The Sybil Attack in Sensor Networks: Analysis & Defenses

Outlines

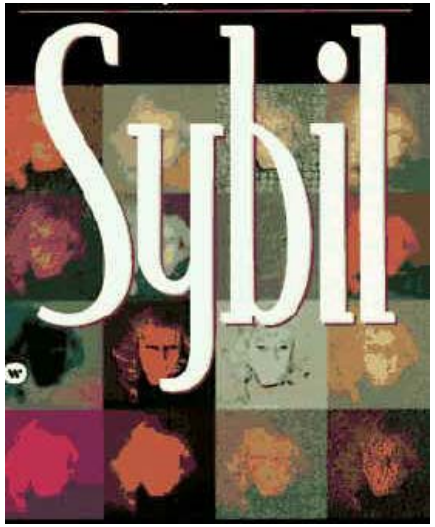
- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Outlines

- **Introduction**
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Introduction

- Security in Sensor Network
  - Wireless network natures
  - Sensor nodes constraints
- Sybil Attacks
  - First described in peer-to-peer networks.
  - An attack against identity.
  - A particularly harmful attack in sensor networks.



Definition of Sybil Attack

- In this paper
  - *A malicious device illegitimately takes on multiple identities.*
  - *The additional identities are called Sybil nodes.*
- Question:
  - How does an attacker create Sybil nodes and use them?

Outlines

- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Sybil Attack Taxonomy

- ***Dimension I*** – Direct vs. Indirect Communication
  - Direct Communication
    - Legitimate nodes can communicate with Sybil nodes directly.
  - Indirect Communication
    - One or more of the malicious devices claims to be able to reach the Sybil nodes.
    - Messages sent to a Sybil node are *routed* through one of these malicious nodes.

Sybil Attack Taxonomy

- ***Dimension II*** – Fabricated vs. Stolen Identities
  - Fabricated
    - Simply create arbitrary new Sybil identities.
  - Stolen
    - Assign other legitimate identities to Sybil nodes.
    - May go undetected if attacker destroys or disable them.
  - Identity Replication Attack
    - The same identity is used many times and exists in multiple places in the network.

Sybil Attack Taxonomy

- ***Dimension III*** – Simultaneity
  - Simultaneous
    - All Sybil identities participate in the network at once.
  - Non-Simultaneous
    - Only act as a smaller number of identities at any given time *by:*
      - Letting different identities join and leave
      - Or only using each identity once.
      - Having several physical devices swap identities.

Outlines

- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Known Attacks

- Distributed Storage
  - Defeat replication and fragmentation mechanisms
- Routing
  - Attack routing algorithm
  - Geographic routing
  - Evade misbehavior detection mechanisms

New Attacks

- Data Aggregation
  - With enough Sybil nodes, an attacker may be able to completely alter the aggregate reading.
- Voting
  - Depending on the number of identities the attacker owns, he may be able to determine the outcome of any vote.
    - Either claim a legitimate node is misbehaving or Sybil nodes can vouch for each other…

New Attacks

- Fair Resource Allocation
  - Using Sybil attack, a malicious node can obtain an unfair share of any resource shard in per-node manner.
  - Consequently, cause DoS to legitimate node, and also give the attacker more resources to perform attacks.
- Misbehavior Detection
  - Sybil nodes could “spread the blame” .
  - Even action is taken to revoke the offending nodes, the attacker can continue using new Sybil identities to misbehave.

Outlines

- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Defenses

- Two types of ways to validate an identity
  - Direct validate
  - Indirect validate

Defenses

- **Previous Defense**
  - *Resource testing*

    By verifying that each identity has as much of the tested resource as a physical device.
    - Computation, storage
    - and communication
    - Unsuitable for wireless sensor networks
      - WHY?

New Defenses in this paper

- Radio Resource Testing
- Random Key Predistribution
- Registration
- Position Verification
- Code Attestation

Radio Resource Testing

- Direct validation
- Assumptions
  - Any physical device has only one radio
  - A radio is incapable of simultaneously sending or receiving on more than one channel.
- The basic idea:
  - A node assigns each of its n neighbors a different channel.
  - By challenging a neighbor node on the exclusively assigned channel, a sensor node can detect Sybil nodes with a certain probability.

Radio Resource Testing with enough channels

- ***Suppose:***
  - *s* Sybil nodes out of *n* neighbors.
  - One channel for each neighbor.
- ***Pr*** (choose a channel is not being transmitted on)

  = $\frac{s}{n}$
- ***Pr*** (not detecting a Sybil node)

  = $\frac{n-s}{n}$.
- Repeat test for r round

  Pr (no Sybil nodes being detected)

  = $\left(\frac{n-s}{n}\right)^r$

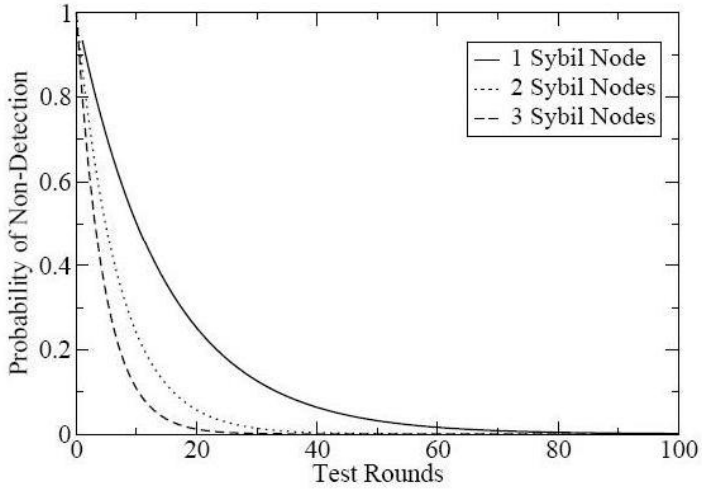


Figure 1: Probability of no Sybil nodes being detected, using the radio defense, with a channel for every neighbor. Assumes 15 neighbors (including Sybil nodes), any number of which could be malicious.

Radio Resource Testing with limited channels

- In case of limited channels, only subset of its neighbors can be tested at one time.

Radio Resource Testing with limited channels

- Repeating this test for *r* rounds

  The probability of a Sybil node being detected is

$$
\begin{aligned}
Pr(detection) &= 1 - Pr(nondetection)^{r}_{1round} \\
&= 1 - (1 - Pr(detection)_{1round})^{r} \\
&= 1 - \left(1 - \sum_{allS,M,G} \frac{\binom{s}{S}\binom{m}{M}\binom{g}{G}}{\binom{n}{c}} \frac{S-(m-M)}{c}\right)^{r}
\end{aligned}
$$

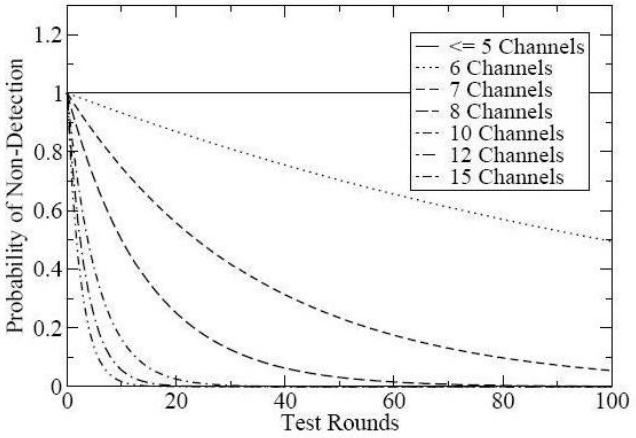


Figure 2: Probability of no Sybil nodes being detected, using the radio defense, with fewer channels than neighbors. Assumes 5 correct neighbors, 5 malicious neighbors, and 5 Sybil neighbors.

Random Key Predistribution

- **Random Key Predistribution**
  - Each node is assigned a random set of keys or key-related information.
  - In key set-up phase, each node can discover or compute the common key it shares with its neighbors…
  - Node-to-node secrecy.

Random Key Predistribution

- **Key ideas:**
  - Associating the node identity with the keys assigned to the node.
  - Key validation, i.e., the network being able to verify part or all of the keys that an identity claims to have.
    - *Direct or Indirect Validation?*
- **Different variants**
  - Key pool
  - Single-space pairwise key distribution
  - Multi-space pairwise key distribution

Key Pool

- An Extension
  - Let $\hat{\Omega}(ID) = \{K_{\beta_1}, K_{\beta_2}, \ldots, \tilde{K}_{\beta_k}\}$ be the set of keys assigned to *ID*,
    - ID is the identity of the node, anc β_i is the index of its i[th] key in the key pool,
  - The set of keys that node *ID* possesses are determined by:

$$\beta_i = \bar{PRF}_{H(ID)}(i),$$

    - where *H* is a hash function, and *PRF* is a pseudo random function.
  - The index of a node's *i[th]* key, β_i is determined by a *pseudo random function* with *H(ID)* as the function's key, and *i* as its input.

Key Pool

- *An example*
  - Node ID = 30
  - Key set = { K_1, K_8, K_{12}, K_{78}, …}
  - Rule: pick the 3rd indices
  - How to validate this node ID (= 30) ??
    - Test whether $PRF_{H(30)}(3) = 12$??
  - What properties does this scheme have?

Key Pool

- **What can the attacker do?**
  - Capture legitimate nodes and read off the keys,
  - Build up a compromised key pool *S*,
  - Fabricate *usable Sybil identities ID'* to use in Sybil attack, which means *ID'* must be able to pass the validation by other nodes.
- ***Question:***
  - Given a set of compromised keys S
  - How difficult for an attacker to generate a usable Sybil identity?
  - How to evaluate the difficulty?

Key Pool

- How to evaluate the difficulty?
  - The time complexity to generate a usable Sybil node ID given a set of compromised nodes could be expressed in terms of the probability *p* that a random identity is a usable Sybil identity.
  - So, the expected number of times an attacker has to try to find a usable Sybil identity is *1/p*.

Random Key Predistribution

- In contrast, Pairwise key distribution
  - Assigns a unique key to each pair of nodes…
  - Single-space Pairwise Key Distribution
  - Multi-space Pairwise Key Distribution

Multi-space Pairwise Key Distribution

- To further enhance the security of single-space…
- In this scheme, each sensor node will be assigned *k* out of the *m* key spaces.
- Key computation
  - Use single-space scheme, if they have one or more key spaces in common.

Summary of Random Key Predistribution

- Key Pool
  - One-way function
  - Indirect validation
- Single-space pairwise key distribution
  - λ-secure property
  - Direct validation ensures globally consistent outcome.
- Multi-space pairwise key distribution

Other Defenses

- Identity Registration
  - Based on a trusted central authority
  - However,
    - Attacker may be able to control the good list.
    - Need maintain the deployment information
- Position Verification
  - Assume network is immobile.
  - Verify the physical position of each node.
  - How to securely verify a node's exact position is still an open question.
  - Mobile attacker's identity needs to be verified simultaneously.

Other Defenses

- Code Attestation
  - Code running on a malicious node must be different form that on a legitimate node.
  - The technique is not readily applicable to wireless network.
    - High cost
    - Energy consumption

Outlines

- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Comparison and Discussion

- All these Sybil Defenses…

Defense	Who Can Validate	Remaining Sybil Vulnerabilities
Radio	Neighbors	Indirect Com., Non-Simult.
Position Verification	Neighbors	Indirect Com.*
Registration	Anyone	Stolen IDs
Key Predistribution	Anyone w/shared keys	Stolen IDs**
Code Attestation	Anyone	None***

* Assume that nodes can only verify the position that they directly communicate with;

** Key predistribution can not stop an attacker from using stolen identities… but it does make it more difficult for the attacker to steal identities in the first place.

Outlines

- Introduction
- Three Dimensions of Sybil Attack Taxonomy
- Attacks
  - Known & New attacks
- Defenses
  - Radio Resource Testing
  - Random Key Predistribution
  - Other Defenses
- Discussion
- Conclusion

Conclusions

- The first paper that systematically analyzes the Sybil attack and its defenses in sensor networks.
- It introduces a taxonomy of the different forms of the Sybil attack.
- Several new defenses are proposed.

Conclusions

- *In radio resource testing*
  - Based on the assumption that each node has only one channel and can't send and receive simultaneously on more than one channel.
  - How a sensor node assigns the radio channels to its neighbors?
  - The testing process may consumes a lot of battery power
- *In random key predistribution*
  - If some keys are compromised, the attacker may be able to falsely claim the identities of many non-compromised sensor nodes.
  - It's not practical in a mobile wireless network environment.
- *Other defenses*
  - Have their own drawbacks and not very applicable in wireless sensor networks…