Effectiveness Metrics for Intrusion Detection in Wireless Sensor Networks

Andriy Stetsko and Vashek Matyas
{xstetsko, matyas}@fi.muni.cz
Wireless sensor network (WSN) consists of:
- sensor nodes
- base station(s)

Sensor node is equipped with:
- microcontroller
- wireless transceiver
- battery
- sensor(s)

Sensor nodes are resource-constrained
Area of deployment is not physically protected
Sensor nodes are not tamper-resistant
Attacks on WSNs

Jamming attack
- malicious device interferes with physical transmission and reception of wireless communication

Selective forwarding attack
- refuses to forward certain packets and drops them

Packet alternation attack
- modifies packets that it forwards for neighbours

Sinkhole attack
- attracts all traffic from a particular area towards itself
We consider a distributed IDS that consists of IDS agents.
We assume that each sensor node runs an IDS agent.
We assume that sensor nodes employ single-channel transceivers.

Local monitoring technique

Watchdog monitoring technique
Monitoring nodes might cooperate with each other since:

- single monitoring node may not have enough audit information

Cooperation should be limited only to monitoring nodes in close neighbourhood since:

- communication causes energy consumption overhead

**Trade-off between detection accuracy and energy consumption should be found for each particular application**
Conceptual proposal for a suite of tools

1: administrator sets network characteristics and anticipated attacks
2a: Framework generates a possible configuration
2b: sets initial values of parameters of detection modules
3: evaluates effectiveness of the configuration
4: “improves” the values of parameters and repeats steps 3 & 4
5: provides the administrator with optimal IDS agent
The proposed metrics

Metrics of detection
- Number of false negatives
- Number of false positives
- Energy usage

Metrics of attack impact
- Number of lost packets
- Number of modified packets
- Energy usage

We assume that all metrics are measured within a period of time of duration Q, which should be specified by a network operator.
Number of false negatives (1)

\[ F(t) = \sum_{i=1}^{\lvert A \rvert} \sum_{j=1}^{\lvert B_i \rvert} f(a_i, b_{ij}, t) + \sum_{i=1}^{\lvert A \rvert} z(a_i, t) \]

\[ A = \{a_1, \ldots, a_n\} \quad \text{– set of malicious nodes} \]

\[ B_i = \{b_{i1}, \ldots, b_{im}\} \quad \text{– set of neighbours of malicious node} \ a_i \]

If node performs an attack and it is not detected by a node that runs an IDS agent then: \( f(a_i, b_{ij}, t) = 1 \)

else: \( f(a_i, b_{ij}, t) = 0 \)

If node performs an attack and a notification is not sent to an administrator then: \( z(a_i, t) = 1 \)

else: \( z(a_i, t) = 0 \)
How to prioritize IDSs that are able to detect attacks within specified time bounds, say, $T$ minutes after attack was started?

We propose to divide evaluation time $Q$ into $p$ non-overlapping, consecutive intervals of duration $T$ and to count false negatives for each of them and then add the obtained numbers together.

![Diagram showing the evaluation time $Q$ divided into intervals of $T$ and the detection rates for IDS 1 and IDS 2. IDS 1 has a detection rate of B in the first two intervals and B in the next two intervals. IDS 2 has a detection rate of B in all intervals. The malicious node shows detection rates of A in the second and fourth intervals and B otherwise.]
Number of false positives (1)

\[ F(t) = \sum_{i=1}^{C} \sum_{j=1}^{D_i} f(c_i, d_{ij}, t) + \sum_{i=1}^{C} z(c_i, t) \]

\[ C = \{ c_1, \ldots, c_n \} \quad \text{– set of all nodes in a network} \]

\[ D_i = \{ d_{i1}, \ldots, d_{im} \} \quad \text{– set of neighbours of a certain node } a_i \]

If node does not perform an attack but it is considered as malicious by a node that runs an IDS agent then:

\[ f(c_i, d_{ij}, t) = 1 \]

else:

\[ f(c_i, d_{ij}, t) = 0 \]

If node does not perform an attack but a notification is sent to an administrator then:

\[ z(c_i, t) = 1 \]

else:

\[ z(c_i, t) = 0 \]
Number of false positives (2)

How to prioritize IDSs that cause false positives as late as possible during the evaluation time $Q$?

We propose to divide evaluation time $Q$ into $p$ non-overlapping, consecutive intervals of duration $T$ and to count false positives for each of them and then add the obtained numbers together.
Energy usage

\[ \Theta = \sum_{i=1}^{\left| C \right|} E_i \]

\( C = \{ c_1, \ldots, c_n \} \) – set of all nodes in a network

\( E_i \) – energy consumed by a node \( c_i \)

Measures a total amount of energy used by all IDS agents in a network

Does not take into account energy consumption distribution

\[ \Theta = \sum_{i=1}^{\left| C \right|} \varepsilon^{E_i} (\varepsilon > 1) \]

Prioritizes IDSs with energy consumption uniformly distributed over a network
## Metrics of attack impact

<table>
<thead>
<tr>
<th></th>
<th>Packet dropping</th>
<th>Packet delay</th>
<th>Packet mod.</th>
<th>Energy usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selective forwarding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Packet alternation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jamming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinkhole</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Number of lost packets**
- difference between a number of packets that were sent by sensor nodes and a number of packets that were successfully received by a base station

**Number of modified packets**
- number of packets that were modified and accepted by a base station

**Energy usage**
We proposed metrics for evaluation of intrusion detection systems in wireless sensor networks:

- number of false negatives & number of false positives
- energy usage
- number of lost & modified packet