Modeling Adaptive Node Capture Attacks in Multihop Wireless Networks

Patrick Tague & Radha Poovendran
Network Security Lab (NSL)
Dept. of Electrical Engineering
University of Washington

Protocol eXchange Seminar, NPS, Monterey, CA
January 29, 2007

Funding provided by:

ONR YIP

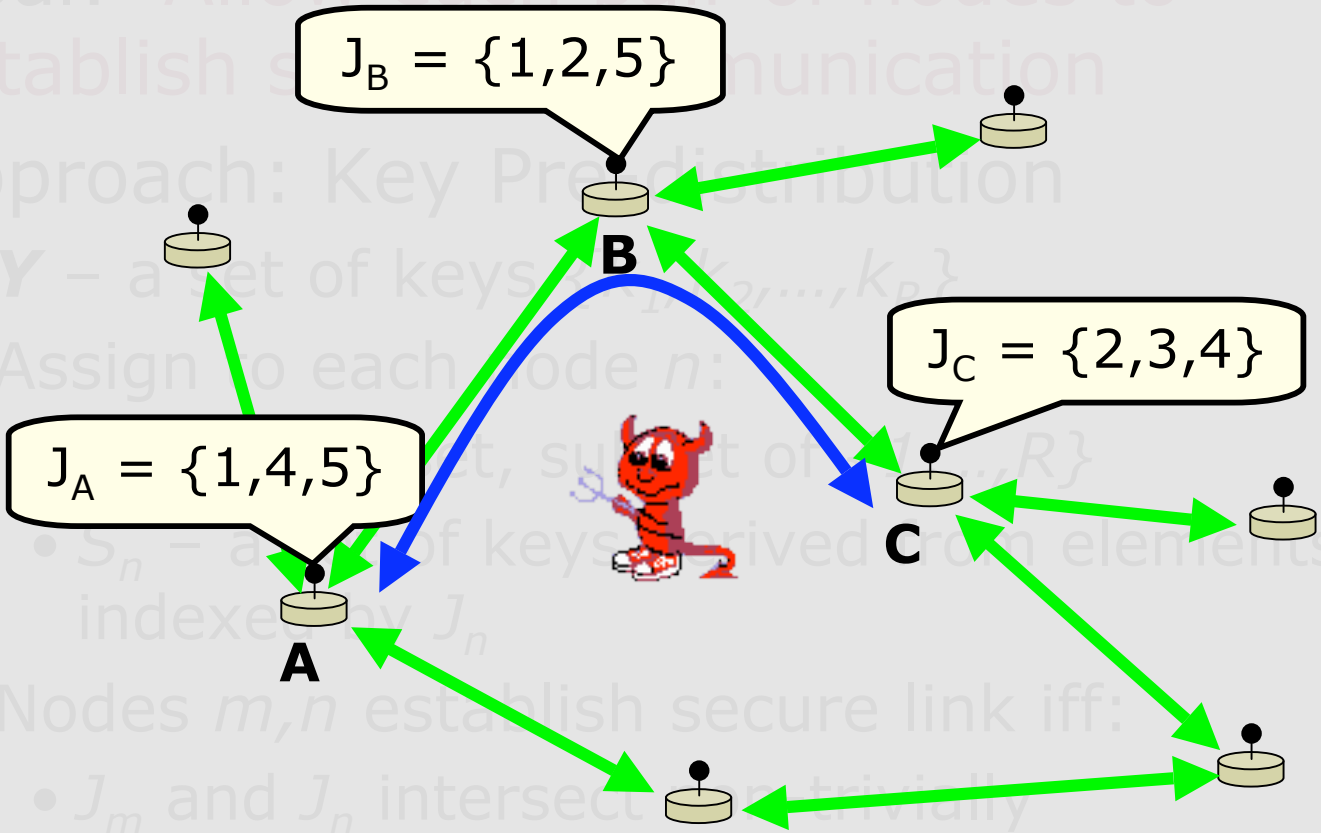
ARO PECASE

DoD/NSA IASP

Outline

- Problem – Node Capture Attacks
- Attack Illustration
- Attack Model
- Model Analysis
- Prevention & Mitigation Techniques
- Case Study

Key Establishment in Wireless Networks



Adversary

- Goal: Recover a set ***Z*** = ***Z(Y)*** of target elements *with minimal resource expenditure* using available information.
  - Assume computationally bounded adversary
  - Instead of attacks on cryptography, capture & access individual nodes intelligently
- Existing research focuses on random node capture

Contributions

- Propose two strategies for node capture
- Model optimal node capture attacks
  - Metrics for optimality
- Analysis
  - Attack complexity
  - Algorithms using efficient heuristics
- Prevention & mitigation techniques
  - Prove that some attacks can’t be prevented
  - Propose new key assignment methods for mitigation purposes

Set Coverage Attack

- Random and Optimal Attacks
  - $\boldsymbol{Y} = \boldsymbol{Z} = \{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, k_{10}, k_{11}, k_{12}\}$

Node	*Assigned Keys*
1	$S_1 = \{k_2,k_5,k_9,k_{10}\}$
2	$S_2 = \{k_3,k_4,k_5,k_9\}$
3	$S_3 = \{k_1,k_5,k_7,k_9\}$
4	$S_4 = \{k_1,k_2,k_4,k_6\}$
5	$S_5 = \{k_4,k_6,k_7,k_{11}\}$
6	$S_6 = \{k_6,k_7,k_8,k_{10}\}$
7	$S_7 = \{k_1,k_3,k_8,k_{12}\}$
8	$S_8 = \{k_5,k_7,k_9,k_{12}\}$
9	$S_9 = \{k_1,k_3,k_5,k_8\}$

The optimal set coverage attack is equivalent to the minimum set cover solution!

Subset Coverage Attack

- It may be the case that all keys are not equally valuable to the adversary.
  - A key may never be used
  - A key may be used more often than others
- Instead of targeting keys, adversary can directly target secure links.
- Let **Z** be a set of subsets of **Y** such that

Node	List of compromised links	Value	
1	(2,3),(2,8),(8,9)	3	0
2	(1,2),(1,8),(1,9),(8,9)	4	0
3	(1,2),(1,8),(1,9),(2,8),(4,7),(4,9),(5,8),(6,8),(8,9)	9	
4	(2,5),(3,7)	2	1
9	(2,7),(3,4),(3,7),(4,7),(6,7)	5	2

This iterative attack may not be optimal, but it is efficient.

Attack Model Requirements

- Metric for attack efficiency
- How to incorporate node heterogeneity
- How to interpret the adversary's goals
- Metric for attack optimality

Attack Model - Metrics

- Cost (resource expenditure):
  - Communication cost – request/collect information from nodes
  - Computation cost – run an algorithm
  - Physical cost – capture nodes & access data
- Benefit
  - What fraction of the target set ***Z*** has been recovered?
  - Metric for comparison between nodes
- Goal: Achieve benefit of 1 with minimum cost.

Node Capture Attack Model

- Define:
  - $a_{i,n}$ – value of node n toward the recovery of the element z_i in $\boldsymbol{Z}$
  - s_i – a threshold parameter which models the sufficiency of recovered information toward the recovery of z_i
  - c_n – the total cost to capture node n
  - x_n – 1 if node n is captured, 0 else
- Attack:
  - Given $\boldsymbol{A}=[a_{i,n}]$, $\boldsymbol{s}=[s_i]$, $\boldsymbol{c}=[c_n]$
  - Find $\boldsymbol{x}=[x_n]$ to minimize $\boldsymbol{c}^T\boldsymbol{x}$ s.t. $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{s}$

Example

Node	*Assigned Keys*
1	$S_1 = \{k_2,k_5,k_9,k_{10}\}$
2	$S_2 = \{k_3,k_4,k_5,k_9\}$
3	$S_3 = \{k_1,k_5,k_7,k_9\}$
4	$S_4 = \{k_1,k_2,k_4,k_6\}$
5	$S_5 = \{k_4,k_6,k_7,k_{11}\}$
6	$S_6 = \{k_6,k_7,k_8,k_{10}\}$
7	$S_7 = \{k_1,k_3,k_8,k_{12}\}$
8	$S_8 = \{k_5,k_7,k_9,k_{12}\}$
9	$S_9 = \{k_1,k_3,k_5,k_8\}$

- Set coverage attack:
  - $z_i = k_i$: $\boldsymbol{Z} = \boldsymbol{Y}$
  - $s_i = 1$: each element of S_n reveals an element of $\boldsymbol{Z}$
  - $a_{i,n} = 1$ if k_i is in S_n, 0 otherwise
- From table at left:
  - $\boldsymbol{A}$ is a 12x9 matrix with 4 ones / column

Analysis of Attacks

- Attacks in the given model are NP-hard – optimal solution is that of an integer program
  - Well-studied greedy heuristic
  - Let A_n be the sum of the n^{th} column of $\boldsymbol{A}$
  - Choose the node n with maximum A_n/c_n, then update $\boldsymbol{A}$, $\boldsymbol{x}$, and $\boldsymbol{s}$
  - Ratio bound computed as a function of $\boldsymbol{A}$

Attack Prevention

- Privacy-preserving protocols:
  - Instead of announcing key indices (J_n), each node broadcasts a challenge
    - Ex: n broadcasts $\alpha_n, E_{k1}(\alpha_n), \ldots, E_{k4}(\alpha_n)$
  - Receiver m infers shared keys but learns nothing about remaining keys ($J_m \setminus J_n$)
- Prevents subset coverage attacks because adversary cannot compute subsets

Attacks with Partial Info

- Feature of heuristic algorithm
  - Choice of n does not depend on individual entries $a_{i,n}$, only column sum A_n
  - **Lemma:** If $|S_n|$ can be computed by the adversary, set coverage attacks (those in which $\boldsymbol{Z} \subseteq \boldsymbol{Y}$) can be performed deterministically (regardless of crypto).
- What if $|S_n|$ is random?

Storage Randomization

- Randomize $|S_n|$: attempt to mitigate set coverage attack
  - Assume probability distribution $P(|S_n|)$ is known
  - Adversary cannot compute $\boldsymbol{A}$ or column sums A_n of $\boldsymbol{A}$
  - Adversary can estimate the value of each node

Probabilistic Attacks

- Let τ_n be the number of elements in S_n known to the adversary
  - This is computable, even under privacy-preserving protocols
- Compute the expected value $\kappa(\tau_n)$ of $|S_n|$ given τ_n
- Probabilistic heuristic
  - At each iteration, choose node *n* to maximize $(\kappa(\tau_n)-\tau_n)/c_n$

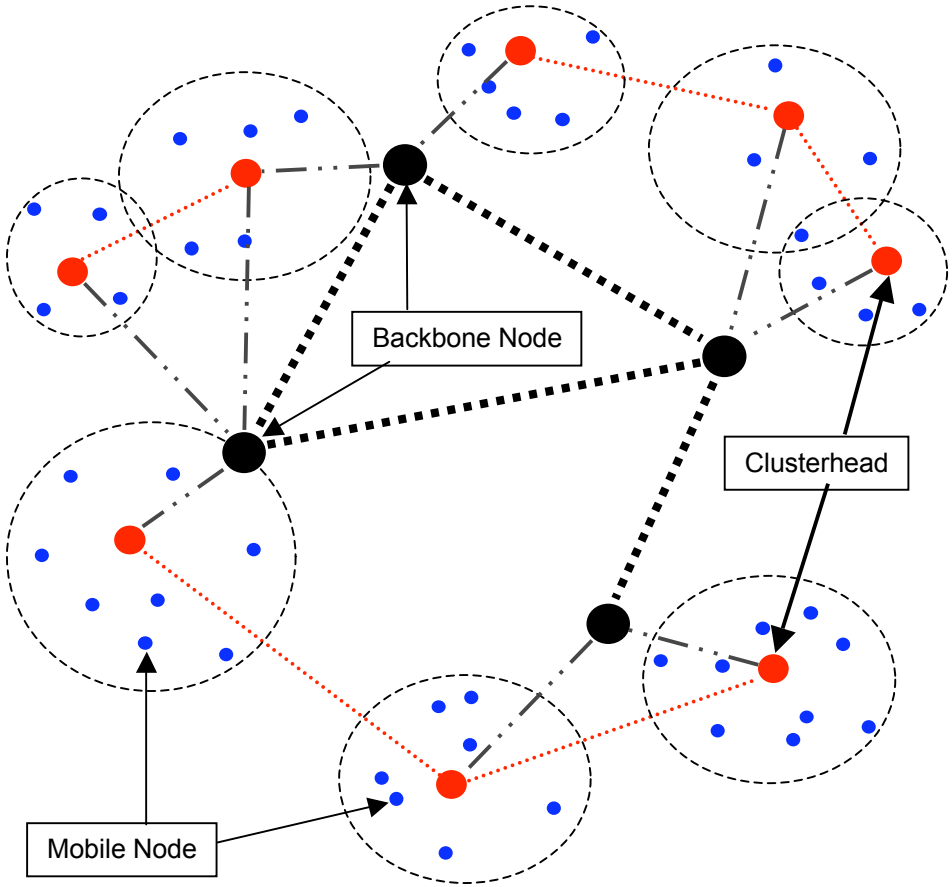
Mitigation of Probabilistic Attacks using Randomization

- **Result:** Even under randomization, the adversary can expect to out-perform random node capture.

- Why?
  - Distribution used for randomization provides information for the attack.
  - Expected performance of the attack is still better than random capture (with or without randomization).

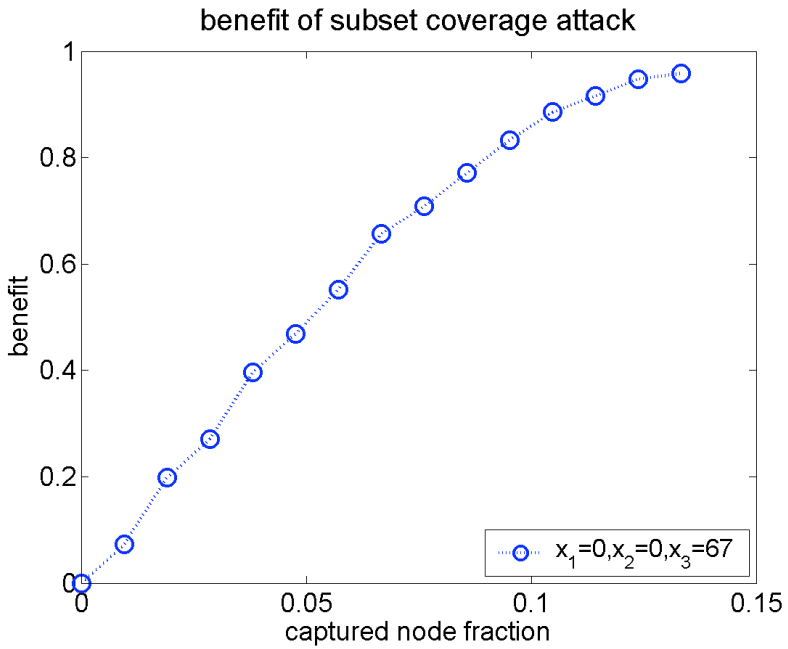
Case Study – Mesh Network

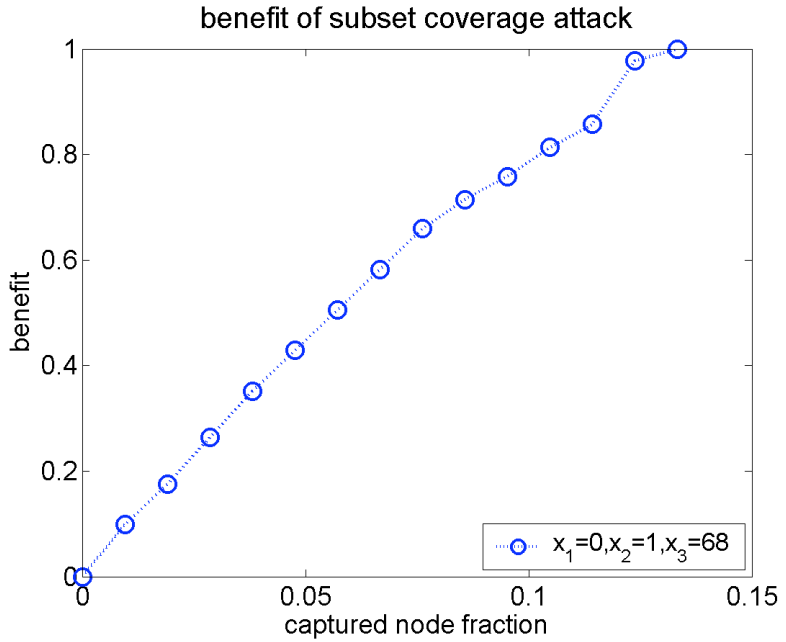
- Simulate a wireless mesh network with a 3-class hierarchy
  - c_n, $P(|S_n|)$ depend on node class
  - Attacks:
    - Subset coverage attack with fixed $|S_n|$ for each class
    - Set coverage attack with randomized $|S_n|$ for each class



Subset Coverage Attack

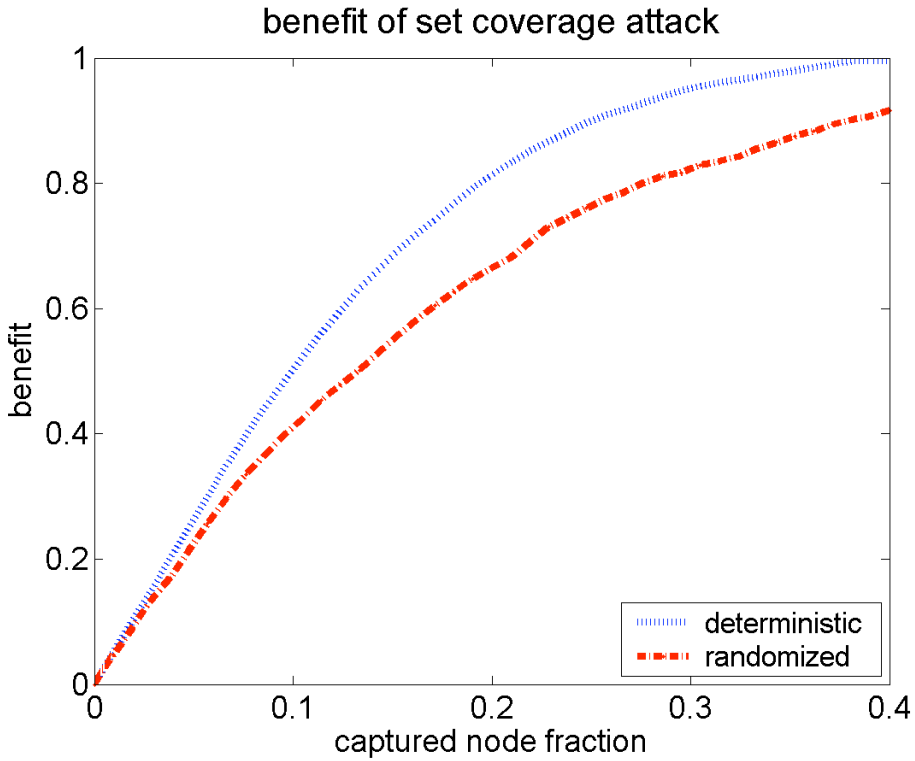
- **Goal**: recover ***Z***, the set of links between backbone and clusterhead nodes, with minimum cost
  - Cost difference between classes is more prominent in first simulation than in second





Probabilistic Set Coverage Attack

- Goal: recover ***Y*** with minimum cost
  - Deterministic assignment included for comparison



Summary of Contributions

- Proposed two strategies for node capture
  - Set coverage, subset coverage
- Provided a model for optimal node capture attacks
  - Proved the optimal attack to be NP-hard
  - Provided heuristic attack algorithms
- Proved that some attacks cannot be prevented
  - Even under privacy-preserving protocols
- Investigated the effect of storage randomization as a mitigation technique

- This work to appear in *Ad Hoc Networks*, 2007