Privacy Risk Analysis of Large-scale Temporal Data

Application to Electricity Consumption Data

Author: Antonin VOYEZ

Supervisors: Élisa FROMONT
Gildas AVOINE
Tristan ALLARD
Pierre CAUCHOIS

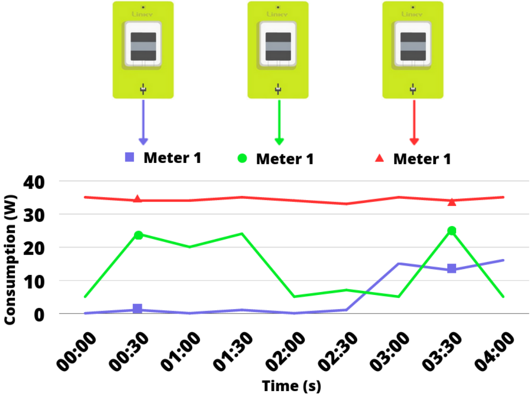
Smart metering

French smart **meter**: Linky

- Principal French electricity distribution operator.
- 33M Linky deployed.

Measurement:

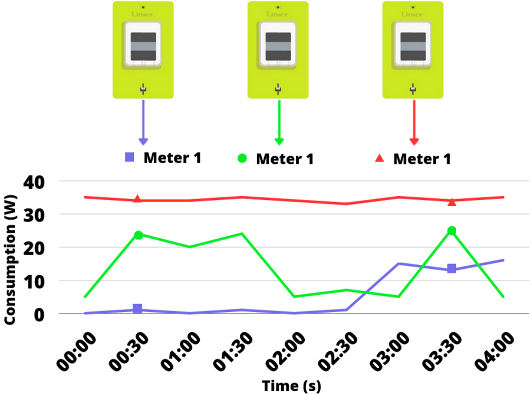
- 1 measurement every **30 minutes**.
- Electricity consumed during the measurement period (Watt).



Time series

Time series:

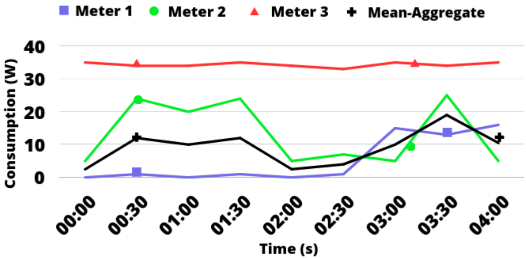
- Sequence of timestamped data.
- Ordered by **time**.
- Time-series **length**: number of timestamps.
- Time-series: $[1; 0; \cdots ; 14; 15]$.



Open data[1]

Publication[2]:

- **Sum** / **mean** multiple measurements per timestamp.
- **Aggregate size:** number of series in the aggregate.
- **Threshold:** aggregate size $\geq$ 5000.
- **Additional information** alongside the aggregate (contract type, ...).



[1] https://data.enedis.fr

[2] Code de l'Énergie, Loi pour une République Numérique

Data usage

- Energy transition:
  - Figure: Energy consumption in Rennes[1].
- Network management (prevision, dimensioning).
- Crossing with other data:
  - Energy data (gaz).
  - Socio-economics data (insee, data.gouv.fr).

[1] https://observatoire.enedis.fr/

Privacy issues

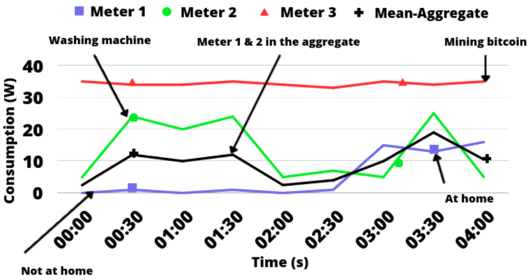
Electricity consumption time series are **personal** data (GDPR).

Inferring properties[1]:

- Devices used?
- Home presence?

Inferring membership (MIA)[2]:

- Is a series participating in the aggregate?



[1] Pascal A. Schirmer and Iosif Mporas. "Non-Intrusive Load Monitoring: A Review". In: *IEEE Transactions on Smart Grid* (2023).

[2] Hongsheng Hu et al. "Membership Inference Attacks on Machine Learning: A Survey". In: *ACM Computing Surveys* (2022).

Objectives of the thesis

- Understand the risks of publishing series and aggregates.
  - **Is the current threshold (5000 series) safe?**
  - What makes a series vulnerable?
- Propose attacks on existing open data aggregates.
  - Experimental approach.
  - Leveraging large-scale real-life electricity consumption data.

Presentation outline

1. Datasets analysis.
   - Exploring the datasets.
   - Uniqueness study: almost everyone is unique.
2. The SubSum attack.
   - Able to infer the appartenance of each member of the aggregate.
3. The STATS attack.
   - Find if a specific series participates in the aggregate.

Datasets analysis

- Presents the datasets.
- Uniqueness study.
- Is it safe to publish pseudonymized series?
    - Pseudonymized: removing identifying information (name).

Datasets

ENEDIS:

- Enedis's real French data.
- Sampling: 30 minutes.
- Duration: 2 years.
- Size:
  - 3M 30 minutes series.
  - **2M residential series.**
- Profiles: Type of series (contract, consumption pattern).

ISSDA[1]:

- Public Irish electricity consumption datasets.
- Sampling: 30 minutes.
- Duration: 1.5 years (2009 - 2010).
- Size: Approx. 4500 series.

[1] CER. *CER Smart Metering Project - Electricity Customer Behaviour Trial. Irish Social Science Data Archive. (Accessed May 11 2022).* https://www.ucd.ie/issda/data/commissionforenergyregulationcer. 2012.

Datasets statistics

Measured values:

- Metering range: [0; 36000] W.
- Actual consumption: **mostly below 1000 W.**
- Peak at 0 W.

Seasonal patterns:

- More electricity consumption in the winter.
- More electricity consumption in the evening.
- The higher the consumption, the higher the dispersion.

Uniqueness study

- Unique individuals are potentially identifiable.
  - **Unique:** an individual is unique if it is the only one possessing a set of values.
  - Sweeney's governor Welds re-identification[2].
  - Narayanan Netflix attack[3].
- The proportion of unique individuals is used as a risk metric.
  - **Uniqueness:** proportion of unique individuals.
  - De Montjoye studies showed high uniqueness on large datasets with little adversarial knowledge[4].

[2] Latanya Sweeney. "Simple demographics often identify people uniquely". In: *Health (San Francisco)* (2000).

[3] Arvind Narayanan and Vitaly Shmatikov. "Robust De-anonymization of Large Sparse Datasets". In: *IEEE Symposium on Security and Privacy (SP)*. 2008.

[4] Yves-Alexandre De Montjoye et al. "Unique in the crowd: The privacy bounds of human mobility". In: *Scientific reports* (2013).

Uniqueness study: methodology

- Uniqueness computed per time window:
  - t: starting timestamp.
  - **k: number of consecutive timestamps.**
- Dataset uniqueness: averaging the uniqueness per timestamp.

	Uniqueness = 60% t = 1, k = 2		
	Timestamp 1	**Timestamp 2**	**Timestamp 3**
Meter 1	0	1	0
Meter 2	0	2	0
Meter 3	2	1	0
Meter 4	2	2	1
Meter 5	0	1	2

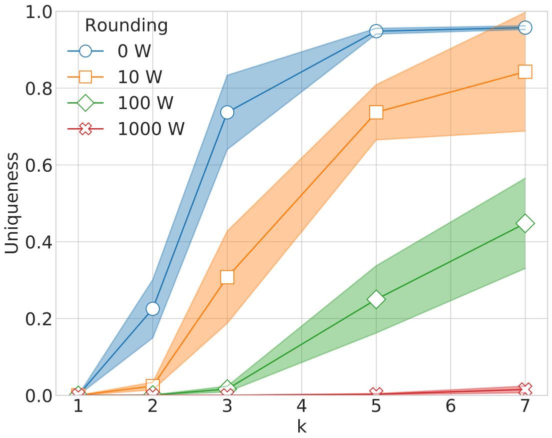
Uniqueness study: methodology

- Uniqueness computed per time window:
  - t: starting timestamp.
  - **k: number of consecutive timestamps.**
- Dataset uniqueness: averaging the uniqueness per timestamp.

		Uniqueness = 60% t = 2, k = 2	
	Timestamp 1	**Timestamp 2**	**Timestamp 3**
Meter 1	0	1	0
Meter 2	0	2	0
Meter 3	2	1	0
Meter 4	2	2	1
Meter 5	0	1	2

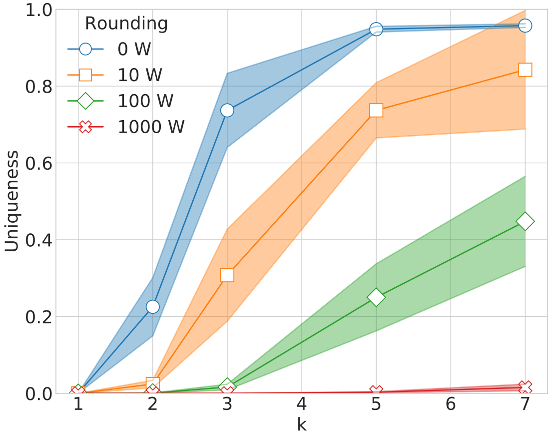
Uniqueness results

- Figure: average uniqueness (95% confidence interval) according to the number of consecutive points (k) and the rounding.
- **High uniqueness** considering only a few timestamps:
  - $>$ 70% for k = 3 (1h30)
  - **$>$ 90% for k = 5 (2h30)**



Uniqueness results: rounding

- **Reducing the measurement's precision (by rounding them) is not enough to protect the series.**
  - Approx. 12k series are unique for $k = 7$ and with rounding to 1 kW.
  - Rounding to 1 kW renders the data useless: 80 % of measurements below 1kW.



Uniqueness study: conclusion

- Enedis dataset: 3M half-hourly series.
- Is it safe to publish pseudonymized series?
  - High uniqueness with minimal adversarial knowledge.
  - Potentially vulnerable to uniqueness-based reidentification attacks.
  - It is **unsafe to publish pseudonymized** electricity consumption time series.
- **Publication (under review):** Nature Scientific Report: Smart Cities.

The SubSum attack

- Attacking aggregates.
- Is it possible to find who is in an aggregate?

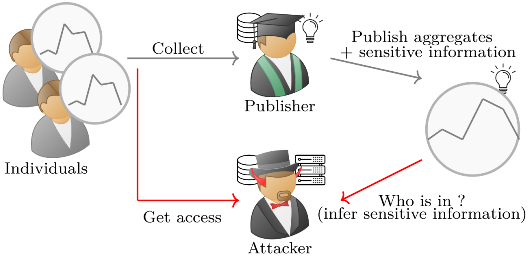
The SubSum attack: problem statement

Objective:

- Find **all** the series participating in the aggregate.

Attacker knowledge:

- Open data aggregate.
- Population larger than the series participating to aggregate.
- In real life: available to a major provider, disclosed by a data breach.



The SubSum attack

Constraints:

- $\forall t \in \mathcal{T}, A_t = \sum_{\forall i \in S} S_{i,t} \cdot X_i$
- A_t: Aggregate value at the timestamp t. $S_{i,t}$: Consumption of the individual i at the timestamp t. X_i: Boolean telling whether or not the individual i is in the aggregate.

	Meter 1	Meter 2	Meter 3	Meter 4	Aggregate (size = 2)
Timestamp 1	10	0	15	10	25
Timestamp 2	5	5	7	10	12

The SubSum attack

Constraints:

- $\forall t \in \mathcal{T}, A_t = \sum_{\forall i \in S} S_{i,t} \cdot X_i$
- A_t: Aggregate value at the timestamp t. $S_{i,t}$: Consumption of the individual i at the timestamp t. X_i: Boolean telling whether or not the individual i is in the aggregate.

	Meter 1	Meter 2	Meter 3	Meter 4	Aggregate (size = 2)
Timestamp 1	10	0	15	10	25
Timestamp 2	5	5	7	10	12

The SubSum attack

Constraints:

- $\forall t \in \mathcal{T}, A_t = \sum_{\forall i \in S} S_{i,t} \cdot X_i$
- A_t: Aggregate value at the timestamp t. $S_{i,t}$: Consumption of the individual i at the timestamp t. X_i: Boolean telling whether or not the individual i is in the aggregate.

	Meter 1	Meter 2	Meter 3	Meter 4	Aggregate (size = 2)
Timestamp 1	10	0	15	10	25
Timestamp 2	5	5	7	10	12

The SubSum attack

Constraints:

- $\forall t \in \mathcal{T}, A_t = \sum_{\forall i \in S} S_{i,t} \cdot X_i$
- A_t: Aggregate value at the timestamp t. $S_{i,t}$: Consumption of the individual i at the timestamp t. X_i: Boolean telling whether or not the individual i is in the aggregate.

	Meter 1	Meter 2	Meter 3	Meter 4	Aggregate (size = 2)
Timestamp 1	10	0	15	10	25
Timestamp 2	5	5	7	10	12

The SubSum attack

Constraints:

- $\forall t \in \mathcal{T}, A_t = \sum_{\forall i \in S} S_{i,t} \cdot X_i$
- A_t: Aggregate value at the timestamp t. $S_{i,t}$: Consumption of the individual i at the timestamp t. X_i: Boolean telling whether or not the individual i is in the aggregate.

	Meter 1	Meter 2	Meter 3	Meter 4	Aggregate (size = 2)
Timestamp 1	10	0	15	10	25
Timestamp 2	5	5	7	10	12

Background on reconstruction attacks

Reconstruction attacks with linear reconstruction[5]

- Applying filters (queries) to an aggregated dataset.
- Build a set of constraints (equations) from the filters.
- Solving the constraints recreates the original dataset.

In practice:

- DIFFIX[6], US census[7]

[5] Irit Dinur and Kobbi Nissim. "Revealing Information While Preserving Privacy". In: *ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. 2003.

[6] Aloni Cohen and Kobbi Nissim. "Linear Program Reconstruction in Practice". In: *Journal of Privacy and Confidentiality* (2020).

[7] Simson Garfinkel, John M. Abowd, and Christian Martindale. "Understanding Database Reconstruction Attacks on Public Data". In: *Communications of the ACM* (2019).

The SubSum attack: experiments

Goal:

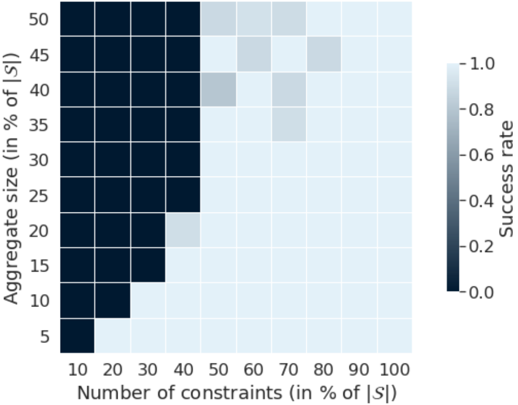
- What are the **conditions** required to make our attack work?
  - Number of series, aggregate size, series length.
- **How long** does it take?

Experimental setup:

- Success: Find **all** the existing solutions in the impaired time.
  - time budget: θ, maximum amount of solutions p
- Data: ISSDA.
- Solver: Gurobi.

Success rate

- Figure: success rate depending on the number of constraints and the aggregate size.
  - Figure parameters: $|\mathcal{S}| = 4500$, $\theta = 24\text{h}$, $p = 100$, 20 repetitions.
- The number of constraints required for a successful attack is of the same order as the aggregate size attacked.



Experimentation time depending on the time budget

- When the number of constraints is too low:
  - Attack fails due to the wall time.
- When the number of constraints is too high.
  - The attack should be a success but the time increases linearly with the number of constraints.

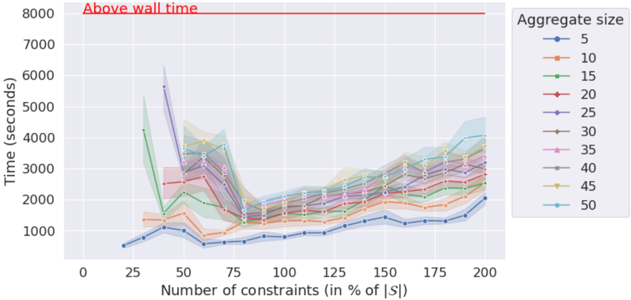


Figure: Experiment's time (s). Parameters: $|\mathcal{S}| = 2000$, $\theta = 8000$s (approx. 3h), $p = 2$, 20 repetitions. Computer: 2 cores and 8 Go RAM.

The SubSum attack: conclusion

- Based on solving the subset-sum problem.
- Able to **find all** the series participating in an aggregate.
- Heavy requirements:
  - Large number of series, and timestamps.
  - **Scaling issues:** time consuming.
- **Publication:** International Conference on Security and Cryptography (SECRYPT) 2022.

The STATS attack

- Reduce the background knowledge.
- Is it possible to find a single individual within an aggregate?
- STATS: **S**hadow **T**raining for **A**ggregated **T**ime **S**eries.

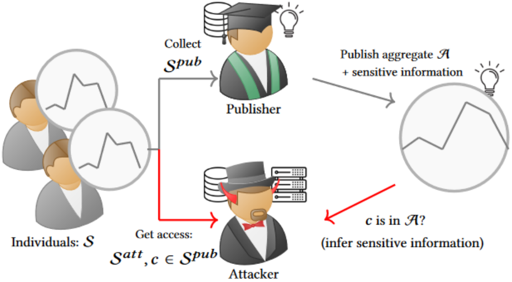
The STATS attack

Objective:

- Find **one** series participating in the aggregate.

Attacker knowledge:

- Open data aggregate.
- The targeted series (c).
- A set of series with **similar statistical properties** to the ones in the aggregates.
- In real life: public data (ISSDA), supplier data, data leaks.



Background: Knock Knock who is there?[1]

Pyrgelis's attack:

- Attacking location aggregate (number of individuals per place and time).
- Method: Shadow training[2].
- Results: small aggregates (≤ 100) are vulnerable.
  - Using simple classifiers (linear regression) and features (PCA).

Our contribution:

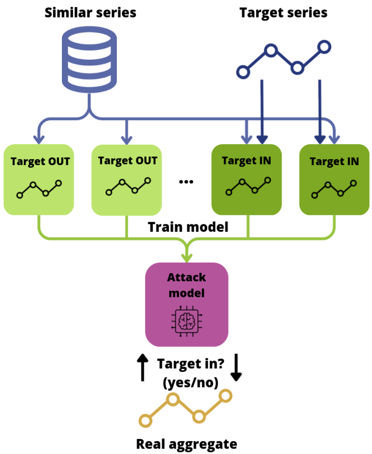
- Increase the attacked aggregate size.
- Adapt the Pyrgelis's attack to efficiently cope with time series.

[1] Apostolos Pyrgelis, Carmela Troncoso, and Emiliano De Cristofaro. "Knock Knock, Who's There? Membership Inference on Aggregate Location Data". In: *Network and Distributed System Security Symposium, NDSS*. 2018.

[2] Reza Shokri et al. "Membership Inference Attacks Against Machine Learning Models". In: *IEEE Symposium on Security and Privacy (SP)*. 2017.

Shadow training algorithm

1. Build a set of fake aggregates (with and without the target).
2. Train a classifier to detect the target in the aggregates.
3. Test the classifier on test aggregates and evaluate using accuracy.
4. Use the attack model against the real aggregate.



Experiments

Goal:

- Find the threshold (aggregate size) such as aggregates are no longer vulnerable.
- What makes a series vulnerable?
- **Vulnerable**: accuracy > 0.6.

Experimental setup:

- Experiments: exploring the parameters space (aggregate size, series length, group, and profile).
- Data: Enedis (June 2021 and June 2022).
- Classifier: MiniRocket[1].

[1] Angus Dempster, Daniel F Schmidt, and Geoffrey I Webb. "Minirocket: A very fast (almost) deterministic transform for time series classification". In: *ACM SIGKDD*. 2021.

Target choice

- **Score:** standard deviation.
  - Captures the impact of individual series on the aggregate (and on the classification).
- **Groups (G):** splitting the score distribution.
  - Forcing the sampling of outliers.
  - Sampling 50 targets per group and profile.

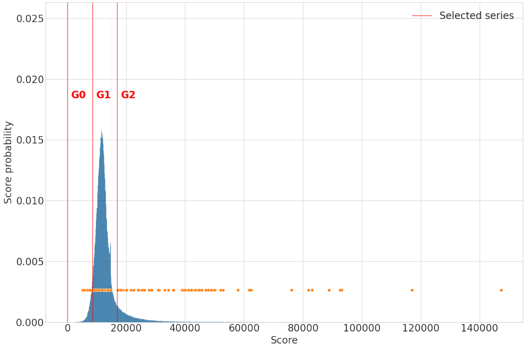
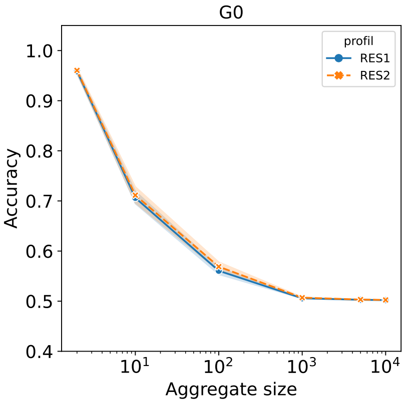


Figure: Scores distribution ENEDIS RES1.

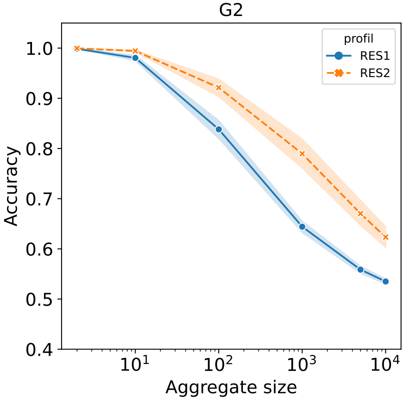
Attack results: G0, train / test in June 2021

- Figure: **attack accuracy** per group, profile and aggregate size.
  - Training and testing in June 2021.
  - RES1: basic pricing, RES2: dynamic pricing.
- **Larger aggregates lead to lower accuracy.**



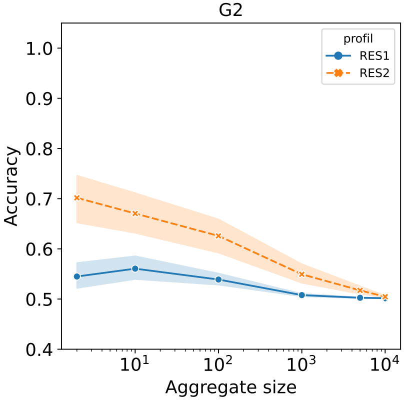
Attack results: G2, train / test in June 2021

- Figure: average **attack accuracy** per group, profile and aggregate size.
  - Training and testing in June 2021.
  - RES1: basic pricing, RES2: dynamic pricing.
- **Atypical series (group G2) are vulnerable.**
  - Aggregated of 5000 series: accuracy > 0.65.
  - Approx. 60k (2%) series are vulnerable.



Attack results: G2, training on historical data

- Figure: **attacker accuracy** per group, profile, and aggregate size.
  - Training in June 2021, testing in June 2022.
- **Lower accuracy** due to the historical data.
  - The **series changes** over time (due to temperature and human behavior).
  - Small aggregates remain vulnerable.
  - Still approx. 10k series (0.05%) are vulnerable against 5000 aggregated series.



The STATS attack: conclusion

- The STATS attack is a time series classification problem.
- Minimal requirements: target series, similar series.
- Is the legal threshold safe?
  - **The legal threshold is vulnerable.**
  - At least against the most atypical series.
- **Publication:** Communication of the ACM. (CACM, under review).

Conclusion: contributions

1. Data analysis and uniqueness study.
   - High uniqueness rate for individual series.
   - The publication of pseudonymized time series is risky.
   - **Publication:** Nature Scientific Report (under review).
2. The SubSum attack.
   - Identify all members of the aggregate when requirements are met.
   - Requires to know, at least, all the aggregate members.
   - **Publication:** BDA 2021, SECRYPT 2022.
3. The STATS attack.
   - Time series classification problem.
   - Identify the presence of outliers in large aggregates (above the legal threshold).
   - **Publication:** Communication of the ACM (under review).

Conclusion

1. Is the current threshold (5000 series) safe?
   - No: the STATS attack.
   - Only the most atypical series are at risk.
   - For most individuals the threshold may be reduced.
2. What makes a series vulnerable?
   - Uniqueness: high uniqueness rate for individual series.
   - Atypical series relative to the population.

Future work: extension

Current contributions:

- Attack unprotected aggregates.
- Access to full series.

Background knowledge:

- Missing values.
- Approximate values.

Protection methods:

- Differential privacy[1].
- Synthetic series (GAN)[2].

[1] Vibhor Rastogi and Suman Nath. "Differentially Private Aggregation of Distributed Time-Series with Transformation and Encryption". In: *ACM SIGMOD International Conference on Management of Data*. 2010.

[2] Theresa Stadler, Bristena Oprisanu, and Carmela Troncoso. "Synthetic Data – Anonymisation Groundhog Day". In: *USENIX Security Symposium (USENIX Security 22)*. 2022.

Future work: attributes inference

Current contributions

- Membership inference attack.

Properties inference:

- Non-Intrusive Load Monitoring (NILM).
- Extract events (e.g., devices used) from electricity consumption time series.
- Few existing works on coarse (half-hourly) time series[1].
  - Home presence? Unemployment? Large devices (EV)?
- Can we extract individual or group properties from series? From aggregates?

[1] Pascal A. Schirmer and Iosif Mporas. "Non-Intrusive Load Monitoring: A Review". In: *IEEE Transactions on Smart Grid* (2023).

Thank you

1. Data analysis and uniqueness study.
   - High uniqueness rate for individual series.
   - The publication of pseudonymized time series is risky.
   - **Publication:** Nature Scientific Report. 2022 (under review).
2. The SubSum attack.
   - Identify all members of the aggregate when requirements are met.
   - Requires to know, at least, all the aggregate members.
   - **Publication:** SECRYPT 2022.
3. The STATS attack.
   - Time series classification problem.
   - Identify the presence of outliers in large aggregates (above the legal threshold).
   - **Publication:** Communication of the ACM. (under review).