Detection of Irregular Power Usage using Machine Learning

May 22, 2018

Patrick Glauner, **Jorge Augusto Meira** and **Radu State**
Interdisciplinary Centre for Security, Reliability and Trust,
University of Luxembourg

Motivation

What is a typical non-technical loss (NTL)?

Motivation

What is a typical non-technical loss (NTL)?

Motivation

What is a typical non-technical loss (NTL)?

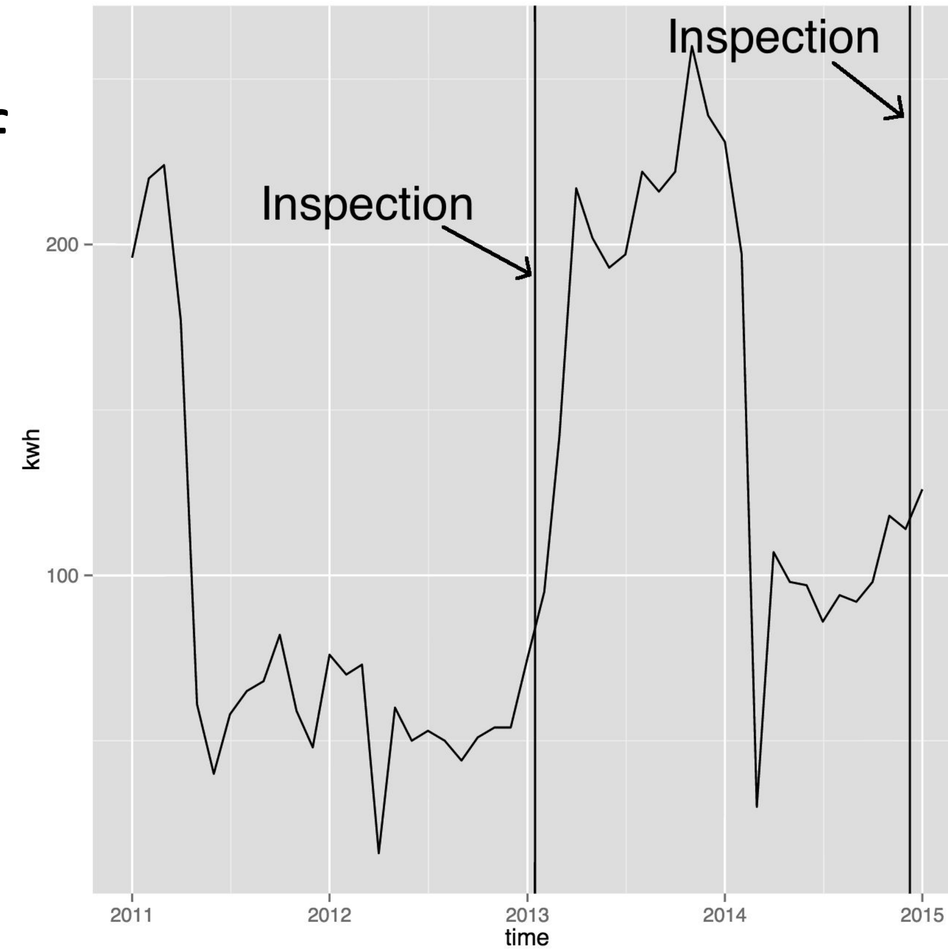
Motivation

Worldwide electric utilities lose $96 billion USD (*) annually to fraud/theft

(*) Electricity Theft and Non-Technical Losses: Global Market, Solutions and Vendors
May 2017 | Northeast group, llc
www.northeast-group.com
(*) World Bank data

Motivation

Example of NTL: Two assumed occurrences of NTL due to significant consumption drops followed by inspections (visualized by vertical bars).



Project overview

- Joint university-industry research project on detection of NTL
- Started in late 2015
- Goal: applied R&D focused on both publishing papers and deploying features in partner company's products

securityandtrust.lu

UNIVERSITÉ DU LUXEMBOURG

Project overview

Published results in: IEEE Innovative Smart Grid Technologies (**ISGT**), Power and Energy Conference at Illinois (**PECI**), International Conference on Intelligent System Applications to Power Systems (**ISAP**), IEEE International Conference on Data Mining (**ICDM**), etc.

securityandtrust.lu

UNIVERSITÉ DU LUXEMBOURG

Power & Energy Society

Project overview

AI could put a stop to electricity theft and meter misreadings

https://www.newscientist.com/article/2148308-ai-could-put-a-stop-to-electricity-theft-and-meter-misreadings/

Project overview

Cited in:

IEEE PES Power & Energy Society®

IEEE

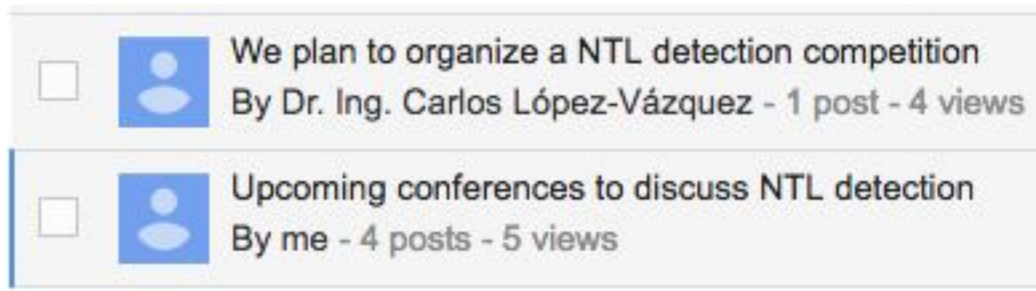
Goals of this tutorial

- Introducing the problem of NTL
- Reviewing the state of the art of NTL detection using machine learning methods
- Discussing the challenges of NTL detection
- Presenting a selection of our works
- Providing a forum for discussions

Interested in NTL?

Join our mailing list:

https://groups.google.com/d/forum/ntl-community



IEEE PES Power & Energy Society®

IEEE

Contents

- Introduction to NTL
- State of the art
- Challenges
- Selection of our works
- Conclusions

Introduction to NTL

- Losses in power systems:
  - Technical
  - Non-technical

Introduction to NTL

- Meter tampering in order to record lower consumptions
- Bypassing meters by rigging lines from the power source
- Arranged false meter readings by bribing meter readers

Introduction to NTL

- Faulty or broken meters
- Un-metered supply
- Technical and human errors in meter readings, data processing and billing

State of the art

The Challenge of Non-Technical Loss Detection Using Artificial Intelligence: A Survey

Patrick Glauner [1], Jorge Augusto Meira [1], Petko Valtchev [12], Radu State [1], Franck Bettinger [3]

Published in:

International Journal of Computational Intelligence Systems (IJCIS), vol. 10, issue 1, pp. 760-775, 2017.

State of the art

- Features
- Models
- Comparison

State of the art: features

- Monthly consumption:
  - Daily averages:

$$x_d^{(m)} = \frac{L_d^{(m)}}{R_d^{(m)} - R_{d-1}^{(m)}},$$

  - Monthly consumption before the inspection
  - Consumption in the same month in the year before
  - Consumption in the past three months

State of the art: features

- Monthly consumption:
  - The customer’s consumption over the past 24 months
  - Average consumption
  - Maximum consumption
  - Standard deviation
  - Number of inspections
  - Average consumption of the residential area

State of the art: features

- Smart meter consumption:
  - Consumption features from intervals of 15 or 30 minutes
  - The maximum consumption in any 15-minute window
  - Load factor is computed by dividing the demand contracted by the maximum consumption
  - Shape factors are derived from the consumption time series including the impact of lunch times, nights and weekends

State of the art: features

- Smart meter consumption:
  - $4 \times 24 = 96$ measurements are encoded to a 32-dimensional space:
    - Each measurement is 0 or positive
    - Next, it is then mapped to 0 or 1, respectively
    - Last, the 32 features are computed
    - A feature is the weighted sum of three subsequent values, in which the first value is multiplied by 4, the second by 2 and the third by 1

State of the art: features

- Master data:
  - Location (city and neighborhood)
  - Business class (e.g. residential or industrial)
  - Activity type (e.g. residence or drugstore)
  - Voltage
  - Number of phases (1, 2 or 3)
  - Meter type

State of the art: features

- Master data:
  - Demand contracted, i.e. the number of kW of continuous availability requested from the energy company and the total demand in kW of installed equipment of the customer
  - Information about the power transformer to which the customer is connected to
  - Town or customer in which the customer is located
  - Type of voltage (low, median or high)

State of the art: features

- Master data:
  - Electricity tariff
  - Contracted power
  - Number of phases
  - Type of customer
  - Location
  - Voltage level
  - Type of climate (rainy or hot)
  - Weather conditions

State of the art: features

- Credit worthiness ranking (CWR):
  - Computed from the electricity provider's billing system
  - Reflects if a customer delays or avoids payments of bills
  - CWR ranges from 0 to 5 where 5 represents the maximum score
  - It reflects different information about a customer such as payment performance, income and prosperity of the neighborhood in a single feature

State of the art: models

- Expert systems and fuzzy systems
- Neural networks
- Support vector machines
- Genetic algorithms
- Rough sets
- Various other methods: optimum path forest, linear regression, etc.

State of the art: comparison

- Accuracy: $$\frac{tp + tn}{tp + tn + fp + fn}$$
- Precision: $$\frac{tp}{tp + fp}$$
- Recall: $$\frac{tp}{tp + fn}$$
- F1: $$2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

State of the art: comparison

Ref.	Model	#Customers	Accuracy	Precision	Recall	AUC	NTL/theft proportion
1	SVM (Gauss)	< 400	0.86	-	0.77	-	-
7	SVM + fuzzy	100K	-	-	0.72	-	-
16	Bool rules	700K	-	-	-	0.47	5%
16	Fuzzy rules	700K	-	-	-	0.55	5%
16	SVM (linear)	700K	-	-	-	0.55	5%
16	Bool rules	700K	-	-	-	0.48	20%
16	Fuzzy rules	700K	-	-	-	0.55	20%
16	SVM (linear)	700K	-	-	-	0.55	20%
17	SVM	< 400	-	-	0.53	-	-
18	Genetic SVM	1,171	-	-	0.62	-	-
19	Neuro-fuzzy	20K	0.68	0.51	-	-	-
22	NN	22K	0.87	0.65	0.29	-	-
23	Rough sets	N/A	0.93	-	-	-	-
24	SOM	2K	0.93	0.85	0.98	-	-

State of the art: comparison

25	SVM (Gauss)	1,350	0.98	-	-	-	-
27	Regression	30	-	-	0.22	-	1%
27	Regression	30	-	-	0.78	-	2%
27	Regression	30	-	-	0.98	-	3%
27	Regression	30	-	-	1	-	4-10%
29	SVM	5K	0.96	-	-	-	-
29	KNN	5K	0.96	-	-	-	-
29	NN	5K	0.94	-	-	-	-
30	OPF	736	0.90	-	-	-	-
30	SVM (Gauss)	736	0.89	-	-	-	-
30	SVM (linear)	736	0.45	-	-	-	-
30	NN	736	0.53	-	-	-	-
33	Decision tree	N/A	0.99	-	-	-	-

Challenges

- Class imbalance and evaluation metric
- Feature description
- Data quality
- Covariate shift
- Scalability
- Comparison of different methods

Class imbalance, evaluation metric

- Imbalanced classes appear frequently in machine learning, which also affects the choice of evaluation metrics.
- Most NTL detection research do not address this property.

Class imbalance, evaluation metric

- In many papers, high accuracies or high recalls are reported:

Ref.	Model	#Customers	Accuracy	Precision	Recall
1	SVM (Gauss)	< 400	0.86	-	0.77
7	SVM + fuzzy	100K	-	-	0.72

- The following examples demonstrate why those performance measures are not suitable for NTL detection in imbalanced data sets.

Class imbalance, evaluation metric

- For a test set containing 1K customers of which 999 have regular use,
  - A classifier always predicting non-NTL has an accuracy of 99.9%
  - While this classifier has a very high accuracy and intuitively seems to perform very well, it will never predict any NTL.

Class imbalance, evaluation metric

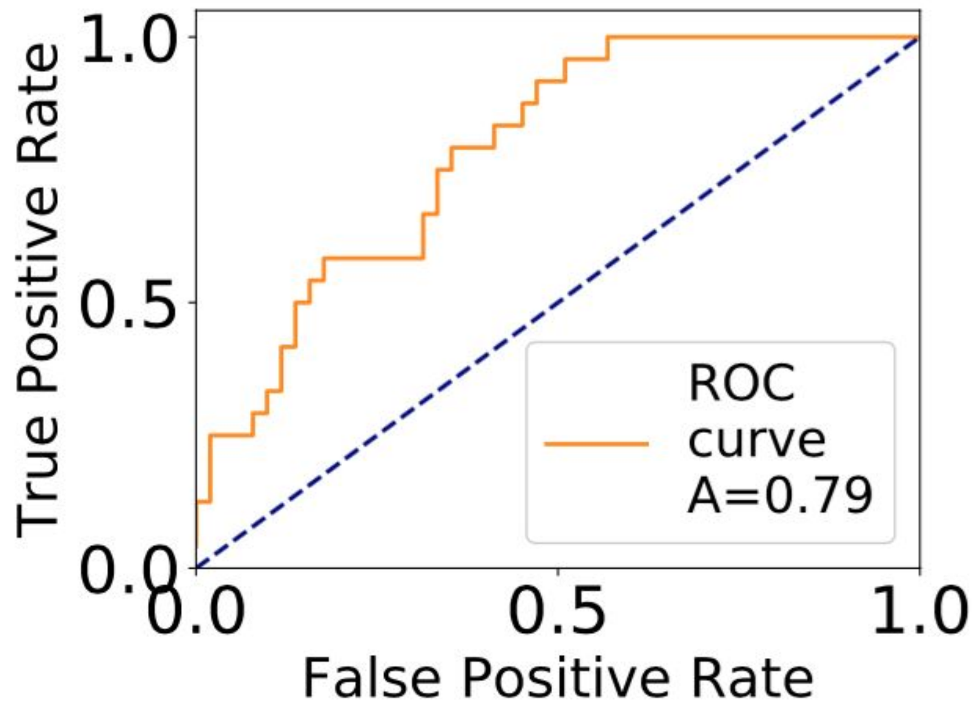
- For a test set containing 1K customers of which 999 have regular use,
  - A classifier always predicting NTL has a recall of 100%.
  - While this classifier will find all NTL, it triggers many costly and unnecessary physical inspections by inspecting all customers.

Class imbalance, evaluation metric

- This topic is addressed rarely in NTL literature.
- For NTL detection, the goal is to reduce the false positive rate (FPR) to decrease the number of costly inspections, while increasing the true positive rate (TPR) to find as many NTL occurrences as possible.

Class imbalance, evaluation metric

- We propose to use a receiver operating characteristic (ROC) curve, which plots the TPR against the FPR.



Class imbalance, evaluation metric

- The area under the curve (AUC) is a performance measure between 0 and 1, where any binary classifier with an AUC > 0.5 performs better than chance

Class imbalance, evaluation metric

Large-Scale Detection of Non-Technical Losses in Imbalanced Data Sets

Patrick Glauner*, Andre Boechat*, Lautaro Dolberg*, Radu State*, Franck Bettinger†, Yves Rangoni† and Diogo Duarte†

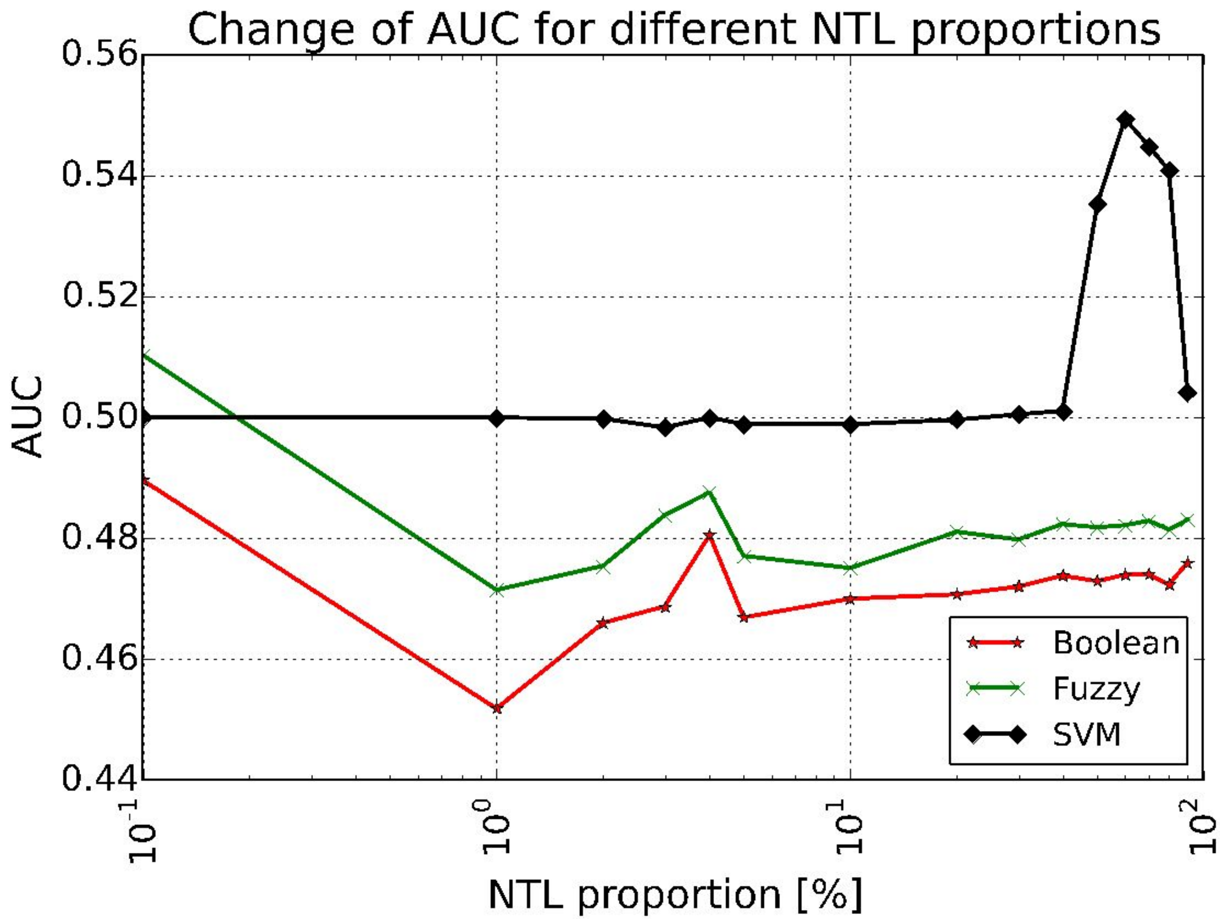
Published in:

Proceedings of the Seventh IEEE Conference on Innovative Smart Grid Technologies (ISGT 2016), Minneapolis, USA, 2016.

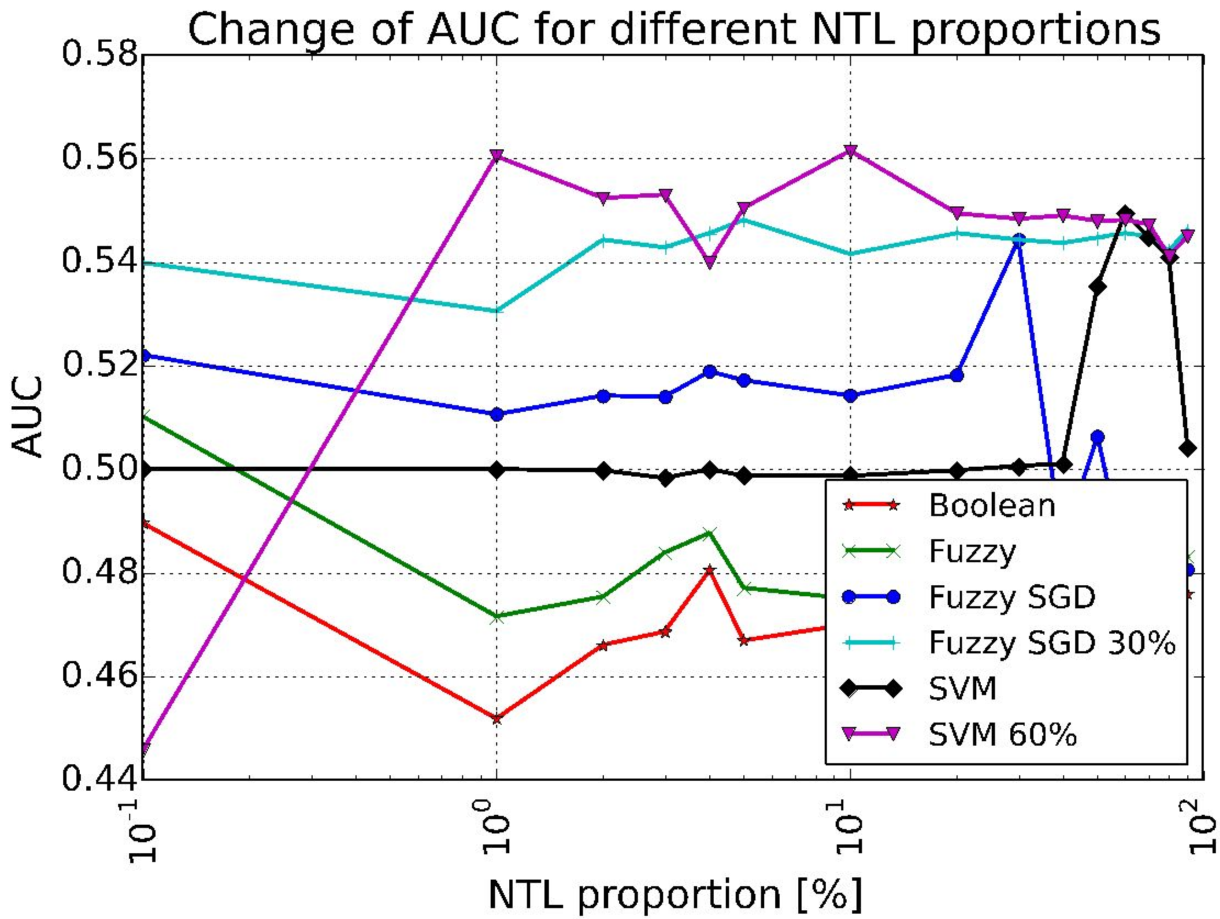
Class imbalance, evaluation metric

Ref.	Model	#Customers	Accuracy	Precision	Recall	AUC	NTL/theft proportion
1	SVM (Gauss)	< 400	0.86	-	0.77	-	-
7	SVM + fuzzy	100K	-	-	0.72	-	-
16	Bool rules	700K	-	-	-	0.47	5%
16	Fuzzy rules	700K	-	-	-	0.55	5%
16	SVM (linear)	700K	-	-	-	0.55	5%
16	Bool rules	700K	-	-	-	0.48	20%
16	Fuzzy rules	700K	-	-	-	0.55	20%
16	SVM (linear)	700K	-	-	-	0.55	20%

Class imbalance, evaluation metric



Class imbalance, evaluation metric



Feature description

- Generally, hand-crafting features from raw data is a long-standing issue in machine learning having significant impact on the performance of a classifier.
- Different feature description methods have been reviewed in the previous section.

Feature description

- They fall into two main categories:
  - Features computed from the consumption profile of customers which are from:
    - Monthly meter readings
    - Or smart meter readings
  - And features from the customer master data.

Feature description

- The features computed from the time series are very different for monthly meter readings and smart meter readings.
- The results of those works are not easily interchangeable. While electricity providers continuously upgrade their infrastructure to smart metering, there will be many remaining traditional meters. In particular, this applies to emerging countries.

Feature description

- There are only few works on assessing the statistical usefulness of features for NTL detection.
- Almost all works on NTL detection define features and subsequently report improved models that were mostly found experiment ally without having a strong theoretical foundation.

Data quality

- We noticed that the inspection result labels in the training set are not always correct and that some fraudsters may be labelled as non-fraudulent.
- The reasons for this may include bribing, blackmailing or threatening of the technician performing the inspection.
- Also, the fraud may be done too well and is therefore not observable by technicians.

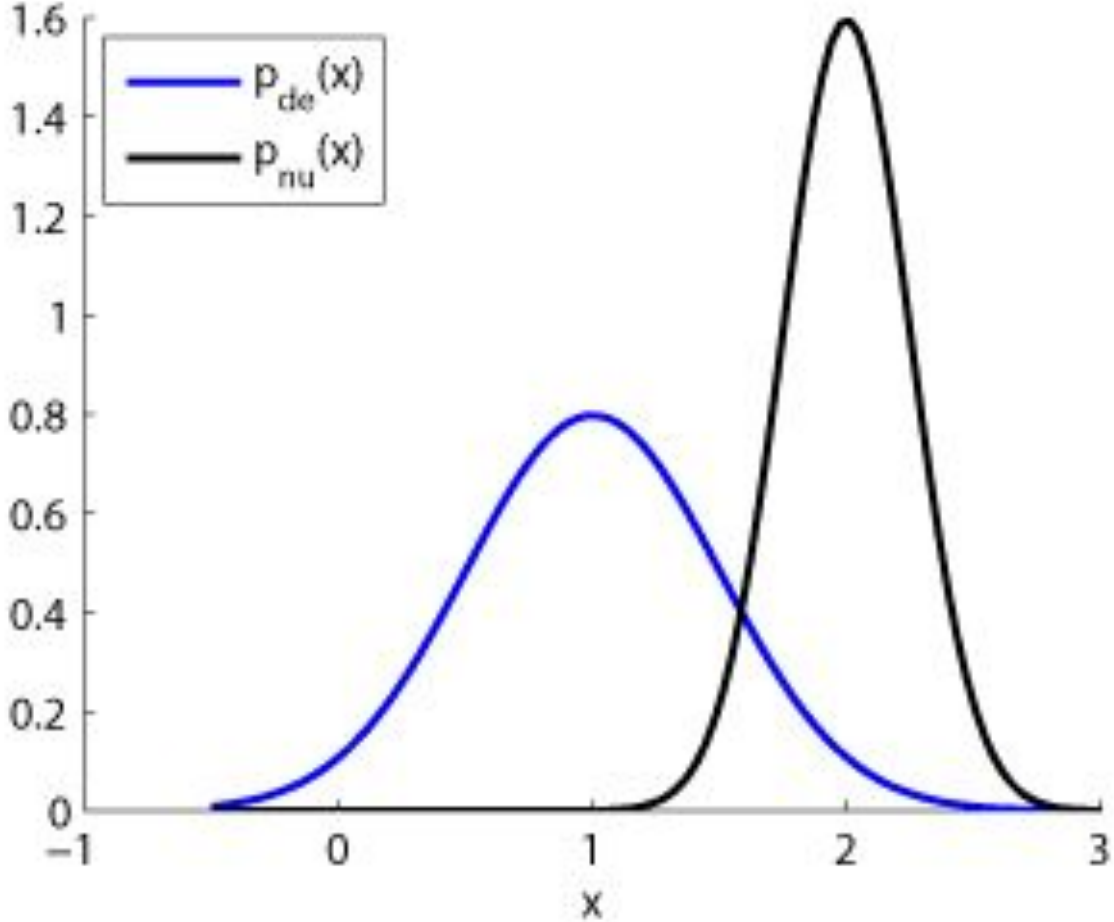
Data quality

- Another reason may be incorrect processing of the data. It must be noted that the latter reason may, however, also label non-fraudulent behavior as fraudulent.
- Most NTL detection research use supervised methods. This shortcoming of the training data and potential wrong labels in particular are only rarely reported in the literature.

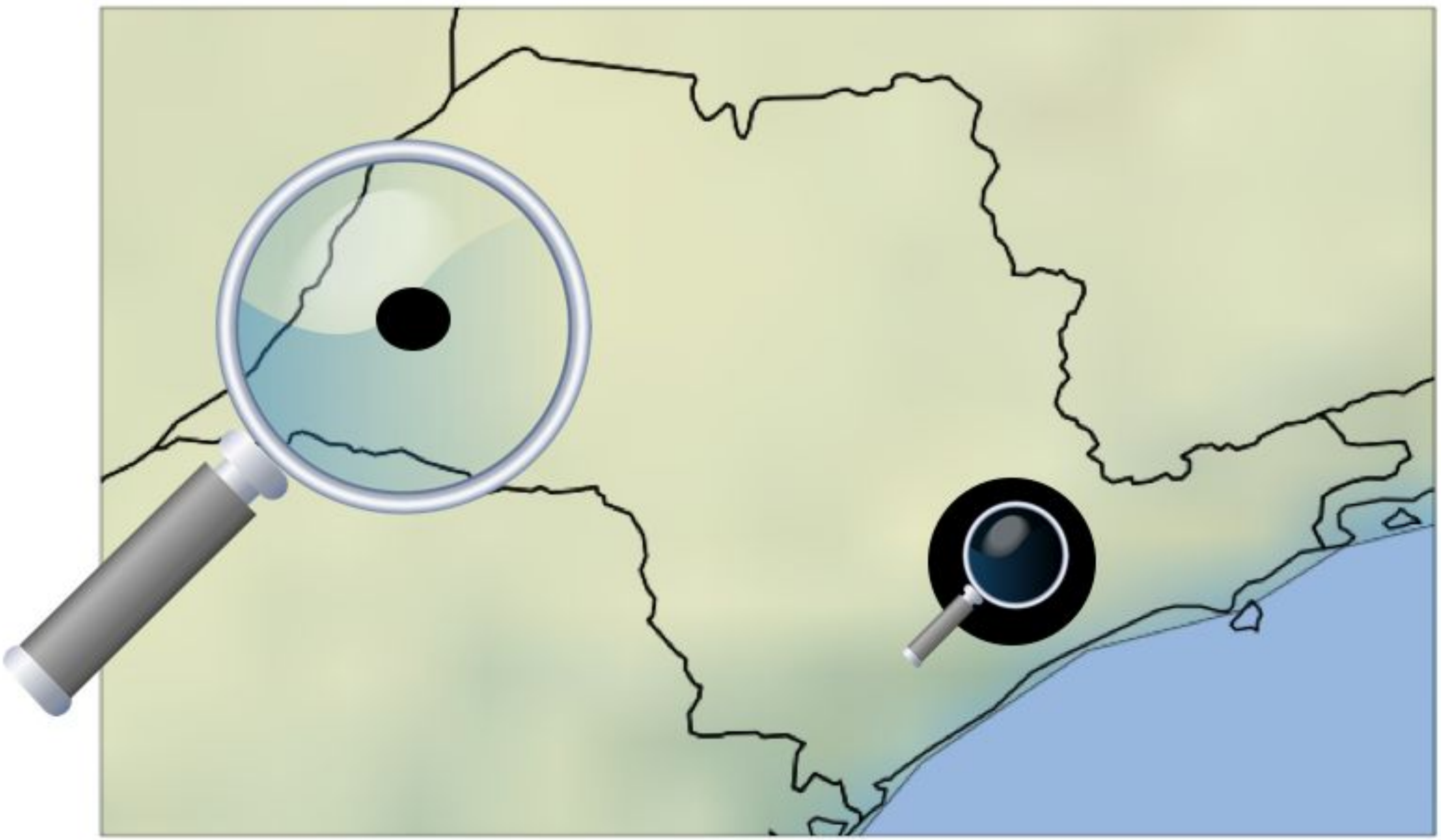
Covariate shift

- Covariate shift refers to the problem of training data (i.e. the set of inspection results) and production data (i.e. the set of customers to generate inspections for) having different distributions.

Covariate shift

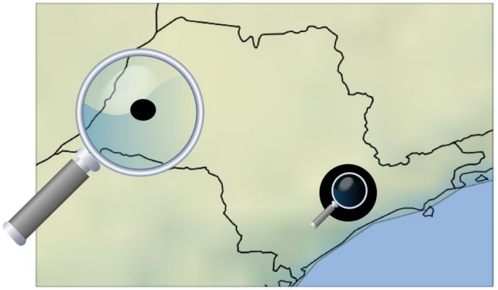


Covariate shift



Covariate shift

The large city is close to the sea, whereas the small city is located in the interior of the country. The weather in the small city undergoes stronger changes during the year. The subsequent change of electricity consumption during the year triggers many inspections. As a consequence, most inspections are carried out in the small city. Therefore, the sample of customers inspected does not represent the overall population of customers.



Covariate shift

- This fact leads to unreliable NTL predictors when learning from this training data.
- Historically, covariate shift has been a long-standing issue in statistics

Covariate shift

- The Literary Digest sent out 10M questionnaires in order to predict the outcome of the 1936 US Presidential election.
- They received 2.4M returns.
- Predicted Alfred Landon to win.

Covariate shift

- Nonetheless, the predicted result proved to be wrong.
- The reason for this was that they used car registrations and phone directories to compile a list of recipients.
- In that time, the households that had a phone or a car represented a biased sample of the overall population.

Covariate shift

- In contrast, George Gallup only interviewed 3K handpicked people, which were an unbiased sample of the population.
- As a consequence, Gallup could predict the outcome of the election very well.

Covariate shift

Is Big Data Sufficient for a Reliable Detection of Non-Technical Losses?

Patrick Glauner*, Angelo Migliosi*, Jorge Augusto Meira*, Petko Valtchev*†, Radu State* and Franck Bettinger‡

Published in:

Proceedings of the 19th International Conference on Intelligent System Applications to Power Systems (ISAP 2017), San Antonio, USA, 2017.

Covariate shift

We propose a robust algorithm for measuring covariate shift in data sets.

Algorithm 1 Quantifying covariate shift.

```
result ← 0
reliability ← 0
selected ← train_data.add_feature(s, 1)
not_selected ← prod_data.add_feature(s, 0)
data ← selected ∪ not_selected
folds ← cv_folds(data, k)
for model in get_model_candidates() do
    mccs ← list()
    for fold in folds do
        X_train, X_test, y_train, y_test ← fold
        classifier ← DecisionTree(model)
        classifier.train(X_train, y_train)
        y_pred ← classifier.predict(X_test)
        mccs.append(MCC(y_test, y_pred))
    end for
    mcc_mean ← mean(mccs)
    if mcc_mean > result then
        result ← mcc_mean
        reliability ← std(mccs)
    end if
end for
return result, reliability
```

Covariate shift

ASSESSED FEATURES.

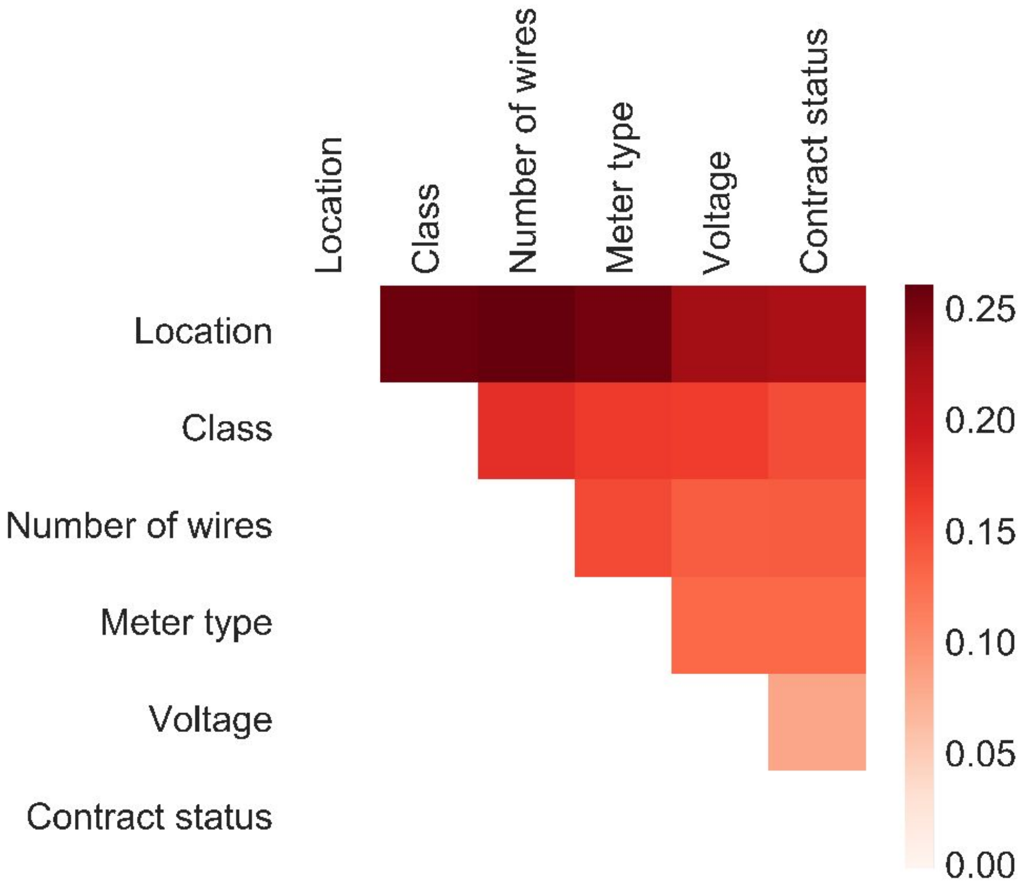
Feature	Possible values
Class	Power generation infrastructure, residential, commercial, industrial, public, public illumination, rural, public service, reseller
Contract status	Active, suspended
Location	Longitude and latitude
Meter type	22 different meter types
Number of wires	1, 2, 3
Voltage	$\leq$2.3kV, $>$2.3kV

Covariate shift

GLOBAL COVARIATE SHIFT OF SINGLE FEATURES.

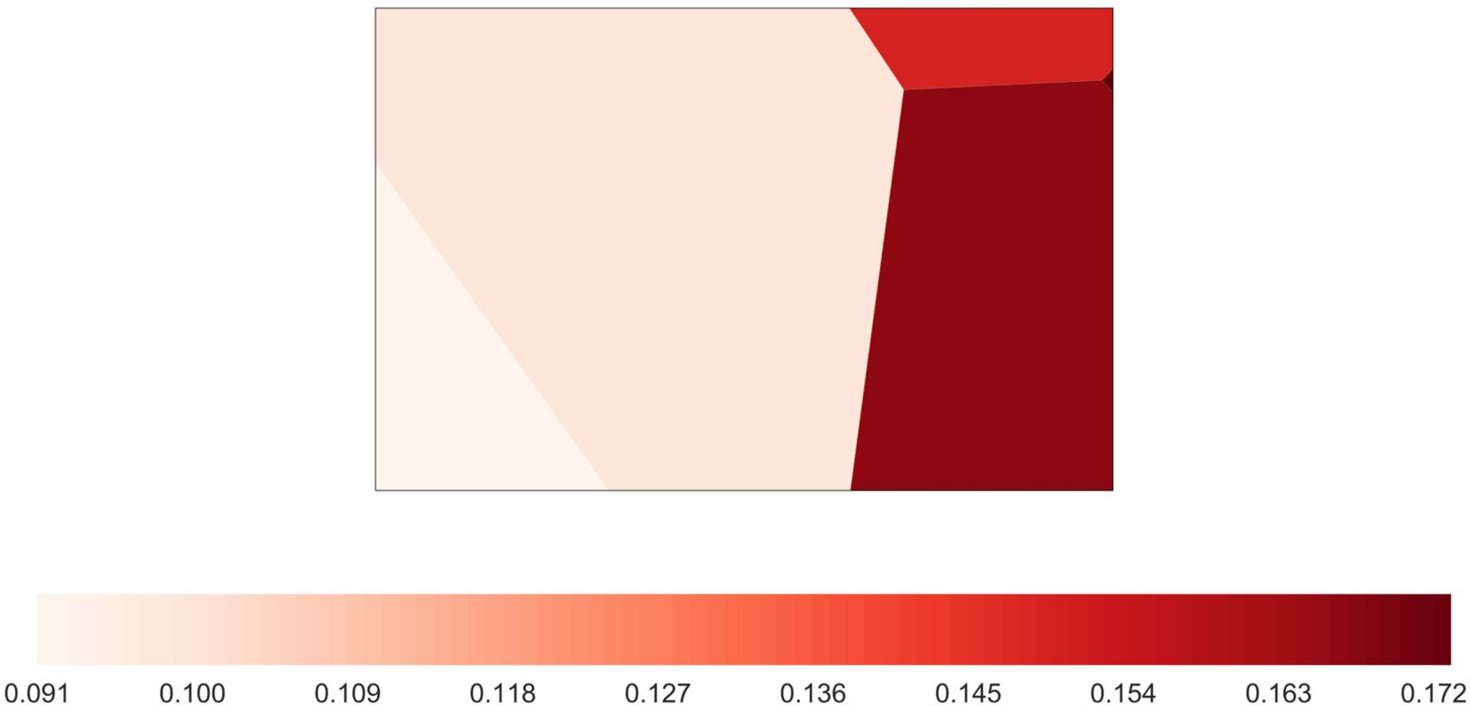
Feature	$\overline{MCC}_{max}$	σ
Location	**0.22367**	0.03453
Class	0.16255	0.01371
Number of wires	0.14111	0.00794
Meter type	0.13158	0.00382
Voltage	0.07092	0.02375
Contract status	0.03744	**0.09183**

Covariate shift



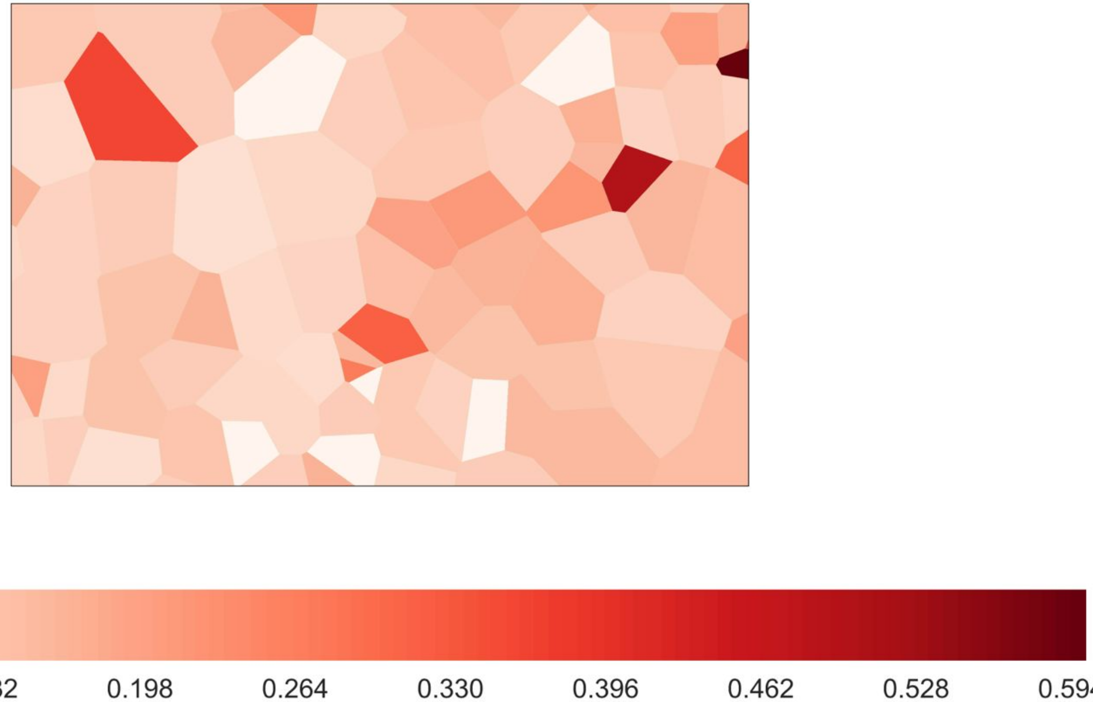
Covariate shift

Regional level:



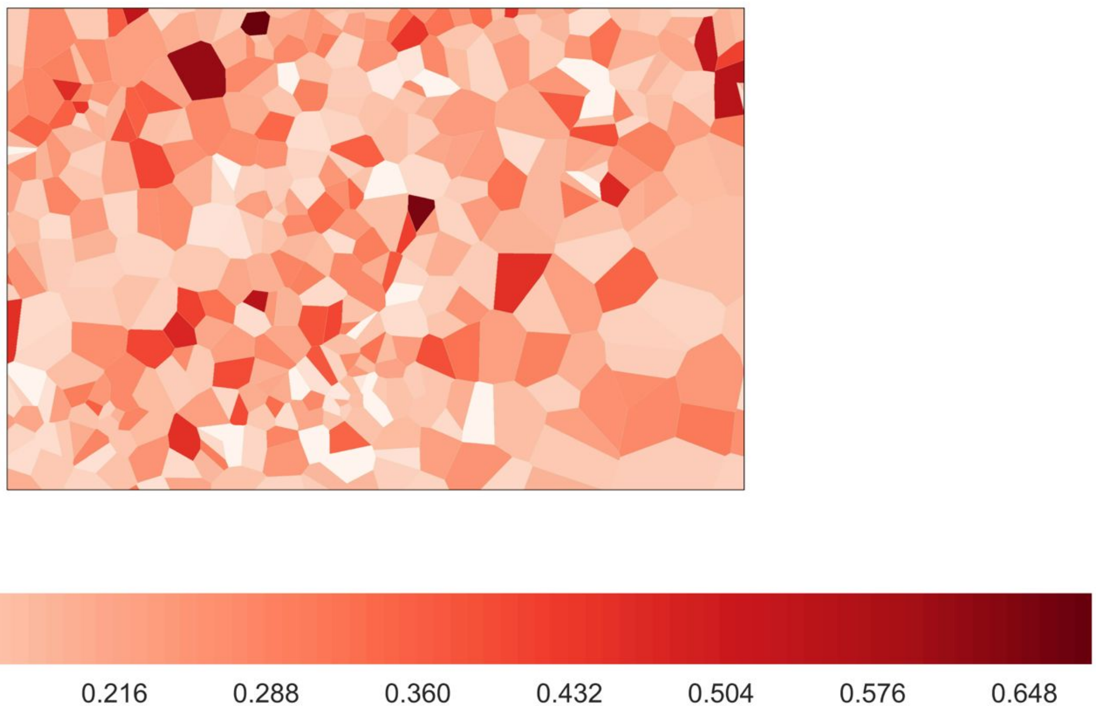
Covariate shift

Municipal level:



Covariate shift

Local level:



Covariate shift

- We have proposed new methods to detect NTL more reliably by reducing covariate shift among other biases.

Covariate shift

ON THE REDUCTION OF BIASES IN BIG DATA SETS FOR THE DETECTION OF IRREGULAR POWER USAGE

PATRICK GLAUNER and RADU STATE

PETKO VALTCHEV

DIOGO DUARTE

To appear in:

Proceedings of the 13th International FLINS Conference on Data Science and Knowledge Engineering for Sensing Decision Support (FLINS 2018), Belfast, UK, 2018.

Covariate shift

Impact of Biases in Big Data

Patrick Glauner[1], Petko Valtchev[2] and Radu State[1]

Published in:

Proceedings of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2018), Bruges, Belgium, 2018.

Scalability

- The number of customers used throughout the research reviewed significantly varies.
- Some papers use less than a few hundred customers in the training.
- Some papers use SVMs with a Gaussian kernels. In that setting, training is only feasible in a realistic amount of time for up to a couple of tens of thousands of customers in current implementations.

Scalability

- Another paper uses the Moore-Penrose pseudoinverse. This model is also only able to scale to up to a couple of tens of thousands of customers.

$$\hat{R} = (H^T H)^{-1} H^T L \qquad (4)$$

where

$$H = \begin{bmatrix} \frac{I_1(t_2)^3 - I_1(t_1)^3}{3s_{1,2}} & \frac{I_2(t_2)^3 - I_2(t_1)^3}{3s_{2,2}} & \dots & \frac{I_n(t_2)^3 - I_n(t_1)^3}{3s_{n,2}} & 1 \\ \frac{I_1(t_3)^3 - I_1(t_2)^3}{3s_{1,3}} & \frac{I_2(t_3)^3 - I_2(t_2)^3}{3s_{2,3}} & \dots & \frac{I_n(t_3)^3 - I_n(t_2)^3}{3s_{n,3}} & 1 \\ \dots & \dots & \dots & \dots & 1 \\ \frac{I_1(t_m)^3 - I_1(t_{m-1})^3}{3s_{1,m}} & \frac{I_2(t_m)^3 - I_2(t_{m-1})^3}{3s_{2,m}} & \dots & \frac{I_n(t_m)^3 - I_n(t_{m-1})^3}{3s_{n,m}} & 1 \end{bmatrix}$$

$$L = [L_2 \quad L_3 \quad \dots \quad L_m]^T$$

$$\hat{R} = [\hat{R}_1 \quad \hat{R}_2 \quad \dots \quad \hat{R}_n \quad l_0]^T$$

Scalability

- A few papers use up to hundreds of thousands or millions of customers.
- An important property of NTL detection methods is that their computational time should scale to large data sets of hundreds of thousands or millions of customers. Most works reported in the literature do not satisfy this requirement.

Comparison of different methods

- Comparing the different methods reviewed in this paper is challenging because they are tested on different data sets.
- In many cases, the description of the data lacks fundamental properties such as the number of meter readings per customer, NTL proportion, etc.

Comparison of different methods

- In order to increase the reliability of a comparison, joint efforts of different research groups are necessary.
- These efforts need to address the benchmarking and comparability of NTL detection systems based on a comprehensive freely available data set.

Comparison of different methods

- Carlos López, Universidad ORT Uruguay, is planning a NTL detection challenge
- Project title: Objective comparison among NTL detection methods in houses

We plan to organize a NTL detection competition

1 post by 1 author

Dr. Ing. Carlos López-Vázquez Jul 26

Hello everybody:

My name is Carlos López, and I am based in Universidad ORT Uruguay (Montevideo). As the subject states, we are applying to our NSF (named ANII) in order to fund a competition intended to objectively compare the performance of various NTL detection methods in a given dataset. Below you will find both the title and summary of the project. The project has been presented last year, and despite a good evaluation some weakness were noticed. ANII stated that we have no evidence that external researchers will be willing to participate in the competition. We kindly request your opinion on the initiative, and if you feel it is possible, send a formal letter stating that you are considering participate in the competition once it is launched (not before mid 2019). If you are interested in the details let me know.

Regards

Carlos

Comparison of different methods

- Goal: Create a simulation environment in which competitors can objectively test and compare their NTL detection methods to the ones of others
- Currently looking for researchers interested in realizing the competition
- Carlos will apply for funding at Agencia Nacional de Investigación e Innovación (ANII)

Comparison of different methods

- Have a look at our mailing list: https://groups.google.com/d/forum/ntl-community
- Details:
  - The comparison will be through a Monte Carlo simulation
  - Unlike typical competitions platforms (e.g. kaggle.com) which just require one classification, the algorithms should be run many times

Comparison of different methods

- Challenges:
  - Derive suitable metrics to assess models
  - Get a sufficiently large dataset that can be put in the public domain

Locality vs. similarity

Is it possible to provide an accurate detection of non-technical losses by using features only derived from provider-independent data?

Locality vs. similarity

Distilling Provider-Independent Data for General Detection of Non-Technical Losses

Jorge Augusto Meira, Patrick Glauner,
Radu State and Petko Valtchev
SnT, University of Luxembourg, Luxembourg

Lautaro Dolberg, Franck Bettinger and
Diogo Duarte
CHOICE Technologies Holding Sàrl, Luxembourg

Published in:

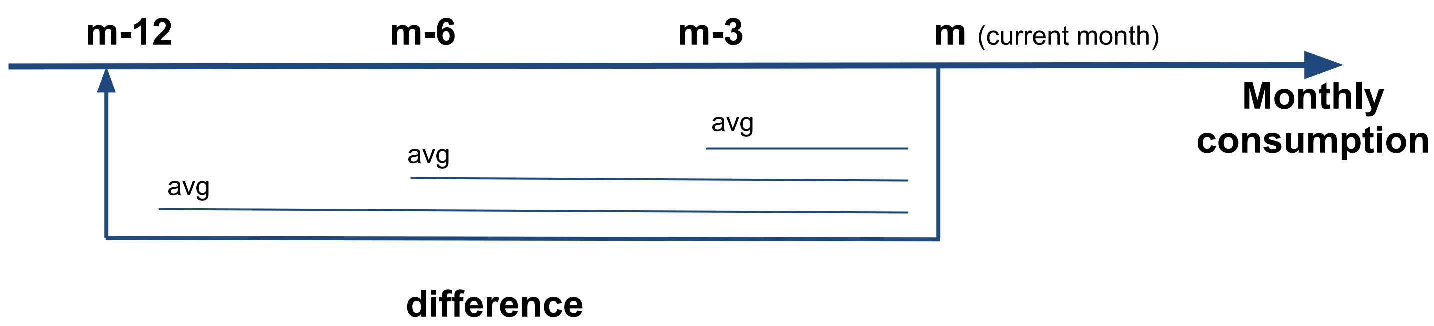
2017 IEEE Power and Energy Conference at Illinois (PECI 2017), Urbana, USA, 2017.

Features

Based on the following categories:

- Temporal: Seasonal, Monthly, Semiannual, Quarterly, Intra Year;
- Locality: Geographical Neighbourhoods;
- **Similarity: k-means clustering using consumption profile**
- Infrastructure: Transformers;

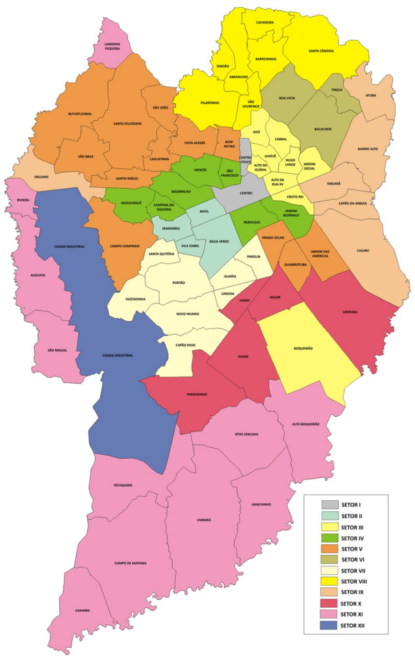
Features

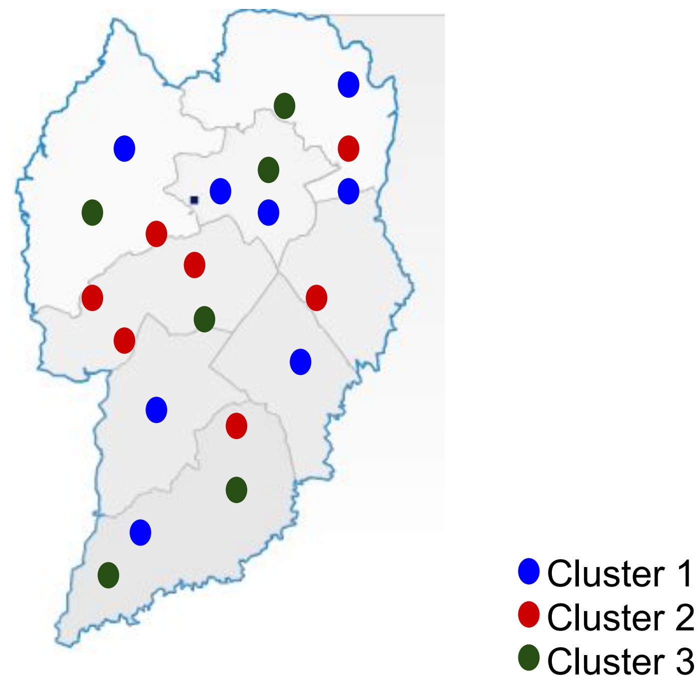


- The temporal features are calculated individually and for each of the three subsequent categories: Locality, Similarity and Infrastructure

Locality vs. similarity

Neighbourhood (code_Neig) vs Consumption Profile (~~code_Neig~~)

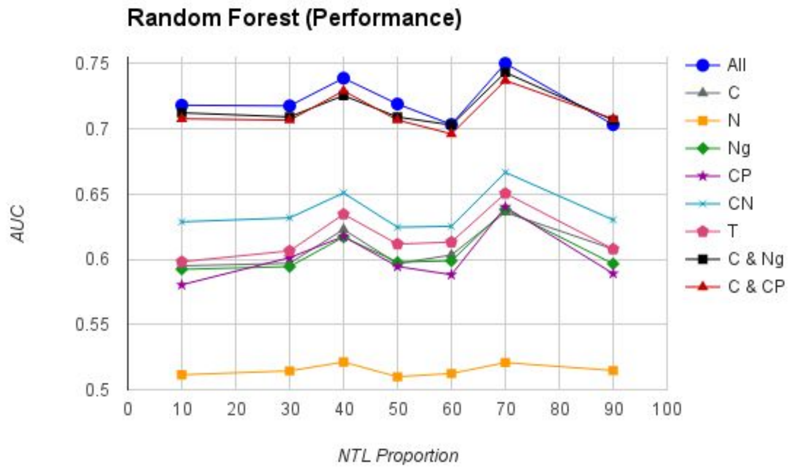


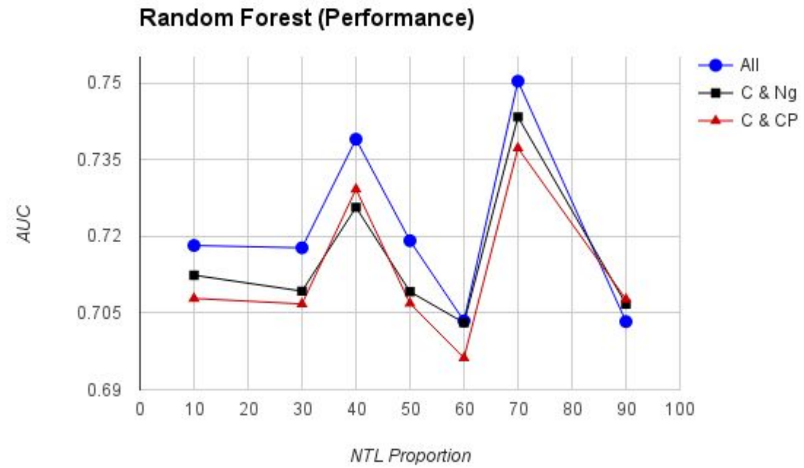


Locality vs. similarity

Set of features	Description
Notes (N)	Meter reader's notes
Consumption (C)	Fixed Interval + Fixed Lag
Consumption & Notes (CN)	Fixed Interval and Notes
Neighbourhood (Ng)	Intra Group (geographical neighbourhood)
Transformers (T)	Intra Group (Transformers)
Consumption Profile (CP)	Intra Group (k-means clustering)
C & Ng	Consumption and Neighbourhood
C & CP	Consumption and Consumption Profile
All	N+C+Ng+CP+T

Locality vs. similarity





Conclusion

- Several sets of features computed using four criteria: temporal, locality, similarity and infrastructure.
- The experimental results show that sets of features supported only by raw consumption data can achieve satisfactory performance when compared with sets composed of ”providers’ dependent features”.

Challenge

A 100% automatic tool

Mixed reality

- How to improve targets selection by taking advantage of the domain expert experience?

Mixed reality

Identifying Irregular Power Usage by Turning Predictions into Holographic Spatial Visualizations

Patrick Glauner*, Niklas Dahringer*, Oleksandr Puhachov*, Jorge Augusto Meira*,
Petko Valtchev†, Radu State* and Diogo Duarte‡

Published in:

Proceedings of the 17th IEEE International Conference on Data Mining Workshops (ICDMW 2017), New Orleans, USA, 2017.

Mixed reality

- How to improve targets selection by taking advantage of the domain expert experience?

HoloLens

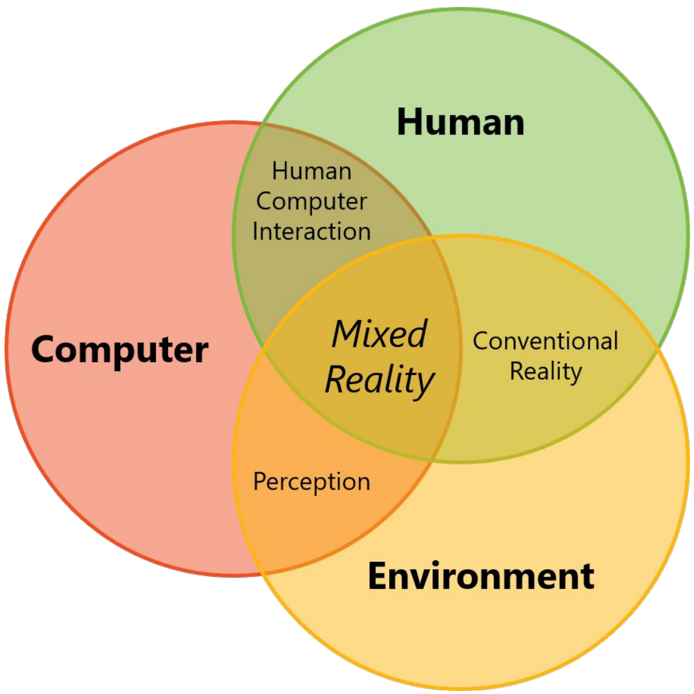
Mixed reality

- In 1994, the paper "A Taxonomy of Mixed Reality Visual Displays" introduced the term *Mixed Reality.*

“... *a particular subset of Virtual Reality (VR) related technologies that involve the merging of real and virtual worlds somewhere along the "virtuality continuum" which connects completely real environments to completely virtual ones.”*

Mixed reality

- In short, MR is the result of blending the physical world with the digital world.

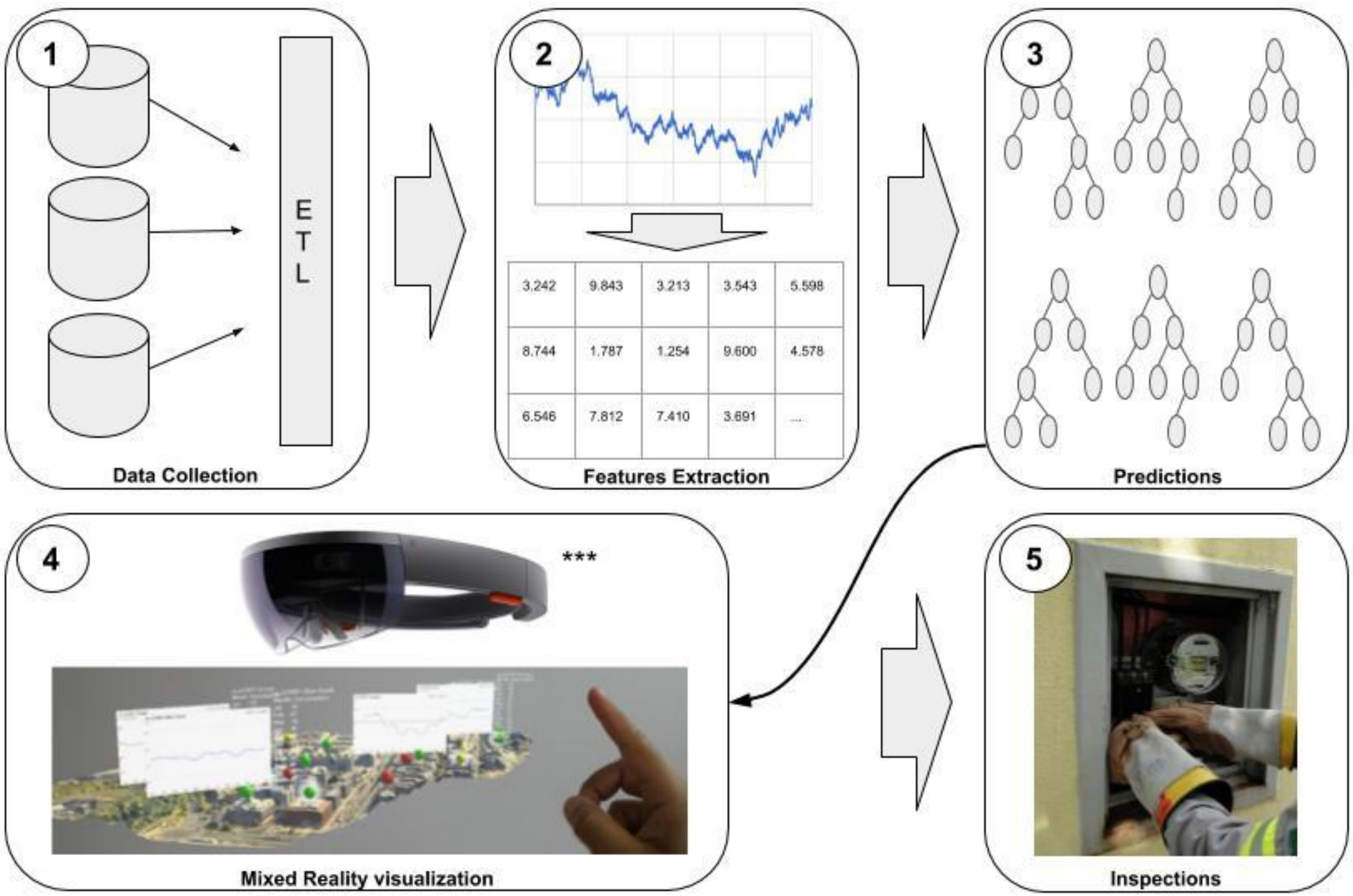


source:
https://developer.microsoft.com/en-us/windows/mixed-reality/mixed_reality

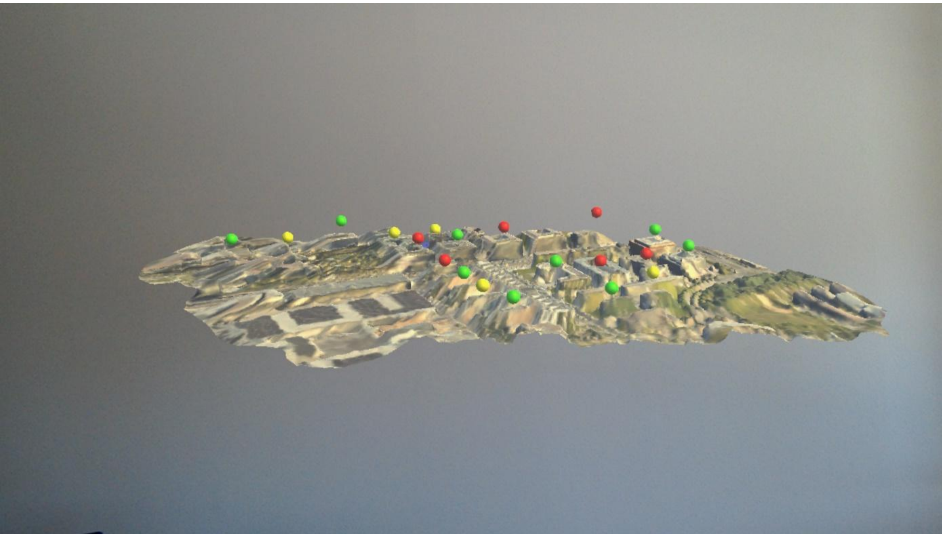
Mixed reality

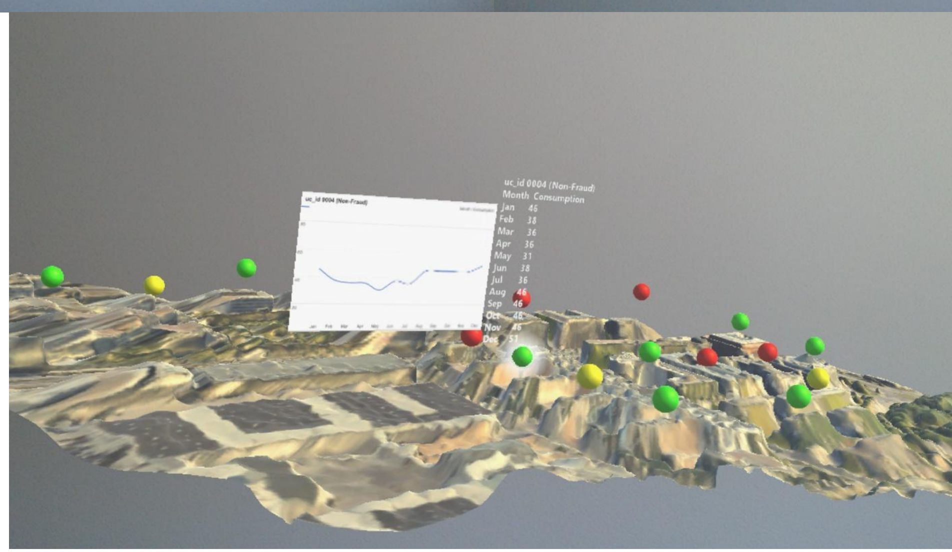
- A novel approach to support to visualize the prediction results in a 3D hologram that contains information about customers and their spatial neighborhood.

Mixed reality



Mixed reality





Mixed reality

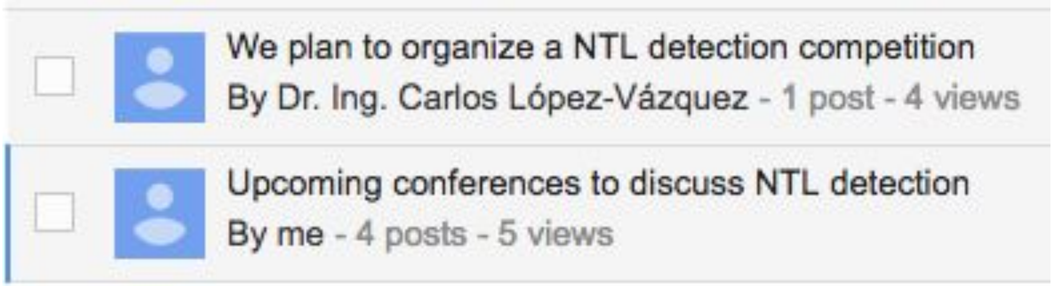
Conclusions

- Non-technical losses (NTL) cause major financial losses to electricity suppliers
- Detecting NTL thrives significant economic value
- Different approaches reported in the literature, superior performance of machine learning approaches compared to expert system
- Many open challenges

Interested in NTL?

Join our mailing list:

https://groups.google.com/d/forum/ntl-community



References

[1] P. Glauner, et al., "**The Challenge of Non-Technical Loss Detection using Artificial Intelligence: A Survey**", International Journal of Computational Intelligence Systems (IJCIS), vol. 10, issue 1, pp. 760-775, 2017.

[2] P. Glauner, et al., "**Large-Scale Detection of Non-Technical Losses in Imbalanced Data Sets**", Proceedings of the Seventh IEEE Conference on Innovative Smart Grid Technologies (ISGT 2016), Minneapolis, USA, 2016.

[3] J. Meira, et al., "**Distilling Provider-Independent Data for General Detection of Non-Technical Losses**", 2017 IEEE Power and Energy Conference at Illinois (PECI 2017), Urbana, USA, 2017.

References

[4] P. Glauner, et al., "**Neighborhood Features Help Detecting Non-Technical Losses in Big Data Sets**", Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing Applications and Technologies (BDCAT 2016), Shanghai, China, 2016.

[5] P. Glauner, et al., "**Is Big Data Sufficient for a Reliable Detection of Non-Technical Losses?**", Proceedings of the 19th International Conference on Intelligent System Applications to Power Systems (ISAP 2017), San Antonio, USA, 2017.

References

[6] P. Glauner, et al., "**Identifying Irregular Power Usage by Turning Predictions into Holographic Spatial Visualizations**", Proceedings of the 17th IEEE International Conference on Data Mining Workshops (ICDMW 2017), New Orleans, USA, 2017.

[7] P. Glauner, et al., "**Impact of Biases in Big Data**", Proceedings of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2018), Bruges, Belgium, 2018.

References

[8] P. Glauner, et al., "**On the Reduction of Biases in Big Data Sets for the Detection of Irregular Power Usage**", Proceedings of the 13th International FLINS Conference on Data Science and Knowledge Engineering for Sensing Decision Support (FLINS 2018), Belfast, UK, 2018.