Secure Information Aggregation for Smart Grids using Homomorphic Encryption

Fengjun Li, Bo Luo, Peng Liu

PRESENTER: EMAN HAMMAD

Submitted in Partial Fulfillment of the Course Requirements for
ECEN 689: Cyber Security of the Smart Grid
Instructor: Dr. Deepa Kundur

Outline

- Motivation, approach
- Related works
- Background:
  - Homomorphic Encryption
  - Honest-but-curious Model
- Secure Information Aggregation
- Example
- Analysis and discussion
- Personal insight and assessment
- References

Motivation, approach

- Instant aggregation of power usage data:
  - At different levels: Neighborhood , subdivision, district, city etc. and at different frequencies.
- Essential for:
  - Monitoring and predicting power consumption.
  - Allocating and balancing loads and resources.
  - Administering power generation, etc.
- Goal: efficient and secure data aggregation for smart grids.
- Approach:
  - In-network distributed aggregation.
  - Homomorphic encryption.

Related Works

- Various in-network data aggregation approaches:
  - For sensor networks, sensors are limited by battery and resources.
  - Sensors in the network are usually trusted, and security is against eavesdroppers and tampering attacks using fake inputs.

- Smart Grids:
  - Power of the smart meter is not a concern, but communication bandwidth is, specially when frequent aggregation is required.
  - Power usage is considered a privacy of the user.

  - *Traditional tree-based aggregation on plaintext does not apply.

Background: Homomorphic Encryption

- Homomorphic encryption:
  - A form of encryption where specific algebraic operation performed on the plaintext is equivalent to another (possibly different) algebraic operation performed on the ciphertext.
  - Given a homomorphic encryption function $E()$, and two messages $x, y \in Z_N$

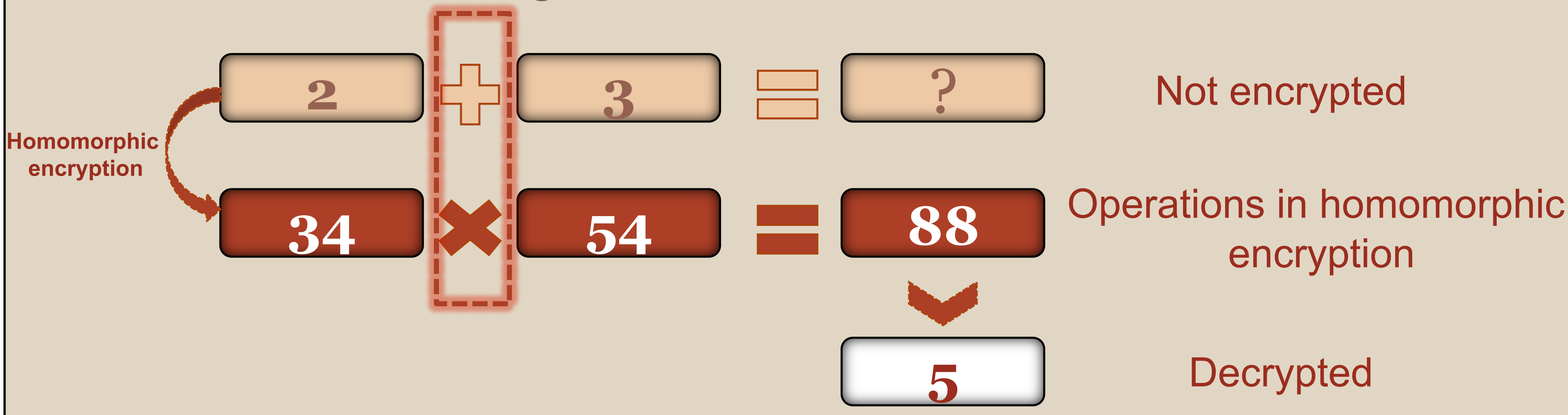
$$E_k(x \star y) = E_{k1}(x) \circ E_{k2}(y)$$

Without knowing the plaintext x, y or the private key.

  - Used for privacy-preserving operations, voting.
  - Known schemes: RSA, El Gamal, Paillier, Naccache-Stern, BGN etc.
  - Paper adopts Paillier scheme.

Homomorphic encryption

- Paillier cryptosystem:
  - Invented in 1999 by Pascal Pailier.
  - Has additive homomorphic property.
    - Given only the public-key and the encryption of m_1 and m_2, one can compute the encryption of $m_1 + m_2$.[2]
  - Indeterministic:
    - the same message will be encrypted into different ciphers using different random blinding factors.

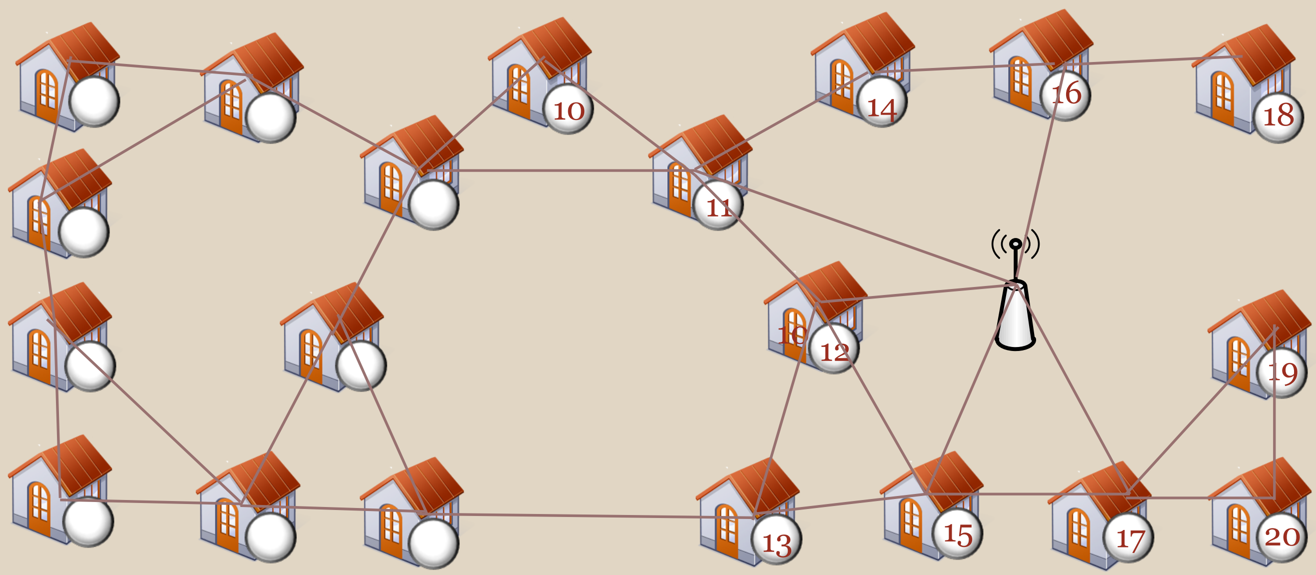


Background: Honest-but-curious Model

- Honest-but-curious model:
  - All parties are assumed to follow protocol properly "honest".
  - Keep all inputs from other parties and all intermediate computation results "curious".
- Honest-but-curious smart meters:
  - Do not tamper with the aggregation protocols
  - Do not drop or distort any source value or intermediate result.
  - Will try to infer others' electricity usage from messages routed through them

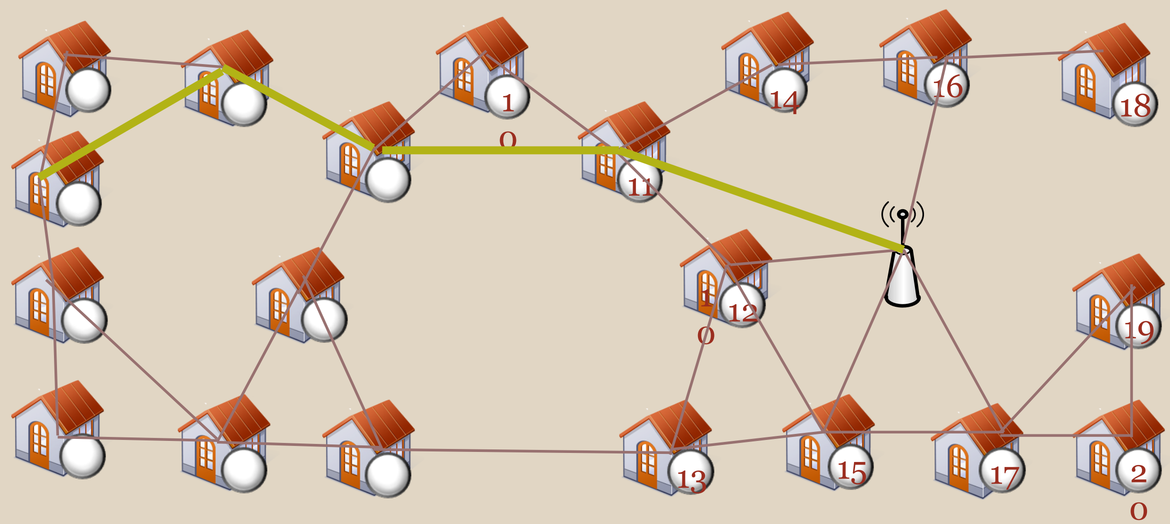
Secure Information Aggregation

- Smart Grid Communication Infrastructure:
  - Most popular: wireless-wired multi-layer architecture.
    - Wireless: smart meters in a neighborhood communication with a collector device.
    - Wired: collector device with the rest of the grid.



Secure Information Aggregation

- Data Aggregation: important type of query in Smart Grids.
  - Example: average power usage of the neighborhood.
  - Traditionally: every smart meter establishes a connection with the collector and uses it exclusively to report its data.
    - Excessive network traffic.
    - Overwhelming demands at the collectors.



Secure in-network incremental aggregation

- Approach:
  - Establish an aggregation tree.
  - Enroute meters to share the channel.
  - Ensure privacy using homomorphic encryption.
    - With reasonable computation overhead.

The Aggregation Tree

- To enable in-network Aggregation:
  - Aggregation path:
    - All smart meters in the neighborhood.
  - For each aggregation task:
    - All or subset of nodes on the aggregation path participate.

The Aggregation Tree

-
- Considering the smart meter network as a graph:
  - Graph $G(V, E)$:
    - V, set of smart meters (vertices).
    - E, set of available wireless links (edges) between any two smart meters.
    - Graph should be connected; every smart meter should have at least one communication path to the collector.

The Aggregation Tree (cont.)

- The Aggregation Tree:
  - A spanning tree of the graph with minimal subset of *E* that connects all the vertices.
  - Always roots at the collector node; which initializes all aggregation tasks and collects the results.
  - Aggregation is recursively calculated in a bottom-up manner; every nodes takes input from itself and its children nodes, aggregate the data and sends the result to its parent node.

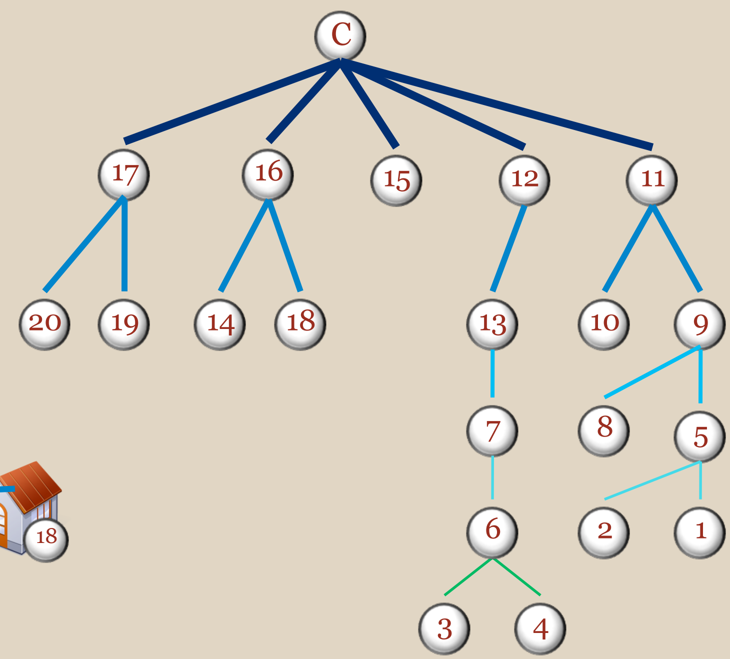
- Collector device:
  - Has the network graph of the entire neighborhood.
  - The aggregation tree is constructed locally at the collector.
  - An aggregation tree remains valid for an extended period of time.

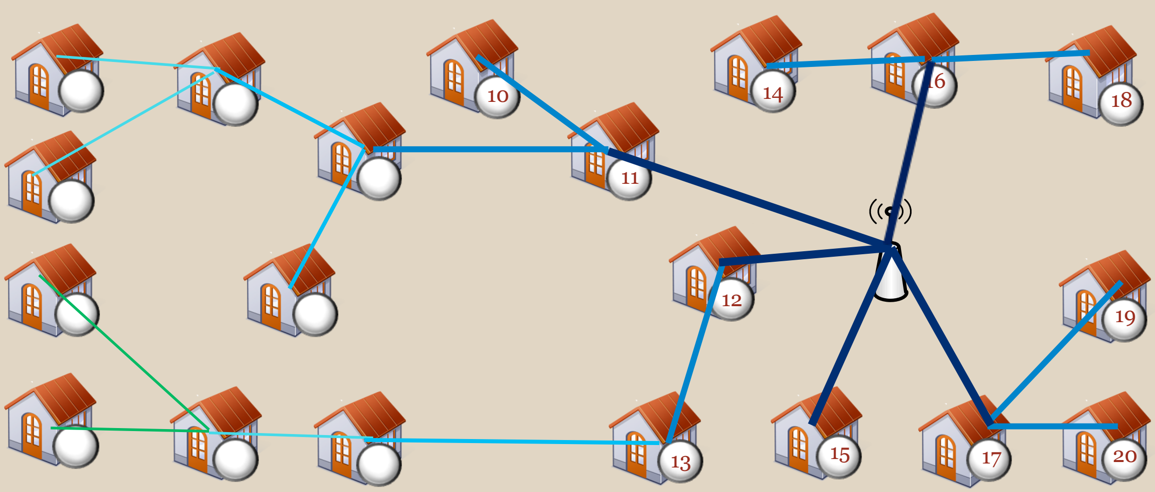
Constructing the Aggregation Tree

- Algorithm goals:
  - Height of the tree should be small.
  - An interior node should not have too many children, to avoid excessive computation and communication load.
- Approach:
  - Breadth-first traversal of the graph, starting at the collector node.
  - If node *K* has too many children rebalance the three.
    - If a child of *K* is connected to a less populated sibling of K, move child to that sibling (will not increase the height of the tree).
    - If *K* still has too many children, check if a child is also connected to another child of K, and move it to that child (may increase the height of the tree).

Example: constructing the Aggregation Tree

Aggregation tree constructed from the graph





Breadth-first traversal of the network graph

In-network aggregation using homomorphic encryption

- Having the aggregation tree:
  - Construct operation plans for participating nodes (smart meters).
  - Deploy the operation plans in a top-down manner.

In-network aggregation using homomorphic encryption

- An operation plan for a smart meter:

$$\{T_{ID}, Trigger, Data, Collect, Operation, Destination, Key\}$$

T_{ID}, arbitrary unique identifier to identify message.

Trigger, defines when the aggregation will be conducted; periodically, upon collector request, or at a particular time. Time of local data reading, important in time-sensitive tasks.

Data, what information from the local smart grid will be collected in the aggregation.

Collect, tells a smart meter to wait for input from its children in the aggregation.

Operation, what operation to be performed; pre-processing, encryption and operations for aggregation.

Destination, the parent node, to whom the output from *Operation* will be submitted.

Key, a public key from the collector to be used to encrypt the local data.

In-network aggregation using homomorphic encryption

- Output message from a participating node:
  - Is constructed as

$$\{T_{ID}, TS, Data\}$$

  - Where TS is the timestamp of local data retrieval. This timestamp is used for synchronizing different occurrences of repeating tasks.

TEXAS A&M UNIVERSITY

Examples

- Example:
  - To calculate the total output power (KW) at time t_0 in the entire neighborhood:
    - Aggregation plan at node 9 is: $\{tid, t_0, power, \{N_5, N_8\}, Enc_K(power) \times I_5 \times I_8, N_{11}, K\}$
    - When node 9 receives the aggregation plan:

1. It retrieves its own power at t_0.
2. It encrypts the reading with K to get local input $C_{p9} = E_K(P_9)$.
3. Node 9 then waits for input from nodes 5,8.
4. After receiving C_{o5} ,C_{o8}, node 9 calculates $C_{09} = C_{p9} \times C_{05} \times C_{08}$.
5. Node 9 submits C_{o9} to Node 11.

TEXAS A&M UNIVERSITY

Analysis

- Comparing:
  - The in-network aggregation with homomorphic encryption to traditional aggregation approach.

- Network:
  - Traditional: messages from all smart meters are routed to collector simultaneously. Let $\bar{h}$ be the average number of hops for each message to the collector, assuming number of nodes to be N, total load on the network will be $\bar{h} * N$.
  - In-network aggregation, total load will be N hops.

Analysis (cont.)

- Scalability, bottleneck and robustness:
  - Overall scalability highly depends on the smart meter network topology.
  - In-network aggregation:
    - For a well designed network, the aggregation tree will be wide and shallow. The longest path in an aggregation process is the graph diameter, grows at $\sqrt{N}$.
    - Almost no bottleneck in the in-network aggregation; since most computations are distributed, and also with the rebalance scheme.
    - If one start meter fails, failure is detected immediately by its parent in aggregation and reported to the collector, the collector updates the aggregation tree and re-issues the query.

Analysis (cont.)

- Security and privacy analysis:
  - The Paillier cryptosystem:
    - Semantically secure: polynomial time adversary who intercepts communication cannot derive significant information about the plaintext from the ciphers and public key.
    - Resilient to dictionary attacks; based on the use of the blinding factor r, same data will be encrypted to different ciphers with different r.
    - WARNING: all homomorphic encryption systems are malleable; given cipher and public key, an adversary could generate another cipher that decrypts to another meaningful plaintext in the same domain as the original plaintext. Hence, a dishonest meter or fake meter could falsify its data causing inaccurate aggregation result. NOT considered by in-network aggregation, can be solved by increasing physical and software security of smart meters.

Analysis (cont.)

- Computation:
  - Asymmetric encryption (homomorphic encryption):
    - Is more computationally expensive than symmetric encryption (AES and triple-DES).
    - Traditional (symmetric):
      - Each smart meter encrypt its message, collector to decrypt N messages.
    - In-network aggregation:
      - Each smart meter encrypt its message once, and the collector decrypts one message (result of aggregation).
      - Distributes the computation of the aggregation from collector to intermediate smart meters (with low overhead).

Personal assessment

- Authors successfully extend aggregation concepts from sensor networks into a smart grid framework, carefully handling smart grid issues(smart meters, privacy, etc).
- Authors fully understand the pros and cons of their proposed system, and include future research plans to cover the shortcomings.
- Authors did not provide a quantitative simulation results that show the gains in savings of computation, and the actual implementation a real smart grid system/subsystem.
- The proposed solution does not handle the Integrity aspect in the C-I-A security framework, since authors tried to carefully limit any overhead computations, yet this should be looked at.

References

1. Homomorphic encryption, http://en.wikipedia.org/wiki/Homomorphic_encryption
2. What is Homomorphic Encryption, and Why Should I Care? http://blogs.teamb.com/craigstuntz/2010/03/18/38566/
3. Fengjun Li; Bo Luo; Peng Liu; , "Secure Information Aggregation for Smart Grids Using Homomorphic Encryption," *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on* , vol., no., pp. 327-332, 4-6 Oct. 2010 doi: 10.1109/SMARTGRID.2010.5622064
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5622064&isnumber=5621989