Hiding the Base Station in WSNs

Ruben Rios¹, Jorge Cuellar², Javier Lopez¹

¹NICS Lab – University of Málaga
²Siemens AG, Munich
Receiver-location privacy is concerned with hiding the location of the BS
- Physical protection
- Strategic information

These problems are extensible to any WSN scenario (e.g., sealife monitoring, smart metering, etc.)
Motivation

- WSN solutions are designed to maximize the lifetime of the network
  - Data is transmitted using single-path routing algorithms as soon as an event is detected

- Routing protocols introduce pronounced traffic patterns because all the data is address to the base station (BS)
  - Nodes transmit shortly after receiving a packet
  - Traffic volume is higher as we approach the BS
Agenda

- Motivation
- Problem Statement
- Hiding Scheme
- Evaluation
- Conclusion
Problem Statement

- Network model
  - Vast deployment area
  - Densely populated network
  - A single base station
  - Event-driven monitoring application
  - Sensor nodes share cryptographic keys

- Adversary model
  - Passive eavesdropper with local vision
  - Cannot decrypt messages
  - Cannot distinguish real from bogus traffic
  - Can move in the field based on
    - Time-correlation (flow direction)
    - Rate-monitoring (traffic volume)
  - Can capture a portion of the nodes
Data transmission

- The idea is to **locally homogenise** the number of packets sent by a node to its neighbours such that
  - Real traffic reaches the BS
  - The attacker gains no information

- Whenever a node has to transmit, it sends **two messages**
  - Real message: follows a biased random walk
  - Fake message: must serve as traffic normaliser
Data transmission

We require three properties to ensure the usability (Prop 1) and security (Prop 2, 3) of the system.

- **Prop 1:** Convergence
  \[
  E(\text{dist}(x', BS)) < E(\text{dist}(x, BS))
  \]

- **Prop 2:** Homogeneity
  \[
  \forall y, z \in \text{neigh}(x) \quad F_{rec_m}(x, y) \simeq F_{rec_m}(x, z)
  \]

- **Prop 3:** Exclusion
  \[
  \forall m, m', x, y, t \quad \text{send}(m, x, y, t) \land m \neq m' \implies \neg \text{send}(m', x, y, t)
  \]
The previous properties can be ensured by means of a computationally inexpensive approach:

- Sorted combinations without repetition of two neighbours
- Select one of the combinations uniformly at random

<table>
<thead>
<tr>
<th>neighs(x)</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$n - 1$</td>
</tr>
<tr>
<td>B</td>
<td>$n - 1$</td>
</tr>
<tr>
<td>C</td>
<td>$n$</td>
</tr>
<tr>
<td>D</td>
<td>$n$</td>
</tr>
<tr>
<td>E</td>
<td>$n + 1$</td>
</tr>
<tr>
<td>F</td>
<td>$n + 1$</td>
</tr>
</tbody>
</table>
Data transmission

- The previous properties can be ensured by means of a computationally inexpensive approach
  - Sorted combinations without repetition of two neighbours
  - Select one of the combinations uniformly at random

<table>
<thead>
<tr>
<th>neighs(x)</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>( n - 1 )</td>
</tr>
<tr>
<td>B</td>
<td>( n - 1 )</td>
</tr>
<tr>
<td>C</td>
<td>( n )</td>
</tr>
<tr>
<td>D</td>
<td>( n )</td>
</tr>
<tr>
<td>E</td>
<td>( n + 1 )</td>
</tr>
<tr>
<td>F</td>
<td>( n + 1 )</td>
</tr>
</tbody>
</table>
Data transmission

- The previous properties can be ensured by means of a computationally *inexpensive approach*
  - Sorted combinations without repetition of two neighbours
  - Select one of the combinations uniformly at random

<table>
<thead>
<tr>
<th>neighs((x))</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(n - 1)</td>
</tr>
<tr>
<td>B</td>
<td>(n - 1)</td>
</tr>
<tr>
<td>C</td>
<td>(n)</td>
</tr>
<tr>
<td>D</td>
<td>(n)</td>
</tr>
<tr>
<td>E</td>
<td>(n + 1)</td>
</tr>
<tr>
<td>F</td>
<td>(n + 1)</td>
</tr>
</tbody>
</table>
Data transmission

- The previous properties can be ensured by means of a computationally inexpensive approach
  - Sorted combinations without repetition of two neighbours
  - Select one of the combinations uniformly at random
The previous properties can be ensured by means of a computationally inexpensive approach:

- Sorted combinations without repetition of two neighbours
- Select one of the combinations uniformly at random

\[
\begin{array}{c|c}
\text{neighs}(x) & \text{distance} \\
\hline
A & n - 1 \\
B & n - 1 \\
C & n \\
D & n \\
E & n + 1 \\
F & n + 1 \\
\end{array}
\]
Data transmission

- Every node receives, on average, the same number of packets.
- Real traffic has been most likely transmitted to nodes closer or at equal distance (A, B, C) to the base station.
  - Although some nodes further (E) might also receive real traffic.

\[
\begin{array}{c|c|c}
\text{neighs}(x) & \text{distance} \\
\hline
A & n - 1 \\
B & n - 1 \\
C & n \\
D & n \\
E & n + 1 \\
F & n + 1 \\
\end{array}
\]
Moreover, recall that the attacker cannot distinguish real from bogus traffic

Therefore, what the attacker sees locally gives him no information about the direction to the base station.
However, this protection mechanism becomes useless if the attacker has direct access to the routing tables of the node

- Node capture attacks are likely due to the unattended nature of WSNs

Routing tables are sorted \((L^C, L^E, L^F)\) to allow the data transmission protocol to ensure the Convergence Property

- Leaks the direction to the BS

<table>
<thead>
<tr>
<th>neighs(x)</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(n - 1)</td>
</tr>
<tr>
<td>B</td>
<td>(n - 1)</td>
</tr>
<tr>
<td>C</td>
<td>(n)</td>
</tr>
<tr>
<td>D</td>
<td>(n)</td>
</tr>
<tr>
<td>E</td>
<td>(n + 1)</td>
</tr>
<tr>
<td>F</td>
<td>(n + 1)</td>
</tr>
</tbody>
</table>
We introduce a routing table perturbation scheme that rearranges the elements of the table

- Still ensure that $\text{Prob}(n \in L^C) > \text{Prob}(n \in L^F)$

An optimisation algorithm is used to perturb the tables to a desired degree ($\text{bias} \in [-1, 1]$)

- Trade-off between security and delivery time
Evaluation: Usability

- Message **delivery time** is affected by the probabilistic nature of the protocol
  \[ x_n = 1 + px_{n-1} + qx_n + rx_{n+1} \]
- The routing table perturbation mechanism also impacts negatively on the delivery time
  - Hop count is below 100 for a bias greater than 0.2
Evaluation: Usability

- The use of fake traffic impacts on the network lifetime.
- The durability of fake traffic is controlled by a parameter, which is dependent on the hearing range ($n$) of the adversary.
  - Discarded after several hops.
- The hearing range of a typical adversary is $n=1$ (local adversary).
Evaluation: Privacy

- We have verified the privacy protection level of our solution for different types of adversaries
  - Passive eavesdroppers should better move at random
  - Active attackers must capture more than 1/10 of nodes to be successful
Conclusion

- The location of the base station is critical for the survivability and privacy of the network

- We present a receiver-location privacy solution capable of countering both passive and active attackers

- The protection mechanism introduce additional overhead and impacts on the delivery time but it includes two parameters to balance between usability and security

- Future work
  - Reduce the overhead caused by fake traffic
  - Protect the topology discovery process
Thanks for your attention!

NICS Lab – University of Málaga
https://www.nics.uma.es/

Extra Slides

NICS Lab – University of Málaga
https://www.nics.uma.es/

Analysis of Potential Limitations

- The **topology** of the network might negatively impact the **convergence** of real packets
  - Theorem: Real messages reach the base station if \( F < \sqrt{2C(S - C)} \)

- Validation on randomly deployed networks