Private Data Aggregation with Groups for Smart Grids in a Dynamic Setting using CRT

IEEE Workshop on Information Forensics and Security

16-19 November, 2015

Dr. Zeki Erkin

Cyber Security Group
Delft University of Technology

TU Delft

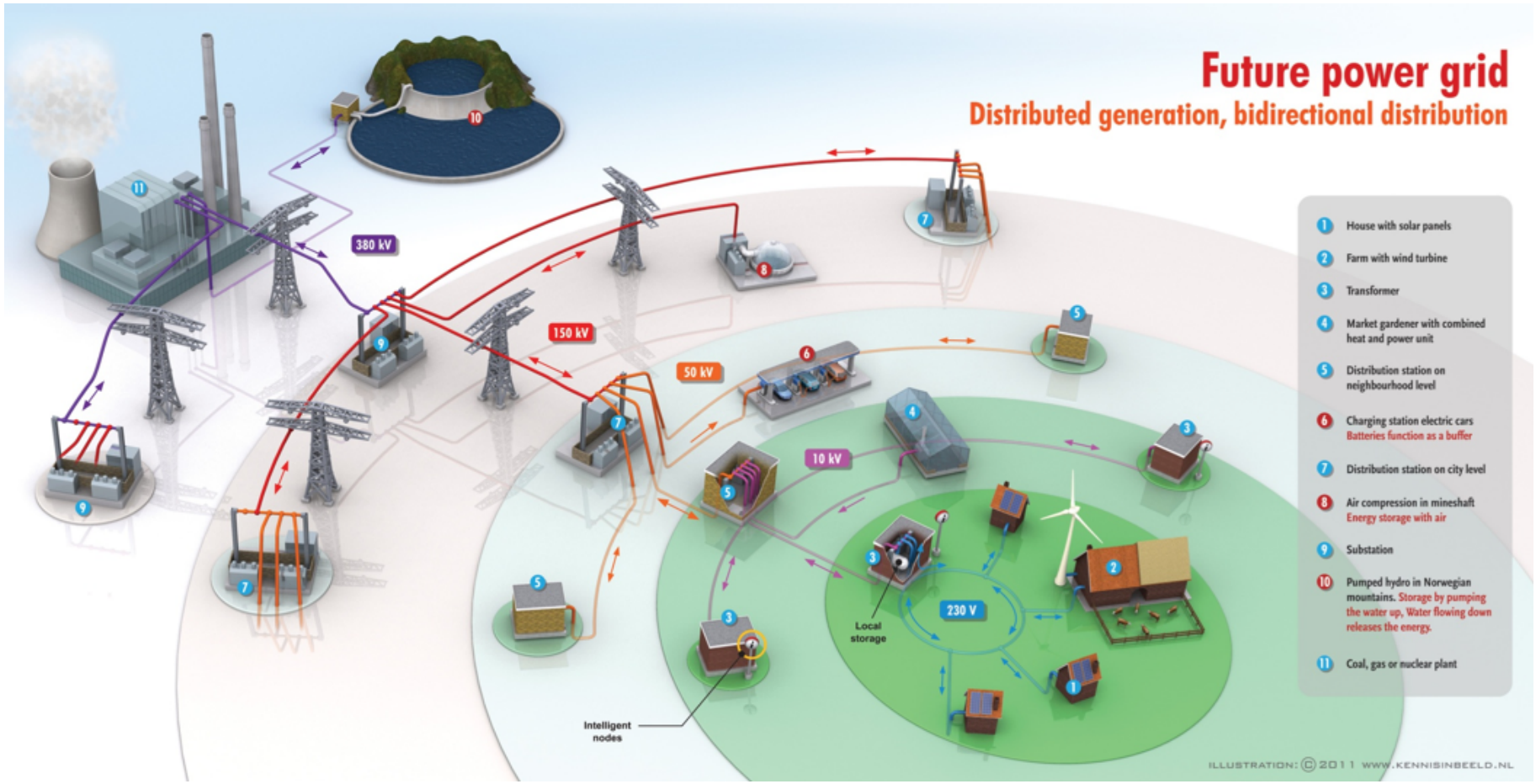
TU Delft

Outline

- Smart Grid
  - Why it is good
  - How it works
  - What we want from it
- Privacy in Smart Grids
  - Goal
  - Challenges
- Private data aggregation
  - State-of-the-art
  - A better protocol in a dynamic setting
- Conclusions

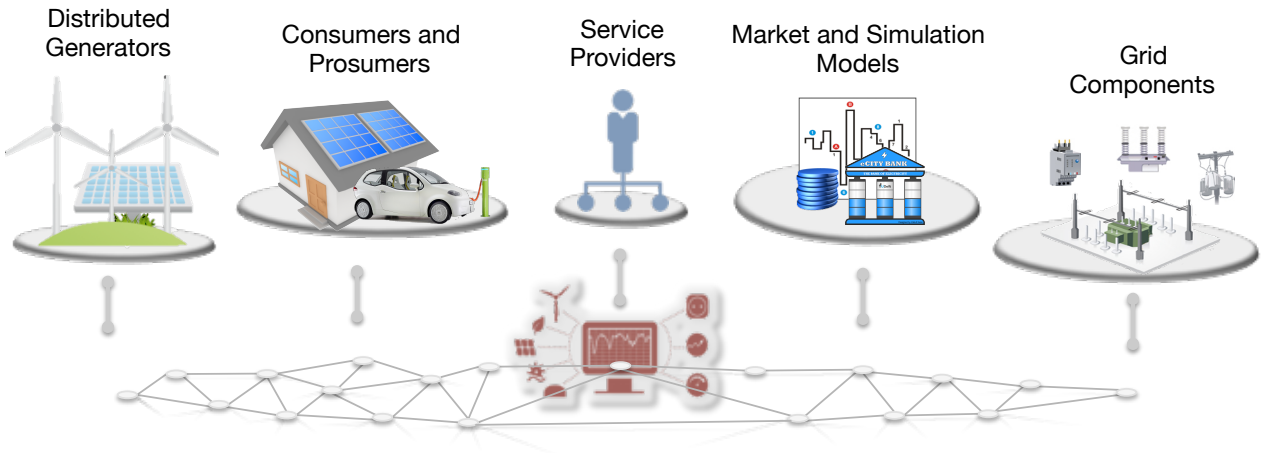
TU Delft

Smart Grids



TU Delft

Business

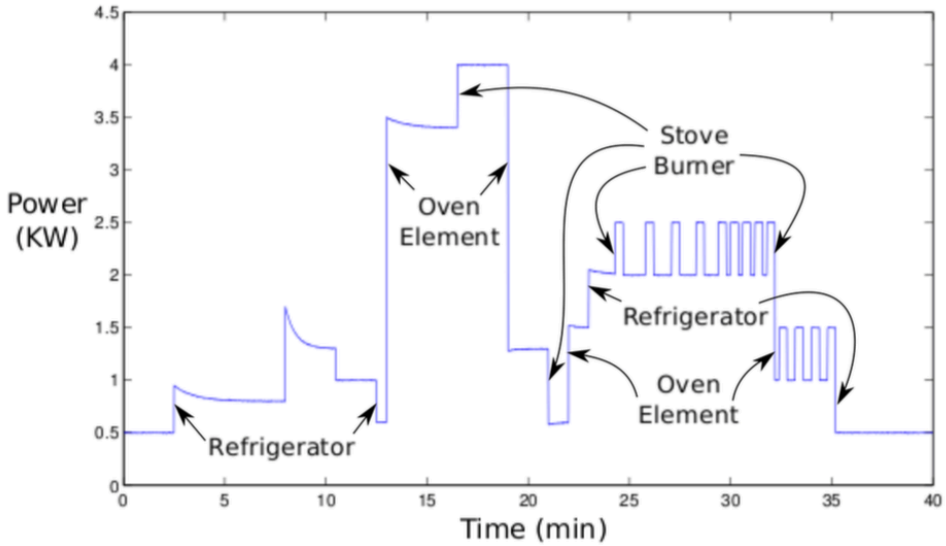


- Monitoring
  - Load balancing, prediction, profiling etc
- Billing
  - Variable pricing
- Data utilization
  - New business opportunities (e-commerce, customization, smart home)
- New concepts
  - Micro-grids, virtual grids, brokers, auctions, dynamic players

TU Delft

Privacy issues

- Smart meters: limited devices
  - Collects data every 15 mins (100 ms by design)
  - Dutch Parliament Bill, 2009: optional deployment
- What is the problem?
  - Long list of security issues (devices, sca, protocols…)
  - Privacy is our focus

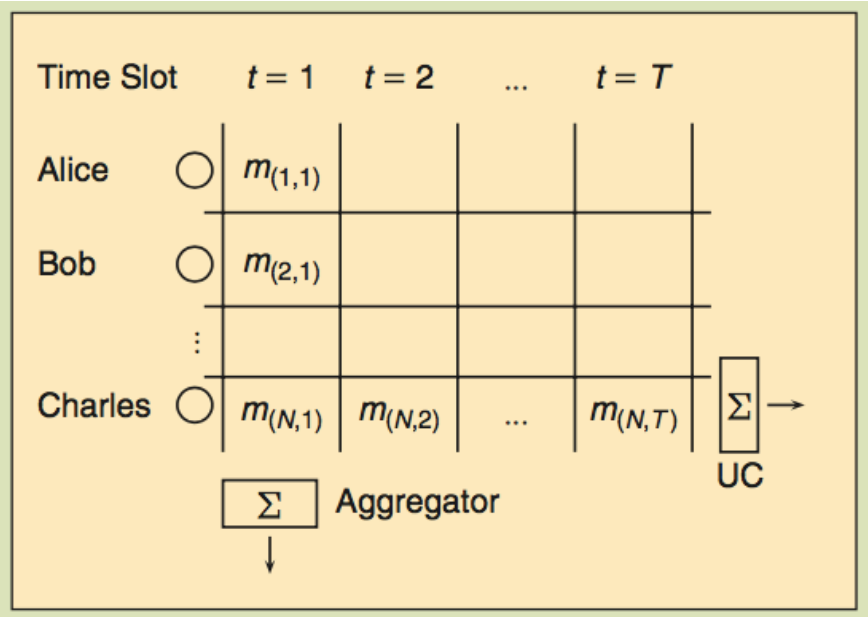


Erkin, Z., Troncoso-Pastoriza, J.R., Lagendijk, R.L., Perez-Gonzalez, F., "Privacy-preserving data aggregation in smart metering systems: an overview," Signal Processing Magazine, IEEE , vol.30, no.2, pp.75,86, March 2013

Data Aggregation

- Utility provider
- Aggregator(s)
- Households

- Can we compute aggregated data without learning individual consumption?
  - Spatial?
  - Temporal?
  - Missing data?



TU Delft

Prior Work

- GJ10: homomorphic encryption and secret sharing
  - Very inefficient
- KDK11: ElGamal
  - A look-up table is necessary and thus, the range
- ET12: Modified homomorphic encryption
  - Requires 3rd party in case of missing data

F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in Proc. 6th Workshop Security and Trust Management (STM 2010), (LNCS), vol. 6710, pp. 226–238.

K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in Privacy Enhanced Technologies Symposium, Waterloo, Canada, 2011, pp. 175–191.

Z. Erkin and G. Tsudik, "Private computation of spatial and temporal power consumption with smart meters," in Proc. Int. Conf. Applied Cryptography and Network Security, Singapore, 26–29 June 2012, pp. 561–577.

TU Delft

Additive Homomorphism

- Some cryptosystems preserve structure after encryption.

$$n = p \cdot q$$

$$\text{Public key: } (g, n)$$

$$\text{Private key: } (\lambda = lcm(p-1, q-1))$$

$$\mathcal{E}_{pk}(m, r) = g^m \cdot r^n \bmod n^2$$

$$(a^n)^\lambda \bmod n^2 = 1$$

Additive Homomorphism (Paillier '99)

$$\mathcal{E}_{pk}(m_1) \times \mathcal{E}_{pk}(m_2) = \mathcal{E}_{pk}(m_1 + m_2)$$

$$\mathcal{E}_{pk}(m)^c = \mathcal{E}_{pk}(m \cdot c)$$

Pascal Paillier, Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. EUROCRYPT 1999: 223-238

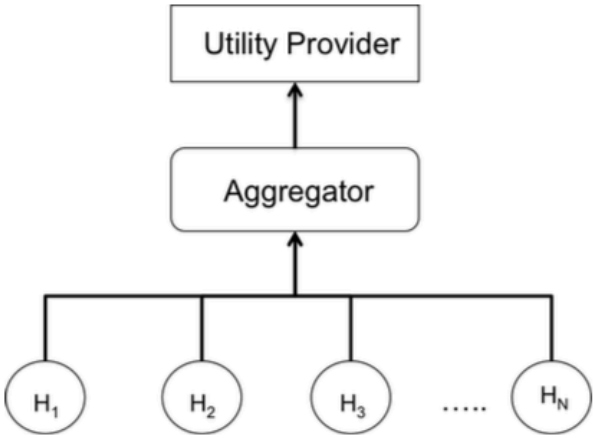
TU Delft

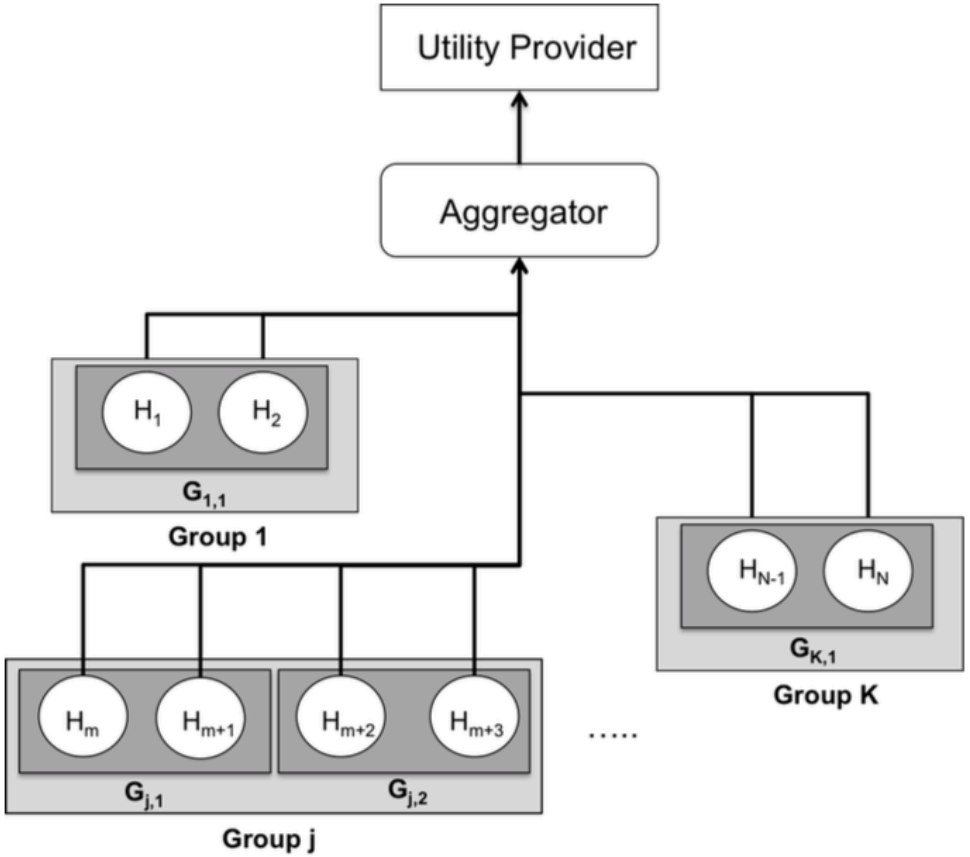
A New Scheme

- Groups
  - Aggregates data over groups and sub-groups
    - Households, schools, shops, hospitals etc
- Dynamic Environment
  - Should cope with missing data
- Efficiency
  - Fast enough for almost real time processing (100ms)

TU Delft

Scenario





TU Delft

Chinese Remainder Theorem

- Chinese army! How to count that many soldiers?
- Group them
  - groups of size 11,13 and 7 and count the remaining soldiers
  - There is a unique solution in 11x13x7
    - X mod 11 = 5
    - X mod 13 = 3
    - X mod 7 = 2
    - X= 5x(13x7x4)+3x(11x7x12)+2x(11x13x5)= 6022 mod 1001= 16

Theorem 1. *Suppose $p_1, \ldots, p_r$ are pairwise relatively prime positive integers and let $a_1, \ldots, a_r$ be integers. Then, the system of r congruences $x \equiv a_i \bmod p_i$ for $1 \leq i \leq r$ has a unique solution modulo $P = p_1 \times \ldots \times p_r$, which is given by*

$$x = \sum_{i=1}^{r} a_i P_i y_i \bmod P \,, \tag{1}$$

where $P_i = P/p_i$ and $y_i = P_i^{-1} \bmod p_i$ for $1 \leq i \leq r$.

TU Delft

Protocol

- **Set-up** UP:
  - Generates a key pair and publishes the public key
  - Generates a prime number for each group and broadcasts
  - Households:
    - Find a partner and create a secret key
- **Protocol**
  - Household:
    - Prepares the input using CRT, encrypts and sends it to A
  - A: aggregates data and sends it to UP
  - UP:
    - Aggregate received encrypted messages
    - Using the primes, compute the group consumptions

$$E_{pk}(m_1', r) = g^{m_1'} \cdot r^n \cdot h^{n-\alpha} \bmod n^2$$

$$E_{pk}(m_2', r) = g^{m_1'} \cdot r^n \cdot h^{n+\alpha} \bmod n^2$$

$$E_{pk}(m, r) = g^m \cdot r^n \bmod n^2$$

TU Delft

Obtaining consumptions per group

$$E_{pk}(T) = E_{pk}(\sum_i m'_i)$$

$$T_{G_1} := \sum_i m'_i \bmod p_k$$

- Correct parameter use is checked.
  - Different ideas: UP or A can perform some additional computations
- Individual measurements cannot be obtained
  - Sub-groups of size 2 (can be generalized)

Dynamic Environment?

- Adding a new household is straightforward
  - Couple 2 new ones
  - One real, one dummy realized by another households
- Removing (malfunction)
  - Single household
    - The other partners data should be dropped too
  - A group is missing
    - UP performs a correction computation

TU Delft

Efficiency

		Multiplication	**Exponentiation**	**Decryption**	**Hashing**	**Communication**
Setup	$\mathcal{UP}$		K			K
Reporting	$\mathcal{A}$				N	
	H_i	3	2		1	1
Aggregation	$\mathcal{UP}$				1	
	$\mathcal{A}$	N				1
Computation	$\mathcal{UP}$			1		

	Garcia&Jacobs [9]		Kursawe *et. al* [5]		Erkin&Tsudik [16]		Ács&Castelluccia [8]		Proposed		
Operations	**SM**	**A**	**SM**	**A**	**SM**	**A**	**SM**	**A**	**SM**	**A**	**UP**
	Paillier (2048 bits)		DH Group (256 bits)		Paillier (2048 bits)		HE (32 bits)		Paillier (2048 bits)		
Encryption	$\mathcal{O}(N)$	-	-	-	$\mathcal{O}(1)$	-	-	-	$\mathcal{O}(1)$	-	-
Decryption	$\mathcal{O}(1)$	-	-	-	-	$\mathcal{O}(1)$	-	-	-	-	$\mathcal{O}(1)$
Multiplication	-	$\mathcal{O}(N^2)$	-	$\mathcal{O}(N)$	-	$\mathcal{O}(N)$	-	$\mathcal{O}(1)$	-	$\mathcal{O}(N)$	-
Exponentiation	-	-	$\mathcal{O}(1)$	-	-	-	-	-	$\mathcal{O}(1)$	-	$\mathcal{O}(K)$
Addition	-	-	-	-	-	-	$\mathcal{O}(1)$	$\mathcal{O}(N)$	-	-	-
Subtraction	-	-	-	-	-	-	-	$\mathcal{O}(1)$	-	-	-
Communication	$\mathcal{O}(N)$	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$	$\mathcal{O}(N^2)$	$\mathcal{O}(1)$	$\mathcal{O}(N)$	$\mathcal{O}(1)$	$\mathcal{O}(N)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$	$\mathcal{O}(K)$

TU Delft

Conclusion

- A data aggregation protocol
  - Groups and sub-groups
  - Dynamic (addition and removing)
  - Efficient (single encryption)
- There is no need to have an external (3rd) party
- Statistical computations can be achieved
- Future work:
  - Implementation on a test-bed
  - More complex functions can be build upon

TU Delft

GJ10
Secret Sharing and HE

$$C_{\text{total}}(t) = m_{1,t} + m_{2,t} + m_{3,t},$$

$$\begin{aligned}
\text{Alice: } m_{1,t} &= m_{1,t}(1) + m_{1,t}(2) + m_{1,t}(3) \bmod \eta, \\
\text{Bob: } m_{2,t} &= m_{2,t}(1) + m_{2,t}(2) + m_{2,t}(3) \bmod \eta, \\
\text{Charles: } m_{3,t} &= m_{3,t}(1) + m_{3,t}(2) + m_{3,t}(3) \bmod \eta,
\end{aligned}$$

Alice Bob Charles

$$\mathcal{E}_{pk_i}(m'_{i,t}) = \prod_{j \neq i} \mathcal{E}_{pk_i}(m_{j,t}(i)) = \mathcal{E}_{pk_i}\left(\sum_{j \neq i} m_{j,t}(i)\right),$$

- Keep one share for yourself, encrypt other two
- UP adds them up (HE), sends you back
- You decrypt and add your share and send it to UP
- UP adds up all data

F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in Proc. 6th Workshop Security and Trust Management (STM 2010), (LNCS), vol. 6710, pp. 226–238.

TU Delft

KDK11

Masking and Brute Forcing

$$\text{Alice} : g^{m_{1,t}+r_1}$$
$$\text{Bob} : g^{m_{2,t}+r_2}$$
$$\text{Charles} : g^{m_{3,t}+r_3}$$

$$r_1 + r_2 + r_2 = 0$$

$$\prod_{j=1}^{3} g_i^{m_{j,t}+r_j} = g_i^{\sum_{j=1}^{3} m_{j,t}+r_j} \bmod p$$

- Assumption: UP roughly knows the total; checks for equality
- 4 protocols to derive random values
  - Secret sharing
  - 3x DH and bilinear maps

TU Delft

K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in Privacy Enhanced Technologies Symposium, Waterloo, Canada, 2011, pp. 175–191.

ET12
Modified HE

$$\text{Alice: } \mathcal{F}_{pk}(m_{1,t}) = g^{m_{1,t}} \cdot r^{n_1} \bmod n^2,$$
$$\text{Bob: } \mathcal{F}_{pk}(m_{2,t}) = g^{m_{2,t}} \cdot r^{n_2} \bmod n^2,$$
$$\text{Charles: } \mathcal{F}_{pk}(m_{3,t}) = g^{m_{3,t}} \cdot r^{n_3} \bmod n^2,$$

$$\prod_i \mathcal{F}_{pk}(m_i) = g^{\Sigma_i m_{i,t}} \cdot r^{\Sigma_i n_i} \bmod n^2$$
$$= g^{\sum_i m_{i,t}} \cdot r^n \bmod n^2 := \mathcal{E}_{pk}\left(\sum_i m_{i,t}\right).$$

- Spatio-temporal consumption
- Time-stamps
- Efficient: Paillier, Hash, PRF
- Cannot deal with missing data (external party needed)

Z. Erkin and G. Tsudik, "Private computation of spatial and temporal power consumption with smart meters," in Proc. Int. Conf. Applied Cryptography and Network Security, Singapore, 26–29 June 2012, pp. 561–577.

TU Delft