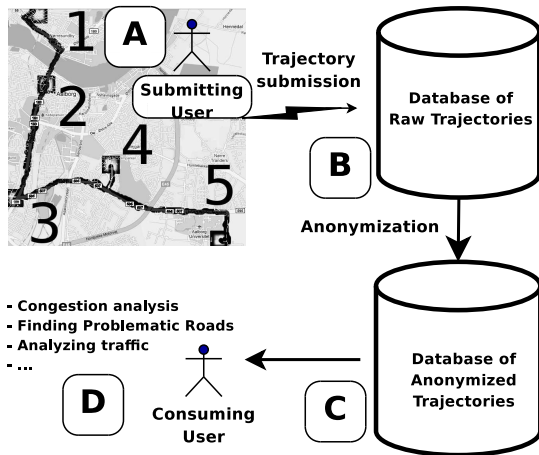# Personalized Protection of Identifiers on Public Trajectories

Jeppe R. Thomsen

Aalborg University
Department of Computer Science

June 24, 2010

Introduction
Privacy Profile
Algorithms
Conclusion

Problem Setting
Goals
Related work

Problem Setting

A Privacy Aware User

B Trusted Server

C Public Untrusted Server

D Service Providers

Introduction
Privacy Profile
Algorithms
Conclusion

Problem Setting
Goals
Related work

Goals

At the service provider:

- Remove all user identifying information from trajectories.
- Preserve usability to users of public dataset

At the users side:

- Provide **Usability**. specifying privacy should be simple.
- Be **Practical**. No user interaction during normal operation.
- Be **Flexible**. Support several ways of defining privacy.

Introduction
Privacy Profile
Algorithms
Conclusion

Problem Setting
Goals
Related work

Related work

Protection of Trajectories

- Collapse trajectories and remove updates
- Only publish edges with k support.

- At each update compute MBR including k-1 updates
- Precompute regions before sending.

- Degrade public dataset so no sub-trajectory can be matched to it.

Introduction
Privacy Profile
Algorithms
Conclusion

Settings
PSR
Time Period
t-anonymity

Privacy Profile

- Settings
- t-anonymity
- PSR
- Protection types and schemes

Introduction
Privacy Profile
Algorithms
Conclusion

Settings
PSR
Time Period
t-anonymity

Settings

Users Can

- Set both globally and locally
  - Temporal sensitivity
  - Spatial sensitivity
- Define a PSR
- Have multiple profiles.

Definition (Privacy Profile)

$(stime, etime, d_s, d_t, \{PSR\})$

PSR

Definition (PSR)

A PSR p is a tuple $(p_{edges}, d_s, d_t, class)$ where p_{edges} is the set of tuples $\{(e, e_{from}, e_{to} | 0 \leq e_{from} < e_{to} \leq e_{length})\}$ which is sensitive. $e \in \mathbf{E}$ and $e_{from}, e_{to}, e_{length} \in \mathbb{R}$. e_{from}/e_{to} specifies on e the start-/end-location covered by p_{cover}. If e is fully included in p_{cover}, e_{from}/e_{to} is equal to $0/p_{length}$. $d_s, d_t, class \in \mathbb{N}$ is respectively the spatial sensitivity, the temporal sensitivity, and the PSR classification

Introduction
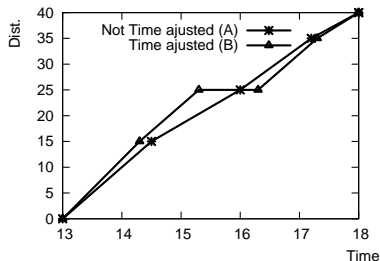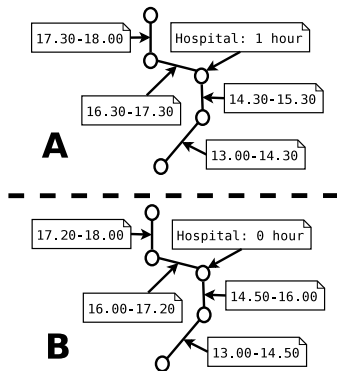Privacy Profile
Algorithms
Conclusion

Settings
PSR
Time Period
t-anonymity

PSR Classes

Classification	Scheme
Public Service Point	AS
House	ASTI,RS
Route w. endpoints	AS, ASTI, RS
Route w/o endpoints	AS, ASTI, RS

Protection Schemes
- AS - Always Sensitive.
- ASTI - Always Sensitive within a time interval.
- RS - Rarely Sensitive.

Introduction
Privacy Profile
Algorithms
Conclusion

Settings
PSR
Time Period
t-anonymity

Time Period

Introduction
Privacy Profile
Algorithms
Conclusion

Settings
PSR
Time Period
t-anonymity

t-anonymity

Definition (t-anonymity)

Given $\mathbf{T}$, the set of trajectories and p_{edges}, the set of edges covering a sensitive part of trajectory γ.

Let $\Gamma \subseteq \mathbf{T}$ be all trajectories which subtrajectories intersect with p_{edges}. $\Gamma' \subseteq \Gamma$ be all trajectories where, for edges intersecting with p_{edges}, at each timestamp of γ their timestamps lie within a time period TP symmetric around the timestamp of γ.

Γ' is said to satisfy t-anonymity with respect to TP and γ iff Γ' contains at least $t - 1$ other trajectories.

Algorithm

---

1 **while** *Sensitive unanonymized edges exist* **do**

2 $\quad$ $\alpha \leftarrow \textbf{Choose}\_\alpha(\textbf{T}, \textbf{PS})$

3 $\quad$ $PSRcand \leftarrow \textbf{FindCand}(\alpha, \textbf{PS})$

4 $\quad$ $calcCand \leftarrow \textbf{CalcCand}(PSRcand, \alpha, D, n)$

5 $\quad$ $sortCand \leftarrow$ Sort $calcCand$ using ordering given by **CompareCand()**

6 $\quad$ $anonData \leftarrow anonData \cup \textbf{AnonCand}(sortCand, \alpha)$

7 $anonData \cup \{\forall t_i \in t | t \in \textbf{T}\}, t_i$ is a subtrajectory that has not been modified or otherwise included during anonymization.

TrajecoriesPSR

> 1 **TrajectoriesPSR**$(P, \mathbf{T}, \alpha)$
> 2 $u \leftarrow$ **PSRtoUser**(P) //user $(id, s, \{t\})$ where $P \in s.\{PSR\}$
> 3 $tSet \leftarrow \{(t, t_{se}, d_t, d_s) | \forall t \in \mathbf{T} \wedge t \cap \alpha.t \neq \emptyset \wedge P.p_{edges} \cup t \neq \emptyset \wedge t_{se} = \alpha.t \cap t \wedge t \in u.t, d_t = P.d_t, d_s = P.d_s \wedge \forall i, j | \alpha.t_{se}[i]_{\tau_s} - \frac{\alpha.d_t}{2} \leq t_{se}[j]_{\tau_s} \leq \alpha.t_{se}[i]_{\tau_s} + \frac{\alpha.d_t}{2}\}$
> 4 **return** $tSet$

Introduction
Privacy Profile
Algorithms
Conclusion

Conclusion
Future Work

Conclusion

- Novel Privacy Profile to specify spatial-temporal sensitivity

- Introduced t-anonymity

- Introduced a way of temporally hiding users movements.

Introduction
Privacy Profile
Algorithms
Conclusion

Conclusion
Future Work

Future Work

- Performance study to determine a threshhold **D** for data integrety, to determine when data is no longer usable by data consumers.

Introduction
Privacy Profile
Algorithms
Conclusion

Conclusion
Future Work

End of Presentation

Thank You For Listening