TRACK ME IF YOU CAN

On the Effectiveness of Context-based Identifier Changes in Deployed Mobile Networks

Network and Distributed System Security Symposium (NDSS)

6 February 2012

Laurent Bindschaedler*

laurent.bindschaedler@epfl.ch

School of Communications and Computer Sciences, EPFL, Switzerland

Co-Authors:

Murtuza Jadliwala, Wichita State University, USA*

Igor Bilogrevic, LCA EPFL

Imad Aad, Nokia Research Center

Philip Ginzboorg, Nokia Research Center

Valtteri Niemi, Nokia Research Center

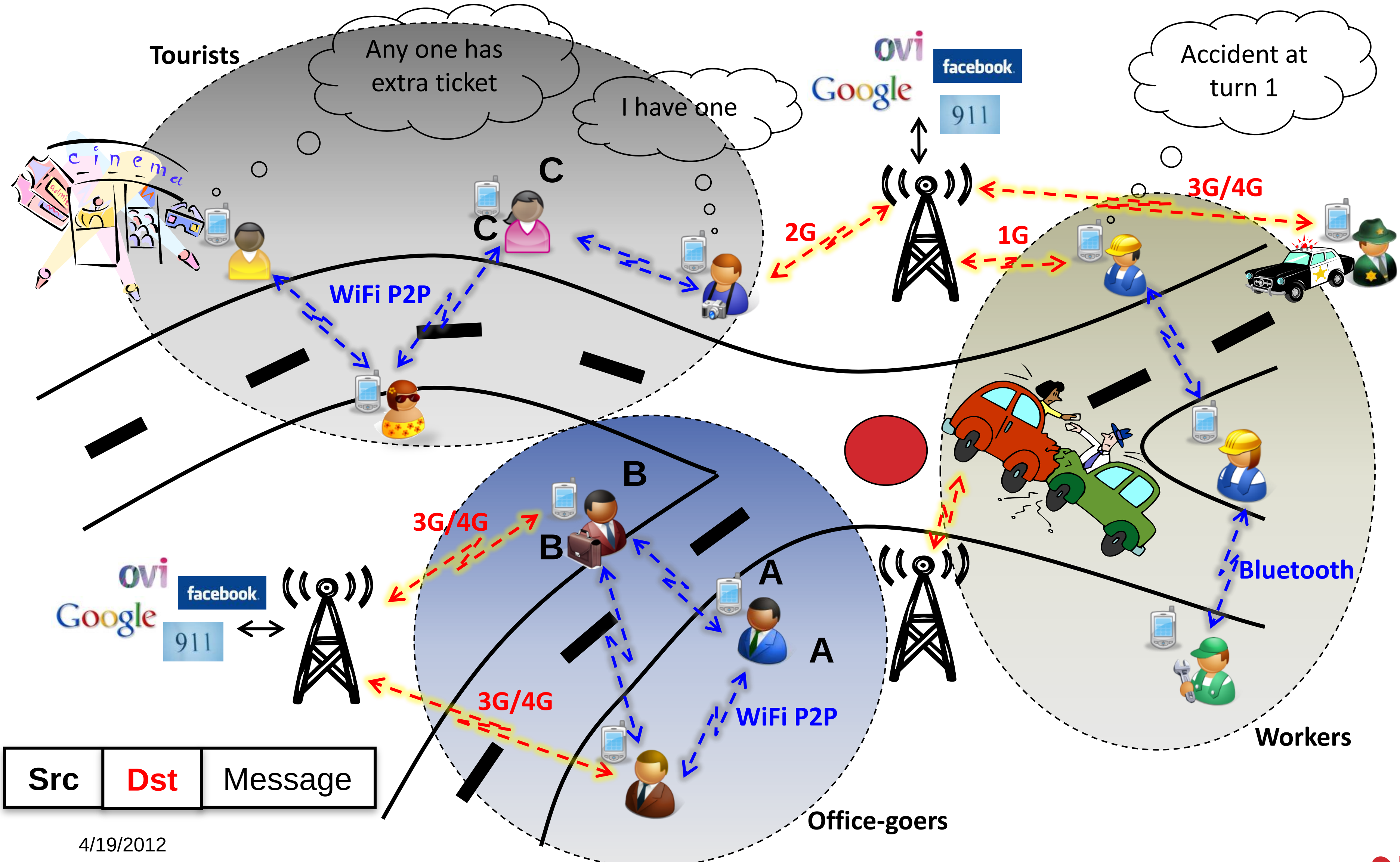
Jean-Pierre Hubaux, LCA EPFL

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

NOKIA Connecting People

** Equally contributing authors. Murtuza was affiliated with EPFL when this work was done.*

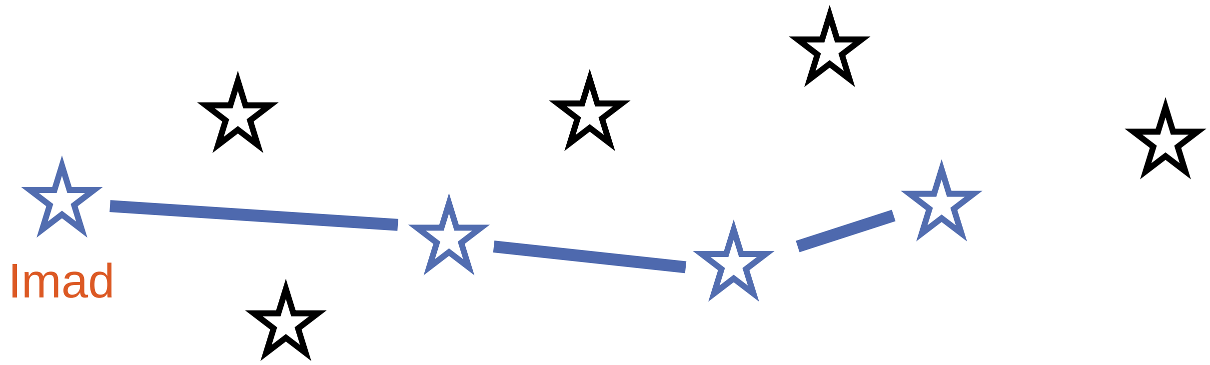
PERVASIVE SOCIAL NETWORKS



PRIVACY THREATS

- Most current devices have static identifiers, which allow service providers or malicious parties to track the location of users

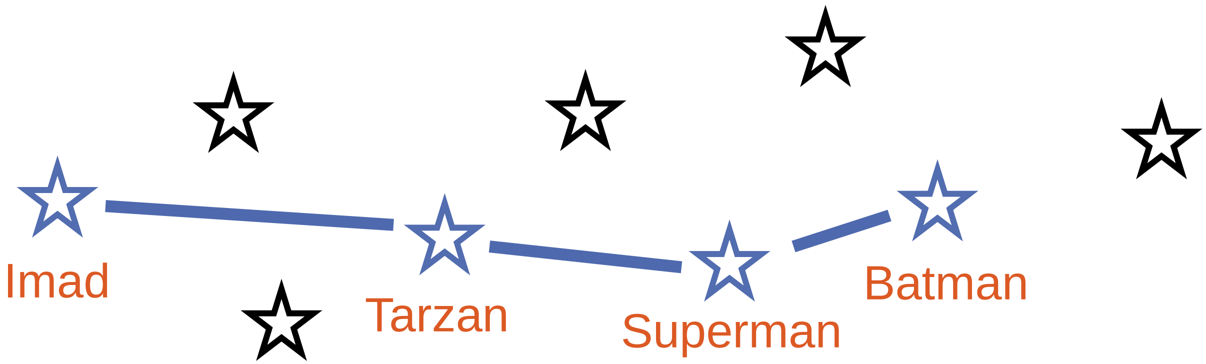
- Protecting the location privacy of users is critical



PRIVACY PROTECTION

- **Commonly deployed scheme to preserve privacy: replace device identifiers by short-lived identifiers or pseudonyms**

- **Mix zones are spatio-temporal regions where pseudonyms of several users are changed (mixed) to provide decorrelation between pseudonyms and devices [BeresfordS2003]**



[BeresfordS2003] Beresford, A.R. and Stajano, F., Location privacy in pervasive computing in *IEEE Pervasive Computing*, 2003.

STATE OF THE ART

- **There have been a few studies on the effectiveness of mix zones and optimal placement [ButtyanHV2007,GerlachG2007, WiedersheimMKP2010, FreudigerSH2009, JadliwalaBH2011]**
- **A majority of these studies focuses on other network models such as Vehicular Ad-hoc NETworks (VANETs)**
  - Difficult to transfer to PerSoNs because human and social factors
- **Due to the difficulty in running large-scale trials, many studies rely on simulated data**

[ButtyanHV2007]	L. Buttyán, T. Holczer, and I. Vajda. On the effectiveness of changing pseudonyms to provide location privacy in VANETs. In *ESAS*, 2007.
[GerlachG2007]	M. Gerlach and F. Guttler. Privacy in VANETs using changing pseudonyms - ideal and real. In *IEEE VTC-Spring*, 2007.
[WiedersheimMKP2010]	B.Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos. Privacy in Inter-Vehicular Networks: Why simple pseudonym change is not enough. In *IEEE/IFIP WONS*, 2010.
[FreudigerSH2009]	J. Freudiger, R. Shokri, and J.-P. Hubaux. On the optimal placement of mix zones. In *PETS*, 2009.
[JadliwalaBH2011]	M. Jadliwala, I. Bilogrevic, and J.P. Hubaux. Optimizing mixing in pervasive networks: a graph-theoretic perspective. In *Computer Security--ESORICS 2011*.

TRACK ME IF YOU CAN

On the Effectiveness

of Context-based Identifier Changes

in Deployed

Mobile Networks

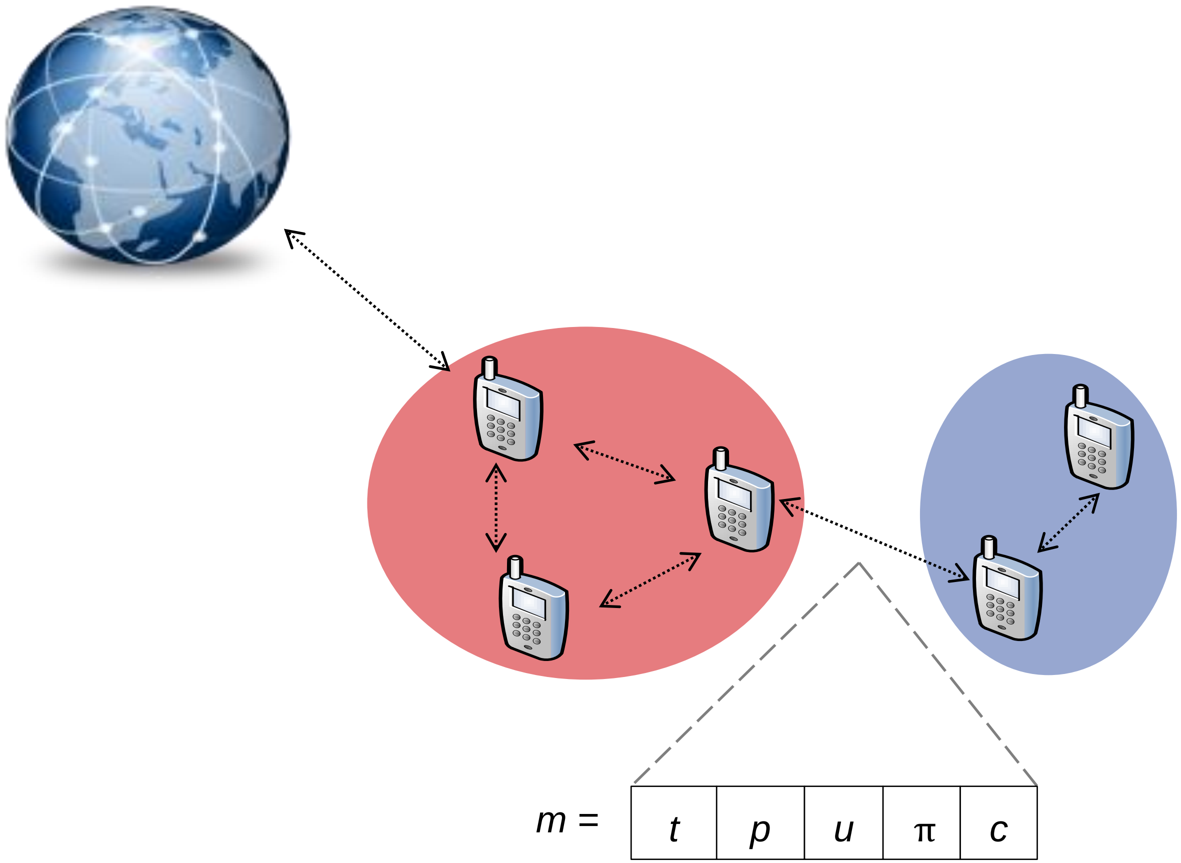
PRESENTATION OUTLINE

1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Conclusion

PRESENTATION OUTLINE

1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Conclusion

SYSTEM MODEL



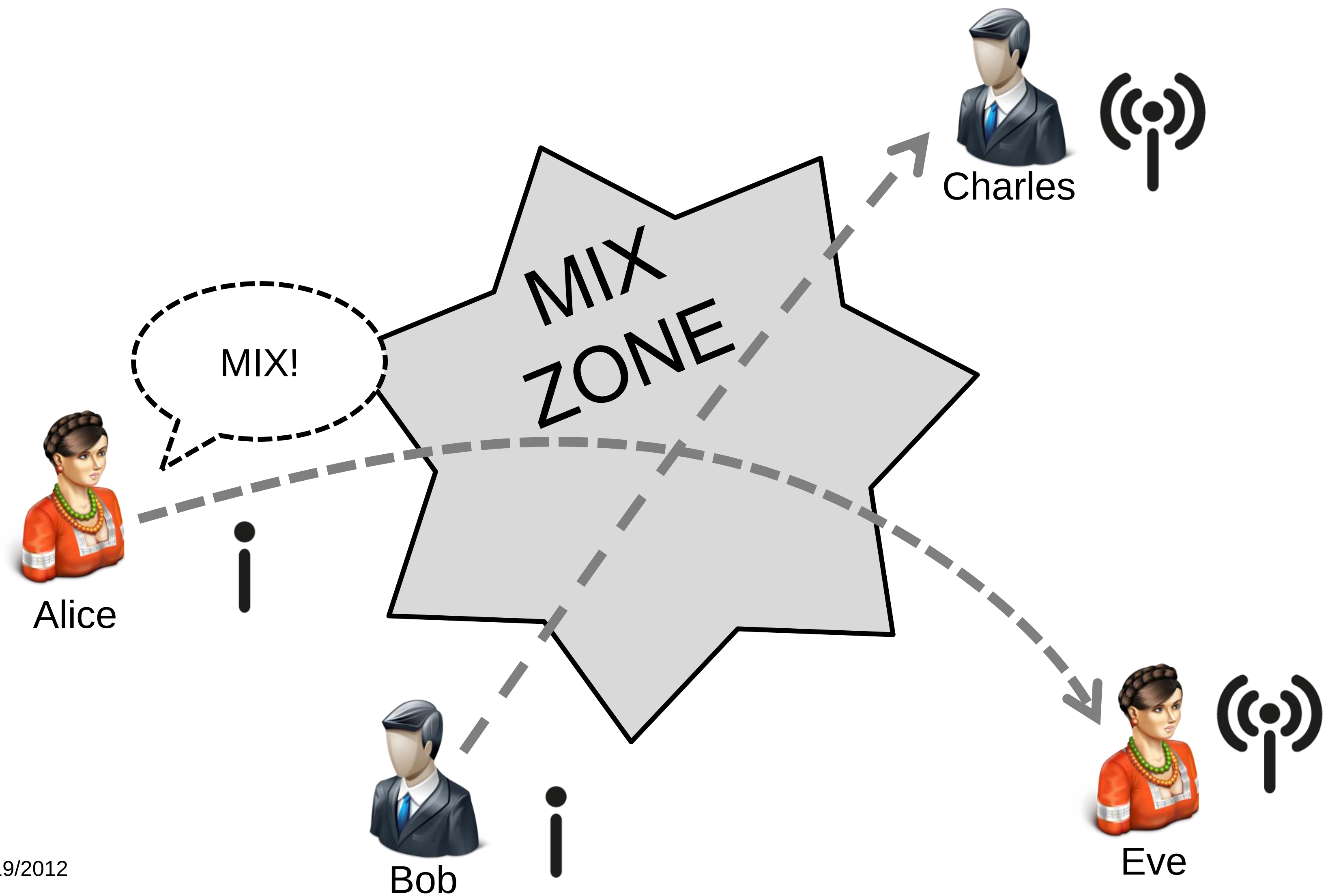
m: *message*
t: *timestamp*
p: *position*
u: *user*
π: *pseudonym*
c: *content*

NIC / NIC TRIAL

- **Nokia Instant Community (NIC)**
  - Multi-hop peer-to-peer network based on IEEE 802.11
  - Publish-subscribe messaging pattern
  - Users organized into communities

- **NIC Trial in a nutshell**
  - EPFL campus
  - March to June 2011
  - 80 participants (students and teachers)
  - Nokia N900 smartphones with NIC preinstalled
  - Log everything: usage, activity, message content, etc.

PSEUDONYM CHANGE ALGORITHM (PCA)



PCA PRINCIPLES

- **Change pseudonym**
  - based on context
  - at fixed (randomized) intervals

- **When a pseudonym change decision is made, the device broadcasts a mix request and changes its pseudonym**

- **Upon receiving a mix request, other devices in the neighborhood also change pseudonyms with some probability**

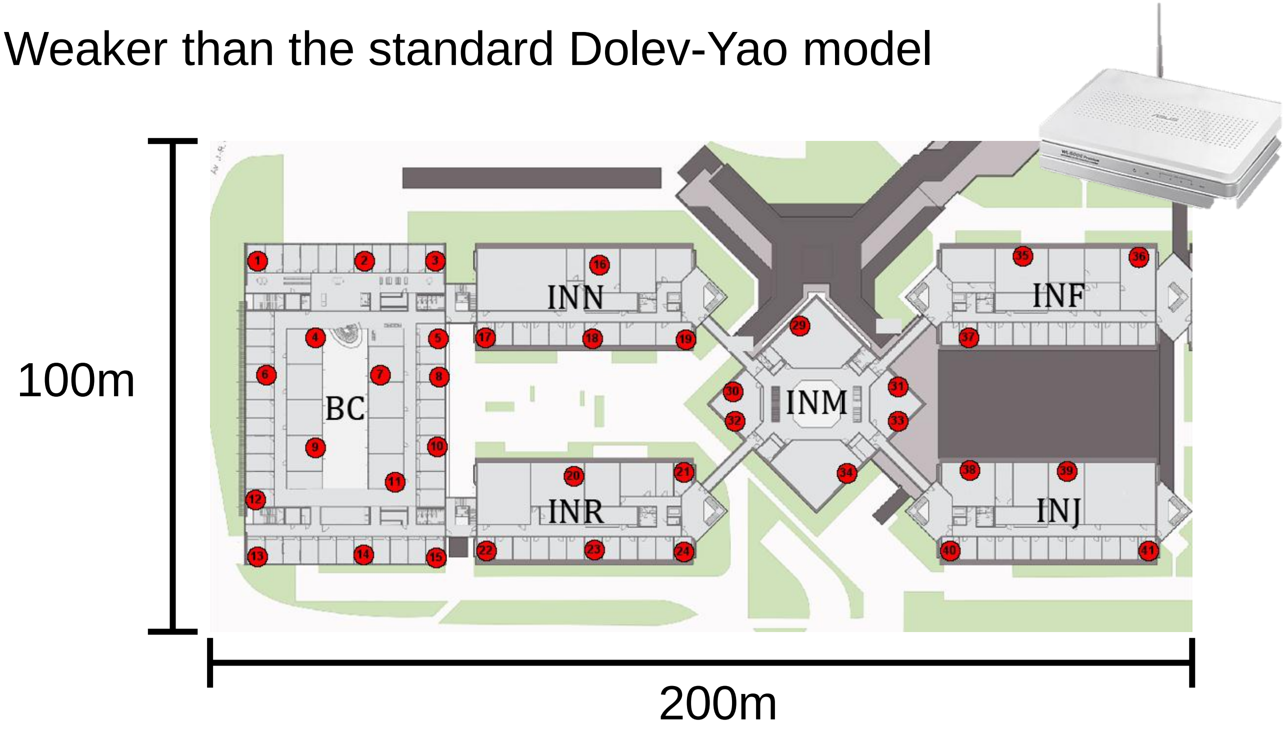
- **A quota is placed on the number of allowed pseudonym changes to prevent network performance collapse**

PCA PARAMETERS

	Cost effective	Intermediate	Privacy sensitive
Forced timer	14400s	7200s	3600s
Context timer	3600s	1200s	300s
Change threshold	7200s	1800s	600s
Neighbor threshold	1	2	3
Daily change quota	5	20	50

ATTACKER MODEL

- **Passive, eavesdrops using static mesh network of sniffing stations**
  - Weaker than the standard Dolev-Yao model



- **Reconstruction attack**

PRESENTATION OUTLINE

1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Conclusion

DATA PROCESSING

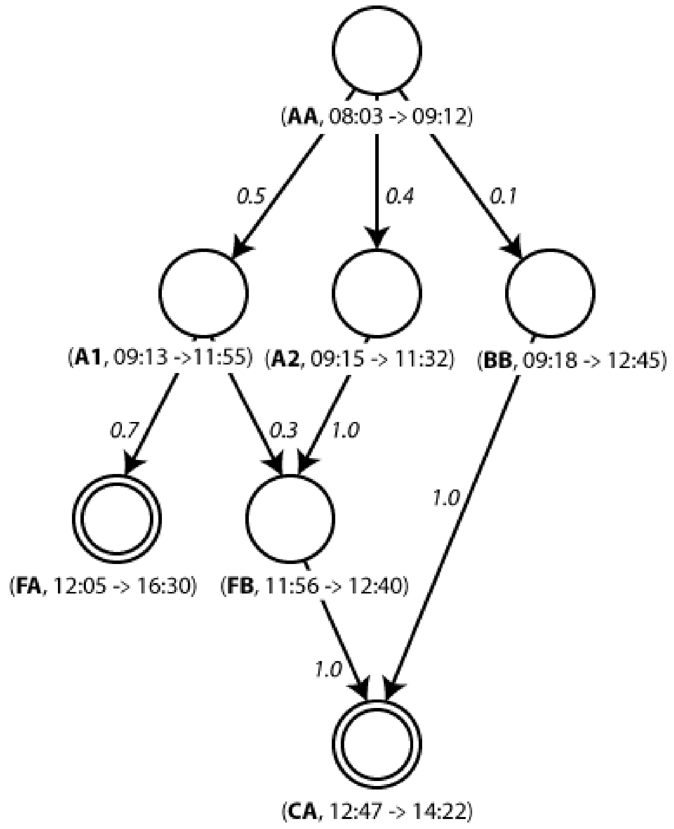
	Step performed by	
	Attacker	Experimenter
• Database insertion	✖	✖
• Groundtruth reconstruction		✖
• Timestamp synchronization	✖	
• Event identification	✖	
• Coordinate mapping	✖	

PRESENTATION OUTLINE

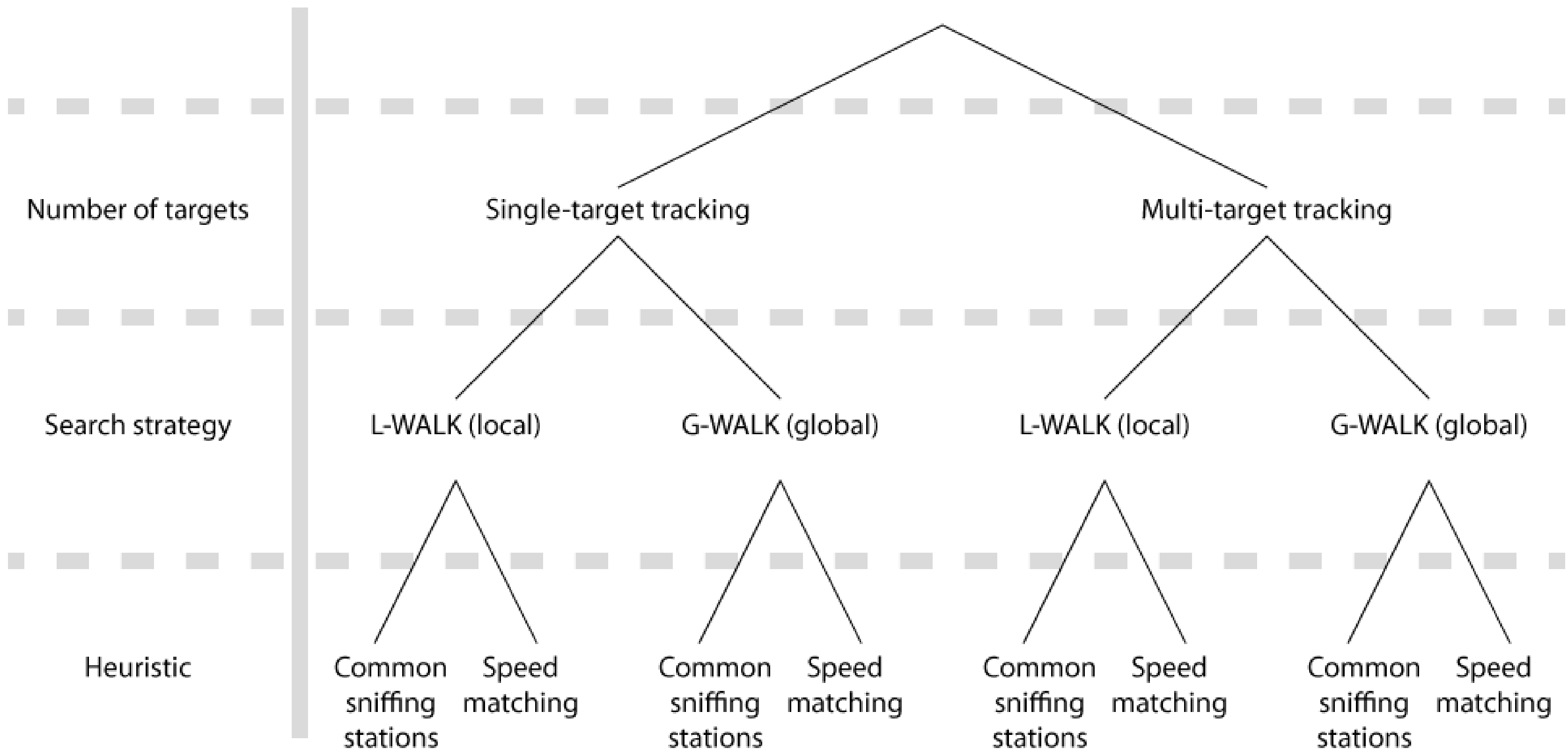
1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Future work

TRACKING MODEL

- **Finite-state first order Markov chain**
- **States S**
  - (pseudonym, first event → last event)
- **Transition probabilities P : S x S → [0,1]**
  - $\sum_{s_j \in S} P(s_i, s_j) = 1 \; \forall s_i \in S$
  - $P(s_i, s_j) = 0 \; \forall s_i, s_j \; with \; t_{end}(s_i) < t_{start}(s_j)$



TRACKING ATTACKS



TARGET TRACKING

- **Single-Target Tracking (STT) is the tracking of a single user in the state space**

- **Multiple-Target Tracking (MTT) is the simultaneous tracking of several users in the state space**

- **MTT can sometimes be more accurate than STT because it has a more global picture**
  - e.g., MTT can avoid collisions

SEARCH STRATEGY

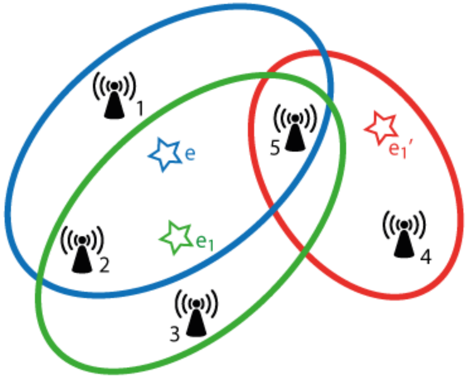
- **L-WALK builds a walk in the state space such that the next state candidate with the highest probability is selected at every step**
  - such a walk is locally optimal

- **G-WALK builds a walk in the state space such that the probability over the entire walk is maximized over all walks**
  - such a walk is globally optimal

HEURISTICS TO ESTIMATE TRANSITION PROBABILITIES

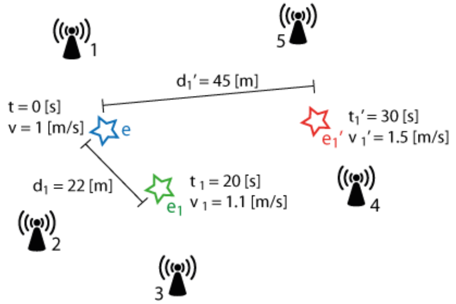
- **Common sniffing stations**
  - «The more sniffing stations in common between the current state and the next state candidate, the more likely the candidate»

$$h_1(s_0, s) = \begin{cases} \dfrac{|o_{end}(s_0) \cap o_{start}(s)|}{\sum_{s' \in C(s_0)} |o_{end}(s_0) \cap o_{start}(s')|} & \text{if } s \in C(s_0) \\ 0 & \text{otherwise} \end{cases}$$



- **Speed matching**
  - «The closer the user speeds between the current state and the next state candidate, the more likely the candidate»

$$h_2(s_0, s) = \begin{cases} \dfrac{1 - \min(\Delta v(s_0, s)/v_{max}, 1)}{\sum_{s' \in C(s_0)} (1 - \min(\Delta v(s_0, s')/v_{max}, 1))} & \text{if } s \in C(s_0) \\ 0 & \text{otherwise} \end{cases}$$



PRESENTATION OUTLINE

1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Conclusion

PRIVACY METRICS

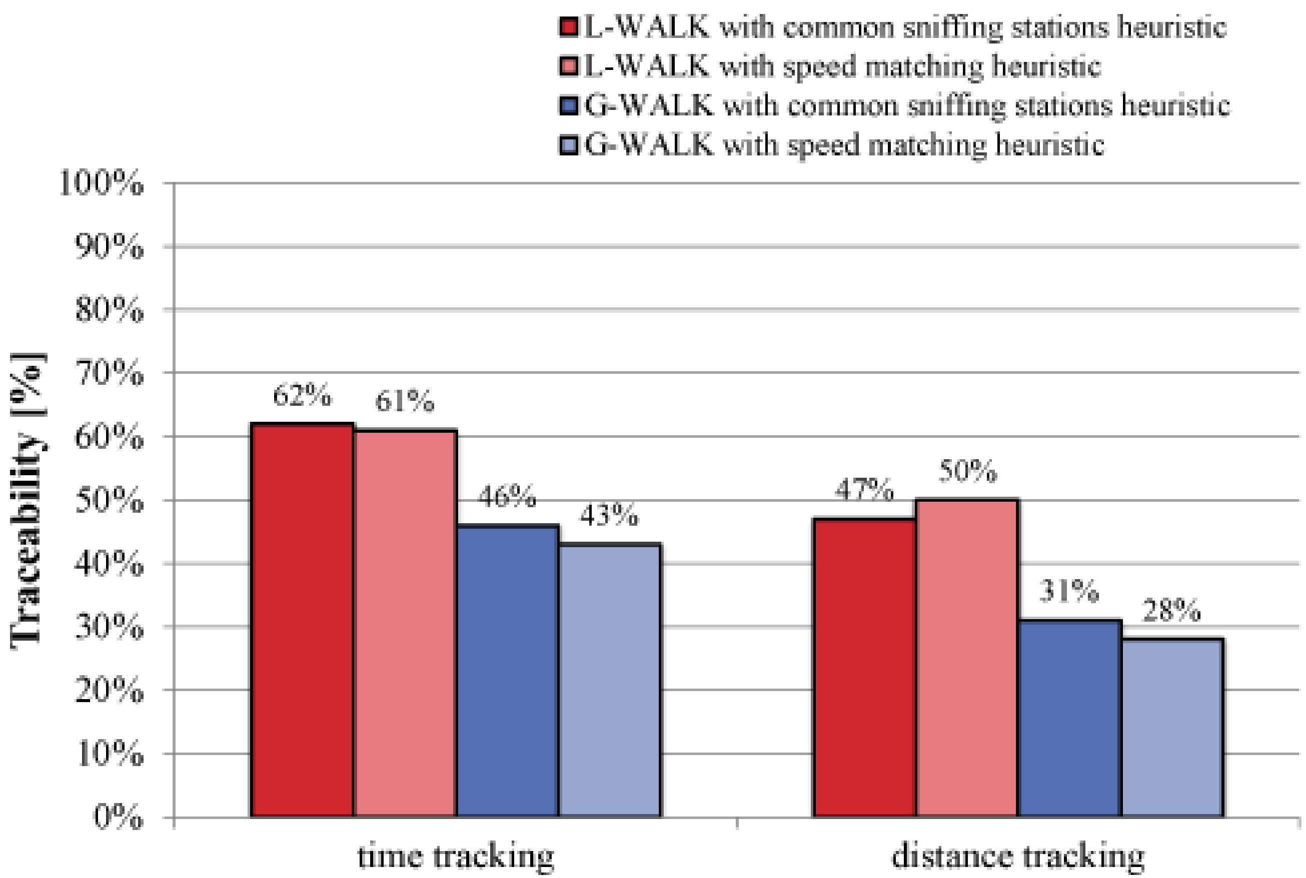
- **Traceability τ-metrics: Capture the extent to which the user can be tracked in time or distance [HohGXA2007]**
- **Uncertainty u-metrics: Capture the uncertainty in the next choice of pseudonym [DiazSCP2002]**
- **Traceability-uncertainty μ-metrics: Capture the extent to which the user can be tracked along with the difficulty (uncertainty) in the tracking (homebrewed)**
- **Clustering c-metrics: Capture the extent to which one user was confused for another [HohG2005]**

[HohGXA2007]	B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Preserving privacy in gps traces via uncertainty-aware path cloaking in *ACM CCS*, 2007.
[DiazSCP2002]	C. Diaz, S. Seys, J. Claessens, and B. Preneel.Towards measuring anonymity in *PET*, 2002.
[HohG2005]	B. Hoh and M. Gruteser. Protecting location privacy through path confusion in *SECURECOMM*, 2005.
[ShokriFJH2009]	Shokri, R. and Freudiger, J. and Jadliwala, M. and Hubaux, J.P., A Distortion-based Metric for Location Privacy in *Proceedings of the 8th ACM workshop on Privacy in the electronic society, 2009.*

GENERAL TRACKING RESULTS

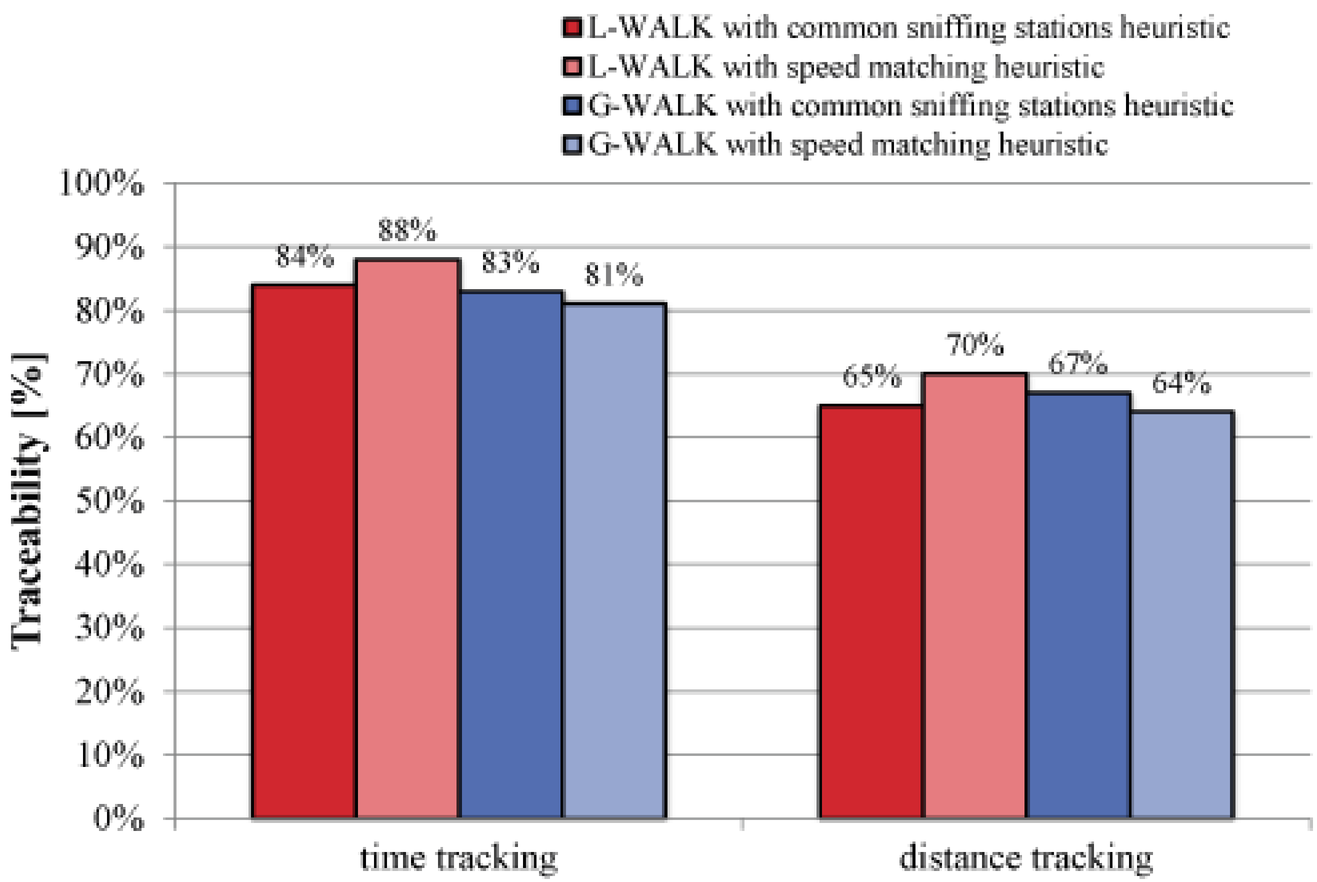
Direct application of tracking algorithms to the original data set

TRACKING RESULTS



Single-user tracking results
(averaged over the three sets of PCA parameters)

TRACKING RESULTS

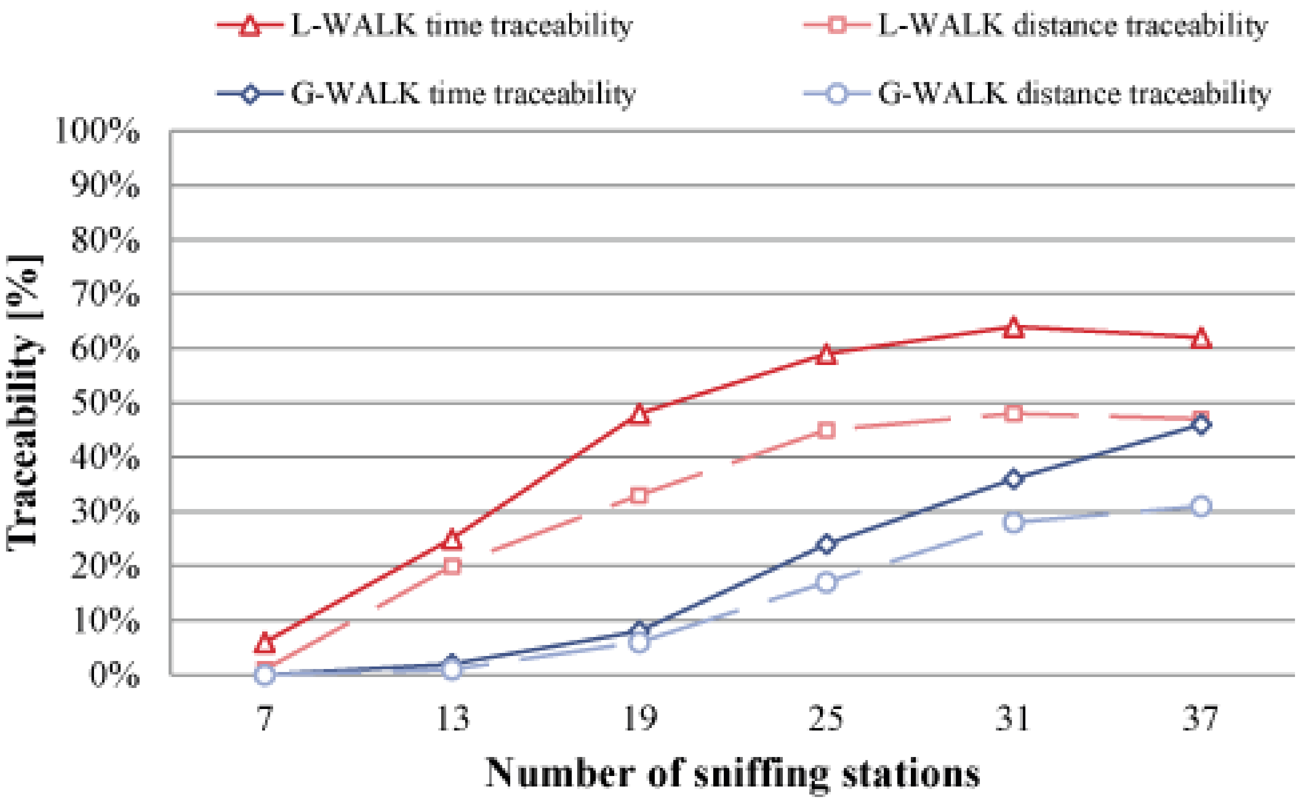


Multi-user tracking results
(averaged over the three sets of PCA parameters)

TRACKING WITH VARYING ADVERSARY STRENGHTS

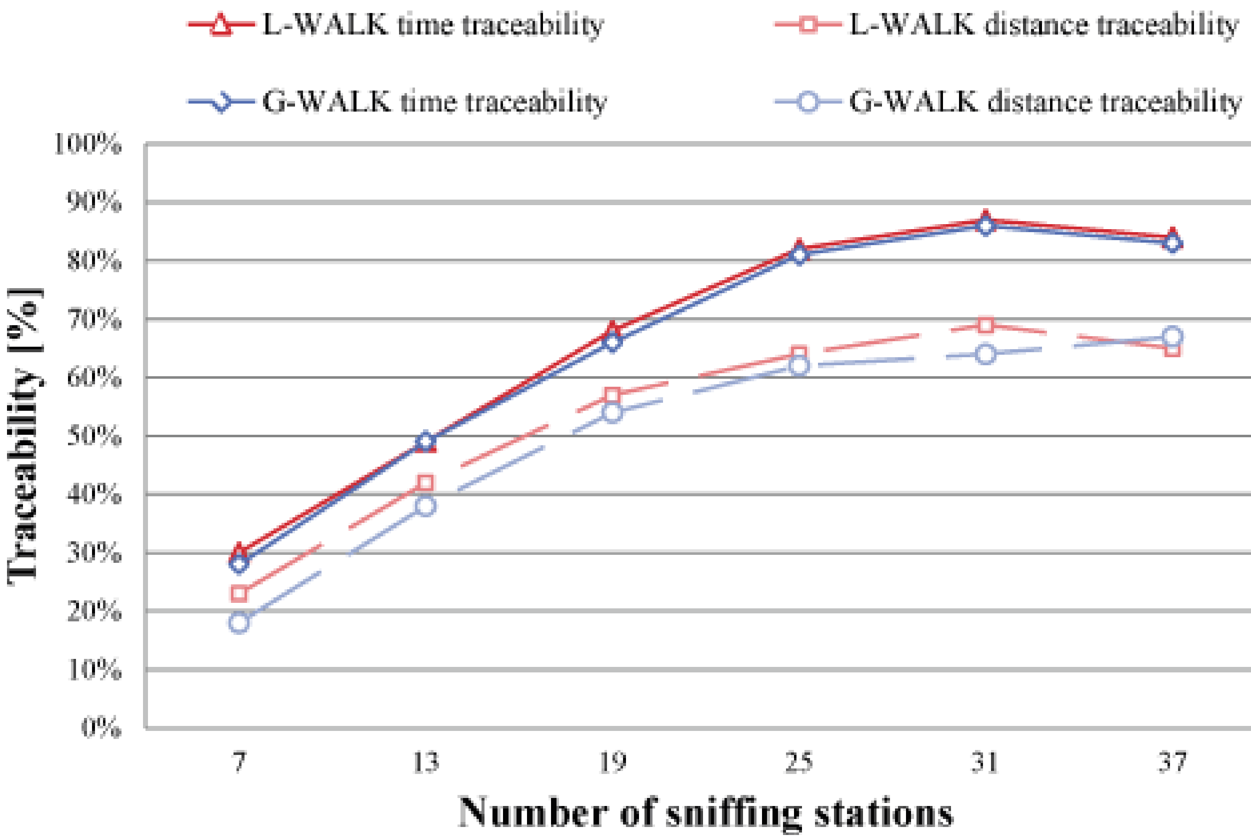
Vary the number of sniffing stations by iteratively removing 6 stations (uniformly selected)

TRACKING RESULTS



Single-user tracking results
with varying adversary strength

TRACKING RESULTS

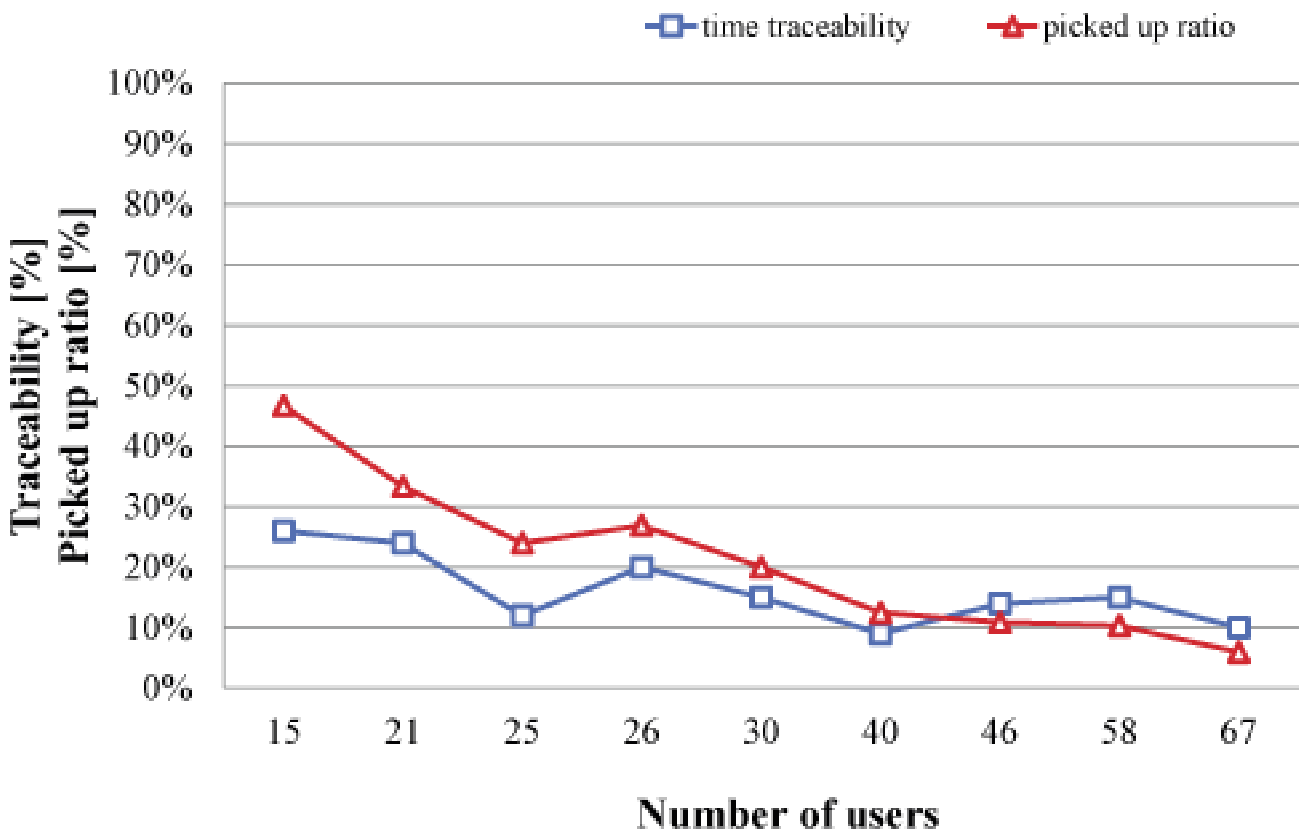


**Multi-user tracking results
with varying adversary strength**

TRACKING IN LARGE USER CLUSTERS

Evaluation of the tracking effectiveness when user density is the highest

TRACKING RESULTS



Traceability in large user clusters

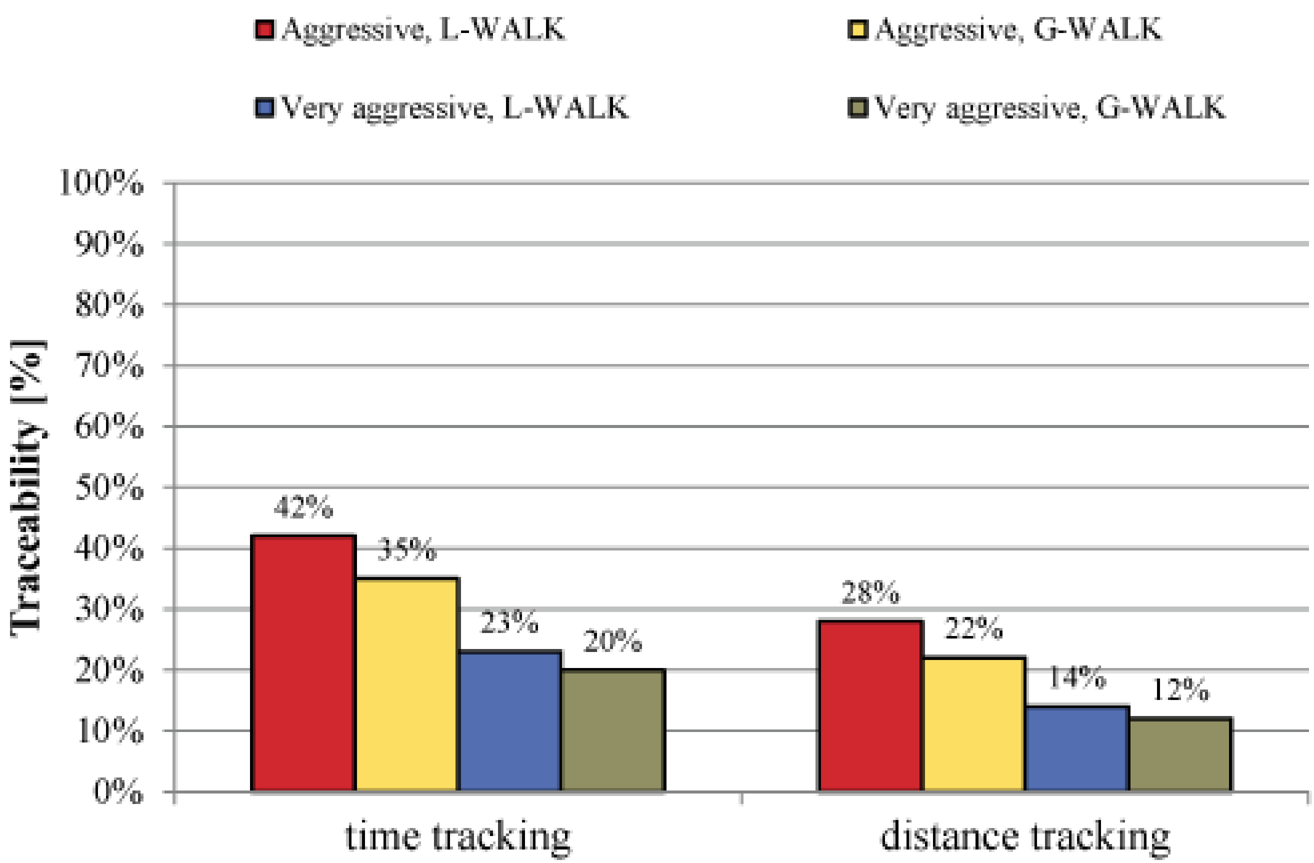
TRACKING WITH AGGRESSIVE PCA

Simulate PCA with more aggressive parameters based on the original data set

AGGRESSIVE PCA PARAMETERS

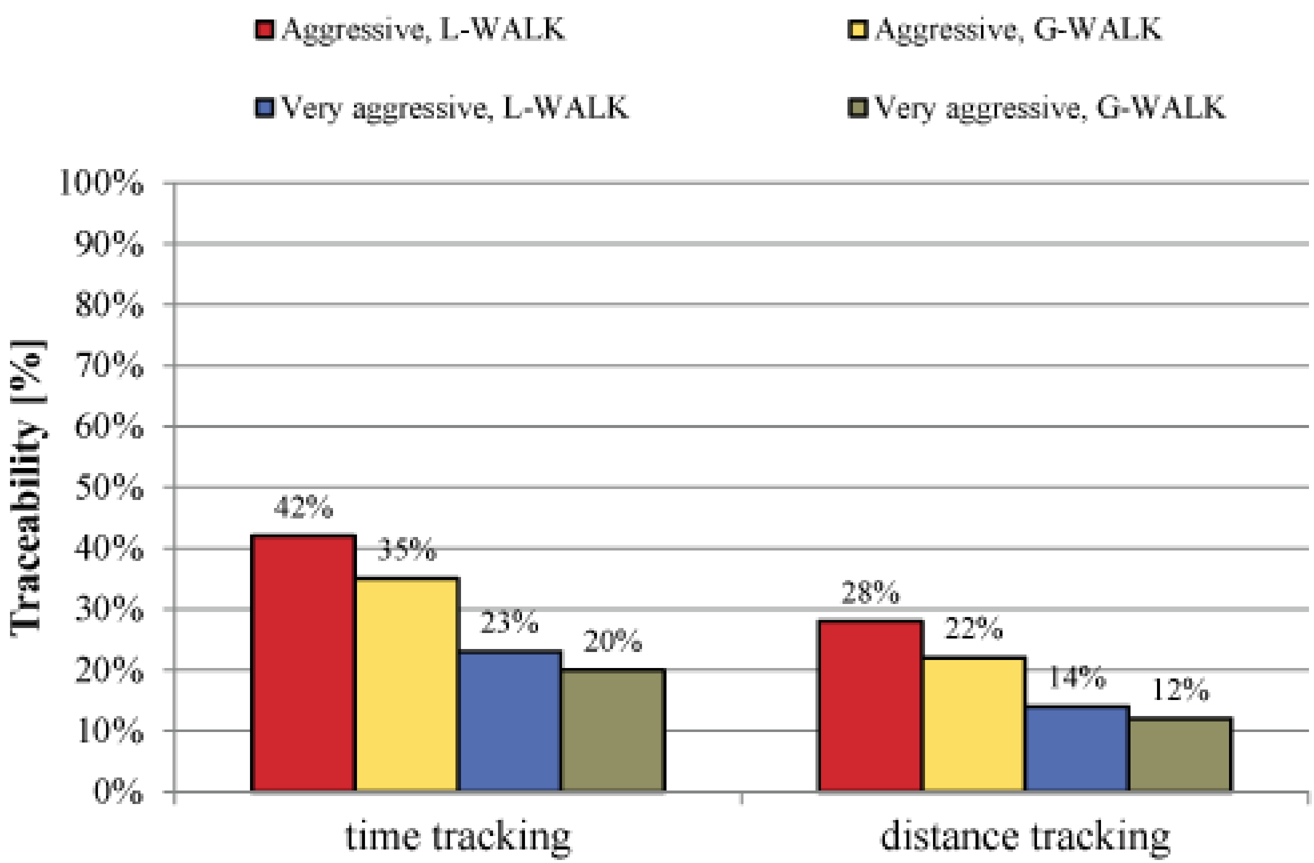
	Privacy sensitive	Aggressive	Very aggressive
Forced timer	3600s	1200 sec	600 sec
Context timer	300s	120 sec	60 sec
Change threshold	600s	300 sec	120 sec
Neighbor threshold	3	3	3
Daily change quota	50	200	500

TRACKING RESULTS



Single-user traceability results with more aggressive PCA

TRACKING RESULTS



Multi-user traceability results with more aggressive PCA

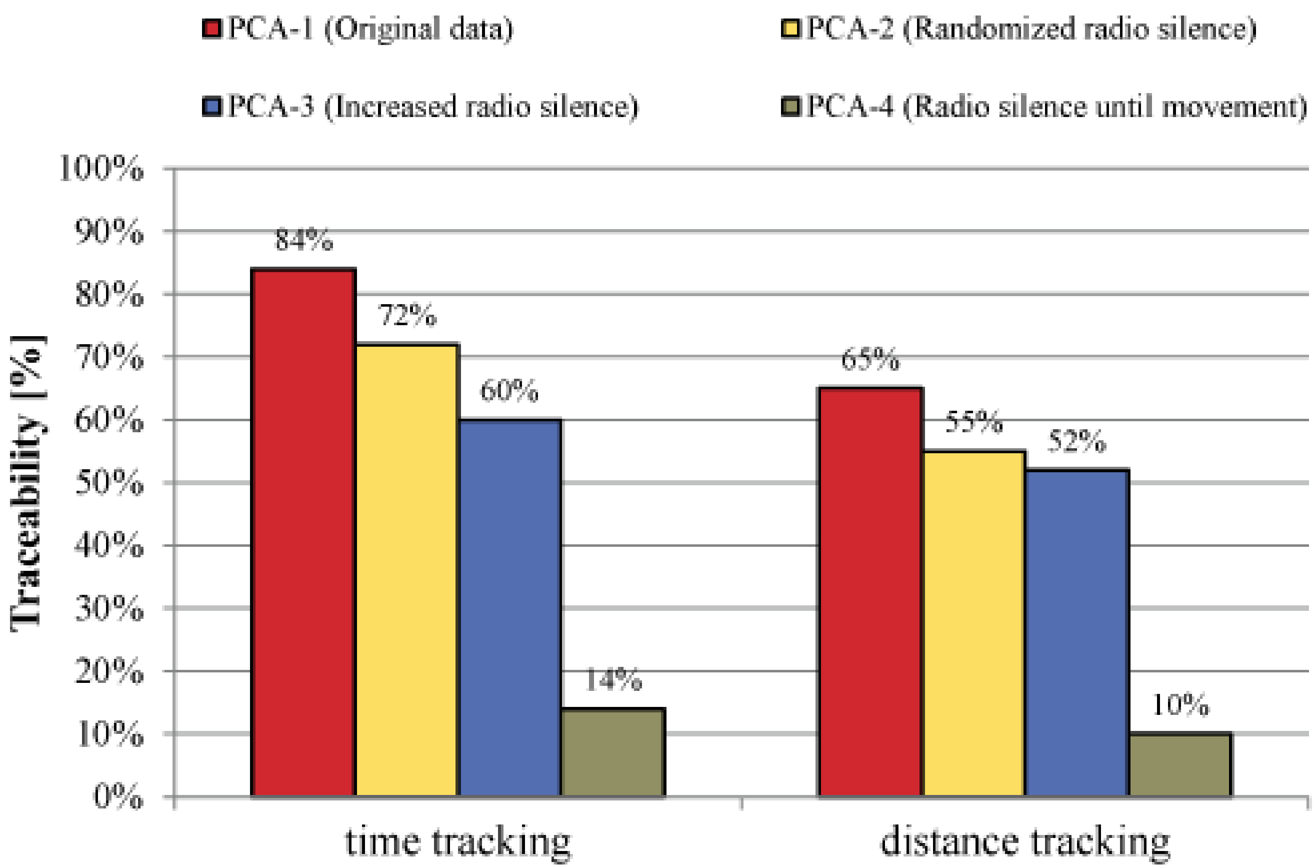
TRACKING WITH IMPROVED PCA

Simulate modified versions of the PCA on the original data set

PCA IMPROVEMENTS

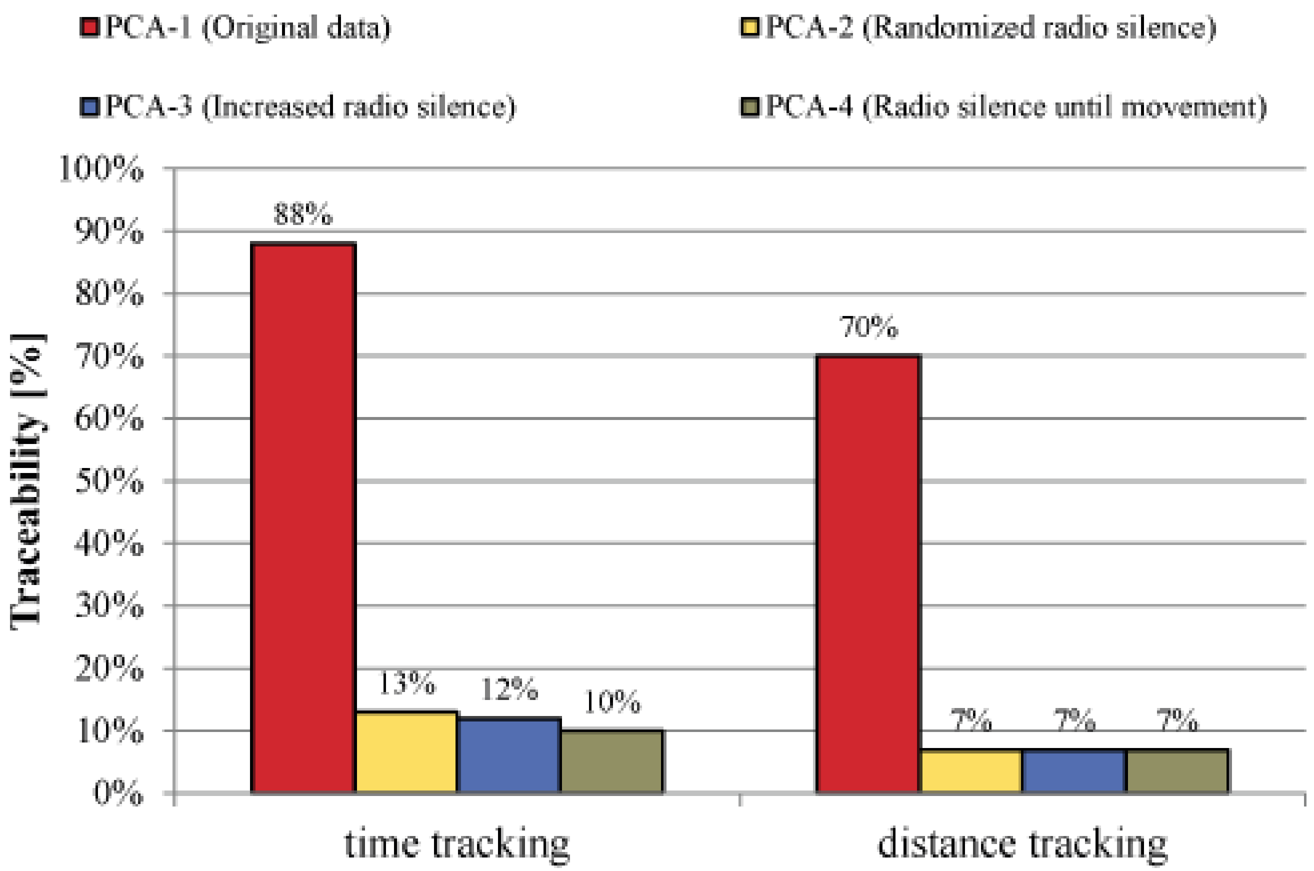
- PCA-2
  - Randomize radio silence over a larger time interval
- PCA-3
  - Observe longer radio silence periods
- PCA-4
  - Maintain radio silence until there has been significant movement of the user

TRACKING RESULTS



Multi-user traceability results with improved PCA (obtained using common sniffing stations heuristic)

TRACKING RESULTS



Multi-user traceability results with improved PCA (obtained using speed matching heuristic)

PRESENTATION OUTLINE

1. System Model
2. Data Collection and Processing
3. Tracking Framework and Algorithms
4. Empirical Results and Evaluation
5. Conclusion

CONCLUSION

- **Even basic tracking strategies can achieve high success in real life**

- **Pseudonym change has an impact on network performance which should not be neglected when designing an algorithm and choosing parameters**
  - tradeoff is required

- **Standard pseudonym change algorithms should be modified to improve protection**
  - e.g., by taking movement into consideration

WANNA PLAY WITH THE NIC TRIAL DATA?

Here's where you can find more info:

https://lausanne.nokiaresearch.com/nic

or

http://bit.ly/nictrial

REFERENCES

[Gartner2009]	A Lapkin. Context-aware computing: Four questions CIOs should be asking. *Gartner*, 2009.
[BeresfordS2003]	Beresford, A.R. and Stajano, F., Location privacy in pervasive computing in *IEEE Pervasive Computing*, 2003.
[ButtyanHV2007]	L. Buttyán, T. Holczer, and I. Vajda. On the effectiveness of changing pseudonyms to provide location privacy in VANETs. In *ESAS*, 2007.
[GerlachG2007]	M. Gerlach and F. Guttler. Privacy in VANETs using changing pseudonyms - ideal and real. In *IEEE VTC-Spring*, 2007.
[WiedersheimMKP2010]	B.Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos. Privacy in Inter-Vehicular Networks: Why simple pseudonym change is not enough. In *IEEE/IFIP WONS*, 2010.
[FreudigerSH2009]	J. Freudiger, R. Shokri, and J.-P. Hubaux. On the optimal placement of mix zones. In *PETS*, 2009.
[HohGXA2007]	B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady. Preserving privacy in gps traces via uncertainty-aware path cloaking in *ACM CCS*, 2007.
[DiazSCP2002]	C. Diaz, S. Seys, J. Claessens, and B. Preneel.Towards measuring anonymity in *PET*, 2002.
[HohG2005]	B. Hoh and M. Gruteser. Protecting location privacy through path confusion in *SECURECOMM*, 2005.
[ShokriFJH2009]	Shokri, R. and Freudiger, J. and Jadliwala, M. and Hubaux, J.P., A Distortion-based Metric for Location Privacy in *Proceedings of the 8th ACM workshop on Privacy in the electronic society, 2009.*