Macrothink Institute™

Network Protocols and Algorithms
ISSN 1943-3581
2010, Vol. 2, No. 1

Traffic Analysis & Modeling in Wireless Sensor Networks and Their Applications on Network Optimization and Anomaly Detection[1,2]

Qinghua Wang[3]

Dept. of Information Technology and Media, Mid Sweden University

SE-85170 Sundsvall, Sweden

Tel: +46-60-148914 E-mail: qinghua.wang@ieee.org

Abstract

Wireless sensor network (WSN) has emerged as a promising technology thanks to the recent advances in electronics, networking, and information technologies. However, there is still a great deal of additional research required before it finally becomes a mature technology. This article concentrates on three factors which are holding back the development of WSNs. Firstly, there is a lack of traffic analysis & modeling for WSNs. Secondly, network optimization for WSNs needs more investigation. Thirdly, the development of anomaly detection techniques for WSNs remains a seldom touched area. Among these three factors, the understanding regarding the traffic dynamics within WSNs provide a basis for further works on network optimization and anomaly detection for WSNs.

Keywords: Anomaly detection, network optimization, traffic analysis, traffic modeling, wireless sensor network.

[1] This work has been based on the author’s PhD dissertation [1].

[2] This work was partially carried out during the tenure of an ERCIM “Alain Bensoussan” Fellowship Programme.

[3] Current address: Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim N-7491, Norway. Tel: +47-73594331.

1. Introduction

Wireless sensor network (WSN) has emerged as a promising technology because of the recent advances in electronics, networking, and information processing. The WSN research was initially driven by military applications such as battlefield surveillance and enemy tracking. Now, many civil applications of WSN have also been proposed, which include habitat monitoring, environmental observation and forecasting systems, health monitoring, etc. In these applications, many low power and inexpensive sensor nodes are deployed in a vast space to cooperate as a network.

Although WSN is a promising technology which can be used in many applications, there are still a few obstacles to overcome before it finally becomes a mature technology. One of the key obstacles is the energy constraint suffered by the most inexpensive sensor nodes, where batteries are the main source of power supply. Given this obstacle cannot be removed in the near future, optimizing the design of WSNs thus the minimum energy will be consumed is very important.

In WSNs, communication is believed to dominate the energy consumption [2]. Energy expenditure is less for sensing and computation. The energy cost of transmitting 1 Kb a distance of 100 meters is approximately the same as that for the execution of 3 million instructions by using a general-purpose processor [3]. Thus, minimizing the energy consumption due to communication is the key for the relief of the energy constraint in WSNs.

Currently, the knowledge about the communication in WSNs is still partial and vague, especially for traffic characteristics and communication patterns. Obviously, the knowledge about the traffic characteristics and communication patterns can aid in the understanding of the energy consumption and its distribution in WSNs. Thus, the investigation of traffic characteristics and communication patterns is a good starting point in the search for more energy-efficient WSNs. Following on from this it will be possible to propose new solutions for the design of WSNs in order to optimize the energy consumption.

Another concern for WSN technology involves security. WSNs will not be successfully deployed if the security issue is not addressed adequately. Security becomes more important because WSNs are usually used for very critical applications. Furthermore, WSNs are very vulnerable and thus attractive to malicious attacks because of their cheap prices, human-unattended deployment and the nature of wireless communication. The existing solutions to the security in WSNs include using key management and authentication [4]. However, these preventive mechanisms cannot deter all possible attacks (e.g. insider attacks possessing the key). Actually, malicious attacks may exhibit anomalous behaviours in WSNs. With regard to communication, malicious attacks can trigger arbitrary communications, while a normal communication must follow protocol specifications and application scenarios. Thus, it should be interesting to investigate the possibility of detecting malicious attacks by identifying the anomalies exhibited within the WSNs' communication traffic.

Because sensor nodes are cheap devices and they can be deployed in harsh environments (e.g. battlefield, forest), they are prone to fail either by themselves or by means of others (e.g.

enemies, animals). Further, it is also common for battery-supported sensor nodes to fail because of energy exhaustion. To provide efficient maintenance for WSNs, those performing this maintenance require instant notifications about the sensor node failures. Because a failed sensor node cannot maintain efficient communication with the other nodes, sensor node failures have the possibility to be instantly noticed by observing the degraded or lost communication in relation to the failed nodes. This strategy has a similarity with the detection of traffic anomalies caused by malicious attacks. Both of them require comprehensive knowledge about the communication traffic before they can identify any traffic anomaly.

The aim of this article is to investigate the communication traffic dynamics and patterns in WSNs and find their applications with reference to network optimization and network anomaly detection. The applications of WSNs are abundant. Because the communication traffic in WSNs is very dependent on the application scenario, only those selected typical WSN scenarios (e.g. surveillance, target tracking) will be investigated. Additionally different types of communication traffic exist, including data traffic, routing discovery traffic, link layer feedback and hello message, etc. This article mainly focuses on data traffic and there is a limited involvement of other traffic types.

In the following, the survey of the works in the field of traffic analysis & modeling, in the field of network optimization and in the field of network anomaly detection will be presented separately. However, particular emphasis will be put on the relationships between traffic analysis & modeling and network optimization, and between traffic analysis & modeling and network anomaly detection throughout the presentation.

2. Traffic Analysis & Modeling for WSNs

WSNs consist of a large number of tiny and cheap sensor nodes that cooperatively sense a physical phenomenon. Existing research results and products have provided the possibility to build effective WSNs for many applications. If the traffic features inside WSNs were better understood then the WSNs could be made to be even more effective. For example, better routing protocols and sensor deployment strategy could be designed if the traffic burden among the sensors was better understood. Better fault and security management could be applied if normal and abnormal traffic could be kept apart according to traffic features.

The traffic dynamics for different types of traditional networks, both wired and wireless, have been investigated in the literature. However, the specialty of WSNs makes a reinvestigation of traffic dynamics necessary. Constructing accurate and analytically tractable models for sensor network traffic will provide a basis for future work on network design, optimization and security. Unfortunately, at the time that this article was written, research regarding traffic modeling and analysis in WSNs was still rather limited. The few studies that do exist include works focusing on data traffic arrival process, sequence relations among general kinds of packets, and data traffic load distribution.

2.1 Data Traffic Arrival Process

Because the data traffic dynamics in different WSN scenarios are quite different, the data traffic modeling and analysis in WSNs will be quite application dependent. In [5], it is suggested that WSN applications can be categorized as *event-driven* or *periodic data generation*. For periodic data generation scenarios, constant bit rate (CBR) can be used to model the data traffic arrival process when the bit rate is constant [6]. When the bit rate is variable, a Poisson process can be used to model the data traffic arrival process as long as the data traffic is not bursty [7]. For event-driven scenarios such as *target detection* and *target tracking*, bursty traffic can arise from any corner of the sensing area if an event is detected by the local sensors. A Poisson process has also been used to model the traffic arrival process in an event-driven WSN [8]. However, there is no solid ground to support the use of a Poisson process in this case. Actually, the widely used Poisson processes are quite limited in their burstiness [9, 10]. Instead of using Poisson processes, the author of this article proposes to use an ON/OFF model (see Figure 1) to capture the burst phenomenon in the source data traffic of an event-driven WSN [11]. Further, the distributions of ON/OFF periods are found to follow the generalized Pareto distribution in his considered WSN scenario. Ref. [12] studies a different WSN scenario - a mobile sensor network (MSN). In an MSN, the node mobility introduces new dynamics to network traffic. In [12], the authors find that the mobility variability of humans (in this case, sensor nodes are attached to humans) and the spatial correlation of the collected information lead to the pseudo-LRD (i.e. long range dependent) traffic, which exhibits characteristics significantly different to that of Markovian traffic.

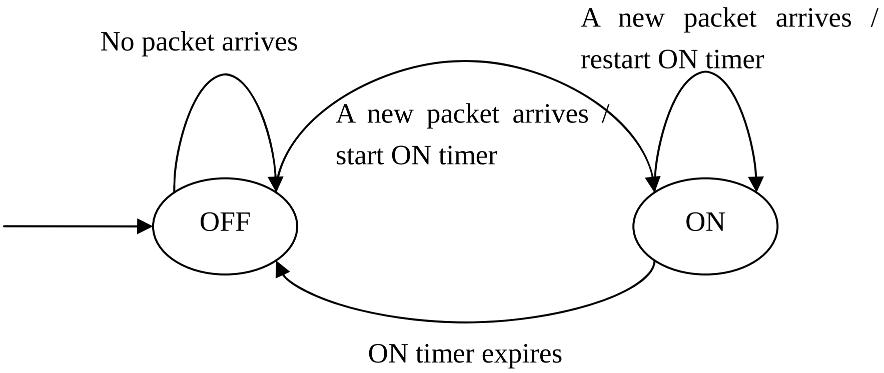
No packet arrives
A new packet arrives / restart ON timer
A new packet arrives / start ON timer
OFF
ON
ON timer expires

Figure 1: ON/OFF state transition diagram

2.2 Sequence Relations among General Kinds of Packets

Sequence relations exist in some kinds of packets. For example, a Routing Reply message always comes after a Routing Request message and that is specified by any ordinary routing protocol. In [13], the authors propose to use a finite state machine (FSM) to specify correct routing behavior for the ad hoc on demand distance vector (AODV) routing [14]. The rationale behind this is that the AODV protocol has specified the sequence relations among different kinds of routing messages and such sequence relations can be depicted by an FSM. The authors in [15] also use FSM to model the correct routing behavior for the dynamic source routing (DSR) [16]. Because the routing protocols AODV and DSR have clearly specified the routing operations, the sequence relations among different kinds of routing packets can be manually abstracted into an FSM. In both [13] and [15], the authors have used their FSMs to validate real-time routing behaviors and detect possible malicious attacks.

In addition to that the sequence relations among some special kinds of packets (e.g. routing messages) are possible to be specified according to protocol specifications, the author of this article suggests that the sequence relations among general kinds of packets can also be learned automatically by on-line training. In [17], the authors firstly classify the arriving packets according to their attributes (e.g. packet type, addresses) and then map the packet arriving sequence to an infinite character string. Afterwards, the on-line learning of the packet sequence relations are conducted by extracting every unique character substring encountered during the window-based scanning process. The learned packet sequence relations can be used to build the normal traffic profile for the node of interest in a static WSN. In a dynamic WSN in which some of the nodes are mobile, the traffic profile learned in this manner will evolve quickly over time and will thus be less meaningful.

2.3 Data Traffic Load Distribution

In a WSN, the data traffic load is not evenly distributed over the nodes. For example, the sensors which are one hop away from the sink relay the entire network's data traffic. This imbalanced data traffic load distribution can degrade the network's lifetime and functionality. Hence, efforts have been devoted to characterizing the data traffic load distribution in WSNs. Ref. [18] proposes an analytical analysis on the data traffic load distribution over a randomly deployed linear WSN. It has been shown that the data traffic load over a node increases the closer it is to the sink, however, a reduction in the data traffic load is expected for sensors that are very close to the sink. In [19, 20], data traffic load is formularized as a function of the distance to the sink in dense planar WSNs. In a similar manner to that in a linear WSN, the data traffic load over a node in planar WSNs also increases as the node moves closer to the sink. For a symmetric sensor network (i.e. all nodes of the same distance from the center of the network are similar) with nodes evenly distributed in the sensing field, the author of this article concludes that the expected data traffic load over a node is in direct proportion to the network radius, in inverse proportion to the mean routing hop length, and independent of the node density [20].

Because the distribution of data traffic load is closely related to the distribution of energy consumption and the latter has a significant impact on the performance of WSNs, the research results concerning the distribution of data traffic load can be used to optimize the performance of WSNs. For example, the author of this article has proposed an optimal energy allocation scheme for WSNs based on the understanding of the data traffic related energy consumption in the network [21].

3. Network Optimization for WSNs

There are many network optimization problems to be solved in WSNs, such as rate control, flow control, congestion control, medium access control, queue management, power control and topology control, etc. [22]. It is difficult to provide a complete overview in relation to all issues relating to network optimization in WSNs. However, it is worthwhile, none the less, to aim for a fairly comprehensive summary of important topics, with particular emphasis on the optimization of energy consumption.

3.1 Energy-Efficient Routing Design

Because communication dominates the critical energy consumption, routing design is usually considered to be the core of sensor network design. Many routing algorithms have been proposed in prior research. The shortest path is the typical and fundamental consideration for network flow routing problems. A simple translation of this consideration in sensor network routing is the minimum hop (MH) routing. The AODV routing is an example of using the number of link hops as its routing metric. However as the limitation of battery power is one of the most fundamental aspects of sensor networks, routing algorithms for sensor networks generally attempt to minimize the utilization of this valuable resource. Many researchers have proposed shortest path algorithms in order to minimize the utilization of energy. For example, the minimum total transmission power routing (MTPR) proposed in [23] and the minimum total energy (MTE) routing introduced in [24, 25] attempt to reduce the total transmission energy per data bit, where the path length is the sum of the energy expended per data bit during its transmission over all links in the path.

It was realized by the sensor network research community that improving the ratio of packets transmitted to energy consumed by the network is, by itself, not a good measure of the efficiency of the network [26]. Ref. [25] proposes an algorithm which attempts to minimize the variation in node energy levels. This metric ensures that all the nodes in the network remain up and running together for as long as possible. A flow augmentation (FA) [24, 27] algorithm incorporates MH, MTE, and other residual energy considered routing algorithms together with adjustable parameters. The maximum residual energy path (MREP) routing [27, 28] is an algorithm based on similar considerations which attempts to postpone the death of the first node by using the maximum remaining energy path.

To provide more insights into the energy-efficient routing design, many theoretical analyses concerning the optimal routing performance have also been conducted. In [28], the authors consider the problem of choosing routes between a set of source nodes and a set of sink nodes of an ad-hoc network so that the time until the first battery expires, is maximized. The authors note that choosing a route which results in minimum total energy expenditure is not always desirable because some of the nodes may have an excessive relaying burden, and hence these nodes may expire too soon. This in turn could lead to a loss of connectivity. To overcome this problem, the authors suggest that the routes should be chosen with the ultimate objective of maximizing the time until the first battery expires. In order to achieve this objective, the minimum energy paths are not necessarily the best choices. In [28], such an energy-efficient routing problem reduces to a linear programming problem which is described as the following:

$$\begin{array}{ll} \texttt{max} & \textit{Lifetime} \\ \texttt{s.t.} & 1.\ \textit{Energy Constraint} \\ & 2.\ \textit{Flow Conservation Constraint} \end{array} \qquad (1)$$

where *Lifetime* is the network operational time till the first battery expires, *Energy Constraint* specifies that the energy expended by sensing, communication and other operations cannot surpass the initial energy reserves, and *Flow Conservation Constraint* specifies that the number of outgoing data flows of each node should be equal to the sum of the number of incoming data flows of that node plus the number of data flows originating at that node. Obviously, the data flows which maximize the *Lifetime* correspond to the optimal routing strategy. The authors of [29] have a similar concern. They also consider the lifetime of a network until the first battery expires, and the network suffers from both the flow conservation constraint and the energy constraint. Besides the performance of the optimal routing, the authors are also interested in the following question: How much improvement in the lifespan of a network can be expected by changing only the routing algorithm? Thus, they have computed explicit bounds on both the minimal and the maximal energy that routings will consume, and used them to bound the lifetime of the network.

However, the fact that the routing strategy is designed in such a way that all nodes die simultaneously (by attempting to postpone the time that the first battery expires) does not automatically imply that the energy utilization is optimal [26]. In reality, the energy possessed by a normal sensor node is very easily exhausted and thus the node fails. For many sensor network applications such as military surveillance, full or guaranteed sensing coverage can still be provided in the case of sensor failures, by leveraging the redundant deployment of sensor nodes. Given that the network can still be useful even after some of sensor nodes have died, the metric attempting to postpone the death of the first node is unable to offer an optimal solution. The authors of [26] have exhibited similar thinking and define the network lifetime as the time elapsed for some fraction of nodes in the network to die, which is more practical than earlier definitions which use the time to the death of the first node as the network's lifetime. Unfortunately, as the network lifetime definition changed, the fundamental performance bound or the reference to the optimal solution also became unclear.

In [30, 31], the author of this article uses a new concept called application-tolerable network run-time information-collecting ability in judging the lifetime of a network, which is more information oriented compared to the definition used in [26]. With this new network lifetime definition, nodes are allowed to die during the network's operational lifetime, which means that the network topology could change during the network operational lifetime and the data transmission between any two nodes could become unstable. All these make it difficult to give a linear programming optimization model similar to those proposed in [24, 28 and 32]. Thus, a relaxed linear programming optimization model which can give a tight upper bound is instead proposed in [30, 31]. In a similar manner to (1), the optimization problem formulated in [30, 31] is described as the following:

max *Lifetime* or *Total Information Collected*

s.t. 1. *Energy Constraint*

2. *Flow Conservation Constraint*

3. *Application-Dependent Requirement on Network Information-Collecting Ability* (2)

In the above, *Lifetime* is the network operational time till the network's information collecting ability falls below the application-dependent requirement. *Total Information Collected* is a performance metric which could be more suitable for information-collecting purpose WSNs. It represents the total information collected by the entire network throughout its lifetime. *Energy Constraint* and *Flow Conservation Constraint* have the same meaning as those in (1). *Application-Dependent Requirement on Network Information-Collecting Ability* specifies the worst network information collecting ability which can be tolerated by the application. It can be also viewed as a translation of the network lifetime definition. The results obtained through this model offer insights for future routing design and can also be used as benchmarks in the evaluation of energy-efficient routing algorithms designed for WSNs.

3.2 Energy-Efficient MAC Design

Compared to routing protocols, medium access control (MAC) protocols provide more direct influence over the utilization of the transceiver which is the largest energy consumer in most sensor nodes. Traditionally, MAC protocols are designed to maximize packet throughput, minimize latency and provide fairness. However, the design of MAC protocols for WSNs focuses on minimizing energy consumption.

It has been identified that the idle mode energy expenditure may spend a considerable amount of energy in WSNs [33]. Because many WSN applications possess a low message rate characteristic, most energy will be wasted by *idle listening* when traditional MAC protocols are used for WSNs: Since a node does not know when it will be the receiver of a message from one of its neighbors, it must maintain its radio in receive mode at all times. If nodes exchange short messages with their neighbors at an average rate of one per second and both the transmitting and the receiving of a short message take 5 milliseconds, then the radio will spend 99% of the time on idle listening [34].

There are several solutions addressing the problem of energy waste due to idle listening. In general, some kind of duty cycle is involved, which allows each node to sleep periodically. TDMA-based protocols are naturally energy preserving. However, allocating TDMA slots is a complex problem that requires coordination. Another way of energy saving is to use an extra radio, which operates on a different frequency to that of the radio used for communication [35]. However, this approach is not appropriate for most wireless sensor nodes currently in use where only a single radio is available on each node. S-MAC [36] is a single-frequency contention-based protocol specially designed for WSNs. It divides the time into fairly large frames. Each frame consists of two parts: an active period and a sleeping period. During the sleeping period, a node turns off its radio in order to preserve energy. During the active period, a node communicates with its neighbors and sends any message queued during the sleeping period. In order to synchronize, the sensor nodes periodically transmit SYNC messages at the beginning of the active period. The SYNC messages allow the sensor nodes to learn of their neighbors' schedules so that they can wake up at the appropriate time. Each sensor node performs a simple contention avoidance algorithm based on a random backoff to limit the number of SYNC message collisions. The T-MAC [34] protocol extends S-MAC by using a timer to indicate the end of the active period instead of relying on a fixed duty cycle schedule. By adaptively ending the active period, T-MAC nodes may save energy by lowering the amount of time they spend on idle listening and also by adapting to changes in traffic conditions.

3.3 In-Network Processing

WSNs are capable of collecting an enormous amount of data over space and time. Often, raw data is transmitted from each sensor node to a central processing location. This may cause a significant drain on communication and energy resources. However, in many applications, the ultimate objective is not merely the collection of “raw” data, but rather an estimate of certain environmental parameters or functions of interest (e.g., source locations, spatial distributions) [37]. Distributed in-network processing, which eliminates the need to transmit raw data to a central point, may significantly reduce the communication and energy resources consumed.

There have been many existing in-network processing approaches many of which are combined with routing algorithms. If the ultimate objective is to compute the average or other quadratic cost functions of all the measurements, the estimate of the objective parameter can be passed and updated along a routing path which passes through all the nodes and visits each node just once [37]. Each node updates the estimate by adjusting the previous value to improve or reduce its local cost and then passes the update to the next node. In the case of a quadratic cost function, one pass through the network is sufficient to achieve the objective. In more general cases, several “cycles” through the network are required in order to obtain a solution. The LEACH protocol presented in [38] is an elegant solution to the data aggregation problem in which clusters are formed in a self-organized manner to fuse data before transmitting it to the base station or sink. In LEACH, a designated node in each cluster, called the clusterhead, is responsible for collecting and aggregating the data from sensors in its cluster and eventually transmitting the result to the base station or sink. In [39], the authors propose a new chain-based protocol called PEGASIS that minimizes the energy consumption at each sensor node. The key idea is that nodes organize to form a chain and each node takes turns in being the leader for communication to the base station or sink. The data is collected by starting from each endpoint of the chain and aggregated along the path to the designated head node. Unlike LEACH, PEGASIS uses a flat topology thereby eliminating the overhead of dynamic cluster formation.

3.4 Load Balancing

In WSNs, the dominating communication pattern is that a large number of sensor nodes deliver their sensed information to one or a few data sinks through multi-hop transmission [30, 31]. This kind of communication pattern causes a drastic imbalance to the traffic load distribution across the network in which the nodes close to a sink experience heavy traffic loads. Since communication is believed to dominate the energy consumption of a sensor node [38] and sensor nodes are usually provided with limited energy resources, the imbalanced traffic load distribution is very harmful and it could cause the nodes close to a sink to die at an earlier stage which thus renders the remainder of the network to be useless.

To counter or alleviate the harm resulting from an uneven traffic load distribution, many researchers have turned their attention to the problem of load balancing. The authors of [27, 40, and 41] realize that the imbalanced traffic load distribution can cause one part of nodes to die earlier than the others, thus degrading the network performance. To counter the negative effect of the imbalanced traffic load distribution on network performance, new routing algorithms which resort to the measure of the remaining energy reserves and other kinds of path capacity measurements are proposed. The authors of [42] consider the load balancing problem of uniformly distributed traffic demands in a unit disk. By deliberately routing traffic along slightly longer paths instead of the shortest paths, the highly congested links are avoided and a particularly flat traffic load distribution is achieved. The authors of [43] address the problem of balancing the traffic load in multi-hop wireless networks with uniformly distributed point-to-point communication. They develop a routing algorithm called Curveball Routing which can avoid the crowded center and provide a performance which is not significantly worse than that of the optimum. The authors of [44] propose an algorithm that makes a decision at each step as to whether to propagate data one-hop towards the sink, or to send data directly to the sink in order to balance the energy consumption over the nodes. If appropriate, data aggregation and in-network processing techniques are also methods for balancing traffic distribution. The adoption of data aggregation not only reduces the total amount of packets being transmitted but also yields a more even traffic distribution.

3.5 Resource Allocation

Fair resource allocation is another approach to counter the harm resulting from uneven traffic load distribution as explained in Section 3.4.

In the category of fair resource allocation, resources (e.g. energy, bandwidth, nodes) are allocated to an object (e.g. a node or an application) according to the workload of that object. The authors of [45] present an optimal energy allocation criterion and thus all clusters have the same exhaustion time in a cluster based WSN. The author of this article has discovered that the performance upper bounds in a WSN linearly increase with the energy reservation in an identified bottleneck zone; thus assigning more available energy resources to the important bottleneck zone can effectively alleviate the bottleneck effect [46]. Radio range adjustment is also proposed to save the energy consumption on a routing path [45, 47 and 48]. However, this must be conducted with caution since assigning shorter relaying ranges for nodes closer to the sink adds more imbalances to the already imbalanced traffic load distribution. In [49], the authors propose to place additional sensor nodes around the sink nodes to mitigate their hot spot problem (i.e. uneven traffic load distribution). Their results show that for certain networks only a limited number of additional nodes are required to fourfold network lifetime. Similar thinking appears in [50] where the authors propose efficient node placement and topology control protocols to balance the power consumption of sensor nodes. More specifically, they propose the allocation of more sensor nodes to the zone closer to the sink and also to assign a smaller packet transmission power to them. In addition, the author of this article proposes a fair energy allocation scheme such that the initial energy resource allocated to a node is proportional to its expected traffic load [21]. Because traffic load is an indicator of the energy consuming rate, the proposed fair energy allocation scheme maximizes the network lifetime by equalizing the expected energy exhaustion time of all sensor nodes.

4. Anomaly Detection for WSNs

4.1 The Necessity of Anomaly Detection in WSNs

WSNs consist of a large number of tiny sensor devices that have limited power and limited sensing, computation, and wireless communication capabilities. Sensor nodes usually operate in unattended and even harsh environments, and as a result, sensor nodes are prone to failures and are vulnerable to malicious attacks. Since it is not possible to avoid the appearance of failures and malicious attacks, it will be essential that these failures and malicious attacks are detected immediately after their appearance. Thus, emergency responses can be made accordingly in order to mitigate the harm due to sensor failures and malicious attacks.

One common point between failures and malicious attacks is that they both cause errors inside the system. Therefore, the system can malfunction due to the errors caused. The difference is that failures cause errors randomly, but malicious attacks are usually done deliberately and will preferentially target the most important component in the system. In addition, failures can exist everywhere in the system and can happen at anytime, but the scope of malicious attacks is subject to the abilities of attackers. In terms of available techniques, there are similarities between the detection of failures and the detection of malicious attacks. Because errors caused by failures and malicious attacks are abnormal events in the system, it should be possible to detect such an event by realizing that there has been a deviation of a system's state to that considered normal. The technique of detecting a system's abnormal events or behaviors by comparing a system's run-time profile to its normal profile is called *anomaly detection*. Of course, the technique of anomaly detection can also be used to detect anomalies other than failures and malicious attacks. For example, a target's behavior change in a target-tracking system can also be detected by anomaly detection.

Although anomaly detection usually suffers from a high false alarm rate in traditional systems, the anomaly detection in WSNs is expected to perform well because the operations of WSNs are less dynamic in comparison to those in traditional counterpart systems like the Internet.

4.2 Packet Traffic in WSNs Serves as the Data Source of Anomaly Detection

Packet traffic has been the most used data source in the anomaly detection for WSNs. The authors of [51] propose that an anomaly in WSNs could violate one of the following rules applied to packet traffic:

1) Interval rule: A failure is raised if the time which passes between the reception of two consecutive messages is larger or smaller than the allowed limits.

2) Retransmission rule: The monitor listens to a message, pertaining to one of its neighbors as its next hop, and expects that this node will forward the received message, which does not happen.

3) Integrity rule: The message payload must be the same along the path from its origin to a destination, considering that in the retransmission process there is no data aggregation by other sensor nodes.

4) Delay rule: The retransmission of a message by a monitor's neighbor must occur before a defined timeout.

5) Repetition rule: The same message can be retransmitted by the same neighbor only a limited number of times.

6) Radio transmission range: All messages listened to by the monitor must have originated (previous hop) from one of its neighbors.

7) Jamming rule: The number of collisions associated with a message sent by the monitor must be lower than the expected number in the network.

By regularly monitoring the violations of the listed rules, network anomalies will be detected.

Ref. [13] proposes a specification-based anomaly detection to detect malicious attacks on AODV routing. In this approach, the authors use an FSM to specify correct AODV routing behavior and use distributed network monitors to detect run-time violation of the AODV specifications. The rationale behind this is that the AODV protocol has specified the sequence relations among different kinds of routing messages, and such sequence relations can be depicted by an FSM. Any violation of the protocol specification will trigger an alert. Ref. [15] also proposes a specification-based anomaly detection to detect routing attacks. In their approach, they use an FSM to specify the DSR routing behavior, instead of the AODV routing behavior.

In addition to the ability to specify the sequence relations among some special kinds of packets (e.g. routing messages) according to protocol specifications, the author of this article suggests that the sequence relations among general kinds of packets arriving at a sensor node can also be learned automatically by on-line training, thus anomalies can be detected by comparing the run-time traffic patterns with those prelearned normal traffic patterns [17]. If a “match” cannot be found between the runtime packet arriving sequence and any prelearned normal packet arriving sequence, the traffic profile has been violated and an alarm can be launched.

Ref. [52] uses another traffic feature instead of packet sequence relations. It records the arrival time of each observed packet and checks the mean and the standard deviation of the interarrival times of the packets in a long term receive buffer and a short term intrusion buffer. An arrival is considered anomalous if the statistics in these two buffers deviate significantly.

In [1], the author of this article suggests that the anomaly detection can be done in a sensor network for target tracking purpose, of which the traffic modeling results have already been presented in [11]. In [11], an ON/OFF model is used to capture the traffic bursts caused by intermittent target observations in a target-tracking sensor network. Each ON period indicates an event that the neighborhood of the considered sensor node has been visited by the target, while the length of an ON period provides information relating to the duration of a visit. An OFF period corresponds to the idle time period when there is no target observation. It has been found that both the ON and OFF period distributions are steady (i.e. not changing with time) if the target follows a certain random mobility model. Further, both the ON and OFF period distributions have tails which decrease near-exponentially as the period lengths increase. That means, an unusual long ON/OFF period can only be observed with an extremely low probability in the normal situation, and its runtime observation should trigger an anomaly alarm to receive special attention. An ON/OFF state transition diagram for anomaly detection is shown in Figure 2, where the length of the “anomaly ON/OFF timer” is set to be a probabilistic upper length limit.

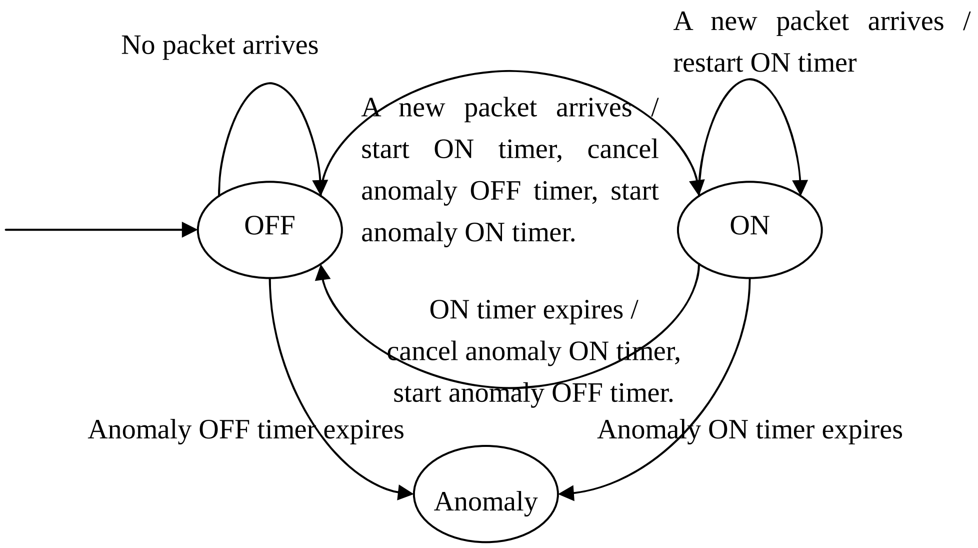
No packet arrives
A new packet arrives / restart ON timer
A new packet arrives / start ON timer, cancel anomaly OFF timer, start anomaly ON timer.
OFF
ON
ON timer expires / cancel anomaly ON timer, start anomaly OFF timer.
Anomaly OFF timer expires
Anomaly ON timer expires
Anomaly

Figure 2: ON/OFF state transition diagram with the function of anomaly detection

In [53, 54], the authors introduce a new data mining method that uses “cross-feature analysis” to capture the inter-feature correlation patterns in normal packet traffic, thus it can make decisions based on multiple traffic features. These patterns can be used as normal profiles to detect deviations (or anomalies) caused by malicious attacks and network failures. More specifically, this approach computes a classifier C_i for each feature f_i using $\{f_1, f_2, ..., f_{i-1}, f_{i+1}, ..., f_L\}$, where $\{f_1, f_2, ..., f_L\}$ is the feature set. C_i can be learned from a set of training data. It predicts the most likely value of f_i based on the values of other features. Based on a set of rules presented, this approach can identify the attack type of several well-known attacks. In some cases the rules can also identify the attacking or misbehaving nodes.

4.3 Evaluating Anomaly Detection Strategies for WSNs

The two commonly used measurements for evaluating the performance of an anomaly detection strategy are the false positive rate (FP) and the false negative rate (FN). FP is defined as the proportion of normal events that are erroneously classified as abnormal. FN is defined as the proportion of abnormal events that are erroneously classified as normal. Obviously, a good anomaly detection strategy should have both a low FP and a low FN. However, a tradeoff is usually to be made between FP and FN, given that these two measurements are usually influenced in opposing ways, by adjusting the threshold parameters used in many anomaly detection strategies. In addition to FP and FN, the overhead introduced by an anomaly detection strategy is also a concern. Considering the extreme resource-constrained specialties of WSNs, a good anomaly detection strategy should introduce as little overhead as possible. Although WSNs are designed for low rate communication, a broad range of real-time applications, such as health care, highway traffic coordination and even multimedia transmission have also been proposed. When an anomaly detection strategy is designed for real-time applications, it should also fulfill the real-time requirement such that it will not cause performance degradation to the applications.

5. Conclusions & Open Issues

WSNs have been identified as one of the most important technologies for the 21st century. In this article, the author has provided a survey of the current works involved in traffic analysis & modeling, network optimization and network anomaly detection for WSNs. Through the presentation of this article, the readers can see that many of the works involved in network optimization and anomaly detection are based on the research results from traffic analysis & modeling. Actually, network traffic and its associated energy consumption play a key role in most of the works relating to network optimization for WSNs. It is also shown that detecting sensor network anomalies through the analysis of network traffic is technically feasible.

As WSNs are still a young research field, much activity is still on-going in order to solve many open issues. For example, traffic dynamics in WSNs are application dependent. For many WSN application scenarios, the traffic dynamics are still very obscure. Network optimization continues to be the prime important research area for WSNs given the constraint of the very limited resources which are unable to be removed in the near future. As more and more WSNs become available for practical deployment, the problems relating to sensor failures and malicious attacks will attract more and more attention. Anomaly detection, which is a promising technique for the immediate detection of any network anomaly such as sensor failure and malicious attack, has as yet been touched upon only rarely.

In the future, traffic analysis & modeling for WSNs should focus on those event-driven WSN scenarios because traffic dynamics in event-driven WSNs are much more uncertain than those in periodic data generation WSNs. Further, as node mobility has been utilized in a few WSN applications such as healthcare monitoring, it will be useful to investigate the traffic dynamics in WSNs when there is node mobility. In-network processing has been viewed as an essential method to reduce and balance the energy consumption within WSNs. Because in-network processing eliminates the need to transmit raw data to a central point, it also changes those familiar traffic patterns in WSNs. Investigating traffic dynamics in WSNs, when different in-network processing strategies are applied, will be very necessary.

In the future, network optimization for WSNs will continue to be of prime importance given the inherent nature of limited resources. For those WSNs with node mobility and in-network processing, the fundamental performance bounds are still not clear. More network optimization models could be built to investigate the fundamental performance bounds of such WSNs, and the provision of accurate traffic models will be a pre-condition for this option. For those well-investigated WSNs without mobility and in-network processing, the optimal performances which are achievable by centralized coordination algorithms are already known. The research focus should shift to the development of distributed coordination algorithms which are more practical in a real implementation. As to the optimal resource allocation for WSNs, there have already been schemes developed for simple WSN scenarios. The development of more resource allocation schemes for more general WSN scenarios should be very useful.

In the future, anomaly detection for WSNs will become more and more important as more and more WSNs become available for real deployment. This article argues that packet traffic is a good source for anomaly detection in WSNs. This argument requires more support in the future. As malicious attacks will be low probability events in many WSNs, high false alarm rates are not tolerable in these WSNs. Designing an anomaly detection system with an extremely low false alarm rate will be a challenge. After a network anomaly is detected, either a person is required to be sent to the identified problem region or the network must take some measures to automatically recover from the possible damage. The development of such accompanying emergency response strategies will be necessary for future anomaly detection in WSNs.

References

[1] Wang, Q. (2010). *Traffic Analysis, Modeling and Their Applications in Energy-Constrained Wireless Sensor Networks - On Network Optimization and Anomaly Detection*. PhD thesis, No. 78, Dept. of Information Technology and Media, Mid Seden University, Sundsvall, Sweden.

[2] Doherty, L., Warneke, B. A., Boser, B. E., and Pister, K. S. J. (2001). Energy and performance considerations for smart dust. *International Journal of Parallel and Distributed Systems and Networks*, 4(3):121--133.

[3] Pottie, G. J. and Kaiser, W. J. (2000). Wireless integrated network sensors. *Communications of the ACM*, 43(5):51--58.

[4] Wang, Q. and Zhang, T. (2009). A survey on security in wireless sensor networks. In Zhang, Y. and Kitsos, P., editors, *Security in RFID and Sensor Networks*, chapter 14, pages 293--320. CRC Press, Taylor & Francis Group.

[5] Demirkol, I., Alagoz, F., Delic, H., and Ersoy, C. (2006). Wireless sensor networks for intrusion detection: Packet traffic modeling. *IEEE Communications Letters*, 10(1):22--24.

[6] Cui, S., Madan, R., Goldsmith, A. J., and Lall, S. (2005). Joint routing, mac, and link

layer optimization in sensor networks with energy constraints. In *Proc. of IEEE International Conference on Communications (ICC'05)*, pages 725--729.

[7] Ma, Y. and Aylor, J. H. (2004). System lifetime optimization for heterogeneous sensor networks with a hub-spoke topology. *IEEE Transactions on Mobile Computing*, 3(3):286--294.

[8] Tang, S. (2006). An analytical traffic flow model for cluster-based wireless sensor networks. In *Proc. of 1st International Symposium on Wireless Pervasive Computing*.

[9] Paxson, V. and Floyd, S. (1995). Wide-area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3:226--244.

[10] Øverby, H. and Stol, N. (2004). Effects of bursty traffic in service differentiated optical packet switched networks. *Optics Express*, 12(3):410--415.

[11] Wang, Q. and Zhang, T. (2008). Source traffic modeling in wireless sensor networks for target tracking. In *Proc. of the 5th ACM International Symposium on Performance Evaluation of Wireless Ad-Hoc, Sensor, and Ubiquitous Networks (PE-WASUN'08)*, pages 96--100.

[12] Wang, P. and Akyildiz, I. F. (2009). Spatial correlation and mobility aware traffic modeling for wireless sensor networks. In *Proc. of IEEE Global Communications Conference (Globecom'09)*.

[13] Tseng, C., Balasubramanyam, P., Ko, C., Limprasittiporn, R., Rowe, J., and Levitt, K. (2003). A specification-based intrusion detection system for AODV. In *Proc. of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks*.

[14] Perkins, C. E. and Royer, E. M. (1999). Ad-hoc on-demand distance vector routing. In *Proc. of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'99)*, pages 90--100.

[15] Yi, P., Jiang, Y., Zhong, Y., and Zhang, S. (2005). Distributed intrusion detection for mobile ad hoc networks. In *Proc. of the 2005 Symposium on Applications and the Internet Workshops (SAINT-W'05)*.

[16] Johnson, D. B., Maltz, D. A., and Broch, J. (2001). DSR: The dynamic source routing protocol for multi-hop wireless ad hoc networks. In Perkins, C. E., editor, *Ad Hoc Networking*, chapter 5, pages 139--172. Addison-Wesley.

[17] Wang, Q. and Zhang, T. (2007). Detecting anomaly node behavior in wireless sensor networks. In *Proc. of the 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, pages 451--456.

[18] Noori, M. and Ardakani, M. (2008). Characterizing the traffic distribution in linear wireless sensor networks. *IEEE Communications Letters*, 12(8):554--556.

[19] Subramanian, R. and Fekri, F. (2006). Sleep scheduling and lifetime maximization in sensor networks: Fundamental limits and optimal solutions. In *Proc. of the 5th International Conference on Information Processing in Sensor Networks (IPSN'06)*, pages 218--225.

[20] Wang, Q. and Zhang, T. (2009). Characterizing the traffic load distribution in dense sensor networks. In *Proc. of the 2nd International Workshop on Wireless Sensor Networks: theory and practice (WSN'09).*

[21] Wang, Q. and Zhang, T. (2010). Fair energy allocation in wireless sensor networks: theory and practice. In *Prof. of IEEE Global Communications Conference (Globecom'10).* Submitted.

[22] Ren, H., Meng, M. Q.-H., and Chen, X. (2006). Investigating network optimization approaches in wireless sensor networks. In *Proc. of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2015--2021.

[23] Toh, C. K. (2001). Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks. *IEEE Communications Magazine*, 39:138--147.

[24] Chang, J.-H. and Tassiulas, L. (2004). Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Transactions on Networking*, 12:609--619.

[25] Singh, S., Woo, M., and Raghavendra, C. S. (1998). Power-aware routing in mobile ad hoc networks. In *Proc. of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98)*, pages 181--190.

[26] Zhang, Y., Ramkumar, M., and Memon, N. (2004). Information flow based routing algorithms for wireless sensor networks. In *Proc. of IEEE Global Telecommunications Conference (Globecom'04)*, pages 742--747.

[27] Chang, J.-H. and Tassiulas, L. (2000). Energy conserving routing in wireless ad-hoc networks. In *Proc. of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies ({INFOCOM}'00)*, pages 22--31.

[28] Chang, J.-H. and Tassiulas, L. (1999). Routing for maximum system lifetime in wireless ad-hoc networks. In *Proc. of the 37th Annual Allerton Conference on Communication, Control, and Computing*.

[29] Alonso, J., Dunkels, A., and Voigt, T. (2004). Bounds on the energy consumption of routings in wireless sensor networks. In *Proc. of the 2nd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'04).*

[30] Wang, Q., Zhang, T., and Pettersson, S. (2007). Bounding the information collection performance of wireless sensor network routing. In *Proc. of the 5th Annual Conf. on Communication Networks and Services Research (CNSR'07)*, pages 55--62.

[31] Wang, Q., Zhang, T., and Pettersson, S. (2008). An effort to understand the optimal routing performance in wireless sensor network. In *Proc. of the IEEE 22nd Int. Conf. on Advanced Information Networking and Applications (AINA'08)*, pages 279--286.

[32] Ordonez, F. and Krishnamachari, B. (2004). Optimal information extraction in energy-limited wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 22:1121--1129.

[33] Shih, E., Cho, S., Ickes, N., Min, R., Sinha, A., Wang, A., and Chandrakasan, A. (2001). Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. In *Proc. of the 7th Annual ACM International Conference on Mobile Computing and Networking (MobiCom'01)*, pages 272--286.

[34] van Dam, T. and Langendoen, K. (2003). An adaptive energy-efficient mac protocol for wireless sensor networks. In *Proc. of the International Conference on Embedded Networked Sensor Systems (SenSys'03)*, pages 171--180.

[35] Singh, S. and Raghavendra, C. (1998). Pamas: Power aware multi-access protocol with signalling for ad hoc networks. *ACM SIGCOMM Computer Communication Review*, 28(3):5--26.

[36] Ye, W., Heidemann, J., and Estrin, D. (2002). An energy-efficient mac protocol for wireless sensor networks. In *Proc. of the 21st International Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, pages 1567--1576.

[37] Rabbat, M. and Nowak, R. (2004). Distributed optimization in sensor networks. In *Proc. of the 3rd International Symposium on Information Processing in Sensor Networks*, pages 20--27.

[38] Heinzelman, W. R., Chandrakasan, A., and Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *Proc. of the 33rd Hawaii International Conference on System Sciences*.

[39] Lindsey, S., Raghavendra, C., and Sivalingam, K. M. (2002). Data gathering algorithms in sensor networks using energy metrics. *IEEE Transactions on Parallel and Distributed Systems*, 13(9):924--935.

[40] Huang, S.-C. and Jan, R.-H. (2004). Energy-aware, load balanced routing schemes for sensor networks. In *Proc. of the 10th International Conference on Parallel and Distributed Systems (ICPADS'04)*, pages 419--425.

[41] Teixeira, I., de~Rezende, J.~F., and de~Castro P.~Pedroza, A. (2004). Wireless sensor network: Improving the network energy consumption. In *Proc. of the XXI Simposio Brasileiro de Telecommunicacoes (SBT'04).*

[42] Hyytiä, E. and Virtamo, J. (2007). On traffic load distribution and load balancing in dense wireless multihop networks. *EURASIP Journal on Wireless Communications and Networking*. Article ID 16932.

[43] Popa, L., Rostamizadeh, A., Karp, R. M., Papadimitriou, C., and Stoica, I. (2007). Balancing traffic load in wireless networks with curveball routing. In *Proc. of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc'07)*, pages 170--179.

[44] Efthymiou, C., Nikoletseas, S., and Jose, R. (2006). Energy balanced data propagation in wireless sensor networks. *Wireless Networks*, 12(6):691--707.

[45] Tang, S. and Li, W. (2006). Qos supporting and optimal energy allocation for a cluster based wireless sensor network. *Elsevier Computer Communications*, 29(13-14):2569--2577.

[46] Wang, Q. and Zhang, T. (2009). Bottleneck zone analysis in energy-constrained wireless sensor networks. *IEEE Communications Letters*, 13(6):423--425.

[47] Gao, Q., Blow, K. J., Holding, D. J., Marshall, I. W., and Peng, X. H. (2006). Radio range adjustment for energy efficient wireless sensor networks. *Ad Hoc Networks Journal*, 4(1):75--82.

[48] Song, C., Liu, M., Cao, J., Zheng, Y., Gong, H., and Chen, G. (2009). Maximizing network lifetime based on transmission range adjustment in wireless sensor networks. *Elsevier Computer Communications*, 32(11):1316--1325.

[49] Rivas, H., Voigt, T., and Dunkels, A. (2006). A simple and efficient method to mitigate the hot spot problem in wireless sensor networks. In *Proc. of Performance Control in Wireless Sensor Networks*.

[50] Chang, C.-Y., Shih, K.-P., Chang, H.-R., and Liu, H.-J. (2006). Energy-balanced deployment and topology control for wireless sensor networks. In *Proc. of IEEE Global Telecommunications Conference (Globecom'06)*, pages 1--5.

[51] da Silva, A. P., Martins, M., Rocha, B., Loureiro, A., Ruiz, L., and Wong, H. (2005). Decentralized intrusion detection in wireless sensor networks. In *Proc. of the 1st ACM International Workshop on Quality of Service & Security in Wireless and Mobile Networks (Q2SWinet'05)*, pages 16--23.

[52] Onat, I. and Miri, A. (2005). A real-time node-based traffic anomaly detection algorithm for wireless sensor networks. In *Proc. of Systems Communications*, pages 422--427.

[53] Huang, Y., Fan, W., Lee, W., and Yu, P. S. (2003). Cross-feature analysis for detecting ad-hoc routing anomalies. In *Proc. of the 23rd International Conference on Distributed Computing Systems*.

[54] Huang, Y. and Lee, W. (2003). A cooperative intrusion detection system for ad hoc networks. In *Proc. of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks*, pages 135--147.