Security and Privacy-Preserving Communication in Hybrid Ad Hoc Networks

Srdjan Capkun, Jean-Pierre Hubaux and Markus Jakobsson

Paper Outline

- Introduction
- System Model
- Privacy Goals and Challenges
- Overview of the Solution
- Privacy Preserving Routing
- Security and Performance
- Related Work / Conclusion

Introduction

 Objective is to provide both routing security and privacy preservation for hybrid ad hoc networks

Hybrid ad hoc network

- wireless ad hoc network + dual-homed (wireless/ wired) access points
- Access points provide connection to wired infrastructure (therefore reach & scalability)
- E.g. multi-hop Wi-Fi or cellular networks



Privacy features

- Anonymity
 - "the state of being not identifiable within a set of subjects called the anonymity set"
- Location Privacy
 - "ability to prevent other parties from learning one's current and past locations"

 Goal is to keep a node's identifier and location private from other network nodes

Introduction (cont'd)

Approach

- Use node pseudonyms and change frequently
 - Nodes should avoid being identified by:
 - the locations they visit
 - the type of traffic they generate
- Enforce user accountability via dynamic, but verifiable, cryptographic keys
 - Same keys that provide confidentiality, integrity, and authentication

Introduction (cont'd)

Contents of the Paper

- Present an overview of privacy threats
- Propose a scheme for secure and privacypreserving communication
- Present a quantitative analysis of privacy



<section-header><section-header><list-item><list-item><list-item><list-item><list-item><list-item><list-item>

Network Model (cont'd)

Assumptions

- All comms between nodes, and between a node and an AP, are wireless
- APs and mobile nodes have the same power range
- All links are bi-directional, i.e. any two communicating nodes must be in each others' power range
 - Some nodes will need to user other nodes as relays to reach an AP

Network Model (cont'd)

 All communicating nodes access the backbone in a multi-hop fashion

 Source node (S) transmits message (m) to destination node (D) via an access point (BS)

- S -> BS_s : uplink
- BS_s -> BS_D: inter-station
- BS_D -> D : downlink

Network Model (cont'd)

- Both uplink and downlink protocols are multi-hop, i.e. they require the participation of nodes on the route
 - These nodes are typically peers of the source and destination nodes

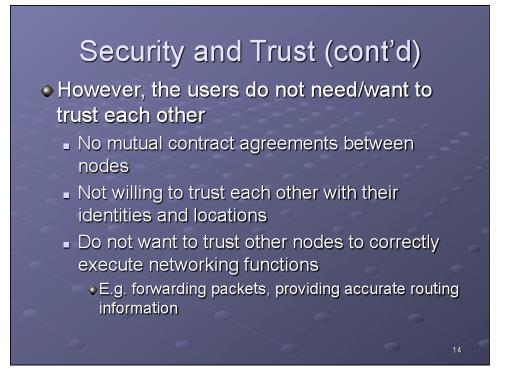
 All nodes in the control area are loosely time synchronized

Security and Trust

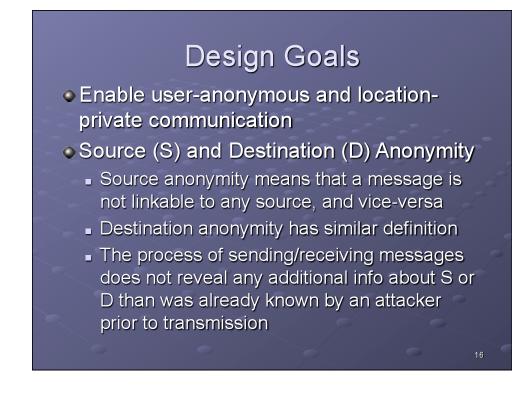
- Each mobile node has
 - A unique identifier
 - A secret key
- Both are known by the operator(s) of the BSs, but not by the other mobile nodes
- Contractual agreement between nodes and network operator
 - Access points monitor node behavior
 - Misbehavior can lead to service/network exclusion

Security and Trust (cont'd)

- Network membership includes:
 - Certificate of membership
 - In order to provide proof of membership to other nodes
 - Ability to uniquely sign a message
 - Other nodes can verify a legitimate node signed it
 - But only the network operator can identify who signed it
- This allows protocols to be secure and anonymous while holding users accountable for their behavior



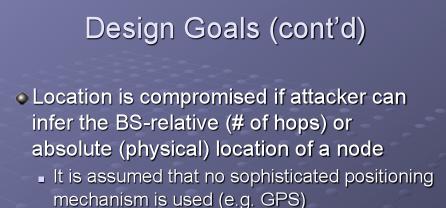




Design Goals (cont'd)

Strictly a need-to-know basis

- S needs to know the identity of D, but not its location
- The BSs need to know who S and D are (to verify membership) and their location (to route messages successfully)
- Nobody else (incl. the nodes on the route between S/D) should be able to infer identity or location of S/D



Design Goals (cont'd)

- Anonymity metrics
 - Anonymity set
 - Max. degree of anonymity is proportional to the size of the list of registered nodes
 - Assume a sufficiently large anonymity set
 - Entropy
 - Computed based on probabilities assigned to each identity
 - E.g. the probability that a given user is the message source
 - Both metrics are used in the analysis



Privacy Challenges (cont'd)

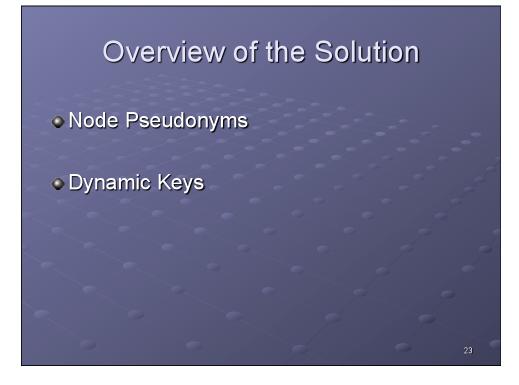
Threats (cont'd)

- Untrusted network operators
 - Can easily trace users and/or reveal their true identity
- Unique network/interface addresses & cryptographic keys
 - Use of static/unique addresses (e.g. MAC, IP) or crypto keys/certificates (e.g. Public Key) can facilitate user tracking



• Threats (cont'd)

- Radio fingerprinting
 - Radio transceivers emit signals with unique fingerprints that could be used for tracking
 - Also, static S/N can facilitate pseudonym mapping



Node Pseudonyms

- Each node shares a secret key with the BS
- Only the central authority (and the node itself) knows this key <u>and</u> the true identity of the node
- Node identity is protected via a pseudonym which changes over time:

$P_{S}(t) = HMAC_{K_{S}}(ID_{S}, t)$

Note that *t* is a time step design parameter, and different than a device timestamp

Dynamic Keys

A privacy-preserving key management scheme is proposed

 Control area-wide secret key schemes can protect identity but completely fail if a single node is compromised

Misbehavior is hard to isolate as well

Dynamic Keys

Dynamic public key scheme

- Each node holds a set of key pairs ...
 - $(PK^{1}_{A}/PrK^{1}_{A}, ..., PK^{n}_{A}/PrK^{n}_{A})$
- ... and certificates
 Cert^k_A = [PK^k_A, SIG_{PrKAuth}(PK^k_A)]
- Nodes use key pairs to establish symmetric secret keys with neighbors.
- Each time node changes pseudonym, it changes key pairs and symmetric keys.

Dynamic Keys

Update frequency

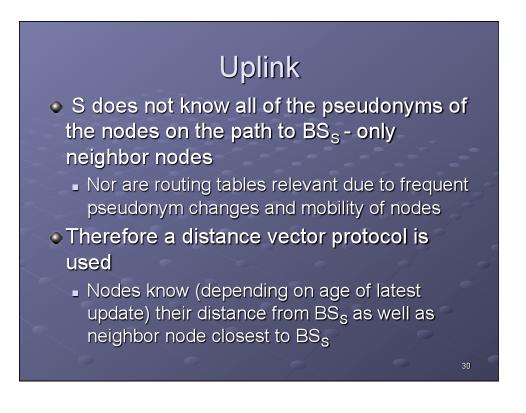
- Frequency of pseudonym and key changes is a design parameter (arbitrary)
- Can be temporal or event-driven (e.g. start of new session)
- Other factors that determine degree of privacy include node mobility and attacker strength
- Authors conclude that 1/min is sufficient for their scenario

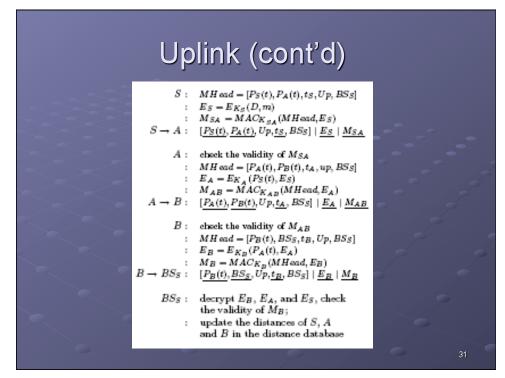
Privacy Preserving Routing

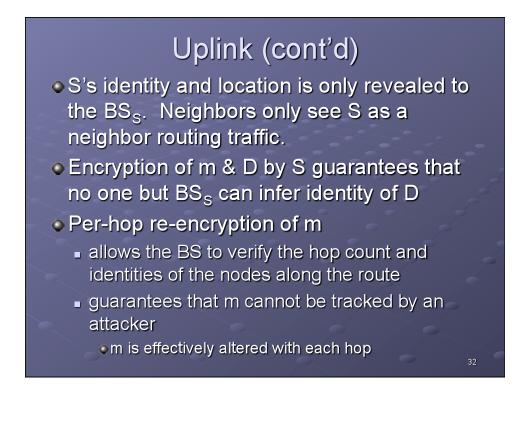
- Protocol Overview
- Uplink
- Downlink
- Inter-station Protocol
- Book-keeping

Protocol Overview

- Four sub-protocols are described
 - Uplink
 - Routing from S to BS_S
 - Downlink
 - ${\color{black}{\bullet}} \operatorname{Routing}$ from ${\color{black}{\mathsf{BS}}}_{{\color{black}{\mathsf{D}}}}$ to ${\color{black}{\mathsf{D}}}$
 - Inter-station Protocol
 - Routing between BSs
 - Book-keeping
 - Used by BSs to track node locations, pseudonyms, and network topology





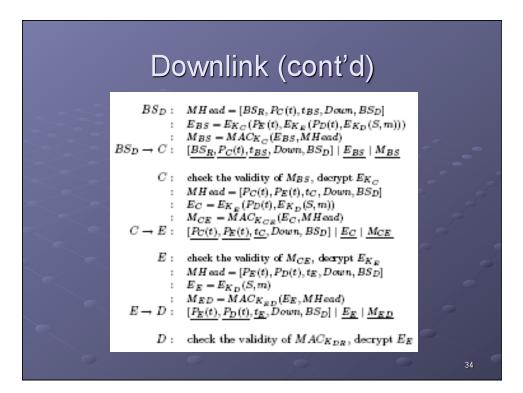


Downlink

 BS_D knows the optimal route to D hence a source routing protocol is used

BS_D performs the following:

- computes the current pseudonyms of the nodes on the route
- includes them in the packet
- sends the packet to the first node on the route



Downlink (cont'd)

Similar to Uplink

- D's identity and location is not revealed.
 Neighbors only see D as a neighbor routing traffic.
- Encryption of S & m by BS_D guarantees that no one but D can infer identity of S
- Per-hop packet content changes guarantees that m cannot be tracked by an attacker
- If the route is broken and delivery fails, it is reported to the BS – which updates route info and re-sends



 If S/D does not trust BS_S/BS_D, they will use their home networks, HN_S and HN_D respectively, to protect their identities

Inter-station Protocol (cont'd)

If D does not trust BS_D

- S's message will be first sent to HN_D (by BS_s ?)
- HN_D computes D's pseudonym and sends packet to the appropriate, but untrusted, BS_D
- BS_D than creates and routes the downlink packet to D, using D's pseudonym as the destination address

Inter-station Protocol (cont'd) If S does not trust BS_s Issue: How does S prove to BS_s and neighbor nodes that it is a legitimate node without

- revealing it's true identity ?
 - \bullet S uses existing dynamic public keys that are certified by HN_{S}
 - HN_S's public key needs to be certified by the untrusted network, NU
 - Since NU trusts HN_S (at least for charging purposes), S can be considered a legitimate node without revealing its identity
 - Alternatively, NU can issue a short-term cert to S

Book-keeping

 BSs keep records of the time, distances, identities, and pseudonyms of the nodes in their control areas

Associated key topics:

- Secure and Private Topology Discovery
- Topology Update
- Secure Time Synchronization

Book-keeping (cont'd)

• Secure and Private Topology Discovery

 Topology discovery is initiated by the BS via a discovery request

 $BS \rightarrow *: TREQ, rid, BS, t \mid SIGp_{rK_{BS}}(TREQ, BS, t)$

 Each receiving node forwards it to its neighbors if it has not seen the same request previously

Book-keeping (cont'd)

 Secure and Private Topology Discovery (cont'd)

- Receiving nodes then perform:
 - neighborhood discovery/update
 - neighbor authentication and key establishment
 - generates an encrypted neighbor list (pseudos, PKs) and sends it back to the node that forwarded the request
- Intermediate nodes merge the received information with their own and pass it on

Book-keeping (cont'd)

- Secure and Private Topology Discovery (cont'd)
 - BSs then perform the following:
 Verify the signatures of the nodes
 Match the PKs to users' real identities
 Reconstruct network topology
 - Note that only BS can decrypt neighbor lists successfully. Therefore intermediate notes can not observe or modify the topology information
 - Compromised node attacks must be mitigated by the BSs detecting topology inconsistencies₂

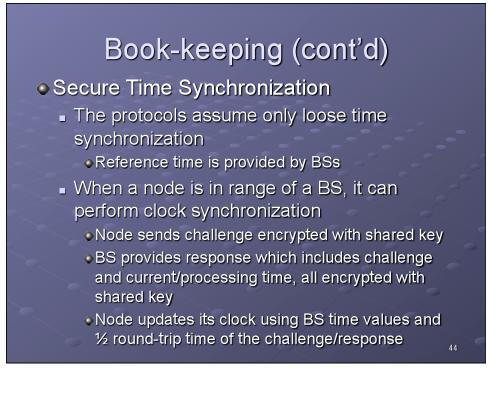
Book-keeping (cont'd)

Topology Update

Maintenance

 Nodes determine their distances from BS by collecting distance information from neighbors
 Protected by timestamps and shared secret keys

- Uplink
 - When BSs receive uplink packets, it will note the route taken and can update topology accordingly.
- Downlink
 - When nodes receive downlink packets, they can update their topology if the BS piggy-backs believed distances for the nodes on the route.



Book-keeping (cont'd)

Secure Time Synchronization

- Nodes can use neighbors to update clock as well
 - Node sends similar challenge to all of its neighbors
 - False time info can be detected unless majority of neighbors are compromised
 - Node can complain about other nodes providing false time info to BS

Since pseudos and PKs do not need to be changed very frequently, node clock differences can be as high as several seconds

Security and Performance

- Now to analyze the privacy-preserving scheme for performance and resistance to various attacks
- Topics include:
 - Attacker Model
 - Anonymity
 - Location Privacy
 - Security of Routing
 - Performance Analysis

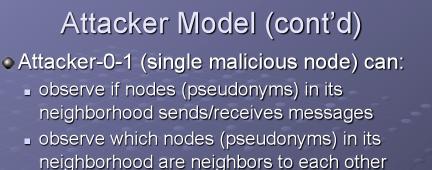
Attacker Model

Malicious node

 node controlled by a malicious adversary and cannot authenticate to a BS (& other nodes ?)

Compromised node

- Node controlled by a malicious adversary and can authenticate to a BS
- Undistinguishable from an honest node until misbehavior is detected
- Notation: Attacker-C-M



- observe signal-to-noise (S/N) ratios of the devices in its neighborhood and try to link each S/N ratio with a given node pseudonym
- detect signal watermarks of the devices in its neighborhood and link them with node pseudonyms

•24

Attacker Model (cont'd)

- Attacker-0-1 can also:
 - estimate how distant nodes in its neighborhood are from the access point (in term of number of hops), based on its physical distance to the access point.

Attacker-1-0 (compromised) can also:

- observe accurate pseudonym distances to the access point of the nodes (pseudonyms) in its neighborhood
- modify network traffic or generate traffic to infer nodes' locations or real identities.



Attacker-c-m can also:

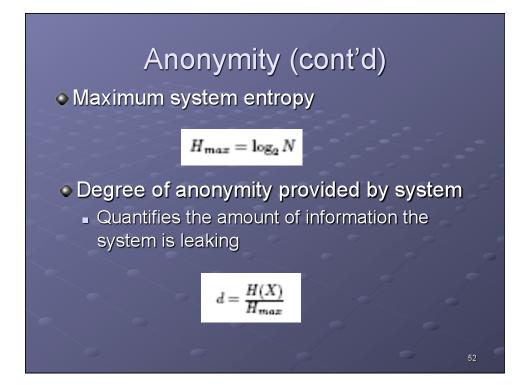
- observe/generate intelligence-gathering traffic on a wider network area
 - would further facilitate inference of users' real identities and locations
- May be able to send a message to D and track the message to find D's location
 However if BS_S <> BS_D this is pretty hard

Anonymity

 Analyze the level of source and destination anonymity achieved by the scheme based on entropy

$$H(X) = -\sum_{t=1}^{N} p_t \log_2 p_t$$

H(x) = the entropy of the system after attack $p_i = Pr(X=i)$, X is a discrete random variable i = an element of the anonymity set (a node) N = size of the node set



Anonymity (cont'd)

 Recall S/D anonymity as the property that a particular message is not linkable to any S/D, and vice-versa

Two important aspects to be analyzed

- Anonymity of node pseudonyms
 - linkability of messages to pseudonyms and their PKs

Mutual linkability of node pseudonyms

 ability for an attacker to link two or more pseudonyms to a particular node



Anonymity of node pseudonyms

- Attacker-0-1 can observe behavior of a transmitting neighbor node, P_A(t) -> P_B(t), and determine:
 - $\circ P_A(t)$ is either the source or just a forwarding node
 - $\bullet P_B(t)$ is not the source
 - Other neighbor nodes are probably not the source

Anonymity (cont'd)

 Denoting p(X=P_A(t)) = p_A, an attacker could assign source probabilities as follows:

- p_A = 1/s, s <= N, N = the # of possible sources
- ∎ p_B = 0
- $p_1 = \dots = p_{k-1} = 0$, k = # of attacker's neighbors
- $p_i = (1-p_A) / (N-k-1)$

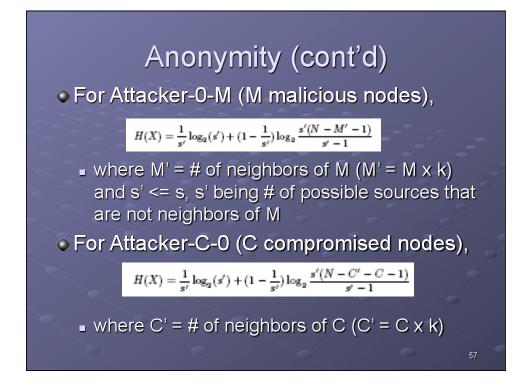
Anonymity (cont'd)

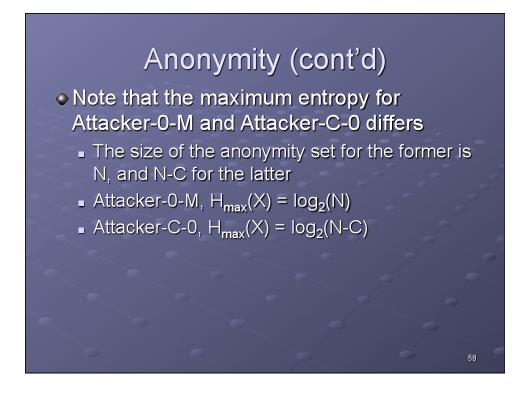
 If P_A(t) is located close to BS, almost any node in the control area could be source

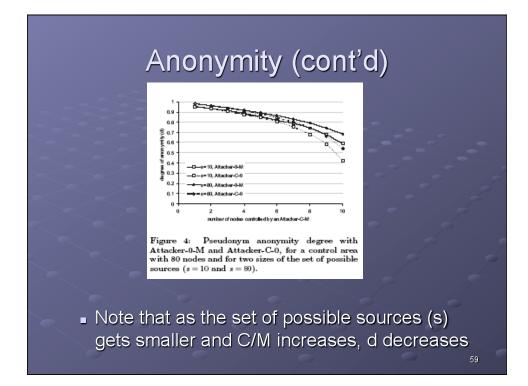
 If P_A(t) is located on the edge, only a few nodes could be source

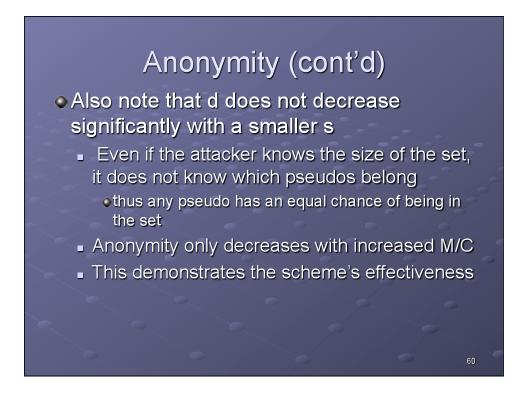
• For Attacker-0-1,

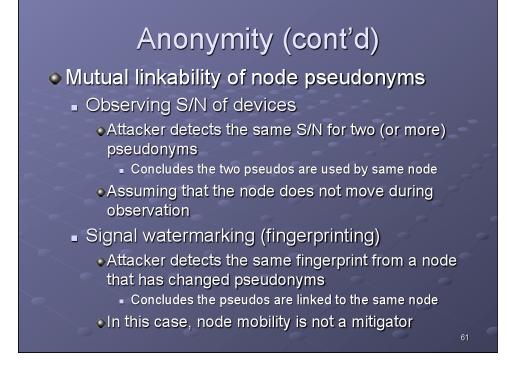
$$H(X) = \frac{1}{s} \log_2(s) + (1 - \frac{1}{s}) \log_2 \frac{s(N - k - 1)}{s - 1}$$











Location Privacy

 Both S/N and fingerprinting can also be used to track node locations

- Once attackers can map node movements to pseudonyms
- Signal watermark randomization can mitigate fingerprinting
- By installing a large # of nodes across the control area, the attacker can track pseudos and correlate them by location

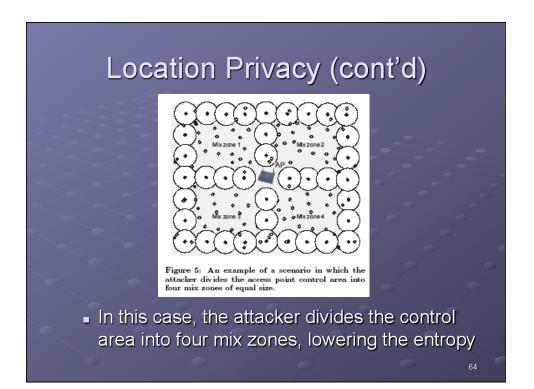


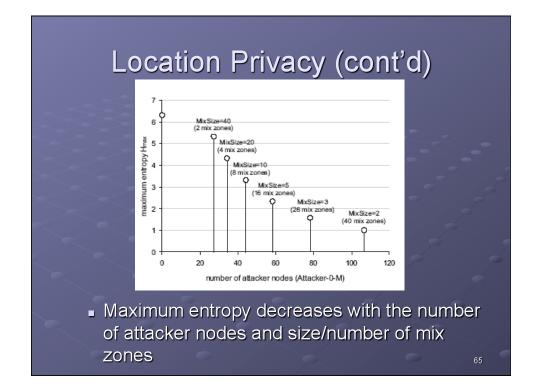
Mix zones

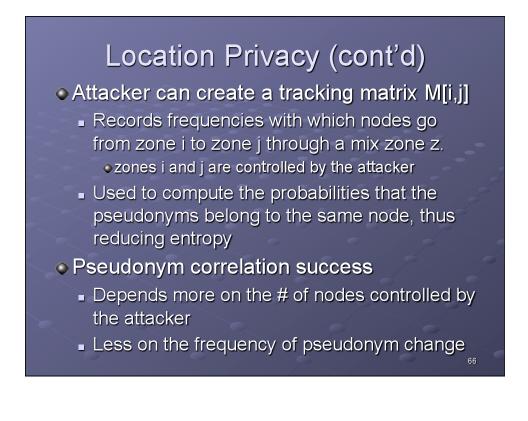
- A connected spatial region of a maximum size in which none of the nodes are in the power range of any of the nodes controlled by the attacker
- For each pseudo pair $P_X(t_1)$ and $P_Y(t_2)$, the probability X = Y is:

$Pr(P_{\chi}(t_1), P_{\gamma}(t_2) : X = Y) = 1/MixSize$

 If the attacker can divide control area into smaller mix zones, the entropy drops and it is easier to correlate pseudonyms







Location Privacy (cont'd)

Frequency of pseudonym change

- Needs to be only 2x higher than the average frequency a node moves from an attackercontrolled zone to a mix zone
- Estimated to be 1/t(r)

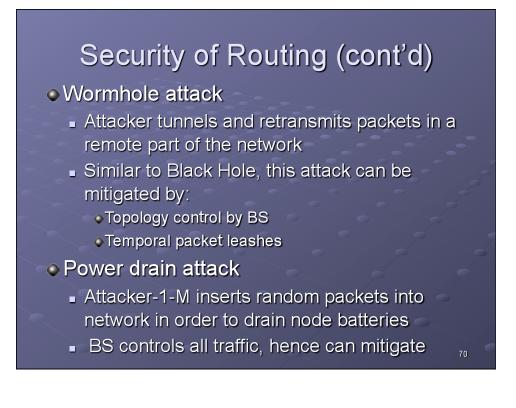
•where t(r) is the average time that it takes a node to cross the distance equivalent to the power range.

Security of Routing

- How resistant is the protocols to various attacks ?
- False distance information dissemination
 - Attacker claims it is closer to or further from BS than it really is
 - Cannot be performed by Attacker-0-M
 - Attacker-1-M will be easily detected as nonneighbor nodes will report different distances
 - Attacker-C-M could be successful if C is sufficiently large to fake the whole topology without being detected by the BS

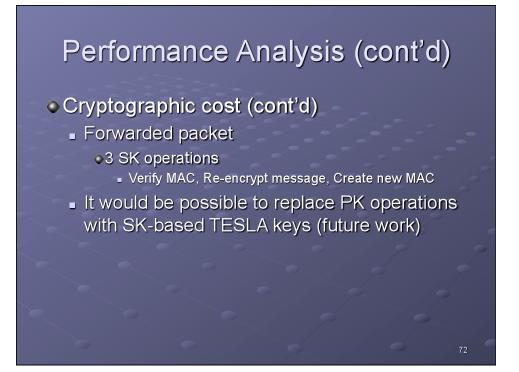
Security of Routing (cont'd)

- Black Hole attack
 - Attacker advertises close proximity to BS, then gathers/drops packets
 - Can be detected similarly to false distance
 - Can also be mitigated by nodes randomizing their choice of next uplink hop
 - Black hole can't paralyze an entire hybrid network as it can for MANETs
 - Only a fraction of its neighboring nodes in its control area



Performance Analysis

- Analyze the cryptographic and communication costs associated with the scheme
- Cryptographic cost
 - Routing is secured by symmetric key (SK)
 - Dynamic key establishment is by public key (PK)
 - SK establishment between 2 nodes
 1 PK signature & 1 PK signature verification per node
 1 PK signature verification of authority's certificate
 - SK updates are of minimal impact
 Fixed cost, seldom performed (1/min or less)



Performance Analysis (cont'd)

Communication cost

- Dynamic Key Update cost
 - PK update cost depends on frequency
 - Low, for this scheme
 - BS sends one certificate to each node at the same frequency at which keys/pseudos are updated
- Secure and Private Routing cost
 - Low, a single MAC is added to each message on its way to and from BS

Related Work

- Existing research efforts related to:
 - Hybrid ad hoc networks
 - Secure Routing
 - Anonymity and Location Privacy
 - Anonymous Credentials

See paper for details

Conclusion

 Proposed a scheme to secure and protect the privacy of communication in hybrid ad hoc networks

- Both security and privacy preservation can be integrated in the same protocol
- Privacy preservation provided by the use of pseudonyms and dynamic key renewal
- Detailed description of the Privacy Preserving Routing protocol and associated overhead/ robustness

Discussion

- Performance analysis is high-level and theoretical.
 - Simulation would provide harder data
- No discussion of how nodes determine destination node identity (D)
 - Perhaps a service database is available that uses pseudos instead of true identity

Discussion (cont'd)

• All message traffic must go through BSs

No trust mechanism is provided between S & D

S & D do not need to share a key
All trust is provided via BS/HN

