Analysis of Cikhaj Experiment 2013

PA018 Term project report

Author: Ondřej Koutský, 359295

1 Cikhaj Experiment description

Cikhajj Experiment is an experiment which was made in scope of a project for development of security software platform for wireless sensor networks. The purpose of this experiment was to get as much data as possible about behavior of real nodes with ProtectLayer[1] application uploaded and enabled. Main goal was to field-test intermediate version of prototype application.

The basic scenario of the experiment is to build a wireless sensor network that would serve for a simulation of network for motion detection.

1.1 Experiment settings

Network for the experiment was composed of 29 wireless sensor nodes. Each node had more application components deployed and running: user application, ProtectLayer application, Logger (serves as IDS) and application for simulation of motion detection. Nodes were physically organized into 3 concentric circles in whose center was special node called Base Station (BS) located. Graphical visualization of the network is shown in Figure 1.

Duration of whole experiment was approximately 65 minutes. In the first 60 minutes the network was calm and in the remaining time there were 4 attacks simulated (description of the attacks in subsection 1.2).

Task of each node was to inform BS with 2 types messages:

- *"Still alive"* message (SA) – generated regulary every 5 seconds informing BS that the node is still actively connected to the network
- *"Movement detected"* message (MD) – generated once whenever the node detects attacker's movement

Nodes do not address their messages to the BS directly. Both SA and MD messages were delivered along a fixed static routing tree whose structure is indicated by arrows in Figure 1.

The last functionality implemented into nodes was logging network information. Every node was required to store every received (either addressed for this or different node) message into EEPROM memory.

[1] ProtectLayer is application which is being developed at Laboratory of Security and Applied Cryptography at Faculty of Informatics, Masaryk University. Application's purpose is to provide basic security features for wireless sensor networks: encryption of transmitted messages, basic IDS (Intrusion Detection System), network nodes authentication, etc.

Besides the regular nodes there were 4 additional nodes placed in the network. These nodes are called sniffers and their purpose was to monitor network traffic in promiscuous mode in order to provide additional debugging information.

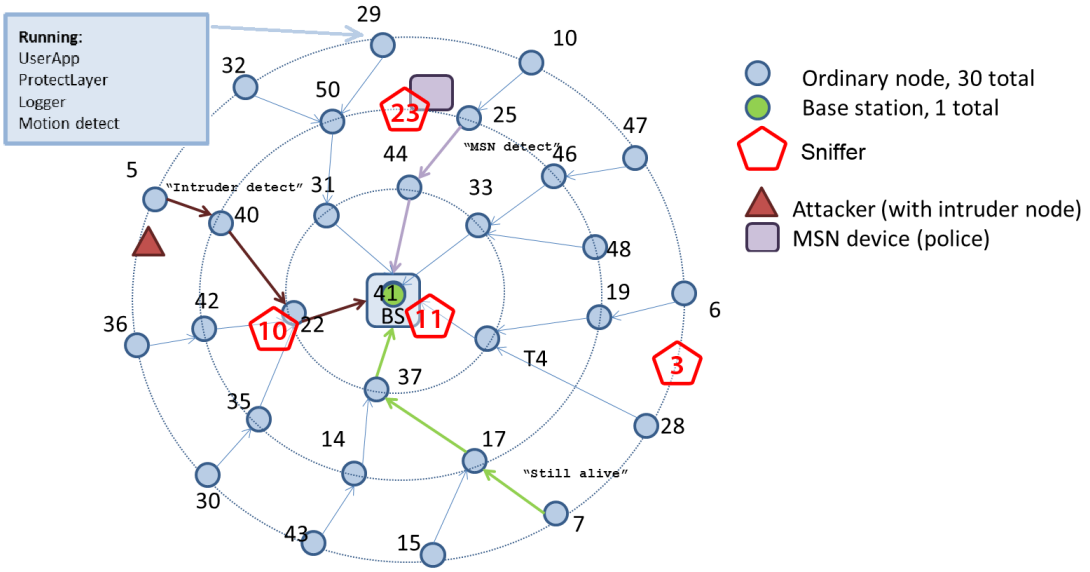

Figure 1: Map of spatial distribution of nodes and sniffers for the experiment.

1.2 Intrusion scenarios

Experiment was designed as a simulation of sensor network for motion detection. Attacker movement detection was simulated by radio proximity. Whenever any node in the network captured a signal produced by an attacker, it informed BS by sending MD message.

During the experiment there were 4 attacks simulated. Each attack differs in direction and speed of attacker's movement. All of the intrusion scenarios are described in Figure 2.

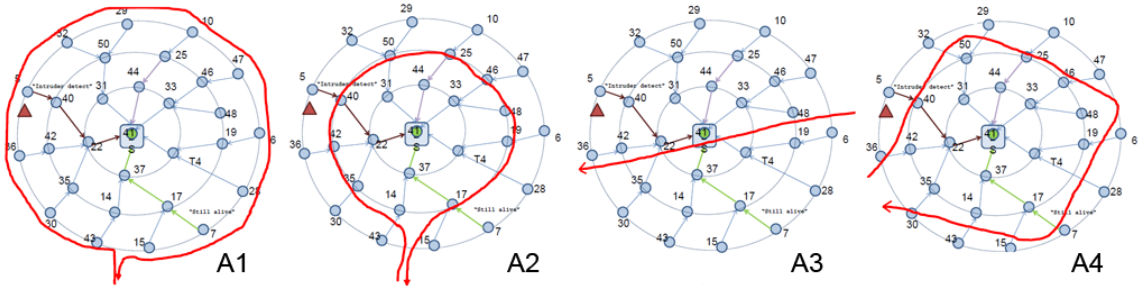

Figure 2: Description of attackers movement in space. **A1 (Attack 1)**: outer circle, slow movement. **A2 (Attack 2)**: inner circle, fast run. **A3 (Attack 3)**: fast run through the middle. **A4 (Attack 4)**: fast run in the square pattern.

2 Term project goals

The goal of this term project was to perform data analysis on the set of logs from single nodes and from the sniffers. The analysis was performed in 3 steps. First step in the analysis was a design of suitable data structure for storing logging information and its implementation. The structure allows easy searching over stored data and provide data for statistical analysis. The second steps was the analysis on retrieved data and the last step covers interpretation of the results of this analysis.

2.1 Expected outputs of analysis

There are 3 areas of expected outputs of performed data analysis:

1. Graphs of packet loss

   Based on information from nodes' logging information fraction of SA and MD message packets delivered and lost during the routing process should be inspected. Visualization of processed data in graphs of packet loss (on the first and the second hop in the routing tree and on the BS level) can locate possible communication bottlenecks in the network.

2. Description of what section was captured by sniffers

   Sniffers are important part of the experiment. The main purpose for their usage is to determine what communication is actually transfered via network.

3. Description of attacker's movement in time

   How many nodes detected him? Did it happen that attacker wasn't detected (lost of a packet)? How quickly and reliably was attacker detected?

3 Analysis results

In following paragraphs analysis results will be provided and discussed. Beginning of this section describes problems that occurred during the data processing and in the remaining parts data analysis results are presented.

3.1 Encountered problems

During the process of data analyzing a number of different problems appeared. Dealing with these problems helped to point out few implementation and deployment weaknesses in experiment settings. During the data processing, some of these problems were easy to bypass; However, some other problems shown to be critical for accuracy of the analysis and their solution is not so straightforward.

Basically, the encountered problems can be divided into 2 categories:

1. Implementation and deployment problems:

   - Initial rubbish in memory – At the beginning of each node's log there are few ($\pm$ 10) logs that do not follow appropriate structure and is obvious that they should not be considered as valid data.
   - Anomalous behavior of few nodes:

(a) Node 40 – BS log: 0 messages initiated by this node. Sniffers log: 3 messages.

(b) Node 5 – Node 40 log: 0 messages initiated by this node. Sniffers log: out of reach.

(c) Node 6 – Node 23 log: 0 messages. Sniffers log: 679 messages (sniffer 23 is physically very close).

(d) Node 7 – Node 17 log: 2 messages. Sniffers log: 8 messages.

(e) Node 29 – Node 50 log: 1 message. Sniffers log: 1 message.

(f) Node 32 – Node 50 log: 4 messages. Sniffers log: 1 message.

(g) Node 19 – Node 4 log: 6 messages. Sniffer 10 log: 367 messages. Other sniffers: 0–3 messages.

(h) Node 10 – Node 10 log: 0 messages. Sniffers log: 0 messages. There is no way how to determine, whether this node was even active.

Position of all nodes with anomalous behavior is visualized in graph in Figure 3.

Anomalous behavior of nodes listed above is probably caused either by high rate of interferences with another node transmitting in the same time or by inappropriate rotation of node's anthena.

Data from these nodes' logs are not included in following analysis.

Problems included in this category cause only partial inaccuracies and if omitted, it does not influence global characteristics of the network. However, it is impossible to recover missing data from corrupted nodes and behavior of these can not be inspected.

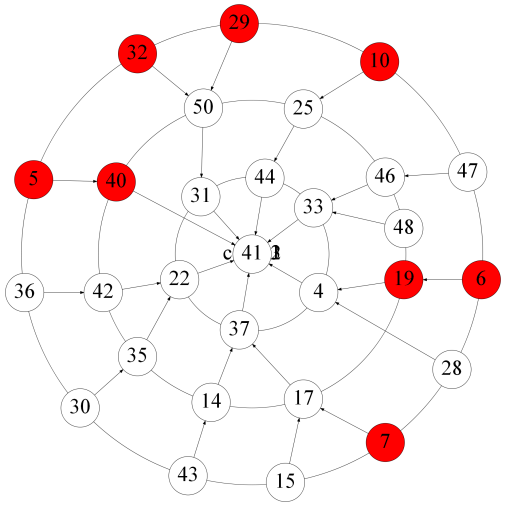

Figure 3: Graph showing position of nodes with anomalous behavior (red).

2. Overladed nodes

It turned out during the analysis of data from nodes' logs that most of the nodes in the network were getting overloaded during the experiment. Nodes' internal CPUs were probably unable deal with both sending messages and routing them and with logging received messages.

When comparing logging information from nodes and from sniffers it is possible to estimate, that approximately 62.5 % of received messages is being properly logged only.

Due to the problem with overloaded nodes would using of their logs for analyzing network characteristics lead to very distorted results. There is therefore need to try to reconstruct as much network traffic from the sniffers as possible and process analysis over this data.

3.2 Area covered by sniffer nodes

As it has already been explained in subsection 3.1, logging information from sniffer nodes shown to be crucial for proper analyzing network characteristics. It is therefore important to determine what area of network and what nodes were covered within reach of which sniffer.

Logging data from each sniffer was processed separately at first. Based on files containing logged network traffic it was analyzed, how many different messages were capture from each node. Numerical results were visualized as graphs similar to one in Figure 4. Nodes within reach of given sniffer (sniffer node is filled with green color) are filled with color. It holds that the darker the filling color is, the more messages coming from given node were captured. Graphs showing data from independent nodes are not presented in this report; However, they were processed and can be found at [1].

More valuable information about the area covered by sniffer nodes provides graph in Figure 4. Data it contains were collected from all sniffers together with duplicate logs removed (same messages could be captured by more sniffers). Data coming from this analy-sis were later used for analyzing fractions of packets loss and also for describing attackers movement in time.

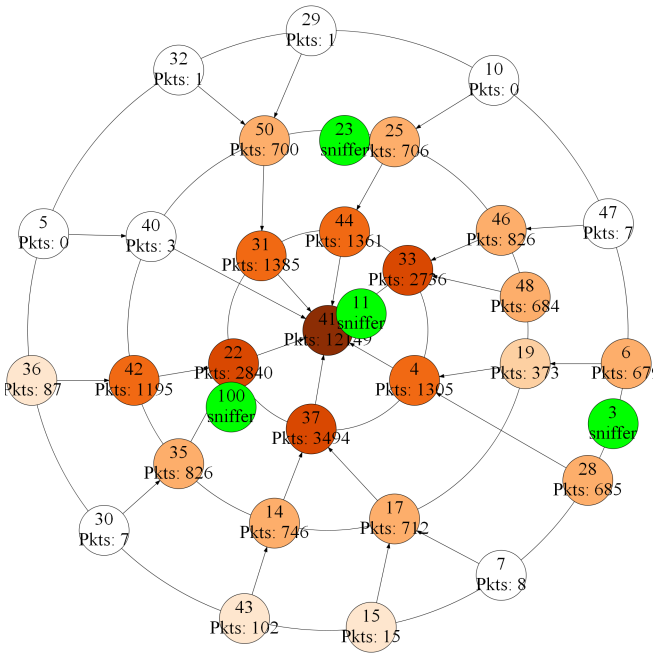

Figure 4: Description of what section of network was captured by sniffers. Graph contains data collected from all sniffers with duplicate logs removed.

Coloring of nodes in graph in Figure 4 indicates, that sniffers do not cover whole are of network. Most nodes situated on the outer ring are filled with either white color or with the color of light tint which means, that they are not located within reach of any sniffer.

However, from the perspective of following analysis the important fact is, that the area of nodes which are closer to the BS node is covered properly. It is therefore possible to determine how many messages were successfully delivered to the BS and what fraction of them was lost during the transmission or routing process.

3.3 Analysis of packet loss

In section 3.1 problems connected with analyzing data from nodes' logs were discussed. Nodes do not have enough computing power for both sending messages and storing received packets. Approximately 62% of all received messages only are properly logged. Necessary step towards retrieving at least some reliable data about packet loss in the network was therefore to focus on analyzing data from sniffer nodes.

However, neither using sniffers' logging data is ultimate solution for analyzing all of the network characteristics. The area of nodes which were sufficiently covered by sniffer nodes is relatively small. In order to retrieve reliable data it is necessary to analyze communication which was transmitted between BS and nodes situated at the inner circle of the network only. Less reliable but still meaningful and quit accurate data were retrieved from communication between inner and middle ring of network nodes.

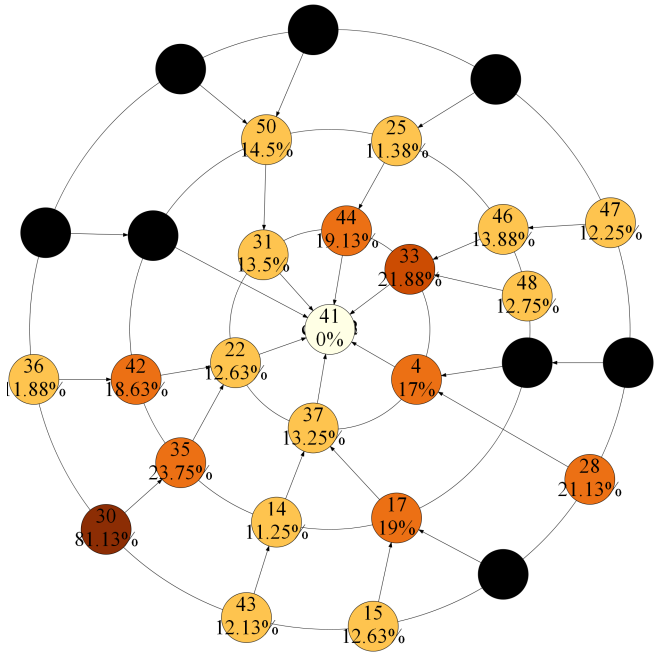

Figure 5: Graph showing fractions of packet loss on each node.

First provided output of analysis of packet loss is a graph shown in Figure 5. Each node in a graph is labeled with number containing an information about what fraction of messages initiated by this node were successfully delivered into BS. This number was computed as a ratio of messages logged by sniffers and a value of counter of last delivered message. This counter has value 800 and corresponds properly to the number

of messages expected to occur during the experiment (SA messages are transmitted 12 times per minute and the running time of the experiment was cca. 65 minutes).

Nodes in the graph in Figure 5 are filled with different colors. The darker the tint of filling color is, the higher ratio of lost packets is associated with given node is. Nodes colored in black are those connected with anomalous behavior and are excluded from packet loss analysis (see section 3.1).

Provided graph shows, that median value of packet loss is approximately 15% and that the vaules are distrubuted over nodes quit uniformly (except node 30).

Second provided output of analysis of packet loss is a graph shown in Figure 6. There is inner circle of network nodes with BS visualized. Each node is labeled with number informing about what fraction packets was not routed and delivered by given node.

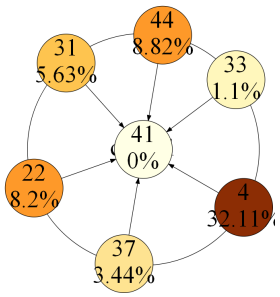

Figure 6: Graph showing fractions of packet loss on nodes situated in the inner ring.

The highest fraction of packets were lost at node 4. This result is unexpected since the node is responsible for routing packets from one other node only.

3.4 Detection of attacker's movement in time

Nodes in the network are not equipped with synchronized clock and it is therefore impossible to determine the exact time of MD message delivery. The only analysis we are able to process is thus comparing number of MD messages successfully delivered to BS with number of messages emitted by the attacker.

- Number of emitted messages: 214
- Number of MD messages received at the BS: 99
- Fraction of delivered messages: 46.3%

The results of the analysis show that more than a half MD messages were not delivered properly.

4 Suggested improvements for future experiments

Target of performed data analysis was not only to provide numerical and graphical results, but also based on these result try to localize weak points in experiment settings and implementation. In the following section few improvements for future re-running the experiment are suggested and described.

1. Introduction of internal clock into the nodes application.

   It is not necessary to completely synchronize these clocks in all of network nodes. Supplying logging information with proper timestamp should give accurate information about experiment duration (helps to determine expected number of messages transmitted over the network) and for example say what is the average time of transmission of individual messages into BS (determines a speed of informing BS about moving attacker).

2. Logging both sent and received messages.

   Introducing this functionality would lead much easier and much more accurate analysis of network characteristics. When all messages sent by each node were logged in this experiment, analysis of for example packet loss would be more accurate.

3. Deal with overloaded nodes problem.

   Solving overloaded nodes problem is crucial for future re-running of the experiment and for the following analysis. Without all information properly logged is the analysis almost impossible. Better synchronization of sending messages and logging should help with solving this problem.

4. Verify that all nodes and all network devices are working properly.

   During the analysis many problems concerning broken or wrongly implemented nodes were encountered. Data from these nodes are useless for the analysis.

5 Summary

In this report analysis of Cikhaj Experiment 2013 is described. The report provides some basic information about the experiment itself, describes its settings and all the scenarios it was working with. The output of this experiment is set of files containing logs of all the messages that were transmitted in the network during this experiment. The goal of author's term project was to perform data analysis on the set of logs from single nodes and from the sniffers. In this report there are presented results of this analysis.

6 References

[1] http://fi.muni.cz/~xkoutsky/PA018/graphs