Towards Understanding Source Location Privacy in Wireless Sensor Networks Through Fake Sources

Arshad Jhumka, Matthew Bradbury and Matthew Leeke

Department of Computer Science,
University of Warwick, Coventry
United Kingdom, CV4 7AL
{arshad, csujbt, matt}@dcs.warwick.ac.uk

Monday 25th June 2012

Some Quick Definitions

- Wireless Sensor Network (WSN) is a collection of Wireless Sensor Nodes
- Nodes are equipped with the sensors necessary for the network's application
- Nodes communicate wirelessly
- There are lots of applications for WSNs
  - Tracking
  - **Monitoring**

Introduction

- Privacy threats can be classified as either content-based or **context-based**
- Content-based threats have been widely addressed (using cryptography) [Perrig et al., 2004]
- Context-based threats are varied
- We focus on protecting the location context of broadcasting nodes

Important Considerations

- Wireless Sensor Nodes are energy constrained
- Sending messages is the most expensive task
- Receiving messages is the next most expensive task [Shnayder et al., 2004]

The Problem of Source Location Privacy (SLP)

Given:

- A WSN that detects valuable assets
- A Node detecting a valuable asset broadcasting information

Found:

- An attacker can find the source node by backtracking the messages through the network
- So by deploying a network to monitor a valuable asset we have provided a way for it to be captured

The Problem:

- Panda-Hunter Game
- Difficult

Related Work

- [Ozturk et al., 2004]
- [Kamat et al., 2005]

System Model

- Network modelled as a graph $G = (V, E)$
- A link exists between two nodes m and m' if both can communicate with each other
- Two nodes $m \in V$ and $m' \in V$ are 1-hop neighbours iff $\{m, m'\} \in E$
- There exists a sink node $S \in V$ that collects data
- Other nodes $v \in V \setminus \{S\}$ route data to the sink for collection
- The network is event-triggered when an object of interest is sensed by a node that then begins broadcasting

Attacker Model

- We consider a single mobile *distributed eavesdropping* attacker
- Relevant System Assumptions:
  - Messages sent by a source are encrypted and include node ID
  - Only the sink can tell a nodes location from the ID
- Assumptions:
  - The attacker can tell the direction the message came from
  - The attacker can move at any speed and has no power limitations
  - The attacker knows of (i) sink location (ii) network topology (iii) routing algorithm

Fake Sources

- Select a subset of nodes in the network to act as fake sources that simulate the real source
- Two types: (i) permanent (ii) temporary
- Both modelled using a duration parameter and a rate parameter
- Aim to broadcast messages to lure the attacker away from the real source
- Fake messages are similar to real messages to prevent the attacker distinguishing between them

Contribution 1: Formal Definition of SLP

- A fake source F_i is a determined by the *duration* d_i
- If d_i is (small and) finite, then the fake source is *temporary*, else it is *permanent*

Definition 1 (SLP): Given a graph $G = (V, E)$, a set F of fake source tags $\{F_1 \ldots F_f\}$, each F_i associated with duration d_i, a relation $\prec$ on V, a set N of m nodes $\{n_1 \ldots n_m\} \subset V$, a deadline τ and a routing strategy R, assign tags in T to nodes in N to obtain an m-node schedule σ for T that meets the deadline τ under R and obeys $\prec$.

Contribution 2: SLP Complexity

Theorem 1 (SLP Complexity): The Source Location Privacy problem is NP-complete.

Proof: We reduce the multiprocessor scheduling with precedence constraints (MSPC) to SLP. We first define the MSPC problem, and then identify the mapping between MSPC and SLP.

Contribution 2: SLP Complexity

Instance: The MSPC problem is as follows: Given a set $T = \{T_1 \dots T_n\}$ of tasks, with task T_i having execution time e_i, a set P of $m \in \mathbb{Z}^+$ processors, partial order $\lessdot$ on T, and a deadline $D \in \mathbb{Z}^+$.

Question: Is there an m-processor schedule σ for T that meets the overall deadline D and obeys the precedence constraints, i.e., such that $T_i \lessdot T_j$ implies that $\sigma(T_j) \geq \sigma(T_i) + e_i$.

Contribution 3: Heuristic and Parameters

- Heuristic algorithm developed that uses fake sources
- These fake sources are used to "pull" the attacker towards them
- They need to be created in the direction away from the real source
- Selection of fake sources in the algorithm is important
- Can vary the *duration*, *rate* and *probability* of fake sources
- Heuristic uses temporary fake sources try to first "lure" attacker away, and then permanent fake sources ensure the attacker is kept away from the real source

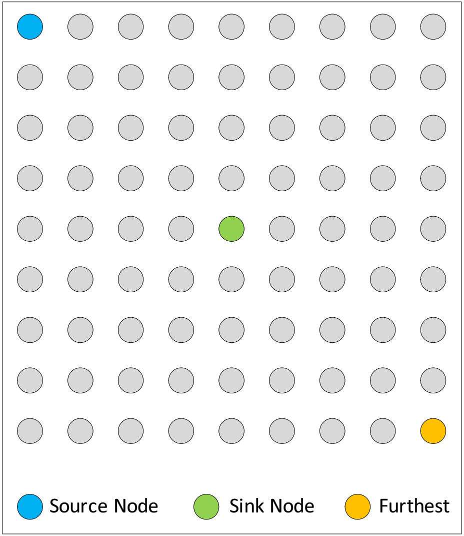
Experimental Setup: Simulation Environment

- JProwler
- Rayleigh Radio model used as it assumes node mobility
- Gaussian Radio model was the alternative

Experimental Setup: Network Configuration I

- Square grid network of 11^2, 15^2 and 21^2 nodes
- Nodes 28 meters apart
- Sink node positioned at the center of the grid
- Source node position in one of the four corners or at a random location along the network edge
- Attacker starts at the location of the sink
- 500 Repeats

Experimental Setup: Network Configuration II



Experimental Setup: Protocol Configuration

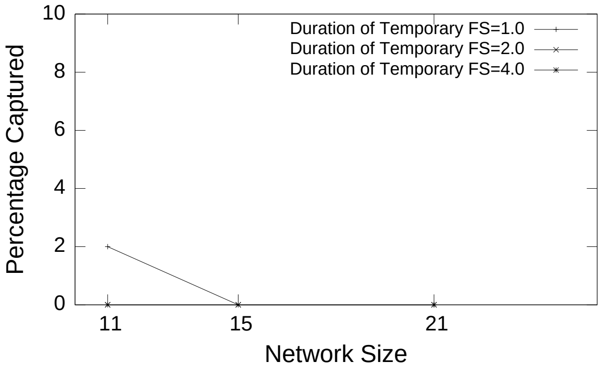
- Flooding Protocol used as a baseline
- Temporary Fake Source Duration $\in \{1, 2, 4\}$ seconds
- Real Source Rates $\in \{1, 2, 4\}$ messages per second
- Fake Source Rates $\in \{2, 4, 8\}$ messages per second
- Simulations never run where real source message rate is greater than fake source message rate

Experimental Setup: Safety Period

- Calculated for each network size and source rate
- Defined as twice the amount of time it took an attacker to capture the source when no protection was in place
- Twice the time allows an attacker to go to the opposite end of the network and back
- A bounded safety period allows bounded simulation time

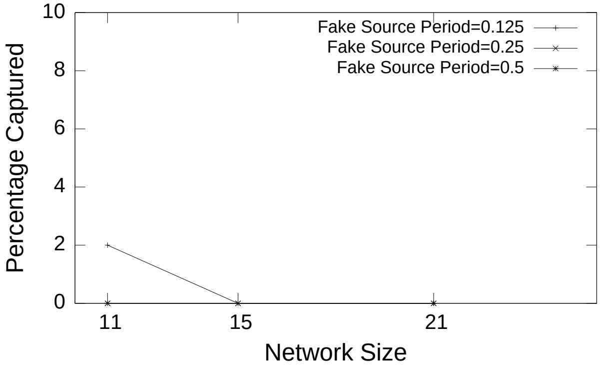
Results: Impact of Temporary Fake Source Duration

- Increasing temporary fake source duration leads to a decreased capture ratio



Results: Impact of Message Rates

- Intuitively, the higher the rate of the fake source , the greater the "pull" it has
- Increasing the message rate of fake sources leads to a decreased capture ratio
- However, with high transmission rates more collisions can occur increasing the capture ratio



Results: Limitations

- Algorithm designed for a grid network structure
- Assumed that attackers follow a message when it has received one

Summary

- Developed a new way of modelling SLP using permanent and temporary fake sources
- Formalised SLP and shown the problem to be NP-complete
- Identified message rates and fake source duration as important parameters
- Provided a heuristic which has been shown provide optimal privacy under certain parameter settings

Further Work

- Investigate other network configurations and shapes
- Consider fake sources in the context of a more perceptive attacker
- Investigate optimising the algorithm to reduce energy usage by making the algorithm adaptive by taking advantage of run-time knowledge

The End

Any Questions?

References I

Kamat, U., Zhang, Y., and Ozturk, C. (2005).
Enhancing source-location privacy in sensor network routing.
In *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems*, pages 599–608.

Ozturk, C., Zhang, Y., and Trappe, W. (2004).
Source-location privacy in energy-constrained sensor network routing.
In *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pages 88–93.

Perrig, A., Stankovic, J., and Wagner, D. (2004).
Security in wireless sensor networks.
Communications of the ACM - Special Issue on Wireless Sensor Networks, 47(6):53–57.

Shnayder, V., Hempstead, M., Chen, B., Allen, G. W., and Welsh, M. (2004).
Simulating the power consumption of large scale sensor network applications.
In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pages 188–200.