Privacy-Enhanced Participatory Sensing with Collusion Resistance and Data Aggregation

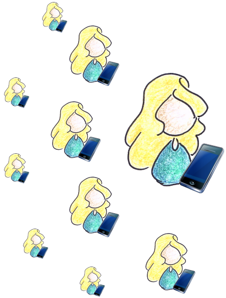
Felix Günther, Mark Manulis, and Andreas Peter

Technische Universität Darmstadt, University of Surrey, University of Twente

Participatory Sensing

or: Urban/Opportunistic/People-Centric Sensing

TECHNISCHE UNIVERSITÄT DARMSTADT



Smartphones

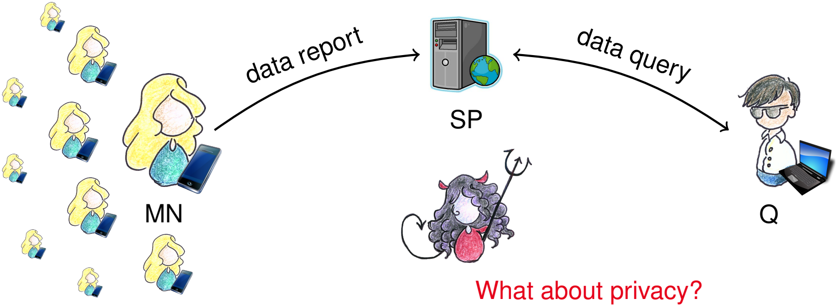
- $\gg$ 1 billion worldwide
- highly mobile
- powerful
- always connected
- embedded sensors
  GPS, motion, temperature, ...

Thanks to *Giorgia Azzurra Marson* for the drawings.

Participatory Sensing

or: Urban/Opportunistic/People-Centric Sensing

TECHNISCHE UNIVERSITÄT DARMSTADT



What about privacy?

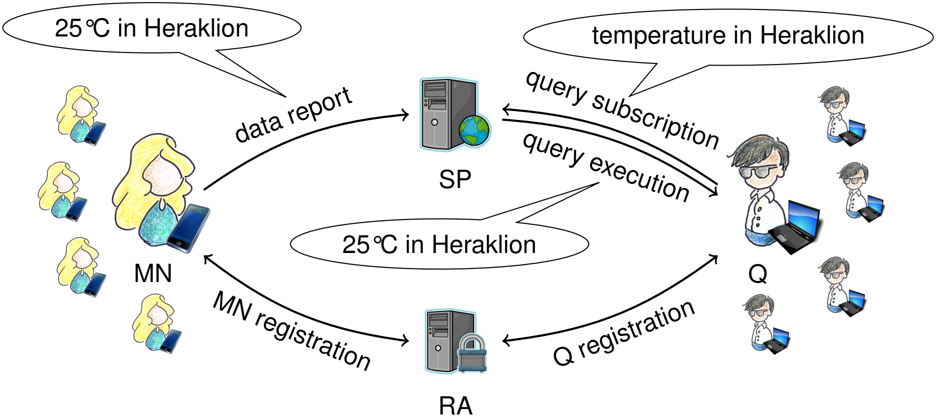
Thanks to *Giorgia Azzurra Marson* for the drawings.

Previous Approaches (selection)

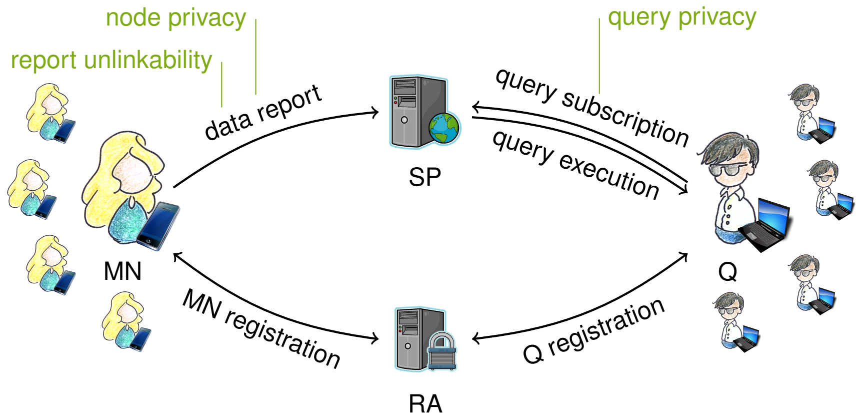
TECHNISCHE UNIVERSITÄT DARMSTADT

- AnonySense — Cornelius et al. @ MobiSys 2008
  - k-anonymity, mix networks, multiple semi-trusted servers
  - extension to l-diversity — Huang et al. @ Computer Comm. 33(11), 2010
  - no confidentiality wrt. servers
- PEPPeR — Dimitriou et al. @ MobiSys 2012
  - querier privacy (only)
  - crypto tokens based on blind signatures
  - communication overhead MN $\leftrightarrow$ querier
- **PEPSI** — De Cristofaro and Soriente @ WiSec 2011
  - first cryptographically provable security
  - privacy for both mobile nodes and queriers
  - simple architecture with trusted key generation, but untrusted service provider

PEPSI Architecture



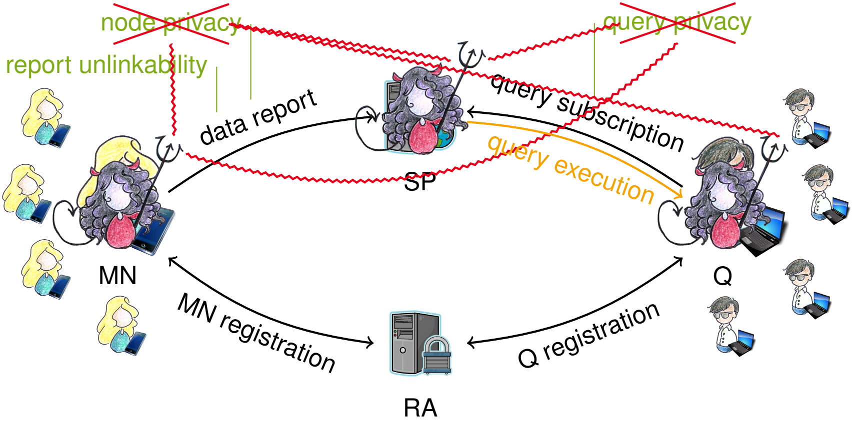
PEPSI Architecture



instantiation based on modified Boneh–Franklin identity-based encryption

(identity $\hat{=}$ “temperature in Heraklion”)

Limitations of PEPSI



no **collusion resistance**
(realistic thread)

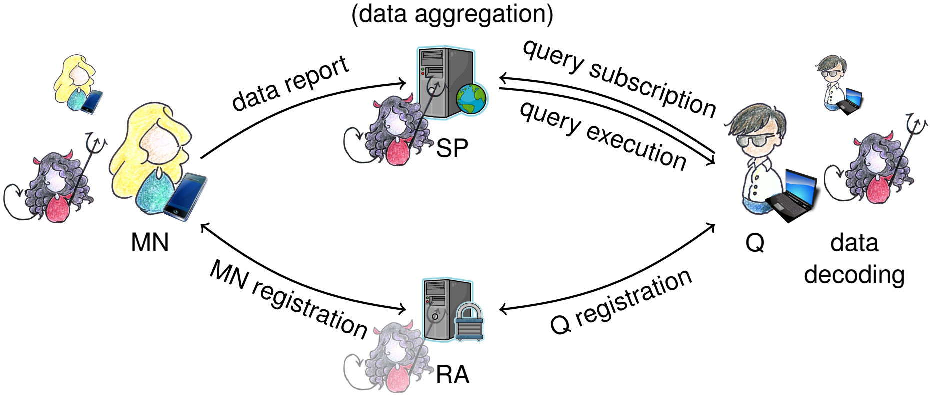
no **data aggregation**
(added performance + privacy)

PEPSICo

Revised Security Model for Participatory Sensing

PEPSI architecture

+ formal model + collusion resistance + data aggregation (optional)

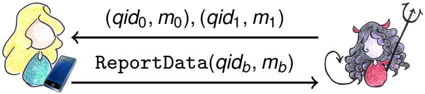


PEPSICo

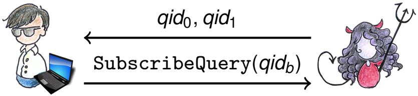
Node Privacy, Query Privacy & Report Unlinkability (idea)

TECHNISCHE UNIVERSITÄT DARMSTADT

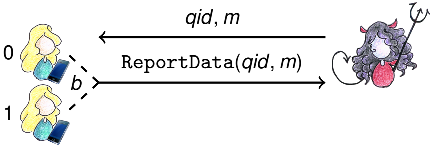
Node Privacy: hides both message and query identity of a report from SP, unauth. Qs and other MNs, even colluding.



Query Privacy: hides query identity of a subscription from SP, MNs and other Qs, even colluding.



Report Unlinkability: prevents linkage of two reports as originating from same MN by any other party, even colluding and including RA.



PEPSI: insecure PEPSICo instantiation, collusion attacks on node + query privacy

A Generic Solution

TECHNISCHE UNIVERSITÄT DARMSTADT

Preliminaries

- **Identity-Based Encryption (IBE)**
  - Setup(1^n) $\rightarrow$ (mpk, msk)
  - Extract(mpk, msk, id) $\rightarrow$ sk_{id}
  - Enc(mpk, id, m) $\rightarrow$ c
  - Dec(mpk, sk_{id}, c) $\rightarrow$ m

- **Security Notions for IBE**
  - indistinguishability (of message encryptions) — IND-ID-CPA/-CCA
  - anonymity (of identites used to encrypt) — ANO-ID-CPA/-CCA
  - indistinguishability + anonymity — ANO-IND-ID-CPA/-CCA

A Generic Solution

$\mathrm{PI}_{\mathrm{IBE}}$ Scheme

TECHNISCHE UNIVERSITÄT DARMSTADT

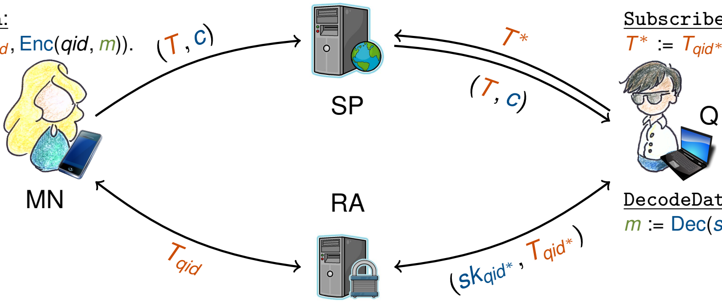
Ingredients: IBE scheme $\mathcal{E}$, pseudorandom function (PRF) $f: \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^n$.

ExecuteQuery: If $T = T^*$ return (T, c).

AggregateData: $(T', c') := (T, c_1 \circ \cdots \circ c_\ell)$.

ReportData:
$(T, c) := (T_{qid}, \mathrm{Enc}(qid, m))$.

SubscribeQuery:
$T^* := T_{qid^*}$.



DecodeData:
$m := \mathrm{Dec}(sk_{qid^*}, c)$.

Setup: RAsk := (msk, k), RApk := mpk. RegisterMN: $T_{qid} := f_k(qid)$. RegisterQ: (sk_{qid^*}, T_{qid^*}).

A Generic Solution

$\mathsf{PI}_{\mathsf{IBE}}$ Scheme

TECHNISCHE UNIVERSITÄT DARMSTADT

Security Analysis

CPA/CCA flavor

- **Node Privacy**, if
  - $\mathcal{E}$ is ANO-IND-ID-CPA/-CCA (hides message)
  - f is pseudorandom (hides query identity)
- **Query Privacy**, if
  - f is pseudorandom (hides query identity)
- **Report Unlinkability**
  - unconditional (no MN-specific information)

Concrete Instantiations

TECHNISCHE UNIVERSITÄT DARMSTADT

With Boneh–Franklin IBE Scheme ($\mathsf{PI}_{\mathsf{BF}}$)

- $\mathsf{Enc}(qid, m) := (g^r, m \oplus H_2(e(H_1(qid), \mathsf{mpk})^r))$, $\quad sk_{qid} := H_1(qid)^{\mathsf{msk}}$
- secure under Bilinear Diffie–Hellman (BDH) assumption in the ROM
- same high practical performance as PEPSI

Standard Model Instantiations

- proofs for generic construction are in standard model
- plug in any secure scheme in standard model (e.g., Boyen–Waters, Gentry)
- usually less efficient

Anonymous MN/Querier Registration

- use oblivious PRF + blind IBE

What about aggregation?

Adding Data Aggregation

TECHNISCHE UNIVERSITÄT DARMSTADT

Additively Homomorphic IBE Scheme (AIBE)

- based on Boneh–Franklin IBE scheme, secure under Decisional BDH in ROM
- messages are poly-size set $\mathcal{M} = \mathbb{Z}_M = \{0, \ldots, M-1\} \subseteq \mathbb{Z}_q$

- $\mathrm{Enc}(id, m) \rightarrow (g^r, \bar{g}^m \cdot e(H(id), \mathrm{mpk})^r)$, $\quad sk_{id} := H(id)^{\mathrm{msk}}$
- $\mathrm{Dec}(sk_{id}, c) \rightarrow \log_{\bar{g}}(c_2 / e(sk_{id}, c_1))$
  - needs to compute discrete log
  - but only for poly-size $\mathcal{M}$ and by querier, not MN
    — feasible even for full 32bit integers (<1sec on Intel i7 @2.9GHz)

- additive homomorphism (in $\mathbb{Z}_q$):

$$
\begin{aligned}
c_1 \cdot c_2 &= (g^{r_1} \cdot g^{r_2}, \bar{g}^{m_1} \cdot e(H(id), y)^{r_1} \cdot \bar{g}^{m_2} \cdot e(H(id), y)^{r_2}) \\
&= (g^{r_1+r_2}, \bar{g}^{m_1+m_2} \cdot e(H(id), y)^{r_1+r_2}) = \mathrm{Enc}(id, m_1 + m_2 \mod q)
\end{aligned}
$$

The $\mathsf{PI}_{\mathsf{AIBE}}$ Instantiation with Data Aggregation

TECHNISCHE UNIVERSITÄT DARMSTADT

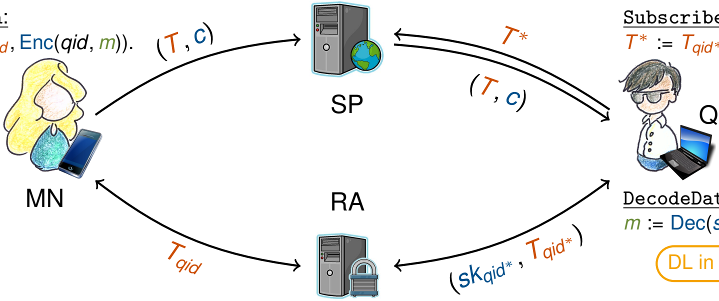
Ingredients: AIBE scheme, pseudorandom function (PRF) $f \colon \{0,1\}^n \times \{0,1\}^* \rightarrow \{0,1\}^n$.

`ExecuteQuery`: If $T = T^*$ return (T, c).

`AggregateData`: $(T', c') := \left(T, \left(\prod_{i=1}^{\ell} c_{i,1}, \prod_{i=1}^{\ell} c_{i,2}\right)\right)$.

`ReportData`: $(T, c) := (T_{qid}, \mathsf{Enc}(qid, m))$.

`SubscribeQuery`: $T^* := T_{qid^*}$.



`DecodeData`: $m := \mathsf{Dec}(sk_{qid^*}, c)$.

DL in $\mathbb{Z}_M$

`Setup`: RAsk := (msk, k), RApk := mpk. `RegisterMN`: $T_{qid} := f_k(qid)$. `RegisterQ`: (sk_{qid^*}, T_{qid^*}).

Performance Comparison

PEPSI vs. $\mathrm{PI}_{\mathrm{BF}}$ vs. $\mathrm{PI}_{\mathrm{AIBE}}$

	Computation			Communication		
Algorithm	PEPSI	$\mathrm{PI}_{\mathrm{BF}}$	$\mathrm{PI}_{\mathrm{AIBE}}$	PEPSI	$\mathrm{PI}_{\mathrm{BF}}$	$\mathrm{PI}_{\mathrm{AIBE}}$
`Setup`	2E	1E	1E	–	–	–
`RegisterMN`	–	1f	1f	n	n	n
`RegisterQ`	1E	1f+1E	1f+1E	2G	1G+n	1G+n
`ReportData`	1E+1P+2H	2E+1P+2H	3E+1P+1H	2n	1G+2n	2G+n
`SubscribeQuery`	1P+1H	–	–	n	n	n
`ExecuteQuery`	–	–	–	2n	1G+2n	2G+n
`DecodeData`	1P+1H	1P+1H	1P+1DL	–	–	–
`AggregateData`	n/a	n/a	≈ 0	n/a	n/a	–

E	modular exponentiation in $\mathbb{G}$ or $\mathbb{G}_T$	G	group element in $\mathbb{G}$ or $\mathbb{G}_T$
P	pairing evaluation	n	message length, Hash/PRF output length
H	hash function evaluation		
f	PRF evaluation		
DL	computation of discrete logarithm		

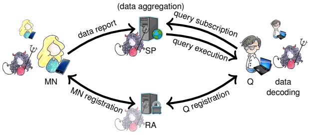
- $\mathrm{PI}_{\mathrm{BF}} \approx$ PEPSI wrt. computation and communication cost
- $\mathrm{PI}_{\mathrm{AIBE}}$: DL computation on decode, but aggregation is cheap
  + saves factor ℓ for decode and communication

Summary

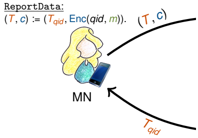
participatory sensing: privacy is important, collusion attacks are a realistic threat

We

- propose a revised model for privacy-enhanced participatory sensing with collusion resistance



- provide a generic solution and concrete instantiations with practical performance



- enable data aggregation in the model with an additively homomorphic IBE scheme

$$c_1 \cdot c_2 = (g^{r_1} \cdot g^{r_2}, \bar{g}^{m_1} \cdot e(H(id), y)^{r_1} \cdot \ldots$$

Thank You!