Optimization of intrusion detection systems for wireless sensor networks using evolutionary algorithms

Martin Stehlík
xstehl2@fi.muni.cz
Faculty of Informatics
Masaryk University
Brno, Czech Republic

Abstract
Wireless sensor networks (WSNs) are a type of ad hoc wireless networks with several specifics. WSNs consist of sensor nodes that monitor some environment and send measurements hop-by-hop to a base station (BS). Sensor nodes are low-cost devices highly restricted in their resources regarding to the size of memory, the microcontroller performance and last but not least the energy supply limited by the battery capacity. The concept of the WSNs opens an area for novel attacks. Active attacks, where the attacker manipulates communication, can be detected by an intrusion detection system (IDS) running on the sensor nodes in a distributive and collaborative manner. Since the IDS necessary brings an overhead to the computation and wireless communication of the sensor nodes, it should be deployed with respect to the restricted WSNs. Our goal is to optimize IDS parameters towards solutions suitable for various scenarios and to deal with trade-off between IDS accuracy and WSN performance and lifetime. We assume that a network operator of the WSN will be able to set IDS parameters according to current requirements. In this work, we show how multi-objective evolutionary algorithms can provide optimized solutions and assist with IDS settings.

Keywords: intrusion detection system, IDS, wireless sensor network, WSN, simulator, optimization, evolutionary algorithm.

1 Introduction
Recent advances in wireless communications and low-cost electronic devices enabled the development of low-cost and high-performance wireless networking technologies. Apart from widely used cellular wireless networks known from mobile phones and wireless local area networks (WLAN), there are also ad hoc wireless networks operating without any given and fixed infrastructure, where the connections are established on demand in ad hoc manner [1].

Wireless sensor networks (WSNs) can be considered as a type of ad hoc wireless networks with many specifics. The main difference against the “ordinary” ad hoc wireless networks is that the WSNs consist of a large number of usually homogeneous and resource restricted low-cost sensor nodes (also called motes). Their goal is to monitor some area and transfer the values of measurements of physical parameters like temperature, humidity, intensity of light or movement detection to a base station (BS; also called a sink) for further processing. However, the sensing and communication tasks have to be performed by sensor nodes that collaboratively build an ad hoc infrastructure. The sensor nodes can also pre-process the collected data.
Various examples of applications of WSNs can be shown as emergency operations, battlefield monitoring in a military, habitat monitoring, precision agriculture, home automation, health care, logistics and environmental applications, such as forest fire and flood detection.

However, since WSNs are often deployed in physically open and sometimes even hostile environments, they can be subjects of various security attacks. An adversary may be able to eavesdrop the communication, insert a malicious node into the network or to capture a node to reveal sensitive information. We believe that such malicious activities necessary bring anomalous behavior that can be observed by an intrusion detection system (IDS).

To enable intrusion detection even in the remote parts of the network, the intrusion detection agents should be deployed in all parts of the network to monitor the malicious events locally. Thus, they are assumed to run on the ordinary sensor nodes [2, 3]. Nevertheless, the sensor nodes are limited in their resources. To give an example of sensor node capabilities, MICAz – a typical sensor node – is equipped with the 8 MHz Atmel Atmega128L microcontroller, 4 kB RAM, 512 KB flash memory, 802.15.4 compliant Texas Instruments CC2420 transceiver and two AA batteries.

There is a trade-off between the IDS accuracy and the WSN performance because of the mentioned limitations. WSNs feature many aspects that are subject to optimization, including the IDSs. An IDS optimization framework [4] was proposed to optimize IDS parameters and we aim to extend the framework towards the design of robust solutions that enable detection of more complex attacks.

The text above, as well as some parts of the following text mainly comes out from the PhD thesis proposal [5].

2 Optimization framework for WSNs

There exist several active attacks on WSNs that can disrupt the functionality of the network. The attacker can alter, drop, selectively forward, replay and inject packets or jam legitimate transmissions. Several techniques to detect these attacks have been proposed so far. However, the detection techniques usually do not consider optimization of their parameters with respect to IDS accuracy and with WSN performance and lifetime.

We implemented a framework incorporating a simulator and an optimization engine that utilizes evolutionary algorithms [4]. New solutions (IDS configurations) are generated in evolution process and provided to the simulator for evaluation. Based on the evaluation (simulation statistics) in the simulator, new solutions are proposed by the optimization engine. This process continues iteratively until some stopping criterion is fulfilled.

2.1 Simulator

Simulators are highly used by WSN research community. They enable fast evaluation of proposed protocols or techniques and can speed-up the real time – WSNs running for months can be simulated in minutes or hours (depending on the complexity of the simulator). Last but not least, the results can be verified and the experiments can be repeated. In the past, we compared several simulators thoroughly with the reality [6]. Currently, we use MiXiM framework [7] based on OMNeT++ in our on-going research for evaluation of the IDS.

2.2 Evolutionary algorithms

Since search space of possible IDS configurations is very large, various metaheuristics can be used to find a good approximation in a relatively short time. As is shown in [4], evolutionary algorithms showed to be
a suitable tool and were able to find even optimal solutions. However, the disadvantage is that they produce a single solution based on weights given to each of the objectives. If the network operator changes his/her priorities, the optimization process should run again. Hence, multi-objective approach brings an advantage in such a way that the network operator can switch between different optimized settings according to the current requirements.

2.3 Multi-objective evolutionary algorithms

Multi-objective evolutionary algorithms (MOEAs) can be useful for optimization of the IDSs and other aspects in the WSNs because they can provide the network operators with a set of non-dominated solutions. Then the network operator can choose between, e.g., an optimized solution A with better IDS accuracy at the cost of higher energy consumption and another optimized solution B with lower energy consumption at the cost of worse IDS accuracy. The set of non-dominated optimal solutions is called Pareto front and the goal of the MOEAs is to find a good approximation of the true Pareto front.

3 Considered aspects

The main goal is to provide a working IDS framework that allows the network operators to design intrusion detection solutions for their WSNs based on the purpose of the network, area of deployment, network stack and security issues. The IDS framework will provide detection techniques for specified attacks and optimize their parameters. Optionally, the IDS framework could also assist the operators in design and configuration of the network stack for given scenarios based on provided models.

3.1 Attacks and detection techniques

The network operator considers the vulnerabilities and specifies the attacks exploitable by an attacker in order to compromise the network. We focus on selective forwarding, delay and modification attacks, where the exchanged packets are monitored and thus storage overhead is generated. The detection techniques can be optimized to obtain sufficient accuracy at the cost of reasonable consumption of the resources. Another considered attack is jamming, where parameters such as carrier sensing time (time of waiting for the channel to become idle) or numbers of retransmissions are monitored and corresponding thresholds can be optimized. Last but not least, considered attack is Sybil attack, where the attacker’s node changes its identity. The framework is extendable to cover other techniques and to optimize them.

3.2 Environment

WSNs can be utilized in specific types of environment ranging from open clear spaces to complicated indoor spaces divided into subareas by obstacles with different throughput of the radio signal. Hence the solution should be optimized for a specific environment, yet robust enough for the deviations caused by changes of the wireless channel. Several models have been proposed but log-normal shadowing is the widely used conventional one implemented across the simulators as we investigated in [6]. In our work, the environment is simulated primarily using the log-normal shadowing model. The parameters of the model are provided by the network operator and the robustness of the found solution should be validated (or ensured) mainly by changing of the path loss exponent with adjacent values within a specified interval. The robustness can be maintained either during the optimization at the cost of higher computational complexity or afterwards at the risk of inappropriate behaviour in case of slightly different environmental parameters.
3.3 Network stack

The performance of individual nodes including the IDS and the efficiency of the overall WSN is influenced by the protocols used at different layers. Application runs at the highest level and depends on the purpose of the WSN. Regarding the IDS, the application influences particularly the number of packets that are sent by a corresponding node and the types of the packets. For the IDS framework, the application provides input parameters for optimization of the lower level protocols and of the IDS. The selection of a protocol at the network layer influences the routing and adaptation on changes in nodes’ topology. Regarding the IDS, the network layer provides a list of parent nodes that a packet should be forwarded to. Protocols at the link layer (MAC) maintain connections between nodes and observe the channel. The overheard communication providing information to the IDS is captured at this layer. MAC protocols are also responsible for the collision avoidance, retransmissions and duty cycles. Their parameters are also subject of optimization. Measurements at the physical layer provide information about the emitted signal that can be utilized to detect, e.g., jamming or hello flood attack. We investigate options for configurations of the whole network stack using our optimization framework.

4 Results and conclusion

We explored the possibilities of multi-objective evolutionary algorithms for optimization of IDS parameters. Our recent research covered optimization of an intrusion detection technique for selective forwarding attack. In our technique, IDS nodes buffer packets that are supposed to be forwarded by their neighbours. Size of the buffer and number of monitored nodes improve IDS accuracy but require higher amount of memory. The higher amount of memory is consumed, the better IDS accuracy is reached. Thus, we had memory, number of false positives and number of false negatives as objectives. The goal is to minimize all of them.

We were able to found solutions fairly spread on the Pareto front (that we computed in exhaustive search) using approx. 5,000 simulation evaluations that is much more time efficient comparing to evaluation of 25,000,000 simulations needed in a case of exploring the objective space using exhaustive search. Detailed results are submitted for publication at European Conference on Artificial Life.

We showed that multi-objective approach is beneficial since it provides the network operator with a set of optimized solutions. He/she can select a solution according to the purpose of the WSN and change the IDS settings according to current needs. However, the computation using exhaustive search is extremely time-demanding and mostly infeasible for large networks. MOEAs proved to be a good compromise between the quality of Pareto front approximation and the optimization time.

References


