in most cases. The protocol overhead can become the dominating and limiting performance factor depending on the peak traffic load and the traffic pattern.

In the case that more powerful sensor nodes are available, it is possible to aggregate incoming data on the one hand, and to filter duplicate data on the other hand. The filtering of duplicate data reduces the energy consumption, the traffic load and the delay. Thus, the filtering of duplicate packets should always be done if possible. Data aggregation and packet filtering mechanisms can be implemented in an efficient way in protocols which already support the clustering of nodes [36] since these protocols already provide the needed infrastructure.

Over-emitting

Over-emitting represents another aspect of wireless communication which increases the energy consumption in WSNs. It describes the event when a node is transmitting a packet to another node which is currently not listening to the radio channel. The impact of over-emitting scales with the node density in the network. However, over-emitting may reduce the lifetime of a WSN in a significant way if the MAC protocol is based on wake-up functions. A large number of MAC protocols uses preambles or busy signals to indicate a future data transmission. Thus, nodes wake up and stay awake for a certain period of time in order to sense the medium. Therefore, over-emitting can increase the duty-cycle of a node which shortens its lifetime. Moreover, over-emitting increases the utilization of the medium. As a result of the higher utilization, the medium access of other nodes is delayed such that they have to listen for a longer period to the medium. Over-emitting can be reduced by synchronizing the nodes in the network [37]. Nonetheless, synchronization is a complex task which requires a significant amount of bandwidth and computational power due to the frequent exchange of synchronization messages. Thus, the designer of MAC protocols have to decide whether the advantages of the synchronization outweighs the required network resources. Another way to reduce over-emitting in duty-cycled networks is represented by sleep schedule exchange [38]. Sleep schedule exchange is a very practical solution in wireless networks with low or medium node density where nodes do not frequently change their sleep scheduling.

2.2 Taxonomy of MAC Protocols

MAC protocols can be classified according to many different performance metrics, e.g. energy consumption, scalability, delay, packet loss or throughput. Nonetheless, the performance of the protocols strongly depends on the hardware limitations and the sensing capabilities of the transceiver. Therefore, the MAC protocols are classified in the following according to their used medium access mechanism.

2.2.1 Random Access

In the early age of wireless communication, transceivers were very limited in their functionality. Thus, the first MAC protocol ALOHA [39] follows the most simple approach where every node is allowed to transmit in a pure random-access manner. The ALOHA protocol was developed by Norman Abramson in 1970 for the ALOHAnet [40] in order to allow packet-based wireless communication. It is clear that this strategy will result in a high collision probability which limits the maximum achievable throughput depending on the traffic pattern. The second generation of MAC protocols took advantage from the reduced switching time between receive mode and sending mode of the transceivers. The short switching time opened the way for CSMA-based protocols in the wireless communication domain. These protocols have in common that they allow to access the medium at any time provided that the transceiver has sensed a free channel. This type of MAC protocols is in the following referred to as random access protocols.

2.2.2 Slotted Access

Sensor networks consist of a large number of nodes which compete for the medium access. Thus, the radio channel represents a resource which is shared by all competing nodes. Many MAC protocols partition the capacity of the channel into smaller units in order to support fairness. The partitioning is often done in the time domain which results in a Time Division Multiple Access (TDMA) scheme. However, other access schemes like Frequency Division Multiple Access (FDMA) and Code Division Multiple Access (CDMA) are also considerable for WSNs. However, TDMA represents the most common solution due to simplicity reasons. There are two groups of protocols that take advantage from partitioning in the time domain. The first group divides the time into intervals which are referred to as slots. Each node is only allowed to access the medium at the beginning a time slot. Therefore, this group of protocols builds the group of slotted access protocols.

2.2.3 Frame-based Access

Frame-based MAC protocols go one step further such that they divide the time into frames containing a fix number of slots. The advantage of the frame-based access lies in the optimization of the communication in terms of throughput, delay, and energy efficiency. In general, the frames start with a short preamble field which allows new sensor nodes to synchronize with the network. Moreover, many protocols use a contention period where nodes can request the access to a slot in one of the next frames. The contention period is then typically followed by an access map which allows nodes to go to sleep if they are not part of a transmission. However, frame-based protocols often require a central node with high performance which covers the tasks of an access point and is responsible for efficient scheduling. Thus, there are high demands on the access point in respect to energy consumption and computational power. The scheduling becomes even more complicated if the nodes transmit with variable data rate or have to communicate with the access point over several hops.

2.2.4 Hybrid Access

Hybrid access MAC protocols attempt to take advantage of the simplicity of the random access and the efficiency of slotted and frame-based access. The protocols transmit a request for a certain slot during a short unsynchronized contention period. The data is then transmitted in the corresponding reserved slot which increases the efficiency of the protocol. Due to the scheduled access, these type of protocols are capable to adapt themselves quickly to changes in the traffic load. All nodes which want to transmit a packet have to send a request during the short contention period which results in a high collision probability of requests in case of event-driven traffic in dense wireless networks.

The performance of hybrid access protocols can be improved if the nodes have some knowledge regarding their future bandwidth requirements. Typical WSN applications like structural health monitoring only need to exchange data during short mission critical time periods, e.g. during the start and the landing phase of a plane. Moreover, nodes do not transmit any data unless the sensed values exceed a predefined threshold. If the values exceed the threshold, the application layer of the nodes generates data packets periodically for a certain interval which represents a constant bit rate traffic. Thus, the nodes are aware of their future bandwidth requirements during the next seconds.

This knowledge can be used to increase the performance of hybrid access protocols by following one of the two presented strategies. Nodes can aggregate several data packets before sending a request in order to reduce the number of requests in the case that the sensed data is not time critical. As a consequence of the smaller number of requests, the collision probability decreases which might also decrease the medium access delay. The second strategy requires more intelligent sensor nodes or a central node which is responsible for the scheduling on the air interface. Some MAC protocols, like the IEEE 802.16 [41] standard, support a feature called bandwidth requests which enables nodes to request a certain amount of data over a short time interval. Thus, the nodes are able to reserve several slots with a single bandwidth request. This mechanism is called “aggregated bandwidth request” and is mainly responsible for the high performance of the IEEE 802.16 MAC protocol.

2.2.5 Polling

The most popular protocol which uses polling as medium access strategy is Bluetooth [42]. It is used by wide range of everyday devices, like headsets, mobile phones, cameras, keyboards, printers, and game controllers, for short range connectivity.

The medium access mechanisms which were described in the previous subsections require either time synchronization or have to rely on the CSMA capabilities of the transceiver in order to schedule the medium access. The polling access strategy does not require synchronization nor a high performance transceiver. A node in the network is only allowed to transmit on the radio channel if it has received the permission to transmit from the master node. Thus, the master node has to transmit an access message to a slave node to allow the slave node to access the medium.

Different kinds of polling strategies exist to optimize the medium access. However, all polling strategies suffer from two major drawbacks. The first one is represented by the high power consumption of the master node which has a high duty-cycle as a consequence of the frequent polling message transmission. In addition, the slave nodes have to listen to requests which are dedicated for other nodes which increases their own power consumption. It is clear that this strategy only achieves a high performance in terms of medium access delay if the node density is low. Otherwise, a large period of time is required to query every single slave node. The third drawback of the conventional polling strategy is that it requires a single-hop network since all nodes must be in transmission range of the master to gain the medium access. Nevertheless, the advantage of simplicity makes polling a very practical solution for short range wireless networks.

2.2.6 Token-Passing

A complete different idea is followed by MAC protocols which make use of the token-passing mechanism. The basic idea is that a node is only allowed to access the medium if it is the current owner of the token. The token is a unique message which is passed from one node to another after the transmission of one or more data packets. The idea of token-passing was mainly used in wired networks. The most popular protocol which is using the token-passing mechanism is the Token-Ring standard IEEE 802.5 [43]. It is also possible to apply the mechanism to a wireless network which enables the nodes to build a logical ring as shown in Fig. 2.1. Due to the shared characteristic of the radio channel, the token-passing

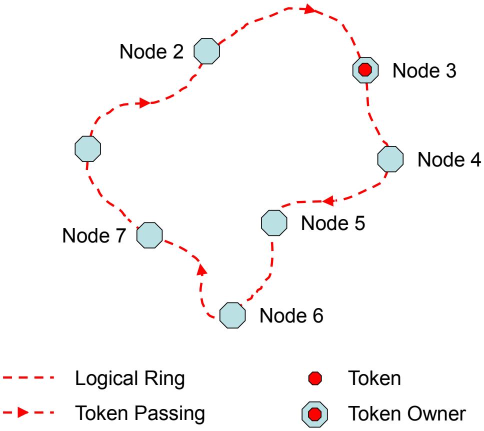


Figure 2.1: *Logical Ring Example*

mechanism requires primitives for joining and leaving a logical ring. The protocol also has to provide mechanisms which recover the network in the case that the token is lost or duplicated. Protocols which are based on token-passing achieve a high throughput if the traffic load is evenly distributed in the network due to the fact that the nodes usually gain ownership of the token according to the round robin principle. The disadvantage of the token mechanism from the perspective of a WSN is that the duty-cycle of the nodes corresponds to the token-passing frequency. The token-passing frequency and the token hold time have a large impact on the medium access delay, especially where the node density is high.

2.3 A Survey on MAC Protocols

The number of MAC protocols is still increasing very fast since most of them are optimized for a certain scenario or application. In the previous sections, the different aspects of communication were discussed and the protocols were classified according to their used medium access scheme. Moreover, the problems of the different access schemes were highlighted and some solutions were introduced. The focus of this section lies on a more detailed description of a selection of popular MAC protocols which apply different medium access schemes and energy saving strategies. The introduced protocols provide a basis for the majority of MAC protocols which are designed for WSNs. In the following, the advantages and the drawbacks of each protocol will be discussed.

2.3.1 CSMA

CSMA is the most popular MAC protocol. It was originally designed for wired networks but was soon recognized as a practical solution for unsynchronized wireless networks. The basic idea of the protocol is to sense the medium before transmitting any traffic on the shared medium. There exist a large number of variations of the protocol. Moreover, the majority of MAC protocols apply a modified version of CSMA in order to schedule the medium access. The different types of CSMA are described in the following paragraphs of this subsection.

p-persistent CSMA

The p-persistent CSMA protocol is almost similar to the 1-persistent CSMA type. The difference between the two types lies in the fact that the nodes only start their transmission with the probability p if the medium becomes idle. The medium access is delayed by one slot duration with a probability of $(1-p)$. In the next time slot the node restarts its medium access procedure. If the medium becomes busy it follows the same procedure as the 1-persistent CSMA protocol and waits a uniform distributed number of slots. Otherwise, it starts its transmission with a probability p. The p-persistent CSMA protocol version achieves a high performance in dense networks. However, the transmission probability has to be adapted to the number of competing nodes in order to optimize the throughput and the medium access delay.

1-persistent CSMA

A node senses the medium if it wants to transmit any data. In the case that the medium is busy, the node keeps on sensing the medium until it becomes idle. Then the node starts its data transmission with a probability one. The node switches its transceiver to receive mode after the end of its transmission in order to sense the medium. If the node senses a busy medium after the transmission, it assumes a collision and waits a random period of time before it restarts the medium access procedure.

non-persistent CSMA

Nodes which use the non-persistent CSMA protocol sense the medium and immediately start their data transmission, similar to the 1-persistent CSMA protocol, if the medium becomes idle. The difference to 1-persistent CSMA-based protocols is that non-persistent CSMA based protocols wait a random period of time if the medium is busy instead of sensing the medium continuously. After waiting a random period, the nodes restart the medium access procedure. The random waiting period helps to spread the traffic load which decreases the number of collisions.

CSMA-CD

The CSMA-Collision Detection protocol became very popular with the outcome of the IEEE 802.3 [33] standard. The basic idea of the protocol is to add mechanisms which allow the detection of collisions. These mechanisms differ depending on the underlying physical layer and the characteristics of the medium. During the early years of the Ethernet protocol, nodes were mainly connected with hubs over half duplex wired links. Thus, the minimum packet size was defined according to the signal propagation delay τ and the transmission delay t_x. A node is able to detect a collision with a transmission of another node, if the medium is busy after it has finished its own transmission. In order to allow recognizing a collision, the minimum packet size has to be chosen such that the transmission delay is longer than twice the signal propagation delay plus the minimum sensing time of the receiver. A drawback of the CSMA-CD protocol is that it can only be applied to wired networks or one hop wireless networks due to the hidden-node problem. The hidden-node and the exposed-node problem will be described in the following paragraph since the CSMA-CA protocol provides several mechanisms to mitigate the impact of both problems.

CSMA-CA

It soon became clear that the CSMA protocol had to be extended to meet the requirements of wireless networks. Wireless networks differ from wired networks due to the fact that nodes are not able to sense the channel during the transmission. Therefore, the CSMA-CD protocol cannot be applied in a practical way since wireless transceivers require a large period of time to switch between receive and transmit mode. Moreover, the hidden-node problem illustrated in Fig. 2.2 has to be addressed in wireless networks.

The hidden-node problem is caused by the limited transmission and sensing range of wireless transceivers. Fig. 2.2 shows a string topology of four nodes A, B, C, and D. The nodes are placed such that they are only able to communicate with their direct neighbors. Thus, node A may only interfere with node B while

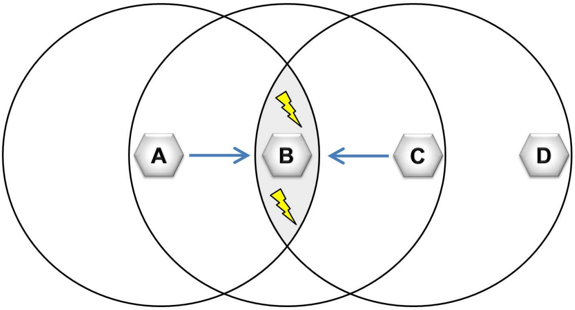


Figure 2.2: *Hidden-Node Problem*

node B may interfere with node A and node C. Let us assume that node A transmits data to node B. Furthermore, node C wants to transmit data to node B, as well. For this reason, it senses the medium and recognizes a free radio channel since it is not able to detect the transmission from node A. Node C will then start its transmission due to the fact that it assumes the medium to be idle. As a result, the transmissions of node A and node C will collide at node B since none of them is aware of the potential competitor.

In addition, the limited transmission and sensing range leads to another problem called exposed-node problem which decreases the overall throughput in wireless networks. Fig. 2.3 shows a typical example scenario of the exposed-node problem. Again, the nodes are placed in a string topology and their transmission and sensing range is limited such that they can only communicate with their direct neighbors. In this scenario, node B continuously transmits to node A. Node C recognizes a certain event and wants to forward the information to node D. Thus, node C switches its transceiver to receive mode and senses the medium. Node C senses a busy medium due to the transmission of node B. Therefore, node C would wait until the medium becomes idle before starting its own transmission. However, the transmission of node C would not interfere with node A since it is out of transmission and interference range of node A. Moreover, the

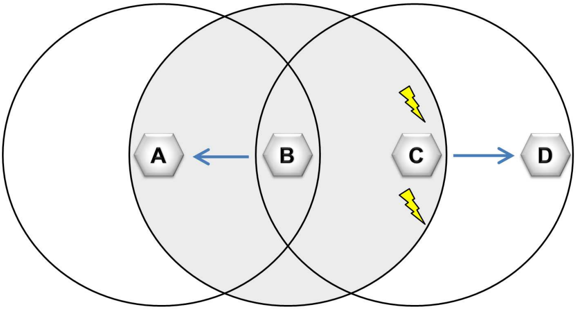


Figure 2.3: *Exposed-Node Problem*

transmission of node B would not interfere with the transmission of node C to node D either. This situation, which limits the overall throughput in a wireless network, is called exposed-node problem. Nonetheless, it is possible to mitigate the problem by using the following collision-avoidance mechanism.

The collision-avoidance mechanism is supplemented by a Request-To-Send (RTS) - Clear-To-Send (CTS) handshake. If a node wants to send a packet, it senses the medium before it transmits the RTS message. The RTS message serves two different functions. The first one is to inform the neighbor nodes of the transmitter about the wish to transmit one or more data packets in the near future. The second purpose is to inform the destination node about the transmission. The destination node responds with a CTS message. Thus, the neighbors of the destination node are aware of the transmission if they have received the CTS message. In addition, the originating node can now be sure that the destination node is ready to receive the transmission. The RTS-CTS handshake solves the hidden-node problem in many cases. However, the handshake cannot guarantee collision-free medium access due to asymmetric links and different transmission ranges of the nodes. Furthermore, the RTS-CTS handshake can only solve the exposed-node problem if the nodes are synchronized. A node may assume that it is an exposed-node in the case that it hears a RTS message from one of its neighbors but does not receive the corresponding CTS message. Nevertheless, the problem may occur that the sender of the RTS message will not receive the corresponding CTS or acknowledgment if the falsely assumed exposed-node starts its own transmission.

CSMA-CR

The CSMA-Collision Resolution protocol follows a different approach which can only be applied to wired networks. Nonetheless, it is possible to modify its medium access mechanism such that it offers interesting possibilities for wireless networks, as well. The basic idea of CSMA-CR is to use bit arbitration. Bit arbitration requires a medium which has a dominant and a recessive state. Furthermore, the nodes in the network have to be able to transmit and receive at the same time.

The collision resolution mechanism is applied by the Controller Area Network (CAN)-BUS [44] which uses an arbitration on message priority medium access scheme. All nodes that want to transmit a message access the bus at the same time after the synchronization bit. The synchronization bit indicates the start of a new frame. The second field is an arbitration field which contains the address of the destination or the originator. Due to the fact that the nodes are connected to the bus via a wired-AND fashion, the nodes are able to pull down the signal on the bus. Thus, a bus level of zero is dominant. A transmitting node compares the state of the bus with its own transmission. The node stops its transmission if it recognizes a difference between the bus level and its transmitted signal. As a result, the message will be send which is dedicated or originated by the node with the highest priority.

CSMA-PS

The traffic load in WSNs is low compared to other wireless networks since nodes sleep most of the time to reduce their energy consumption. For this reason, nodes switch off their transceivers as often as possible since the transceiver usually is the most power-consuming part of a sensor node. Moreover, sensor nodes are often unsynchronized due to the high clock drift of the micro controllers. The CSMA-Preamble Sampling [45, 46] protocol was introduced by El-Hoiydi in 2002. The nodes in the network periodically activate their transceiver in order to listen to the medium for a short time interval. If a node senses a busy channel, it stays awake until the current data transmission has finished. Otherwise, the node switches off its transceiver and waits for the next wake-up interval. Therefore, a node transmits a preamble before its data transmission. The duration of the preamble has to be longer than the wake-up time interval to be sure that the destination node is listening to the medium. A medium access example of the CSMA-PS protocol with acknowledgments is shown in Fig. 2.4.

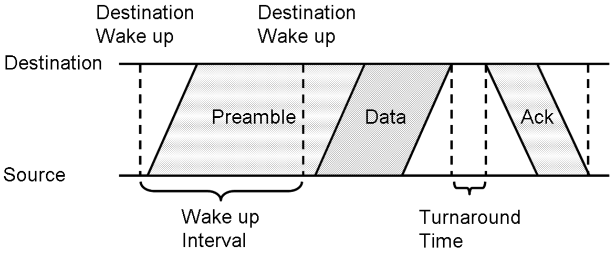


Figure 2.4: *Medium Access Example - CSMA-PS with Acknowledgment*

Acknowledgments are still required and strongly recommended for reliable data exchange due to the fact that hidden-nodes may still interfere the communication. Furthermore, neighbor nodes could also disturb the current transmission if they start their own transmission during the gap between the reception of the last data packet and the transmission of the acknowledgment. The minimum gap duration is represented by the turnaround time of the transceiver. The idea of CSMA with preamble sampling is adopted by a large number of protocols to prolong the lifetime of WSNs. Nonetheless, the performance of CSMA-PS based protocols is strongly affected by the network characteristics, the hardware limitations, and the traffic pattern. Especially, the duty-cycle and the turnaround time have a large impact on the performance of the protocol.

2.3.2 S-MAC

The Sensor-MAC (S-MAC) [38] was introduced by Ye et al. in 2002. It is one of the first protocols that takes advantage from sleep schedules and is specifically designed for the wireless communication of sensor nodes. The primary goal of the protocol is to trade-off energy consumption versus latency while maintaining simplicity. The authors identified idle listening, packet collision, overhearing, and control overhead as the four major sources of energy consumption. They try to optimize their protocol by using periodic listening, sleep scheduling, virtual clustering, collision and overhearing avoidance. Moreover, the authors make some assumptions regarding the communication in WSNs. First of all, they assume that nodes can take advantage in terms of energy consumption from short-range multi-hop communication. Furthermore, it is assumed that nodes are randomly placed and have to be able to configure themselves without the need of previous configuration or additional nodes with less energy-constraints and a higher computational power.

Sleep Scheduling and Synchronization

Each node maintains a sleep schedule table of its neighbor nodes. Before a node joins the network, it needs to choose its own sleep schedule and has to exchange it with its neighbors in order to allow synchronized communication. Thus, the node listens to the medium for a certain interval which has to be longer than the actual duty-cycle of the nodes in the network. In the case that the joining node does not receive any schedule message of one of its neighbors, it chooses a random sleep time and broadcasts a synchronization message with its chosen schedule. This synchronization message holds the duration t in seconds when the node will go to sleep. The joining node covers the function of a synchronizer due to the fact that it has not received any schedule message by other nodes. If a joining node receives schedule information during the joining process, it adapts the schedule and rebroadcasts this schedule after a small random delay t_d. The rebroadcasted message is modified such that the time in the message is set to the difference between the received time t and the chosen random delay t_d. Nodes which adapt the scheduling of another node are referred to as followers. Depending on the configuration of the S-MAC protocol nodes may adapt several sleep schedules such that they are awake several times during a duty-cycle. However, the authors of the protocol propose different mechanisms - which are not discussed in this work - to deal with the problem of multiple sleep schedules. The nodes in the network have to periodically exchange synchronization messages with their neighbors in order to stay synchronized. Synchronization messages are exchanged every tens of seconds which is quite sufficient since the protocol only requires loose synchronization.

Medium Access

The authors propose to adapt the RTS/CTS handshake medium access mechanism of the IEEE 802.11 protocol to minimize the collision probability due to the hidden-node problem. Therefore, the traffic in the WSN can be divided into two types of traffic. On the one hand, the nodes have to transmit synchronization messages which are of high importance to the performance of the protocol. On the other hand, nodes transmit RTS/CTS messages and data traffic. Thus, S-MAC divides the listening period into two intervals as shown in Fig. 2.5. The figure shows the medium access of three different nodes. Node 1 only wants to exchange its periodic synchronization message with the receiver. For this reason, it starts to sense the medium during the synchronization interval of the destination node. The node starts to transmit its synchronization message after a short backoff if the medium is idle. Furthermore, Node 1 may switch off its transceiver after the synchronization interval due to the fact that it has no packets to transmit. Node 2 intends to send data to the receiver. It listens to the medium during the synchronization interval in order to learn about the scheduling of new nodes or to update its sleep scheduling table of the known neighbor nodes. Then, it starts to transmit an RTS message in the second interval which is dedicated for RTS

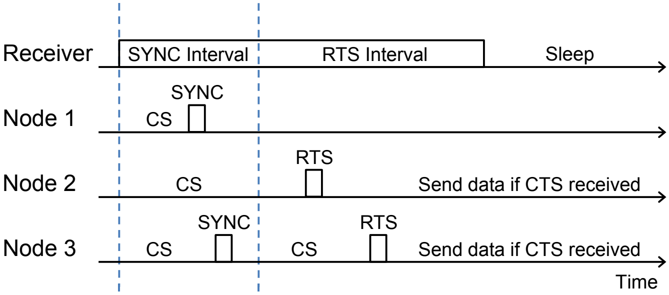


Figure 2.5: *S-MAC - Medium Access*

/ CTS messages and data packets. As mentioned earlier in this subsection, the access to the medium during the RTS interval follows almost the same procedure as the IEEE 802.11 standard. Node 3 wants to perform two activities. First, the node wants to broadcast its sleep scheduling. Thus, it senses the medium and transmits its synchronization message during the synchronization interval. In addition, Node 3 has one or more packets ready for transmission. Therefore, it switches its receiver back to rx mode after the transmission of the synchronization message. After the transmission of the synchronization message, it follows the same procedure as Node 2 in order to transmit the data packets.

Data Transmission

The S-MAC protocol is based on random access. For this reason, multiple nodes may want to send data to a receiver within one RTS interval. The RTS / CTS mechanism is slightly adopted to address the hidden-node problem in an energy efficient way. The protocol attaches a duration field to each packet such that other competing nodes are aware of the remaining duration of the transmission. This feature improves the energy efficiency since the other competing nodes may switch off their transceiver during the transmission. Moreover, S-MAC fragments large packets into smaller ones. In addition, each packet is acknowledged by the receiver. The acknowledgment also covers the function of a CTS message to minimize collisions caused by the hidden-node problem. Thus, smaller packets minimize the collision probability due to the more frequent transmission of acknowledgments.

2.3.3 Sift

Sift [31] is a medium access protocol that is designed for WSNs with a high node density and event-driven traffic. The protocol uses random access and is based on the non-persistent CSMA protocol with a fix size contention window. Most random access protocols use a uniformly distributed backoff to reduce the peak utilization of event-driven traffic which is mainly responsible for collisions in WSNs. The basic idea of Sift is to use a non-uniform backoff distribution in order to sift a winner of the contention resolution if a large number of nodes try to access the medium at the same time. Thus, the protocol divides the time after a transmission into contention slots. Each node which wants to access the medium chooses a random interval between one and a maximum number of slots. A node starts its transmission if the medium was idle during the chosen number of contention slots. In the case that a node senses a busy channel during the backoff, it aborts its medium access procedure and waits for the end of the current transmission.

Collisions may only occur if two or more nodes chose the same number of contention slots. Fig. 2.6 shows a simplified example of the medium access procedure with acknowledgments in which nodes 1 to 4 compete for the medium access in order to transmit to node 0. The short lines which follow each acknowledgment represent the number of contention slots which are chosen by the nodes that want to transmit data. Note that the slot duration is set according to the hardware limitations of the transceiver in order to minimize the collision probability caused by the turnaround time and the CCA delay. Thus, collisions only occur if two or more node select the same number of contention slots.

It is clear that the collision probability increases with the number of compet-

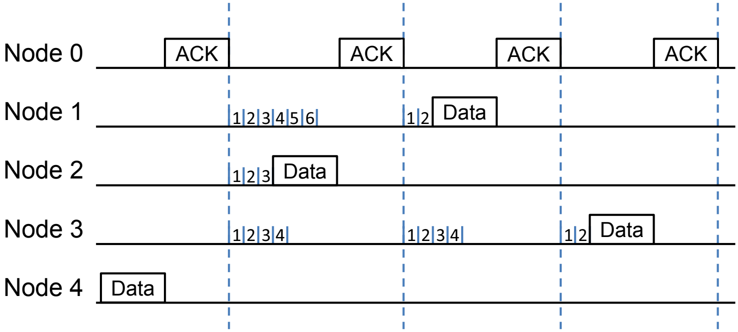


Figure 2.6: *Sift - Medium Access*

ing nodes. For this reason, Sift uses a non-uniform distribution which can be optimized in respect to the number of competing nodes and the available number of contention slots. The contention slots are selected according to the truncated increasing geometric distribution shown in Eqn. 2.1

$$p_r = \frac{(1-\alpha)\alpha^{CW}}{1-\alpha^{CW}} \cdot \alpha^{-r} \qquad for \quad r = 1, ..., CW. \tag{2.1}$$

Variable r represents the selected slot number while p_r is the probability of choosing a backoff duration of r slots. The congestion window which indicates the maximum number of backoff slots is represented by parameter CW. The slot selection probability increases exponentially with r if the parameter α is chosen between 0 and 1. Therefore, the later slots are selected with a higher probability. As a result, the number of nodes which compete for the medium access during the first backoff slots is small which reduces the collision probability in a significant way. Fig. 2.7 shows the backoff distribution of the Sift protocol for a contention window of 32 slots depending on the parameter α. The graphs in the figure indicate that the skewness of the backoff distribution can be modified by varying the parameter α. The skewness of the distribution has to be chosen according to the

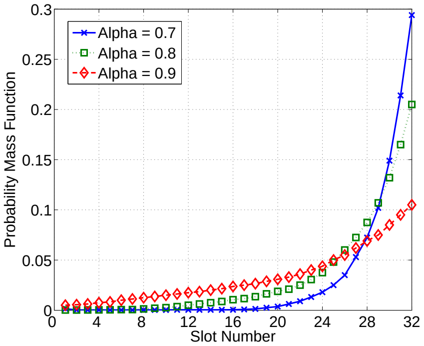


Figure 2.7: *Sift - Backoff Distribution*

number of competing nodes. Moreover, Sift achieves a high success probability of its contention resolution even in the case that the number of competing nodes is underestimated since the majority of competing nodes will choose a high slot number. Thus, the number of nodes that compete for the medium access during the first slots remains small as a consequence of the skewness of the backoff distribution. However, the collision probability becomes very high if the number of competing nodes is underestimated since most of them will choose a high contention slot number.

For this reason, Jamieson et al. recommend to limit the number of nodes which respond to a certain event. They discuss an event suppression mechanism that only allows a predefined number of nodes to respond to a single event. Nodes only try to respond to the event if less than the predefined number of nodes has transmitted a corresponding message. Thus, the number of nodes which try to access the medium at the same time in case of an event remains high such that the Sift protocol achieves a high success probability of its contention resolution.

2.3.4 Wise-MAC

The Wireless Sensor MAC (Wise-MAC) [47] protocol was developed by the Swiss Center for Electronics and Microtechnology as part of the WiseNET platform [48]. The protocol is optimized for energy efficiency in low traffic WSNs. The medium access is based on synchronized preamble sampling. In addition, the protocol is designed for infrastructure communication where more powerful and less energy-constraint nodes cover the task of base stations.

Nodes that are energy-constraint only communicate directly with the base station. In the following, these nodes are referred to as subscribers or subscriber nodes. If a subscriber node wants to transmit a packet to another node, it sends the packet to the base station. The base station transmits the packet to the destination node if the destination node is registered at this base station. Otherwise, the packet is forwarded to the corresponding base station where the destination node is registered.

In infrastructure networks, different MAC protocols and different radio channels can be used for the downlink and for the uplink since a base station will not switch off its transceiver in contrast to the subscriber nodes. Therefore, the downlink - from the base station to the subscriber nodes - represents the challenging part in low-power infrastructure WSNs due to the asynchronous sleep scheduling of the subscriber nodes. Wise-MAC is designed to optimize the downlink in terms of energy consumption and delay. It is based on preamble sampling like many other MAC protocols [45, 46]. However, the difference to other protocols lies in the fact that the base station learns the sampling schedule of its neighbor nodes. Thus, the idle listening time of the subscribers can be reduced if the base station starts to transmit the wake-up preamble in respect to the wake-up period of the corresponding subscriber. The medium access of the Wise-MAC protocol is shown in Fig. 2.8.

Subscriber nodes sense the medium with a wake-up period of T_W. If a base station wants to transmit data to one of its subscriber nodes, it starts to transmit the wake-up preamble right before the wake-up period of the subscriber node.

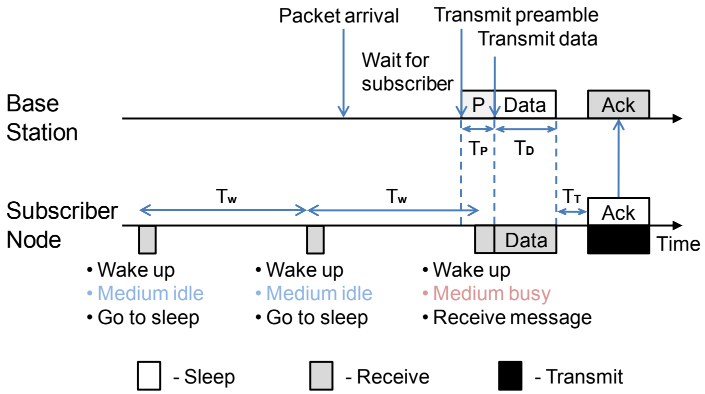


Figure 2.8: *Wise-MAC - Medium Access*

The transmission of a data frame is started as soon as the base station is assured that the subscriber is listening. Note that a frame may contain one or more data packets. The frame starts with the address of the subscriber. Thus, other subscribers can switch off their transceivers in order to avoid idle listening caused by overlapping wake-up intervals. The address field is followed by a data field which holds one data packet. Each frame ends with a frame pending bit to signalize to the subscriber station whether the base station has additional data frames pending for it. As a result, the energy efficiency of the protocol is increased since the subscriber is able to switch off its transceiver as soon as possible. The subscriber node responds with an acknowledgment to the base station in the case that the base station has indicated that no additional frames are pending. The acknowledgment of the subscriber contains the information about the remaining time until the subscriber senses the medium again. This information is then used by the base station to keep its sampling scheduling information table up-to-date. The base station also stores the time when the acknowledgment was received in order to take the clock drift of the oscillator of the micro controller into account.

2.3.5 X-MAC

The X-MAC [49] protocol is designed for asynchronous low-power duty-cycled WSNs. It uses strobed preambles to achieve a better performance than ordinary Low-Power-Listening (LPL) based protocols. The short strobed preambles are used instead of a single large preamble. Moreover, the short preambles contain the address of the destination. Thus, a destination node may recognize its own address immediately and transmit an acknowledgment in the next gap after the preamble which reduces the medium access delay since the originator does not need to transmit all short preambles. Fig. 2.9 shows the difference between the medium access of LPL and X-MAC.

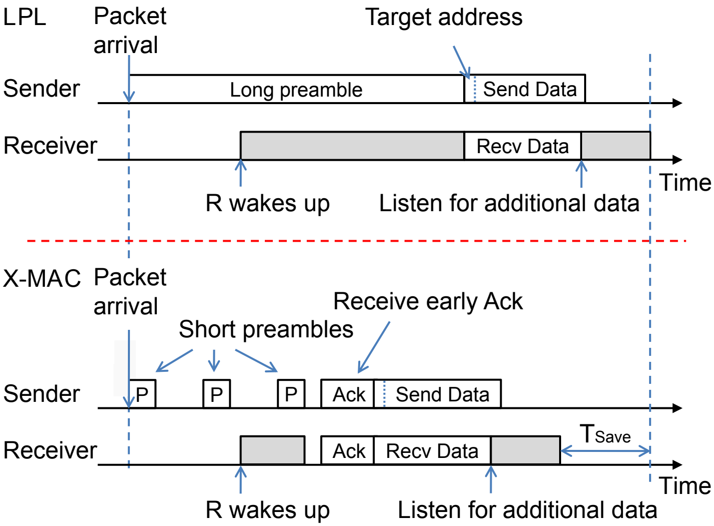


Figure 2.9: *X-MAC - Medium Access*

The advantage of X-MAC over LPL is that the destination node can respond immediately instead of listening to the whole preamble. The originating node stops the preamble transmission and starts its data transmission after receiving the early acknowledgment from the destination node during one of the gaps.

As a result, the medium access delay is reduced by approximately 50% even in the case that there is no contention on the radio channel. The difference may become larger depending on the preamble duration, the traffic load, and the packet size. The efficiency of the protocol depends on the CCA delay and the switching time of the transceiver between rx and tx mode since these hardware limitations are responsible for the length of the short preamble and the duration of the gaps. In addition, the medium access delay is strongly affected by the hardware limitations due to the fact that they also limit the length of the duty-cycle.

The protocol takes advantage from data sniffing. A destination node stays awake a short time after it has received a data transmission. Therefore, it can respond quickly with an early acknowledgment if another node wants to send packets to it. This feature may look unimportant at first glance. However, traffic patterns in WSNs are typically data-centric and event-driven. For this reason, data sniffing significantly affects the performance of the X-MAC protocol. Moreover, the acknowledgment covers the function of a CTS message if received by a node which is not the originator of the preamble. Thus, it reduces the collision probability in multi-hop networks caused by the hidden-node problem. The protocol is able to improve its energy efficiency depending on the traffic load since a node switches off its transceiver if it receives a preamble or an acknowledgment which is not dedicated for it. As a result, the corresponding node safes energy which prolongs its lifetime.

2.4 Communication Issues of Low-Power Transceivers

Typical state-of-the-art low-power transceivers have specific characteristics which affect the performance of WSNs. They are able to transmit data between 32 kB/s and 256 kB/s which limits the possibilities of MAC protocols to exchange information since the number of nodes and the node density are usually very high in WSNs. Therefore, the majority of the MAC protocols which are designed for WSNs rely solely on the sensing capability of the transceiver in order to support random access by using the carrier sense functionality of the chip. The sensing capabilities of low-power transceivers are very limited, especially in the case that small chip antennas are used [50]. As a consequence of the limited sensing capabilities, the packet loss rate in WSNs is very high compared to other wireless networks like IEEE 802.11.

Two communication issues are mainly responsible for the low performance of MAC protocols in WSNs. The first issue is represented by the interval that low-power transceivers require to switch between receiving and transmitting and vice versa. Thus, the switching time which is in the following referred to as turnaround time, specifies the time between the arrival of a packet and the beginning of the corresponding response [51]. During this time interval the transceiver is not able to detect the start of other transmissions.

The second issue is called CCA delay. The CCA delay specifies the interval that a transceiver requires to detect a busy medium provided that the transceiver is already in receive mode. A transceiver is not able to reliably detect the transmission of another node if the transmission has been started within an interval that is shorter than the CCA delay. A closer look is taken on the impact of the turnaround time and the CCA delay on the MAC performance in the following two subsections.

2.4.1 Impact of the Turnaround Time

The turnaround time of transceivers has a direct impact on the efficiency of MAC protocols. However, the impact on the performance depends on the medium access procedure which is used by the MAC protocol. The importance of the turnaround time was first addressed in [52] by Pablo Brenner. In this work, he evaluated the wireless access method and physical specification of the IEEE 802.11 standard. The same topic is discussed in more detail by Johnson et al. [51]

and Diepstraten [53] who describe the effect on the performance caused by several switching aspects. Diepstraten outlines the impact that the turnaround time has on the protocol overhead. The overhead increases especially in the case that a quick mutual exchange of messages, e.g. RTS-CTS messages, data packets and acknowledgments, between the transmitter and the receiver is required. The impact of the turnaround time on a medium access procedure with RTS-CTS handshake and acknowledgment is illustrated in Fig. 2.10. The figure shows the whole data exchange process starting with the recognition of a certain event by the sender. First, the sender has to sense the medium for duration of t_{cs}. The transceiver turnaround time is represented by t_t and the signal propagation delay between the sender and the receiver is represented by t_s. Moreover, we assume that the RTS message, the CTS message and the acknowledgment are of the same size - 8 Byte - in order to simplify the following overhead estimation. The duration which is required to send one of these messages is represented by t_m. The packet transmission time is in the following referred to as t_d. Thus, we can calculate the period of time T which is required by a node to exchange a data packet depending on the CCA delay and the turnaround time according to Eqn. 2.2.

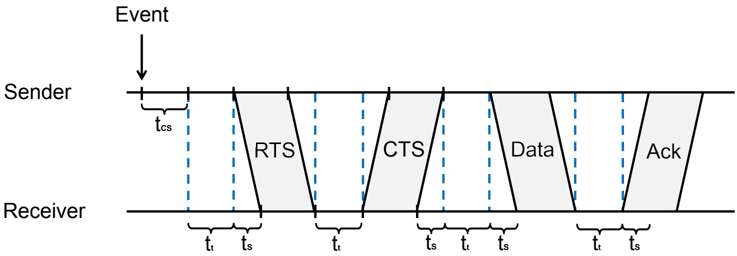


Figure 2.10: *Impact of the Turnaround Time on the Medium Access Procedure*

$$T = t_{cs} + 4t_t + 4t_s + 3t_m + t_d \tag{2.2}$$

Eqn. 2.2 can be further simplified if we assume that the turnaround time has the same duration as the period of time which is required to sense the medium t_{cs}. Note that this assumption does only apply to the CC2420 [20] since the latest generation of low-power transceivers, like the CC1100E [54], require a shorter period of time to switch between receive mode and transmit mode.

Nonetheless, it represents an acceptable approximation for many popular low-power transceivers. The transmission range of wireless sensor nodes is typically very short due to the low transmission power and the characteristics of the small chip antennas. For this reason, t_s is much smaller than t_t and t_m which allows us to remove t_s from the equation. Thus, Eqn. 2.2 can be approximated by Eqn. 2.3

$$T = 5t_t + 3t_m + t_d. \tag{2.3}$$

Fig. 2.11 presents the maximum utilization of the medium access procedure shown in Fig. 2.10 depending on the turnaround time and the packet size provided that the transceiver achieves a data rate of 256 kb/s. The figure illustrates the great impact of the turnaround time and the packet size on the performance on the utilization of the medium. The utilization of the medium may even drop below 20 % if the packet size is smaller than 64 byte and the turnaround time is larger than 128 μs. Furthermore, the figure points out that the turnaround time reduces the utilization of the medium by almost 30 % if typical low-power transceivers try to access the medium according to a RTS-CTS based medium access procedure with acknowledgments. The impact of the turnaround time increases with the data rate of the transceiver as indicated by the results of Fig. 2.12. As a consequence, the turnaround time might become the performance limitation factor for next generation low-power wireless transceivers. In addition, the turnaround time affects the length of the backoff slot duration since the length of a slot should be equal to the total of the CCA delay, the turnaround time, the transceiver power-up time, and the maximum medium propagation delay. Moreover, Brenner [52] and Diepstraten [53] conclude that the turnaround parameter of the IEEE 802.11 physical layer should be stressed as much as possible even if this would limit the number of transceivers which can be used. Note that transceivers with a shorter turnaround

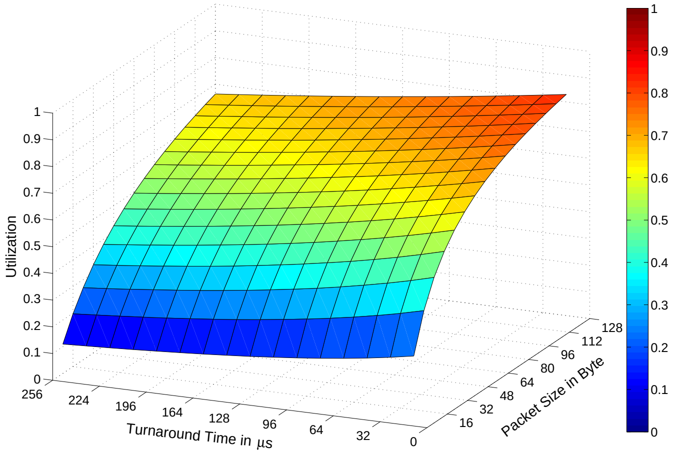


Figure 2.11: *Utilization of the Medium - RTS-CTS based Medium Access Procedure with Acknowledgments using a Transceiver with a Maximum Data Rate of 256 kb/s*

time have to use the same slot duration as those with a longer turnaround time in order to be backward compatible.

2.4.2 The Problem of Clear Channel Assessment Delay

CCA is a logical function which returns the current state of the wireless medium. It is provided by almost any low-power transceiver for WSNs in order to support CSMA functionality to the MAC layer. However, the transceivers require a certain period of time depending on their current state to reliably determine the state of the medium. Moreover, the time that a transceiver requires to switch from receive to transmit mode represents a vulnerable period for MAC protocols which rely on the CSMA functionality since transceivers are not able to detect

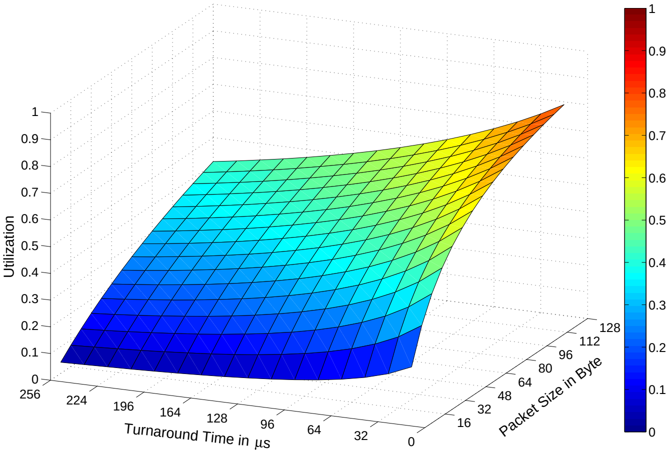


Figure 2.12: *Utilization of the Medium - RTS-CTS based Medium Access Procedure with Acknowledgments using a Transceiver with a Maximum Data Rate of 1Mb/s*

any transmissions that start during the switching period [55, 56].

The CCA delay becomes the dominating performance limitation factor [32] for low-power transceivers which have a relatively high CCA delay compared to IEEE 802.11 transceivers. Typical low-power transceivers, like the CC2400 [20] and the CC2520 [21] (Texas Instruments) or the AT86RF231 [57] (ATMEL), have to listen to the medium for a duration of 8 symbol periods to reliably detect an ongoing transmission. The chips average the Received Signal Strength Indication (RSSI) over the last 8 symbols in order to decide whether the channel is assumed to be busy or idle.

Technical aspects, like the CCA delay of low-power transceivers which have a large influence on the performance of wireless communication in sensor networks, are usually neglected. The problem of CCA delay is only addressed by a small number of papers since standard models from network simulators, e.g. ns-2 [58] or OPNET [59], simplify the physical layer by assuming optimal transceivers which do not need any time to sense the radio channel or to switch between rx and tx mode. The impact of CCA delay on IEEE 802.15.4 [37] networks is described by Kiryushin et al. [32]. The focus of their work lies on real world performance of WSNs and describes the impact of different kinds of communication aspects. Bertocco et al. [60] have shown that the performance of a wireless network can be improved by minimizing the CCA threshold. Nevertheless, the minimization of the threshold requires great knowledge of the radio channel, e.g interference and background noise, since a too small threshold will result in false positives which will significantly decrease the throughput. Thus, nodes will not transmit any data due to the fact that they falsely assume the channel to be busy. The latest generation of low-power transceivers supports different kinds of CCA methods. An intelligent cross-layer approach which takes advantage from different CCA methods is introduced by Ramachandran and Roy [50]. Their idea is to dynamically adapt the CCA method and parameters depending on the current channel conditions and the upper layer parameters.

2.5 Backoff Preamble-based MAC Protocol with Sequential Contention Resolution

The BPS-MAC protocol is optimized for reliability in scenarios with a high node density and highly correlated event-driven data traffic. It does not require synchronization or a large amount of memory which makes the protocol most applicable for sensor nodes with low computational power and limited transceiver sensing capabilities.

In random access MAC protocols all nodes compete for the medium access. A collision occurs if two or more sensor nodes try to access the medium within a time interval which is shorter than the CCA delay of the used transceiver. Backoff algorithms are only able to reduce the probability that two or more nodes access the medium at the same time by spreading the traffic load. Nevertheless, a node can never know whether another node is starting its transmission during the next CCA time interval due to the fact that it cannot listen to the air interface while switching from rx to tx mode.

The BPS-MAC protocol follows a different approach in order to deal with the problem of CCA delay. The basic idea of the protocol is to send a backoff preamble with variable length before transmitting the data. The length of the preamble has to be a multiple of the CCA delay to maximize the reliability of the protocol. Furthermore, the protocol uses a slotted contention resolution due the fact that the backoff preamble is a multiple of the CCA delay. In the case that two nodes send a preamble with different length, the node with the shorter preamble is able to detect the occupation of the medium by the other node.

2.5.1 Single-Sequence Medium Access Procedure

First, a closer look is taken on the single-sequence contention resolution which represents the contention resolution as used by the standard Backoff Preamble-based MAC Protocol (BP-MAC) protocol which we have introduced in [2]. In the following, the term slot is used instead of CCA delay duration since it is more related to the context of contention resolution. Moreover, the term collision probability represents the probability that two or more nodes start their data transmission simultaneously after a backoff transmission which represents an unsuccessful contention resolution.

A node senses the medium for a duration of 3 backoff slots if it wants to transmit a packet. The transceiver is switched from rx to tx in the case that the medium is free for 3 consecutive slots in order to transmit the backoff preamble. The duration of the preamble is chosen according to a uniform distribution between one and a maximum backoff window. The preamble covers the function of a reservation signal. Thus, a longer preamble increases the probability of gaining medium access.

The node senses the medium after the transmission of the preamble. If the medium is busy after the transmission, the node waits between two and maximum backoff window number of slots until it restarts the access procedure described above. Otherwise, the node is allowed to access the medium. Thus, it switches its transceiver from rx to tx mode, which takes a duration of an additional backoff slot. As a consequence, the medium is idle for two slots after the transmission of a backoff preamble. For this reason, a node senses the medium for 3 slots in order to be sure that there is no ongoing contention resolution. The contention resolution in case of synchronous medium access is shown in Fig. 2.13.

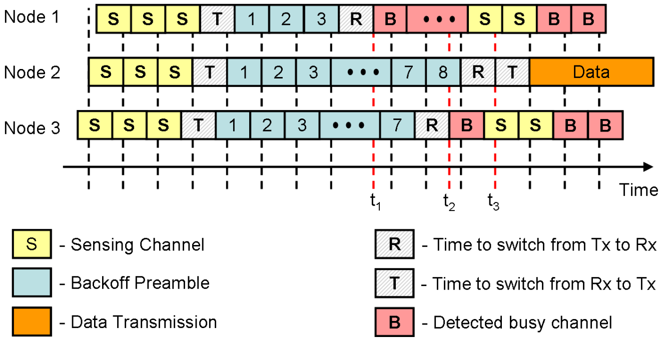


Figure 2.13: *Contention Resolution - Synchronous Access*

A slot is marked with the letter S if the transceiver of a node is able to sense medium for the whole slot duration. In the case that a sensing node recognizes a busy medium in the current slot, the slot is marked with a B. Numbers are used to indicate the corresponding backoff slot of the transmitted preamble. The letter R is used to illustrate that a node switches its transceiver from transmit to receive mode. The slot duration has to be chosen such that a node is able switch the transceiver mode and to detect a busy medium within a single slot duration. The letter T indicates that a node switches its transceiver from rx to tx mode. It is assumed that the node is not able to detect a busy medium during this time slot.

Collisions may still occur, as mentioned earlier in this section. However, collisions only occur if two or more nodes start to access the medium within one slot and choose the same backoff preamble duration. An example of a collision is shown in Fig. 2.14.

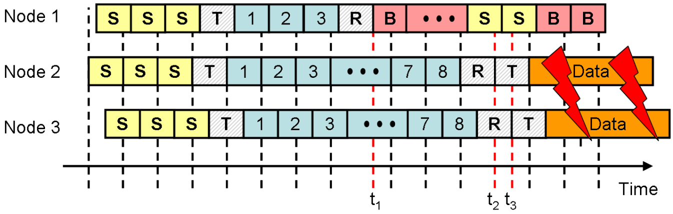


Figure 2.14: *Contention Resolution - Collision*

The figure indicates that the collision probability can be reduced if a longer backoff preamble is used provided that the number of competing nodes is independent from the backoff duration which is typically not the case.

2.5.2 Multiple-Sequence Medium Access Procedure

The single-sequence medium access procedure can be improved if multiple short backoff preambles are used instead of a single long preamble. However, the CCA delay and the turnaround time of the transceiver have to be taken into account if multiple backoff sequences are used. The proposed sequential contention resolution procedure requires a gap of two backoff slots between two consecutive sequences as shown in Fig. 2.15 and Fig. 2.16. The gap is needed in order to switch the transceiver from tx to rx and to sense the medium.

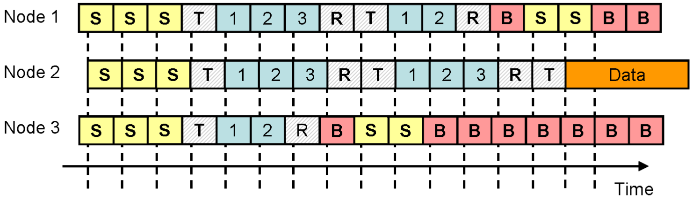


Figure 2.15: *Sequential Contention Resolution - Synchronous Access*

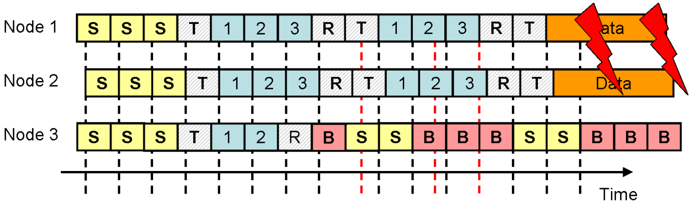


Figure 2.16: *Sequential Contention Resolution - Collision*

In the following, the medium access process of the BPS-MAC protocol shown in Fig. 2.17 is described in more detail. Higher layer packets are put into the waiting queue if a backoff is already pending. In the case that no backoff is pending the medium access process is started. The protocol initializes a sequence counter and an access counter. Note that the access counter is used to count the number of free consecutive backoff preamble slots while the sequence counter represents the number of transmitted backoff preambles. Furthermore, the access counter starts with an initial value of zero in contrast to the sequence counter which starts with an initial value of one. After the initialization of the counters is completed, the protocol switches the transceiver to receive mode and starts to sense the medium. If the medium is busy, the node waits between zero and EBW slots before the medium is sensed again. In addition, the access counter is set back to zero. The transceiver might be switched off during the waiting period depending on the

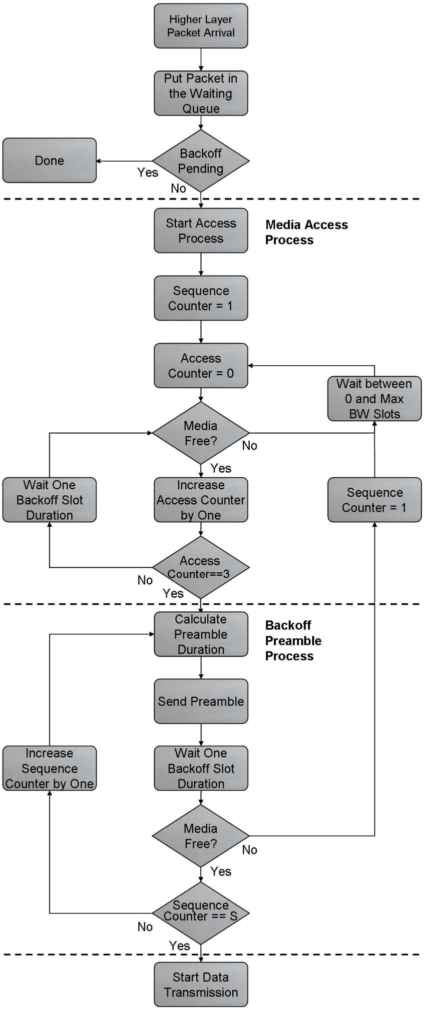


Figure 2.17: *Flow Diagram of the Medium Access of the BPS-MAC Protocol*

energy-constraints of the node. The access counter is increased by one if the medium is idle and checks whether the counter is equal to 3 which indicates that the medium has been idle for duration of 3 consecutive backoff slots. If the value of the access counter is smaller than 3, the protocol waits one backoff slot until it follows the procedure described above. The protocol calculates the preamble duration depending on the sequence counter and starts to send the backoff preamble after the medium has been idle for duration of 3 backoff slots. This mechanism allows the modification of each backoff sequence, e.g. a different EBW size or a different backoff distribution. Then it switches the transceiver back to receive mode, which requires the duration of one backoff slot. If the node senses a busy channel after the preamble transmission, it resets the access and the sequence counter and waits between 0 and EBW slots before it senses the medium again in order to restart the access process. In the case that the medium is idle after the backoff transmission, the node checks whether the sequence counter has reached the maximum number of backoff sequences S. If the value is smaller than S, the counter is increased by one and the preamble process is started again. The node is allowed to start its data transmission if the medium is idle after the transmission of S backoff preambles.

2.5.3 Single-Sequence - Analysis of the Contention Resolution

The probability that two or more nodes access the medium within one slot duration depends on the traffic load and the traffic characteristics. For this reason, every scenario has to be analyzed individually if the reliability of a WSN has to be calculated in advance. Nevertheless, it is possible to calculate the collision probability if the number of nodes which transmit synchronously and the maximum backoff preamble duration is known. Thus, the achievable reliability can be calculated for worst-case scenarios if we assume that all nodes in the WSN start the medium access procedure at the same time.

First of all, we have to specify a collision from probability calculus point of view. A collision of two or more transmissions occurs if two or more nodes select the same number of backoff slots provided that they have selected the highest number of slots in this contention resolution phase.

We have shown in [2] that the number of nodes which are part of a collision in a certain backoff preamble sequence can be calculated as follows. Let *m* be the number of nodes which access the medium at the same time while the maximum number of backoff slots is *n*. The number of nodes which are part of a collision is *c* which corresponds to the value of the discrete random variable C. The probability that the contention is resolved is given by $P(C = 1)$. Moreover, $b(i)$ is the probability that a node transmits a backoff preamble with a duration of *i* slots.

Thus, we can formulate the probability mass function of random variable C according to Eqn. 2.4

$$P(n,m,c) = \begin{cases} \dfrac{m!}{c!(m-c)!} \displaystyle\sum_{i=1}^{n-1} b(i+1)^c \left(\sum_{k=1}^{i} b(k) \right)^{m-c} & 1 \le c < m \\ \displaystyle\sum_{i=1}^{n} b(i)^c & c = m \\ 0 & otherwise. \end{cases} \tag{2.4}$$

Eqn. 2.4 can be simplified to Eqn. 2.5 if we assume that the nodes select each backoff slot with a probability of $b(i) = \frac{1}{n}$ according to a uniform distribution.

$$P(n,m,c) = \begin{cases} \dfrac{m!}{c!(m-c)!} \left(\dfrac{1}{n} \right)^m \displaystyle\sum_{i=1}^{n-1} i^{m-c} & 1 \le c < m \\ \left(\dfrac{1}{n} \right)^{m-1} & c = m \\ 0 & otherwise \end{cases} \tag{2.5}$$

The probability that the transmissions of two or more nodes are part of a collision

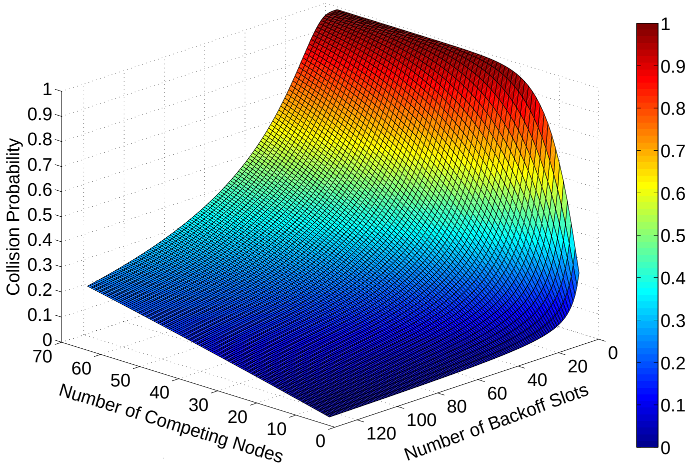


Figure 2.18: *Collision Probability for Simultaneous Medium Access depending on the Number of Competing Nodes and the Number of Backoff Slots*

during a single backoff sequence can be calculated according to Eqn. 2.7 by using the completeness axiom from Eqn. 2.6

$$\sum_{c=1}^{\infty} b(c) = 1, \tag{2.6}$$

$$P(n,m,C>2) = 1 - P(n,m,C=1) = 1 - \sum_{i=1}^{n-1} \frac{m \cdot i^{m-1}}{n^m}. \tag{2.7}$$

Fig. 2.18 shows the results of Eqn. 2.7 in order to give a better impression of the impact that the backoff duration and the number of synchronously transmitting nodes have on the collision probability. The results show that the single-sequence contention resolution is able to reduce the collision probability in a significant way depending on the number of available backoff slots. However, we should keep in mind that the contention resolution reduces the possible overall throughput since no data can be transmitted during the preamble transmission. Thus, the utilization of the air interface and the traffic pattern have to be taken into account when choosing the maximum number of backoff slots.

If the signal strength of two simultaneous transmitting nodes differs by more than 3 dBm [32], the packet with the higher signal strength is received correctly while the other packet is lost due to bit errors. Nevertheless, we focus on the worst-case scenario in which both packets are disturbed such that the bit errors cannot be corrected. In order to calculate the mean packet loss rate a new discrete random variable T is introduced which represents the number of lost packets during one contention resolution period. The random variable T is calculated as described in Eqn. 2.8

$$P(T=t) \;=\; \begin{cases} \dfrac{m!}{t!(m-t)!}\left(\dfrac{1}{n}\right)^{m}\displaystyle\sum_{i=1}^{n-1} i^{m-t} & 2 \leq t < m \\ \left(\dfrac{1}{n}\right)^{m-1} & t = m \\ 0 & otherwise. \end{cases} \tag{2.8}$$

Thus, we are able to calculate the mean of the packet loss due to collisions during the contention resolution according to Eqn. 2.9

$$\begin{aligned} E[T] &= \sum_{t=2}^{m-1} t\,\frac{m!}{t!(m-t)!}\left(\frac{1}{n}\right)^{m}\sum_{i=1}^{n-1} i^{m-t} + m\left(\frac{1}{n}\right)^{m-1} \\ &= \frac{m}{n}. \end{aligned} \tag{2.9}$$

2.5.4 Multiple-Sequence - Analysis of the Contention Resolution

The fact that the mean of the number of lost packets during a single contention phase is described by the fraction of n and m encouraged us to think about a new sequential backoff resolution called BPS-MAC which is described and analyzed in the following paragraphs of this subsection. Short consecutive backoff preambles are able to reduce the number of competing nodes step by step. Therefore, just a small number of nodes will compete for the medium access in the last backoff preamble sequence. The overhead increases with the number of backoff sequences. For this reason, the duration of backoff sequences can only be shortened to some extend, which becomes obvious by taking a look at the proposed sequential contention resolution procedure shown in Fig. 2.15 and Fig. 2.16.

Competing nodes switch their transceivers to rx after the transmission of the first backoff preamble. If they sense a busy channel, the nodes abort their current medium access process and wait between zero and EBW slots before sensing the medium again. In the case of an idle channel, the nodes switch their transceivers back to tx and transmit the next backoff preamble. Thus, collisions may only occur if two or more nodes start their medium access procedure within one CCA delay interval and choose the same number of backoff slots in every backoff sequence. The maximum duration of each backoff preamble sequence and the number of sequences defines the maximum medium access delay for a single contention resolution.

Let y be the medium access delay in number of backoff slots and s the number of sequences. The EBW of sequence i is denoted as n_i. Furthermore, it is assumed that a node senses the medium for 3 slots before it switches its transceiver to tx mode - which requires an additional duration of one slot - in order to start the backoff preamble transmission. The maximum medium access delay can be calculated according to Eqn. 2.10 provided that the gap between two consecutive preambles is two slots.

$$y \leq 4 + \sum_{i=1}^{s} n_i + 2s \tag{2.10}$$

The minimum medium access delay is achieved if a node chooses the first backoff slot in every backoff sequence while no other node is competing for the medium access. The lower bound of the medium access delay is given by Eqn. 2.11

$$y \geq 4 + 3s. \tag{2.11}$$

Many applications for WSNs need guarantees in terms of maximum medium access delay since the generated data is often mission-critical. In the following it is assumed that a certain number of nodes have to transmit a small amount of data if they recognize an event. Thus, the maximum allowed medium access delay in number of backoff slots can be calculated assuming that the amount of data per node, the transmission rate, and the maximum number of nodes responding to an event are known in advance. The BPS-MAC protocol can be easily optimized for a particular application in the case that the maximum allowed medium access delay is known. First, the boundaries of the number of backoff sequences have to be specified according to the maximum allowed delay. The boundaries of the number of backoff sequences can be calculated according to Eqn. 2.12 provided that the smallest allowed value of the EBW is two slots.

$$1 \leq s \leq \frac{y}{4} - 1, s \in N \tag{2.12}$$

The next question which has to be answered is that of defining the length of each individual backoff sequence. Let *n* be the number of available slots for the preamble transmission for a predefined number of backoff sequences *s*. Thus, *n* can be calculated according to Eqn. 2.13

$$n = \sum_{i=1}^{s} n_i. \tag{2.13}$$

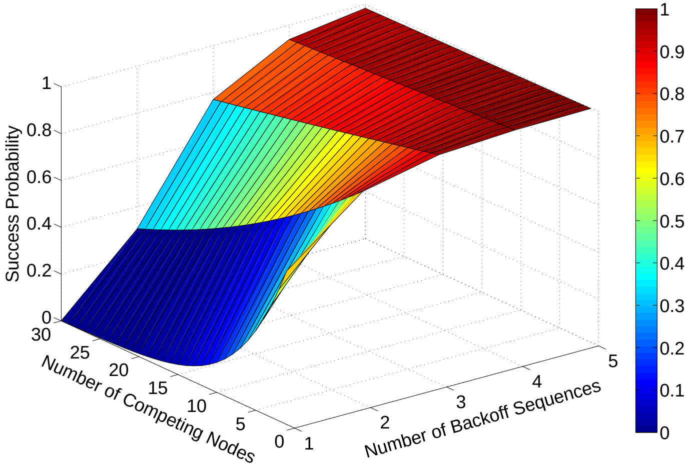


Figure 2.19: *Uniform Distribution - Probability of Successful Contention Resolution depending on the Number of Competing Nodes*

By taking a look at Eqn. 2.9 it becomes clear that the duration of the backoff sequences n_i have to be chosen such that $\prod_{i=1}^{s} n_i$ is maximized. Therefore, the duration of each backoff sequence should be chosen as short as possible in order to maximize the product. However, the number of available backoff slots *n* depends on the number of backoff sequences if an upper limit for the maximum allowed medium access delay *y* is given. The highest probability of a successful contention resolution is achieved if n is a multiple of four. Due to the gap between two consecutive backoff sequences, a length of 4 slots represents the best trade-off between overhead and success probability. In the following it is assumed that *n* is always a multiple of 4.

Fig. 2.19 shows the probability of successful contention resolution as a function of the number of competing nodes and the number of backoff sequences. The results shown in Fig. 2.19 are very promising, especially in the case that three or more backoff sequences are available. However, in some cases it is only possible to use up to two sequences in order to maintain within given boundaries of the medium access delay.

2.5.5 Backoff Optimization

The collision probability can be minimized even in the case in which only a small number of short backoff sequences can be used due to medium access delay boundaries. A solution for this optimization problem is given by Tay et al. [61] which evaluated the performance of non-uniform distributions for slotted contention resolution. They introduced an algorithm which calculates the optimum distribution for a given number of backoff slots provided that the number of competing nodes is known in advance. However, their optimized solution only achieves a high success probability if the number of competitors does not differ much from the assumed number of competitors. Thus, they recommend to use a truncated geometric distribution which is less affected by the number of competitors.

Fig. 2.20 shows the probability of successful contention resolution for a single backoff sequence with maximum backoff duration of four slots depending on the used distribution from Table 2.1 and the number of competitors. The results of Fig. 2.20 suggest that the uniform distribution represents the best choice if only two nodes compete for the medium access. Nonetheless, the probability of successful contention resolution decreases rapidly with the increasing number of competitors. Therefore, the uniform distribution is not the first choice for networks with high node density and correlated traffic. The truncated geometric distribution can be optimized for a fix number of competitors. However, the optimization of the distribution for more than three nodes is not really practical since the performance degrades significantly if the number of competing nodes is overestimated. This behavior becomes clear by taking a look at the different dis-

Table 2.1: *Distributions - Backoff Slot Selection*

Distribution / Probability	Slot 1	Slot 2	Slot 3	Slot4
Optimimized_3	0.534	0.217	0.148	0.101
Optimimized_8	0.766	0.086	0.078	0.070
Optimimized_16	0.884	0.040	0.039	0.037
Uniform	0.250	0.250	0.250	0.250

tributions shown in Table 2.1. The optimized distributions achieve a high success probability of the contention resolution due to their skewness. As a consequence of the skewness, the majority of the competing nodes choose one of the first slots while the minority of the nodes compete in the rest of the available slots. This explains the high success probability of the optimized distributions even in the case that the number of competitors is underestimated. Nevertheless, the skewness reduces the success probability if the number of nodes is smaller than the number for which the distribution is optimized. The number of competing nodes

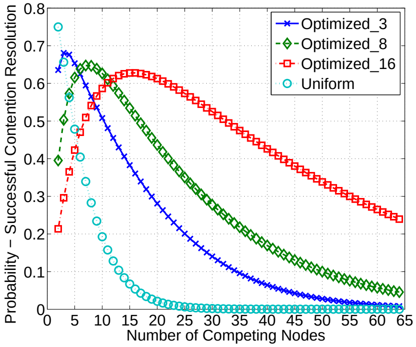


Figure 2.20: *Probability of Successful Contention Resolution for a four Slot Backoff Procedure depending on the Number of Competing Nodes*

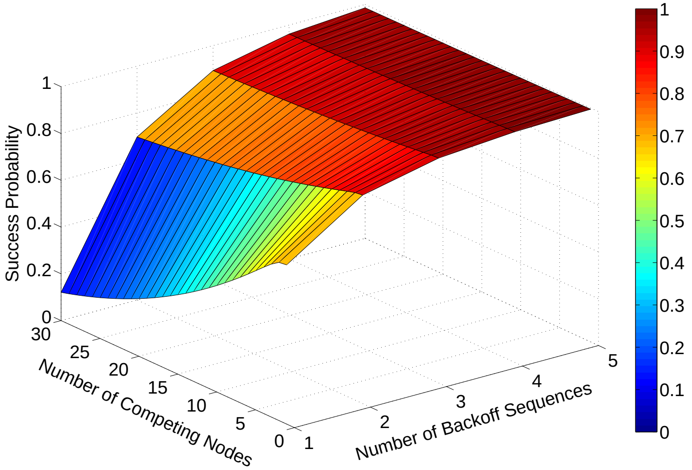


Figure 2.21: *Optimized_3 - Probability of Successful Contention Resolution depending on the Number of Competing Nodes and the Number of Backoff Sequences*

always decreases from a maximum - which depends on the node density and the traffic pattern - to one. Therefore, an optimized distribution for three competing nodes represents the best choice for most scenarios. Recall that an underestimation of the number of competitors only has a small impact on the packet loss in contrast to an overestimation which increases the collision probability in a significant way. The success probability of the Optimized_3 distribution for an EBW of four depending on the number of competing nodes and the number of backoff sequences is shown in Fig. 2.21.

The results of Fig. 2.21 point out that the success probability for the Optimized_3 distribution for the first two backoff sequences is much higher compared to the uniform distribution shown in Fig. 2.19. However, the difference becomes smaller with an increasing number of backoff sequences. The uniform distribution even represents a slightly better solution if the number of sequences exceeds four and the number of competing nodes is less than 32. This behavior is the consequence of the stepwise reduction of the number of competing nodes. The probability that only two nodes compete in a single backoff sequence increases with an increasing number of backoff sequences. Due to the fact that the uniform distribution is the optimum distribution for the case of two competing nodes, its performance increases more with the number of backoff sequences than the performance of the Optimized_3 distribution.

Now that the impact of the backoff distribution on the successful contention resolution is discussed, a closer look is taken on the average number of collisions per backoff. A distribution may achieve a higher success probability than another distribution but it may have a higher packet loss if the average number of nodes which are part of a collision is higher. Thus, the question is, how many nodes are part of a collision in case of an unsuccessful contention resolution. Moreover, the impact of the backoff distribution and the number of backoff sequences on the collision probability has to be evaluated.

Let c_0 be the number of nodes which compete for the medium access in the first backoff preamble sequence and c_i the number of nodes that collide in the i*th* sequence. Moreover, n_i represents the number of backoff slots in the i*th* backoff sequence while s represents the number of backoff sequences. The function $p(var_1, var_2, var_3)$ is an extension of Eqn. 2.4 whereas the parameters n, m, and c are freely configurable. Variable var_1 corresponds to the maximum number of backoff slots n while variable var_2 represents the number of competing nodes m. The number of nodes c which are part of a collision is indicated by var_3. Thus, the average number of nodes which are part of a collision after s backoff sequences can be calculated according to Eqn. 2.14 by using Eqn. 1

$$\begin{aligned} E[C,1] &= \sum_{c_1=2}^{c_0} c_1 p(n_1,c_0,c_1), \\ E[C,2] &= \sum_{c_1=2}^{c_0} \sum_{c_2=2}^{c_1} c_2 p(n_1,c_0,c_1) p(n_2,c_1,c_2), \\ \vdots \quad & \qquad \vdots \qquad\qquad\qquad \vdots \\ E[C,s] &= \sum_{c_1=2}^{c_0} \cdots \sum_{c_s=2}^{c_{s-1}} c_s p(n_1,c_0,c_1) \cdots p(n_s,c_{s-1},c_s). \end{aligned} \tag{2.14}$$

Figures 2.22-2.25 show the average number of collisions per backoff for the optimized truncated geometric distributions for 3 and 8 competing nodes as well as for the uniform distribution depending on the number of 4 slot backoff sequences. Furthermore, the Opt3_Uniform graphs represent the results of a hybrid approach where the Optimized_3 distribution from Table 2.1 is used in the first sequence while the uniform distribution is used for the consecutive sequences.

The first thing which can be mentioned for the results of the single backoff sequence shown in Fig. 2.22 is that the average number of collisions per backoff increases linearly with the number of competing nodes for the uniform distribution. The uniform distribution only offers the best performance for two competing nodes while the Optimized_3 distribution represents the best solution for three to 10 competing nodes. If the number of competing nodes exceeds 10 the Optimized_8 distribution shows a better performance. It is interesting to notice that the Optimized_8 distribution does not achieve the lowest packet loss for 8 competing nodes though its success probability is optimal for 8 competing nodes. The answer is given by the Optimized_8 distribution function. Due to the high probability of the first slot there is a noticeable probability that all nodes choose the first backoff slot in one sequence. Therefore, the average number of collisions is increased in a significant way. Fig. 2.23 shows that the average number of collisions

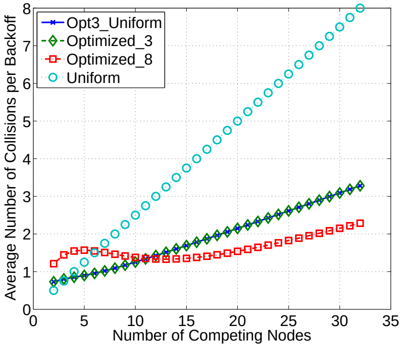


Figure 2.22: *Single Backoff Sequence Contention Resolution - Average Number of Nodes which are Part of a Collision depending on the Backoff Distribution*

can be approximately quartered if the BPS-MAC protocol uses two consecutive backoff preambles to resolve the contention. This effect can be recognized for the uniform, Optimized_3 and Opt3_Uniform distributions. The performance of the Optimized_8 distribution does not represent a good solution for scenarios with less then 32 competing nodes. As a consequence of its heavy-tailed characteristic, the probability is high that less then 10 nodes compete for the medium access in the second backoff preamble sequence. Thus, there is a big chance that the remaining competitors collide in one of the first slots in the second sequence.

If the BPS-MAC protocol uses 3 consecutive backoff preamble sequences the performance of the Optimized_8 distribution degrades even more which is shown by the results of Fig. 2.24. The highest reliability for scenarios with more than three competitors is achieved by the Opt3_Uniform approach. The Optimized_3 distribution reduces the number of competing nodes in the first sequence such that uniform distribution becomes the best choice for the consecutive backoff sequences. For applications with low or medium data rates and less constraints

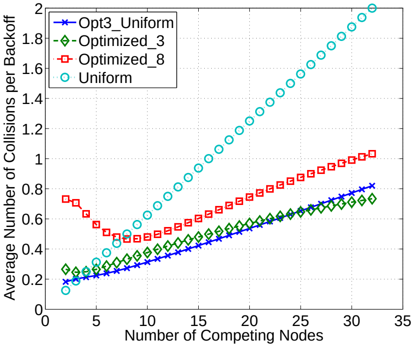


Figure 2.23: *Two Backoff Sequences Contention Resolution - Average Number of Nodes which are Part of a Collision depending on the Backoff Distribution*

regarding the delay, it may be a considerable choice to use four or more consecutive backoff preamble sequences. The Opt3_Uniform approach should be used for these scenarios which is indicated by the results of Fig. 2.25.

2.5.6 Simulative Performance Comparison of the CSMA and the BPS-MAC Protocol

In this subsection the reliability and the delay of the CSMA and the BPS-MAC protocol are simulated under different conditions, e.g. the number of competing nodes, the utilization of the air interface, and the traffic pattern. Due to the fact that the BPS-MAC protocol is designed for WSNs, the focus lies on networks with data-centric traffic pattern and highly correlated traffic. The simulated scenarios distinguish in the traffic load and the number of traffic sources in order to get a better picture of the performance of the BPS-MAC protocol.

In all scenarios, the sensor nodes are in transmission range of each other and the signal strength between the nodes differs by less then 3 dBm. Moreover, all

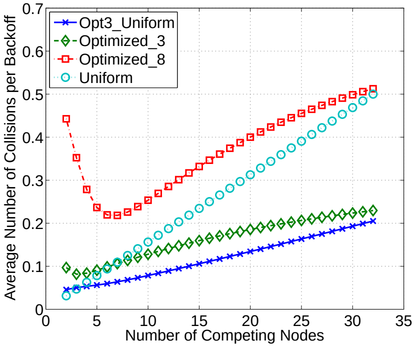


Figure 2.24: *Three Backoff Sequences Contention Resolution - Average Number of Nodes which are Part of a Collision depending on the Backoff Distribution*

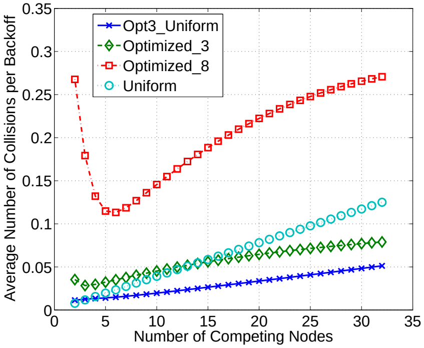


Figure 2.25: *Four Backoff Sequences Contention Resolution - Average Number of Nodes which are Part of a Collision depending on the Backoff Distribution*

packets of simultaneous transmissions are lost which represents the worst-case for collisions [32]. The nodes are positioned according to a uniform distribution within a square of 10 by 10 meters. Three different configurations of the protocol are simulated. The first two configurations of the BPS-MAC protocol use a backoff resolution with 3 backoff sequences. The duration of each sequence is four slots which results in a maximum duration for the first contention resolution of 22 slots. The maximum backoff window is set to 32 slots to spread the traffic load after the first contention resolution in case of an unsuccessful medium access. The first configuration uses the uniform distribution for the contention resolution and is in the following referred to as BPS-4SEQ3-Uni. The second configuration uses the Optimized_3 distribution and is thus abbreviated as BPS-4SEQ3-Opt_3.

In contrast to the previous configurations, the third configuration uses a single-sequence backoff with uniform distributed slot selection. The duration of the EBW and the maximum backoff window are set to 32 slots to get comparable results. The abbreviation BPS-32SEQ1-Uni is used for this configuration.

Due to the fact that the CSMA protocol is the most common protocol in WSNs, its performance is compared with the performance of the BPS-MAC protocol using the three specified configurations. The simulated CSMA protocol uses a Truncated Binary Exponential Backoff Algorithm (TBEBA) to resolve contention on the channel. The duration of a CSMA backoff slot is set to 30.51 μs which corresponds to the 32 kHz clock cycle of the micro controller. The TBEBA uses a SBW and an EBW of 9 to calculate the number of backoff slots. The number of backoff slots is then chosen uniformly distributed between zero and two to the power of the current backoff window. The algorithm increases the backoff window if a busy medium is detected after the current backoff is transmitted which indicates an unsuccessful medium access. The configuration of the CSMA-TBEBA protocol was chosen from a large set of simulated scenarios with more than 10 nodes and data-centric event-driven traffic pattern. In addition, the simulation results were verified by measurements in a testbed.

If not further mentioned, a CCA delay of 128 μs is assumed, which represents the Received Signal Strength Intensity (RSSI) average time of 8 symbol periods.

The data rate of the transceiver is set to 256 kb/s and the backoff slot duration of the BPS-MAC protocol is set to the duration of the CCA delay. The results are calculated from 20 simulation runs with different seeds which is quite sufficient due to the low variance of the collected statistics. Each simulation was run for 1100 seconds and the statistics were collected after a transient phase of 100 seconds. The traffic model started to generate packets after 50 seconds and stopped 10 seconds before the end of the simulation in order to allow the transmission of all waiting packets. The simulated traffic patterns are shown in Table 2.2. The delay results represent the average of the 99 % quantile while the reliability plots show the mean reliability of the simulation runs. Due to the fact that the variance of the collected statistics was very low, the results show only the mean instead of the corresponding confidence intervals.

Table 2.2: *Traffic Pattern*

Pattern Name	Parameter	Distribution	Range / Values
Burst	Burst IAT	uniform	[9.9995; 10.0005] s
	Packets per Burst	constant	3
	Packet IAT	uniform	[0.0000; 0.0010] s
	Packet Size	constant	1024 bit
	Number of Sources	-	[10;20;30;40;50; 60;70;80;90;100]
CCA Delay	Packet IAT	uniform	[0.0950; 0.1050] s
	Packet Size	constant	1024 bit
	Number of Sources	constant	10

Performance of the Protocols depending on the CCA Delay

The BPS-MAC protocol was designed to successfully resolve contention on the medium independent from the sensing capabilities of the transceiver. In fact, the probability of a successful contention resolution is not directly affected by the CCA delay. Nonetheless, the backoff slot duration of the BPS-MAC protocols has to be chosen with respect to the duration of the CCA delay and the period of time which is required by the transceiver to switch between receive and transmit mode. The maximum of both factors represents the minimum threshold of the backoff slot duration. As a result, a higher CCA delay will usually increase the duration of contention resolution since the CCA delay is often higher than the switching time of the receiver. Therefore, the probability increases that more nodes compete for the medium access during a single contention resolution. However, it is clear that the impact of this effect varies according to the utilization of the medium, the number of competing nodes, and the traffic pattern.

In the following, a closer look is taken on the reliability and the delay of the protocols depending on their configuration and the duration of the CCA delay. Typical transceivers for WSNs, like the CC2420 or the AT86RF231, have a high CCA delay between 128 μs and 192 μs.

For this reason, the reliability and the delay of the protocols are simulated while increasing the CCA delay from 32 μs to 256 μs in steps of 32 μs in order to evaluate the potential of the next generation of low-power transceivers. The scenario consists of 10 source nodes which transmit traffic to a data sink according to the burst traffic pattern described in Table 2.2.

Fig. 2.26 shows the reliability of the protocols depending on the CCA delay. The reliability of the CSMA-TBEBA protocol decreases linearly from 99 % to 96 % if the CCA delay is increased from 32 μs to 256 μs. This result is no real surprise since the protocol solely relies on the sensing capabilities of the transceiver. Nevertheless, the reliability seems only to be high on the first look since the number of nodes is relatively small and the utilization of the air interface is approximately 40 %. The reliability of the BPS-MAC protocol with 3 backoff sequences and the uniform distributed backoff slot selection is less affected by a higher CCA delay. However, its reliability decreases from 99.8 % to 99.4 % as a consequence of the higher number of competing nodes per contention resolution due to the higher CCA delay. The same behavior can be recognized for the BPS-32SEQ1-Uni configuration. The higher CCA delay increases the number of competing nodes per contention resolution which increases the collision probability. The Optimized_3 distribution enables the BPS-MAC protocol with 3 backoff se-

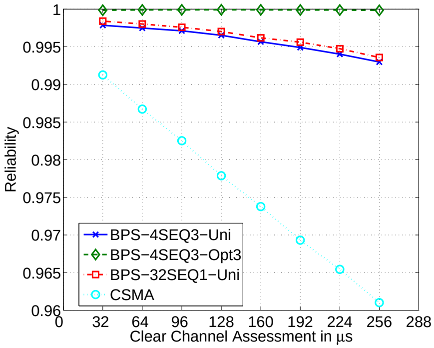


Figure 2.26: *Reliability depending on the CCA Delay of the Transceiver*

quences to achieve a reliability of almost 100 % in most of the simulation runs. The high reliability is the consequence of the stepwise contention resolution and the Optimized_3 distribution which significantly minimizes the number of competing nodes already in the first sequence. Thus, the last two sequences can be used for the remaining competing nodes.

The results of Fig. 2.27 show that the delay of the CSMA protocol is not affected by the CCA delay. The delay of the protocol remains nearly constant since its backoff slot duration is independent from the CCA delay. The very slight decrease of the delay results from the lower reliability in scenarios with a higher CCA delay. The lowest delay in all scenarios is achieved by the BPS-4SEQ3-Uni configuration. The slope of the curve reveals that the BPS-MAC protocol takes a large benefit from transceivers with a low CCA delay. The non-linear increase of the delay again results from the higher number of competing nodes due to the longer contention resolution caused by the higher CCA delay. The BPS-4SEQ3-Opt3 and the BPS-32SEQ1-Uni achieve a lower delay compared to the CSMA protocol as long as the CCA delay remains below 224 μs. A high increase of the delay of the BPS-4SEQ3-Opt3 configuration can be recognized if the CCA

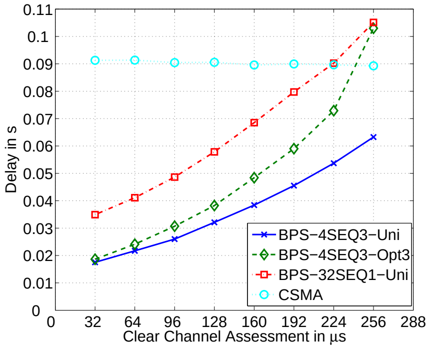


Figure 2.27: *99 % Quantile of the Delay depending on the CCA Delay of the Transceiver*

delay reaches 256 μs. The delay of the BPS-4SEQ3-Opt3 is more affected by the higher number of competitors due to the skewness of the optimized truncated geometric distribution. The delay of the protocol becomes very low if the CCA delay is below 192 μs which is achieved by many low-power transceivers.

Performance of the Protocol depending on the Number of Sources

The results presented in the previous section have shown that the number of competing nodes affects the reliability and the delay of the protocols in a different way. Therefore, the focus of this section lies on the performance of the protocols depending on the number of competing nodes. The number of source nodes is increased from 10 to 100 in steps of 10 nodes in order to simulate the impact of the number of competitors. In addition, the nodes send traffic according to the burst traffic pattern shown in Table 2.2. Thus, the nodes transmit 3 packets approximately every 10 seconds which results in high peaks of the utilization. The burst traffic pattern is similar to the event-driven traffic of structural health monitoring applications. The CCA delay and the slot duration of the BPS-MAC protocol are set to 128 μs which reflects the CCA delay of typical low-power transceivers, like the CC2420 (Texas Instruments) and ATMELs AT86RF231.

Fig. 2.28 shows the impact that the number of competing nodes have on the reliability of the protocols to successfully transmit a packet. The reliability of the CSMA protocol varies between 94.8 % and 92.4 %. It achieves the highest success probability for the scenario with 10 source nodes. Then the reliability drops sharply to 92.6 % if 20 nodes send traffic periodically to the sink. However, it is interesting to notice that the reliability increases again to 93.6 % if the number of nodes is increased to 100. The increase is the consequence of the large SBW and EBW which spreads the traffic load such that the collision probability is decreased to some extend. Moreover, the probability that a node finds the medium occupied increases with a higher number of sources nodes. Therefore, the medium access becomes more deterministic. The BPS-32SEQ1-Uni and the BPS-4SEQ3-Uni are affected in a similar way. Nevertheless, their reliability remains on a high level between 98.8 % and 99.4 % independent from the number of competing nodes. The reliability of the BPS-4SEQ3-Opt3 is less affected by the number of source nodes. A reliability of 99.9 % was achieved even in the 100

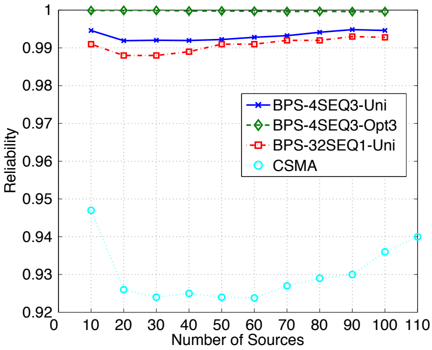


Figure 2.28: *Reliability depending on the Number of Burst Sources*

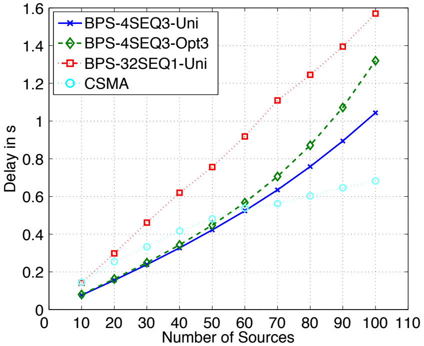


Figure 2.29: *99 % Quantile of the Delay depending on the Number of Burst Sources*

node scenario. Note that the average number of nodes which access the medium within one CCA delay interval is far below 100 even in scenarios with correlated traffic. Therefore, the BPS-MAC protocol is able to achieve this very high success rate with only 3 backoff sequences. A fully synchronized medium access would not generate any meaningful results for the CSMA protocol.

The results of Fig. 2.29 indicate that the delay of the protocols is affected in a different way by the number of source nodes. The BPS-4SEQ3 configurations achieved the lowest delay if the number of sources is less than 60. The lowest delay in scenarios with more than 60 nodes is achieved by the CSMA protocol. However, the price of the lower delay is paid by the higher packet loss rate of the protocol. Note that acknowledgments and retransmissions would increase the delay of the CSMA protocol beyond the delay of the BPS-MAC protocol. The highest delay can be recognized for the BPS-32SEQ1-Uni configuration. The reason for its high delay compared to the other BPS-MAC configurations lies in its very long backoff preamble. As a consequence of the sequential backoff resolution, the probability is high that nodes detect a busy medium after the transmission of the first short preamble sequence. Therefore, the delay of the BPS-MAC protocol configurations with multiple backoff sequences is lower than the delay of the single backoff configuration. The delay of the BPS-4SEQ3-Opt configuration increases more than the delay of the other protocols if the number of source nodes exceeds 80. This increase results from the skewness of the Optimized_3 distribution. The duration of the backoff sequences increases due to the higher number of competing nodes. Nonetheless, the delay is still acceptable for most applications since 99 % of the packets arrive within 1.3 seconds. Note that the 95 % quantile of the delay - which is not plotted - remains lower than one second even for the 100 source node scenario.

2.6 Summary

In this chapter, a closer look was taken on typical communication aspects of WSNs. Several aspects were identified which are mainly responsible for the energy consumption and the performance limitation. Furthermore, communication strategies were introduced which mitigate the impact of each aspect on the network performance. Additionally, different classifications of MAC protocols were discussed. The protocols are then classified according to their used medium access mechanisms due to the fact that the medium access mechanism represents a characteristic which does not depend on network parameters, e.g. the number of nodes or the node density. The taxonomy of the MAC protocols is followed by a survey of state-of-the-art MAC protocols for WSNs. However, the number of MAC protocols for WSNs is almost countless [62]. Thus, the presented survey only represents a small selection of MAC protocols which are optimized for WSNs. The protocols were selected with respect to their used medium access procedures and energy saving strategies. Moreover, they are representative for a large number of WSN MAC protocols since they provide the basis for many medium access strategies.

Most MAC protocols neglect communication issues caused by low-power transceivers. For this reason, the impact of the turnaround time and the CCA delay on the network performance were discussed. These issues become the limiting performance factors in dense wireless networks with event-driven correlated traffic patterns. We encountered the problem caused by a long turnaround time and a high CCA delay in a structural health monitoring project in which we attached sensor nodes to a metal plate in order to measure the stress of the plate. The nodes were configured such that they transmit their sensed values to a central node for a short interval if the sensed values exceeded a certain threshold. Therefore, they started to transmit almost synchronously if the plate was stressed due to the fact that they were placed within a small area. Furthermore, we relied on the CSMA capabilities of the transceiver. Thus, the nodes were allowed to access the medium if the CCA pin was indicating a free medium. As a consequence of the synchronous transmission approximately 15 % of the packets were part of a collision. Due to the fact that a high packet loss is unacceptable for real-time monitoring applications, we started to develop the BPS-MAC protocol which was introduced in Section 2.5. The BPS-MAC protocol directly addresses the communication issues caused by low-power transceivers. It can be easily configured to achieve an extraordinary high reliability for point-to-point communication without the need of using acknowledgments. In addition, the protocol will take more advantage of next generation low-power transceivers compared to CSMA-based protocols since its performance improves with shorter CCA delays.

3 Routing in Wireless Sensor Networks

Sensor networks have high requirements on the routing protocol due to the high node density and the limited hardware resources. Scalability represents an important issue in WSNs since the nodes have to solve the problem of routing table explosion [19]. Beside the problem of scalability, routing protocols have to deal with a large number of other challenges which arise from the characteristics of WSNs.

First of all, links in WSNs are often unreliable [13, 14] and not stable as a consequence of the low transmission power, the chip antenna design, and the fact that the nodes are often placed randomly distributed on the ground [63]. Furthermore, link breaks may be caused by e.g. asynchronous sleep times, interference, moving obstacles, energy exhaustion, node failure or mobility [64].

Thus, the topology of the network changes frequently which represents a challenging problem for the routing protocol design since the sensor nodes are also very limited in their computational power and memory. Moreover, the available bandwidth of the low-power transceivers and the high node density of WSNs have to be taken into account when designing a routing protocol.

The primary goal of routing protocols which are designed for WSNs is to maintain energy efficient and reliable paths between different nodes in the network without generating a lot of overhead. Note that sustaining a route from a source to a destination may consume more bandwidth than is required to support the data traffic flow. In order to optimize the communication, it is important to know the characteristics of the traffic in advance.

Some protocols [65–67] can be configured to achieve good performance in different scenarios. However, the capability of the protocols to adapt themselves to different network characteristics often comes at the price of increased complexity. For this reason, the majority of the protocols is optimized for a certain network architecture and application with a characteristic traffic pattern in order to maximize the performance and the lifetime of the network.

As a consequence, the number of routing protocols is still increasing quickly since many researchers and developers modify existing protocols or design new protocols to meet the requirements of their network. In the following, different taxonomies are discussed to give a better overview of the large number of protocols and their characteristic routing mechanisms. Moreover, the strategies and characteristics of routing protocols are introduced in Section 3.1 to allow a further classification.

The protocols have to fulfill certain tasks, e.g. route establishment, route maintenance, route repair, packet forwarding or dissemination of routing information, which are independent from the strategy that is followed by the routing protocol. These tasks are described in more detail in Section 3.2.

The topology which is generated by the protocols mainly depends on the used routing metric. Thus, a closer look is taken on different routing metrics and their impact on the network performance in Section 3.3. Furthermore, a short survey of a selection of popular routing protocols is given in Section 3.4. The protocols were selected with respect to their characteristic routing mechanisms in order to highlight the different routing approaches.

Then, the Statistic-Based-Routing protocol is introduced in Section 3.5 which outperforms a large number of popular routing protocols while maintaining simplicity. We designed the protocol to meet the requirements of low-power WSNs. Thus, the focus was laid on energy efficiency, low overhead, high reliability and simplicity. The key features of the protocol are its continuous routing metric and its flexibility. The protocol supports different routing strategies and can be optimized for a wide range of applications. Finally, the chapter is concluded with a summary.

3.1 Classification of Routing Protocols

Routing protocols have a large number of characteristics, e.g. the routing technique, the route establishment procedure, and the protocol operation, which can be used for detailed classification. Often routing protocols are also classified according to the generated network structure [68] since the network structure directly affects performance parameters, like scalability and overhead. The routing protocols are then categorized into flat, hierarchical, and location-based protocols.

Another possibility is to classify the routing protocols with respect to the protocol operation. The protocol operation can be multi-path-based, query-based, negotiation-based, QoS-based, or coherent-based. In this monograph, the classification according to the protocol operation is adopted from [68] since it represents a very clear way to classify the routing protocols.

Routing protocols which are designed for mesh and ad hoc networks are usually categorized according to the way they establish routes in the network [69]. This kind of classification is also very practicable for WSN routing protocols. There exist three different categories. The first one is called proactive. Proactive protocols try to establish and maintain routes before they are needed. The second category is represented by reactive protocols which follow the contrary approach where routes are only established or computed on demand. The last group consists of hybrid protocols which combine the ideas of reactive and proactive route establishment. Fig. 3.1 gives an overview of the routing taxonomy which is used in this monograph.

3.1.1 Route Establishment

Routing protocols can follow different strategies to enable connectivity between the nodes in the networks. The used strategy has a large impact on the performance and the lifetime of the network. In the following, a closer look is taken on the three different route establishment strategies. Moreover, their performance under various conditions is discussed.

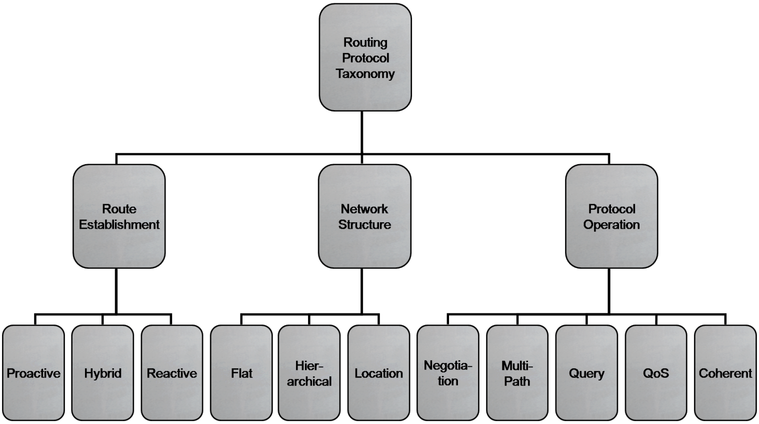


Figure 3.1: *Taxonomy of Routing Protocols*

Reactive

Reactive protocols [65, 70–72] try to establish a route between two nodes in the network if data has to be transmitted. Thus, they are usually the first choice for low data-rate networks with a dynamic topology since they do not generate any routing overhead in the absence of traffic [73]. Nonetheless, this strategy comes with certain disadvantages, e.g. high delay due to the fact that a new route has to be established before it is possible to start the data transmission. In addition, reactive protocols should only be used in static or low mobile networks since they are not able to detect link breaks as quickly as their proactive counterparts [74].

The majority of reactive routing protocols, like the Ad hoc On-Demand Distance Vector (AODV) [65] protocol, use timeouts in order to detect broken links. Furthermore, most of them support a local repair feature to minimize the routing overhead in case of link breaks. Note that the routing overhead of reactive protocols may become higher than the routing overhead of proactive protocols if they have to frequently re-establish routes as a consequence of node failures or link breaks [73]. The protocols assume a link to be broken if no data has been received from a neighbor along the route for a certain duration. Data-traffic in WSNs is usually event-driven and highly correlated. Thus, the medium access delay can become very high in the case that a large number of nodes try to access the medium at the same time. There exist a large number of strategies on different layers to mitigate the impact of synchronous medium access [31, 75, 76]. In the case that the delay becomes higher than the timeout interval, reactive protocols will try to repair the existing route or establish a new one since they assume the route to be broken. For this reason, the protocols have to be configured carefully to minimize the number of false positives. However, a large timeout interval decreases the capability of the protocols to detect links breaks which might lead to a very low performance.

The traffic pattern has a large impact on the performance of protocols which follow a reactive strategy. Routes are usually removed from the routing table of the nodes after a certain time interval in order to minimize the memory consumption. The protocols will establish a new route for every packet if the packet inter-arrival time is longer than the route timeout. Most of the protocols use flooding mechanisms to establish the routes. Thus, the routing overhead may become even a multiple of the data traffic.

Proactive

Proactive routing protocols [4, 67, 77, 78] are most common in ad hoc and mesh networks since these networks have higher demands on the delay compared to WSNs. The advantage of proactive route establishment is that a lower delay can be achieved due to the fact that all routes are established in advance. Nonetheless, this kind of protocols should only be applied if there are no constraints regarding the energy efficiency and the available bandwidth since they periodically transmit routing information [8]. It is clear that the energy consumption depends on the duration of the transmission interval. Shorter intervals increase the routing overhead and the energy consumption.

Another advantage of proactive protocols is that they are able to quickly detect link breaks as a consequence of the frequent network probing. However, the consequence of frequent probing is high routing overhead which has to be taken into account especially in large networks. Thus, scalability represents a serious issue for proactive routing protocols. For this reason, proactive routing protocols are mainly used in hierarchical or cluster-based networks to minimize the routing overhead which is caused by the dissemination of routing information. Nevertheless, proactive strategies can also be applied to sensor networks [79] if the number of nodes which forward routing information is limited. Furthermore, the broadcasting frequency can be reduced due to the fact that WSNs have fewer requirements regarding the delay and the detection of broken links. Moreover, the protocols may provide an attractive alternative to reactive protocols [80] if the periodic transmission of routing information is constricted to short time intervals during which communication is required.

Hybrid

Hybrid protocols [4, 81–85] combine different routing mechanisms and techniques. Therefore, they do not fit into a single category. In the following, the term hybrid protocol is used for protocols which use both reactive and proactive mechanisms to maintain existing routes or to establish new routes. The majority of hybrid protocols can be divided into two groups.

The first group does not transmit any routing information if no route is required. Only the source and the destination of an active route periodically transmit routing information in order to keep the existing routes up-to-date. The advantage of this approach compared to reactive strategies is that the routing protocol is able to quickly detect link breaks in active routes. In addition, the frequent probing of the network allows the protocol to keep track of other paths between the source and the destination which can be used immediately if the primary route becomes unavailable. Thus, this kind of protocols achieve a low delay without generating a lot of overhead due to the fact that nodes stop broadcasting routing information if they are not further part of an existing route. The initiating packet has a higher delay than consecutive packets since it has to be buffered until a route between the source and the sink is established.

The efficiency of hybrid protocols strongly depends on the traffic pattern and the configuration of the protocols. Routes in the network may have to be re-established very often if the protocols use a too short route timeout. As a result, the delay of the protocol increases since a large number of packets have to be stored in the waiting queues of the nodes until a new route is available. Hybrid protocols which follow this strategy generate almost the same overhead as proactive protocols in the case that the route timeout interval is set to a too large value. Therefore, nodes in the network keep on sending routing information for a long time even if the route is not further required [10]. WSNs are usually data-centric which results in data aggregation along the route from the sources towards the sink. For this reason, the sink frequently receives traffic from the nodes in the network after a certain duration. Therefore, hybrid protocols provide a practical solution to sensor networks in the case that the number of nodes which broadcast routing information is limited [8].

The second group of hybrid protocols use proactive routing mechanisms for short range communication and reactive routing techniques for long range communication [85, 86]. Thus, they use periodic broadcast mechanisms to establish and maintain routes to nodes which are reachable within two or three hops. The protocols switch to a reactive mechanism similar to AODV if the destination is further away. A similar approach is introduced in [87] where the authors propose a routing protocol which adapts the link state exchange frequency depending on the distance to the destination. As a consequence, the overhead is significantly reduced since these kinds of protocols minimize the periodic flooding of routing information.

3.1.2 Network Structure

The network structure represents one of the key characteristics of WSNs since it is mainly responsible for the efficiency of the network. Routing protocols can follow different strategies to build a stable and efficient topology. The most common network structures are flat, hierarchical, and location-based. Flat structured networks represent the simplest solution and are often applied in small or middle-size networks where each node covers the same functions and responsibilities. Hierarchical protocols try to cluster the nodes in the network in order to minimize the routing overhead. Thus, they require more powerful nodes which are able to fulfill the function of cluster heads. Location-based protocols follow a different approach to structure the network. Nodes that are placed within a certain area are grouped instead of using a unique address for each node. Therefore, the networks scale with their size and not with the number of nodes. However, all three approaches have advantages and disadvantages which make them only a practical solution for specific applications. In the following paragraphs, the three network structures are introduced and discussed in more detail from the perspective of WSNs.

Flat

Sensor nodes are often randomly deployed on the ground or in areas where the signal propagation is very limited. Therefore, the connectivity between the nodes differs a lot and is also very hard to predict in advance. Moreover, the majority of applied sensor networks consist of less than 200 nodes. For this reason, routing protocols which build a flat network structure [4, 70, 72, 88–91] represent the most applied type of protocols in WSNs since they provide a practical solution for small and medium-size networks without the need of a complex network structure algorithm. It is obvious that a flat network structure does not provide the best solution for large or dense wireless networks due to scalability issues. The inefficiency of the dissemination of routing information [92] can become a serious problem in multi-hop sensor networks with a flat network structure due to the limited bandwidth of the low-power transceivers.

Hierarchical

Sensor networks take advantage from a hierarchical network structure as soon as the number of nodes and the node density becomes high. Hierarchical networks provide an optimal structure for the efficient exchange of routing information. Furthermore, hierarchical protocols [28, 36, 64, 71, 93] are highly scalable since they can build a multiple-tier architecture to simplify the addressing of the nodes by minimizing the routing table. In addition, data traffic is usually aggregated along the path from the sensing nodes to the data sink. Therefore, the nodes along the route consume more energy as a consequence of the frequent forwarding of data packets. This problem is often compensated by nodes with less energy constraints. These nodes cover the function of cluster heads [93] and play a central role in hierarchical protocols since they have to establish and maintain the topology in the network. Cluster heads are also often responsible for synchronization, scheduling and medium access of the other nodes. Hierarchical networks can be easily optimized under different criteria due to the high performance of the cluster heads. Nonetheless, the usage of more powerful nodes increases the deployment costs of the networks. Moreover, the cluster heads represent critical point of failures since the failure of a cluster head will result in large changes of the underlying topology.

Location-based Routing

Scalability represents a serious issue in WSNs, as already discussed in the introduction of this chapter. The size of the routing table has to be considered when deploying a WSN. Hierarchical protocols minimize the routing table by assigning nodes to a cluster head. A similar approach is followed by location-based routing protocols [85, 94–97] which group nodes with respect to their location. This approach is very practical for applications, like agricultural monitoring, environment monitoring or boarder surveillance since these applications are mainly interested in information which was gathered from a particular area. The disadvantage of this approach is that the nodes cannot be deployed randomly if they are not able to specify their current location. For this reason, location-based routing protocols [98] are typically used for outdoor applications where nodes may use low-power GPS modules to measure their current position. However, the usage of GPS modules increases the costs of the network. There exists a large number of other mechanisms to estimate the position of the nodes in the network [99, 100], e.g. the deployment of seed nodes which can be used as anchor points [101] or by using RSSI measurements [102]. Nonetheless, the position information is not always of practical use for the routing protocol as a consequence of the irregular communication range of the nodes. Furthermore, there might be no link between two nodes which are close to each other since obstacles may block the signal propagation.

3.1.3 Protocol Operation

Routing protocols can be also categorized with respect to their protocol operation [68]. This kind of categorization has the advantage of being more application oriented compared to the two previously discussed taxonomies. In the following, the protocols are distinguished in negotiation-based and query-based protocols. Moreover, the protocols are grouped whether they offer multi-path or QoS support and depending on the used data processing technique. However, the protocol-based classification is not as strict as the other classifications. Thus, a routing protocol may fit into more than one category.

Negotiation-based Routing
Negotiation-based routing protocols directly address the problem of duplicate data transmission by using high-level data descriptors. The dissemination of duplicate data represents a serious issue for WSNs since it increases the energy consumption. In the worst case, nodes may even retransmit the duplicate data which will lead to data implosion and overlap. A large number of protocols were introduced in the late 90’s which are using negotiation-based routing techniques. The most popular family of negotiation-based protocols is represented by the SPIN [103] family. The main idea of these protocols is to exchange certain negotiation messages in order to verify the following data transmission. Thus, they prevent the transmission of redundant data which increases the lifetime of the network.

Multi-Path

Multi-path routing protocols are often used to increase the fault tolerance of wireless networks with unreliable data links [104]. The maintenance of multiple routes increases the energy consumption and the traffic generation. However, in some cases multi-path routing can be almost as energy efficient as single-path routing if the data links are very lossy. Routing protocols often optimize the routes between nodes in the network such that most of the traffic is routed along a certain path in the network. Thus, the energy consumption of the nodes along the path is very high which reduces the lifetime of the network. Multi-path routing protocols are able to mitigate this problem by occasionally using a set of near optimal paths [105] to distribute the traffic load in the network. Moreover, it is obvious that multi-path routing decreases the delay in the network since nodes may immediately switch to backup routes if a route breaks or becomes congested [97]. Another advantage of multi-path routing is represented by the energy-efficient recovery from network failures which is often neglected. The resilience to network failures, e.g. link breaks or node failures, can be further improved if the routes are partially disjoint [106].

Query-based Routing

A WSN can be regarded as a large distributed data base in which each node holds a small part of the overall information. This point of view has led to a new group of routing protocols which are based on queries [72, 89, 91, 107]. These protocols try to optimize the communication in the network by using queries to request information from a node. The queries are typically described in high-level languages in order to allow more complex requests. Thus, they minimize the routing overhead due to the fact that even complex queries can be performed or described by a single message. The query-based routing mechanisms are mainly used in a reactive manner such that a node only sends data if it has received the corresponding request.

Quality of Service Support

Quality of Service (QoS) support in WSNs was neglected for a long time since it was assumed that applications for WSNs are very delay and fault tolerant. However, due to advantages in technology a large number of applications, e.g. multimedia, surveillance, industrial process control, structural health monitoring or health care, becomes interesting under economical aspects [108, 109]. Nevertheless, QoS support is very hard to achieve in WSNs as a consequence of the low data rate and the hardware limitations compared to mesh and ad hoc networks. In [110] the authors introduced the most important QoS issues from the perspective of WSNs. Moreover, they showed that the majority of the problems are NP-complete and can thus only be approximated in real-time. For this reason, only a small number of routing protocols for WSNs [96, 97] try to support QoS functionality.

Coherent-based Routing

The transceiver is usually the most energy-consuming part of a sensor node. Therefore, routing protocols which are designed for wireless networks try to aggregate and preprocess data in order to minimize the data transmission. Two different strategies [111] can be followed by the protocols. The first strategy is represented by a non-coherent approach where nodes perform intensive preprocessing of the raw data before forwarding the data to the next aggregator. Moreover, the nodes collect data over longer periods to aggregate as much information as possible. This effectively reduces the protocol overhead if the sensed values are frequently collected. The second strategy is based on a coherent approach. Nodes that follow a coherent strategy only perform minimum processing of the data, e.g. timestamps and duplicate suppression. Further processing of the data is done by major aggregators which usually have a higher computational power, more memory space, and are less energy constraint. Both strategies require an optimized topology such that they can take advantage from preprocessing and data aggregation.

3.2 Routing Tasks

Routing protocols have to perform various tasks which are independent of the route establishment procedure, the network structure and the protocol operation. The following four tasks are mandatory for every routing protocol in order to establish routing functionality.

The first task is represented by the forwarding of data packets. The forwarding of data can be either based on local information, global information or on information which is stored in the packet. Thus, a routing protocol has to provide packet evaluation mechanisms which build the second task. The received information is then used to update the local routing table or the next hop list. Maintaining a valid topology by evaluating incoming routing information is the third task which has to be performed by every routing protocol. In the case that a node has no information how to reach a certain destination, it has to query other nodes which is typically done by broadcasting a request in the network. The dissemination of routing information - the fourth task - is very energy consuming due to the high node density in WSNs. Thus, a large number of strategies [76, 92] were developed to increase the efficiency of dissemination by using intelligent forwarding of routing messages.

3.2.1 Forwarding

The forwarding of data packets is a mandatory task in multi-hop wireless networks. The applied forwarding strategies may distinguish depending on the requirements of the network. Due to the large number of different forwarding techniques, only a short overview of the most popular ones is given. A very practical strategy is to divide the network into active and passive nodes as introduced in [10]. Passive nodes do neither forward routing messages nor data packets. Thus, the active nodes build the backbone of the networks. Active nodes are usually less constraint in terms of energy, memory, and computational power.

The forwarding itself can be either based on local information or on information which is stored in the packet [66]. Local information comes with a slight advantage in mobile networks where frequent topology changes occur as a consequence of the link breaks. Information about the state of local links is in general more up-to-date than the information of distant ones. For this reason, local decision-based forwarding usually represents the better choice. However, it prevents the originating node from choosing a desired route, e.g. the avoidance of certain nodes or a particular area.

3.2.2 Processing

Processing describes how incoming routing messages are evaluated by a protocol. Note that the way a routing protocol evaluates incoming information has a large impact on its performance. First of all, routing protocols can be divided into two groups. The first group is represented by protocols which only evaluate routing messages which are directly addressed for the corresponding node. This kind of protocols are mainly used in duty-cycled WSNs since nodes sleep most of the time and will thus unlikely receive routing messages which are dedicated for other nodes. In addition, these nodes are usually very limited in terms of energy and computational power. Therefore, the network will not be able to take significant advantage from overheard traffic. Nonetheless, setting the MAC protocol to promiscuous mode becomes an interesting choice as soon as the nodes have sufficient energy resources and computational power. Thus, the second group is represented by protocols which take advantage from overheard traffic [9]. The evaluation of packets which are not dedicated for this node, may be used to update the information in the routing table. Due to the fact that traffic in WSNs is typically aggregated on its way towards the sink, it is a good idea to use this traffic to quickly build up routing information. One of the first protocols which takes advantage from overheard traffic is represented by the Dynamic Source Routing (DSR) protocol [66, 70]. During the development of the simulation which was used to compare reactive and proactive protocols [8], we found out that even the performance of proactive protocols, like the Open Link State Routing (OLSR) protocol [67], can be improved by taking advantage from overheard traffic.

3.2.3 Topology

Establishing a valid topology becomes a very challenging task in sensor networks due to the fact that the links are highly unreliable [13–16]. As a consequence of the frequent link breaks caused by node failures, interference or mobility, the protocols have to provide mechanisms to quickly detect these topology changes. Otherwise, the end-to-end connectivity cannot be assured especially in the case that the destination is several hops away from the source. Most protocols use frequent probing [67] to minimize the time to detect topology changes. However, the frequent transmission of routing messages increases the overhead of the protocol which might even lead to congestion. Thus, the routing message interval has to be chosen with respect to the estimated link duration time [112] and the available bandwidth.

3.2.4 Dissemination of Routing Information

The dissemination of routing information represents the most critical task for routing protocols since it directly affects the performance of the protocol in respect to reliability and routing overhead. In case of an ideal dissemination [76], all nodes in the network exchange data along shortest-path routes. Furthermore, the nodes receive each piece of distinct data only once. Thus, the nodes do not waste energy in terms of unnecessary retransmission and reception of packets. Different kinds of strategies [92] can be applied to disseminate the information depending on the network structure, the node density, the network size, and the link reliability. In the following, several popular strategies are described in more detail.

Flooding

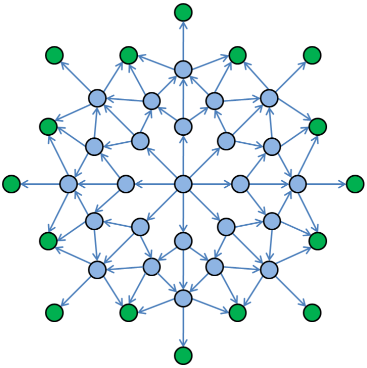
Many routing protocols for WSNs still use ordinary flooding of information due to simplicity reasons. Dissemination strategies, like Multi-point Relay (MPR) [113] and network coding [114], are too complex to be frequently computed on sensor hardware. Note that these algorithms scale with the node density which is often very high in WSNs in order to compensate node failures and unreliably links. Nonetheless, flooding represents the worst solution as a consequence of the large amount of generated overhead. Moreover, simple flooding often leads to temporary contention in WSNs even if the duplicate transmission of information is prevented. The bandwidth of low-power transceivers makes flooding only applicable in small or middle size sensor networks with weak energy constraints. The high number of retransmissions assures that all nodes in the network receive the corresponding information.

Gossiping

Another way to disseminate routing information is represented by gossiping [75, 115] which is a probabilistic-based dissemination scheme. The difference between gossiping and flooding is that nodes retransmit routing messages only with a certain probability. Haas et al. [115] have shown that a retransmission probability of 60 to 80 percent is sufficient in most networks to successfully disseminate routing information among all nodes. Thus, gossiping is more efficient than ordinary flooding since it requires fewer retransmissions. Nevertheless, gossiping should be only considered as dissemination strategy if the node density is very high or if the topology is known in advance. Otherwise, there is a high probability that some nodes in the network may not receive all routing messages.

MPR-based Approach

A very efficient approach to disseminate routing information is used by OLSR [113]. The protocol uses a MPR-based approach to specify the forwarding nodes. A MPR node is a node which is selected by one of its one hop neighbors (MPR selector) to forward all received broadcast messages from this neighbor. The basic idea of the MPR-based approach is to select/calculate a minimum subset of one hop neighbors such that all two hop neighbors receive a routing message if it is forwarded by the minimum subset of one hop neighbor nodes. A comparison of the pure flooding and the MPR-based approach is shown in Fig. 3.2. The figure indicates that the approach is much more efficient than pure flooding. However,



(a) Pure Flooding - 24 Retransmissions (b) MPR - 12 Retransmissions

Figure 3.2: *Example - Flooding vs. MPR-based Forwarding*

the MPR-based approach comes with some drawbacks. First of all it requires precise knowledge of the one and two hop neighborhood. Thus, it requires the frequent exchange of short range routing messages, e.g. hello messages. Moreover, links in WSNs are usually less stable and reliable than links in mesh or ad hoc networks [13] due to the low transmission power and the low gain of the chip antennas. Therefore, a certain number of redundant transmissions are required to assure that all nodes in the network receive the corresponding message. A MPR-based approach is not able to fulfill the task of information dissemination in the case of lossy and unstable links.

Other

There exists a large number of other dissemination schemes besides the ones which were discussed in the three previous paragraphs. Most of them are based on the knowledge of the one hop neighborhood. A very simple scheme is represented by flooding with self-pruning [116]. In this scheme, nodes periodically transmit hello messages which contain a list of all neighbor nodes. Neighbors which receive the hello message only forward it if they reach additional nodes.

Dominant pruning represents another popular neighbor-knowledge based strategy. It uses the greedy set cover algorithm which represents an heuristic of the MPR set calculation. Nodes periodically transmit routing messages which contain a list of their neighbors and another list (subset) of neighbors which are selected to forward the message. The forwarding nodes are chosen according to the greedy set cover algorithm. The algorithm recursively choses one hop neighbors which cover the most two hop neighbors.

Location-based dissemination schemes are also a very practical solution if position information is available. Ni et al. [75] introduced an approach where forwarding nodes are selected depending on the additional coverage area. A threshold is used to minimize the number of retransmissions. However, the approach tends to use high distance links which are usually less reliable [14]. Thus, acknowledgment mechanisms should be applied if this scheme is used.

3.3 Routing Metrics

Routing protocols try to establish and maintain optimized paths between the nodes in the network. The routes are optimized according to the used routing metric. The number of applied routing metrics is almost endless [117] since metrics are also often combined in order to meet the requirements of a certain application.

3.3.1 Classification of Routing Metrics

In the beginning of digital communication routes were optimized with respect to the number of hops. The hop count metric is still popular in nowadays networks due to its simplicity and due to the fact that it is supported by the Internet Protocol (IP). However, links between nodes in the network usually distinguish in bandwidth, delay, reliability or other characteristics of interest. Most of these metrics cannot be described by discrete values. Thus, state-of-the-art routing protocols mainly use continuous metrics to optimize the routes in the network. Furthermore, the majority of routing protocols which are designed for wireless networks combine several metrics in order to find a trade-off between different characteristics.

Discrete

As mentioned in the previous paragraph, discrete routing metrics are used to describe countable characteristics, e.g. hop count or number of neighbors. Thus, they are very limited in reflecting the quality of a link. For this reason, many routes in the network may have the same costs, especially in dense networks, even if the routes distinguish significantly in their performance. This can lead to a low network performance or congestion on a certain path since the metric does not take delay, reliability or available bandwidth into account. Moreover, continuous metrics like remaining energy can only be roughly described by discrete metrics. For this reason, discrete metrics are only applied in small WSNs where nodes are very limited in computational power. Note that even a simple hop count based metric may achieve a better performance than complex link quality metrics in wireless networks as shown in [118].

Continuous

Continuous routing metrics are very common in wireless mesh networks due to the fact that they are less restricted in hardware and bandwidth limitations. Therefore, mesh nodes are able consider more complex metrics, like the expected number of retransmissions or the interference along a certain route. Continuous routing metrics are also discrete from a mathematical point of view due to the bit-based number representation. They are mainly used to describe uncountable characteristics, like interference or delay. In addition, continuous metrics can be based on calculations which use discrete values as input, e.g. the exponential weighted moving average of the node degree [9].

The latest generation of wireless sensor nodes offers enough computational power to allow more complex calculations which make continuous routing metrics a practical solution for them as well. The four most popular continuous routing metrics for WSNs are energy consumption, remaining battery power, reliability and delay. Moreover, cross-layer approaches which take the Signal-to-Noise Ratio (SNR) and other link quality metrics into account become very popular nowadays. The increased popularity results from the fact that the link layer functionality which is provided by low-power transceivers has improved a lot during the last few years.

Combined
Sensor networks are usually designed for a particular application which has specific requirements on the routing protocol. In addition, typical WSN applications require an optimization in respect to different characteristics. For this reason, mathematically combined metrics are used to achieve the desired behavior of the routing protocol. The metrics are also often weighted differently to create the desired characteristic. Another possibility is to simply add or multiply different metrics. However, the mathematical combination of different metrics should be only considered if all side-effects of the metric combination are fully understood [117]. Otherwise, there is a big chance that the established routes are far below the desired optimum [119].

3.3.2 Discussion of Different Routing Metrics

The classification in the previous subsection described the differences between the routing metrics in terms of route differentiation and optimization potential. The focus of this subsection lies on the characteristic of popular metrics and the way they affect the topology in a network. Moreover, techniques are introduced which can be used to deal with the problem of fast changing metrics in order to establish stable routes.

Delay
Delay represents a fast changing metric which is often combined with discrete metrics like hop count to allow further route differentiation. The advantage of

delay-based metrics is that they can be used to balance the load evenly in the network [120] due to the fact that links with low utilization forward routing messages faster than links with higher utilization. As a result, the load is shifted from links with high traffic load to links with low traffic load. Another advantage of delay-based metrics is that they can be applied without generating additional routing overhead since no additional header field is required.

However, due to the fast changing characteristic of the metric, additional filtering mechanisms, e.g. sliding window or weighted moving average, have to be applied [121]. Otherwise, short-term variations would result in an oscillation of the routes. The size of the sliding window should be dynamically chosen depending on the current transmission rate to produce meaningful results.

The problem of filtering mechanisms is that they have to be adjusted to meet the requirements of the network in terms of link change detection and stability of established routes [117]. Delay-based metrics should specify a threshold to minimize too frequent route changes which might have a negative effect on the throughput of flow oriented transport protocols like TCP. This negative impact on the TCP throughput was already recognized in the early stages of the Advanced Research Projects Agency Network (ARPANET) [122].

Reliability

Many routing metrics for WSNs are based on reliability since high reliable links minimize the number of retransmissions which results in less energy consumption. The reliability of a link usually remains on a certain level without much variation. However, short temporary changes in the reliability can be caused by interference of other wireless technologies which use the same frequency spectrum. These changes in reliability occur only sporadically and over a short time interval. Thus, they can be compensated by the filtering mechanisms described in the previous paragraph. The major problem is caused by persistent link breaks as a result from mobility or node failure. For this reason, the applied filtering mechanisms have to be configured such that short term changes are compensated while still maintaining a short reaction time to detect persistent link breaks.

A routing metric which is widely spread in WSNs is represented by the Expected Transmission Count (ETX) [123]. The metric reflects the estimated number of transmissions which are required to assure the successful reception of a packet transmission over a link. It mainly follows the trend of link reliability and is thus greatly affected by short term variations. Therefore, it is strongly recommended to take the variance of the ETX metric [124] into account instead of focusing on the average packet loss ratio.

Energy Consumption

The most complex routing metrics are those which take the energy consumption into account. Energy consumption can be optimized in many different ways. It is possible to optimize the energy consumption for a single node, a route or the whole network. Thus, the metric also has to consider passive energy consumption effects caused by the characteristics of the underlying MAC protocol, like scheduling and overhearing. Overhearing has to be taken seriously since it increases the energy consumption and results in a waste of bandwidth, especially in high density wireless networks [25]. In the following it is assumed that a CSMA-based MAC protocol is used due to the fact that they provide a basis for most MAC protocols. There exist different optimization goals which can be targeted by energy aware routing metrics.

The majority of the metrics tries to prolong the time until the first node in the network runs out of energy [90, 120]. However, this kind of metric is not very practical for two reasons. First of all, the number of nodes in a WSN is typically chosen such that there are sufficient redundant nodes to compensate single node failures. Moreover, the nodes around a source or a sink will always be the first nodes which run out of energy since they receive and forward more traffic than other nodes in the network. Energy aware routing metrics are only able to slightly mitigate the impact of the energy hole problem [125, 126]. Furthermore, this kind of optimization does not necessarily minimize the overall energy consumption.

Another energy aware strategy is to distribute the traffic load evenly in the network by assigning the link costs with respect to the remaining battery power of the corresponding node [120]. This approach is similar to the previous one since the nodes with the lowest remaining energy will be avoided by the routing protocol. Nonetheless, this kind of strategy only performs well if the routing protocol also considers the energy consumption of the nodes which are within the interference range of a link. Otherwise, a node which has only a small amount of energy left will quickly run out of energy in the case that it is surrounded by nodes with high battery power. Note that the energy consumption of this node will remain high since the surrounding nodes will be selected as forwarding nodes. Consider a sensor network where nodes use LPL [45] to access the medium. In this case, the low-energy node will wake up frequently due to ongoing transmissions in its one hop neighborhood. The dissemination of the energy consumption does not necessarily result in the minimization of the overall energy consumption [90] due to the fact that longer routes are taken into account which require more retransmissions.

Other metrics try to maximize the time until the network gets partitioned. This type of metrics require global knowledge of the network in terms of remaining battery power, location of the nodes, interference range, transmission range, characteristics of the underlying mac protocol and the likely traffic pattern. It easily becomes obvious that such a metric can only be approximated since it is far too complex to be calculated by low-power sensor nodes [127].

Hop Count

The hop count metric represents the simplest of all routing metrics. It is applied by large number of popular routing protocols [65–67, 77] not only due to simplicity reasons. Surprisingly the simple hop count metric may outperform complex link quality metrics in a large number of scenarios since it is not affected by minor link quality changes. Thus, the metric generates a more stable topology which may improve the throughput of flow oriented transport protocols. However, the performance strongly depends on the topology, the underlying MAC protocol, the traffic patterns and the mechanisms provided by the routing protocol. In [15] it was shown that Destination-Sequenced Distance-Vector (DSDV) [77] routing takes a large benefit from ETX metrics while Dynamic Source Routing (DSR) [66] is only slightly improved. Their results are based on scenarios where nodes transmit small datagrams over a short period of time. Nonetheless, ETX-based metrics perform better [118] as soon as the duration of the traffic flows increases. Moreover, hop count based metrics tend to use long distance links which are often less reliable [14]. In addition, these metrics do not consider delay and packet loss. As a result, hop count based metrics will try to route traffic through a bottleneck if it represents the shortest path to the destination.

Node Speed

Node speed is one of the less investigated routing metrics due to the fact that nodes are in general not able to measure their speed. However, the link duration time is mainly affected by the speed of the nodes in the network [112]. The link duration time specifies the duration from the point in time when the link was established until it breaks as a consequence of the node movement. A routing protocol which is designed for mobile networks should consider the current speed of a node since it is usually correlated with the number of topology changes. The performance of mobile wireless networks can be improved if fix or slow moving nodes build a backbone for fast moving nodes since they are able to establish a more stable topology. In [9] it was shown that it is possible to reduce the end-to-end packet loss of AODV by approximately 15 percent if the forwarding nodes are chosen in respect to their absolute speed. Furthermore, the results pointed out that even a simple estimation of the relative node speed is able to significantly improve the reliability in the network. A detailed description of this approach is given in Section 3.5 since it is part of the extended functionality of the SBR protocol.

3.4 Survey on Routing Protocols

In this section a short survey of five popular routing protocols which follow completely different strategies is given. The protocols were selected such that they cover almost the whole spectrum of the presented routing taxonomy. The basic mechanisms of the protocols are adopted or used in a slightly modified version by a large number of routing protocols for wireless mesh, ad hoc and sensor networks. Some of the presented protocols were originally designed for wireless mesh and ad hoc networks. However, it was shown in [128] that these protocols outperform their sensor-specific counterparts in typical WSN scenarios.

3.4.1 Ad hoc On-Demand Distance Vector

AODV [65] is a reactive routing protocol which means that it only tries to establish end-to-end routes on demand. Such behavior is very useful in networks with a very low traffic load to keep the routing overhead on a low level. The disadvantage of this strategy is that routes have to be established before data can be transmitted. Therefore, reactive routing protocols have a higher end-to-end delay than proactive protocols which maintain end-to-end routes even in the absence of data traffic. Nevertheless, the maintaining of fresh routes results in a large amount of routing overhead. To keep things short only the basic mechanisms and routing messages of AODV are introduced in this thesis.

Route Requests

Route Request (RREQ) are sent out by nodes to establish a route to a destination. The requests are broadcasted in the network and contain a unique sequence number to allow differentiation. In addition, a hop count field is part of each request to prevent infinite retransmissions. Nodes that receive a request for the first time or with a smaller hop count and do not have a valid route to the destination rebroadcast the request. The forwarding node updates its routing table before broadcasting the request. The node from which the RREQ was received represents the next hop towards the originator of the request. A receiving node keeps track of previously received requests to distinguish new requests from retransmissions. If a destination node recognizes its address in a new received RREQ, it answers the request with a Route Reply (RREP).

Route Replies

RREPs are transmitted by the destination of the RREQ and usually follow the reverse path of the RREQ back to the originator of the request. The RREP also has a hop count field to prevent its infinite retransmission. Intermediate nodes only forward the reply if they have received the reply for the first time and know of a valid route to the destination. Thus, the path of the RREQ and the RREP may distinguish in case of unidirectional links or asymmetric link speeds as indicated by Fig. 3.3. Furthermore, the reply is broadcasted if the hop count in the received reply is smaller than the hop count of the shortest known route to the originator of the reply. Recall that the data source is the originator of the RREQ and the destination of the RREP whereas the data sink is the destination of the RREQ and the originator of the RREP.

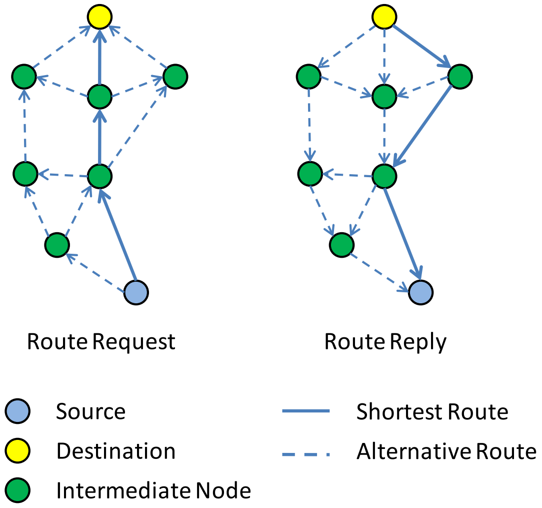


Figure 3.3: *AODV - Dissemination of Routing Messages*

Intermediate Route Replies

Intermediate nodes are also allowed to answer a RREQ directly with an intermediate route reply if they have a valid route to the destination. Intermediate replies help to minimize the routing overhead since the retransmission of RREQ is reduced in a significant way. The node places its distance in number of hops to the destination in the hop count field of the intermediate reply before it retransmits the message.

Route Validity Time

The AODV protocol uses timers to specify the validity time of routes and links. All replies contain a lifetime field to indicate to other nodes how long the route should be considered valid. In addition, each valid route has an active route time out. If no data packet is transmitted via this route for the duration of the active route time out, the route is removed since the node assumes that the route is not further needed.

End-to-End Route Selection

The protocol uses the sequence number in routing messages to differentiate the freshness of the received information. The second value for end-to-end route selection is represented by the number of hops. AODV discards valid RREPs if the number of hops is larger or equal than the current used route. Thus, AODV always uses the shortest end-to-end route for data transmission. In general, more than one shortest path is available in a network. Therefore, the chosen next hop is the one through which the first shortest path RREP is received. For this reason, the topology can be slightly modified if the node delays the forwarding of hello messages [120].

3.4.2 Dynamic Source Routing

The DSR protocol [66, 70] adopts many mechanisms from AODV. The route establishment procedure is very similar in the way that it uses RREQ and RREP messages. The difference to AODV is that a forwarding node appends its address to the routing message in order to provide the destination with the full knowledge of the path between the source and the destination. In addition, the protocol makes use of a large number of mechanisms to improve the performance of the protocol. Instead of using frequent transmission of routing messages, nodes may use passive acknowledgments [66,129] to detect link breaks. Thus, nodes listen to the medium after they have forwarded a routing message. The nodes may be able to hear the transmission of the next hop. If a node does not hear the forwarding of the message within a certain period of time it assumes that the link to the next hop is broken. Another mechanism which is used by DSR to improve its performance is to piggyback small packets to RREQs. This mechanism may increase the throughput in a significant way if the piggybacked packet is an initial synchronization packet opening a TCP connection [130]. Furthermore, DSR takes advantage from using the promiscuous mode of the underlying MAC protocol. Thus, the protocol uses overheard routing messages to update the information in the routing table. Additional mechanisms have been added to the final version of the protocol [66]. However, they are not addressed here since they are out of the scope of this monograph.

3.4.3 Open Link State Routing

OLSR [67] represents one of the most popular proactive link state routing protocols. It was originally designed to meet the requirements of large and dense mobile ad hoc networks with multiple sources and sinks. However, the protocol can be configured and modified such that it also becomes an interesting solution for WSNs [86]. In the following, the focus lies on the basic version of the protocol.

The protocol uses two types of messages to disseminate the routing information in the network. Hello messages are used by a node to exchange information about its one hop neighborhood with its neighbors. The messages are not forwarded since they are only designed to update and exchange local link information. Note that the neighbors of a node are not necessarily able to directly communicate with each other due to asymmetric transmission range [16]. Thus, a node may use the information which is contained within a received hello message to update its two hop neighborhood.

In addition, the protocol uses periodic topology control messages which contain a list of nodes which have selected the originator of the message as MPR node. Topology control messages are only forwarded by MPR nodes which significantly minimizes the routing overhead compared to ordinary flooding especially in dense networks as discussed in Section 3.2.4.

The route to a destination is calculated by each node individually by using the information which is received via the topology control messages. For this reason, the protocol requires a large amount of memory and a high computational power in case of a large network size. As a consequence, the protocol can only be applied to small or middle size WSNs due to the routing table explosion problem.

OLSR uses validity and expiration timers for neighbor and topology entries which are used to calculate the routes. The timers have to be adapted in respect to the mobility of the nodes and the net diameter. If too long time intervals are used in a highly mobile network, the protocol is not able to detect topology changes quickly enough. Thus, the nodes disseminate already outdated information. Moreover, the dissemination of topology control messages becomes inefficient if the MPR set is not calculated correctly due to outdated information. As a result, the outdated information might even lead to a total collapse of the network since a different topology is assumed.

3.4.4 Directed Diffusion

Directed-Diffusion [72, 131] is a data-centric routing paradigm which is adopted by a large number of routing protocols [4, 88, 90]. A data sink periodically broadcasts interest messages. These messages hold the description of a sensing task which can be performed by the network. The interest messages are evaluated and forwarded by the nodes. Furthermore, the nodes set up gradients in the direction from which the messages were received. The sensed data is then routed along the gradients. As a result, it is possible to aggregate data and interests along the routes towards the sink. Due to the periodic broadcasting of interest messages, the protocol is able to discover new routes and link breaks between the source and the sink. The following two paragraphs describe the Gradient Based Routing (GBR) [90] protocol and the Minimum Cost Forwarding Algorithm (MCFA) [88] in more detail which apply mechanisms similar to Directed Diffusion.

Gradient Based Routing

GBR [90] applies almost the same strategy as Directed Diffusion. Thus, a data sink broadcasts messages which are forwarded by the other nodes. A broadcast message contains a hop count field to indicate the distance to the sink. Each time the message is forwarded by a node, the hop count field is increased by one. Therefore, a node simply uses the neighbor with the lowest height as next hop if a message has to be forwarded to the sink.

Moreover, the authors of the GBR protocol introduced three different strategies besides the standard version to optimize the chosen routes. The first scheme is called stochastic-scheme. Due to the fact that the GBR only uses hop count as routing metric there usually exists a large number of routes with equal length. For this reason, they propose to randomly choose one of the shortest paths to distribute the traffic load. The energy-based scheme represents the second scheme which takes the remaining energy of a node into account. If the energy level of a node drops below 50 percent it increases its height in order to minimize the probability of being chosen as next hop by one of its neighbor nodes. The third scheme is stream-based. A node which is part of a data stream tells its neighbors - except the one from which it receives the stream - that its height has increased. Thus, it becomes less attractive for the other nodes to forward their data.

Minimum Cost Forwarding Algorithm

The MCFA [88] sets up gradients similar to the GBR protocol. Furthermore, the sink broadcasts messages which contain a cost field instead of a hop count field. The cost field can hold any information, e.g. hop count, energy consumption or delay. A variable cost field represents a more flexible metric compared to the static hop count metric which is used by GBR. The idea of the MCFA is that messages are only processed and forwarded according to the value which is stored in the cost field. Therefore, it does not require any addressing of the nodes. Each node has only to maintain the minimum cost towards the sink in order to know through which neighbor it can reach the sink. A node only retransmits a broadcast message if the cost field in the message holds a value which is lower than the minimum cost value. Moreover, it increases the value in the cost field by its own costs before the message is forwarded. Thus, all nodes have to set their costs to infinity during the initialization of the network. It is obvious that this forwarding strategy will result in a massive broadcast storm since a node will frequently receive broadcast messages with a cost field which is lower than the current minimum cost value. For this reason, a node delays the forwarding of the message by a time which is proportional to the optimal costs including the costs of the node itself. This mechanism reduces the routing overhead in a significant way and makes MCFA a practical solution to sensor networks where nodes have low computational power and a low amount of memory. The idea of delayed forwarding is adopted by several routing protocols [10, 120] since it can be used to modify the topology as shown in Subsection 3.5.5.

3.4.5 Low-Energy Adaptive Clustering Hierarchy

Flat routing protocols run into scalability problems [132] if the network size becomes too large. First of all, the routing table explosion limits the number of nodes which are supported by a protocol. In addition, the dissemination of routing information becomes very expensive in terms of energy consumption and routing overhead. Low-Energy Adaptive Clustering Hierarchy (LEACH) [36] addresses the problem of scalability by using a clustering mechanism. A static cluster algorithm will results in a short lifetime of the cluster head nodes since they have to aggregate and forward traffic of the nodes which are within their cluster. Therefore, the authors of LEACH propose a randomized rotation of the local cluster heads in order to disseminate the energy consumption equally among the nodes in the network. The protocol uses two phases to setup the network. During the first phase a node which decides to become a cluster head exchanges advertisement messages with its neighbors. The decision is performed in a probabilistic way. Thus, nodes become a cluster head for one round with a pre-defined probability. This probability has to be chosen with respect to the node density such that the number of cluster heads is always close to the optimum. During the second phase, the cluster heads assign a TDMA scheduling to their nodes and setup the network topology. As a consequence of the data aggregation and the clustering, the LEACH routing protocol is able to significantly increase the lifetime of the network compared to the majority of flat routing protocols.

3.4.6 Multipath Multi-Speed Protocol

The Multipath Multi-Speed Protocol (MMSPEED) [97] routing protocol is one of the few routing protocols for WSNs which try to guarantee QoS with respect to reliability and delay. It is based on mechanisms provided by the SPEED [96] protocol which forwards data packets depending on the virtual speed of a link. The virtual speed can be regarded as the physical speed of a data packet in the direction to the destination. Therefore, the protocol requires knowledge of the position of the nodes and the delay of the link towards the next hop. The virtual speed is calculated by the fraction of the change of the distance to the destination and the link delay. An intermediate node only forwards a packet if the virtual speed of a link to one of its neighbors is higher than the minimum speed which is required by the packet. Otherwise, if the node does not know any link which meets the requirements it drops the packet. Thus, the protocol guarantees that almost all packets which are received by the destination meet the QoS requirements. However, dropping packets is not an option if a high reliability is required by an application. For this reason, the protocol sets up multiple paths towards the destination in order to improve the probability that one of the paths is fast enough to achieve the desired QoS in terms of delay and reliability. Due to the fact that the reservation of bandwidth is not applicable in WSNs, the protocol creates multiple speed and reliability layers to minimize the probability that high QoS links get congested by low priority best effort traffic.

3.5 Statistic-Based Routing Protocol

’Simplicity is prerequisite for reliability.’ Edsger Dijkstra

The SBR protocol was originally designed to meet the requirements of low-power WSNs. Its primary goals are energy efficiency, low overhead, high reliability and simplicity. Mechanisms from reactive and proactive protocols [72,78,87,88] were adopted to build a more flexible protocol while maintaining simplicity and low protocol overhead.

Edsger Dijkstra’s statement is correct in many ways. Often, the time to find an optimized configuration is correlated with the complexity of the protocol. Moreover, many features in state-of-the art routing protocols only slightly increase their performance in rare scenarios. Thus, these features represent unnecessary complexity since no advantages are achieved in most of the cases. Therefore, the major focus was laid on simplicity during the design phase of the SBR protocol.

The protocol generates reliable end-to-end routes with low delay in mobile multi-hop wireless networks. It can be configured to operate like a hybrid or a proactive routing protocol depending on the capabilities and the requirements of the network. The protocol uses a continuous adaptive metric to balance the traffic load evenly in the network. Due to the continuous adaptive metric, acceptable performance is achieved even for non-optimized configurations. Almost any metric can be applied by the protocol since it uses the metric to calculate a forwarding delay which allows the modification of the routes in the network. The advantage of the delay-based approach lies in the fact that it only slightly affects the topology. Moreover, the protocol is able to detect link breaks and unidirectional links within a short period of time making it an attractive choice for indoor and mobile scenarios. The generated overhead mainly depends on the number of data sinks [8]. Thus, it achieves a high performance in WSNs without generating a large amount of overhead since the number of (mobile) sinks is usually small.

3.5.1 Basic Functionality

This subsection introduces the basic mechanisms of the SBR protocol starting with the two different types of routing messages. Furthermore, the route establishment process in proactive and in hybrid mode are discussed. In addition, the basic functionality of the routing table is explained by using a small example topology.

Routing Messages

SBR defines two types of routing messages which are passed to the underlying MAC and physical layer. Due to the fact that the latest available sensor nodes are able to support the Internet Protocol [133, 134], we focus in the following on the implementation of the protocol on top of the IP stack. Hello messages and short hello messages are transmitted and received via UDP. Thus, the protocol uses IP addresses to differentiate the nodes in the network. The messages are sent by using the IP broadcast address. The dissemination of the messages is limited by using a Time-To-Live value of one which is set in the IP header. However, it is also possible to use multicast addresses to further limit the flooding.

Route Establishment

The protocol can operate either in proactive or in hybrid mode. In the proactive mode all nodes in the network periodically transmit hello messages which are disseminated in the network. A node receiving a hello message, rates the neighbor through which the message was received using a cumulative function. The cumulative function is discussed in detail in Subsection 3.5.3. The node stores the calculated value in its routing table. If a node wants to transmit a packet to a destination in the network it sends the packet to the neighbor with the highest routing entry. This neighbor is in the following referred to as best neighbor. The best neighbor forwarding mechanism was adopted from the Better Approach to Mobile Ad hoc Networking (BATMAN) protocol [78].

The routing table entries can be interpreted as gradients pointing towards the destination. A cumulative continuous metric is used to distinguish routes with equal hop length. Due to the cumulative metric more reliable links are preferred over less reliable links. Packets are forwarded along the reverse route of the hello messages. Therefore, the used links have to be bidirectional which makes an additional mechanism necessary to provide end-to-end connectivity if unidirectional links are present.

In hybrid mode no hello messages are generated in the absence of data traffic. Thus, the amount of routing overhead in networks with low data rate and event triggered communication is significantly reduced. In the case that a node wants to transmit data packets to another node for which it does not know a next hop it starts to send out hello messages which cover the function of route requests [65]. Intermediate nodes only forward the message if the message is considered to be new and received via the best neighbor. These requesting hello messages contain the address of the destination. If a node recognizes its own address in a hello message it starts to transmit hello messages by itself which represent the reply of the destination. The originating node stops the transmission of requesting hello messages if a hello message from the destination is received. A destination node stops the transmission of hello messages in the case that it does not receive any data packet for a duration longer than active route timeout which is set by default to three times the hello message interval.

Routing Table

The routing table stores values which correspond to the link quality or any other metric depending on the used routing entry increase algorithm. The value of a neighbor is increased each time a new hello message is received via the neighbor. If the entry is increased by one every time a new hello message is received then the stored value directly corresponds to the number of received hello messages. Table 3.1 represents a routing table example of node B resulting from the connectivity graph shown in Fig. 3.4.

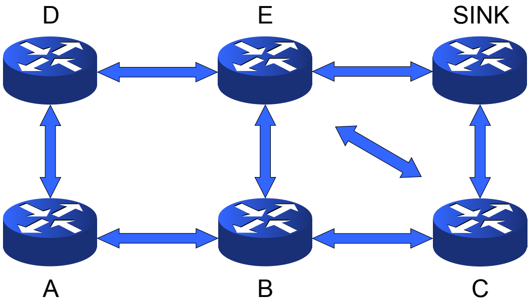


Figure 3.4: *Routing Table - Topology Example*

Table 3.1: *Routing Table Example*

Node B	**Number of Received Hello Messages**				
Originator	A	C	D	E	SINK
A	20	-	-	-	-
C	-	20	-	-	-
D	12	20	-	7	-
E	-	-	-	18	-
SINK	-	8	-	12	-

The columns represent the neighbors through which new hello messages are received. An empty column is the result of the fact that the node is not a neighbor. Multiple values in a row point out that there exist multiple paths between this node and the originator of the hello message. Entries on the diagonal from the upper left to the lower right corner of the table indicate that the corresponding node is a neighbor. The values in the table reflect the corresponding link quality. Thus, node B should forward routing messages to node E if it wants to reach the sink since it receives more messages from the sink via node E. Thus, the path B-C-SINK tends to be slower and/or less reliable than B-E-SINK.

Recapitulate the issue that a node which was the best neighbor for a long time suddenly becomes unreachable. Due to the fact that its routing entry would be very high compared to the other neighbors, it would take a long time for the former second best neighbor to become the neighbor with the highest entry. For this reason, the entries in the routing table are decreased at periodic time intervals to prevent them from increasing to far. The time interval is in the following referred to as Decrease Routing Value Interval (DRVI).

The mechanism reduces the time that the protocol requires to detect link breaks, and thus increases the end-to-end reliability in a significant way in mobile networks. The idea of using a cumulative routing metric in combination with a periodic decrease has not yet been applied by any other routing algorithm. Other protocols try to find a trade-off between a short detection period and a unstable routes by using a sliding window approach or a weighted moving average algorithm [78, 135]. In the next paragraph, a closer look is taken on the change of the entries during a handover in order to give a better picture of the functionality of the routing table.

3.5.2 Topology Changes

The behavior of the protocol during handover or topology changes is discussed in this paragraph. In the following the term handover is used since it is more practical for explanation. However, there is no difference between handover and topology changes from the perspective of the routing protocol as long as the network remains connected after the change.

First, we explain the increase and decrease of the entries in the routing tables when a mobile node X passes two fix nodes A and B. The functions which are used to modify the routing values are called Increase Routing Value Function (IRVF) and Decrease Routing Value Function (DRVF). To keep things as simple as possible a constant increase and decrease value of one is assumed.

Moreover, the DRVI is twice as long as the Hello Message Interval (HMI). Fig. 3.5 shows the trajectory of node X and the transmission range of node A and node B. Important points in time are marked by dashed lines. In addition, it is assumed that no hello messages are lost due to interference and node X moves with constant speed.

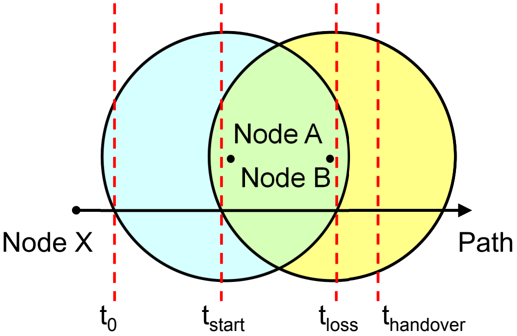


Figure 3.5: *SBR - Handover Example*

The behavior of the entries during the handover is shown in Fig. 3.6. Node X enters the transmission range of node A at time t_o resulting in the increase of the routing value in node A. At time t_{start} node X can be reached by nodes A and B. Therefore, the routing entry in node B increases, too. Node X leaves the communication range of node A at time t_{loss}. As a consequence, hello messages which are transmitted by node X are only received by node B. Thus, the routing value stored in node A decreases whereas the value in node B further increases.

No packets can be routed to node X between t_{loss} and $t_{handover}$ because the routing value of the corresponding entry is still higher in node A. The protocol assumes that node A is the node in charge to reach node X until the value in node B becomes higher. The time period which is needed by the protocol to select the correct node is referred to as downtime. No packets can be forwarded to node X during the downtime.

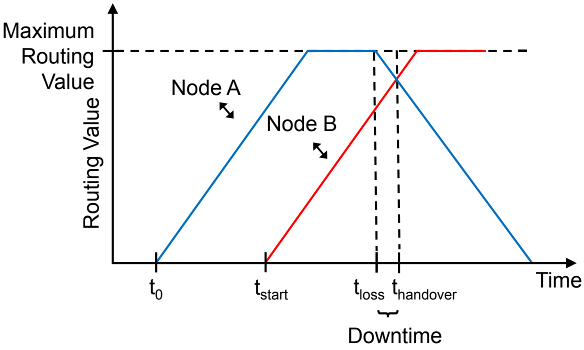


Figure 3.6: *SBR - Linear Routing Value Function*

However, linear functions do not represent the best choice because of their limited capabilities to minimize the downtime [7]. Note that using a higher gradient for the IRVF results in a higher routing value which has to be decreased later on by the DRVF. The only possibility to shorten the downtime with linear functions is to use a higher gradient for the DRVF. In addition, a Maximum Routing Value (MRV) can be applied to further minimize the downtime. Nevertheless, the gradients and the MRV have to be selected in respect to the set HMI and the DRVI. Non-linear functions allow a significant reduction of the downtime if they take the current routing entry value into account. The advantage of a non-linear routing metric is introduced in Subsection 3.5.3.

3.5.3 Routing Metric

Different strategies can be followed to create a trade-off between short reaction time and stable routes. Many popular protocols, like the BATMAN [78] protocol, apply a strategy based on a sliding window to smooth the changes of the routing values which results in more stable routes. Thus, the size of the sliding windows has to be chosen with respect to the variation of the routing metric in order to generate the desired behavior. In addition, the routing values can be weighted according to their freshness [135] which mitigates the impact of short temporary changes of the routing values.

The increase and decrease of the values stored in the routing table offer a way to manipulate the time which is needed by the protocol to adapt itself to topology changes. Instead of using a constant increase and decrease value it is also possible to use functions to estimate the goodness of a path. The following characteristics are required to shorten the downtime. The gradient of the IRVF should be high for low values and low for high values. The gradient of the DRVF should be high for high values and low for low values. In addition, the chosen IRVF has to be asymptotic. Otherwise, a MRV has to be set to limit the routing entry values. Eqn. 3.1 and Eqn. 3.2 are used to increase and decrease the routing entries.

$$I_{n+1} = 2I_n + \frac{4}{I_n^2 + 1} \qquad I_0 = 1, \tag{3.1}$$

$$D_{n+1} = \frac{D_n}{2}. \tag{3.2}$$

The DRVF is usually directly called after the IRVF due to the fact that the DRVI and the HMI are typically set to the same value by default. Thus, the value of an entry asymptotically increases over time if no hello message is lost or faster received through a different neighbor. The used functions have the advantage that new entries increase quickly while entries of links, which break or become unreliable, slowly fade out of the routing table. This behavior decreases the downtime as shown in Fig. 3.7. By comparing Fig. 3.6 with Fig. 3.7, it becomes obvious that the progressive function provides a better performance than the linear function.

The following assumptions are made for the example in order to give a better impression of the characteristics of the functions. The HMI and the DRVI are set to the same duration. The increase is directly followed by the decrease. Hello messages are independently lost and the distribution is chosen such that three percent of the hello messages are discarded. Thus, the decrease function is called slightly more often than the increase function. A typical development of the value

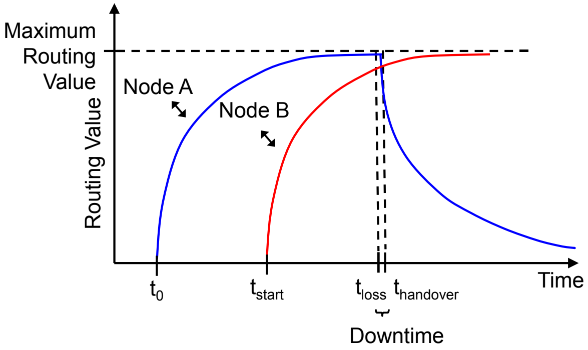


Figure 3.7: *SBR - Progressive Routing Value Function*

within a routing entry is shown in Fig. 3.8. The decrease of the values are the result of lost hello messages. Therefore, the decrease function is called two times after another which decreases the value of the routing entry. The figure indicates that the routing entry recovers quickly after a hello message loss. Moreover, the decrease functions assures that the routing entry value has an asymptotic behavior if no hello message is lost. This behavior is essential to minimize the reaction time of the protocol to detect link breaks since it prevents routing entry values from increasing too high.

3.5.4 Extended Functionality

Extended functionality was added to the protocol to improve its performance in large dense wireless networks where the probability that links have asymmetric quality increases due to higher interference. Overhead represents another serious issue in these networks since the available bandwidth has to be shared among a large number of competitors. Therefore, a passive mode was added to the protocol which is also discussed in this subsection.

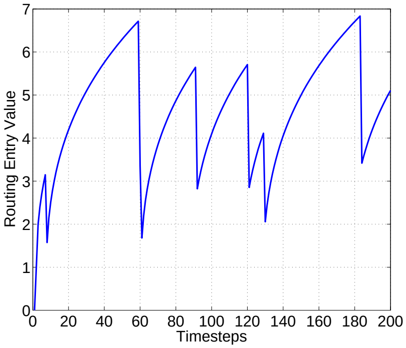


Figure 3.8: *Snapshot of the Routing Value Function*

Unidirectional Links

Short Hello Message (SHM)s are used by the protocol to detect unidirectional links. These messages are similar to ordinary hello messages. The major difference is represented by the TTL value which is set to two. As a consequence, one hop neighbors forward the message whereas two hop neighbors discard the message because of the TTL value. The link to a neighbor is assumed to be bidirectional if the originating node recognizes that the neighbor forwards its hello messages. Thus, SBR takes advantage from passive acknowledgments which have to be supported by the underlying MAC protocol.

The idea of passive acknowledgments was already introduced in [129] and is still used by routing protocols like DSR [66]. The authors of the BATMAN protocol [78] plan to integrate an advanced passive acknowledgment mechanism in the next version of their protocol. The idea of their mechanism is to count the rebroadcasts of a nodes own broadcasts in order to calculate the echo quality. The echo quality to a certain neighbor is given by the fraction of the number of its own broadcasts and the number of rebroadcasts by the neighbor. In the case that the neighbor periodically broadcasts messages, it is even possible to calculate the transmission quality. The transmission quality represents the probability that a packet from the originator is successfully received by the neighbor.

This strategy can only be applied if all or at least the majority of the nodes in the network are periodically broadcasting routing messages. In WSNs, the number of data sinks is usually very small compared to the number of sensor nodes. Moreover, the broadcast frequency which can be used is much lower compared to high data-rate wireless networks. Thus, the advanced passive acknowledgment mechanism is not practical for WSNs since the broadcast frequency is usually too low to estimate the current link quality.

Therefore, the SBR protocol relies on the basic passive acknowledgment approach to detect whether a link to a neighbor is unidirectional or bidirectional. If no routing message of a node is forwarded by one of its neighbor for a certain interval, the link to this neighbor is marked as unidirectional. The interval duration has to be chosen with respect to the HMI and the Short Hello Message Interval (SHMI). Hello messages which are received via a unidirectional marked link are discarded without further evaluation. The support of unidirectional links must be turned off in the case that the MAC protocol does not support working in promiscuous mode. The usage of SHMs is still recommended in mobile networks since it minimizes the time to detect link breaks. The idea of using frequent routing message transmissions for short range topology change detection was inspired by previous work from other researchers [86,87,136]. The authors of [136] introduced a mechanism which sends routing messages depending on the nodes mobility rate. Another interesting approach was presented in [87] where the authors adapt the routing message frequency depending on the distance to the destination.

Passive Nodes

In dense large networks where nodes are very limited in their capabilities due to energy limitation and low data rate of the wireless interface, most of the nodes can be set to passive mode to improve the performance. Passive nodes distinguish from ordinary nodes as they do not contribute in the forwarding of routing messages. The consequence of this passive behavior is that other nodes do not choose these nodes to forward any data traffic. However, even passive nodes broadcast hello messages in proactive mode to enable other nodes to send them data packets. Passive nodes have to be selected such that the network is still connected. The active nodes build then a backbone which can be used by all nodes in the network. The idea of using passive nodes is inspired by previous work [29, 30, 115] of other researchers. They have shown that it is not necessary that all nodes in a dense network participate in the dissemination of routing information in order to reach almost every node in a network.

3.5.5 Delay-based Approach

Routing protocols which use discrete routing metrics, e.g. number of neighbors or hop count, have the problem that the routing metric is not precise enough to clearly identify a single best route towards the destination. This fact may lead to problems in dense networks where there exist a large number of shortest paths between two nodes in the grid. Most protocols choose the first route which is established in order to distinguish between paths with equal costs. Another strategy is to randomly choose one of the best routes to distribute the traffic load evenly in the network [90]. However, in the case that the protocol relies on the first established route, the topology can be slightly modified by delayed forwarding of routing messages [9, 88, 120]. The forwarding delay can be used to apply additional metrics to a routing protocol or to optimize the forwarding process [88]. The advantage of the approach is that it can be integrated in most protocols without the need of major changes. In the following paragraphs, a delay-based approach is introduced which delays the forwarding of routing messages in respect to the change of the neighborhood of a node.

Routing Message Forwarding Delay

The packet loss in wireless networks strongly depends on the mobility of the network. The faster the speed of the nodes, the more frequent link breaks occur in the network [137]. Therefore many metrics for mobile networks focus on the link duration. The link duration represents the duration of the time interval during which two nodes are within transmission range of each other. Previous research concerned the link change rate [138] as routing metric while others used the average link duration [139] as stability criteria. The authors of [140] showed that the high average of the link duration results from a small fraction of links which have a high residual lifetime. Thus, they propose to select the link with the maximum expected residual lifetime according to the gathered statistical data. If the nodes are equipped with a GPS module, it is possible to predict the movement of the nodes in the near future [141]. They show that the duration of the remaining connectivity time between two nodes can be estimated if the motion parameters of two neighbors, e.g. speed, direction, and radio propagation range, are known. Furthermore, the clocks of the nodes have to be synchronized to allow a more accurate estimation.

In general, sensor nodes are not equipped with a GPS module due to the energy consumption of the module and the additional hardware costs. Moreover, many WSNs are designed for indoor scenarios where GPS position information is not available. Therefore, we decided to develop a new approach which neither requires position information nor clock synchronization.

The idea is to take advantage from nodes with correlated movement since the relative movement speed is responsible for link breaks [9] rather than the absolute speed. These nodes should be selected as next hop due to their more reliable links as a consequence of their correlated movement. The problem is to detect the correlated node movement if no position information is available.

Consider a scenario on a highway in one direction. There will be some fast driving cars, some with average speed, and some slow moving trucks. On the one hand, if there are more trucks on the road than other cars, the routing protocol should prefer trucks to forward data since their movement is strongly correlated resulting in a high link duration time and stable paths. On the other hand, fast movement cars are also able to build a stable network if they represent the majority of cars on the highway due to their correlated movement.

Thus, a metric is required which can be used to estimate the current relative movement speed of a node. Each node is able to estimate its speed by keeping track of the presence of its surrounding nodes. A change in its neighborhood indicates that the node has moved away from the other nodes or one of the other nodes has moved away from it.

Now consider another highway scenario in which there are only trucks and one fast driving car. The neighborhood of the fast driving car changes quickly since trucks frequently enter and leave the coverage area of the fast driving car. The trucks only recognize a single change in their neighborhood when the fast car enters and leaves their transmission range.

The basic idea is to delay the forwarding of routing messages depending on the change of the neighborhood. More changes in the neighborhood result in a higher delay of the routing messages. Therefore, the routing protocol chooses nodes with a lower relative node speed since these nodes forward routing information more quickly. It is obvious that a mobile node or a cluster of mobile nodes with correlated movement may have peaks in their neighborhood change if they move through certain areas, e.g. crossing of a road or an area of high node density. Thus, a mechanism is required to reduce the variation of the metric and make it robust against short temporary changes. The exponential weighted moving average algorithm is used to minimize the impact of peaks. The neighborhood change metric is calculated according to Eqn. 3.3

$$\epsilon_\tau = \alpha \cdot \epsilon_{\tau-1} + (1-\alpha) \cdot X_\tau, \alpha = 0.9. \tag{3.3}$$

The chosen smoothing factor α represented the best trade-off between reaction time and peak suppression in the simulated scenarios. X_τ is the number of changes in the neighborhood list during the last observation interval. Note that the number of changes in the neighborhood list reflects the number of detected link changes in the one hop neighborhood of the node.

Neighborhood Change Detection

Changes in the neighborhood list are counted during each observation interval.

The counter is increased by one each time a node is added to the list or removed from it.

Neighborhood List

The neighborhood list consists of neighbor entries. Each entry stores information about the neighbor, e.g. time of last contact. If a node receives a packet for another node it checks whether the originator of the packet is in the list. In the case that the node is not in the list a new entry is created and inserted. Otherwise, the existing corresponding entry is updated.

Neighbor Expiration Interval

The neighbor expiration interval has to be chosen carefully since its optimal duration depends on the traffic pattern. Each node keeps track of its surrounding nodes by periodically transmitting hello messages. To minimize changes in the neighborhood list if some nodes are temporarily unavailable, the neighbor expiration interval is set by default to four times the duration of the hello message interval.

Routing Message Forwarding Delay Calculation

The time a node delays the forwarding of a routing message is in the following referred to as forwarding delay. The delay δ is calculated from the neighborhood change metric ϵ of the last observation interval according to Eqn. 3.4

$$\delta = \frac{\Delta h}{\lambda} \cdot \left(1 - \frac{\phi}{\epsilon_\tau + \phi}\right). \tag{3.4}$$

The quotient of the hello message interval Δh and λ represents the maximum forwarding delay. Thus, λ covers the function of a delay limiter. The second factor of Eqn. 3.4 is influenced by ϕ, and is used to divide the maximum forwarding delay into smaller steps. A smaller ϕ value increases the delay for a smaller number of neighbor changes. A default ϕ value of ten is used since it results in a good accuracy of differentiation in a large spectrum of neighborhood list changes. The additional delay has to be chosen according to the net diameter in number of hops, the underlying medium access layer, and the traffic load of the network. In most scenarios, an additional delay of several milliseconds is quite sufficient to modify the topology of the network. The impact on the end-to-end delay of data traffic is minimized due to the fact that only routing messages are delayed.

Impact of the Delay-based Message Forwarding on AODV and SBR

The OPNET Modeler 14.5 is used to simulate the impact of the delay-based approach on the end-to-end reliability and the selection of the forwarding nodes of AODV and SBR with respect to their current relative speed. Instead of using the OPNET AODV model, the model is implemented as specified in the RFC3561 [65]. The physical layer is replaced by a disc model which limits the radio propagation range to 200 meters. Moreover, the signal strength is calculated according to a free space model in order to minimize side effects caused by asymmetric links. However, communication issues like interference are still considered as long as the corresponding nodes are within the radio propagation range of each other. Furthermore, the transmission data rate is set to 256 kb/s which corresponds to the transmission rate of typical low-power transceivers like the CC2420. The nodes use CSMA as MAC protocol with a simple back-off algorithm. At the beginning of each simulation 100 mobile nodes are randomly placed on a square of 1000 by 1000 meters.

All nodes move according to a random waypoint model [70] which is described in detail in Subsection 4.2.2. The pause time is set to zero seconds to generate continuous movement. In addition, the minimum node speed is set to 1 m/s which shortens the transient phase of the mobility model. The maximum node speed is increased from 2 m/s to 20 m/s in steps of 2 m/s to simulate the performance of the protocols under different node mobility levels. In general, only sinks are mobile in WSNs. However, we have chosen this scenario to evaluate the performance of the routing protocols since it is more challenging for the routing protocols due to the higher number of link breaks. 10 nodes in the net-

Table 3.2: *AODV Configuration*

Active Route Timeout	1.5 s
Hello Message Interval	0.8 s
Net Diameter	12
Net Traversal Time	1.4 s
Node Traversal Time	0.02 s

Table 3.3: *SBR Configuration*

Active Route Timeout	3.0 s
Hello Message Interval	4.0 s
Decrease Routing Interval	4.0 s
Short Hello Message Interval	4.0 s
Short Hello Message TTL	1
Hello Message Time-To-Live	12

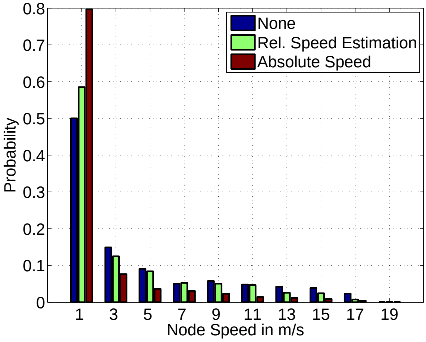
work select a random destination at the beginning of the simulation. These nodes generate packets according to an exponential distribution with a mean value of 2 seconds and a constant packet size of 1024 bits. The packet size is the maximum packet size of typical low-power transceivers like the CC2420. The traffic model is started after 300 seconds to minimize the impact of the transient phase of the random waypoint model.

The duration of the simulations is set to 1400 seconds. Statistics are collected after 400 seconds to allow the stabilization of the network. The results are calculated from 20 simulation runs with different seeds of the traffic and the mobility model. All error bars show the 99 percent confidence level of the collected statistics whereas histograms represent the average of 20 simulation runs. The configurations of the routing protocols are shown in Table 3.2 and Table 3.3. Instead of using the default configuration of the AODV protocol, we decided to use an optimized configuration in order to allow a more meaningful comparison. The default configuration of the AODV protocol uses an active route timeout of 3 seconds and a hello message interval of 1 seconds. These intervals are too long to provide acceptable performance in mobile networks. Thus, we use a shorter timeout and a smaller hello message interval to minimize the period of time which is required by the protocol to detect topology changes. However, the chosen configurations leave enough room for performance improvements of the end-to-end reliability of the protocols such that the impact of the delay-based approach can be investigated.

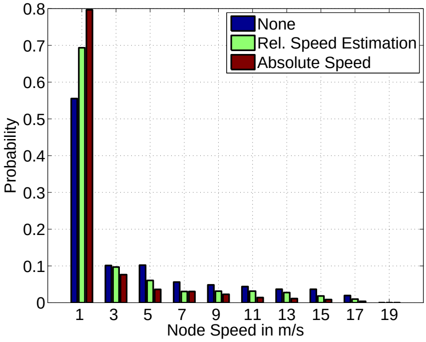
Three different delay metrics are used to simulate the impact that the delayed forwarding of routing messages has on the next hop selection and on the reliability. The first metric forwards the routing messages immediately. The second metric calculates the forwarding delay according to the approach presented in this section and is in the following referred to as Relative Speed Estimation (RSE) metric. The forwarding delay of the third metric is chosen with respect to the current absolute node speed.

The end-to-end reliability of routing protocols is strongly correlated to the movement pattern of the nodes in the network. For this reason, a closer look is taken on the absolute speed of the nodes at the time they forward a packet. Fig. 3.9(a) and Fig. 3.9(b) present the normalized histogram of the absolute speed of the nodes when forwarding a packet. Thus, a sample was collected each time a node forwards a data packet. The results of both figures show that slow moving nodes forward more traffic than faster nodes. This behavior is independent of the protocol and the used metric. The shape of the distribution results from the node speed distribution which is caused by the random waypoint mobility model [112]. A detailed description of the random waypoint mobility model and its characteristic node speed distribution is given in Subsection 4.2.2.

The figures point out that the node speed distribution of the forwarding nodes can be modified by the delayed forwarding of the routing information. However, a lower absolute node speed does not necessarily result in a longer link duration since nodes could still move in opposite directions. If a cluster of nodes move in the same direction with similar the speed their absolute speed can be neglected in contrast to their relative speed.



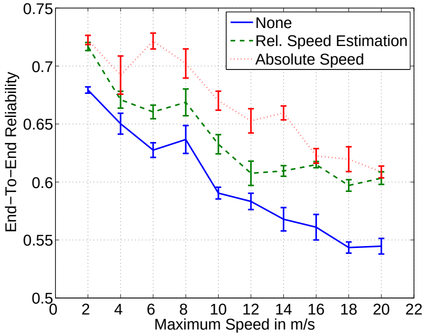
(a) AODV - Forwarding Node Speed Distribution



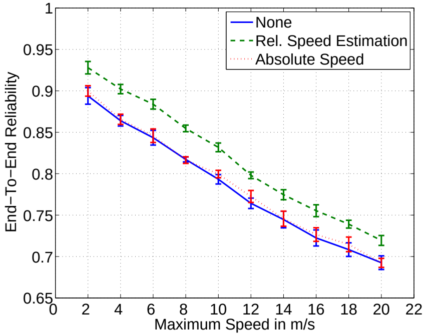
(b) SBR - Forwarding Node Speed Distribution

Figure 3.9: *Delay-based Approach - Forwarding Node Speed Distribution*

The focus of this subsection lies on the reliability since our test scenarios revealed that the delay of routing messages has no significant impact on the average end-to-end delay of data packets. Fig. 3.10(a) and Fig. 3.10(b) show the reliability of the different metrics and protocols depending on the maximum speed.



(a) AODV - Reliability



(b) SBR - Reliability

Figure 3.10: *Delay-based Approach - Reliability Improvement*

The results of Fig. 3.10(a) indicate that the RSE and the absolute speed metric increase the performance of AODV in the case of more mobile scenarios. The higher end-to-end reliability results from more stable routes as a consequence of the longer link duration time. End-to-end reliability describes the probability that a packet is successfully received by the sink via one or multiple hops.

Fig. 3.10(b) points out that the SBR protocol achieves the highest end-to-end reliability if its routing messages are delayed according to the RSE metric which is in contrast to AODV where the absolute speed metric offered a slightly better performance than the RSE metric. This effect can be explained as follows. Hello messages are periodically flooded by the SBR protocol in the network which results in a higher routing overhead. Thus, the nodes have more precise information of their neighborhood. For this reason, the SBR protocol takes more advantage from the delayed forwarding of routing messages than AODV.

3.5.6 Simulative Performance Evaluation of the SBR Protocol

The performance of OLSR, AODV, and the hybrid version of the SBR protocol are compared in this subsection. OLSR and AODV are implemented as specified in the corresponding RFCs. The only exception is represented by the valid time interval of the topology control messages in OLSR which is decreased to topology control interval plus hello interval. This change is necessary in mobile scenarios to minimize the period of time which is required by OLSR to detect topology changes. Furthermore, it prevents the propagation of outdated routing information.

The nodes can transmit data up to 256 kB/s within a radio range of 150 meters. CSMA is used as medium access protocol. 50 mobile nodes are randomly placed on a square of 1000 x 1000 meters. Thus, the node degree in this scenario is smaller than in the previous one in order to minimize the variance of the routing protocol overhead. The movement is generated by a random waypoint model with a minimum node speed of 1 m/s to prevent nodes from moving very slowly for a long period of time if the next waypoint is far away from the current position. Furthermore, the nodes do not stop at a waypoint since the pause duration is set to zero seconds. The maximum speed of the random waypoint model is increased from 1 m/s to 15 m/s in steps of 2 m/s.

All nodes generate packets according to an exponential distribution with a mean value of 10 seconds and a constant packet size of 1024 bits. The higher mean value of the packet inter-arrival time allows us to simulate the impact of the inter-arrival time on the network performance. Therefore, we simulate the performance of the protocols under the same traffic load but with different traffic patterns.

Due to the short radio range and the high mobility, some nodes temporarily have no neighbors. For this reason, 100 percent end-to-end reliability is not achievable. Thus, the collected simulation results represent a relative performance comparison. If a larger amount of overhead is taken into account, the protocols can achieve a slightly higher end-to-end reliability.

Nonetheless, the scenario gives a good picture of how much mobility the protocols can handle. It is obvious that the end-to-end reliability can be increased by decreasing the flooding intervals of proactive protocols. Therefore, only protocol configurations are considered where the total amount of generated overhead is on a reasonable level. The short hello interval and active route timeout of AODV are necessary to compensate the frequent topology changes. Note that the parameters serve different purposes in the protocols. Hello messages in AODV and OLSR are used for neighbor detection and maintaining two hop neighborhood lists whereas SBR disseminates the messages to spread routing information across the whole network. The size of the messages is also different. In contrast to AODV and SBR, OLSR hello messages may become very large in dense networks since they contain a neighborhood list. The configurations of the routing protocols are shown in Tables 3.4 to 3.6.

The duration of each simulation run is 1100 seconds. Statistics are collected after a 100 second transient phase to mitigate the impact of the starting positions of the nodes. A 100 second transient phase is sufficient for this scenario due to the fact that the simulation represents a relative performance comparison.

The results are calculated from 20 simulation runs with different seeds of the mobility model. The seeds of the traffic models are set to constant values to reduce the variance of the simulation results and allow a better comparison of the

Table 3.4: *AODV Configuration*

Active Route Timeout	0.50 s
Hello Interval	0.25 s
Net Diameter	16
Net Traversal Time	0.35 s
Path Discovery Time	0.70 s
Node Traversal Time	0.02 s

Table 3.5: *OLSR Configuration*

Hello Interval	2.00 s
Refresh Interval	2.00 s
Duplication Hold Time	5.00 s
Topology Control Interval	4.00 s
Topology Control Expiration	6.00 s
Max Jitter	0.05 s

Table 3.6: *SBR Configuration*

Active Route Timeout	4.00 s
Hello Message Interval	1.00 s
Decrease Routing Interval	1.00 s
Hello Message Time-To-Live	16

routing protocols. All error bars show the 99 percent confidence level of the collected statistics.

Scenario A

In Scenario A each of the 50 mobile nodes selects a random destination at the beginning of the simulation. Thus, some nodes receive traffic from more than one node which results in a short packet inter-arrival time. The inter-arrival time has a large impact on the performance of reactive protocols, as soon as the inter-arrival time is larger than the active route timeout. If no traffic is received for a time span longer than the active route timeout, the route is marked as inactive. As a result, the route has to be re-established or locally repaired for the next packet transmission. Due to the fact that some nodes are not selected as destination they do not generate routing overhead in reactive and hybrid protocols. The results of Fig. 3.11 show that OLSR achieves the highest end-to-end reliability if the nodes are moving very slowly. As a consequence, the neighborhood of the nodes changes slowly, too. For this reason, the entries in the routing table of the SBR protocol increase to a very high level. Therefore, the time until the values of unreachable

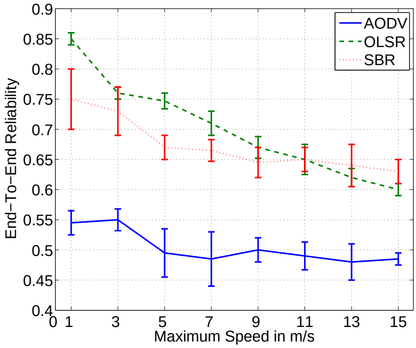


Figure 3.11: *Scenario A - Reliability*

nodes fall behind the values of new neighbors is longer than in more mobile scenarios. Thus, the protocol assumes that some nodes are still reachable because of their high routing entry. The protocol can be modified such that the entries in the routing table decrease faster. A more frequent routing table update allows faster topology change detection in SBR. The routing table update frequency can only be increased to a certain level depending on the hello message interval. Too frequent updates result in low routing entry values which limits the load-balancing and multi-path capability of the protocol.

The low end-to-end reliability of AODV results from the active route timeout. A link break is not detected until the connection to a node along the route expires. Thus, AODV may try to send traffic to a node for the duration of the active route timeout without recognizing that this node is not reachable anymore. Additionally, AODV has no fall back solution similar to SBR. SBR can select the former second best next hop as forwarding node if the next hop is unreachable. In contrast to SBR, AODV tries to re-establish a whole route or locally repair a broken route.

Fig. 3.12 points out that the routing overhead of AODV and SBR remains on the same level independent of the maximum node speed whereas the overhead of OLSR decreases. However, OLSR generates much more routing overhead than AODV and SBR in scenarios with low mobility.

Scenario B

The traffic pattern in Scenario B is changed such that a single node is randomly selected by all other nodes as destination at the beginning of the simulation. The accumulated traffic leads to short packet inter-arrival times at the selected destination. Fig. 3.13 shows that the end-to-end reliability of the protocols in Scenario B is similar to the reliability which is achieved in Scenario A. This behavior is not surprising since the results in Scenario A represent the average accessibility of the nodes in the network. The results in Fig. 3.14 point out that the routing overhead of the SBR protocol is several times smaller than the overhead of AODV and OLSR while maintaining a high end-to-end reliability in this scenario. The

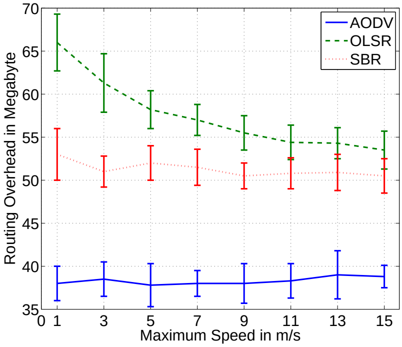


Figure 3.12: *Scenario A - Overhead*

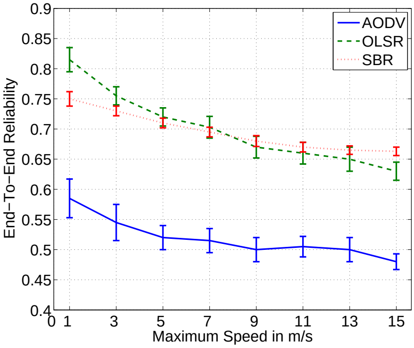


Figure 3.13: *Scenario B - Reliability*

low overhead can be explained as follows. Only data sinks transmit hello messages in SBR. Thus, the generated overhead directly corresponds to the number of data sinks which makes the protocol most applicable in networks where the number of data sinks is small. Another reason for the large difference in routing

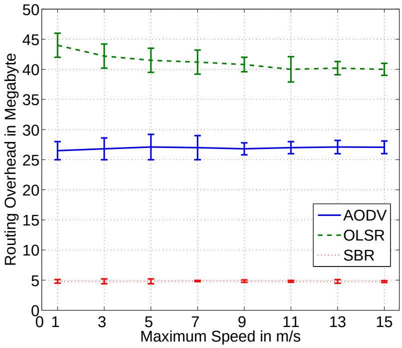


Figure 3.14: *Scenario B - Overhead*

overhead lies in the fact that all nodes take advantage from the hello messages which are transmitted by the data sink since they can use the messages to update their routing table entries.

A closer look at Fig. 3.14 reveals that the overhead generated by AODV in Scenario B is smaller than in Scenario A. Two circumstances are responsible for the reduced routing overhead. In AODV, each data source has to transmit a RREQ to gain knowledge of the route to the destination. For this reason, the nodes around the destination receive and forward route replies which are transmitted by the destination. Therefore, they know how to reach the destination and can answer received RREQs directly with an intermediate route reply instead of forwarding the request to the destination. Thus, the flooding of requests is reduced in a significant way.

Additionally, the nodes in the neighborhood of the destination have to forward data packets more frequently to the destination than nodes further away from the destination. Note that each node transmits packets according to an exponential distribution with a mean value of 10 seconds which is much higher than the chosen active route timeout of AODV. For this reason, less route timeouts occur in the Scenario B since the data traffic accumulates on its way to the destination. As a consequence of the accumulated traffic, the nodes around the destination have recent knowledge about their surrounding nodes. The smaller number of route timeouts results in a smaller number of route requests which further reduces the overhead of AODV. The more accurate knowledge of the nodes around the destination minimizes the number of routing messages which are needed to locally repair a broken route in this particular area of the network.

The routing overhead generated by OLSR remains on a constant level independent of the mobility in Scenario B. Furthermore, the overhead is smaller than in Scenario A. In this case, the different traffic pattern has a great impact on the one hop and two hop neighborhood lists of the nodes. The accumulated traffic allows the nodes which are close to the destination to keep their neighborhood lists up-to-date by listening to the transmissions of their surrounding nodes. Thus, the nodes are able to calculate their MPR set more precisely which decreases the number of retransmissions which are necessary to distribute the topology control messages in the network.

In the previous scenarios, the performance of the SBR protocol was simulated under various levels of mobility. The protocol achieved a high performance in the simulated scenarios while generating a low amount of routing overhead. However, we used optimized configurations for all protocols in order to allow a meaningful comparison. Some protocols like AODV and OLSR have to be configured very carefully since several parameters depend on each other. As a consequence, most users rely on the default configurations of the protocols which are optimized for networks with very low mobility. The strength of the SBR protocol is its simplicity. The HMI and the DRVI are the two most relevant parameters to tune the behavior of the protocols. Moreover, the default configuration provides a good performance in a large range of scenarios due to the adaptive and the cumulative characteristics of the routing metric. A corresponding parameter study is given in Section 4.3 where we simulate the performance of a multimedia application in a mobile network depending on the HMI of the SBR protocol.

3.6 Summary

Three different classifications of routing protocols were discussed in this chapter. It was shown that the protocols can be classified according to the way they establish routes, the network structure or the protocol operation. Moreover, a closer look was taken on the elementary tasks which have to be performed by a routing protocol. Forwarding, processing, topology establishment and the dissemination of routing information are identified as the most important tasks. The topology of a network is always optimized with respect to a certain metric. Routing metrics may have different characteristics which have to be taken into account. Thus, a taxonomy of routing metrics was given which classifies the metrics depending on whether they are discrete, continuous or a combination of several metrics. In addition, the impact of slow changing and fast changing metrics was outlined. Furthermore, a brief overview of techniques was given which mitigate the problem caused by fast changing metrics. Reactive and proactive routing protocols were described in more detail in order to give a better understanding of the different kinds of strategies which can be applied to optimize the performance in a multi-hop wireless network.

The SBR protocol was introduced which combines reactive and proactive mechanisms to achieve optimal performance in various situations. The protocol was originally designed to meet the requirements of low-power WSNs but can be easily configured such that it achieves a high performance in mobile mesh and ad hoc networks as well. The applied routing metric is cumulative and has adaptive characteristics. Thus, the protocol is able to quickly detect topology changes while maintaining stable routes. It can apply any routing metric by deferring the forwarding of routing messages according to the current routing value. This delay-based mechanism was introduced by giving an example how to improve the performance in mobile networks by deferring the forwarding of routing messages with respect to the changes in the neighborhood of a node. Finally, the performance of the SBR protocol was compared with the performance of the reactive AODV and the proactive OLSR protocol.

4 Evaluation of Routing Protocols

Routing protocols are often designed to meet the requirements of a certain application. Thus, the protocols are optimized to achieve a high performance under specific conditions, e.g. reliable links, constant bit rate traffic patterns, and low mobility. However, it is hard to estimate the performance of routing protocols in advance since their performance is not solely influenced by their configuration. Besides the configuration, the performance is strongly affected by the traffic pattern, the movement of the nodes, the spatial node distribution, the underlying MAC and physical layer.

The majority of the protocols show a predictable behavior in standard scenarios [128] as long as their configurations are close to the default settings.It is often assumed that the nodes have no or little mobility. Furthermore, the default configurations do not consider the available bandwidth of the underlying MAC protocol. In this case, minor changes to the configuration result in predictable changes of the protocol performance, e.g. an increase of the hello message interval results in a decrease of the routing overhead.

Detailed knowledge of a routing protocol is required to estimate its performance in non-standard scenarios. Mobility becomes a challenging problem for protocols which try to establish end-to-end connectivity due to the fact that the network topology changes frequently [112]. Moreover, wireless communication issues, like unidirectional and unreliable links, have to be taken into account which also have a great influence on the performance of the protocols [142].

Performance comparison of different routing protocols in a realistic testbed is often not possible. Thus, an estimation of their performance is typically given by analysis or simulation. State-of-the-art routing protocols like DSR or OLSR are very complex which makes it almost impossible to estimate their performance in advance.

In addition, side-effects caused by the unreliable wireless communication and the movement of the nodes can hardly be covered by analysis. Therefore, simulation represents the first choice to estimate the network performance if measurements in a large testbed under realistic conditions are not possible.However, results from stochastic simulations have to be evaluated carefully in order to produce trustworthy results [143]. The comparison of routing protocols requires a software or hardware framework which is able to provide the needed functionality to the protocols, e.g. remaining battery power, link quality or position information.

In this chapter, a modular simulation framework [5, 11] is introduced which was used to evaluate and optimize MAC and routing protocols for WSNs. Furthermore, a closer look is taken on common mobility patterns and their impact on the performance of the network. Moreover, the most important issues of simulative performance evaluation of routing protocols are discussed. Routing protocols can be optimized in many different ways depending on the characteristics of the protocols and the corresponding routing metric. Therefore, different optimization techniques are introduced which can be applied to reactive and proactive protocols. Simulation results should be always validated by measurements - if possible - in order to proof the correctness of the simulation.

However, a single simulation framework might not provide sufficient functionality. Thus, a simulation framework can be extended by other simulations to build a larger more realistic co-simulation framework where different simulation tools interact with each other. Another interesting approach is represented by hardware-in-the-loop simulations which are a trade-off between simulations and testbeds. The different evaluation techniques are described in more detail. Finally, the chapter is concluded with a summary.

4.1 Simulation Framework

The performance evaluation of routing protocols represents a challenging task since their performance is affected by both, higher and lower layers. Thus, the performance of the protocols has to be simulated in various scenarios with different configurations to allow a meaningful performance comparison.

For this reason, a modular framework was implemented in OPNET which consists of several process models. A process model can be regarded as a finite state machine which has a large number of interfaces to interact with other process models. A process model typically covers the functionality of an ISO/OSI layer. The framework consists of three parts which are responsible for different tasks. The first part is represented by a set of process models which build the communication stack of each sensor node. This set of process models is supplemented by a mobility process and a small energy consumption framework. The second part is responsible for collecting global statistics. The global statistic collection is implemented by a central process which receives callbacks from other process models. The third part is represented by process models which offer interface functionality between OPNET and other simulation tools. Moreover, these process models provide gateway functionality between the simulated virtual network and the real network.

The advantage of this modular framework is that a single module can be exchanged without the need of modifying the rest of the framework. Furthermore, modules which are not needed can be easily removed to speed-up the simulation. In addition, the framework has interfaces to interact with other simulations in order to create larger co-simulations. The interfaces also offer the possibility to interact with the real network. Thus, real network traffic can be routed through the simulated virtual network to get a better picture of the perceived quality.

First, a short introduction to the OPNET Modeler is given in this section. The introduction is followed by a description of the basic mechanisms of the framework. Moreover, co-simulation and hardware-in-the-loop functionalities are discussed which are also part of the simulation framework.

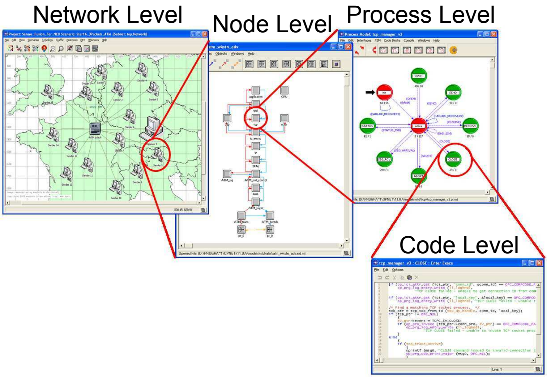
4.1.1 Introduction to the OPNET Modeler

OPNET Modeler is a commercial discrete-event simulator which is optimized for analyzing and designing communication networks and protocols. The software is able to simulate the whole ISO/OSI stack including a simplified physical layer. The standard physical layer uses a free space propagation model [59]. Nonetheless, more complex propagation models, e.g. Longley-Rice [144] or Walfish-Ikegami [145], can be applied which also take terrain information into account. However, complex propagation models which consider signal propagation relevant terrain effects, like multi-path and shadowing, slow down the simulation. For this reason, most simulations use the free space propagation model.

The software uses a hierarchical structure to modify or build a network. The top level is represented by the network level. On this level it is possible to drag an drop existing objects, e.g. workstations, routers, switches, and sensor nodes, into the simulation scenario. Each object is described by a node model which consists of several process models. A process model is responsible for a certain task and may interact with other process models. This modular functionality is often used to build a structure on the node level which reflects the ISO/OSI layers as shown in Fig. 4.1.

The behavior of a process model is defined by a finite state machine. The finite state machine consists of one or more states and transitions. Each state is divided in an enter-state and an exit-state. The enter-state is executed if the finite state machine enters this state while the exit-state of the current state is executed each time the process receives an interrupt. Note that the conditions of transitions are only checked after the execution of the exit-state. Thus, the finite state machine may only change its state if an interrupt was scheduled for it. A process receives an interrupt if the simulation event, which is currently executed, points to this process.

Interrupts are usually triggered by timers or caused by packet arrivals which are then evaluated in the exit-state of the current state or in the enter-state of the following state. A transition always points from an exit-state to an enter-state. The



```
if (op_ici_attr_get (ici_ptr, "conn_id", &conn_id) == OPC_COMPCODE_F
    op_prg_log_entry_write (ll_loghndl,
            "TCP CLOSE failed - unable to get connection ID from com

if (op_ici_attr_get (ici_ptr, "local_key", &local_key) == OPC_COMPCO
    op_prg_log_entry_write (ll_loghndl, "TCP CLOSE failed - unable t

/* Find a matching TCP socket process.  */
tcb_ptr = tcp_tcb_from_id (tcp_dt_handle, conn_id, local_key);
if (tcb_ptr != OPC_NIL)
    {
    ev_ptr->event = TCPC_EV_CLOSE;
    if (op_pro_invoke (tcb_ptr->conn_pro, ev_ptr) == OPC_COMPCODE_FA
        op_prg_log_entry_write (ll_loghndl,
                "TCP CLOSE failed - unable to invoke TCP socket proc
    }
else
    {
    if (tcp_trace_active)
        {
        sprintf (msg0, "CLOSE command issued to invalid connection (
        op_prg_odb_print_major (msg0, OPC_NIL);
```

Figure 4.1: *OPNET - Levels of Design*

functions which are executed inside each state are written in C code which is then used by the OPNET Modeler to build the simulation. The simulation time does not advance during the execution of an enter-state or an exit-state. The simulation time only advances between two consecutive events which is a characteristic of discrete-event simulators.

4.1.2 Basic Framework

The basic framework is a library which consists of several process models. The process models can be combined and configured to simulate wireless nodes. Fig. 4.2 shows a node model which we implemented to compare the performance of routing protocols in mobile wireless networks [8]. The arrows with the solid lines indicate the packet flow while the dashed lines represent statistic-wires. Statistic-wires are used by the OPNET kernel to propagate changes from one module to another module. State changes of the receiver module radio$_{rx}$ and the transmitter module radio$_{tx}$ are forwarded to the MAC module via statistic-wires. The receiver and the transmitter are connected to the same antenna pattern which

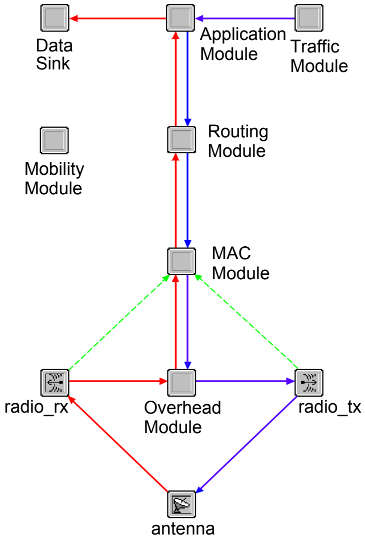


Figure 4.2: *Basic Framework - Node Model*

specifies the antenna gain and the orientation of the antenna. However, it is also possible to use different antenna patterns for the receiver and transmitter modules. Furthermore, a node model may have multiple receiver and transmitter modules. In the following paragraphs the basic modules are described which build the core functionality of our framework.

Traffic Module

The Traffic Module is based on the OPNET standard traffic generation process, but has advanced features. The module offers the possibility to generate single packets and data bursts. Various distribution functions are supported which can be used to generate the packet inter-arrival time, the burst inter-arrival time, the packet size and the number of packets per burst.

In addition, trace files captured by Wireshark [1] or tcpdump [2] can be used to replay captured traffic. This feature was used in [11, 12] to generate data packets according to a previously recorded video trace. Moreover, external network traffic can be captured and routed through the virtual network as long as the simulation runs in real-time. Simulations which interact with a real network or devices are referred to as hardware-in-the-loop simulations. Hardware-in-the-loop simulations have several requirements on the simulation software which are discussed in detail in Subsection 4.1.4.

Application Module

The application module is used to modify the incoming and outgoing data packets in order to simulate the behavior of different applications, like the buffer of a video application or data aggregation. In addition, different kinds of retransmission strategies and acknowledgment mechanisms are implemented to determine which strategy offers the best performance in the simulated scenario. Note that the traffic pattern has a large impact on the performance of reactive and hybrid routing protocols since routes are only established and maintained on demand [11]. Thus, it is essential to simulate the behavior of the application layer to get more accurate results.

Routing Module

The network layer is represented by the routing module. AODV, OLSR, GBR, MCFA and SBR are currently part of the framework. In fact, AODV and OLSR are already included in the OPNET Modeler library. However, both were re-implemented to speed-up the simulation since most of their features, like multiple gateway support, are not required in most WSN scenarios.

MAC Module

The MAC module is implemented such that typical transceiver characteristics,

---

[1] *Wireshark - Network Protocol Analyzer, http://www.wireshark.org/*

[2] *tcpdump - Packet Analyzer, http://www.tcpdump.org/*

like the CCA delay [32, 50] and the turnaround time [51, 146], can be simulated. Statistic-wires are used to connect the transmitter and the receiver with the MAC process model. Thus, the MAC module is informed about their state changes, e.g. busy radio channel, free radio channel or the end of the current transmission. The gathered information can also be used to calculate the energy consumption of a node.

The framework library includes the following MAC protocols, ALOHA, CSMA, CSMA-TBEBA, BP-MAC, BPS-MAC and X-MAC. However, it is possible to replace the data link layer and the physical layer of the framework with the OPNET standard models, e.g. to simulate the IEEE 802.11 MAC and physical layer as shown in [12].

Overhead Module

The overhead module is added to the framework in order to evaluate incoming and outgoing packets. It generates local routing and MAC overhead statistics. In addition, the statistics are forwarded to a central node which generates global protocol overhead statistics. These statistics are forwarded by using an interrupt process routine which is provided by the OPNET kernel. The routine enables process models of different node models to directly communicate with each other [5].

Physical Layer

This paragraph gives a short overview of the radio transceiver pipeline which is used by the OPNET Modeler [59] to simulate the physical layer. The OPNET Modeler divides the physical layer into 14 pipeline stages as shown in Fig. 4.3. A packet is only forwarded to a receiver by the simulation kernel if it has successfully passed all pipeline stages. Stage 0 is executed once at the beginning of the simulation. This stage checks the settings of all receiver and transmitter pairs within the simulation and determines whether communication between a pair is possible. This mechanism speeds up the simulation if the nodes transmit on different channels. The transmission delay is calculated once per transmission since the nodes may have moved or changed the orientation of their antennas. The link

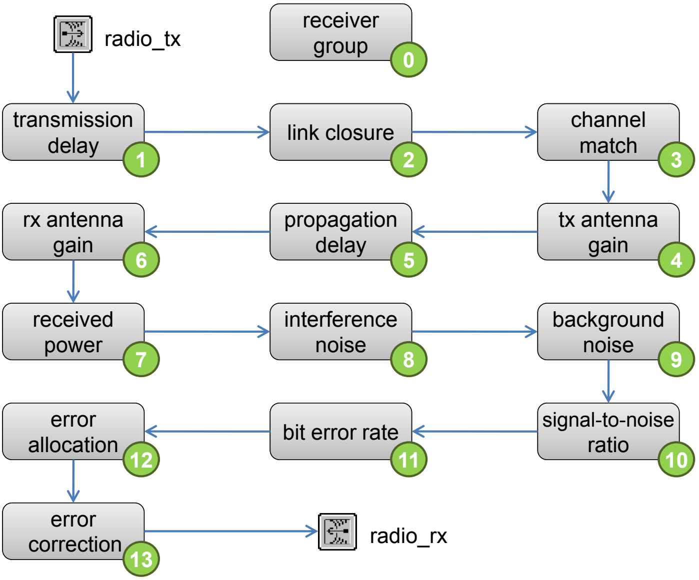


Figure 4.3: *OPNET Modeler - Physical Layer*

closure stage determines whether the terrain or obstacles are blocking the signal propagation such that no communication between this pair of transmitter and receiver is possible. The packet is marked as invalid in the case that it does not pass the link closure stage. Furthermore, the simulation aborts the execution of the pipeline due to the fact that there is no further need to calculate the other stages. Stage 3 compares the settings of the transceiver and the receiver to classify the transmission. The transmission is marked with a valid, a noise or an ignore flag according to the transmission settings. The calculation is aborted if the packet is marked with an ignore flag due to a channel mismatch. Therefore, the channel match pipeline stage represents the optimal stage to modify the simulated signal propagation.

A simple disc model can be simulated if the stage sets the flag of former valid packets to noise if the distance between the receiver antenna and the antenna of the transmitter is longer than the radius of the disc model. Thus, two nodes are only able to exchange packets if the distance between them is shorter than the radius of the disc model. However, the model still considers packets of further distant nodes as noise. The disc model provides a simple way to create the desired node degree in the simulation [137]. Otherwise, the scenario size and the transmission power have to be chosen such that the desired node density is achieved. Moreover, it is possible to reduce the required computational power of the simulation if the interference range is limited. Packets are marked with an ignore flag if the distance between the transmitter and the receiver is longer than the interference range as shown in Fig. 4.4. Note that the simulation kernel aborts the execution of the pipeline stage if a packet is marked with an ignore

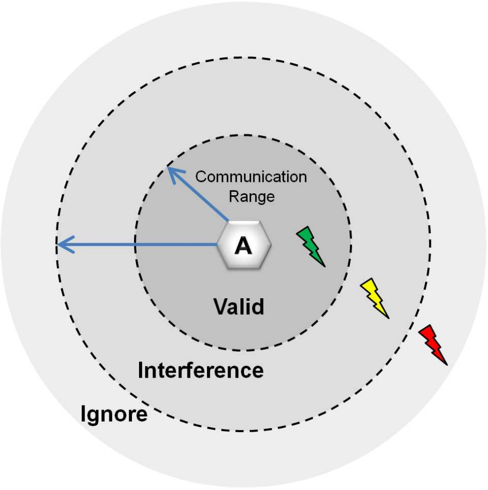


Figure 4.4: *Pipeline Stage - Disc Model*

flag which can be exploited to speed-up the simulation, especially in dense and large network simulations. The tx antenna gain stage calculates the transmission gain depending on the antenna pattern and the orientation of the antenna towards the receiver antenna. Stage 5 is responsible for calculating the propagation delay. The antenna gain of the receiver is calculated in stage 6. Stage 7 returns the received power with respect to transmission power, distance between the antennas, antenna gains, and transmission frequency.

Stages 8 through 12 are called one or several times for a single transmission depending on the interference of other transmissions. Fig. 4.5 shows an example where two transmissions collide as a consequence of the hidden-node problem. Receiver 1 and receiver 2 are very close to each other and within the overlapping transmission range of node 1 and node 2. Moreover, node 1 and node 2 are not within transmission range. Thus, node 2 is not aware of the ongoing transmission and starts its own transmission which results in a collision at receiver 1 and receiver 2. Therefore, stages 8 through 12 are called twice for both transmissions. The allocated errors for each part are accumulated and stored in the corresponding packet. Finally, the packet is passed to the last pipeline stage which is responsible for the error correction. The stage determines whether the packet is forwarded to the receiver of the destination or discarded by the simulation kernel.

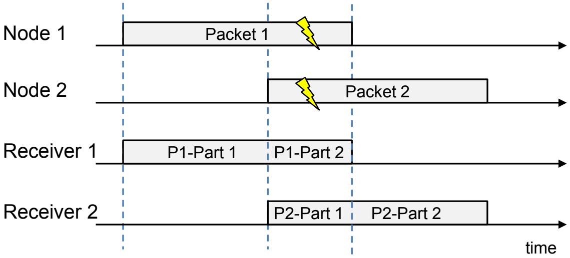


Figure 4.5: *OPNET Modeler - Collision example*

Mobility Module
The mobility module describes the movement of a node within the boundaries of the simulation scenario. Our implementation is based on vectorized movement which has several advantages since the precision of the nodes positions can be set in the simulation kernel. The precision of the position information has a great impact on the computational power which is required by the OPNET Modeler, especially in large networks where all nodes are mobile. The mobility module calculates the direction and the speed which is required to get from the current position to the target position within a certain period of time. The mobility module is able to generate movement according to the random waypoint [70], random walk [147], random direction [148], manhattan [149], and the reference group mobility model [138]. The characteristics of the mobility models are not further described in this section since they are introduced in detail in Section 4.2.

Data Sink Module
The data sink module calculates the delay and the jitter of incoming packets. Moreover, it generates local statistics, e.g. the number of received packets and the received data rate. The collected statistics are also forwarded to a central node which generates global statistics.

4.1.3 Co-operative Simulation

Co-simulation is a simulation technique where different simulation tools are running in a co-operative way. The tools are usually synchronized and are responsible for the simulation of individual components which are part of a larger simulation. Co-simulation offers a wide range of advantages since complex simulations can be divided into smaller ones which exchange information during the simulation. In addition, it allows distributing the simulation tools on different computers in order to speed-up the simulation by taking advantage from higher computational power. Moreover, optimized simulation tools can be used to simulate the individual components regardless of the computing platforms.

However, the simulation time of the individual tools has to be synchronized which requires additional software. The most popular synchronization interface is represented by the High Level Architecture (HLA) IEEE 1516-2000 standard for modeling and simulation [150]. The synchronization and the exchange of information of the individual tools are controlled by the Run Time Infrastructure (RTI) which is a central component. The participating simulation tools may connect to the RTI via UDP or TCP. Therefore, it is guaranteed that the exchange of information is platform independent.

The simulation kernel of the OPNET Modeler has an HLA interface which synchronizes OPNET with the simulation time of the RTI. Furthermore, HLA messages can be directly mapped to virtual data packets which are sent to process models in the OPNET simulation environment. Thus, external simulation tools can communicate directly with the process models. This mechanism can be used, e.g. to modify the position and the orientation of the nodes in the OPNET simulation. The position information can be sent, e.g. from a flight simulator to a MATLAB simulation, which is also participating in the HLA network. The MATLAB software then transforms the gathered information into a format which is supported by the OPNET Modeler and forwards it to the corresponding process model in the OPNET simulation.

4.1.4 Hardware-In-The-Loop

In general, it is not possible to predict the performance of a wireless network in advance due to the fact that the communication stacks are too complex. Analysis and simulation provide the first step to estimate the network performance. Another way is to take measurements from a testbed in order to get a better picture of the impact of the different network characteristics. Nonetheless, testbeds are usually much smaller than the target network. Thus, they are not able to cover all aspects of the target network, e.g. the behavior and the performance of the network during a broadcast or temporary congestion.

The idea of hardware-in-the-loop simulations is to create a virtual network which is able to communicate and to interact with a real network. In the following, the term virtual is used to point out that the corresponding packet or object is part of the software domain while the term real is used to clearly identify an existing piece of hardware or network. It is possible to generate conditions similar to those in a large network by extending a small testbed with a virtual network as shown in Fig. 4.6. The figure shows a standard approach where a node in the network functions as a gateway in order to connect the hardware domain with the software domain. The gateway is divided into two parts. The first part is responsible for the communication in the real network. Thus, the real gateway captures traffic from the real network and forwards it to the virtual gateway. The virtual gateway is a process within the OPNET simulation which maps each captured packet to a virtual packet. Moreover, the virtual packet is transmitted in the simulation via the interface node. The sink modules of the virtual nodes are modified such that they give a feedback to the virtual gateway whether the packet was successfully received or discarded. The virtual nodes may also communicate with the real sensor nodes. In this case, the virtual interface node passes the received virtual packet to the real gateway which creates the corresponding packet and transmits the packet via the real interface node.

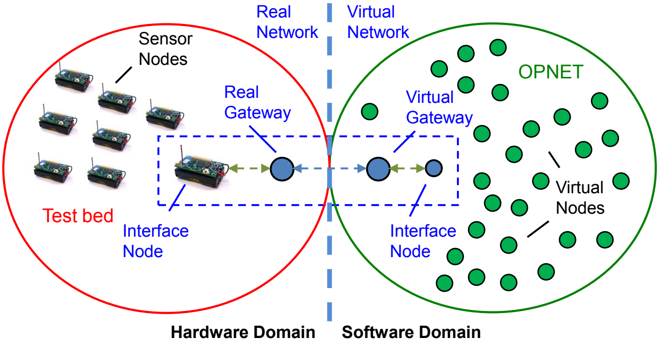


Figure 4.6: *Hardware-In-The-Loop - Testbed Extension*

Network performance is often characterized by standard metrics, like the average delay and the packet loss rate. However, these metrics do not necessarily represent the perceived network performance of a user [151], especially in the case that the received data is further processed. Consider an encoded video transmission where the loss of packets will lead to decoding errors at the decoder/player while delay can cause buffer under-runs. In both cases images are lost at the player which usually freezes the video. Note that modern video-codecs compress the original video by encoding only the differences between consecutive frames. Thus, the loss of a single image results in a distortion of all following images which are encoded based on the lost image. For this reason, it is important to simulate the transmission of the original application traffic over a virtual network as shown in Fig. 4.7.

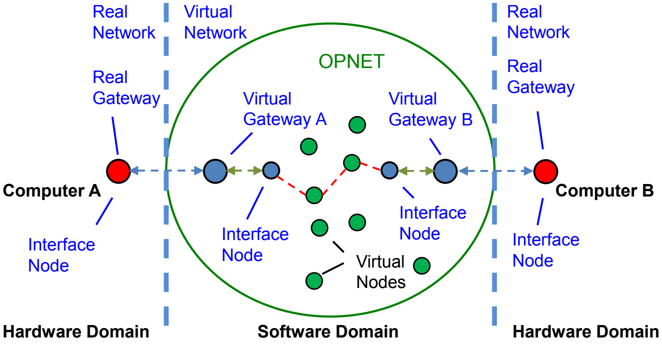


Figure 4.7: *Hardware-In-The-Loop - Virtual Network*

The simulation should run on a computer with at least two network interfaces to minimize the delay caused by the real network traffic such that it can be neglected compared to the virtual delay of the simulated network. Note that the simulation has to run in real time in order to correctly delay packets according to the delay of the simulated virtual network. Otherwise, the simulated delay is either too high in case that the simulation time proceeds slower than real-time or too low if the simulation is executed faster than real-time. The simulation discards real packets which are not successfully transmitted in the virtual network. Thus, the perceived application quality on computer A and computer B is affected by the performance of the simulated network. Therefore, it is possible to get a direct feedback from the simulation.

4.1.5 Video Quality Evaluation

The standard method to assess the performance of video transmission systems is to calculate the Peak Signal-to-Noise Ratio (PSNR) between the source and the received (possibly distorted) video sequence. It is a differential metric which is calculated image-wise and very similar to the well-known SNR but correlating better with the human quality perception [152]. The PSNR calculation yields a quality indicator for each image of the video sequence in relation to the original image. Thus, this metric is only meaningful if the quality of the original image sequence is high in terms of human perception which is not necessarily the case.

The compressed video will be already distorted if the video sequence is passed through a state-of-the-art video encoder to reduce the bit-rate since modern video-codecs – like MPEG-4 or H.264 – are usually lossy. A loss of packets will lead to decoding errors at the decoder/player while delay can cause buffer underruns. Both will ultimately cause the loss of images at the player. Since modern video-codecs make extensive use of the temporal redundancy (encoding only the differences) in most videos, the loss of single images also leads to the distortion of all following images that are differentially encoded based on the lost image. Lost frames usually will cause the video player to "freeze" or to show the last successfully received and decoded image. It is important for an image-by-image metric to reproduce this behavior in case of transmission losses or delay in order to avoid alignment issues between the source and the received video. For a better illustration of the meaning of quality measures for non-experts, the ITU-R developed a quality indication scale which is tied to the quality impression of human observers [153]. This scale is shown in Table 4.1.

ITU-R recommendation BT.500 [153] further describes a methodology to gain these quality indicators by subjective assessment series (by a group of humans). Such a scale is often called Mean Opinion Score (MOS) and is used in several quality assessment systems. A mapping of PSNR values to MOS values is introduced in [154] which can be used to roughly estimate the human quality perception for videos with relatively low motion. This mapping from PSNR to MOS is shown in Table 4.2 and used in this paper. A MOS value is assigned to each image according to Table 4.2 which is based on the PSNR values that are calculated for every single image of a received video sequence. These values are averaged over all images of a sequence to produce a single quality indicator for a video transmission as proposed by the methodology described in [153].

4.1.6 Extended Framework

As a consequence of rapid improvements in technology and miniaturization, sensor networks become an attractive solution for a large number of new applications [109]. At the moment, we can recognize a trend towards sensor nodes which are equipped with high data rate wireless interfaces. The interfaces enable the nodes to transmit multimedia content in a multi-hop wireless network. However, sensor nodes with Bluetooth or other high data rate wireless interfaces are hardly available at the moment. Thus, simulation is the most common approach to estimate the performance of WMSNs.

Table 4.1: *ITU-R quality and impairment*

Scale	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Table 4.2: *PSNR[dB] to MOS conversion*

PSNR	MOS
> 37	5 (Excellent)
31 - 37	4 (Good)
25 - 31	3 (Fair)
20 - 25	2 (Poor)
< 20	1 (Bad)

Simulation tools, like ns-2 [58] or OPNET Modeler [59], only support standard network performance metrics like packet delivery ratio, delay and jitter which do not necessarily represent the perceived quality by the user. If we take a closer look e.g. at video encoding and decoding, we recognize that the video quality over a lossy link strongly depends on the way packets are lost. It is important to know whether consecutive packets are lost as a consequence of topology changes, single packets which might be caused by interference, or a low signal to noise ratio. For this reason, we want to focus on quality metrics which reflect the perceived quality of a video transmission in order to evaluate the performance of our network.

Therefore, we decided to extend the framework by modifying the traffic module and the data sink module in the OPNET Modeler simulation such that it can read and write files according to the format used by the EvalVid [155] video evaluation framework. In the following, a brief description of video quality evaluation is given. Furthermore, the Evalvid software is introduced which is integrated in the extended simulation framework to evaluate wireless networks in terms of QoE.

EvalVid

EvalVid [155] is a video evaluation framework which comes with a large library to analyze the quality of video transmissions. The framework has a lot of advantages compared to other video evaluation tools, e.g. [156] and [157]. Both tools are commercial and mainly focus on the video evaluation. Moreover, they do not offer an interface to connect them directly to other simulation tools like ns-2 or OPNET. Another well known tool is represented by the Video Quality Metric (VQM) Software [158]. This tool offers a large number of different metrics in order to evaluate the quality of videos. However, just as its commercial counterparts it does neither offer the functionality to create trace files nor an interface to interact with network simulation tools for performance studies. The video quality evaluation tool Aquavit [159] offers almost the same functionality as EvalVid but is not further developed. Therefore, it does not support state-of-the-art codecs like the H.264 [160]. For this reason, we decided to connect the EvalVid framework to the OPNET simulation. The framework is used to compare the quality of the source (encoded and already slightly distorted video) with the received video quality in order to evaluate the performance of the simulated network. A detailed flow diagram of the EvalVid framework is shown in Fig. 4.8.

A trace of the original video file is recorded, containing size and type of each video packet which is transmitted over the Real-time Transport Protocol (RTP). The video can be either in raw format or already encoded. Note that most state-of-the-art encoders compress the information such that a slight quality loss is taken into account. Therefore, the video source file is encoded and passed through a decoder if the source video is in raw format. This step is necessary to calculate the perceived quality of the video with encoder losses but without network loss. Thus, it is possible to calculate the degradation of the video quality which is caused by the loss of information during the video encoding.

The trace file of the encoded source video is then used to generate packets in the OPNET simulation. In addition, the OPNET simulation generates a receiver dump file which represents the input for the EvalVid video evaluation tool. The comparison of the sender and the receiver dump file enables EvalVid to calculate the delay, jitter and the loss of packets and frames.

However, the trace file can also be used to replay traffic over a real or virtual network to measure the perceived video quality. In the latter case, a setup as shown in Fig. 4.7 is used where the source is represented by computer A, while the destination of the video is represented by computer B. The advantage of this setup is that the simulation gives a direct insight in the perceived video quality of the simulated network. Nonetheless, the simulation has to run in real-time to ensure that packets are exactly delayed as determined by the simulation.

The data sink modules of the OPNET framework write a dump file of the received packets which includes the creation time of the packet, the time of reception, the sequence number, the frame type and the size of the packet. The dump file is generated at the end of each simulation run and is used as input for the evaluation tool to calculate packet/frame loss and delay statistics as well as

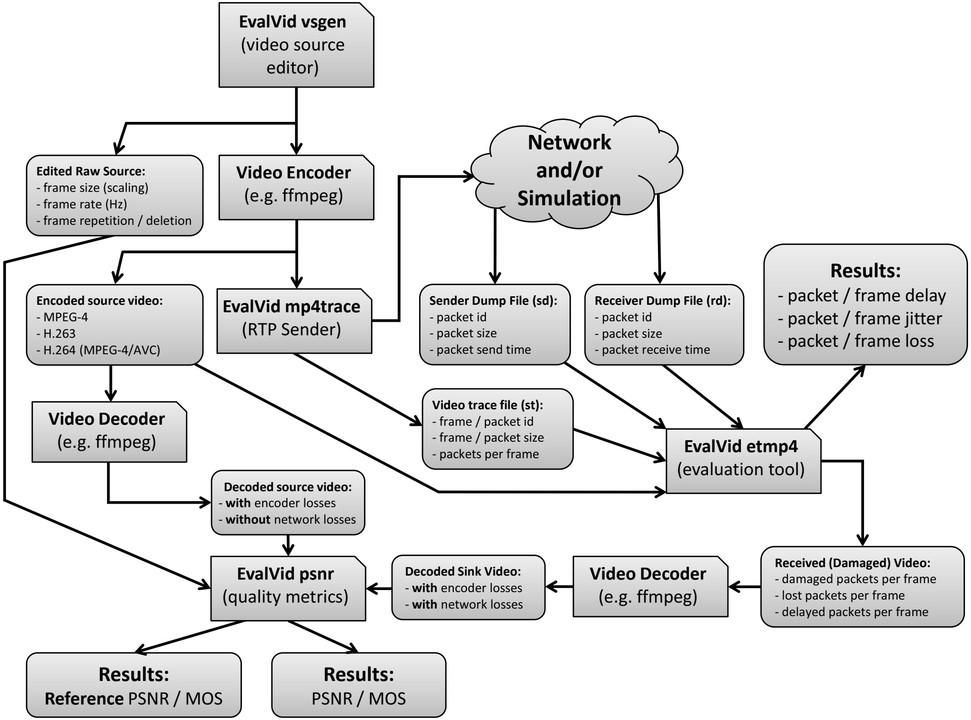


Figure 4.8: *Evalvid - Flow Diagram*

reconstructing the received video files. The received videos are decoded using FFmpeg [1] in order to be able to calculate the statistics of the video quality in terms of PSNR and MOS. The quality of the received video is affected by the encoder losses and the network losses which reflects the perceived video quality.

4.2 Mobility Patterns

Mobility issues become more important nowadays since the number of mobile communication devices is quickly increasing. The performance of a wireless network and its topology is depending on the mobility pattern of the nodes. Thus, it is essential to understand the characteristics of the synthetic mobility models [137, 161, 162] in order to create a realistic simulation environment.

It is often assumed that the nodes in a network should move fully randomly within the simulation area to provide optimal performance evaluation conditions. But what is random movement? Most researches would say that a mobility model generates random movement if the node density and the node speed are uniform distributed. However, studies of real-world traces of mobile phone users have shown that human mobility patterns do not follow these assumptions. Nowadays, human mobility patterns become more important for WSNs since the number of applications, like health monitoring or fire rescue, increases where users are carrying one or more sensor nodes with them. It was shown in [163] that human trajectories have a high degree of temporal and spatial regularity. Moreover, each individual has a time-independent characteristic movement pattern and a significant probability to travel between a few highly frequented locations [164] which results in correlated movement. Thus, it is questionable if such a movement pattern should be declared as random mobility.

This type of movement can be divided into different movement states as indicated in Fig. 4.9. The figure shows snapshots of the Marienplatz in Munich which is one of the most crowded areas in the city center. By observing the movement of the pedestrians over a longer period of time, we found out that their movement

---

[1] FFmpeg - Multimedia Framework, `http://ffmpeg.org`

(a) Hotspot Nodes (b) Hopping Nodes

(c) Transit Nodes (d) Fix Nodes

Figure 4.9: *Snapshot of Marienplatz in Munich*

can be divided into four different groups. In the following, the pedestrians are regarded as nodes since the majority of the pedestrians are assumed to carry mobile devices which are able to build a network.

The first group is represented by nodes which are moving in the area around a hotspot. These nodes only occasionally leave the area around a hotspot and move to another hotspot. Furthermore, their node speed is very slow while being close to a hotspot. The second group consists of hopping nodes which are moving directly between different hotspots without remaining a long period of time at a single hotspot. The third group of nodes is represented by transit nodes. Transit nodes bypass crowded areas in order to move quickly from one point to another. The last group are fix or very slow moving nodes, e.g. waiting pedestrians or people which are visiting a pavement café. These observations point out that the movement of humans has a high spatial regularity which is in coincidence with the observations made in [163].

In addition, the movement of nodes is usually bound or at least strongly influenced by the road network. The infrastructure and obstacles affect the human movement such that it cannot be covered by simple mobility models like random waypoint [165]. The usage of synthetic mobility models often results in an overestimation of the performance of wireless networks [166] which is caused by the speed decay problem [167, 168]. The problem is not further explained here since it is described in detail in Subsection 4.2.2.

4.2.1 Characteristics of Mobility Patterns

Mobility models have a great impact on the performance of routing protocols in wireless networks. Synthetic mobility models [161], like the random waypoint [70], the random walk [147], and the random direction [148] model, do not generate random movement since their behavior is strongly affected by their configuration and the shape of the used scenario.

The most substantial argument against the usage of the standard mobility models is that these models do not reflect typical movement of humans, which brings us back to the question of the previous subsection. The best counter question is, what is typical human movement and how could it be defined? Obviously, typical movement is not reflected by a particular scenario [169], e.g. movement of students on a campus, cars on the road, customers in a shopping mall or pedestrians in the city center. The nodes in these scenarios show different movement patterns which have a different impact on the network connectivity. Therefore, it is necessary to define independent mobility characteristics in order to allow a meaningful comparison between synthetic mobility models and real-world traces. Different mobility characteristics can be used to compare and to evaluate the movement of mobility models. In the following, we focus on six mobility characteristics to give a more detailed insight of the generated movement of the most popular synthetic mobility models.

Link Duration

The link duration is the period of time during which two nodes are within transmission range of each other. Link duration is an interesting connectivity metric since it is directly affected by the mobility pattern of the nodes. Different approaches were introduced in the last few years which try to optimize the chosen links by concerning the average link duration [139], the link change rate [138] or the probability density function of the link duration [140].

The estimation of the probability density function of the link duration is a very promising approach. However, it is also the most complicated one. Moreover,it cannot be applied to all sensor networks since it requires the frequent transmission of data packets or routing messages to quickly detect topology changes in the network. In addition, the nodes have to keep track of the link duration of links in the past to estimate the probability density function of the link duration. Note that sensor nodes have very limited hardware resources. Thus, the solution that was presented in [140] will become more attractive to the next generation of sensor nodes which will have a higher computational power and a larger amount of memory. Another analytical approach was presented in [170] where the authors derive the probability density function of the link duration by focusing on the relative movement speed and the transmission range of the nodes.

Spatial Node Distribution

The spatial node distribution in a wireless network is correlated with the node degree. Therefore, the network topology and the network performance are strongly affected by the spatial node distribution. Furthermore, the nodes are usually not randomly distributed which has a great impact on the network performance. The spatial node distribution which results from the used mobility model has to be considered [171, 172] in order to allow a meaningful performance comparison of different routing protocols. The authors of [171] showed that the distribution is influenced by the mobility parameters and the shape of the scenario. For this reason, the size and the shape of a scenario should be chosen with respect to the mobility model.

Transient Phase

Most synthetic mobility models generate a characteristic spatial node distribution which depends on the mobility parameters and the shape of the scenario as described in the previous paragraph. Thus, the spatial node distribution at the beginning of a simulation should be chosen according to the characteristic spatial node distribution of the used mobility model to minimize the duration of the transient phase. A more detailed description of different characteristic spatial node distributions is presented in Subsection 4.2.2.

However, another aspect is often neglected when using synthetic mobility models. The movement which is generated by synthetic mobility models, e.g. random waypoint and random direction, can be divided into a movement phase and a pause phase. The duration of the movement phase mainly depends on the selected node speed since the travel distance is either limited by the algorithm of the mobility models or the simulation plane. Thus, nodes which have chosen a low speed remain in the movement phase for a long duration due to the fact that they require a long period of time to reach the next decision point or the border of the scenario. This behavior leads to the average node speed decay problem [167, 168]. For this reason, the performance of a network increases towards the end of a simulation if the results are collected before the mobility model is in a steady-state since the average node speed decreases during the transient phase.

Node Speed Distribution

Standard mobility models typically choose the node speed randomly distributed between a minimum and a maximum value. Thus, it is often assumed that the distribution of the speed of the nodes in the simulation reflects the distribution which was used to select the speed of each individual node. However, this is not necessarily the case since the node speed distribution is mainly affected by the algorithm which is used by the mobility model to select the next destination and the travel duration [9, 168]. The impact of the used algorithm on the node speed distribution is discussed in detail in Subsection 4.2.2 which also gives an overview of the most popular synthetic mobility models.

The node speed distribution is a good indicator for the end of the transient phase [168]. Some mobility models, e.g. random waypoint, suffer from the node speed decay problem. The number of slow moving nodes increases while the number of fast moving nodes decreases over time until the mobility model is in a steady-state. This behavior is indicated by the change of the node speed distribution during the transient phase. The duration of the transient phase mainly depends on the configuration of the mobility model and is discussed in more detail in the following subsection.

Correlated Movement

Correlated movement is a typical characteristic of human mobility [163] which is neglected by the majority of developers of routing protocols since it is hard to detect without position information. The term correlated movement is used in this work if one or more nodes move in a similar direction with similar speed such that they are able to communicate directly with each other over a longer period of time. Therefore, temporary correlation of the movement is generated by all mobility models. Nevertheless, the degree of the correlation depends on the mobility model, its configuration and the shape of the simulation plane. Routing protocols may take advantage from correlated movement by selecting neighbor nodes with similar movement pattern as next hop [9].

Two factors are mainly responsible for correlated movement in real networks. The first factor is represented by social relationships. These social ties [164] can be regarded as a measure of the likelihood of geographic co-location. Moreover, the authors of [164] introduced a synthetic mobility model which considers the social interaction to generate more realistic movement. They showed that the correlated movement also has an impact on the link duration, and thus on the network performance. Movement restrictions are the second factor which further increases the correlation of human movement. In real networks the movement is limited by the road infrastructure and obstacles which force the movement into certain directions and areas [165]. Thus, the possibility that two nodes remain within the communication range of each other increases.

Group Mobility

Group mobility models [138, 173] became very popular several years ago. The models are inspired by mobile ad hoc networks where the collaboration of members of the same group is common, e.g. police patrol, avalanche rescue, military battlefield communications, medical assistance or fire fighter scenarios. The majority of synthetic group mobility models classify the nodes into two categories [161]. The first one is represented by the group leaders which move according to a standard mobility model, e.g. random waypoint, random walk or random direction. The second category are fellow nodes which follow the movement of a group leader. Fellow nodes are only allowed to move within a certain range around their group leader. Thus, the movement of nodes of the same group is highly correlated. The generated movement is also very challenging for routing protocols since they have to distinguish between nodes of the same group and other nodes in order to establish stable routes [11].

4.2.2 Survey on Mobility Patterns

This subsection gives a brief survey of a selection of the most popular synthetic mobility models [137, 161] which are mainly used to evaluate mobile networks. Note that the generated movement of the nodes does not necessarily reflect human mobility patterns [174]. However, the models are well investigated and can be configured such that they generate different degrees of mobility in order to evaluate the performance of routing protocols under different conditions. It has to be kept in mind that there is no single model which covers all aspects of realistic human movement due to the fact that the movement depends on the infrastructure and the environment [169]. Nevertheless, standard mobility patterns, like random waypoint and random direction, can be easily modified [175, 176] to generate more human like movement.

In the following paragraphs the random waypoint, the random walk and the random direction mobility model will be introduced in detail. Moreover, a closer look is taken on the characteristics of the generated movement. In addition, the configuration parameters are discussed which have to be chosen with respect to the shape and the size of the scenario to minimize unwanted characteristics.

Random Waypoint

Random waypoint mobility model was developed by Johnson and Maltz [70] in 1996. They developed the model in order evaluate the performance of the DSR protocol in a mobile environment. The model provided a basis for a large number of other synthetic mobility models and is still the most popular model due to its simplicity. The model has certain characteristics which have to be considered if it is used to evaluate the performance of a wireless network [167, 172]. The algorithm of the model can be divided into five steps as shown in Algorithm 1.

```
1:  Select a random destination within the scenario
2:  Select a random speed    $speed \in [speed_{Min}; speed_{Max}]$
3:  Move until the destination is reached
4:  Wait a random period of time    $pause \in [pause_{Min}; pause_{Max}]$
5:  Goto step 1
```

Algorithm 1: Random Waypoint

At first glance, the algorithm seems to generate random movement since the destination and the speed are randomly chosen. Thus, one may assume that the mobility model will generate a uniform spatial distribution of the nodes. However, neither the node speed nor the spatial node density follow a uniform distribution. The non-uniform distribution of both characteristics results from step 3 of the algorithm. In the case that a node selects a low speed it requires a longer period of time until it reaches the next destination compared to faster moving nodes. Therefore, the fraction of slow moving nodes increases over time while the fraction of fast moving nodes decreases. Furthermore, there is a high probability that a node has to cross the center of a scenario to reach its next destination. As a consequence, the node density in the center area of a scenario is higher than the density along the border and in the corners [148].

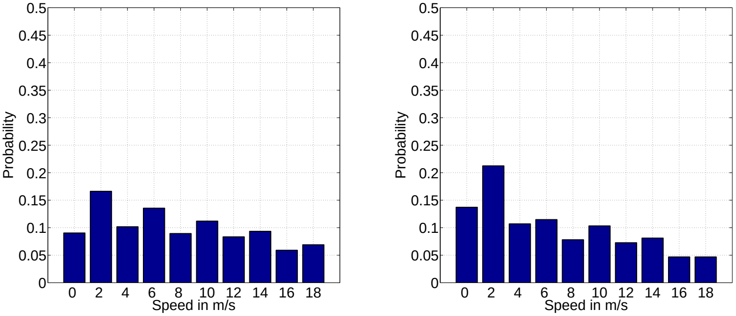
Due to the fact that some characteristics are hard to understand without an

Table 4.3: *Random Waypoint - Configuration*

speed_{Min}	1 m/s
speed_{Max}	20 m/s
pause_{Min}	0 s
pause_{Max}	0 s

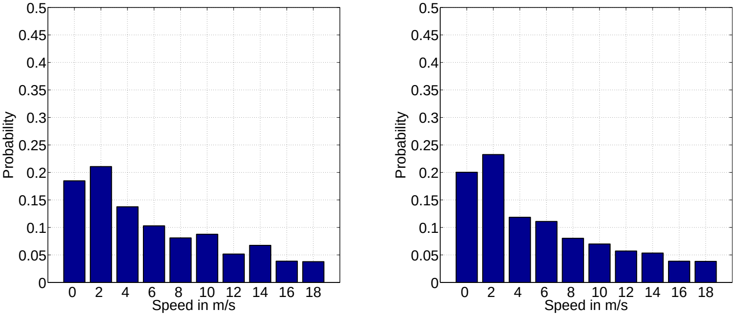
example, we evaluate the random waypoint model by using the framework presented in Subsection 4.1.2. The size of the scenario is 1000 by 1000 meters. 100 nodes are placed evenly distributed in the scenario. In the following, a closer look is taken on the generated movement after 100 s, 200 s, 400 s and 800 s. The random waypoint mobility model is configured according to the settings in Table 4.3 The pause parameters are set to 0 s to create continuous movement. The minimum speed of 1 m/s was chosen to prevent nodes from moving very slowly for a long period of time. Consider a node which selects a speed close to zero. This node would remain almost immobile for the whole simulation. For this reason, the minimum speed should be set to a value higher than zero. However, in the worst case a node which is in a corner of the scenario may choose the opposite corner. Thus, the node will require more than 1400 s in our scenario to reach the next destination if it selects a speed close to the minimum speed of 1 m/s.

This issue becomes clear by taking a look at the results of Fig. 4.10 which shows the current node speed histogram for simulation durations between 100 s and 800 s. The results confirm the previous statement that the percentage of slow moving nodes increases with the advancing of the simulation time. Moreover, the percentage of nodes with a speed of less than 2 m/s is lower than the percentage of nodes with a speed between 2 m/s and 4 m/s which is in contrast to the previous statement. This exception results from the configuration of the random waypoint model. Note that the minimum speed is 1 m/s. Thus, the probability that a node choses a speed between 2 m/s and 4 m/s is twice as high than selecting a speed of less than 2 m/s. For this reason, a longer simulation duration is needed until the



(a) 100 Seconds

(b) 200 Seconds



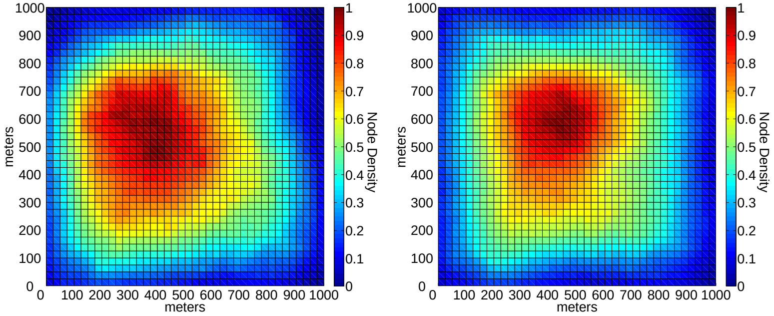
(c) 400 Seconds

(d) 800 Seconds

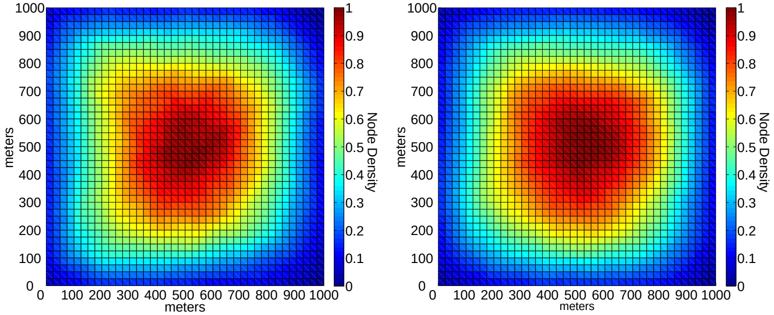
Figure 4.10: *Random Waypoint - Node Speed Histogram*

fraction of nodes with a speed of less than 2 m/s becomes higher than the fraction of nodes with a speed between 2 m/s and 4 m/s.

The difference between the figures indicates that the random waypoint model has a long transient phase. In addition, the results point out that the node speed distribution becomes more and more stable with the advancing of the simulation time. This behavior is reflected by the decreasing average node speed also known as the node decay problem [167, 168]. The node speed distribution becomes almost stable after 800 s which coincides with the observation made in [161] where the authors recommend to discard the first 1000 s of the simulation.



(a) 100 Seconds (b) 200 Seconds



(c) 400 Seconds (d) 800 Seconds

Figure 4.11: *Random Waypoint - Node Density*

The random waypoint model has a very characteristic spatial node distribution as shown in Fig. 4.11. The figures show the normalized node density. The density is calculated such that the lowest density is represented by a value of 0 while the highest density is represented by a value of 1. The results indicate that there is the highest node density in the center of the scenario. The density is decreasing towards the border of the scenario which coincides with the results presented in [176]. This effect becomes more dominant with the advancing of the simulation time.

The characteristic spatial node distribution also affects the node degree and the link duration which both have a great influence on the network performance. In addition, the non-uniform spatial node distribution results in an oscillation of the node degree (density waves) [148] from the perspective of a single node since its node degree changes with its location. The node degree reaches the highest values during the time when the node is close to the center of the scenario while the lowest node degree is recognized in the areas close to the border of the scenario [177]. The density waves are very challenging for routing protocols since they have to deal with frequent link breaks during this time period [148].

The evaluation of the random waypoint model has shown that the generated movement results in a non-uniform spatial node distribution with the highest density in the center of the simulation plane. The non-uniform spatial node distribution represents a challenging environment for routing protocols due to the fact that the node degree changes depending on the position of the node. Therefore, the random way point model is a practical solution for the performance evaluation of protocols in mobile networks since it is simple to implement and generates a challenging environment. Nevertheless, the random waypoint model has to be configured with respect to the node decay problem. Furthermore, the long duration of the transient phase has to be considered to obtain meaningful results.

Random Walk

The random walk or drunkard mobility model was first mathematically analyzed and explained by Einstein [147] in 1905. He described the motion of particles suspended in a fluid at rest which follow the Brownian motion. The model has a large number of interesting characteristics, e.g. it was proven that an object which moves in a one dimensional or two dimensional space according to the Brownian motion always returns to its starting point [178]. Thus, no scenario boundaries are required to assure that a mobile node remains in a certain area. Nevertheless, this is not a practical solution since it may take a long time until a node returns to its starting point. In addition, there is a high probability that the node moves far away from its starting point until it returns which can result in a low node density or even to a partitioning of the wireless network. Moreover, the model generates memoryless movement [179] since it does not retain any knowledge regarding its past locations and speed values. The model became very popular in computer science due to its simplicity and is used and modified in many works [180–182]. The generated movement also represents a very efficient search and data collecting pattern [182, 183]. Therefore, it is often applied in WSNs where a mobile sink gathers data from fix nodes. Furthermore, the model is considered in routing protocols to forward data due to its load-balancing characteristics [184].

The random walk model can be configured such that the nodes move according to the Brownian motion. The generated movement almost follows the Brownian motion if the distance between two consecutive movement steps is close to zero and the nodes move continuously. However, in this case the calculation of the model requires a lot of computational power which becomes clear by taking a closer look on Algorithm 2. At the beginning of the movement phase, the nodes

```
1:  Select a random speed      $speed \in [speed_{Min}; speed_{Max}]$
2:  Select a random direction  $direction \in [0; 2\pi]$
3:  Move into that direction
        a. for a predefined period of time
        b. for a certain distance
        c. if the border of the scenario is reached,
           select a new direction (Bouncing rule)
4:  Wait a random period of time   $pause \in [0; pause_{Max}]$
5:  Goto step 1
```

Algorithm 2: Random Walk Mobility Model select a random speed and a random direction. Then the nodes move into that direction either for a predefined period of time or for a certain distance. The duration of the movement period and the travel distance affect the average node speed. If a node reaches the border of the scenario, it selects a new direction in order to stay within the scenario boundaries. The movement phase ends after the movement period or if the node has traveled the predefined distance. After the movement phase is completed, the model waits a random time interval until it starts the next movement phase and thus jumps back to step 1.

The random walk mobility model generates different movement patterns depending on the duration of the movement period. The decision whether the duration of the movement phase is based on a predefined time interval or a predefined travel distance mainly affects the node speed distribution and the duration of the transient phase. Therefore, the characteristics of the time-based and the distance-based random walk model are evaluated and compared in this subsection. Again, a scenario size of 1000 by 1000 meters is used. 100 nodes are placed evenly distributed in the simulation plane. The mobility models use the configurations shown in Table 4.4 and Table 4.5.

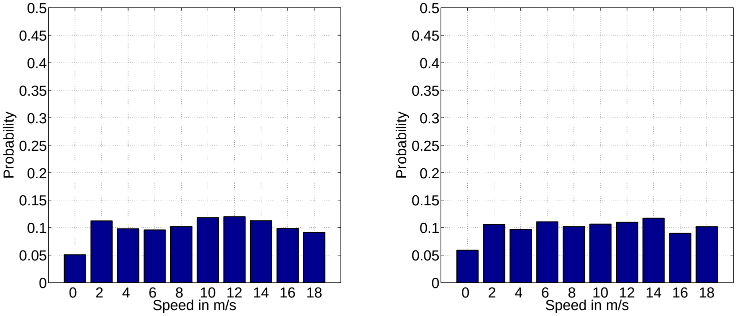
First, we take a look at the node speed distribution of the time-based and the distance-based random walk mobility models which are shown in Fig. 4.12 and Fig. 4.13 respectively. The results of Fig. 4.12 point out that the node speed of the time-based random walk model is uniform distributed. Note that the minimum

Table 4.4: *Random Walk - Configuration A - Time-based*

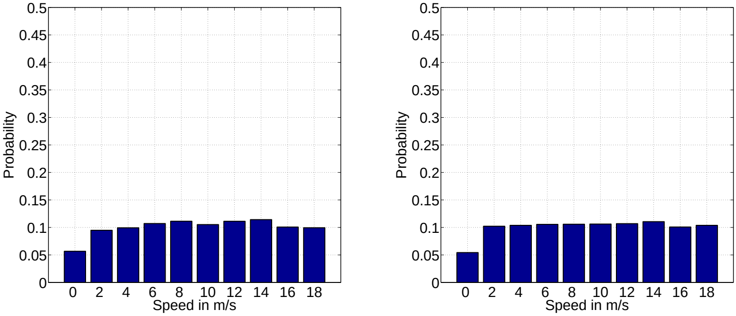
speed$_{Min}$	1 m/s
speed$_{Max}$	20 m/s
pause$_{Min}$	0 s
pause$_{Max}$	0 s
Movement	time-based
Movement Duration	10 s

Table 4.5: *Random Walk - Configuration B - Distance-based*

speed$_{Min}$	1 m/s
speed$_{Max}$	20 m/s
pause$_{Min}$	0 s
pause$_{Max}$	0 s
Movement	distance-based
Travel Distance	200 m



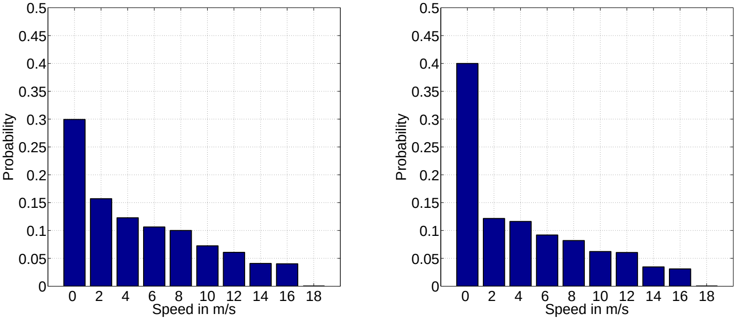
(a) 100 Seconds (b) 200 Seconds



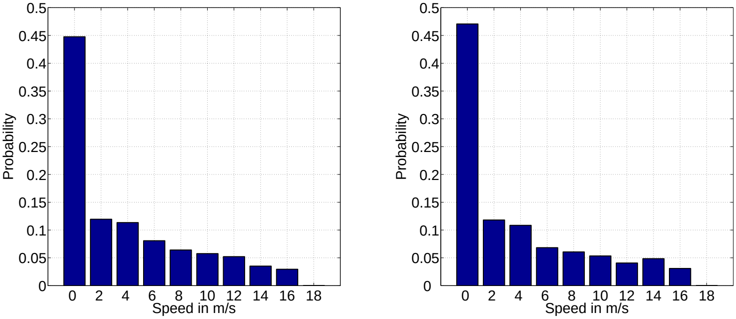
(c) 400 Seconds (d) 800 Seconds

Figure 4.12: *Random Walk(time) - Node Speed Histogram*

node speed is 1 m/s. For this reason, the probability that a node moves with a speed between 2 m/s and 4 m/s is twice as high than the probability that a node moves with a speed of less than 2 m/s. Thus, the simulated node speed distribution reflects the distribution which is used to select the node speed which is the consequence of the time-based movement.



(a) 100 Seconds (b) 200 Seconds



(c) 400 Seconds (d) 800 Seconds

Figure 4.13: *Random Walk(distance) - Node Speed Histogram*

The impact of the duration of the movement phase becomes obvious by evaluating the results shown in Fig 4.13. The histogram reveals similarities with the node speed histogram of the random waypoint mobility model shown in Fig. 4.10 since both models suffer from the fact that the duration of the movement phase is influenced by the distance towards the next destination. The percentage of slow moving nodes in the distance-based random walk scenario is higher compared to the random waypoint scenario.

However, the main difference between both models is that nodes which apply the distance-based random walk model always move the same distance during each movement phase while the travel distance of the random waypoint model depends on the current location of the node and the location of the next destination. Thus, the movement period of the random waypoint model is affected by two random variables since the travel distance of the random waypoint model can be regarded as an additional random variable. As a result, the distribution of the movement duration of the distance-based random walk model and the random waypoint model are different.

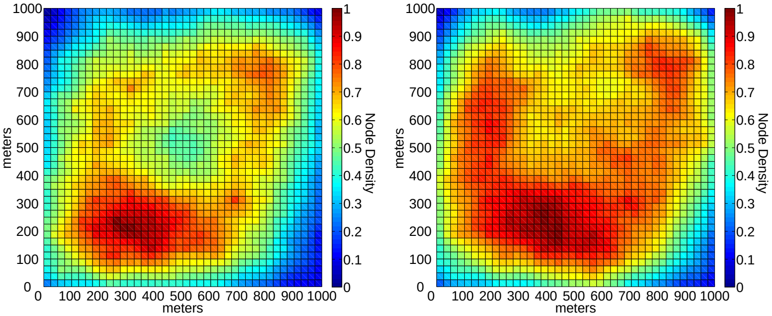
The histogram of the distance-based random walk model shows more variation than the histogram of the random waypoint model. Furthermore, the histograms indicate that the random waypoint model has a shorter transient phase compared to the distance-based random walk model.

The longer transient phase results from the fact that the distance-based random walk model requires more time to generate a stable spatial node distribution as shown in Fig. 4.14 and Fig. 4.15 compared to its time-based counterpart.

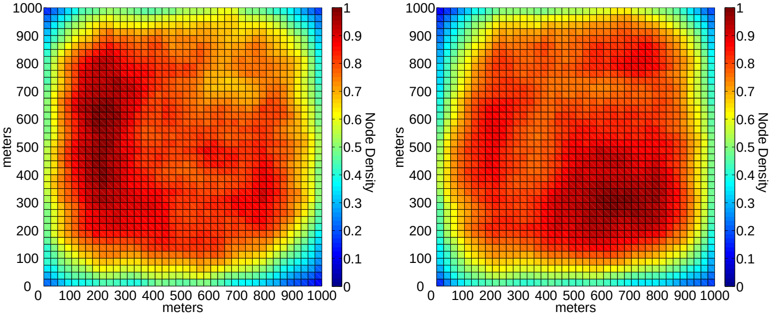
The long transient phase is caused by the high variation of the duration of the movement phase. The duration of the movement phase mainly depends on the current location of a node since the nodes in the center of the scenario may only chose travel-distances up to $\frac{1}{2}\sqrt{2}\sqrt{1000}$ while nodes in a corner may travel twice the distance during a single movement phase. Thus, the node speed distribution becomes stable as soon as the spatial node distribution has stabilized.

Random Direction

The random direction mobility model was introduced in [148] as an alternative to the popular random waypoint model which has certain unwanted characteristics, like the high node density in the center of the scenario and the long transient phase. The authors were looking for a mobility model which generates an homogeneous spatial node distribution in order to determine the optimum node density.



(a) 100 Seconds (b) 200 Seconds

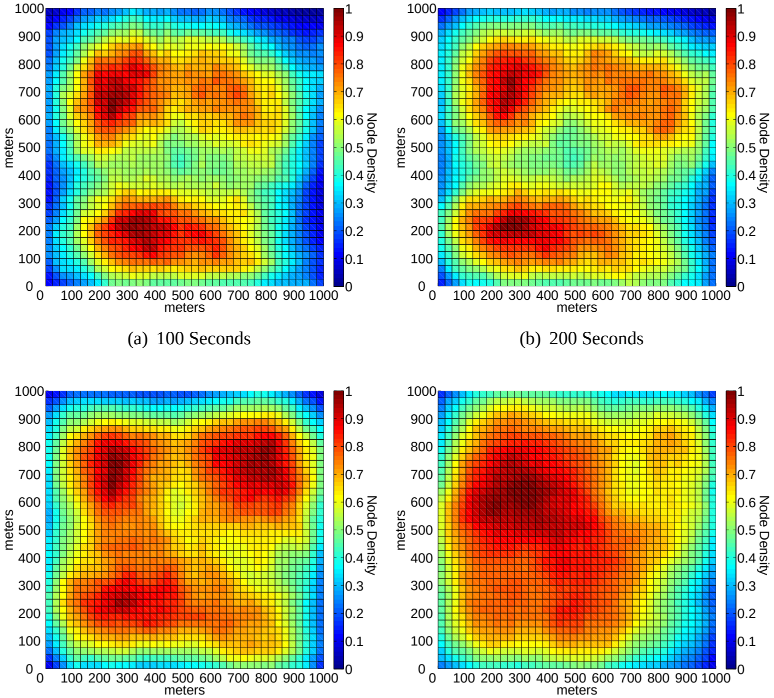


(c) 400 Seconds (d) 800 Seconds

Figure 4.14: *Random Walk (time) - Node Density*

An homogeneous spatial node distribution minimizes the variation of the node degree [148] compared to the random waypoint model where nodes periodically move through areas of high node density. The random direction model has some distinctive features depending on the used bouncing rule in Algorithm 3.

In the first step of the algorithm, the node selects a random direction. If the node is located at the border of the simulation plane the new direction is chosen



(a) 100 Seconds (b) 200 Seconds (c) 400 Seconds (d) 800 Seconds

Figure 4.15: *Random Walk (distance) - Node Density*

such that the node remains within the scenario. The node then choses a random speed and moves into this direction until it reaches the border of the scenario. After reaching the border of the simulation plane, the node waits a random period of time until it selects a new direction and speed.

In [176], the authors introduced two additional bouncing rules. They propose to delete a node which reaches the border of the simulation plane and to replace

```
1: Select a random direction    $direction \in [0; 2\pi]$
   (such that the node does not leave the scenario)
2: Select a random speed    $speed \in [speed_{Min}; speed_{Max}]$
3: Move until the border of the scenario is reached
4: Bouncing rule:
       a. Wait a random period of time
          $pause \in [pause_{Min}; pause_{Max}]$
       b. Delete the node and replace it with a new node
       c. Place the node at the other side of the simulation plane
5: Goto step 1
```

Algorithm 3: Random Direction Mobility Model

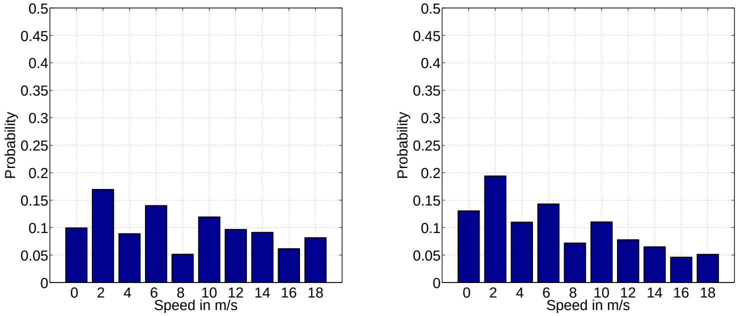
Table 4.6: *Random Direction - Configuration*

speed$_{Min}$	1 m/s
speed$_{Max}$	20 m/s
pause$_{Min}$	0 s
pause$_{Max}$	0 s

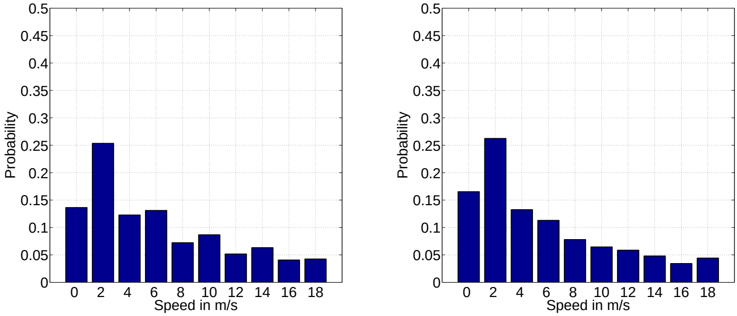
it with a new node. The position of the new node should be chosen with respect to the initial spatial node distribution. The second rule tries to emulate a boundless scenario by placing the node at the opposite side of the simulation plane if it reaches the border of the scenario. However, both bouncing rules do not represent an optimal choice for the simulation of a wireless scenario since former established links will break at once after the corresponding bouncing rule is applied. For this reason, we focus on the original random direction mobility model presented in [148] where nodes wait a random period of time if they reach the border of the simulation plane.

Again, a simulation plane with a size of 1000 by 1000 meters is used in order to allow a meaningful comparison with the other mobility models. Moreover, 100 nodes are placed at the same starting position as in the previous scenario. The configuration of the random direction model is shown in Table 4.6.

The evaluation of the random waypoint and the random walk model have shown that the node speed distribution is mainly influenced by the travel distance. Note that the travel distance of the random direction model only depends on the location of the node and the chosen direction. As a consequence, the duration of the movement phase is influenced by the node speed which is reflected by the node speed distribution shown in Fig. 4.16



(a) 100 Seconds (b) 200 Seconds



(c) 400 Seconds (d) 800 Seconds

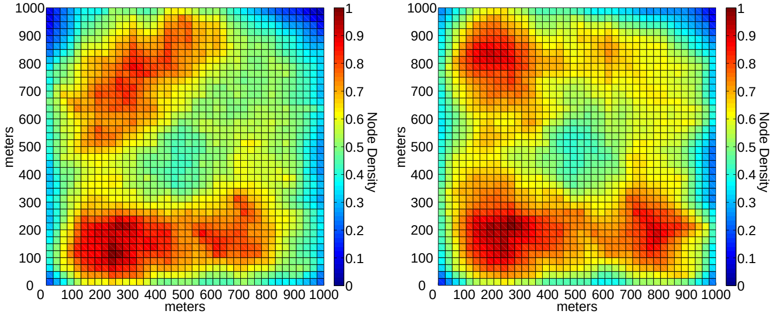
Figure 4.16: *Random Direction - Node Speed Histogram*

The results point out that the percentage of slow moving nodes is increasing due to the fact that slow nodes need more time to reach the next destination. Therefore, the node speed distribution of the random direction model after 800 s is similar to the corresponding node speed distribution of the random waypoint and the distance-based random walk mobility model.

The random direction model generates a characteristic spatial node distribution which has some differences compared to the generated spatial node distribution of the other presented synthetic mobility models. Fig. 4.17 shows the spatial node distribution after 100 s, 200 s, 400 s and 800 s. The results indicate that the lowest node density is in the center of the simulation plane. A higher density can be recognized near the border and in the corners. The higher node density in the corners is the consequence of the fact that the nodes choose a random direction rather than a next destination. Nodes near a corner only have an angle of slightly more than 90 degrees to leave the corner. Thus, there is a high probability that a node which is near a corner needs more than one movement phase to leave it. For this reason, the simulation plane should have the shape of a circle if an even spatial node distribution is desired [185]. The corner effect becomes more dominant if the nodes pause for a certain time period after reaching the border. Therefore, we decided to simulate continuous movement.

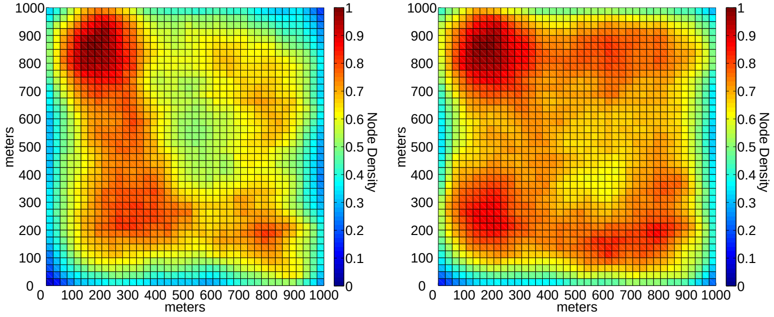
The impact of the corner effect can be mitigated if a delete and replace bouncing rule is applied as proposed in [176]. Thus, a node which reaches the border of the simulation plane is deleted and a new node is inserted at a random position in the scenario. This bouncing rule affects the spatial node distribution such that a slightly higher node density can be recognized in the center of the scenario. The node density becomes lower towards the corner of the simulation plan. However, the delete and replace bouncing rule is not a practical solution for the simulation of wireless networks due to the fact that the frequent change of the topology leads to an underestimation of the network performance.

The results of Fig. 4.17 show that the spatial node distribution of the standard random direction mobility model is still changing significantly after 400 s which indicates that a long transient phase has to be taken into account.



(a) 100 Seconds

(b) 200 Seconds



(c) 400 Seconds

(d) 800 Seconds

Figure 4.17: *Random Direction - Node Density*

4.2.3 Impact of Mobility on Routing

Routing protocols are affected differently by mobility depending on whether they are reactive, proactive or hybrid. In addition, routing protocols which are based link-state routing show a different behavior compared to distance-vector based routing protocols. Proactive and hybrid protocols have a significant advantage over reactive protocols since they are able to quickly detect topology changes due to the frequent probing of the network. Reactive protocols only detect link breaks if they do not receive a response for a certain period of time. Thus, reactive protocols have to use short timeouts in mobile scenarios. Otherwise, the protocols require too much time to detect topology changes which results either in a high packet loss rate or in a high delay if lost packets are retransmitted.

Overhead

The majority of reactive protocols applies route repair mechanisms in order to re-establish broken routes. Therefore, the routing overhead of reactive protocols usually increases with increased mobility since topology changes become more frequent [73, 186]. However, the routing overhead is also affected by the active route timeout which defines the valid time period of an active route. Note that reactive protocols, like AODV, establish a new route for every single packet if the packet inter-arrival time is longer than the active route time out. In this case, the mobility has no significant impact on the routing overhead which is reflected by the results in Subsection 3.5.6.

The overhead of proactive and hybrid protocols is not directly affected by the mobility of the nodes in the network since both types of protocols periodically transmit routing messages to maintain existing routes [73]. Nevertheless, the dissemination of routing information is influenced by the movement especially in the case of OLSR which uses a MPR heuristic to select the forwarding nodes. Thus, the calculated set of MPR forwarding nodes may contain errors which results in a less accurate dissemination of routing information. Moreover, errors in the MPR forwarding set may even lead to a suppression of routing information if links to nodes which are part of the MPR set are broken. The probability of errors in the MPR forwarding set increases with increased mobility. As a result, the routing overhead of OLSR decreases with increased mobility as shown in Subsection 3.5.6. This is in contrast to the hybrid SBR protocol where the routing overhead is not affected by the mobility since the forwarding of routing information is done hop-wise.

Delay

Reactive routing protocols are greatly affected by dynamic topology changes since they have to frequently re-establish or repair broken routes. This is typically done by flooding routing information to detect new routes. Some reactive protocols, like AODV, also make use of local repair mechanisms which limit the flooding to a predefined number of hops. The mechanism which is used by AODV is known as expanding ring search technique [65]. Its basic idea is to increase the flooding range step by step since it is assumed that a flooding of the whole network is not required to find an alternative route around the broken link. This mechanism greatly minimizes the overhead of the routing protocol and thus the network utilization. However, the delay increases in both cases since packets have to be buffered until a new route is established or the existing route is repaired [187]. The delay of proactive and hybrid protocols is only affected if the established routes are based on inaccurate routing information. The time which is needed to detect topology changes only depends on the network probing frequency which is independent from the mobility in the network.

Reliability

The majority of WSNs is tolerant against packet loss since the high node density results in redundancy which makes the network robust against minor loss of information. However, packet loss is only acceptable up to a certain rate due to the fact that the retransmission of packets is usually not an option as a consequence of the low data rate and energy constraints in WSNs.

Reactive protocols are only able to detect link breaks if they do not receive any data via the broken link for a certain period of time. Thus, the packet loss rate of reactive routing protocols mainly depends on the duration of the route timeout and the link duration. The protocol has no influence on the link duration since it is not able to modify the movement of the nodes. For this reason, reactive protocols have to use short active route timeout intervals to quickly detect broken links. In this case, the protocols are able achieve an acceptable reliability at the cost of a higher routing overhead [8].

Proactive and hybrid protocols are also able to increase their reliability in mobile networks by transmitting routing information more frequently [11]. Nevertheless, the transmission frequency has to be chosen with respect to the node density and the available bandwidth. Otherwise, the dissemination of routing information may result in congestion or collisions on the radio channel.

4.2.4 Real World Traces vs. Mobility Patterns

The introduction and discussion of the mobility models in the Subsection 4.2.2 has shown that the movement which is generated by synthetic mobility models does not reflect all aspects of human movement. Extensive studies of real world traces of mobile phone users [163] have shown that human movement is affected by a large number of different factors, e.g. the available infrastructure, the transportation system between points of interests and individual factors(social bounds). Thus, the movement pattern of different individuals differs greatly which makes it very complicated to find similarities that can be used to develop a realistic synthetic model. Their observations have shown that the movement pattern of different individuals remains almost constant over a longer period of time. The recorded traces revealed that the majority of users only travels over short distances while others move regularly over more than hundred kilometers.

It was shown in many research papers that movement of humans and animals can be approximated by Levy Flights [188]. Levy Flights represent a special case of random walks where the travel distance between two movement steps follows a heavy-tailed probability distribution. Thus, on the one hand the probability is high that a node moves only a short distance between two consecutive movement steps. On the other hand, there is a low probability that a node moves into one direction for a long period of time. The heavy-tailed distribution of the travel distance results in a different mobility pattern compared to the time-based and the distance-based random walk models.

However, as discussed in the beginning of this section, the movement of individuals is influenced by many factors. Therefore, it is questionable whether a single mobility model can cover all aspects of human movement. Furthermore, the question has to be answered whether a realistic mobility model is required to allow a meaningful simulative performance evaluation of different routing protocols.

From the perspective of a routing protocol, it does not matter if the movement itself is realistic or not as long as the environment which is generated by the mobility model, e.g. link duration and node density, reflects the characteristics of the target scenario. For this reason, it is recommended to evaluate the performance of routing protocols in a large set of different mobile scenarios in order to get meaningful benchmark which reflects the capabilities of the protocols. A performance comparison in scenarios with different mobility models is necessary since the performance of routing protocols varies greatly depending on the used mobility model [189].

4.3 Performance Studies

In this section, we focus on different simulation aspects, e.g. the scenario, the wireless communication, the protocol parameters and the evaluation, which have to be considered to get meaningful results from a simulation. Moreover, we demonstrate how a small testbed can be used to validate a simulation.

4.3.1 Scenario

Scenarios for performance evaluation of routing protocols have to be selected and configured carefully since a large number of issues affect the performance of the protocols. The characteristics of the underlying MAC protocol have a great impact on the routing protocol due to the fact that the majority of routing protocols uses the broadcast mechanism of the MAC protocol to disseminate routing information. Thus, the dissemination of routing information may become a serious problem if the messages have to be transmitted via a shared medium.

The data rate of low-power transceivers for wireless sensor nodes is usually below 250 kb/s which has to be considered when setting up a simulation scenario. The routing protocol has to be configured with respect to the high node density and low data rate. Moreover, the mobility pattern of the nodes has to be taken into account since the generated movement results in different network characteristics, e.g. the spatial node distribution and the link duration. In addition, the shape of the simulation plane has an impact on the mobility pattern and the topology which is often neglected. Furthermore, the generated movement may lead to a temporary partitioning of the network. Thus, the movement should be evaluated in advance to verify whether the movement which is generated by the synthetic mobility model meets the given requirements.

4.3.2 Wireless Communication

Many simulation tools, like ns-2 or OPNET Modeler, come with simplified propagation models, e.g. free space or disc model, which neglect most of the characteristics that have great impact on the communication in a multi-hop wireless network [190]. Often, these models are even further simplified to allow the simulation of large-scale networks within a justifiable period of time. The authors of [142] summarized the typical assumptions which are often made by simulations, like circular transmission area, equal transmission range, and symmetric links. Their research group showed in [191] that it is important to compare the results from real experiments with the results of the simulation.

Technical characteristics of the simulated hardware, e.g. the CCA delay and the turnaround time of transceivers, are also often neglected. The simulative performance evaluation of the BPS-MAC protocol in Subsection 2.5.6 has shown that the performance of a wireless network strongly depends on the capability of the transceiver in terms of CCA delay and turnaround time. For this reason, the validation of a simulation, which uses a simplified signal propagation model, with a real testbed is indispensable [190].

4.3.3 Protocol Parameters

Different performance metrics can be used to compare the performance of routing protocols. Typical metrics are overhead, packet loss and delay. However, these metrics are influenced by the scenario, the signal propagation, the underlying MAC protocol and the mobility of the nodes. The standard configuration of routing protocols is usually optimized for scenarios with a low node density. Furthermore, it is often assumed that nodes have no or little mobility. Moreover, the majority of links are considered to be symmetric and reliable which is typically not the case in real WSNs [32].

Therefore, the comparison of routing protocols which use the default configuration in non-standard scenarios will not lead to meaningful results. Instead, optimized configurations have to be used in order to allow a fair comparison of the protocols in terms of overhead, packet loss and delay. However, finding optimized configurations is not an easy task since many routing protocols, e.g. AODV and OLSR, have a large number of configuration parameters which have to be chosen with respect to each other. Thus, the optimization of the configuration of the protocols requires a detailed knowledge of each parameter and its impact on the network performance. The protocols should be configured such that they achieve the same performance for a particular metric, e.g. routing overhead. The protocols should then be compared with respect to delay and packet loss.

4.3.4 Mobility

Besides asymmetric and unreliable links, mobility represents the most challenging problem for routing protocols. Mobility leads to frequent topology changes which have to be detected quickly by the routing protocol in order to achieve high reliability since wireless sensor nodes are usually not able buffer a large number of packets. The link duration and the spatial node distribution are mainly responsible for the link change rate. Therefore, the configuration of the routing protocol has to be chosen with respect to these mobility characteristics. In addition, the route establishment routine of the protocols has to be taken into account due to the fact that reactive protocols are not able to detect topology changes as quickly as their hybrid and proactive counterparts.

The results in Subsection 4.2.1 showed that the characteristic spatial node distribution of synthetic mobility models varies depending on their configuration. Furthermore, the shape of the simulation plane and the applied bouncing rule affect the spatial node distribution. Thus, the generated movement should be evaluated in order to verify that the network does not get partitioned during the simulation as a consequence of the used mobility model. Otherwise, the gathered statistics, e.g. the simulated packet loss cannot be compared with scenarios where nodes are connected during the whole simulation.

4.3.5 QoE-based Evaluation

In the beginning of Section 4.3, we outlined the importance of validating a simulation by using measurements from a real testbed. The testbed scenario has to match with the used simulation scenario to validate and calibrate the simulation. In this subsection we measure and simulate the performance of a video application over a high data rate multi-hop wireless network in order to demonstrate a QoE-based performance evaluation of the SBR protocol.

Therefore, we take a look at the perceived video quality instead of focusing solely on the packet delivery ratio since the QoE is the target optimization metric for this application. Furthermore, we discuss the parameters of the SBR protocol and focus on their impact on the video quality during topology changes. Again, the extended framework which was introduced in Subsection 4.1.6 is used to evaluate the perceived user quality.

Implementation

The programming language Java was used to implement the routing protocol to allow its usage on different Operating Systems (OS). Most common OSs, e.g. Linux and Windows, come with tools which allow the modification of their routing table without much effort. Therefore, the Java routing application has to detect which OS is used in order to know which commands are supported by the OS. This enables us to manipulate routes in the table without the need of notifying other applications.

The implementation consists of three major packages. The first one is represented by the network package which is used to receive and transmit data packets via the IEEE 802.11 interface. We use the Jpcap 0.7 [2] library which is based on WinPcap [3] to grab packets from the interface. The second package covers configuration, routing table, and time management functions, e.g. timer and statistic tasks, which are then used by the routing protocol. The behavior of the routing protocol and the used messages build the third package.

Incoming packets are detected and evaluated by a receiver task which sends a callback to the routing task to further evaluate the packet. The routing task then decides what actions have to be performed according to the content of the packet, e.g. modification of the routing table, changing of routing entries, forwarding or dropping of the packet. Additionally, periodic tasks, like the hello message transmission timer or the routing entry decrease timer, send callbacks to the routing application.

A filter class was added to the network package which can be used to limit the topology. Thus, we can restrict the topology of the testbed according to the topology in the simulation. Furthermore, time triggered topology changes can be used to study the behavior of the protocol to deal with link breaks depending on its configuration.

Video Sequences

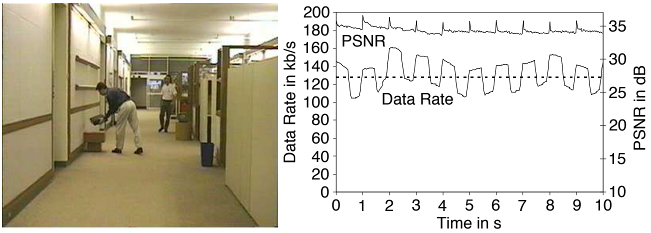
We selected one of the standard video sequences which is used by a variety of video encoding and transmission studies by, e. g. the Video Quality Experts Group. This video sequence is called "Hall Monitor" and consists of 300 frames in CIF resolution (352x288 pixel) with 30 Hz frame rate. It is a relatively low-

---

[2] Jpcap - Java-based Packet Capturing Library,`http://netresearch.ics.uci.edu/kfujii/jpcap/doc/`

[3] WinPcap - Packet Capturing Library,`http://www.winpcap.org/`

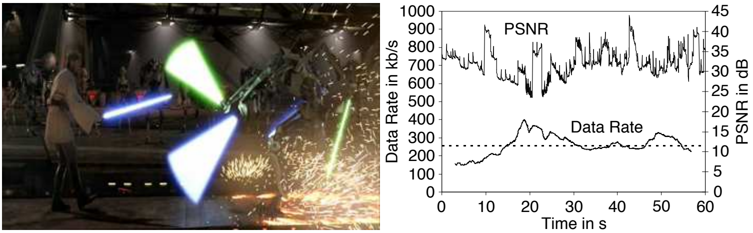
motion sequence such that the PSNR to MOS mapping shown in Table 4.2 can be applied. Due to the fact that it is only 10 s long, we concatenated the sequence six times. Since the video is recorded with a static camera and there is little motion in the scene, the influence of this concatenation on the video encoder performance is low – even at the junctions. The resulting one minute long video was then encoded with the state-of-the-art H.264 video encoder x264 [160] with an average target bit-rate of 128 kb/s. A key-frame was encoded every second in order to have a good balance between coding efficiency and error recovery capabilities. To give a better impression of the video sequence used, Fig. 4.18 displays a sample image together with the bit-rate profile and the PSNR between the encoded and the original video. A second video sequence with more motion was selected to stress the performance evaluation methodology with content appropriate for entertainment applications.



(a) Sample Image (b) Video Profile & Quality

Figure 4.18: *Profile of the Hall Video Clip*

Fig. 4.19 shows a sample image from the one minute long scene from the Movie “Star Wars III”. The resolution is 360x216 pixels and the frame rate is 25 Hz. With current video encoding technology it is not possible to achieve an acceptable PSNR with an average target bit-rate of 128 kb/s. Consequently, the video clip was encoded with a target bit-rate of 256 kb/s. The different content of the selected clips is also reflected in the variations of the size of the encoded frames. While the only variations in the Hall clip are basically the different sizes



(a) Sample Image (b) Video Profile & Quality

Figure 4.19: *Profile of the SW3 Video Clip*

of the I and P frames, the frame size fluctuations in the SW3 clip are much higher.

We performed a set of test runs in a specific scenario in order to calibrate the simulation with the measurements. Five wireless nodes are initially connected to each other and transmit the encoded Hall video sequence using RTP [192] from node 1 to node 5 as shown in Fig. 4.20. The nodes in the simulation and the testbed are configured such that they are forced to build a string topology. Thus, they are only able to receive messages from their direct neighbors. Node 5 represents an exception since it temporarily connects to the other nodes. We have chosen this extraordinary example due to the fact that it is the worst case scenario for the routing protocol. A description of the connectivity during the simulation and the measurement is given in Fig. 4.20. Note that u_1, u_2, and u_3 are random variables which are selected at the beginning of the simulation according to a uniform distribution between -1 s and 1 s. The variables are required to shift the disconnection times in order to avoid the alignment with I frames.

At the end of each interval, the direct connections between node 5 and node 1, 2, 3, and 4, respectively were detached which caused the system to find a new route. This represents a relatively harsh scenario since abrupt disconnections represent the worst case for real-time applications. The most relevant routing protocol parameters are the hello message interval and the routing decrease interval which are both set to 1 s. Due to the fact that the other parameters have no signif-

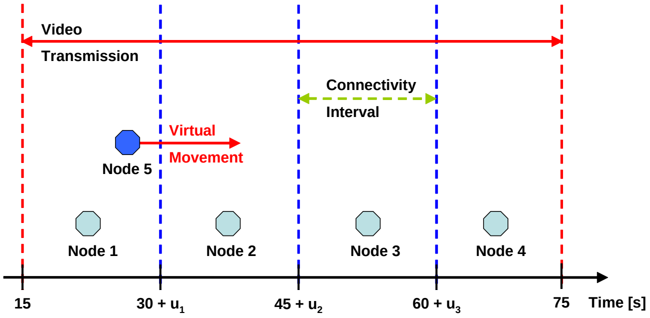
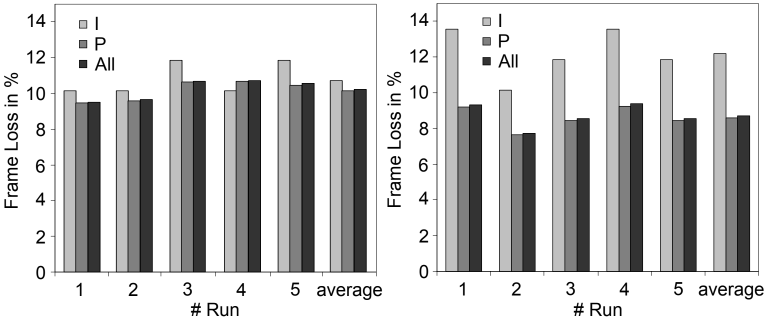


Figure 4.20: *Illustration of the Connectivity during the Video Transmission*

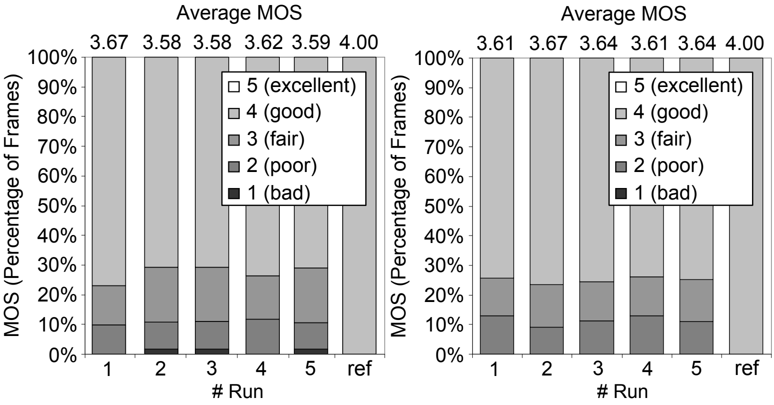
icant impact in our scenario we skip their description since a detailed description of all parameters is given in Section 3.5.

Using EvalVid, traces of the video files were generated, containing the size and type of each video packet transmitted over RTP. Additionally, IP-level packet traces were created using Wireshark at the transmitting and receiving node. These traces were used by EvalVid to calculate packet and frame loss figures as well as reconstructing the received (possibly distorted) video files. The received videos were then decoded using FFmpeg to be able to calculate the PSNR and MOS figures for the video quality evaluation. Fig. 4.21 compares the frame loss of the measurements and the simulations while Fig. 4.22 shows the corresponding MOS values for the received video. The overall frame loss is slightly higher for the measurements which is caused by single packet losses due to interferences, multi-path propagation, and moving obstacles. Moreover, the percentage of I frames lost in the simulation was slightly higher which was quite surprising. A closer look at the trace files revealed that the starting times of the disconnections were varying more during the measurements due to the human reaction time. Against, in the simulations the disconnection interval was quite stable and accidentally always during an I frame transmission. This effect is avoided in the following parameter study by equally distributing the disconnection intervals.



(a) Measurement (b) Simulation

Figure 4.21: *Hall - Frame Loss*



(a) Measurement (b) Simulation

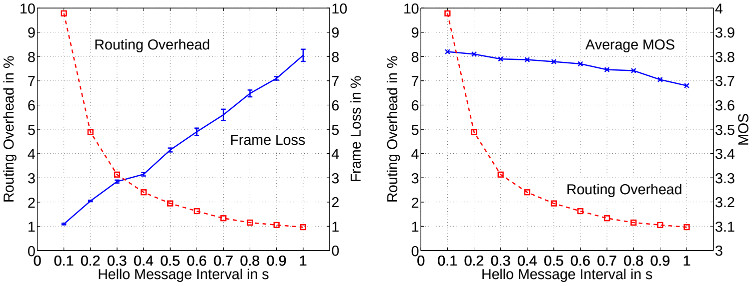
Figure 4.22: *Hall - Comparison of MOS*

The bars in Fig. 4.22 show the percentage of frames with a certain MOS in comparison to the reference videos (rightmost) MOS distribution. The reference video reflects the coding loss and consists of 100% frames with a MOS of 4 (good). In contrast to the raw frame loss these results include the quality degradation caused by frames which could not be correctly decoded due to losses of previous frames. Though the I frame loss in the simulations was higher, the quality of the video was worse in the measured scenario. This is caused by the rare random single packet losses during the measurements which influence all following P frames after the packet loss until the next I frame. The impact that single packet losses have on the MOS and PSNR depends on the used encoder and its configuration, e.g. the I frame rate. Moreover, the type of video which is encoded, e.g. action sequence or landscape stills, has a great impact.

Considering the differences between the measurements and the simulation, the loss and MOS statistics are similar enough such that we can focus on the simulation. In the following we want to demonstrate how to use the simulation for performance evaluation and parameter optimization of the routing protocol to achieve an acceptable video quality even in the case of abrupt disconnections. Thus, we varied the hello message interval of SBR from 1.0 s down to 0.1 s in steps of 0.1 s and transmitted the video 100 times for each setting. Again, the multi-hop scenario with abrupt disconnections every 15 s is simulated. The exact disconnection times were equally distributed in a window of ± 1 s to avoid the exact alignment with an I frame.

Fig. 4.23 shows the resulting average frame loss as well as the average MOS against the overhead of the routing protocol in percent of the video traffic for the Hall clip. Though the frame loss varies between 1 % and 8 % the average MOS only varies between about 3.7 and 3.8. The reason for this is that each lost frame can influence the following frames up to the next I frame. Thus, a loss of consecutive P frames only has a slightly higher impact than the loss of a single P frame. Moreover, the frames are fragmented in up to three IP packets in this scenario. A frame is dropped by the decoder if one or more of its fragments are missing. Therefore, the loss of single fragment has the same impact on the MOS than the loss of the whole frame.

Fig. 4.24 shows the percentage of frames with a certain MOS. Due to the fact that the expressiveness of the average MOS is limited in case of longer videos.



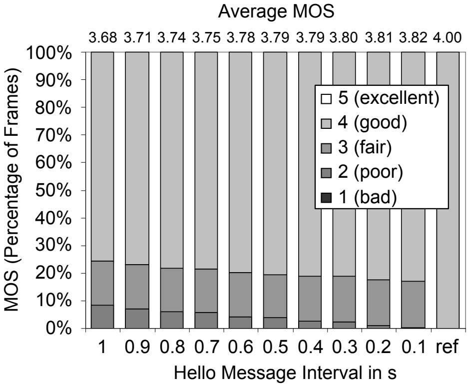
(a) Frame Loss (b) Average MOS

Figure 4.23: *Hall - Frame Loss (a) and Average MOS (b) against the Routing Overhead*

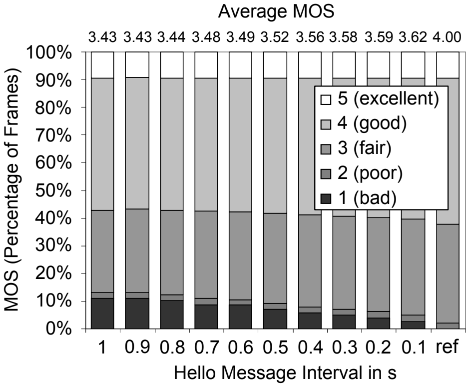
Fig. 4.25 shows the frame loss and MOS statistics for the high-motion scene SW3. Although the frame loss rate is not higher than in the low-motion Hall clip, the average MOS is suffering more from the losses. This results from the higher differences between adjacent frames which lead to a higher sensitivity to lost frames. Another factor is the appearance of frames with a very low MOS (1-2). In fact, the disturbances of the video quality are short in both investigated cases.

Fig. 4.24 shows the number of frames with a certain MOS in comparison to the undistorted reference. In contrast to the average MOS curves in Fig. 4.23(b) and Fig. 4.25(b), it is shown here that the quality impact on the SW3 clip is much smaller than on the Hall clip. This results from the faster recovering in case of losses due to the higher number of intra-coded parts.

The overhead of the routing protocol rises exponentially with the downsizing of the hello message interval. It is acceptable up to around 2-3 % of the application traffic, since this is in the range of the protocol overhead of RTP (1.7 % in this scenario). The relative routing overhead is lower for the SW3 scenario due to the fact that the bit-rate is higher than in the Hall scenario. The MOS distribution bars in Fig. 4.24 show that the difference in quality between the hello message



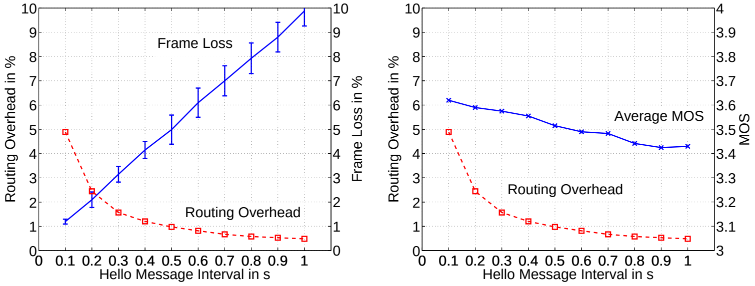
(a) Hall Video Clip



(b) SW3 Video Clip

Figure 4.24: *Percentage of Frames with certain MOS Values depending on the Hello Message Interval*

interval of 0.3 s and 0.4 s is not noticeable by human observer. Considering the smaller overhead, a hello message interval of 0.4 s of the SBR protocol would be optimal in this scenario regarding the perceived video quality.



(a) Frame Loss

(b) Average MOS

Figure 4.25: *SW3 - Frame Loss (a) and Average MOS (b) against the Routing Overhead*

4.4 Summary

In this chapter, we introduced our performance evaluation framework for wireless networks. The framework can be used to evaluate and compare the performance of routing protocols in terms of overhead, packet loss, delay, jitter and other performance metrics like QoE. External simulations can be attached to the OPNET simulation in order to increase the functionality of the framework. Furthermore, the framework allows the exchange of packets between virtual nodes in the software domain and real sensor nodes in the hardware domain. This kind of simulation is known as hardware-in-the-loop simulation and is supported by the framework as long as the simulation can be run in real-time. Additionally, the functionality of the extended framework was introduced and demonstrated by evaluating the performance of a wireless network in terms of packet loss and QoE.

The characterization and evaluation of synthetic mobility models is another central issue of this chapter. The evaluation of a selection of the most popular mobility models has shown that the generated movement has a great impact on the spatial node distribution and the speed distribution. Moreover, the node degree and the link duration were identified as key characteristics of the mobility models. The influence of these characteristics on the performance of routing protocols was outlined. Furthermore, a brief discussion of the differences between real world traces and synthetic mobility pattern was given.

The third topic of this chapter is represented by the performance evaluation of routing protocols in wireless networks. Therefore, the most important simulation parameters and assumptions were discussed, e.g. the configuration of the protocols and the simplification of the signal propagation model. Finally, an example of a performance study of the SBR protocol in a mobile network was given. The focus of the study was laid on the hello message interval which has a direct impact on the perceived video quality.

5 Conclusions

The development of the BPS-MAC protocol and the SBR protocol were motivated by two major trends. The first trend is represented by WMSNs which close the gap between typical WSNs and more powerful ad hoc networks. Sensor nodes in WMSNs are usually equipped with a high data rate interface, like IEEE 802.11 or Bluetooth. Moreover, the nodes have a larger amount of memory available - compared to their WSN counterparts - in order to enable audio and video communication. These networks are often optimized for short-term monitoring applications where nodes start to transmit multimedia data as soon as an event is recognized.

The second trend is driven by the technological advance in energy harvesting and energy storing techniques which prolong the lifetime of sensor networks. In the early stages of WSNs energy harvesting techniques were mainly focused on solar panels. Nowadays, techniques which focus on vibration and temperature differences to harvest energy make sensor networks a practical solution for a large number of long-term monitoring applications for the automotive and the avionic industry. Especially the latter is interested in WSNs which are designed for Structural Health Monitoring applications since wireless solutions provide a practical way to minimize the weight of airplanes.

In this thesis, we introduced and discussed performance issues of MAC and routing protocols in the context of WSNs. The key communication challenges were outlined which result from the special characteristics of WSNs, e.g. limited hardware resources and high node density. Contention resolution represents a performance critical task in dense WSNs since many MAC protocols solely rely on the carrier-sense capabilities of low-power transceivers. Typical transceivers require a long period of time to detect a busy radio channel, especially in the case that the transceiver has been switched off or has to be switched from transmit to receive mode. The duration of the switching phase is often referred to as turnaround time. Transceivers are not able to sense the medium during the switching phase which leads to a large number of collisions in dense networks with correlated event-driven traffic load. Moreover, the CCA delay - which is the period of time that is required by a transceiver to detect the state of the medium - was identified as performance limitation factor for MAC protocols. The impact of the CCA delay and the turnaround time on the network performance depends on the node density and the correlation of the traffic. For this reason, CSMA protocols with random access do not represent an optimal choice for structural health monitoring applications since they have high demands in terms of delay and reliability. Therefore, CSMA-based protocols have to apply acknowledgment mechanisms in order to assure reliable connectivity. However, the transmission of acknowledgments and the retransmission of packets increase the utilization of the medium. As a result, the number of retransmissions may further increase due to a higher collision probability as a consequence of the higher utilization. This behavior may even lead to a collapse of the network depending on the retransmission strategy.

The BPS-MAC protocol was developed to directly address the limitations of low-power transceivers. Its new sequential preamble-based contention resolution reduces the number of competing nodes step by step which makes the protocol attractive to Structural Health Monitoring applications where the correlated event-driven traffic represents a serious issue. It is able to deal with a high number of competing nodes due to the stepwise contention resolution. The contention resolution mechanism of the protocol can be tuned in order to achieve the desired trade-off between delay and reliability. The medium access procedure is independent from the hardware capabilities of the transceiver and thus can be applied on almost any sensor platform. Furthermore, the protocol will take more advantage of next generation low-power transceivers compared to CSMA-based protocols since its performance improves with shorter CCA delays.

Routing protocols are usually developed for a set of scenarios with a certain traffic pattern and network architecture. However, the capability of the protocols to adapt themselves to different network conditions is essential in the context of WSNs. New trends, like WMSNs, come with additional requirements on the routing protocol apart from the typical WSN requirements.

The SBR protocol provides new mechanisms and an adaptive routing metric in order to achieve high performance in WSNs and WMSNs. The protocol is designed such that it is able to deal with many of the challenges in mobile wireless networks without the need of complex algorithms or a large amount of memory. Its adaptive cumulative routing metric minimizes the time which is required by the protocol to detect topology changes. The detection time is short enough to support a high perceived video quality even in the presence of frequent topology changes due to unreliable links or mobility. SBR can operate in a proactive or in a hybrid mode. The hybrid mode generates less routing overhead than the proactive mode since no routing messages are transmitted in the absence of data traffic. Thus, the hybrid mode should be used in typical WSN scenarios where nodes have very limited energy resources. The routing overhead of the protocol scales with the number of data sinks since routing messages are not generated by other nodes. This makes the protocol a good choice for networks with a small number of sinks. In addition, nodes can be set to a passive mode. Passive nodes may use existing routes but do not forward any routing messages which minimizes their energy consumption. This mechanism allows to assign the energy consuming task of routing message forwarding to less energy constraint nodes. Furthermore, these nodes are then used by the protocol to forward data packets which prolongs the lifetime of the network. The proactive mode is dedicated for high data rate wireless networks where routes should be established in advance to minimize the delay in the network. Thus, the nodes periodically transmit routing messages even if no data packets have to be transmitted. The periodic transmission of routing messages in combination with its adaptive cumulative metric allow the protocol to quickly detect topology changes which results in a high perceived quality of multimedia applications.

List of Acronyms

AODV	Ad hoc On-Demand Distance Vector
BATMAN	Better Approach to Mobile Ad hoc Networking
BP-MAC	Backoff Preamble-based MAC Protocol
BPS-MAC	Backoff Preamble-based MAC Protocol with Sequential Contention Resolution
CCA	Clear Channel Assessment
CDMA	Code Division Multiple Access
CSMA	Carrier Sense Multiple Access
CTS	Clear-To-Send
DRVF	Decrease Routing Value Function
DRVI	Decrease Routing Value Interval
DSDV	Destination-Sequenced Distance-Vector
DSR	Dynamic Source Routing
ETX	Expected Transmission Count
FDMA	Frequency Division Multiple Access
GBR	Gradient Based Routing
HLA	High Level Architecture
HMI	Hello Message Interval
IRVF	Increase Routing Value Function
LEACH	Low-Energy Adaptive Clustering Hierarchy
LPL	Low-Power-Listening

MAC	Medium Access Control
MCFA	Minimum Cost Forwarding Algorithm
MMSPEED	Multipath Multi-Speed Protocol
MOS	Mean Opinion Score
MPR	Multi-point Relay
MRV	Maximum Routing Value
OLSR	Open Link State Routing
PSNR	Peak Signal-to-Noise Ratio
QoE	Quality of Experience
QoS	Quality of Service
RREP	Route Reply
RREQ	Route Request
RSE	Relative Speed Estimation
RSSI	Received Signal Strength Intensity
RTP	Real-time Transport Protocol
RTS	Ready-To-Send
S-MAC	Sensor-MAC
SBR	Statistic-Based Routing
SHM	Short Hello Message
SHMI	Short Hello Message Interval
SNR	Signal-to-Noise Ratio
TBEBA	Truncated Binary Exponential Backoff Algorithm
TDMA	Time Division Multiple Access
Wise-MAC	Wireless Sensor MAC
WMSN	Wireless Multimedia Sensor Network
WSN	Wireless Sensor Network

Bibliography

Bibliography of the Author

— Journals and Books —

[1] A. Klein, *Performance Comparison of Different Wimax Configurations - Impact of Scheduling and Contention Resolution on Quality of Service in Wimax Networks*. Saarbrücken, Germany: VDM Verlag Dr. Muellller, June 2008.

[2] A. Klein, J. Klaue, and J. Schalk, “BP-MAC: a high reliable backoff preamble MAC protocol for wireless sensor networks,” *Electronic Journal of Structural Engineering (EJSE): Special Issue on Sensor Network for Building Monitoring: From Theory to Real Application*, vol. -, pp. 35–45, December 2009.

— Conference Papers —

[3] A. Klein, R. Pries, and D. Staehle, “Performance study of the WiMAX FDD mode,” in *OPNETWORK 2006,* (Washington D.C., USA), OPNET, August 2006.

[4] A. Klein and P. Tran-Gia, “A statistic-based approach towards routing in mesh networks,” in *Proc. IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems MASS 2007,* pp. 1–6, October 2007.

[5] A. Klein and P. Tran-Gia, “Energy consumption framework for wireless sensor networks,” in *OPNETWORK 2007,* (Washington D.C., USA), OPNET, August 2007.

[6] A. Klein and P. Tran-Gia, “Poster abstract - a co-simulation framework for performance comparison of wireless sensor networks,” in *EWSN 2007,* (Delft, Netherlands), January 2007.

[7] A. Klein and P. Tran-Gia, “Handover in sensor networks using statistic-based routing,” in *Proc. 6.Fachgespräche Sensor Networks (FGSN)*, (Aachen, Germany), pp. 83–86, July 2007.

[8] A. Klein, “Performance comparison and evaluation of AODV, OLSR, and SBR in mobile ad-hoc networks,” in *Proc. 3rd International Symposium on Wireless Pervasive Computing ISWPC 2008*, pp. 571–575, May 2008.

[9] A. Klein, “How to take advantage from correlated node movement in mobile sensor networks,” in *Proc. 7.Fachgespräche Sensor Networks (FGSN)*, (Berlin, Germany), pp. 77–80, September 2008.

[10] A. Klein, “Statistic-based routing protocol,” Tech. Rep. 452, University of Wurzburg, Wurzburg, Germany, December 2008.

[11] A. Klein and J. Klaue, “Performance evaluation framework for video applications in mobile networks,” in *Proc. Second International Conference on Advances in Mesh Networks MESH 2009*, (Athens, Greece), pp. 43–49, June 2009.

[12] A. Klein and J. Klaue, “Performance study of a video application over multi-hop wireless networks with statistic-based routing,” in *Networking 2009*, vol. 5550/2009, (Aachen, Germany), pp. 955–966, Springer Berlin / Heidelberg, May 2009.

General References

[13] J. Zhao and R. Govindan, “Understanding packet delivery performance in dense wireless sensor networks,” in *SenSys ’03: Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, (Los Angeles, USA), pp. 1–13, ACM, November 2003.

[14] A. Woo, T. Tong, and D. Culler, “Taming the underlying challenges of reliable multihop routing in sensor networks,” in *SenSys ’03: Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, (Los Angeles, USA), pp. 14–27, ACM, November 2003.

[15] D. S. J. De Couto, D. Aguaya, J. Bicket, and R. Morris, “A high-throughput path metric for multi-hop wireless routing,” in *MobiCom ’03:*

Proceedings of the 9th Annual International Conference on Mobile Computing and Networking, (San Diego, USA), pp. 134–146, ACM, September 2003.

[16] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker, "Complex behavior at scale: An experimental study of low-power wireless sensor networks," UCLA CS Technical Report UCLA/CSD-TR 02-0013, University of California (UCLA), Los Angeles, 2002.

[17] J. Polastre, R. Szewczyk, and D. Culler, "Telos: Enabling ultra-low power wireless research," in *Proc. Fourth International Symposium on Information Processing in Sensor Networks IPSN 2005*, pp. 364–369, Apr. 15, 2005.

[18] M. M. Holland, R. G. Aures, and W. B. Heinzelman, "Experimental investigation of radio performance in wireless sensor networks," in *Proc. 2nd IEEE Workshop on Wireless Mesh Networks WiMesh 2006*, pp. 140–150, Sept. 25–28, 2006.

[19] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, pp. 102–114, August 2002.

[20] Chipcon, "Chipcon SmartRF CC2400 datasheet rev. 1.3, http://www.ti.com/," March 2006.

[21] Chipcon, "Chipcon SmartRF CC2520 datasheet rev. 1.3, http://www.ti.com/," 2007.

[22] Moteiv, "Datasheet, Tmote Sky - ultra low power IEEE 802.15.4 compliant wireless sensor module," November 2006.

[23] T. Instruments, "Datasheet, ultra-low power microcontroller MSP430," September 2004.

[24] K. Langedoen, *Medium Access Control in Wireless Networks*, vol. 2, ch. 20, pp. 535–560. Nova Science Publishers, July 2008.

[25] D. Kim, J. J. Garcia-Luna-Aceves, K. Obraczka, J.-C. Cano, and P. Manzoni, "Routing mechanisms for mobile ad hoc networks based on the energy drain rate," *IEEE Transactions on Mobile Computing*, vol. 2, pp. 161–173, April 2003.

[26] L. Krishnamachari, D. Estrin, and S. Wicker, “The impact of data aggregation in wireless sensor networks,” in *Proc. 22nd International Conference on Distributed Computing Systems Workshops*, pp. 575–578, July 2002.

[27] Tang and Körpeoğlu, “Power efficient data gathering and aggregation in wireless sensor networks,” *SIGMOD Rec.*, vol. 32, no. 4, pp. 66–71, 2003.

[28] S. Lindsey and C. S. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in *Proc. IEEE Aerospace*, vol. 3, pp. 1125–1130, 2002.

[29] M.-J. Lin, K. Marzullo, and S. Masini, “Gossip versus deterministic flooding: Low message overhead and high reliability for broadcasting on small networks.”

[30] P. Kyasanur, R. R. Choudhury, and I. Gupta, “Smart gossip: An adaptive gossip-based broadcasting service for sensor networks,” in *Proc. IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, pp. 91–100, Oct. 2006.

[31] K. Jamieson, H. Balakrishnan, and Y. Tay, “Sift: A MAC Protocol for Event-Driven Wireless Sensor Networks,” in *Proc. of the Third European Workshop on Wireless Sensor Networks (EWSN)*, (Zurich, Switzerland), February 2006.

[32] A. Kiryushin, A. Sadkov, and A. Mainwaring, “Real-world performance of clear channel assessment in 802.15.4 wireless sensor networks,” in *Proc. Second International Conference on Sensor Technologies and Applications SENSORCOMM ’08*, pp. 625–630, August 2008.

[33] IEEE, “IEEE 802.3-2008 IEEE standard for information technology-specific requirements - part 3: Carrier sense multiple access with collision detection (CMSA/CD) access method and physical layer specifications,” 2008.

[34] E. H. Callaway, *Wireless Sensor Networks: Architectures and Protocols*. AUERBACH, August 2003.

[35] IEEE, “IEEE 802.11-2007 IEEE standard for information technology-telecommunications and information exchange between systems-local and metropolitan area networks-specific requirements - part 11: Wireless lan medium access control (MAC) and physical layer (PHY) specifications,” 2007.

[36] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS)*, pp. 10–20, January 2000.

[37] IEEE, “IEEE 802.15.4-2006 IEEE standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks– specific requirements part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (LR-WPANs).” 2006, 2006.

[38] W. Ye, J. Heidemann, and D. Estrin, “An energy-efficient MAC protocol for wireless sensor networks,” in *Proc. IEEE Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM 2002*, vol. 3, pp. 1567–1576, June 2002.

[39] N. Abramson, “The ALOHA system,” in *Proc. of AFIPS 1970 Fall Joint Computer Conference*, vol. 37, pp. 281–285, 1970.

[40] N. Abramson, “Development of the ALOHANET,” *IEEE Transactions on Information Theory*, vol. 31, pp. 119–123, March 1985.

[41] IEEE, “IEEE 802.16e-2005 IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands.,” 2005.

[42] “Specification of the Bluetoothsystem, volume 0: Core, v3.0.” Bluetooth SIG, April 2009.

[43] IEEE, “IEEE 802.5, 1998 edition (ISO/IEC 8802-5:1998) IEEE standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements–part 5: Token ring access method and physical layer specification,” 1998.

[44] Bosch, “Robert Bosch GmbH, CAN specification 2.0,” 1991.

[45] A. El-Hoiydi, “Aloha with preamble sampling for sporadic traffic in ad hoc wireless sensor networks,” in *Proc. IEEE International Conference on Communications ICC 2002*, vol. 5, pp. 3418–3423, May 2002.

[46] A. El-Hoiydi, “Spatial TDMA and CSMA with preamble sampling for low power ad hoc wireless sensor networks,” in *Proc. Seventh International Symposium on Computers and Communications ISCC 2002*, pp. 685–692, July 2002.

[47] A. El-Hoiydi and J.-D. Decotignie, “WiseMAC: An ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks,” in *Proc. Ninth International Symposium on Computers and Communications ISCC 2004*, vol. 1, pp. 244–251, June 2004.

[48] C. C. Enz, A. El-Hoiydi, J.-D. Decotignie, and V. Peiris, “WiseNET: An ultralow-power wireless sensor network solution,” *Computer*, vol. 37, pp. 62–70, August 2004.

[49] M. Buettner, G. V. Yee, E. Anderson, and R. Han, “X-MAC: A short preamble MAC protocol for duty-cycled wireless sensor networks,” in *SenSys ’06: Proceedings of the 4th International Conference on Embedded Networked Sensor Systems*, (New York, NY, USA), pp. 307–320, ACM, 2006.

[50] I. Ramachandran and S. Roy, “WLC46-2: On the impact of clear channel assessment on MAC performance,” in *Proc. IEEE Global Telecommunications Conference GLOBECOM ’06*, pp. 1–5, November 2006.

[51] E. E. Johnson, M. Balakrishnan, and Z. Tang, “Impact of turnaround time on wireless MAC protocols,” in *Proc. IEEE Military Communications Conference MILCOM 2003*, vol. 1, pp. 375–381, October 2003.

[52] P. Brenner, “The importance of the tx-rx switching time on the mac protocol,” 1993.

[53] W. Diepstraten, “The importance of short rx-tx turnaround time,” September 1993.

[54] Chipcon, “Chipcon SmartRF CC1100E datasheet, http://www.ti.com/,” April 2009.

[55] G. Chalhoub, N. Hadid, A. Guitton, and M. Misson, “Deference mechanisms significantly increase the MAC delay of slotted CSMA/CA,” in *ICC*, pp. 1–5, June 2009.

[56] A. Koubaa, M. Alves, and E. Tovar, “A comprehensive simulation study of slotted CSMA/CA for IEEE 802.15.4 wireless sensor networks,” in *IEEE International Workshop on Factory Communication Systems (WFCS)*, pp. 1–10, July 2006.

[57] Atmel, “AT86RF231 datasheet rev. b, http://www.atmel.com,” February 2009.

[58] ns 2, “Network simulator, http://nsnam.isi.edu/nsnam/,” May 2010.

[59] OPNET, “Technologies, inc., OPNET modeler, university program, http://www.opnet.com/services/university/,” May 2010.

[60] M. Bertocco, G. Gamba, and A. Sona, “Experimental optimization of CCA thresholds in wireless sensor networks in the presence of interference,” in *Proc. of IEEE EMC Europe 2007 Workshop on Electromagnetic Compatibility*, June 2007.

[61] Y. C. Tay, K. Jamieson, and H. Balakrishnan, “Collision-minimizing CSMA and its applications to wireless sensor networks,” *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1048–1057, August 2004.

[62] K. Langedoen, “The MAC alphabet soup served in wireless sensor networks, http://www.st.ewi.tudelft.nl/ koen/macsoup/,” 2006.

[63] K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” *Ad Hoc Networks*, vol. 3, pp. 325–349, 2005.

[64] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, “A two-tier data dissemination model for large-scale wireless sensor networks,” in *MobiCom ’02: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking*, (Atlanta, USA), pp. 148–159, ACM, September 2002.

[65] C. Perkins, E. Royer, and S. Das, “Ad hoc on-demand distance vector (AODV) routing.” RFC 3626, pp. 1-75, July 2003.

[66] D. Johnson, Y. Hu, and D. Maltz, “The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4.” RFC 4728, pp. 1-107, February 2007.

[67] T. Clausen, P. Jacquet, C. Adjih, A. Laouiti, P. Minet, P. Muhlethaler, A. Qayyum, and L.Viennot, “Optimized link state routing protocol (OLSR).” RFC 3626, pp. 1-75, October 2003. Network Working Group.

[68] J. N. Al-Karaki and A. E. Kamal, “Routing techniques in wireless sensor networks: A survey,” *IEEE Wireless Communications*, vol. 11, pp. pp. 6–28, June 2004.

[69] E. M. Royer and C.-K. Toh, “A review of current routing protocols for ad hoc mobile wireless networks,” *IEEE Personal Communications*, vol. 6, pp. 46–55, April 1999.

[70] D. B. Johnson and D. A. Maltz, “Dynamic source routing in ad hoc wireless networks,” in *Mobile Computing*, pp. 153–181, Kluwer Academic Publishers, 1996.

[71] A. Manjeshwar and D. P. Agrawal, “TEEN: a routing protocol for enhanced efficiency in wireless sensor networks,” in *Proc. 15th International Parallel and Distributed Processing Symposium*, pp. 2009–2015, April 2001.

[72] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, “Directed diffusion for wireless sensor networking,” *IEEE/ACM Transactions on Networking*, vol. 11, pp. 2–16, February 2003.

[73] L. Viennot, P. Jacquet, and T. H. Clausen, “Analyzing control traffic overhead versus mobility and data traffic activity in mobile ad-hoc network protocols,” *ACM Wireless Networks Journal (Winet)*, vol. 10, no. 4, pp. 447–455, 2004.

[74] C. Mbarushimana and A. Shahrabi, “Comparative study of reactive and proactive routing protocols performance in mobile ad hoc networks,” in *Proc. 21st International Conference on Advanced Information Networking and Applications Workshops AINAW ’07*, vol. 2, pp. 679–684, May 2007.

[75] S. Ni, Y. Tseng, Y. Chen, and J. Sheu, “The broadcast storm problem in a mobile ad hoc network,” in *MobiCom ’99: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, (Seattle, Washington, USA), pp. 151–162, ACM, August 1999.

[76] W. Heinzelman, J. Kulik, and H. Balakrishnan, “Adaptive protocols for information dissemination in wireless sensor networks,” in *MobiCom ’99:*

Proceedings of the 5th Annual International Conference on Mobile Computing and Networking, (Seattle, Washington, USA), ACM, August 1999.

[77] C. E. Perkins and P. Bhagwat, “Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers,” *SIGCOMM Comput. Commun. Rev.*, vol. 24, no. 4, pp. 234–244, 1994.

[78] A. Neumann, C. Aichele, M. Lindner, and S. Wunderlich, “Better Approach To Mobile Ad-hoc Networking (B.A.T.M.A.N.).” Internet-Draft, pp. 1-24, April 2008. Network Working Group.

[79] D. F. Macedo, L. H. A. Correia, A. L. dos Santos, A. F. Loureiro, and J. M. S. Nogueira, “A pro-active routing protocol for continuous data dissemination in wireless sensor networks,” in *Proc. 10th IEEE Symposium on Computers and Communications ISCC 2005*, pp. 361–366, June 2005.

[80] A. Koliousis and J. Sventek, “Proactive vs reactive routing for wireless sensor networks,” Tech. Rep. TR-2007-293, University of Glasgow, Department of Computing Science, Dept of Computing Science, University of Glasgow, UK, 2007.

[81] A. Manjeshwar and D. P. Agrawal, “APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks,” in *Abstracts and CD-ROM Parallel and Distributed Processing Symposium., International, IPDPS 2002*, pp. 195–202, April 2002.

[82] R. Sivakumar, P. Sinha, and V. Bharghavan, “CEDAR: a core-extraction distributed ad hoc routing algorithm,” *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1454–1465, August 1999.

[83] V. Ramasubramanian, Z. J. Haas, and E. G. Sirer, “SHARP: a hybrid adaptive routing protocol for mobile ad hoc networks,” in *MobiHoc ’03: Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking & Computing*, (New York, NY, USA), pp. 303–314, ACM, 2003.

[84] Z. J. Haas, “A new routing protocol for the reconfigurable wireless networks,” in *IEEE 6th International Conference on Universal Personal Communications Record Conference Record*, vol. 2, pp. 562–566, October 1997.

[85] Y. Yu, R. Govindan, and D. Estrin, “Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks,”

Tech. Rep. Technical Report UCLA/CSD-TR-01-0023, UCLA Computer Science Dept., May 2001.

[86] C. Adjih, E. Baccelli, T. Clausen, P. Jacquet, and R. Rodolakis, "Fish eye OLSR scaling properties," *IEEE Journal of Communication and Networks (JCN), Special Issue on Mobile Ad Hoc Wireless Networks*, pp. 343–351, 2004.

[87] G. Pei, M. Gerla, and T.-W. Chen, "Fisheye state routing: a routing scheme for ad hoc wireless networks," in *Proc. IEEE International Conference on Communications ICC 2000*, vol. 1, pp. 70–74, June 2000.

[88] F. Ye, A. Chen, S. Lu, and L. Zhang, "A scalable solution to minimum cost forwarding in large sensor networks," in *Proc. Tenth International Conference on Computer Communications and Networks*, pp. 304–309, October 2001.

[89] D. Braginsky and D. Estrin, "Rumor routing algorthim for sensor networks," in *WSNA '02: Proceedings of the 1st ACM international Workshop on Wireless Sensor Networks and Applications*, (New York, NY, USA), pp. 22–31, ACM, 2002.

[90] C. Schurgers and M. B. Srivastava, "Energy efficient routing in wireless sensor networks," in *Proc. Communications for Network-Centric Operations: Creating the Information Force. IEEE Military Communications Conference MILCOM 2001*, vol. 1, pp. 357–361, October 2001.

[91] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *ACM SIGMOD Record*, vol. 31, no. 3, pp. 9–18, 2002.

[92] B. Williams and T. Camp, "Comparison of broadcasting techniques for mobile ad hoc networks," in *MobiHoc '02: Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing*, (Lausanne, Switzerland), pp. 194–205, ACM, 2002.

[93] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, pp. 366–379, October 2004.

[94] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *MobiCom '00: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, (Boston, USA), pp. 243–254, ACM, August 2000.

[95] Y. Xu, J. Heidemann, and D. Estrin, “Geography-informed energy conservation for ad hoc routing,” in *MobiCom ’01: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, (Rome, Italy), pp. 70–84, ACM, July 2001.

[96] T. He, J. A. Stankovic, C. Lu, and T. Abdelzaher, “SPEED: a stateless protocol for real-time communication in sensor networks,” in *Proc. 23rd International Conference on Distributed Computing Systems*, pp. 46–55, May 2003.

[97] E. Felemban, C.-G. Lee, and E. Ekici, “MMSPEED: Multipath Multi-SPEED protocol for qos guarantee of reliability and. timeliness in wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 5, pp. 738–754, June 2006.

[98] M. Mauve, J. Widmer, and H. Hartenstein, “A survey on position-based routing in mobile ad hoc networks,” *IEEE Network*, vol. 15, pp. 30–39, November 2001.

[99] J. Hightower and G. Borriello, “Location systems for ubiquitous computing,” *Computer*, vol. 34, pp. 57–66, August 2001.

[100] K. Langendoen and N. Reijers, “Distributed localization in wireless sensor networks: a quantitative comparison,” *Computer Networks*, vol. 43, no. 4, pp. 499–518, 2003.

[101] C. Savarese, J. Rabaey, and J. Beutel, “Locationing in distributed ad hoc wireless sensor networks,” in *Proc. 2001 Int’l Conf. Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 4, pp. pp. 2037–2040, IEEE, Piscataway, NJ, May 2001.

[102] A. S.-I. Noh, W. J. Lee, and J. Y. Ye, “Comparison of the mechanisms of the zigbee’s indoor localization algorithm,” in *Proc. Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing SNPD ’08*, pp. 13–18, August 2008.

[103] J. Kulik, W. Heinzelman, and H. Balakrishnan, “Negotiation-based protocols for disseminating information in wireless sensor networks,” *Wireless Networks*, vol. 8, pp. 169–185, 2002.

[104] S. Dulman, T. Nieberg, J. Wu, and P. Havinga, “Trade-off between traffic overhead and reliability in multipath routing for wireless sensor networks,” in *Proc. IEEE Wireless Communications and Networking WCNC 2003*, vol. 3, pp. 1918–1922, March 2003.

[105] R. C. Shah and J. M. Rabaey, “Energy aware routing for low energy ad hoc sensor networks,” in *Proc. WCNC2002 Wireless Communications and Networking Conference 2002 IEEE*, vol. 1, pp. 350–355, March 2002.

[106] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, “Highly-resilient, energy-efficient multipath routing in wireless sensor networks,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 4, pp. 11–25, 2001.

[107] N. Sadagopan, B. Krishnamachari, and A. Helmy, “Active query forwarding in sensor networks (ACQUIRE),” *Journal of Ad Hoc Networks*, vol. 3, pp. 91–113, January 2005.

[108] I. F. Akyildiz, T. Melodia, and K. R. Chowdury, “Wireless multimedia sensor networks: A survey,” *IEEE Wireless Communications Magazine*, vol. 14, pp. 32–39, December 2007.

[109] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “Wireless multimedia sensor networks: Applications and testbeds,” *Proceedings of the IEEE*, vol. 96, pp. 1588–1605, October 2008.

[110] Z. Wang and J. Crowcroft, “Quality-of-service routing for supporting multimedia applications,” *IEEE Journal on Selected Areas in Communications*, vol. 14, pp. 1228–1234, September 1996.

[111] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, “Protocols for self-organization of a wireless sensor network,” *IEEE Personal Communications*, vol. 7, pp. 16–27, October 2000.

[112] M. Gerharz, C. de Waal, M. Frank, and P. Martini, “Link stability in mobile wireless ad hoc networks,” in *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks (LCN)*, (Tampa, USA), pp. 30–39, November 2002.

[113] P. Jacquet, A. Laouiti, P. Minet, and L. Viennot, “Performance analysis of OLSR multipoint relay flooding in two ad hoc wireless network models,” in *The second IFIP-TC6 NETWORKING Conference*, (Pisa), May 2002.

[114] S. Crisostomo, J. Barros, and C. Bettstetter, “Flooding the network: Multipoint relays versus network coding,” in *Proc. 4th IEEE International Conference on Circuits and Systems for Communications ICCSC 2008*, pp. 119–124, May 2008.

[115] T. J. Haas, J. Y. Halpern, and K. Li, “Gossip-based ad hoc routing,” *IEEE/ACM Transactions on Networking (TON)*, vol. 14, pp. 479–491, June 2006.

[116] H. Lim and C. Kim, “Multicast tree construction and flooding in wireless ad hoc networks,” in *MSWIM ’00: Proceedings of the 3rd ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, (Boston, USA), pp. 61–68, ACM, 2000.

[117] R. Baumann, S. Heimlicher, M. Strasser, and A. Weibel, “A survey on routing metrics,” Tech. Rep. 262, Computer Engineering and Networks Laboratory ETH Zürich, ETH-Zentrum, Switzerland, February 2007.

[118] R. Draves, J. Padhye, and B. Zill, “Comparison of routing metrics for static multi-hop wireless networks,” in *SIGCOMM ’04: Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, (Portland, USA), pp. 133–144, ACM, September 2004.

[119] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh networks,” in *MobiCom ’04: Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*, (Philadelphia, USA), pp. 114–128, ACM, October 2004.

[120] K. Naito, K. Mori, and H. Kobayashi, “Evaluation of power-aware routing for sensor networks with forwarder nodes,” *Journal on Systemics, Cybernetics and Informatics (JSCI)*, vol. 6, no. 5, pp. 87–92, 2008.

[121] S. Floyd, J. Handley, J. Padhye, and J. Widmer, “Equation-based congestion control for unicast applications,” in *SIGCOMM ’00: Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, (Stockholm, Sweden), pp. 43–56, ACM, August 2000.

[122] A. Khanna and J. Zinky, “The revised ARPANET routing metric,” *SIGCOMM Comput. Commun. Rev.*, vol. 19, no. 4, pp. 45–56, 1989.

[123] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *SenSys '09: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, (New York, NY, USA), pp. 1–14, ACM, 2009.

[124] C. E. Koksal and H. Balakrishnan, "Quality-aware routing metrics for time-varying wireless mesh networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1984–1994, Nov. 2006.

[125] S. Olariu and I. Stojmenovic, "Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting," in *Proc. 25th IEEE International Conference on Computer Communications INFOCOM 2006*, pp. 1–12, Apr. 2006.

[126] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660–670, October 2002.

[127] Q. Li, J. Aslam, and D. Rus, "Online power-aware routing in wireless adhoc networks," in *MobiCom '01: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, (Rome, Italy), pp. 97–107, ACM, June 2001.

[128] Y. Gadallah, "A comparative study of routing strategies for wireless sensor networks: Are MANET protocols good fit?," in *Ad-Hoc, Mobile, and Wireless Networks*, vol. 4104/2006, (Ottawa, Canada), pp. 5–18, Springer Berlin / Heidelberg, August 2006.

[129] J. Jubin and J. D. Tornow, "The DARPA packet radio network protocols," *Proceedings of the IEEE*, vol. 75, pp. 21–32, January 1987.

[130] J. B. Postel, "Transmission control protocol." RFC 793, pp. 1-84, September 1981.

[131] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*, (New York, NY, USA), pp. 56–67, ACM, 2000.

[132] L. Alazzawi and A. Elkateeb, “Performance evaluation of the WSN routing protocols scalability,” *Journal of Computer Systems, Networks, and Communications*, vol. 2008, pp. 1–9, December 2008.

[133] G. Montenegro, N. Kushalnagar, J. Hui, and D. Culler, “Transmission of IPv6 Packets over IEEE 802.15.4 Networks.” RFC 4944, pp. 1-30, September 2007. Network Working Group.

[134] N. Kushalnagar, G. Montenegro, and C. Schumacher, “IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals.” RFC 4919, pp. 1-12, August 2007. Network Working Group.

[135] M. Reineri, C. Casetti, and C.-F. Chiasserini, “Routing protocols for mesh networks with mobility support,” in *Proc. 6th International Symposium on Wireless Communication Systems ISWCS 2009*, pp. 71–75, Sept. 7–10, 2009.

[136] S. Basagni, I. Chlamtac, V. R. Syrotiuk, and B. A. Woodward, “A distance routing effect algorithm for mobility (DREAM),” in *MobiCom ’98: Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, (New York, NY, USA), pp. 76–84, ACM, 1998.

[137] E. R. Cavalcanti and M. A. Spohn, “Estimating the impact of mobility models’ parameters on mobility metrics in manets,” *IEEE Network Computing and Applications*, vol. 0, pp. 227–233, July 2009.

[138] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang, “A group mobility model for ad hoc wireless networks,” in *MSWiM ’99: Proceedings of the 2th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, (Seattle, Washington, United States), pp. 53–60, ACM, 1999.

[139] J. Boleng, W. Navidi, and T. Camp, “Metrics to enable adaptive protocols for mobile ad hoc networks,” in *In Proceedings of the International Conference on Wireless Networks (ICWNŠ02)*, pp. 293–298, 2002.

[140] M. Gerharz, C. de Waal, P. Martini, and P. James, “Strategies for finding stable paths in mobile wireless ad hoc networks,” in *Proc. 28th Annual IEEE International Conference on Local Computer Networks LCN ’03*, pp. 130–139, October 2003.

[141] W. Su, S.-J. Lee, and M. Gerla, “Mobility prediction and routing in ad hoc wireless networks,” *International Journal of Network Management*, vol. 11, no. 1, pp. 3–30, 2001.

[142] D. Kotz, C. Newport, R. S. Gray, J. Liu, Y. Yuan, and C. Elliott, “Experimental evaluation of wireless simulation assumptions,” in *MSWiM ’04: Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, (Venice, Italy), pp. 78–82, ACM, 2004.

[143] K. Pawlikowski, H.-D. J. Jeong, and J.-S. R. Lee, “On credibility of simulation studies of telecommunication networks,” *IEEE Communications Magazine*, vol. 40, pp. 132–139, January 2002.

[144] A. G. Longley and P. L. Rice, “Prediction of tropospheric radio transmission over irregularterrain, a computer method-1968.,” Tech. Rep. ERL 79-ITS 67, U.S. Dep. of Commerce, Boulder, CO, USA, 1968.

[145] D. J. Cichon and T. Kurner, *Propagation Prediction Models Digital Mobile Radio Toward Future Generation Systems*. No. Chapter 4, http://www.lx.it.pt/cost231/: COST 231 Final Report, European Comission, 1999.

[146] A. Chandra, V. Gummalla, and J. O. Limb, “Effect of turn-around times on the performance of high speed ad-hoc MAC protocols,” in *NETWORKING ’00: Proc. of the IFIP-TC6 European Commission International Conference on Broadband Communications, High Performance Networking, and Performance of Communication Networks*, (London, UK), pp. 507–517, Springer-Verlag, January 2000.

[147] A. Einstein, “On the motion, required by the molecular-kinetic theory of heat, of particles suspended in a fluid at rest,” *Ann. Phys*, vol. 17, pp. 549–560, 1905.

[148] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser, “An analysis of the optimum node density for ad hoc mobile networks,” in *Proc. IEEE International Conference on Communications ICC 2001*, vol. 3, pp. 857–861, June 2001.

[149] J.-K. Chen, C. Chen, R.-H. Jan, and H.-H. Li, “Expected link life time analysis in manet under manhattan grid mobility model,” in *MSWiM ’08: Proceedings of the 11th International Symposium on Modeling, Analysis*

and Simulation of Wireless and Mobile Systems, (New York, NY, USA), pp. 162–168, ACM, 2008.

[150] IEEE, “IEEE standard for modeling and simulation (MS) high level architecture (HLA) - framework and rules,” September 2000.

[151] M. Koivisto and A. Urbaczewski, “The relationship between quality of service perceived and delivered in mobile internet communications,” *Information Systems and E-Business Management*, vol. 2, pp. 309–323, September 2004.

[152] M. Riley and I. Richardson, *Digital Video Communication*. Norwood, MA, USA: Artech House, Inc., 1997.

[153] ITU-R, “Recommendation BT.500-10, methodology for the subjective assessment of the quality of television pictures,” 2000.

[154] J.-R. Ohm, *Multimedia Communication Technology*. Springer, 2004.

[155] J. Klaue, B. Rathke, and A. Wolisz, “Evalvid - a framework for video transmission and quality evaluation,” in *Proceedings of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, pp. 255–272, September 2003.

[156] MSU, “MSU video quality measurement tool, http://compression.ru/video/quality/.”

[157] Sarnoff Corporation, “Measuring image quality: Sarnoffs JNDmetrix technology, Sarnoff JNDmetrix technology overview, http://www.sarnoff.com/.”

[158] S. Wolf and M. Pinson, “Video quality measurement techniques,” NTIA Report 02-392, Institute for Telecommunication Sciences, 2002.

[159] H. Kasai and M. Nilsson, “The implementation of an audiovisual transcoding system,” in *Proceedings of the International Conference on Multimedia and Expo (ICME)*, pp. 357–360, August 2001.

[160] L. Merritt and R. Vanam, “X264: A high performance H.264 / AVC encoder,” tech. rep., University of Washington, 2007.

[161] T. Camp, J. Boleng, and V. Davies, “A survey of mobility models for ad hoc network research,” *Wireless Communication and Mobile Computing (WCMC): Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, no. 2, pp. 483–502, 2002.

[162] D. Djenouri, E. Nekka, and W. Soualhi, “Simulation of mobility models in vehicular ad hoc networks,” in *SOMITAS ’08: Proceedings of the 2008 Ambi-Sys workshop on Software Organisation and MonIToring of Ambient Systems*, (Quebec City, Canada), pp. 1–7, Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (ICST), 2008.

[163] M. C. González, C. A. Hidalgo, and A. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, pp. 779–782, June 2008.

[164] M. Musolesi and C. Mascolo, “Designing mobility models based on social network theory,” *SIGMOBILE Mob. Comput. Commun. Rev*, vol. 11, no. 3, pp. 1559–1662, 2007.

[165] A. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri, “Towards realistic mobility models for mobile ad hoc networks,” in *MobiCom ’03: Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, (San Diego, CA, USA), pp. 217–229, ACM, 2003.

[166] R. Baumann, S. Heimlicher, and M. May, “Towards realistic mobility models for vehicular ad-hoc networks,” in *Proc. Mobile Networking for Vehicular Environments*, pp. 73–78, May 2007.

[167] J. Yoon, M. Liu, and B. Noble, “Random waypoint considered harmful,” in *Proc. INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, vol. 2, pp. 1312–1321, April 2003.

[168] G. Lin, N. G, and R. Rajaraman, “Mobility models for ad hoc network simulation,” in *Proc. INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 1, Mar. 7–11, 2004.

[169] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong, “On the levy-walk nature of human mobility,” in *Proc. INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 924–932, April 2008.

[170] Y.-T. Wu, W. Liao, C.-L. Tsao, and T.-N. Lin, “Impact of node mobility on link duration in multihop mobile networks,” *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 2435–2442, June 2009.

[171] D. Yu and H. Li, “Influence of mobility models on node distribution in ad hoc networks,” in *Proc. International Conference on Communication Technology ICCT 2003*, vol. 2, pp. 985–989, April 2003.

[172] C. Bettstetter, G. Resta, and P. Santi, “The node distribution of the random waypoint mobility model for wireless ad hoc networks,” *IEEE Transactions on Mobile Computing*, vol. 2, pp. 257–269, July 2003.

[173] M. Sánchez and P. Manzoni, “ANEJOS: A java based simulator for ad hoc networks,” *Future Generation Computer Systems*, vol. 17, no. 5, pp. 573–583, 2001.

[174] C. Tuduce and T. Gross, “A mobility model based on WLAN traces and its validation,” in *Proc. IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM 2005*, vol. 1, pp. 664–674, Mar. 13–17, 2005.

[175] A. Rojas, P. Branch, and G. Armitage, “Validation of the random waypoint mobility model through a real world mobility trace,” in *Proc. TENCON 2005 IEEE Region 10*, pp. 1–6, November 2005.

[176] C. Bettstetter, “Mobility modeling in wireless networks: Categorization, smooth movement, and border effects,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 3, pp. 55–66, 2001.

[177] B. Ishibashi and R. Boutaba, “Topology and mobility considerations in mobile ad hoc networks,” *Ad Hoc Networks*, vol. 3, no. 6, pp. 762–776, 2005.

[178] E. W. Weisstein, *The CRC Concise Encyclopedia of Mathematics*. CRC Press, 1998.

[179] B. Liang and Z. J. Haas, “Predictive distance-based mobility management for PCS networks,” in *Proc. IEEE Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM ’99*, vol. 3, pp. 1377–1384, March 1999.

[180] A. N. Mian, R. Baldoni, and R. Beraldi, “An adapting random walk for ad hoc and sensor networks,” in *Proc. 29th IEEE International Conference on Distributed Computing Systems Workshops ICDCS Workshops ’09*, pp. 153–157, June 2009.

[181] L. Lima and J. Barros, “Random walks on sensor networks,” in *Proc. 5th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops WiOpt 2007*, pp. 1–5, April 2007.

[182] K. K. Rachuri and C. Murthy, “Energy efficient search in sensor networks using simple random walks with level biased steps,” in *Proc. 29th IEEE International Conference on Distributed Computing Systems Workshops ICDCS Workshops ’09*, pp. 178–185, June 2009.

[183] F. Bartumeus, M. G. E. Da Luz, G. M. Viswanathan, and J. Catalan, “Animal search strategies: A quantitative random-walk analysis,” *Ecology*, vol. 86, pp. 3078–3087, November 2005.

[184] I. Mabrouki, X. Lagrange, and G. Froc, “Random walk based routing protocol for wireless sensor networks,” in *ValueTools ’07: Proceedings of the 2nd international Conference on Performance evaluation Methodologies and Tools*, (Nantes, France), pp. 1–10, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007.

[185] J. Lu and Y. Liu, “Research on the node probability distribution of random direction model in manet,” in *Proc. 5th International Conference on Wireless Communications, Networking and Mobile Computing WiCom ’09*, pp. 1–4, September 2009.

[186] F. Bai, N. Sadagopan, and A. Helmy, “Important: a framework to systematically analyze the impact of mobility on performance of routing protocols for ad hoc networks,” in *Proc. INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, vol. 2, pp. 825–835, April 2003.

[187] C. P. Agrawal, M. K. Tiwari, and O. P. Vyas, “Evaluation of AODV protocol for varying mobility models of MANET for ubiquitous computing,” in *Proc. Third International Conference on Convergence and Hybrid Information Technology ICCIT ’08*, vol. 1, pp. 769–774, November 2008.

[188] A. M. Edwards, R. A. Phillips, N. W. Watkins, M. P. Freeman, E. J. Murphy, V. Afanasyev, S. V. Buldyrev, M. G. E. Luz, E. P. Raposo, H. E. Stanley, and G. M. Viswanathan, “Revisiting levy flight search patterns of wandering albatrosses, bumblebees and deer,” *Nature*, vol. 449, pp. 1044–1048, October 2007.

[189] G. Jayakumar and G. Ganapathi, “Reference point group mobility and random waypoint models in performance evaluation of MANET routing protocols,” *Journal of Computer Systems, Networks, and Communications*, vol. 2008, Article ID 860364, pp. 1–10, 2008.

[190] A. Muneb, U. Saif, A. Dunkels, T. Voigt, K. Römer, K. Langendoen, J. Polastre, and Z. A. Uzmi, “Medium access control issues in sensor networks,” *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 2, pp. 33–36, 2006.

[191] R. S. Gray, D. Kotz, C. Newport, N. Dubrovsky, A. Fiske, J. Liu, C. Masone, S. McGrath, and Y. Yuan, “Outdoor experimental comparison of four ad-hoc routing algorithms,” in *MSWiM ’04: Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, (Venice, Italy), pp. 220–229, ACM, 2004.

[192] H. Schultzrinne, S. Casner, and V. Frederick, R. Jacobson, “RTP: A transport protocol for real-time applications.” RFC 3550, pp. 1-104, July 2003.

ISSN 1432-8801