Resilient and Private Aggregation

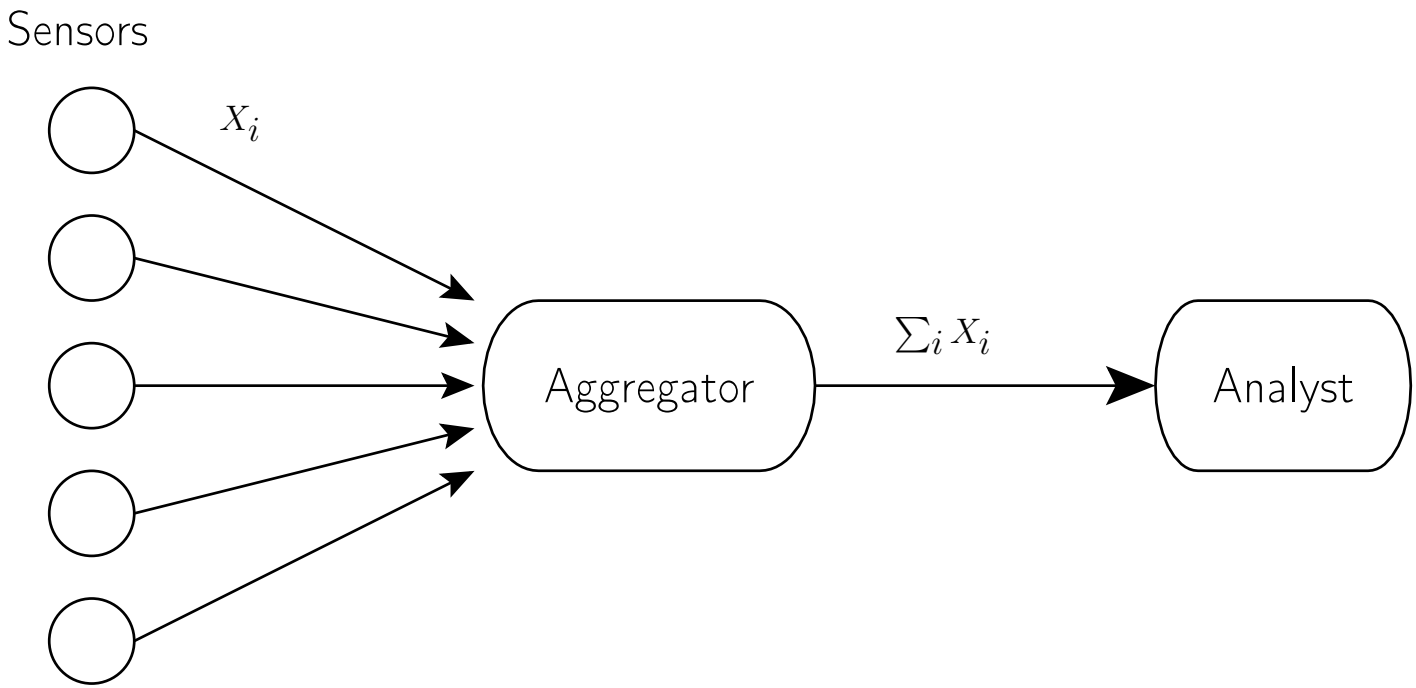
Mathieu Cunche, Cédric Lauradoux, Marine Minier

INSA-Lyon/CITI Lab., INRIA

March 21, 2013

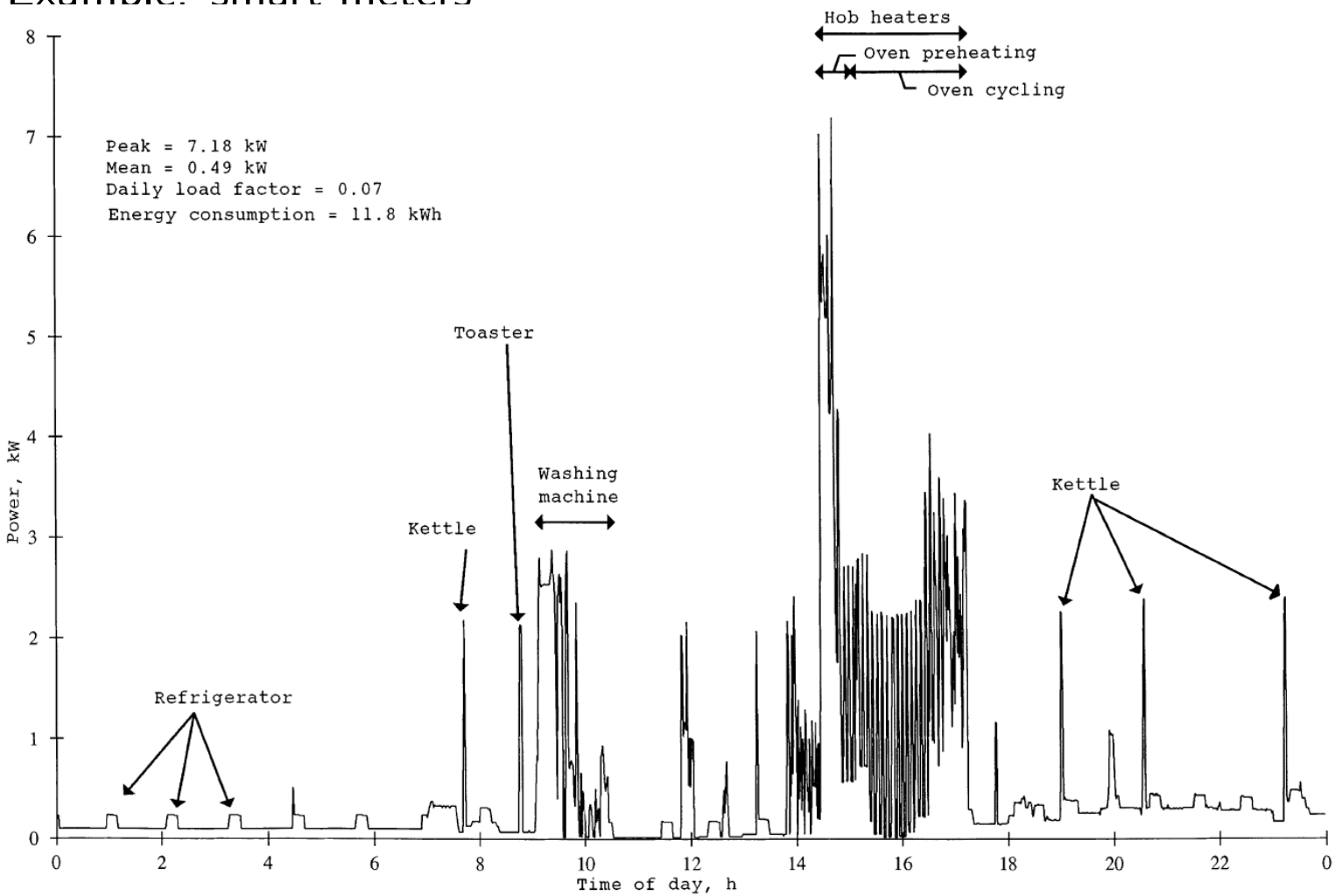
Sensor networks I

- Sensor network aggregation system:
  - Set of data sources (sensors)
  - An aggregator and an analyst (data consumer)
  - An aggregation function (in this work *average* or *sum*)



Privacy in sensor network I

- Data collected by sensors can lead to privacy breach
  - Example: smart-meters



Privacy in sensor network II

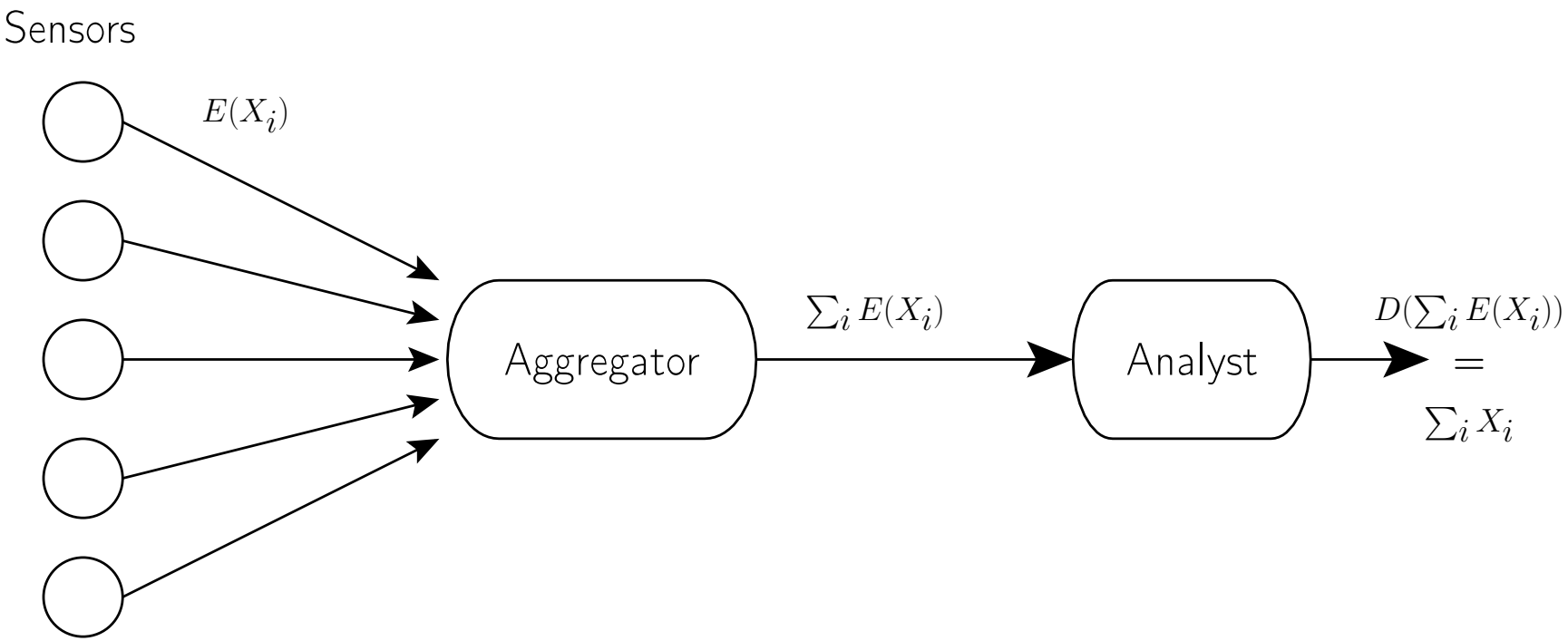
Definition (Private aggregation)

The aggregation system should prevent any privacy breach. Individual sensed values should not be available in clear outside of the sensor.

- Solutions for Privacy-preserving aggregation
  - Differential privacy [AC11]
  - Homomorphic encryption [GJ11, CCMT09]
  - Secure Multiparty computation [BSMD10]
  - And many others ...

Privacy in sensor network III

- Private aggregation using homomorphic encryption
  - Homomorphic encryption : $E(X+Y) = E(X) \times E(Y)$
  - Introduction of a new party: the aggretator
  - System description:
    1. Sensors send encrypted report $E(X_i)$ to aggregator
    2. Aggregator perform private aggregation : $\sum_i E(X_i) = E(\sum_i X_i)$
    3. Analyst decrypt the aggregated data $D(E(\sum_i X_i)) = \sum_i X_i$



Resilient aggregation I

- Sensors can be faulty
  - Send bogus data instead of the real measurements
- Sensors can be controlled by an attacker
  - Send data chosen by the attacker
  - Attacker objective: make the aggregate value deviate as much as possible from its real value
  - Attacker motivation: DoS, trigger a reaction (heating system)

Definition (Resilient aggregation)

The system must be robust against faulty/compromised sensors. A small number of bogus data must not make the aggregate significantly deviate from its real value.

Resilient aggregation II

- "Resilient aggregation in sensor networks" [Wag04]
  - Introduction of the resilient aggregation problem
  - Study the resilience of common aggregation functions: *Avg*, *Min*,*Max*, *Median*, etc
  - Propose solution to *Robustify* aggregation functions
- Truncation for a resilient average
  - Verify that reported values fall in a valid range $[minV, maxV]$
  - Discard bogus values
  - Attacker must control a large number of sensors in order to significantly deviate the aggregate

Private and Resilient aggregation I

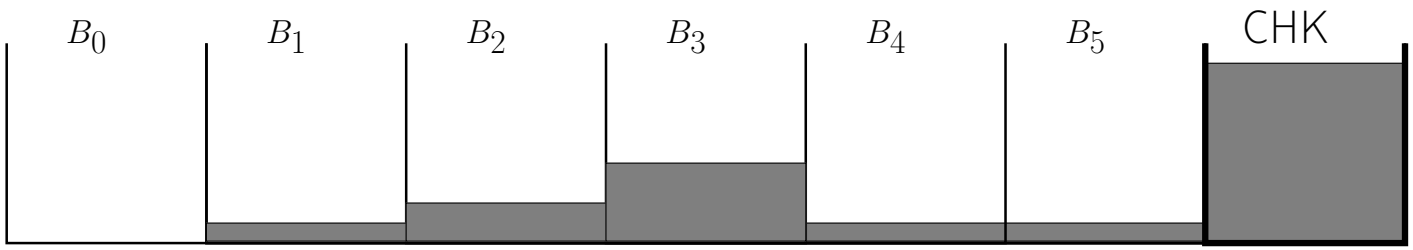
- Need for Private and Resilient aggregation system
  - More and more sensor system interacts with human activity
  - Sensors are cheap: prone to fault and wide range of attacks
- Privacy and Resilience are antagonistic goals
  - Bogus report filtering require the aggregator to have the plaintext information
  - Privacy requirement do not allow for individual reports to be available in plaintext

Private and Resilient aggregation II

- "Private and Resilient" aggregation systems have already been proposed
- But they rely on different definition of *Resilience*
  - Resilience against faulty links
  - Resilience against compromised/faulty aggregator
  - etc ...
- No Private and Resilient aggregation system has been proposed so far
  - Where *Resilient* means robust against faulty/compromised sensors

ABBA I

Figure: *ABBA: A Balls and Bins Approach to Secure Aggregation in WSNs* [CS08] by Castelluccia et al



ABBA II

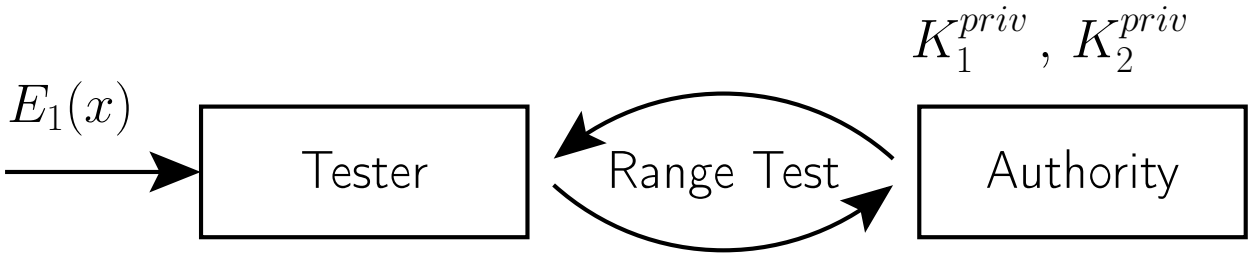
- Privacy preserving aggregation
  - Homomorphic encryption
- Analyst can detect bogus data in the aggregate
  - Checksum-based verification
- A corrupted aggregate must be discarded
  - Cannot filter out bogus values
  - Vulnerable to DoS attack

A Private and Resilient aggregation

- We propose a new system that provides:
  - Privacy: individual sensor data
  - Resilience: limited impact of faulty/compromised sensors on the aggregated result
- Core idea: filter out bogus values
  - Genuine values lies in a valid interval
  - Ex.: temperature sensed in a building should lie between 0 and 40 degrees
  - Rely on a Private Range Test ...

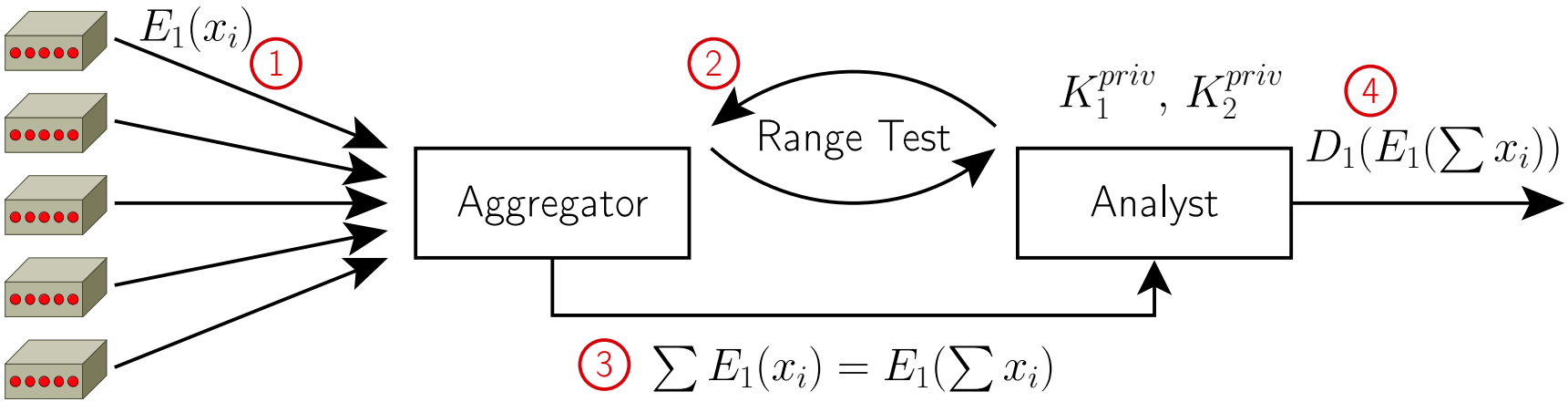
Private Range-Test

- Range-Test on encrypted values
  - Given $E(x)$, $I = [a, b]$, can we answer $x \in I$? without decrypting $E(x)$
- The *Private Range-Test protocol* [PBDO06]
  - Two additive homomorphic cryptosystems $(E_1, D_1), (E_2, D_2)$
  - Two parties: a *Tester* and an *Authority*.
  - An interactive test $RT(E_1(x), [a, b])$
    - Return TRUE if $x \in [a, b]$, FALSE otherwise
    - Does not reveal any other information on x



Proposed system

- A Private and Resilient Aggregation system
  - A privacy-preserving aggregation with a homomorphic cryptosystem
  - A Range-Test to filter out bogus data
- Protocol description
  1. Sensor send encrypted data
  2. Aggregator and Analyst perform a range test to filter out bogus values
  3. Aggregator send the encrypted sum of "valid" values
  4. Analyst decrypt the aggregated result



Performances I

- System: n sensors, a valid range $[a, b]$, average $X = \sum x_i / n$
- Resilience: maximum deviation of the aggregate
  - a single bogus sensor $\Delta_1 = (b - a)/n$
  - t bogus sensors $\Delta_t = t\,(b - a)/n$
- Privacy: sensor values remain private
  - Homomorphic encryption ensure data confidentiality
  - Analyst has the decryption key, but only receive the aggregated result

Performances II

- Complexity:

Table: Processing cost for a single value

	Enc.	Dec.
Data producer	1	-
Aggregator (A_1)	$2 + 18t$	-
Analyst (A_2)	$8t$	$24t$

  - Limited cost for the sensor: only one encryption
  - Potentially high cost for $A_1/A2$: multiple Enc/Dec required
- Aggregator and Analyst: standard computers
  - One Range Test runs in 216 ms on an Intel i7
- Data producer: sensor with limited ressources
  - Elliptic Curve El-Gamal feasible on sensors [UWL+07]
  - TinyECC implementation [LN08]

Aggregated Range-Test I

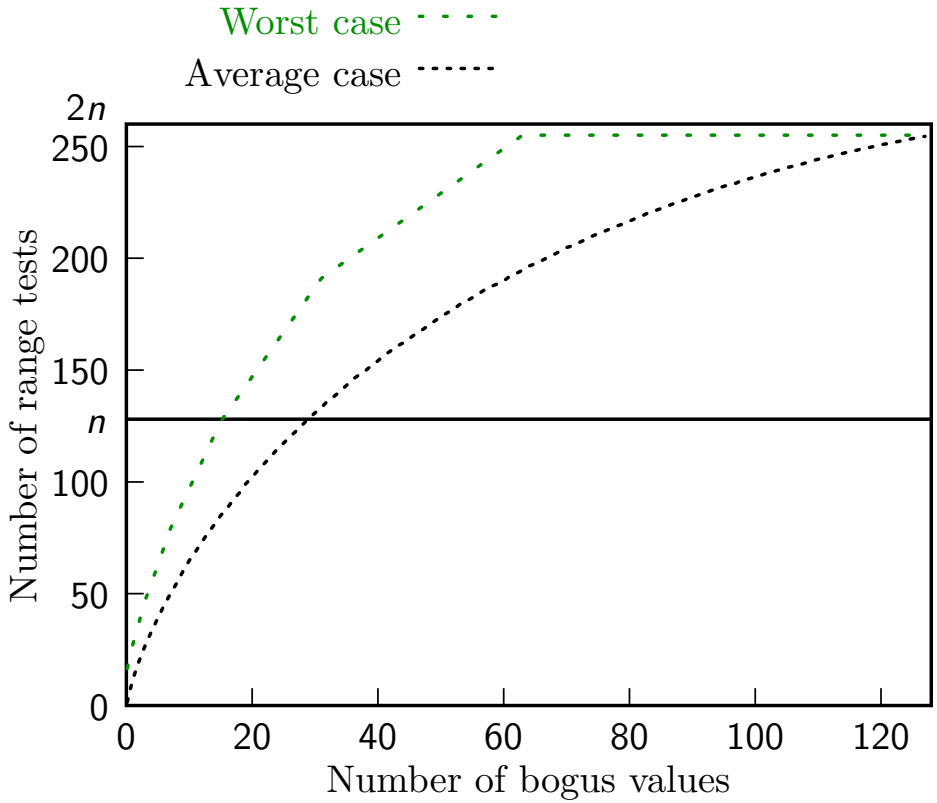
- Aggregated Range-Test
  - Aggregate sub-group of sensor values
  - Perform Range-Test on an sub-aggregate of k values

$$RT\left(\sum_{i \in [pk\,..\,(p+1)k-1]} E_1(x_i), [ka, kb]\right)$$

  - If aggregated RT return false, then perform a dichotomic search to identify bogus values

Aggregated Range-Test II

- Complexity: reduce the load of the Aggregator and the Analyst



- If no bogus value detected: reduce by a factor k the number of Range-Test
- If bogus values detected: more efficient up to a certain fraction of bogus values

Aggregated Range-Test III

- Resilience: maximum deviation of the aggregate
  - a single bogus sensor $\Delta_1 = k(b-a)/n$
  - t bogus sensors $\Delta_t = t\,k\,(b-a)/n$
- Privacy: identical to non-aggregated RT
  - No additional information obtained from aggregated Range-Test
  - Even when combining results of multiple aggregated Range-Tests

Conclusion

- Privacy and Resilience are not necessarily antagonistic goals
- A Private and Resilient aggregation system for the average
  - Based on a Private Range-Test protocol
  - Optimisation: aggregated Range-Test
- Future work
  - Other aggregation functions
  - Other data types with different validity domains

Bibliography I

Gergely Ács and Claude Castelluccia.

I have a DREAM!: differentially private smart metering.
In *International conference on Information hiding - IH'11*, Lecture Notes in Computer Science 6958, pages 118–132, Prague, Czech Republic, 2011. Springer-Verlag.

Martin Burkhart, Mario Strasser, Dilip Many, and Xenofontas Dimitropoulos.

Sepia: privacy-preserving aggregation of multi-domain network events and statistics.
In *Proceedings of the 19th USENIX conference on Security*, USENIX Security'10, pages 15–15, Berkeley, CA, USA, 2010. USENIX Association.

Claude Castelluccia, Aldar C-F. Chan, Einar Mykletun, and Gene Tsudik.

Efficient and provably secure aggregation of encrypted data in wireless sensor networks.
ACM Trans. Sen. Netw., 5(3):20:1–20:36, June 2009.

Claude Castelluccia and Claudio Soriente.

Abba: A balls and bins approach to secure aggregation in wsns.
In *6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops, WIOPT 2008*, pages 185–191, Berlin, Germany, March 31 - April 4 2008.

Flavio D. Garcia and Bart Jacobs.

Privacy-friendly energy-metering via homomorphic encryption.
In *Proceedings of the 6th international conference on Security and trust management*, STM'10, pages 226–238, Berlin, Heidelberg, 2011. Springer-Verlag.

An Liu and Peng Ning.

Tinyecc: A configurable library for elliptic curve cryptography in wireless sensor networks.
In *Proceedings of the 7th international conference on Information processing in sensor networks*, IPSN '08, pages 245–256, Washington, DC, USA, 2008. IEEE Computer Society.

Bibliography II

Kun Peng, Colin Boyd, Ed Dawson, and Eiji Okamoto.

A Novel Range Test.

In Lynn Margaret Batten and Reihaneh Safavi-Naini, editors, *Australasian Conference on Information Security and Privacy ACISP 2006*, volume 4058 of *Lecture Notes in Computer Science*, pages 247–258. Springer, 2006.

Osman Ugus, Dirk Westhoff, Ralf Laue, Abdulhadi Shoufan, and Sorin A. Huss.

Optimized Implementation of Elliptic Curve Based Additive Homomorphic Encryption for Wireless Sensor Networks.

In *2nd Workshop on Embedded Systems Security, WESS'2007*, Salzburg, Austria, October 4 2007.

David Wagner.

Resilient aggregation in sensor networks.

In *ACM workshop on Security of ad hoc and sensor networks - SASN '04*, pages 78–87, Washington DC, USA, October 2004.