Location-Based Security Design for Wireless Sensor Networks

Yanchao Zhang

Department of Electrical & Computer Engineering

New Jersey Institute of Technology

Roadmap

- Introduction
  - Wireless sensor networks
  - Security requirements & challenges
- Security issues to tackle
- Our location-based solution
- Conclusion & future work

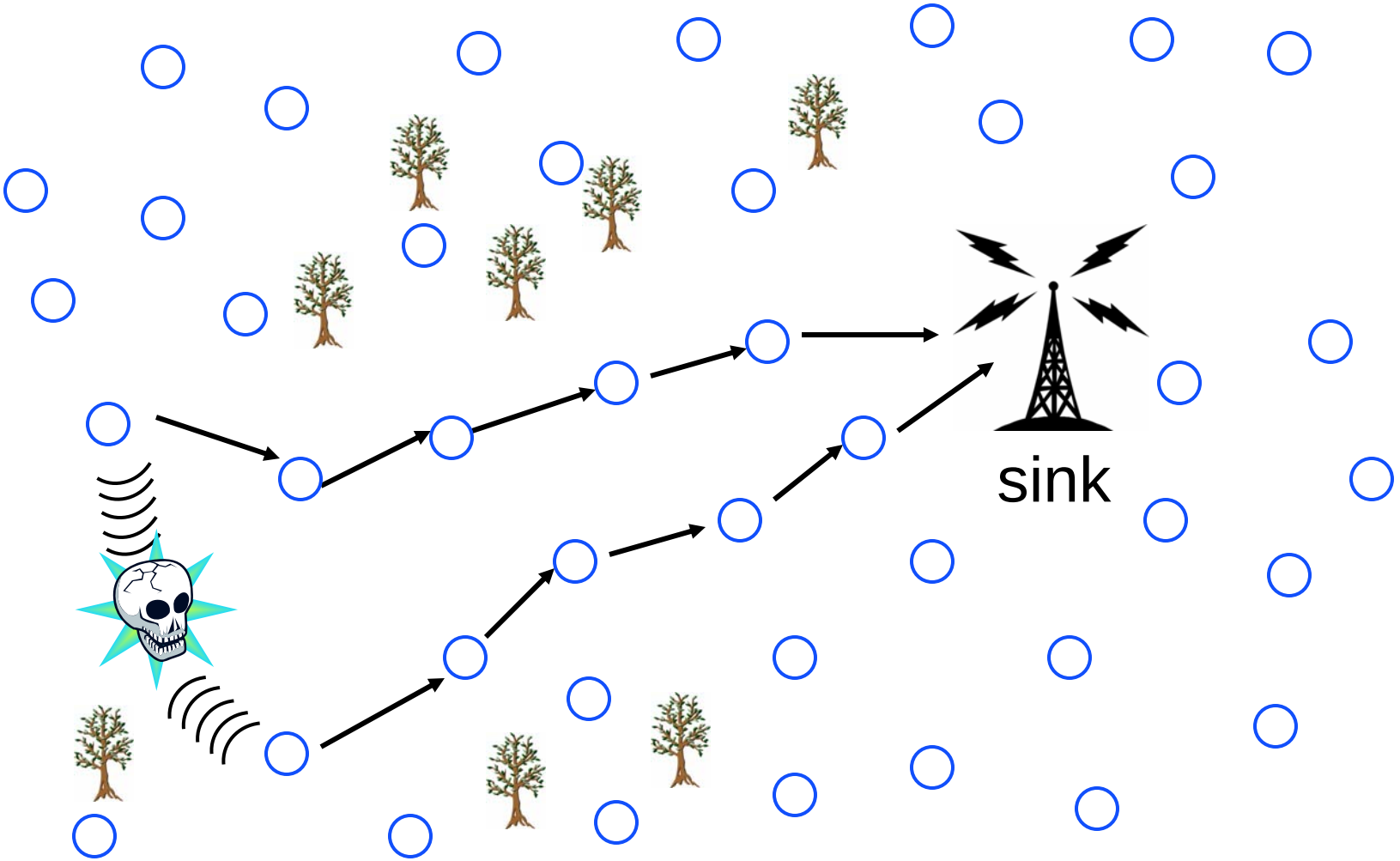
What is a Wireless Sensor Network?

- A wireless sensor network (WSN) is composed of a large number of low-cost sensor nodes randomly deployed to monitor the field of interest
- Sensor nodes
  - Limited in energy, computation, and storage
  - Sense/monitor their local environment
  - Perform limited data processing
  - Communicate untethered over short distances
- Sink
  - Gather data from sensor nodes and connect the WSN to the outside world

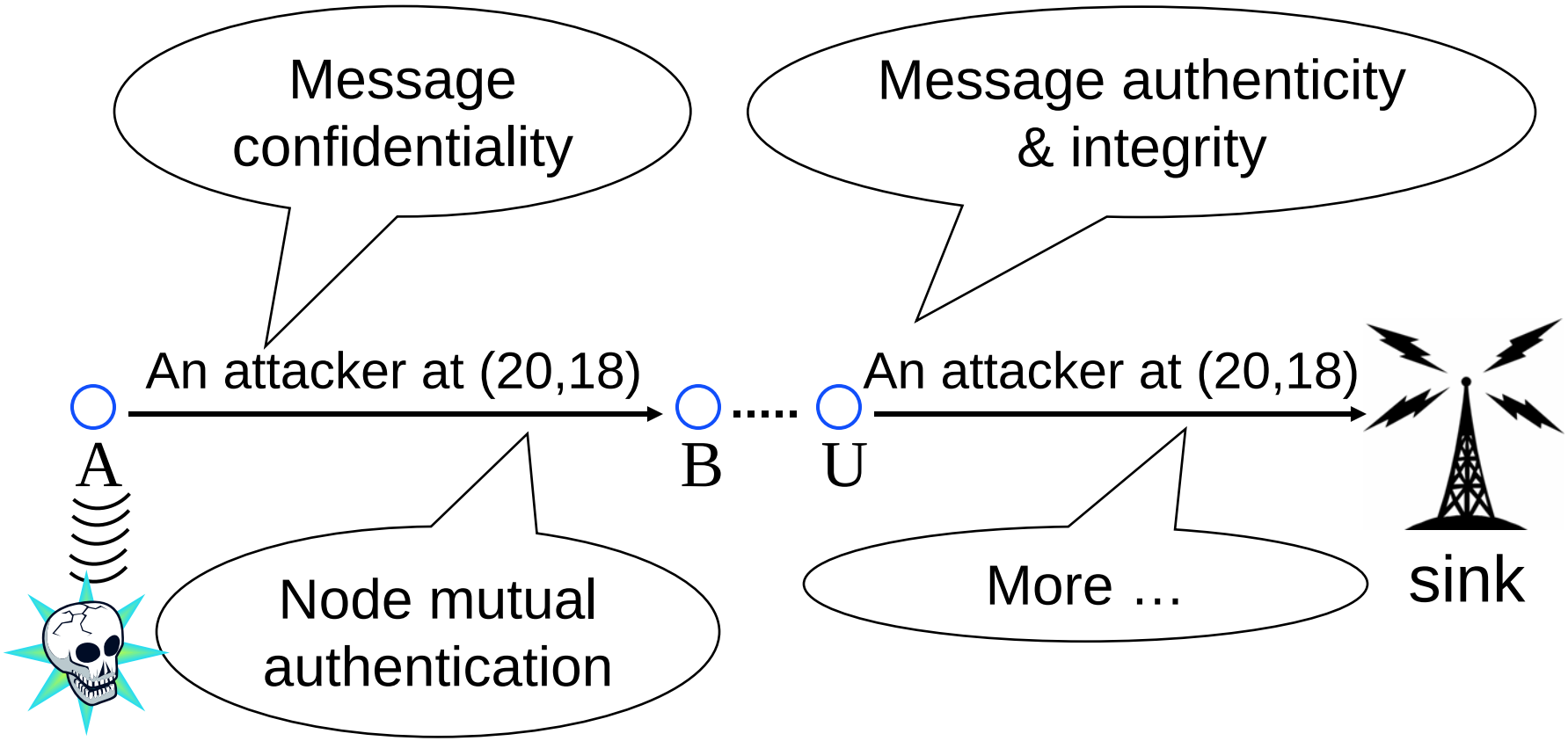
Wireless Sensor Networks

- Applications
  - Physical security for military operations
  - Indoor/outdoor environmental monitoring
  - Seismic and structural monitoring
  - Industrial automation
  - Bio-medical applications
  - Health and wellness monitoring
  - Inventory location awareness
  - Future consumer applications, e.g., smart homes
  - …

A Sample Wireless Sensor Network



Security Requirements

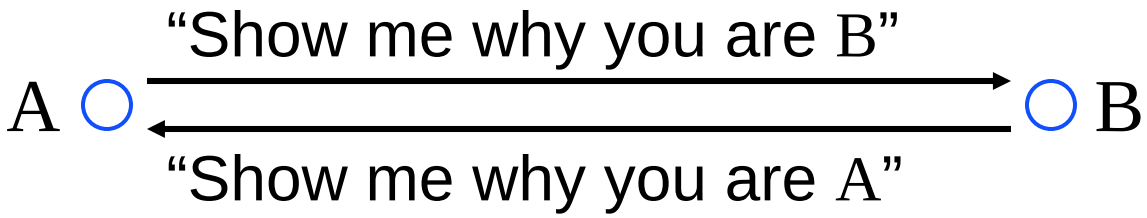


Research Challenges

- Shared wireless channel
  - Facilitate message eavesdropping & injection
- Resource constraints of sensor nodes
  - Battery, memory, computation, communication …
- Very large network scale (n*100 or n*1000)
  - Impossible to monitor each individual node
  - Nodes are subject to attacks such as captures
- Vulnerable protocol design
  - Security is often overlooked
- ...

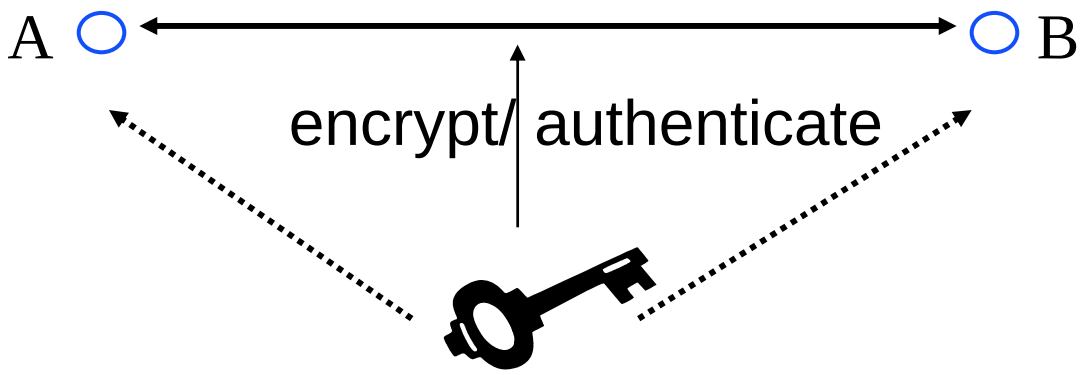
#1 Neighbor-to-Neighbor Authentication

- Two neighboring nodes verify that the other party is who it claims to be
  - Chan et al. (SP'03)
- Otherwise, attackers can
  - Inject false data reports via good nodes
  - Distribute wrong routing information
  - Impersonate good nodes to misbehave



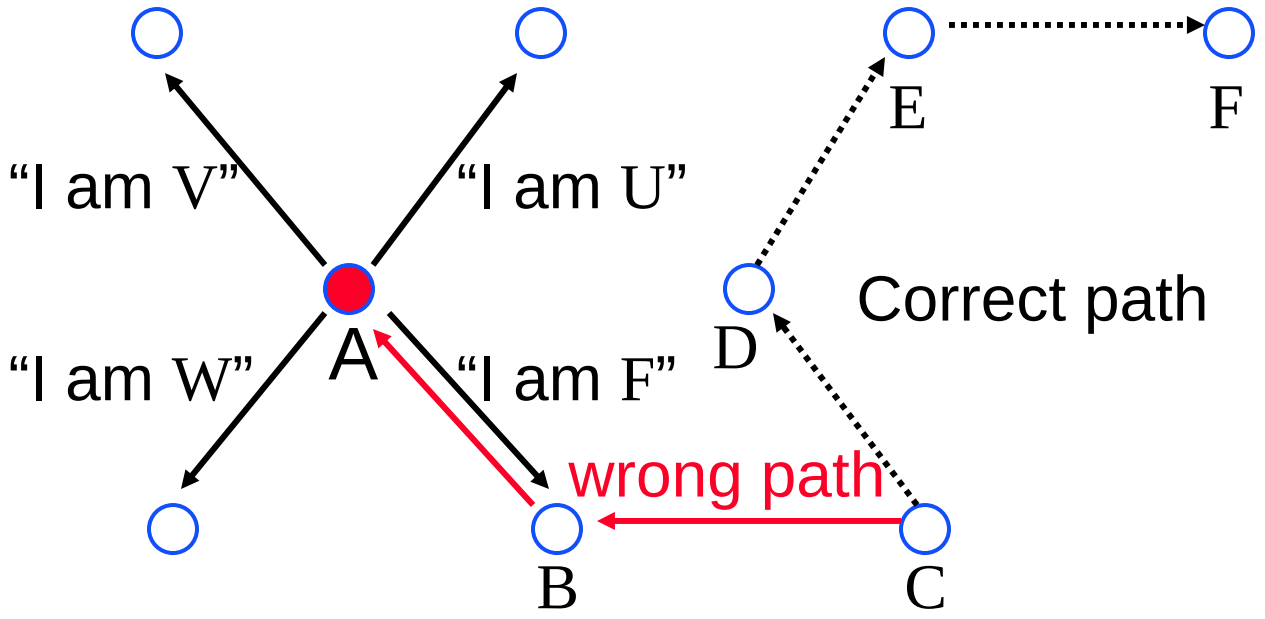
#2 Key Agreement

- Two neighboring nodes establish a shared secret key known only to themselves
  - Eschenauer and Gligor (CCS'03), Chan et al. (SP'03), Liu and Ning (CCS'03), …
- The shared key is a prerequisite for
  - Message encryption/decryption
  - Message authentication



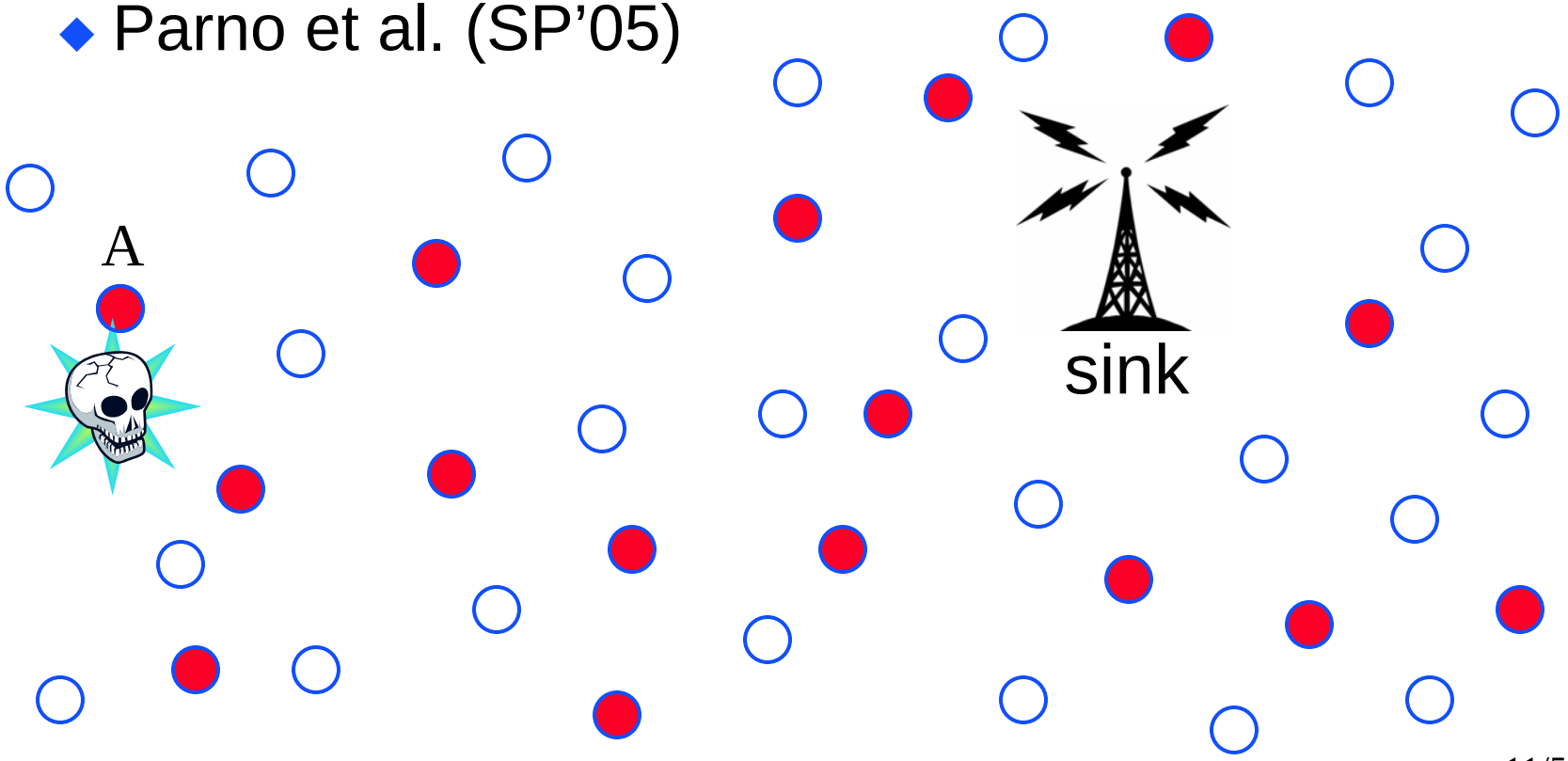
#3 Sybil Attack

- A malicious node claims multiple identities
  - Severely interrupt routing, fair resource allocation, distributed storage, misbehavior detection …
  - Douceur (IPTPS'02), Newsome et al. (IPSN'04)



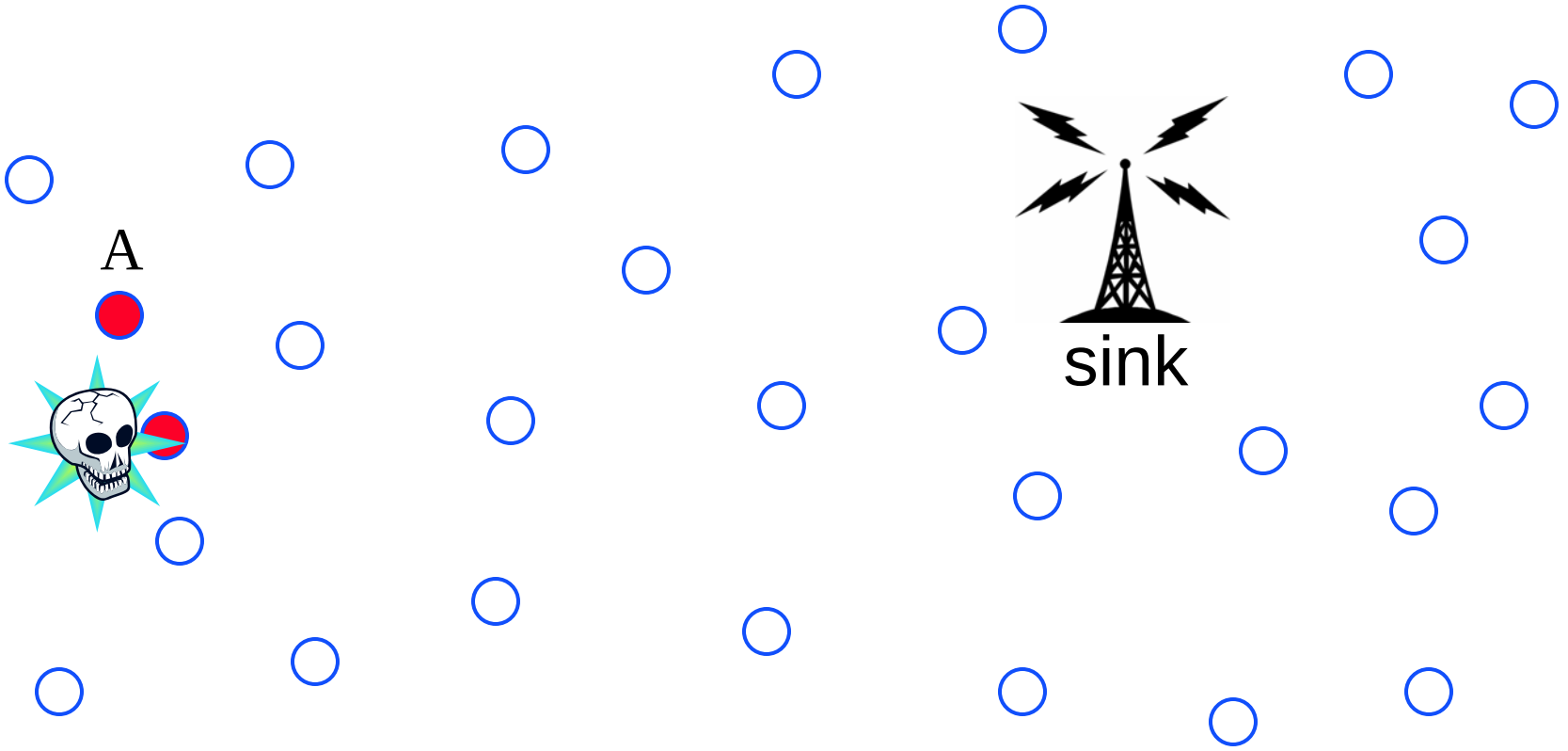
#4 Node Duplication Attack

- The attacker put clones of a captured node at random or strategic locations in the network
  - Parno et al. (SP'05)



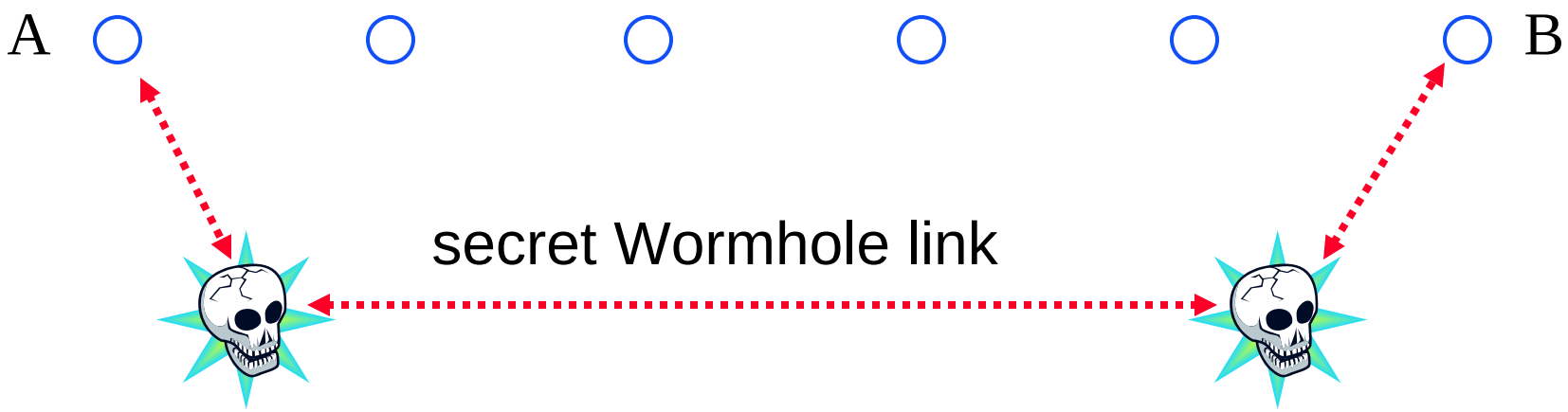
#5 Random Walk Attack

- The attacker uses secret information of a captured node to roam in the network



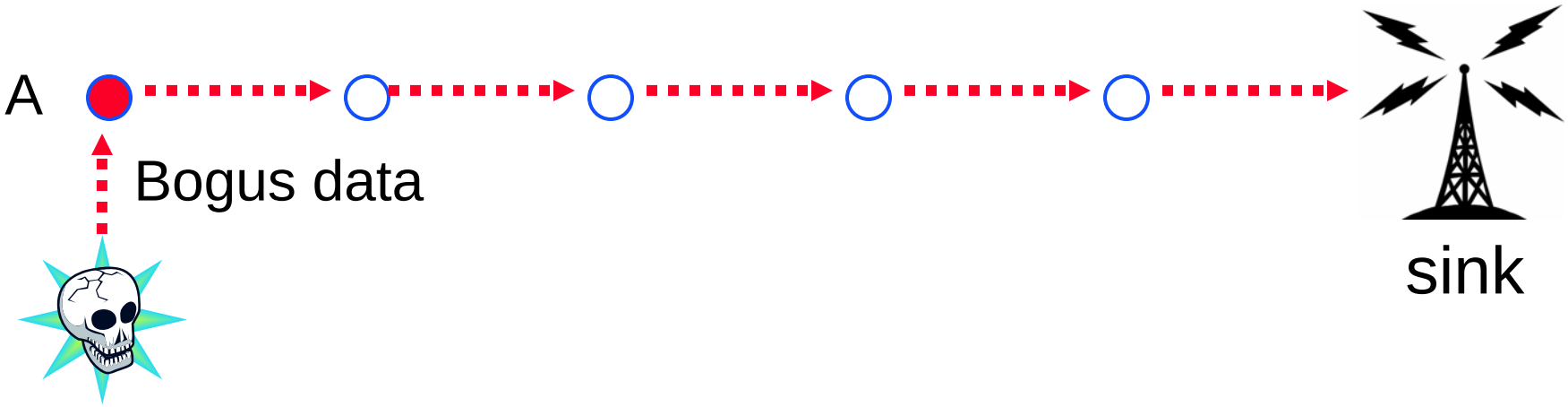
#6 Wormhole Attack

- Attackers tunnel packets received at one location to another distant network location
  - Hu et al. (INFOCOM'03), Karlof et al. (SNPA'03)
- Allowing the attacker to
  - Disrupt routing, selectively drop packets, …



#7 Data Injection Attack

- The attacker continuously injects bogus data into the network via a captured node
  - Ye et al. (INFOCOM'04), Zhu et al. (SP'04)
- Allowing the attacker to
  - Deplete scarce energy of sensor nodes
  - Cause network congestion & false alarms



Drawbacks of Prior Solutions

- Many separate solutions exist, but
  - Difficult to combine due to different or even conflicting underlying assumptions
  - Even if possible, far too complex a solution stack
  - Most prior solutions do not work when a small number of nodes are captured by attackers

Roadmap

- Introduction
  - Wireless sensor networks
  - Security requirements & challenges
- Security issues to tackle
- Our location-based solution
  - A unified, lightweight, attack-resilient solution that can limit the damage caused by captured nodes
- Conclusion & future work

Motivation

- Almost all WSN applications are location-dependent and require sensor nodes to know their own locations
  - E.g., military sensing and tracking
- Sensor nodes are stationary once deployed
  - Can be identified by their IDs plus locations
- Sensor nodes have a limited comm. range
  - Can only directly communicate with others inside their communication range

Notation

ID_A : node A's ID

L_A : node A's physical location

q : a large prime (≥ 160 bits)

G_1, G_2 : two cyclic groups of order q

s : a network master secret, $1 \leq s \leq q-1$

W : an arbitrary generator of G_1

W_p : $W_p = sW \in G_1$

H_1 : mapping inputs to non-zero elements in G_1

H_2 : mapping inputs to fixed-length outputs, e.g., SHA-1

R : common communication range of sensor nodes

Notation (cont'd)

$f : G_1 \times G_1 \rightarrow G_2$ (pairing), such that, $\forall U, V, S, T \in G_1$,

$$f(U+V, S+T) = f(U,S)\, f(U,T)\, f(V,S)\, f(V,T) \quad \text{(bilinear)}$$

$$\Downarrow$$

$$\forall a, b \in [1, q-1]$$

$$\begin{cases} f(aU, bV) = f(aU, V)^b = f(U, bV)^a = f(U, V)^{ab} \cdots & \text{(bilinear)} \\ f(U, V) = f(V, U) & \text{(symmetric)} \end{cases}$$

Boneh and Franklin (CRYPTO'01), Barreto et al. (CRYPTO'02)

Location-based Security Solution

- Location-based authentication
  - Neighbor-to-neighbor authentication
  - Key agreement
  - Sybil attack
  - Node duplication attack
  - Random walk attack
  - Wormhole attack
- Location-based threshold-signing
  - Data injection attack

Location-Based Keys

- Conventional way: ID-based keys
  - Name a node merely with its ID
  - Bind sensor nodes' keys only to their IDs
  - Vulnerable to many attacks, e.g., node duplication
- Our method: location-based keys (LBKs)
  - Name a node with both its ID and location
    - Grace@NJIT is more specific than Grace!
  - Bind sensor nodes' keys to both IDs and locations

Location-Based Keys

- Assume a secure way to decide node locations
  - Zhang, et al., JSAC'06
- Node A's LBKs:

$$\begin{cases} \text{Public key: } ID_A @ L_A \\ \text{Private key: } K_A = sH_1(ID_A @ L_A) \in G_1 \end{cases}$$

  - Given ($ID_A@L_A$, K_A), it is infeasible to derive s, as the Discrete Logarithm Problem is hard in G_1.
- Each node only knows its unique LBK pair, and has no knowledge of s

Neighbor-to-Neighbor Authentication

- Purpose
  - Discover and perform mutual authentication with neighboring sensor nodes
- Criteria
  - Check if the candidate is within the comm. range and has the correct location-based private key

Neighbor-to-Neighbor Authentication

Node A: $K_A = sH_1(ID_A @ L_A)$

Node B: $K_B = sH_1(ID_B @ L_B)$

$$\xrightarrow[\text{broadcast}]{ID_A @ L_A, n_A}$$

$$?\left\|L_B - L_A\right\| \le R$$

$$k_{B,A} = f(K_B, H_1(ID_A @ L_A))$$

$$\xleftarrow[\text{unicast}]{ID_B @ L_B, n_B, H_2(n_A \| n_B \| 1 \| k_{B,A})}$$

$$?\left\|L_A - L_B\right\| \le R$$

$$k_{A,B} = f(K_A, H_1(ID_B @ L_B))$$

$$?\, H_2(n_A \| n_B \| 1 \| k_{A,B}) = H_2(n_A \| n_B \| 1 \| k_{B,A})$$

$$\xrightarrow[\text{unicast}]{H_2(n_A \| n_B \| 2 \| k_{A,B})}$$

$$?\, H_2(n_A \| n_B \| 2 \| k_{B,A}) = H_2(n_A \| n_B \| 2 \| k_{A,B})$$

Neighbor-to-Neighbor Authentication

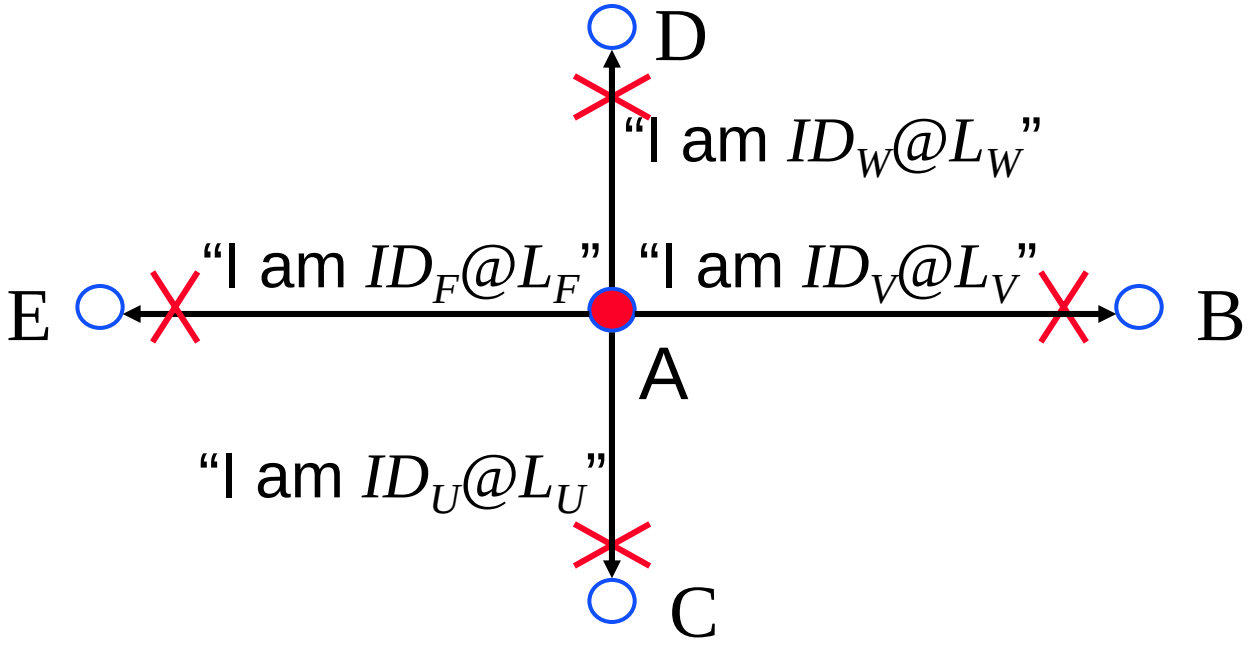
Node A: $K_A = sH_1(ID_A @ L_A)$ Node B: $K_B = sH_1(ID_B @ L_B)$

$k_{A,B} = f(K_A, H_1(ID_B @ L_B))$ $k_{B,A} = f(K_B, H_1(ID_A @ L_A))$

$$
\begin{aligned}
k_{A,B} &= f(K_A, H_1(ID_B @ L_B)) \\
&= f(sH_1(ID_A @ L_A), H_1(ID_B @ L_B)) \\
&= f(H_1(ID_A @ L_A), sH_1(ID_B @ L_B)) \quad \longrightarrow f \text{ is bilinear} \\
&= f(H_1(ID_A @ L_A), K_B) \\
&= f(K_B, H_1(ID_A @ L_A)) \quad \longrightarrow f \text{ is symmetric} \\
&= k_{B,A}
\end{aligned}
$$

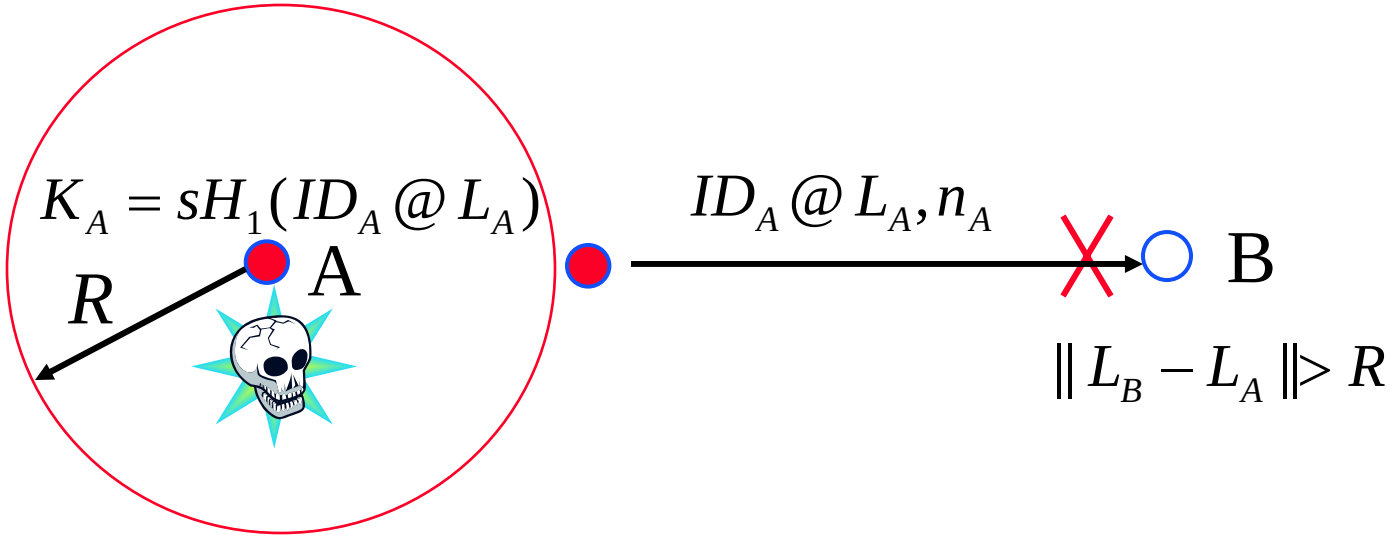
$$?\, H_2(n_A \| n_B \| 1 \| k_{A,B}) = H_2(n_A \| n_B \| 1 \| k_{B,A})$$
$$?\, H_2(n_A \| n_B \| 2 \| k_{B,A}) = H_2(n_A \| n_B \| 2 \| k_{A,B})$$

Resilience to Sybil Attack



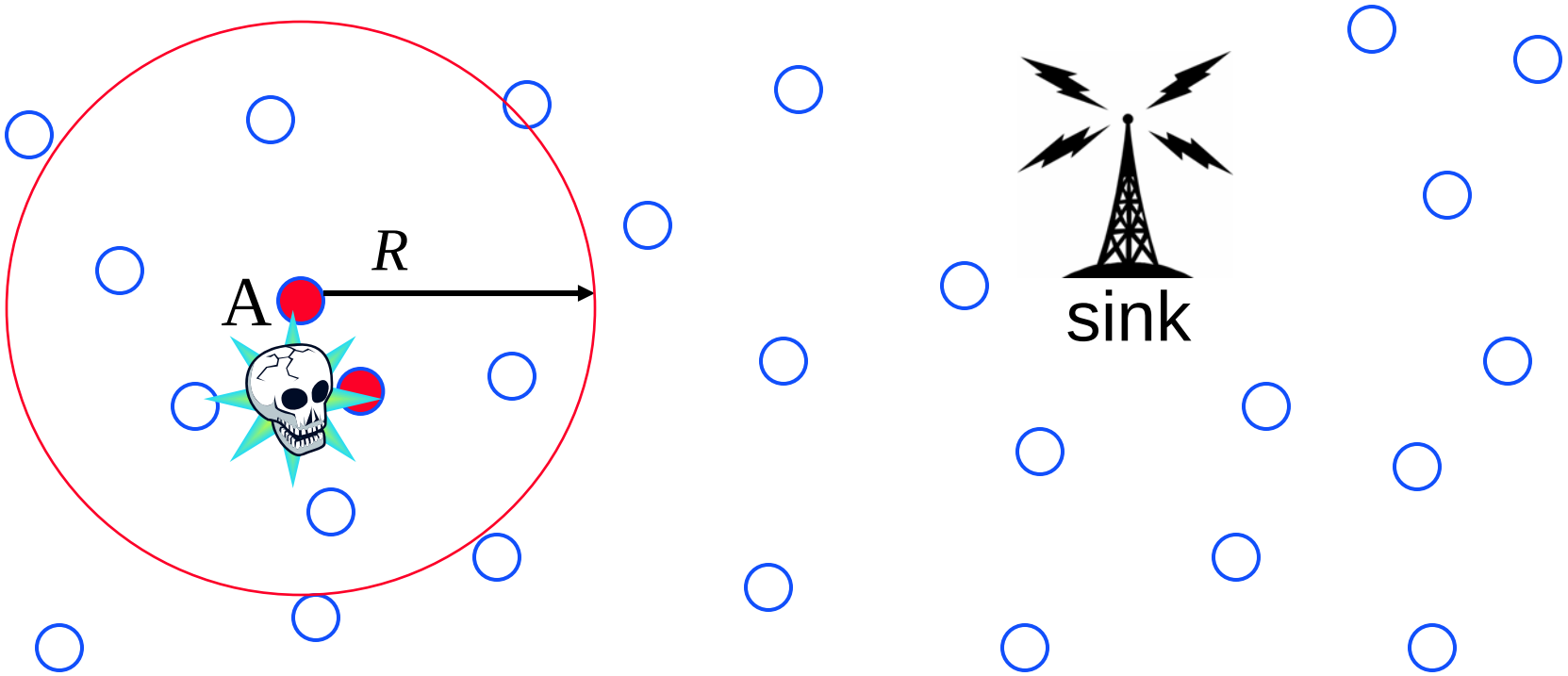
- The captured node doesn't have the correct location-based private keys of the nodes it claims to be
- Comparison to Newsome et al. (IPSN'04)
  - Our solution has much higher network scalability

Resilience to Node Duplication Attack



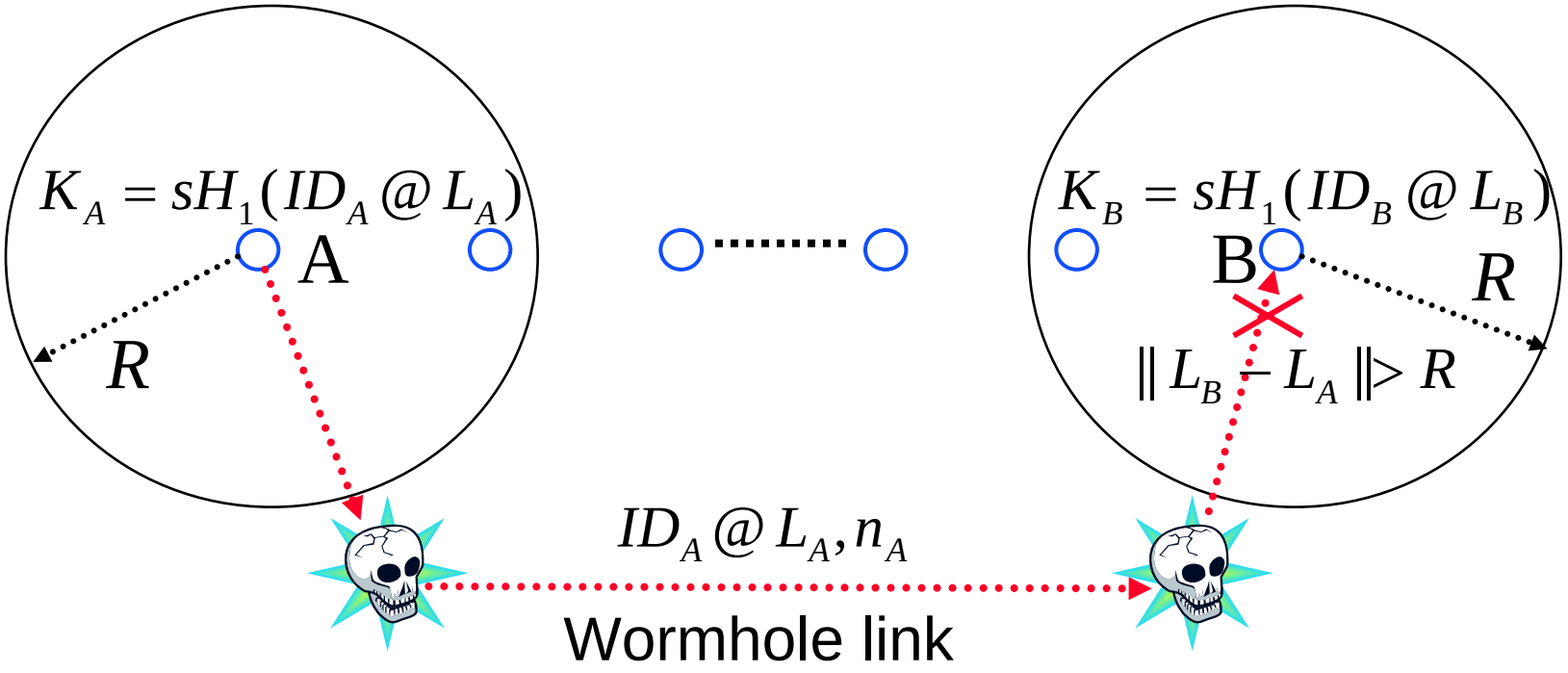
- A duplicate will be detected if talking to good nodes outside the communication range of node A
- The impact range of a captured node is reduced from the whole network to a small circle of radius < R
- Comparison to Parno et al. (SP'05)
  - Our solution is much more efficient in both communication and computation

Resilience to Random Walk Attack



- The impact range of a capture node is reduced from the whole network to a small circle of radius < R

Resilience to Wormhole Attack



- The wormhole attack is completely defeated
- Comparison to Hu et al. (INFOCOM'03)
  - Our solution has no stringent requirement on sensor hardware and time synchronization

Comparison to Prior Solutions

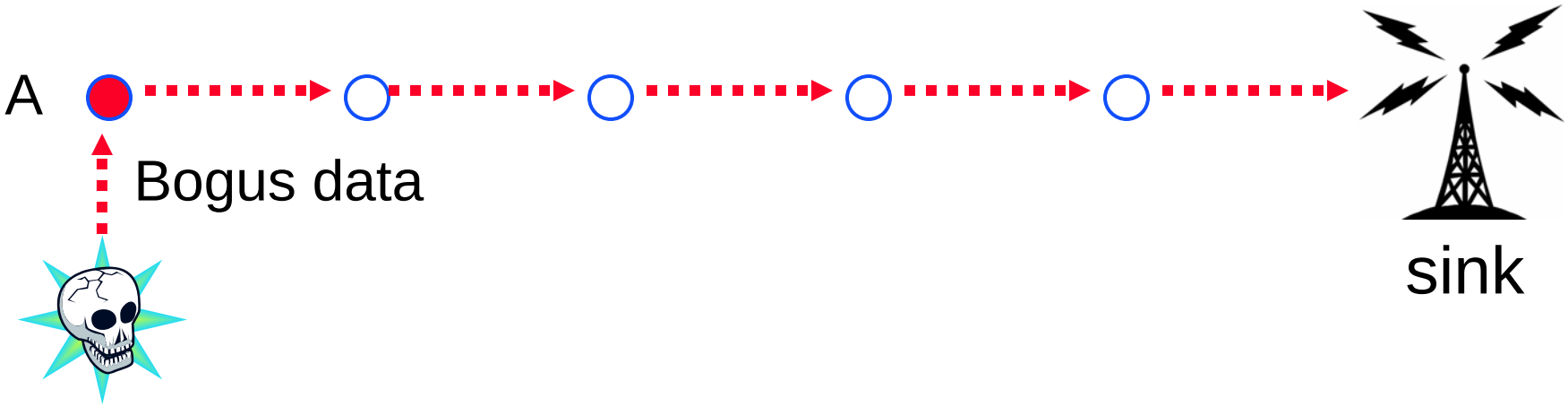
	Our scheme	**Eschenauer'02, Chan'03, Du'03, Liu'03 …**
Key agreement	Deterministic	Probabilistic
Neighborhood authentication	Yes	No or very limited
Support for digital signatures	Yes	No
Storage cost	Low	High
Network scalability	High	Poor
Attack resilience	High	Poor
Communication overhead	Low	High
Computation overhead	High	Low
Comm. + Comput. overhead	Low	High

Location-based Security Solution

- Location-based authentication
  - Neighbor-to-neighbor authentication
  - Key agreement
  - Sybil attack
  - Node duplication attack
  - Random walk attack
  - Wormhole attack
- Location-based threshold-signing
  - Data injection attack

Data Injection Attack

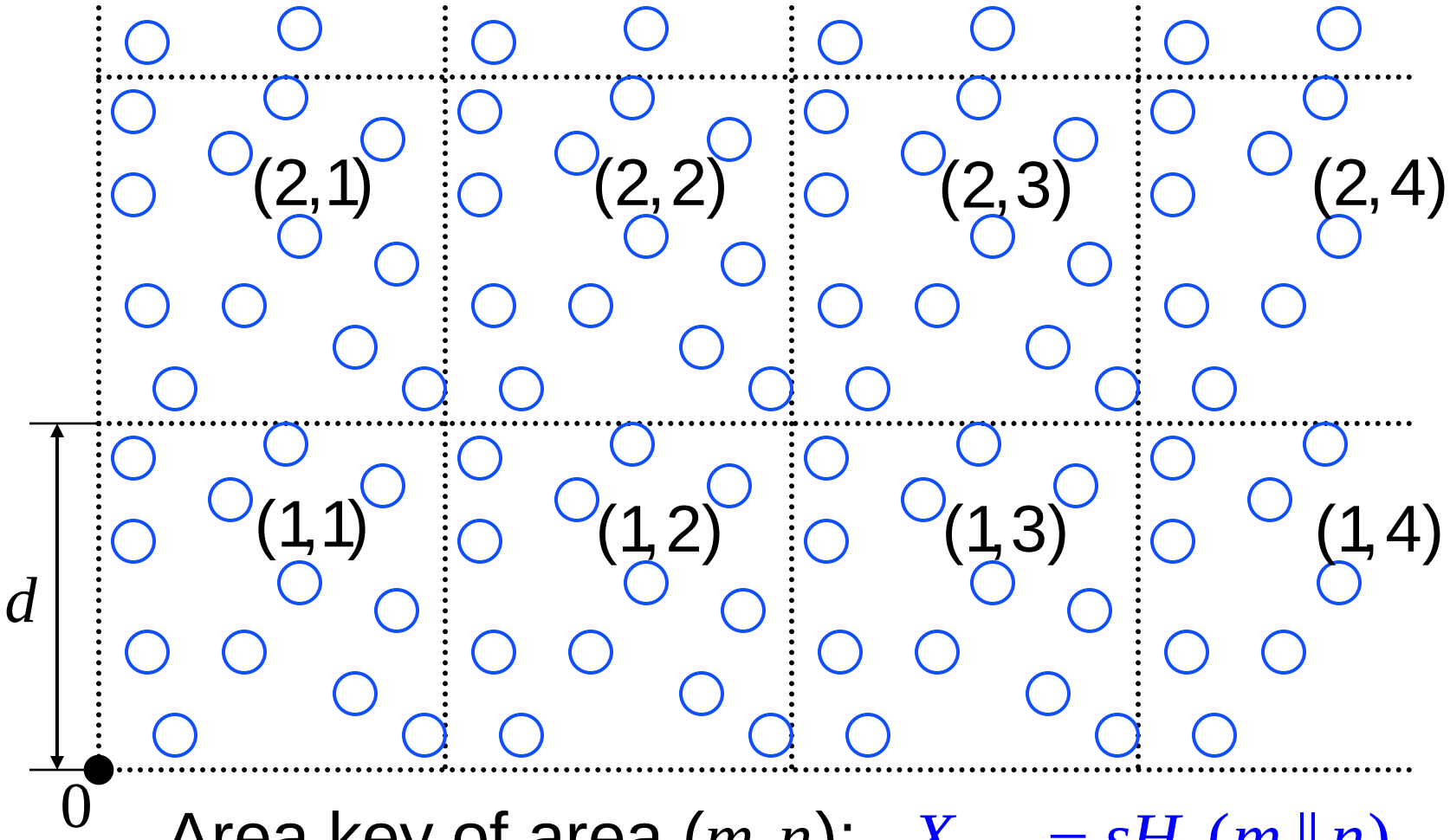
- The attacker continuously injects bogus data into the network via a captured node
  - Ye et al. (INFOCOM'04), Zhu et al. (SP'04)
- Allowing the attacker to
  - Deplete scarce energy of sensor nodes
  - Cause network congestion & false alarms



Location-based Threshold-Signing

- Observation
  - Each point in the sensor field should be covered by at least *k* sensor nodes; or each point should be within the sensing distance r of at least k nodes
  - The k-coverage problem (Kumar et al., MOBICOM'04)
- Basic idea
  - Each data report should be co-signed by t sensing nodes which generate it, where $1 \leqslant t \leqslant k$
  - Intermediate nodes drops data reports without correct threshold signatures

Cell Keys



Area key of area (m,n): $X_{m,n} = sH_1(m \| n)$

Secret-Sharing of Cell Keys

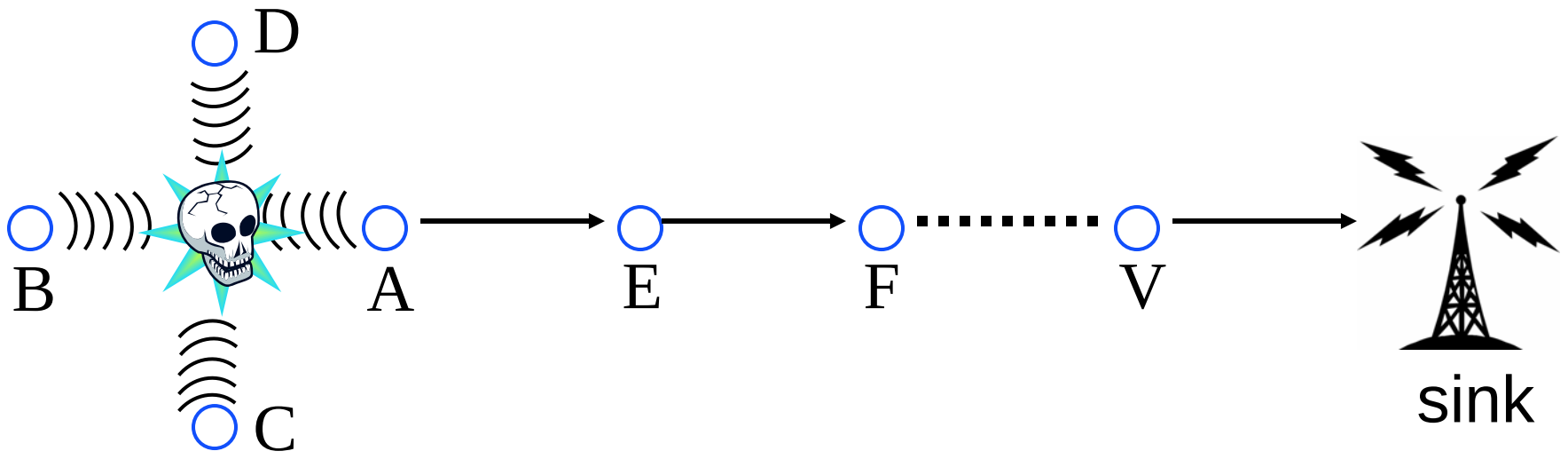
- Threshold secret-sharing of $X_{m,\,n}$
  - Each node in cell (m, n) holds a share of $X_{m,n}$
  - Any t nodes can recover $X_{m,n}$ to co-sign a data report originating from cell (m, n)
  - Any fewer than t nodes cannot do so

$$F_j \in G_1^*,\ j = 1, \cdots, t-1$$

$$X_{m,n}^i = X_{m,n} + \sum_{j=1}^{t-1} H_1(F_j \,\|\, m \,\|\, n)(ID_{m,n}^i \,@\, L_{m,n}^i)^j \in G_1$$

$$X_{m,n} = \sum_{i \in \Omega} \lambda_i X_{m,n}^i, \quad \text{where } \lambda_i = \prod_{k \in \Omega \setminus \{i\}} \frac{ID_{m,n}^k \,@\, L_{m,n}^k}{ID_{m,n}^k \,@\, L_{m,n}^k - ID_{m,n}^i \,@\, L_{m,n}^i}$$

Threshold-Signing



- $t = 4$
- The attacker is simultaneously detected by nodes A, B, C, D, all in cell (m, n)
- Nodes agree on an event report
- A is the local group leader

Threshold-Signing (cont'd)

Node A Node B or C or D

Select a random integer α

$$\theta = f(W,W)^{\alpha} \xrightarrow[\text{broadcast}]{}$$

$$\xleftarrow[\text{Unicast}]{} U_{m,n}^{B} = X_{m,n}^{B} H_2(\text{report} \| \theta)$$

$$\begin{aligned} U_{m,n} &= \lambda_A U_{m,n}^{A} + \lambda_B U_{m,n}^{B} + \lambda_C U_{m,n}^{C} + \lambda_D U_{m,n}^{D} + \alpha W \\ &= (\lambda_A X_{m,n}^{A} + \lambda_B X_{m,n}^{B} + \lambda_C X_{m,n}^{C} + \lambda_D X_{m,n}^{D}) H_2(\text{report} \| \theta) + \alpha W \\ &= X_{m,n} H_2(\text{report} \| \theta) + \alpha W \end{aligned}$$

Send <report, $U_{m,n}$, $H_2(\text{report} \| \theta)>$ to the sink

En-route Filtering of Bogus Reports

- Each intermediate node:

1. Deduce (m,n) from <report, $U_{m,n}$, $H_2(\text{report} \| \theta)$ >
2. Compute $\theta' = f(U_{m,n}, W)\, f(H_1(m \| n), -W_{pub})^{H_2(\text{report}\|\theta)}$
3. $\begin{cases} H_2(\text{report}\|\theta) = H_2(\text{report}\|\theta') \Rightarrow \text{the report is authentic} \\ H_2(\text{report}\|\theta) \neq H_2(\text{report}\|\theta') \Rightarrow \text{the report is bogus} \end{cases}$

En-route Filtering of Bogus Data

$$
\begin{aligned}
\theta' &= f(U_{m,n}, W)\, f(H_1(m \| n), -W_{pub})^{H_2(\text{report}\|\theta)} \\
&= f(X_{m,n} H_2(\text{report} \| \theta) + \alpha W, W)\, f(H_1(m \| n), sW)^{-H_2(\text{report}\|\theta)} \\
&= f(X_{m,n} H_2(\text{report} \| \theta) + \alpha W, W)\, f(sH_1(m \| n), W)^{-H_2(\text{report}\|\theta)} \\
&= f(X_{m,n} H_2(\text{report} \| \theta), W)\, f(W, W)^{\alpha}\, f(X_{m,n}, W)^{-H_2(\text{report}\|\theta)} \\
&= f(X_{m,n}, W)^{H_2(\text{report}\|\theta)}\, f(W, W)^{\alpha}\, f(X_{m,n}, W)^{-H_2(\text{report}\|\theta)} \\
&= f(W, W)^{\alpha} \\
&= \theta
\end{aligned}
$$

Probabilistic En-route Filtering

- Should a node always verify the report?
  - If the report is bogus → save energy
  - If the report is real → waste energy
- Solution: probabilistic en-route filtering
  - Each node verify a report with probability p_f
  - The sink always performs the verification
- On average, a bogus report passes

$$\overline{\mu} = \sum_{j=1}^{\infty} j p_f (1 - p_f)^{j-1} = 1/p_f \quad \text{hops}$$

Optimal Filtering Probability

L_n : original packet length (bytes)
L_o : packet overhead of our scheme (bytes)
E_{tr} : energy consumed to transmit & receive one byte
E_p : energy consumed by a pairing operation
β : average hops a real report travels towards the sink
ρ : ratio of bogus data traffic to real data traffic
p_f : en-route filtering probability

The normalized energy consumption of our scheme is

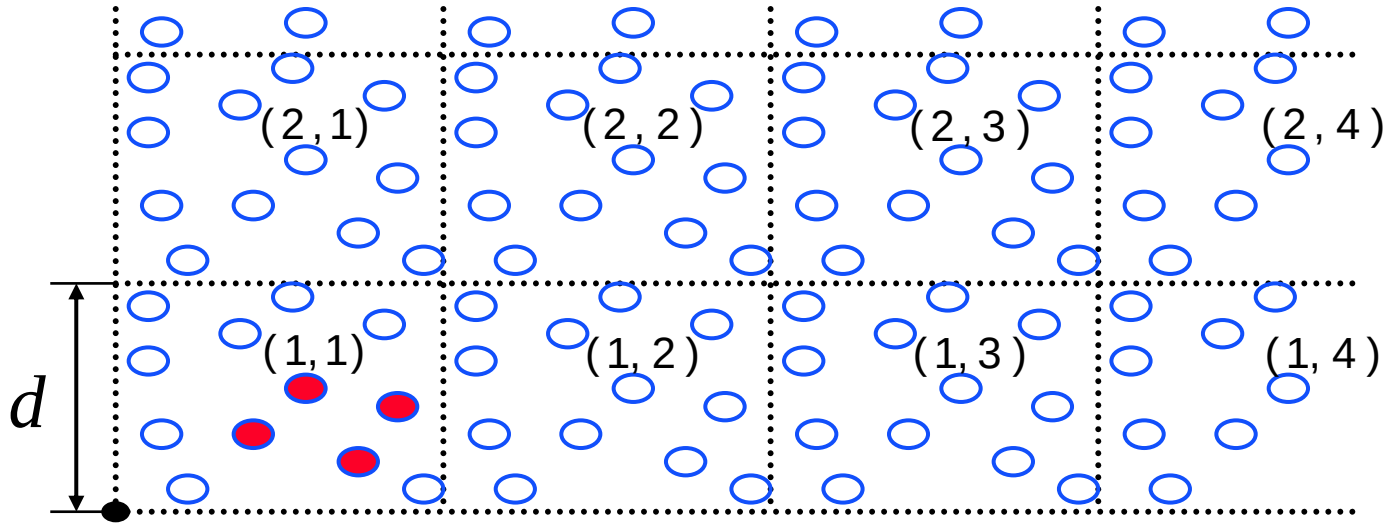
minimize $E'_{sum} = (L_n + L_o)E_{tr}(\beta + \rho\bar{\mu}) + (\beta p_f + \rho)E_p$

$$= (L_n + L_o)E_{tr}(\beta + \rho\frac{1}{p_f}) + (\beta p_f + \rho)E_p$$

$$\geq (L_n + L_o)E_{tr}\beta + \rho E_p + 2\sqrt{(L_n + L_o)E_{tr}\rho\beta E_p}$$

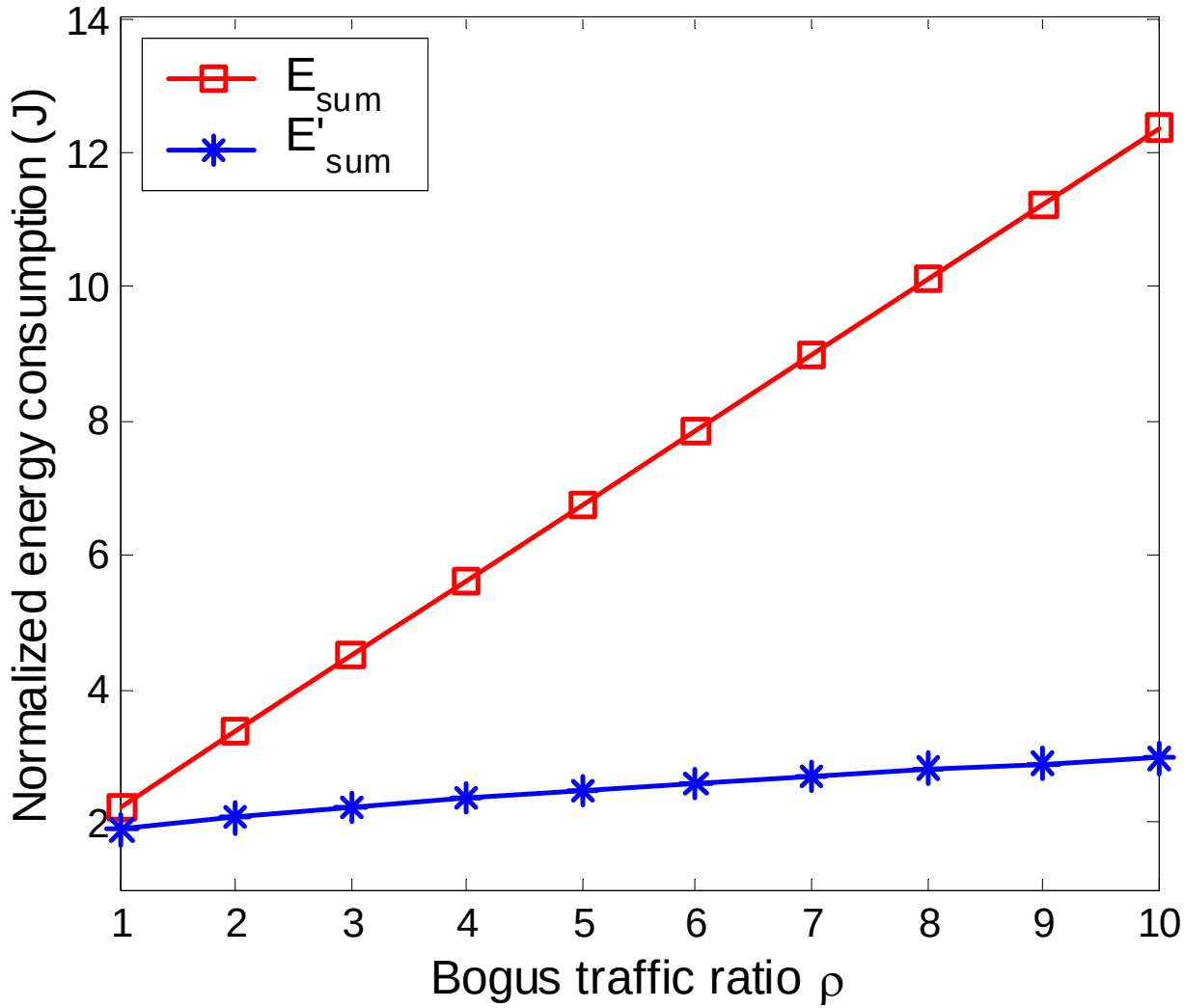
with equality iff $p_f = \sqrt{\frac{(L_n + L_o)E_{tr}\rho}{\beta E_p}}$

Security Analysis



- To inject bogus reports seeming to originate from cell (m, n), attackers must capture $\geqslant t$ nodes there
- Attackers cannot use a compromised cell key to fake reports seeming to originate from other cells
- Comparison to Ye et al. (INFOCOM'04), Zhu et al. (SP'04)
  - In both schemes, attackers can fake reports from any network place after capturing any t nodes in the whole network

Energy-Saving Performance



Conclusion & Future Work

- Proposed a location-based unified solution to
  - Neighbor2Neighbor authentication, key agreement, Sybil attack, node duplication attack, random walk attack, wormhole attack, data injection attack
- Plan to explore the applications of LBKs to
  - Intrusion detection
  - Secure distributed storage
  - Secure routing
  - Secure target tracking