Wireless Network Security
Spring 2016

Patrick Tague

Class #13 – Network Privacy & Anonymity

Class #13

- Privacy risks at the wireless network layer

- Several different approaches in different systems / scenarios

Privacy and Anonymity

Network Privacy Issues

- Network layer interactions in wireless networks often expose information about identity, context, content, relationships, etc.
- In certain cases, cryptographic protections can help, but not always
- In certain cases, pseudonyms help, but not always

ID Matching

- Network IDs/addresses can facilitate tracking, profiling, inference, etc.
  - Ex: a network service provider sees device A connect to a network in Pgh, then to another network in DC, then to another network in SF → the service provider can create a profile of the device owner
  - Ex: an eavesdropper sees device A show up and connects to a network at the same time every day → the eavesdropper can temporally profile the user to learn when they will be away from home

Traffic Analysis

- A curious or malicious party can observe network traffic and analyze flow patterns to infer relationships
  - Plaintext IDs can make this pretty easy
  - Something like “conservation of flow” can allow traffic flow decoupling
  - Inference capability depends on several factors:
    - Network visibility – global or local view?
    - Traffic density – dense or sparse traffic distributions?
    - …

Timing Analysis

- Since network operations are typically at least somewhat delay sensitive, there are end-to-end correlations between transmission events
  - Ex: node A transmit 10 packets, then neighboring node B transmits 10 packets of similar size → maybe B is relaying A's traffic
  - Depending on visibility and density, very little other information is needed (e.g., strong hop-by-hop packet re-encryption doesn't prevent timing analysis)

Understanding the Risks

- What type of network? Services? Etc.?
  - WLAN, cellular, VANET, WSN, …
- What is the attacker's goal / purpose?
  - Real-time tracking, recovering past traces, …
  - Robbery, personal safety, blackmail, mal-marketing, surveillance, …
- What granularity is needed for attack success?
  - Relational, location-specific, region-specific, …

Privacy Challenges

1. Understanding the privacy goals
   - What needs to be protected?
   - What are the rules to be enforced?

2. Understanding the threat
   - What are attackers goals, capabilities, methods, …?
   - Practicality of attacker assumptions?

3. Metrics
   - How to measure privacy protection and enforcement?
   - How to evaluate and incorporate risk?

Different Privacy Concerns

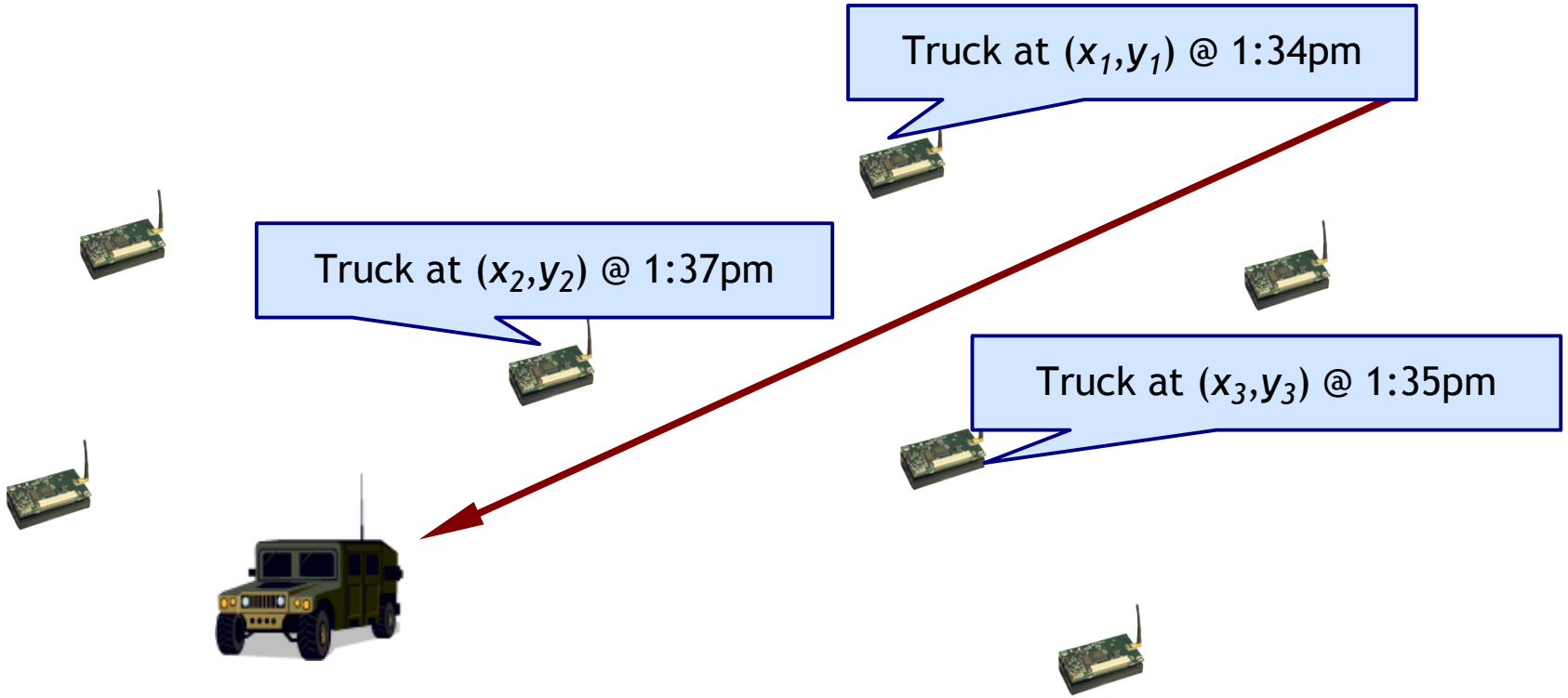
- Profiling and tracking WiFi users
  - We already talked about this one a bit
- Event/object inference in WSN
- Unauthorized user/car tracking in VANET

Traffic Anonymization

- In multi-hop networks (MANET/WSN), transmission linking can expose what path is used for a session
  - Traffic analysis:
    - Analyzing the flow of packets through a network (with global knowledge) allows decomposition into individual flows
  - Local traffic analysis:
    - Without global knowledge, timing information can expose flow decomposition in a neighborhood

WSN Location Privacy

- In sensor networks, we're usually not concerned with protecting sensor locations, but what they're sensing may be more sensitive



Source Location Privacy

- One of the common goals in WSN is to hide the location of the sensed event from an observer
  - But, the traffic generated will immediately expose any singular event
  - Commonly called the "Panda Hunter Problem"
    - Sensors in a wildlife area are used to track/study pandas
    - Whenever a panda walks by a sensor, it generates traffic
    - A hunter can track the traffic to find the panda

Panda Hunter Problem

- Objective of the WSN / defender:
  - Properly / quickly collect panda mobility info
  - Hide the location information from the panda hunters that can eavesdrop on WSN traffic but not decrypt

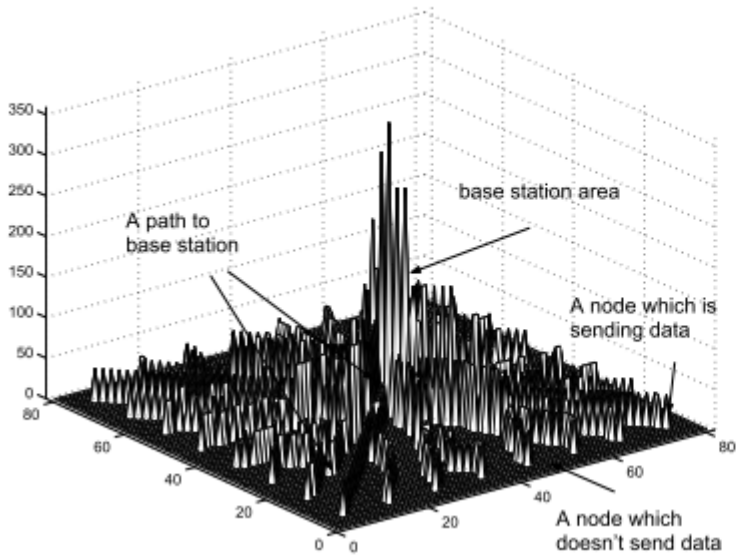
- Objective of the panda hunters:
  - Learn the location of the data source (and thus the panda) by analyzing traffic flow statistics

Panda Hunter Strategies

- Two approaches:
  - Choose one location in the network to monitor traffic
    - Wait for the panda to walk somewhere that creates traffic flows through the chosen location, then find the panda
    - Probably takes a long time depending on the area, but better than naïve hunting
  - Find the base station and monitor all network traffic
    - More work to find the base station, more traffic to analyze all at once, but any panda-related traffic goes here

Anti-Analysis Methods

- In the Panda Hunter context, there are two ways to mitigate the attack:
  - Prevent the hunter from finding the base station (i.e., destination location privacy)



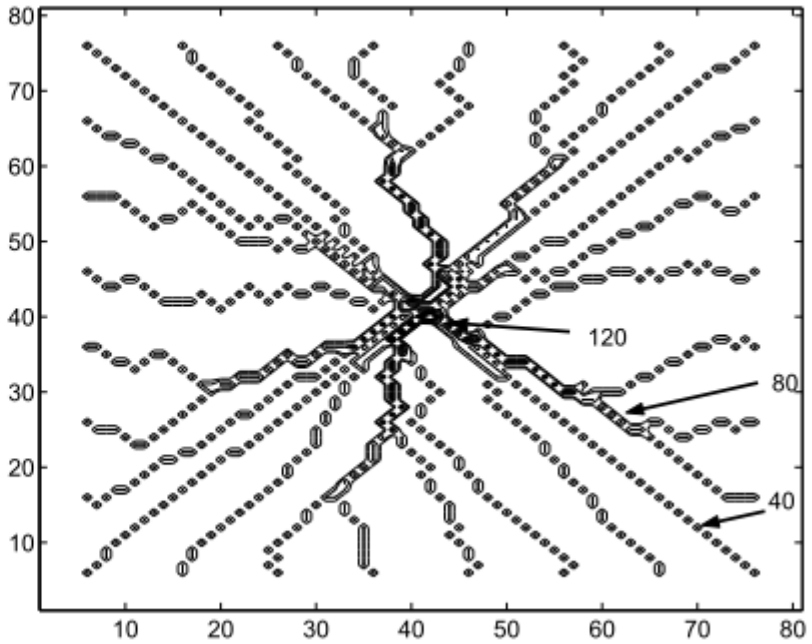


Image from [Deng et al., PMC 2006]

Flooding

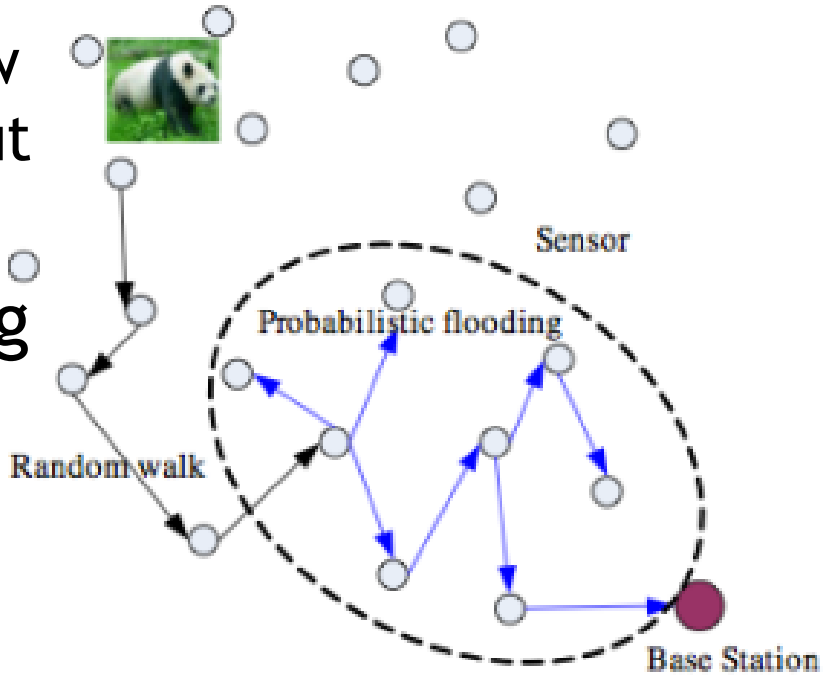
- One common approach is to hide the actual event data in dummy ("chaff") traffic
  - Flooding the network with dummy traffic prevents the attacker from figuring out what is real
    - If it looks like the panda is everywhere, where is it?
  - Of course, flooding dummy traffic is a lot of work for very little reward

Probabilistic Flooding

- Trade-offs can be made between the overhead of flooding and the resulting location privacy by instructing each node to forward dummy traffic only with probability p
  - Less dummy traffic slightly degrades privacy
  - Less dummy traffic means lower overhead
  - Nodes need to be able to distinguish dummy from real traffic, or also drop real traffic w.p. (1-p)

Random Routing

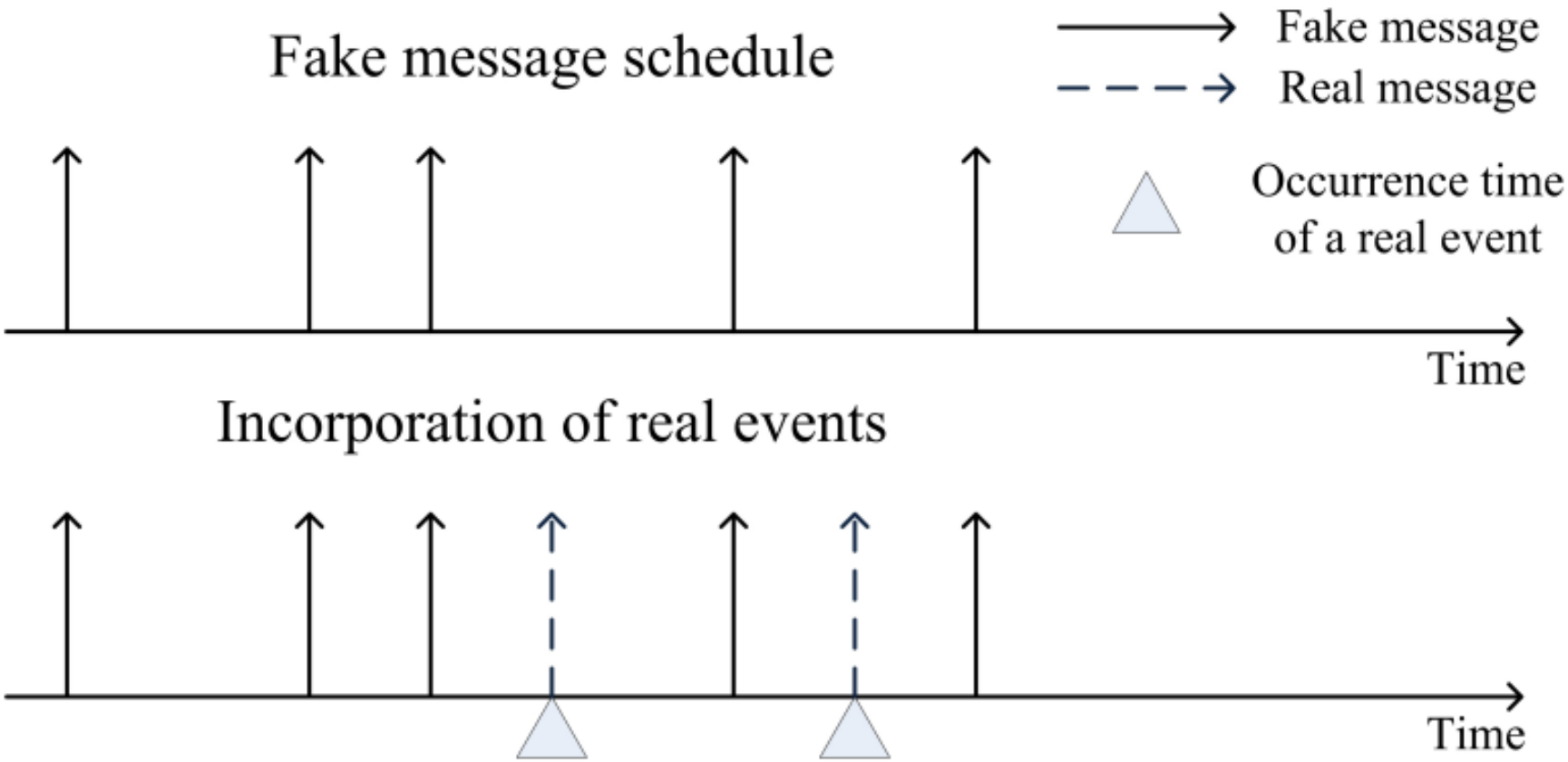
- Another technique to mitigate traffic analysis is random routing
  - Next hop ← rand({neighbors})
  - Non-deterministic packet flow makes the analysis harder, but increases delay
- Can combine random routing with prob flooding
  - Phantom Routing:



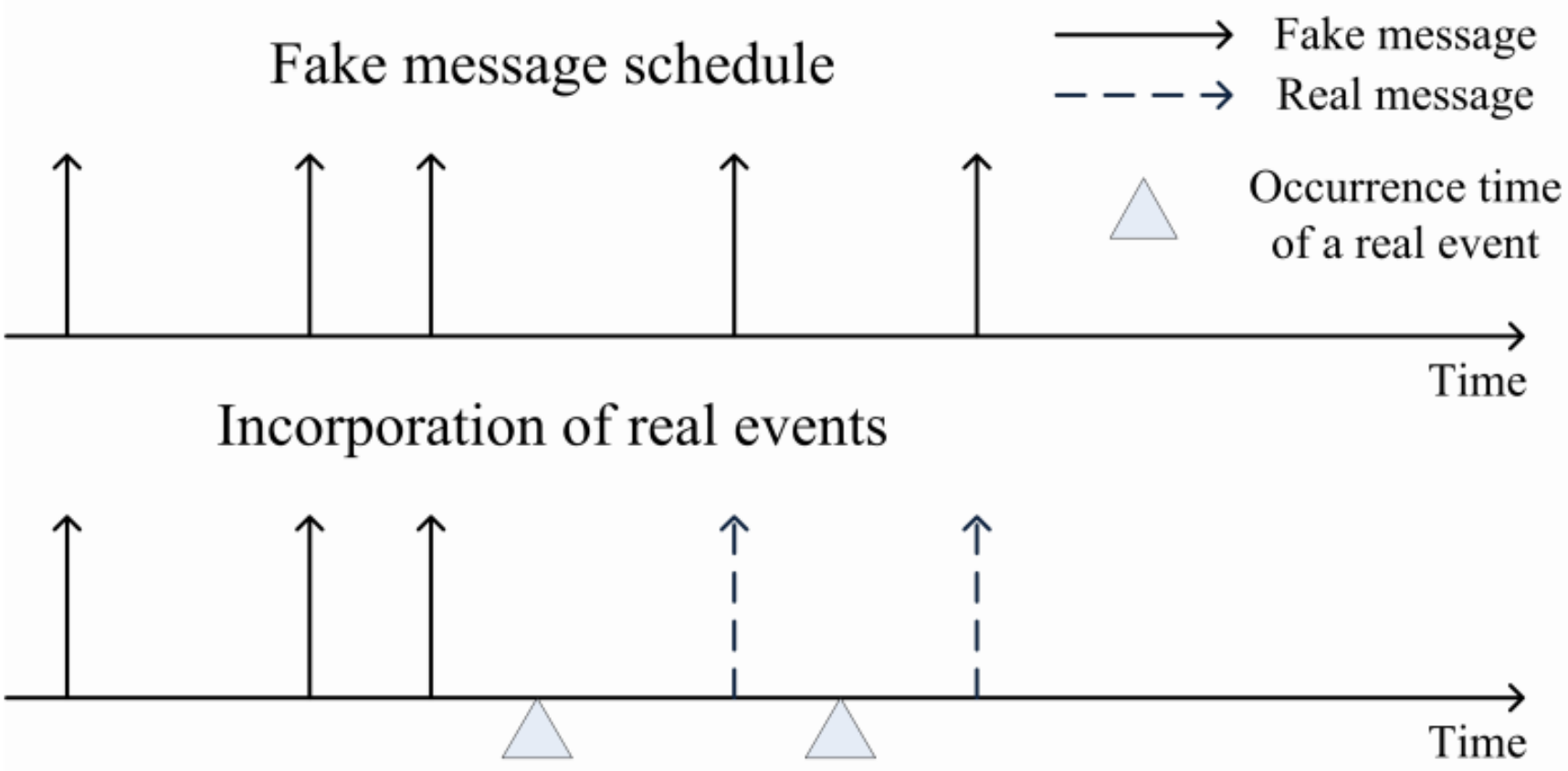
Transmission Correlation

- To make things harder, attackers can analyze timing at a node to further decompose flows at a point
  - Sequence of transmissions by two neighboring nodes can indicate re-transmissions → data on same path
  - Q: how to make re-transmissions statistically uncorrelated with original transmissions?
    - (e.g., [Alomair et al., Globecom 2010])

Simple Approach

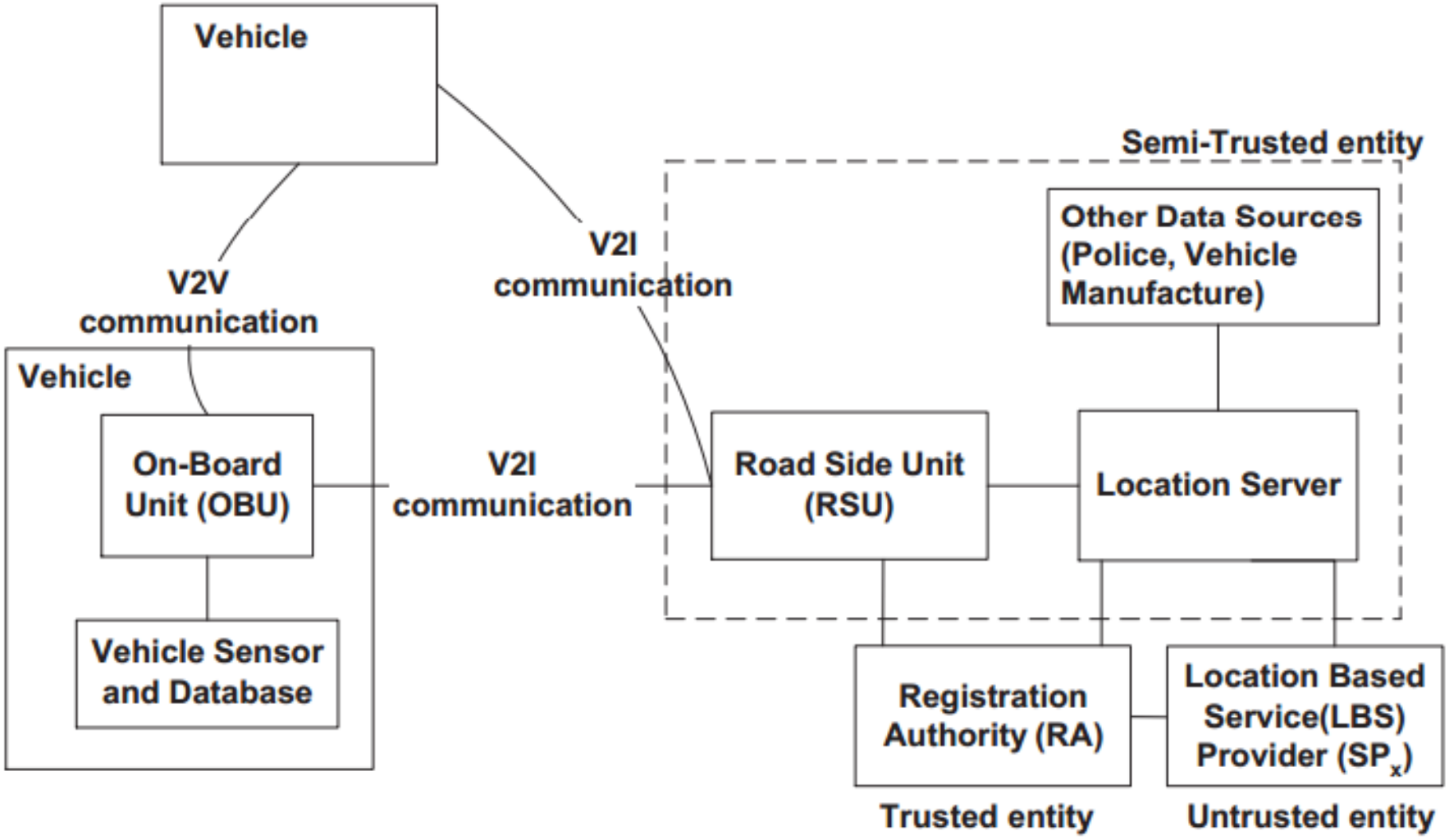


Better Approach



What about location privacy issues in mobile networks (e.g., VANETs)?

LBS in VANET



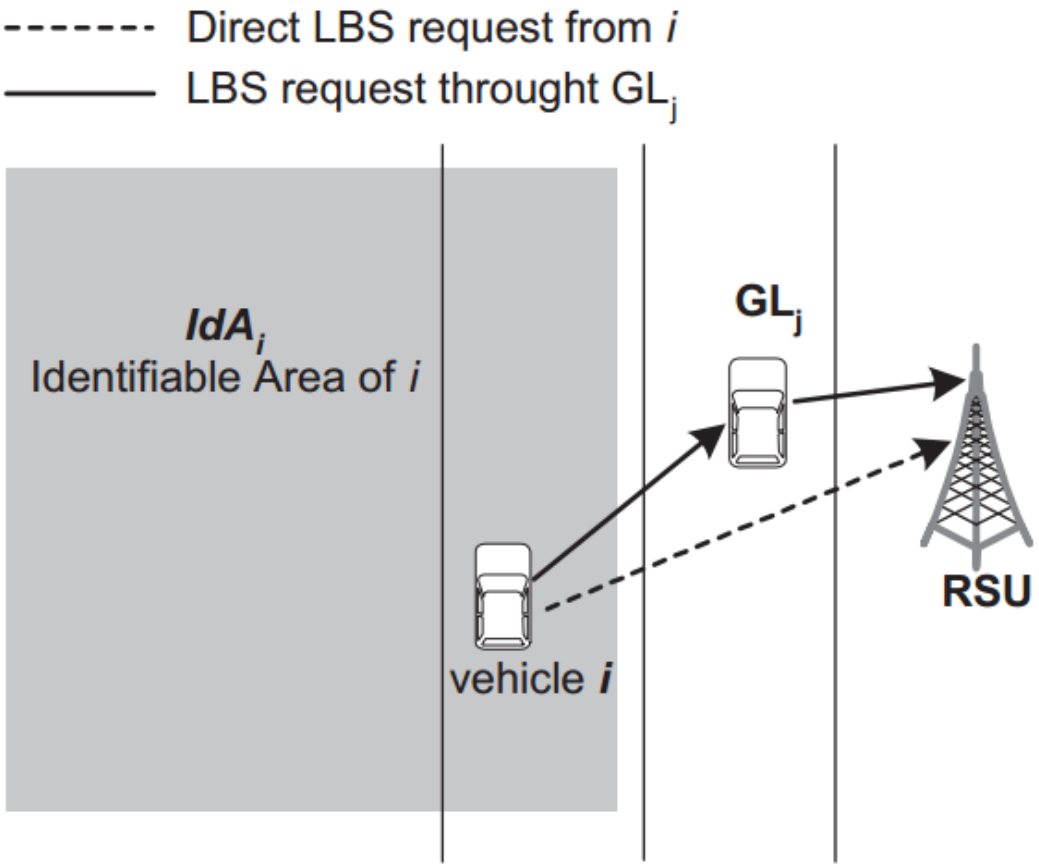
How to prevent the untrusted LBS from tracking vehicles?

AMOEBA

- Pseudonyms + group identity → location privacy among vehicles on the highway
  - Groups increase anonymity and reduce linkability
  - Pseudonym updates and silence at opportune times further reduce linkability
  - Power control allows group communication without infrastructure eavesdropping

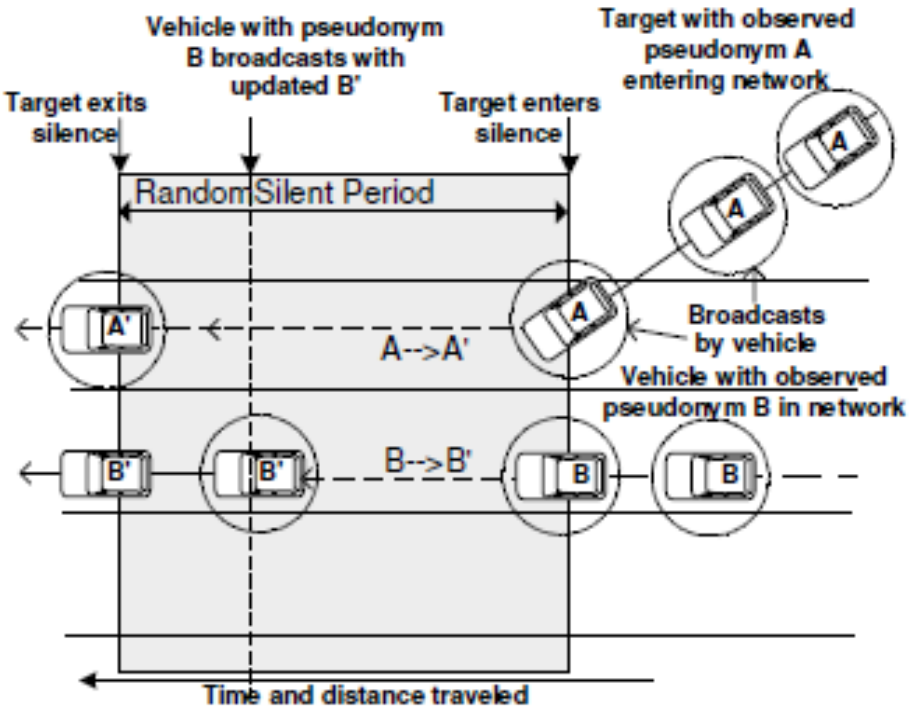
V2I → G2I

- Protect anonymity by grouping network traffic
  - Allow vehicles to form ad hoc groups
  - Group leader communicates to RSU
  - Rotate group leader randomly

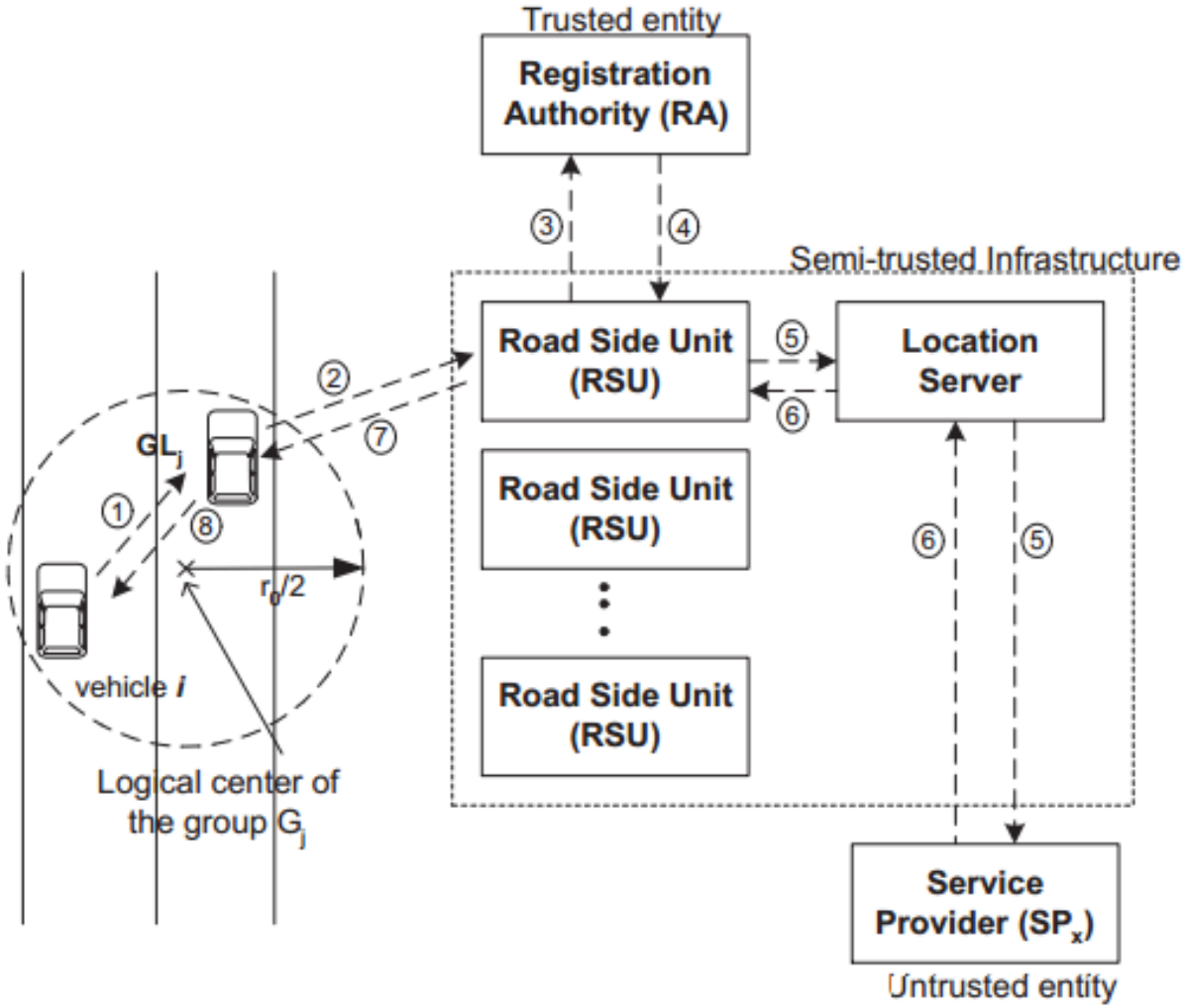


Leveraging Silence

- Road structure → pseudonyms not enough
  - Random silent period with pseudonym update reduces linkability, but causes safety problems
  - Rely on silent periods during times of high driver attentiveness, e.g., while changing lanes or merging



Privacy and LBS



Some Issues

- Trusted group leader?
  - Compromised group leader → no privacy
  - Rotation helps, but doesn't solve
- Trusted group?
  - Malicious group members can expose info to LBS, spoof LBS requests, etc.
- Lack of end-to-end control in V2I/LBS
  - Pay services?
  - No control over vehicles in data flow
  - Malicious leader could interfere

Summary

- We saw some unique location privacy issues in very different wireless systems
  - Additional location privacy issues exist in other domains / contexts, but no time to cover them all

- As systems continue to emerge / evolve, new privacy issues will arise

Mar 1:

Trust and Reputation

Mar 3:

NO CLASS