Utilities Policy 73 (2021) 101304

Contents lists available at ScienceDirect

Utilities Policy

journal homepage: www.elsevier.com/locate/jup

Full-length article

A data-driven approach to reduce electricity theft in developing countries

Ahmad Nadeem [a], Naveed Arshad [b,*]

[a] Department of Electrical Engineering, Lahore University of Management Sciences, Lahore, Pakistan
[b] Department of Computer Science, Lahore University of Management Sciences, Lahore, Pakistan

ARTICLE INFO

Keywords:
Big data analytics
Automated metering infrastructure
Electricity theft

ABSTRACT

Theft of electricity is a problem in many developing countries. But AMI is paving the way for data-centric architecture to help in theft detection. However, a smart grid or even AMR is a long shot for many developing countries due to the costs involved in its large-scale deployment. This paper presents a technique to detect outliers among electricity users that further investigates electricity theft using data analytics on monthly usage data available to every utility company. Using this technique, we have reduced the search space for theft identification to as low as 3.4% of the total customer base.

1. Introduction

Developing countries, such as Pakistan, have a highly regulated national grid feeding electricity to all significant load centers. The country's overall electricity load comprises residential, commercial, industrial, and agricultural sectors, with the residential sector having the largest share. The distribution system is divided into ten utility companies to satisfy the electricity demand. These companies are responsible for providing electricity to the customers and billing and maintenance of the distribution infrastructure. Furthermore, the distribution companies are also in charge of limiting the technical and non-technical losses of the distribution system (Ullah, 2013).

Technical losses within the distribution system result from old and inefficient electric components such as transformers, circuit breakers, relays, and conductors. These losses can be reduced by developing efficient designs and regular maintenance but can never be abolished entirely. On the other hand, non-technical losses occur due to the unmetered use of electricity. The utility companies register this electricity use as a non-technical loss, as it never gets paid for providing this electricity. Many ways of outsmarting or damaging an analog and digital electricity meter are presented in (Hussain et al., 2016) (Czechowski and Kosek, 2016), Hussain et al. and Czechowski et al. respectively.

As the primary objective of the few troublesome customers is to reduce the billed amount, another way is to collude with a utility company officer. Generally, distribution companies in Pakistan have employees that manually note the meter reading of each customer monthly. The customer and employee may conspire to steal electricity either with the employee reporting incorrect electricity consumption information to the companies. Misreporting allows the customer to steal electricity without damaging the electricity meter (Jamil and Ahmad, 2014), or the employee does not report any meter manipulation to the concerned authorities. The utility companies ordered the concerned employees to take snapshots of meter readings on every visit to address this problem (Phoneworld, 2015). The employees resisted this move at a considerable cost to the utility company (Hamariweb, 2016). However, even after the implementation of this policy, electricity theft has not reduced. Using loopholes in the system, the meter reading officers eventually find new ways of dodging the system, e.g., using dummy meters to take snapshots (Dawn, 2016).

Moreover, due to the loss of electricity in theft, utility companies accumulate a large amount of circular debt, which is the amount utility companies cannot pay to the generation companies for buying electricity. The energy sector's circular debt was Rs.1362 billion in January 2019, increasing every year (DAWN, 2019a). Due to this, companies increase the overall price of electricity, which unfairly places the burden of theft on the honest customers who pay their due bills. Another factor that causes a decrease in electricity usage by legitimate customers is forced outages. The distribution companies have adopted this strategy of discouraging electricity theft by forcing outages in targeted areas with high non-technical losses. The forced outages decrease the non-technical losses for the distribution companies and reduce revenues as the legitimate customers are denied the electricity supply. This crude way of decreasing losses means that many legitimate customers also suffer due to the illegal activities of a few (Kazmi et al., 2019).

A more advanced system is needed to overcome such losses in the distribution system and improve efficiency to benefit both the utility

* Corresponding author.
E-mail addresses: ahmad.nadeem@lums.edu.pk (A. Nadeem), naveedarshad@lums.edu.pk (N. Arshad).

https://doi.org/10.1016/j.jup.2021.101304
Received 9 August 2019; Received in revised form 29 September 2021; Accepted 29 September 2021
Available online 26 October 2021
0957-1787/© 2021 Elsevier Ltd. All rights reserved.

companies and the consumers. One particular technology is the Advanced Metering Infrastructure (AMI). This technology provides a communication layer within the electric system, which assists the utilities with efficient administration and surveillance, leading to higher quality customer service (Gungor et al., 2011). AMI also allows the utilities to have an extra layer of security where inefficiencies within the system are widespread. Electricity consumption information is recorded and sent to the utilities after a pre-defined time interval. Similarly, if meter tampering is detected, the utility company is immediately notified, thus further reducing the chances of electricity theft (Yip et al., 2018).

However, regardless of the benefits of AMI within the system, the implementation of such infrastructure on a wide scale is not economically viable for developing countries. AMI deployment within Islamabad Electricity Supply Company (IESCO) and Lahore Electric Supply Company (LESCO), collectively comprising 6.6 million customers, requires $1 billion (Recorder, 2019a). For a distribution company already in debt, the development of such a vast infrastructure is not feasible.

Due to these problems, we propose a data-driven method of detecting outliers using monthly electricity consumption. This approach will allow utility companies to identify customers with anomalous load profiles and prioritize smart meter installation for these customers. Such a technique allows targeted AMI deployment to a subset of the susceptible population, reducing non-technical losses and creating a more stable distribution system that is easy to manage. It will also provide the utility companies with a gradual and smooth transition to better infrastructure without much financial burden.

The remainder of this paper is divided as follows. Section 2 discusses the existing literature regarding multiple ways of detecting electricity theft. Section 3 explains the proposed method and the technique used to identify the outliers within the system. The proposed technique is then implemented in Section 4 on the data used in this study. The policy to introduce AMI in developing countries is discussed in Section 5. Finally, Section 6 concludes this paper.

1.1. Literature review

There are several novel techniques discussed in the literature to detect electricity theft. In (Zehng et al., 2019), Zheng et al. detect electricity theft using two techniques. The first technique focuses on calculating the correlation between the load profiles of customers and non-technical losses. The second technique uses clustering to compare load profiles of customers with each other. This electricity theft detection method was applied to the Irish smart meter dataset with a granularity of half an hour. The electricity theft data was synthesized by modifying the load profiles of randomly selected customers to measure the accuracy of their proposed detection method.

In (Gao et al., 2019), Gao et al. used a real-life smart meter dataset at a sampling rate of 1 h to detect electricity theft using a modified linear regression model. The fraudulent data was created by tampering with selected load profiles randomly. The results show a negative residual from regression for dishonest customers. Machine learning techniques including random forest, support vector machine and neural networks were used in (Razavi et al., 2019) for electricity theft detection, where it was concluded that gradient boosting machine outperforms the rest. Combined with their feature engineering architecture, these techniques were a great way to detect dishonest customers in AMI with at least a half-hour data granularity.

In (Angelos et al., 2011), the method adopted for electricity theft detection is reasonably practical, which requires only six attributes from the historical consumption of the consumer. A two-step algorithm was implemented in this approach. In the first step, C-means-based fuzzy clustering allowing customers with similar consumption patterns to lie within a similar cluster. The second step comprises calculating the Euclidean distance to the cluster's center and obtaining a normalized index score. Potential fraudulent customers or users with irregular patterns turn out to have the highest score.

In addition, several other studies use smart meter data to detect electricity theft using state of the art machine learning algorithms such as generative Gaussian mixture model, gradient boosting theft detection and implementing 2-D electricity consumption data (Zhang et al., 2018), (Singh et al., 2018), (Punmiya and Choe, 2019), (Krishna et al., 2018), (Zheng et al., 2018). All these studies assume that AMI meters are installed and electricity consumption data, at least at half-hour granularity, are available to be used for electricity theft detection.

The representation of the electricity theft model in developing countries was discussed in (Jamil and Ahmad, 2019), where they highlight that one of the main reasons for electricity theft is the collusion of customers and utility employees. Electricity theft becomes feasible when the associated risk of being charged is less than the gain from committing this crime. Moreover, this paper also has several policy recommendations to reduce electricity theft by controlling socio-economic variables such as wages of utility employees, electricity tariffs, and punishment for committing electricity theft. A similar study (Yurtseven, 2015) conducted in Turkey also endorses that education, income, and population are significant determinants of electricity theft. The study also provides empirical results to support their claim.

In (Zorn), Zorn provides excellent insights into the pros and cons of AMI. It further proposes solutions regarding the problems a utility company might face while implementing AMI. However, it also argues that AMI is a technology that has been evolving over the years, where new and better versions of smart meters will be available. Thus, distribution companies should consider these only as commodities. Moreover, several heterogeneous metering systems will be operational simultaneously, and sticking to one technology is not recommended.

In most countries, to reduce the financial burden of an initial investment regarding the integration of AMI, utility companies force customers above a certain threshold of energy consumption to install government-approved smart meters themselves before a specific deadline. This policy will decrease the financial burden on the government as most of these smart meters are subsidized. The grid operators also learn from this minute AMI integration and prepare for a larger scale adaptation. Germany forced customers with annual electricity consumption of more than 10,000 kWh to install smart meters until 2017 and then reduced the threshold to 6000 kWh until 2020 (Smart Energy International).

However, the introduction of AMI in developing countries should be approached differently. These countries share a common problem of high electricity theft. The information provided from past studies might provide practical solutions for developed countries with AMI already installed, but the situation in developing countries is quite different. With almost no AMI integration, only monthly electricity consumption data is available. Our proposed solution uses monthly data and provides a practical solution for introducing AMI in developing countries. Besides better command over the grid, AMI will also help developing countries eradicate electricity theft by prioritizing the installation of smart meters in households with anomalous energy consumption. To the best of our knowledge, no one has proposed such a solution before.

A similar approach to ours is proposed in (Huang, 2021). The encoder tries to reconstruct the inputs using an autoencoder to identify important factors from electricity theft data. Data from a building with electricity theft returns high reconstruction error due to the anomalies. Particle Swarm Optimization (PSO) is used to optimize the hyper-parameters of the autoencoder, and receiver operating characteristic curve is proposed to calculate the optimal error threshold.

Another innovative technique to reduce non-technical losses is the electricity prepayment billing system (EPBS). In (Mwaura, 2012), Mwaura provides a case study of Uganda where such a system was adopted. The customers have to pay in advance to receive credits translated into units of electricity by the smart meter installed in households. The system is effective in reducing electricity theft but has the same problems as AMI. It is an expensive undertaking and requires deployment in targeted localities. However, it has an additional disadvantage: it increases inconvenience for the consumers and requires extensive training and awareness campaigns to adopt this new way of billing.

In all the research reviewed for this analysis, electricity theft is discussed from the load side. Various techniques are discussed to detect electricity theft in residential, commercial, and industrial buildings consuming electricity with anomalous behavior. However, in (Ismail, 2020), the authors discuss the manipulations of distributed generators to misreport their electricity injected into the grid. Recurrent neural networks are used to detect such electricity theft with high accuracy and low false positivity rate.

2. Methodology

In this paper, we propose a new technique that will help the utility company in two ways 1) Utility company will be able to close in on the most likely suspects for the theft of electricity. 2) It will amortize the utility's resources for smart meter roll-out by gradually installing the smart meters first on the outlier customers. Our algorithm determines outliers with erratically fluctuating electricity usage throughout the observation horizon by analyzing residential customers' historical monthly electricity consumption data in a sub-division.

In Pakistan, the distribution system of the electric grid is classified into divisions, which are further classified into subdivisions. Each subdivision is a region that has several customers in close vicinity. Depending on the density of the population, the number of customers varies between 1000 and 2000. The customers of the same subdivision share several characteristics, such as the same neighborhood, similar lifestyle, and weather patterns, which considerably impact electricity consumption. As seen in Fig. 1, the general pattern of energy consumption in Pakistan (A and "State of Industry, 2019) is dependent on temperature (Trading Economics and "Pakis, 2019). With a Pearson correlation value of 0.93 between electricity demand and average temperature, we can say that most electricity consumers in Pakistan typically follow a similar pattern of low electricity consumption when the temperature is low and vice versa.

In the case of Pakistan, neighbors also influence the energy consumption patterns of a household. In (Khalid and Sunikka-Blank, 2017), the authors surveyed ten households in Lahore, Pakistan, recording detailed interviews and analyzing the choices made by the residents that affect electricity consumption in the household. The paper shows how social and cultural acceptance and expectations affect electricity consumption within a neighborhood.

Therefore, it can be concluded that the majority of the residential electricity consumers in the same subdivision have similar electricity consumption patterns over the year. For this reason, the below-mentioned analysis is done separately for each subdivision as electricity consumers from different subdivisions are not comparable.

The electricity consumption of n customers of the same substation for a single month is defined as

$$E = (e_1,\ e_2 \ldots . e_n) \tag{1}$$

where $e_i \ \varepsilon \ N$

e_i is defined as the energy consumption of a customer in a month. These set of integers are clustered using K-means, which is an unsupervised technique that creates pre-defined. k

Number of clusters with each cluster containing observations with the nearest mean for each month (Hartigan and Wong, 1979). These clusters are defined as follows

$$C = (c_1,\ c_2 \ldots .. c_k) \tag{2}$$

where $c_i \varepsilon \ E$

Using each cluster's mean $\overline{c}_i$, we can sort these clusters in order to create ranks for customers as defined in equation (3).

$$R = (r_1,\ r_2 \ldots .. r_k) \tag{3}$$

where $\overline{r}_1 = \overline{c}_{\min} \Big| \overline{r}_k = \overline{c}_{\max}$

The customers in the cluster with the lowest mean $\overline{c}_{min}$ will be ranked lower, i.e., r_1, having a mean of $\overline{r}_1$ and customers in the cluster with the highest mean $\overline{c}_{max}$, will be ranked higher, i.e., r_k. If this is done for m number of months, it creates a matrix **R**, of ranks with dimensions $k * m$, which is illustrated as follows

$$R = \begin{bmatrix} r_{11} & r_{12} \ldots . r_{1m} \\ r_{21} & r_{22} \ldots . r_{2m} \\ . & . \\ . & . \\ . & . \\ r_{k1} & r_{k2} \ldots . r_{km} \end{bmatrix}$$

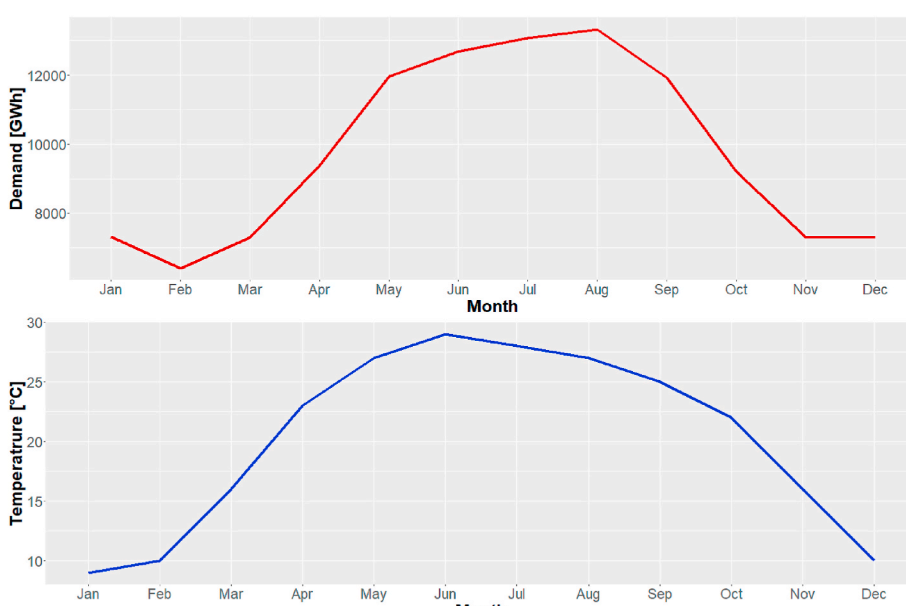


Fig. 1. Monthly electricity demand and average temperature pattern of Pakistan. Pearson correlation between the two is 0.93.

where each element in the matrix represents a rank that is shared by several customers of the same subdivision with electricity consumption nearest to the mean of that rank for a particular month. From this, the ranks of a particular customer Q can be derived and defined as

$$Q = (q_1,\ q_2.....q_k) \tag{4}$$

where $q_{ij} \varepsilon \mathbf{R}$

Such that q_{ij} is the i^{th} rank in the j^{th} month. The rank of a customer will not change much over the months if the customer's electricity consumption follows the load profile of the subdivision. However, if the customer's load profile deviates from the expected behavior of the subdivision, the change in ranks will be more prominent. A range in rank for every customer can be defined as

$$\Delta Q = \max(Q) - \min(Q) \tag{5}$$

$$\Delta Q > \sigma \tag{6}$$

σ is the threshold regarding outlier detection. After setting an appropriate threshold for ranking variation, any customer crossing this threshold can be marked as an outlier, as explained in equation (6). The choice for the threshold is directly proportional to k, i.e., the number of clusters pre-defined for k-means clustering. A level of strictness can be set using σ. If the value of σ is increased, the system will become more lenient. On the other hand, if the σ is decreased, the system would become more strict.

This method for detecting outliers assumes that there are more regular customers than anomalies in a substation. As the majority sets the trend, if the colluders are more in number, the pattern will be disrupted, and the suggested method will be ineffective in detecting outliers. However, this is a realistic assumption as it is observed that there are always more honest customers than colluders.

All the outliers detected through this technique are not necessarily electricity thieves; there will be several false positives. For example, a customer identified as having a suspicious pattern because of a drop in cluster ranking might be due to them being on vacation and away from home. On the other hand, a customer detected having unusual behavior because of an increase in ranking can be due to some new electrical appliance at home, such as an air conditioner, or perhaps the addition of some temporary residents that cause increased electricity consumption for the customer.

This method provides a practical way of detecting outliers as monthly electricity consumption data are evaluated readily available to every utility company. The technique provides us with a small number of customers targeted for smart meter installations. Using this technique, utility companies get a cost-effective approach for AMI integration into the system.

3. Case study

To test our proposed method of detecting outliers from monthly electricity consumption data, we selected LESCO (Lahore Electric Supply Company), which provides electricity to Lahore, Punjab, and its outskirts. LESCO has around 4.5 million registered customers and has vast AT&C losses, and regular disruptions in the grid are increasing their operations and maintenance costs (The Nation, 2019). Like most utility companies, LESCO has substantial circular debt and cannot afford AMI smart meters installation on a larger scale. However, by using our scheme, both of these issues can be resolved together.

For this study, we got monthly electricity consumption data of

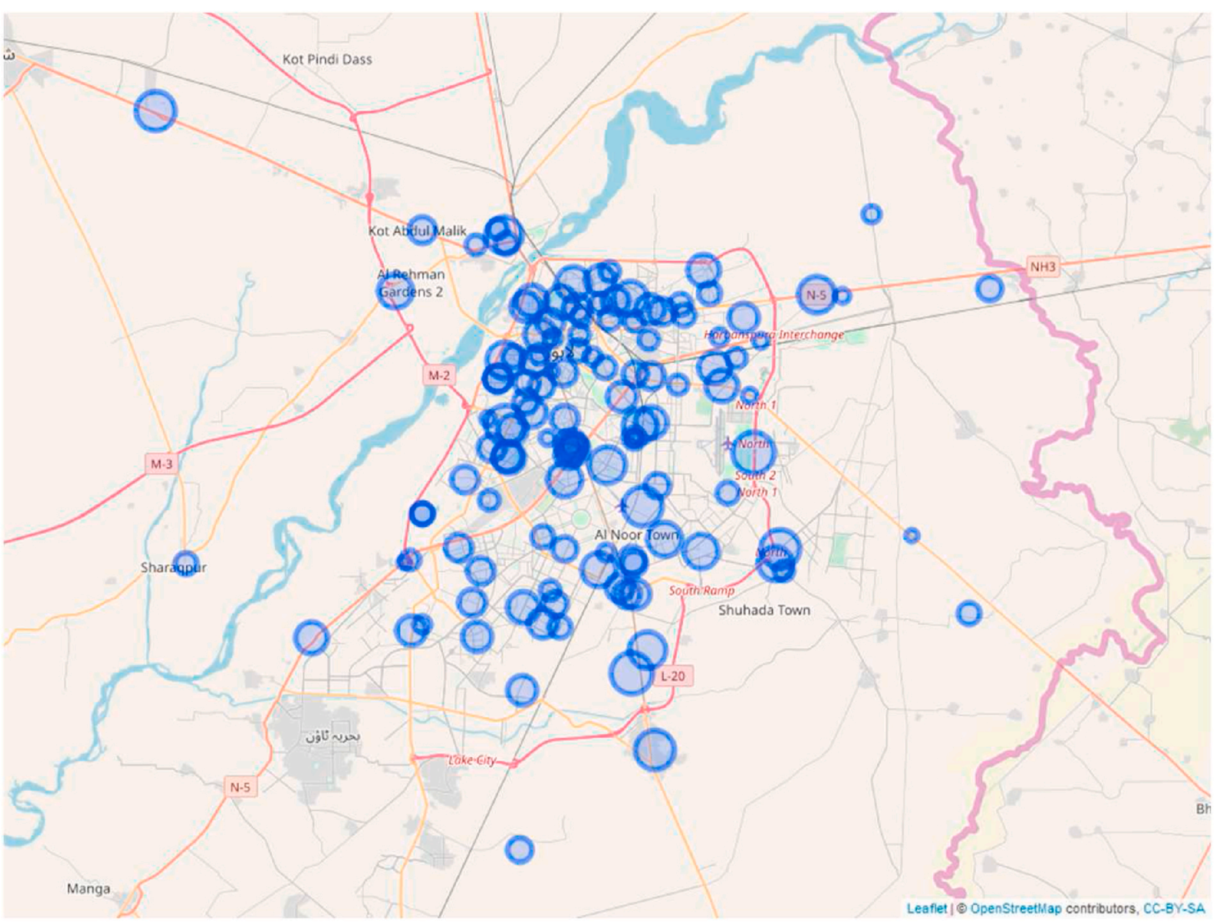


Fig. 2. GEO-LOCATIONS OF SUBDIVISIONS IN LAHORE. EACH CIRCLE IS A SUBDIVISION, WHERE THE SIZE OF THE CIRCLE IS DIRECTLY PROPORTIONAL TO THE NUMBER OF ITS CUSTOMERS.

100,000 residential customers from LESCO for a whole year. The data were randomly selected from 178 subdivisions of LESCO, which gives us approximately 560 customers per subdivision. The geo-location of these subdivisions over the city of Lahore is as shown in Fig. 2, where the size of the circles represents the number of customers in the respective subdivision and a larger circle represents more customers in that particular subdivision.

4. Regular customers

In Fig. 3, the three lines represent the demand profiles of three customers of the same subdivision within LESCO, while the bars represent the average demand of the subdivision for the whole year to which the customers belong. The line labeled "HIGH" represents a customer load profile that consumes more electricity than the average of that subdivision. The demand of customer "HIGH" remains more than the average for the whole year. The load profile labeled 'AVG' is an average customer where the demand for 'AVG' remains closer to the subdivision average. The line labeled 'LOW' represents a customer with less consumption than the average of that subdivision for the whole year.

5. Outlier detection

In Pakistan, the middle of the year is when high electricity consumption is expected because of the high usage of air conditioners during these months, i.e., May till September. Fig. 4 shows an anomalous load profile that deviates from the general trend of the subdivision. The customer followed the general pattern at the beginning of the year. As the average consumption of the subdivision was increasing, it was expected that the customer would consume more electricity. However, his consumption dropped during those months, and then the load profile regains the general trend once the summer months were over. This type of behavior is quite common among customers that commit electricity theft. We expect to filter out these customers as outliers so that smart meters can be installed in their homes first and monitored closely for their electricity consumption.

Another type of anomalous behavior is shown in Fig. 5. These customers do follow the general trend in the beginning but consume more electricity than expected. From January till May, the electricity consumption of this particular customer is less than the subdivision's average, which increases above average from June till October, after which it reduces back to below average. One of the common ways of electricity theft is by secretly tapping the power lines of a neighbor. In such a way, the neighbor is billed for extra electricity consumption. This type of outlier can be a victim of such an act.

6. Results and discussion

The method explained above was applied to the given data of a whole year. For each month, ten clusters were created for each subdivision using k-means clustering and sorted to form ranks. Customers with the highest electricity consumption profiles in their subdivision get a ranking of ten, and customers with the lowest get a ranking of one. The threshold is set to three ranks.

Fig. 6 below shows the ranks assigned to the five customers discussed above for each month. The three customers with high, average, and low electricity consumption have ΔQ of less than three. However, the ranking of Outlier 1 varies from two to six, which gives this customer a ΔQ of four. Similarly, Outlier 2 has a ΔQ of six. As the threshold σ is set to three, both outliers are labeled as anomalies.

Fig. 7(a) shows the geo-location of all the customers selected randomly in LESCO. In contrast, Fig. 7(b) shows the detected outliers in each subdivision, where the size of the circles indicates the number of outliers in every subdivision. With the parameters mentioned above, 3398 customers were found to be outliers, which is 3.4% of the given data. The detected outliers are likely to be colluding and require further investigation by the utility company. This analysis provides a good starting point for the installation of AMI smart meters.

Since the given data is not labeled, it is difficult to know which customers tried to steal electricity from the outliers. However, there are indirect ways to check the effectiveness of this method. LESCO has published its AT&C loss data for each feeder in its distribution system (LESCO, 2019). As each subdivision can have multiple feeders, the average of AT&C losses at the feeder level is calculated using the data provided by LESCO (LESCO, 2019). These calculations help correlate the number of outliers detected to the losses for each subdivision, as shown in Table 1. The results show that subdivisions with higher losses have a higher number of cases detected as outliers.

6.1. Policy implications

The utility companies in developing countries have limited resources to upgrade their technology. In the absence of new technologies, data-driven decision-making helps them to strategize their limited investments. The new approach to data-driven decision-making is getting much attention.

It is possible to obtain optimal solutions for problems using the state of the art techniques. Many times, this is possible with the available data. In this paper, the clustering-based technique of anomaly detection using monthly electricity consumption data can be used for two purposes reducing electricity theft and systematic rolling out of AMI smart meters with budget constraints. The policy implications for both

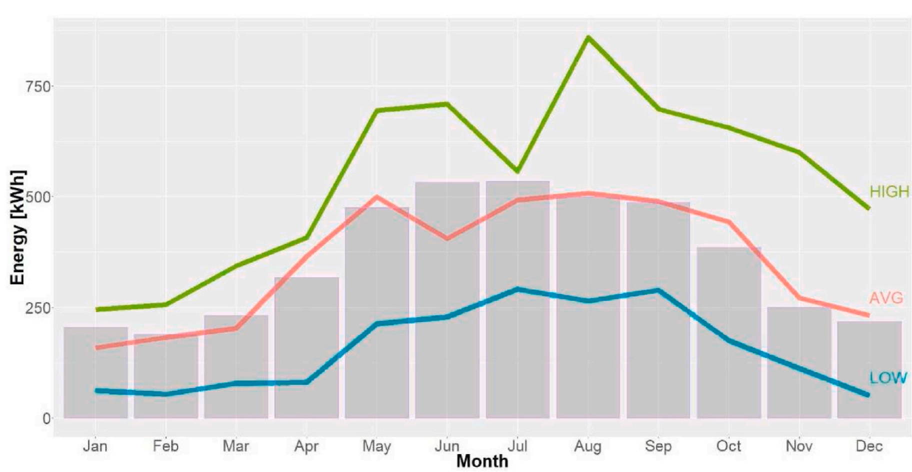


Fig. 3. Customer electricity demand relative to average demand of the subdivision.

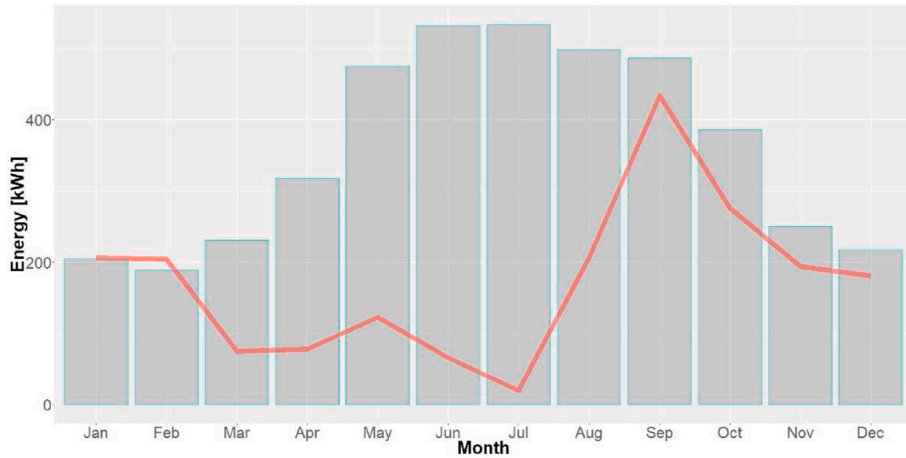


Fig. 4. Suspicious customer deviating from normal trends of the subdivision.

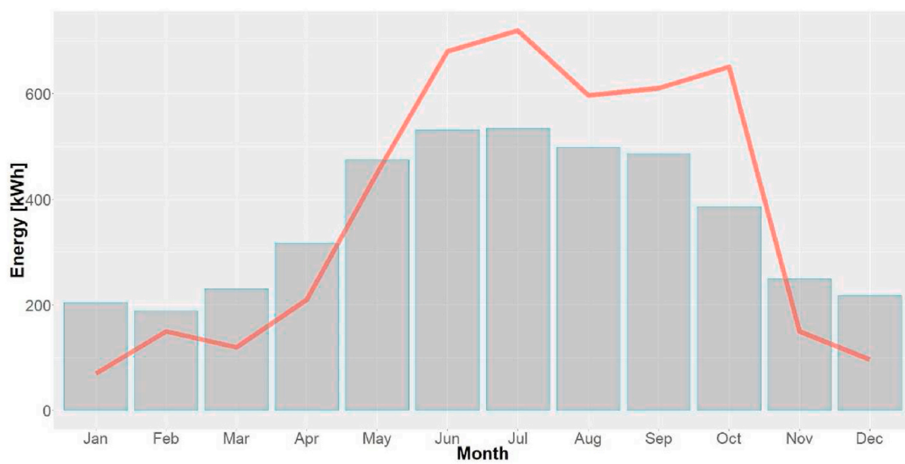


Fig. 5. Suspicious customer consuming more electricity.

purposes are listed below:

6.2. Identifying suspected customers for potential theft

The technique described in this paper effectively closes in on the suspected customers responsible for stealing the electricity. As the number of suspected customers is small, the utility should arrange for a physical inspection of the electricity connection. There are more than twenty ways to steal electricity by tampering with the electricity meter in different ways discussed in detail by Khan et al. (2016). Various tampering methods include the following:

- Tampering meter body or using bogus seal
- Creating a hole in the meter cover to reach the reset button
- Reversing through the terminal
- Looping/Shunting in terminal
- Changing phase polarity
- Using neutral wire of another meter
- Tilting meter
- Damaging by soaking in water, acid
- Using a microwave oven to damage meter circuitry
- Removing EPROM from the meter
- Installing remote-controlled device to start/stop meter whenever desired

Beyond meter tampering, the inspection team should also be looking for direct connections bypassing the meter. Similarly, the meter's wiring should be closely inspected to see if a neighbor is stealing electricity.

In Pakistan, some customers are categorized as lifeline customers who get electricity at a significantly reduced rate and use 300 units or less in a month. Through inspection, the team can find if the customer is truly a lifeline customer or not. Many customers, through stealing, fall in the lifeline category and use the low rate; however, they consume much more. Punishment should be carried out according to applicable laws to deter people from further committing this crime.

6.3. Strategize effective smart meter roll out

The smart meter roll-out is an expensive proposition. As mentioned before, complete smart metering requires a tremendous amount of funding. More than 60% of consumers in Pakistan are lifeline consumers, which means they consume 300 units or less monthly, indicating that these customers may not be able to self-fund smart meters because of their lower socioeconomic status.

Therefore, the utility has to find a way to install smart meters cost-effectively. Typically, smart metering strategies are first applied to high revenue customers. However, we believe that this strategy may not work in Pakistan as effectively as it may work in other countries. Based on the available budget for smart meter roll-out, the following strategy may work for utility:

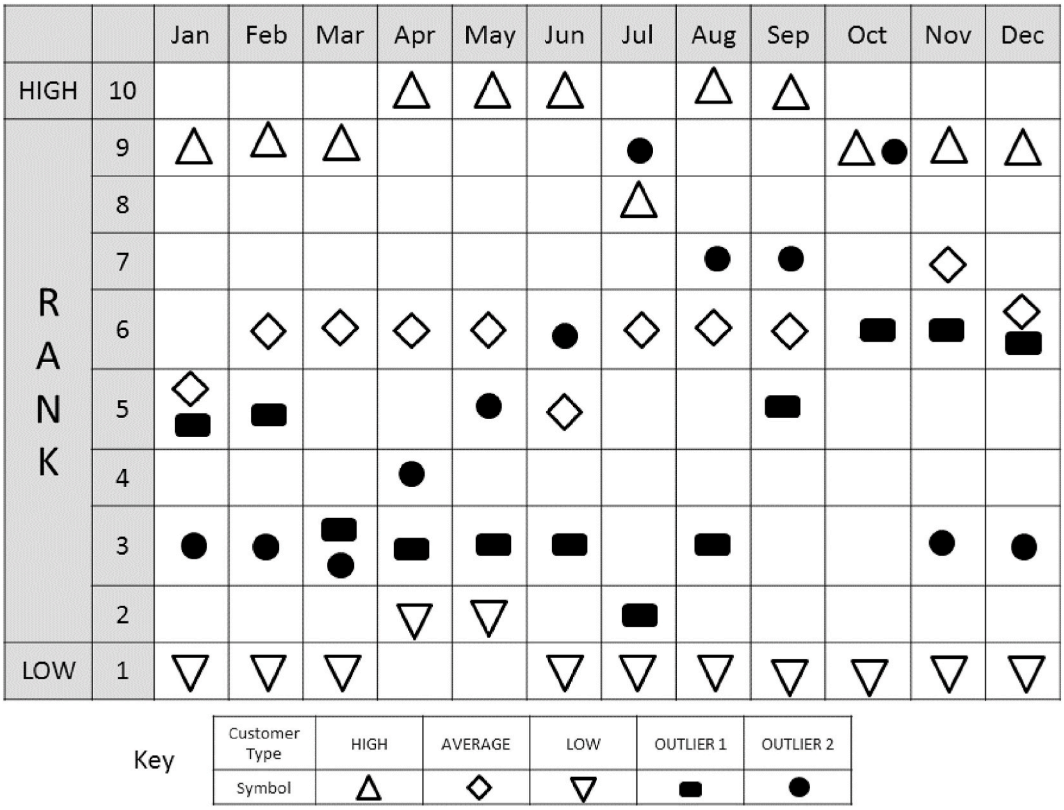
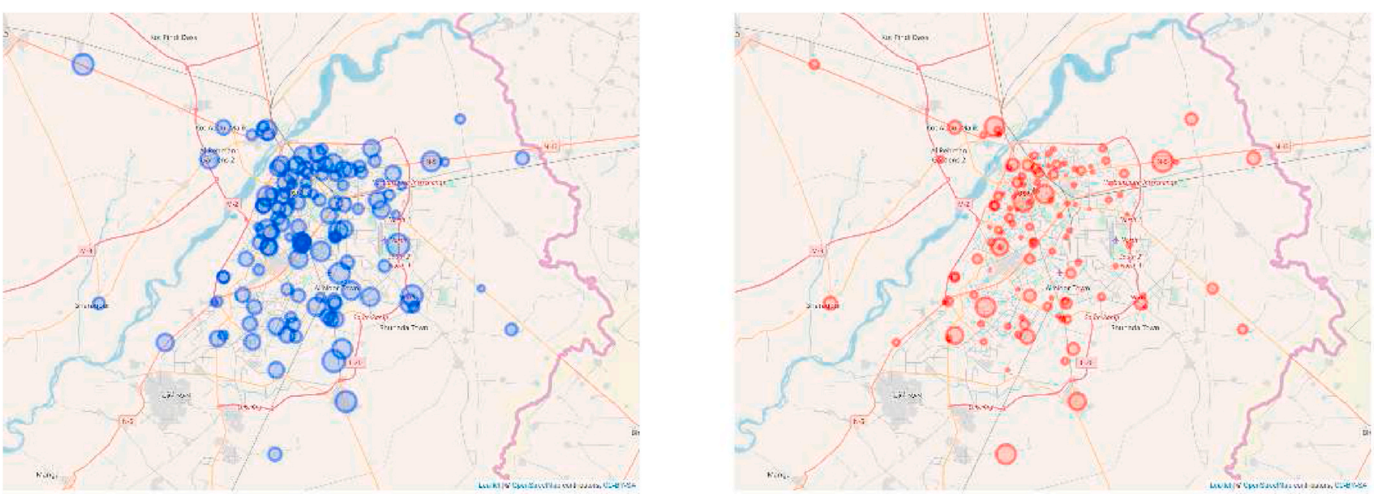


Fig. 6. Rank variation of five different customers.



(a) Geolocation of all customers in LESCO

(b) Geolocation of outliers in LESCO. Size of circle represent the number of outliers in each subdivision

Fig. 7. Map showing locations of customers and associated outliers.

- The distribution companies should calculate ΔQ for all their customers. The customers with the highest ΔQ, i.e., customers with anomalous behavior, should be targeted first. For these targeted customers, the government should provide a fund for the utility company to install the AMI smart meters. Most such customers are never going to pay for their installation. In the initial stages of the roll-out of AMI, the value of σ should be lenient so that there are fewer outliers on which to focus. As the distribution company keeps installing the AMI meters, the value of σ will decrease to include new customers who are relatively less likely to be outliers or suspects of electricity theft. To further facilitate the transition to AMI, a group of low-revenue customers can have a smart meter installed collectively at the start. Later, customers can be individually catered for smart meter installation.
- If the budget allows, the sub-divisions with several customers having high ΔQ value should be covered entirely with smart meters.
- Any new meter installed should be a smart meter.

Table 1
Correlation between subdivision and outliers.

# Of Subdivision	# Of Outliers	% Of Outliers of Total Customers	% Average Loss
108	1523	2.6	Less than 20
61	1492	4.7	Between 20 & 25
8	351	10	Between 25 & 30
1	32	7.1	More than 30

- Faulty meters or meters that are replaced at the end of their life (10–15 years) should be replaced with smart meters.
- High revenue customers can be given the option to install self-funded smart meters to manage their electricity consumption efficiently.

Using the strategy above, the smart meter in utility companies will roll out by reaping the most benefit.

6.4. Further discussion

The smart meters will be installed using the proposed method for both types of outliers, i.e., consumers that themselves are suspected of stealing electricity and those who are victims of it. However, it is difficult to determine the culprits due to collusion between customers and meter reading officers. These colluding officers may also provide illegal connections through the power lines of legitimate customers without their knowledge (Jamil, 2018). Once detected as outliers, the victims or legitimate customers will become more vigilant, further discouraging electricity theft as the legitimate customers will make sure that no one is illegally tapping into their house's power lines. Moreover, AMI smart meters will eventually be installed for all customers, so early installations for these outliers also serve the long-term goal.

Another way to discourage electricity theft can be by rewarding electricity consumers who have low ΔQ. These customers follow the expected electricity consumption pattern, and it is highly anticipated that they are not involved in electricity theft. By rewarding such customers, utility companies can encourage fair electricity usage, keep such customers from committing electricity theft, and discourage the culprits.

AMI is the best solution to tackle electricity theft besides providing opportunities for other demand-side management techniques. It eliminates the need for meter reading officers as AMI smart meters can be remotely monitored and even controlled. AMI has a tree infrastructure, i.e., smart meters are installed at every node in the distribution system to help track electricity leakage. Furthermore, the installation of AMI is inevitable as it resolves several other problems faced by the utility companies.

To create a profitable distribution system, utility companies have to reduce the non-technical losses, which would not be possible without changing their policies. A complication of electricity theft is that it is not considered unethical by the customers, even after knowing it is illegal. The customers should be informed about the problems that others face due to their negligence, i.e., more load shedding hours and increased prices. To discourage such illegal activities, the government of Pakistan has passed strict laws (Recorder, 2019b). However, these laws are weakly implemented, due to which the utility companies are still facing high non-technical losses. To overcome such problems, the government should publicly announce the convicts of electricity theft, which will help to discourage this illegal and unethical act.

During the transition of AMI integration into the distribution system, the utility companies would also require data centers for data acquisition and storage and data processing units for data analytics. The research literature provides several methods for the acquisition, storage, and processing of this big data. It is expected that energy systems will generate the second-largest amount of data after the internet. The current techniques are not enough to handle such a massive amount of data; new innovative and efficient methods need to be developed (Zhou et al., 2016), (Yu et al., 2015).

The utility company itself should play the most prominent role in reducing electricity theft. By keeping strict surveillance on the officers assigned to reading the meter, theft from meter tampering and illegal connections (DAWN, 2019b) can be reduced. Since the primary motivation for an officer to collude is monetary, electricity theft can be reduced by increasing their wages. However, officers that are still found guilty should be penalized.

Moreover, there is a huge communication gap among various utility companies. By sharing experiences, these companies can guide each other and learn from others' mistakes. The utility companies should also seek external help from universities with innovative solutions to utility company problems. By sharing resources, utility companies can optimize their operations and make better decisions. Such as choosing from several AMI options available in the market right now; selecting one would be a tedious job. However, with the help of other utility companies and academia, the decision can be made efficiently.

Customers can also play a vital role by alerting the authorities about electricity theft happening in their vicinity. However, the customer's anonymity that alerts concerned authorities should be guaranteed for his safety and also rewarded. In this way, the customers would be encouraged to help the utility companies. Electronic and print media can also help utility companies convey their message by reporting electricity theft incidents and advertisements discouraging electricity theft.

7. Conclusion

Commonly it is thought that any data analytics technique requires an exorbitant amount of data. However, in this paper, it can be seen that insights can be found even with the data available to the utility companies. In the absence of automated metering infrastructure and smart meter data, the proposed approach is a practical solution for utility companies, especially those in the developing world. It will help the utility companies emerge from the vicious cycle of circular debt accumulated due to the non-recovery of supplied electricity. However, please note that the technique presented in this paper is neither a panacea for electricity theft nor a complete solution for this problem. Electricity theft is a widely spread problem, and anybody with mal intentions will find a way to deceive the system. Data analytic techniques can only reduce the potential theft to a much smaller size, including false positives and false negatives. Since theft detection does not constitute a ground truth, any technique to find theft depends on the best close approximation. Since there are non-technical losses data at the feeder level for LESCO, the proposed technique has a better probability of finding potential theft. We plan to use several big data techniques and tools to identify many other behaviors on elusive datasets in the future. Analytics offers the best way forward in the absence of smart metering data, even on small-size datasets.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

NEPRA, 2019. State of Industry Report 2019 [Online]. Available: https://nepra.org.pk/publications/State//of//Industry//Reports/State//of//Industry//Report//2019.pdf. (Accessed 25 April 2019). Accessed.

Angelos, E.W.S., Saavedra, O.R., Cortés, O.A.C., de Souza, A.N., 2011. Detection and identification of abnormalities in customer consumptions in power distribution systems. IEEE Trans. Power Deliv. 26 (4), 2436–2442.

Czechowski, R., Kosek, A.M., 2016. The most frequent energy theft techniques and hazards in present power energy consumption. In: Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids. CPSR-SG), 2016.

Dawn, 2016. 'Dummy Meter' Scam: Reading Anomalies Continue to Cast a Slur on Lesco's Image [Online]. Available: https://www.dawn.com/news/1300954. (Accessed 25 April 2019). Accessed.
DAWN, 2019a. Circular Debt Has Risen [Online]. Available: https://www.dawn.com/news/1455026. (Accessed 25 April 2019). Accessed.
DAWN, 2019b. Power Theft Not Possible without Staff Connivance [Online]. Available: https://www.dawn.com/news/1449298. (Accessed 25 April 2019). Accessed.
Gao, Y., Foggo, B., Yu, N., 2019. A physically inspired data-driven model for electricity theft detection with smart meter data. IEEE Trans. Industrial Informatics 15, 5076–5088.
Gungor, V.C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., Hancke, G.P., 2011. Smart grid technologies: communication technologies and standards. IEEE Trans. Industrial Informatics 7 (4), 529–539.
Hamariweb, 2016. Lesco Fails to Implement Meter Reading Project [Online]. Available: https://hamariweb.com/enews/lesco-fails-to-implement-meter-reading-project_nid1752906.aspx. (Accessed 25 April 2019). Accessed.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a k-means clustering algorithm. J. Royal StatisticalSoc. C (Appl.Stat.) 25 (1), 100–108.
Huang, Y.a.Q.X., 2021. Electricity theft detection based on stacked sparse denoising autoencoder. Int. J. Electr. Power Energy Syst.
Hussain, Z., Memon, S., Shah, R., Bhutto, Z.A., Aljawarneh, M., 2016. Methods and techniques of electricity thieving in Pakistan. J. Power Energy Eng. 4 (9), 1.
Ismail, M.e. a., 2020. Deep Learning Detection of Electricity Theft Cyber-Attacks in Renewable Distributed Generation. IEEE Transactions on Smart Grid.
Jamil, F., 2018. Electricity theft among residential consumers in Rawalpindi and Islamabad. Energy Pol. 147–154.
Jamil, F., Ahmad, E., 2014. An empirical study of electricity theft from electricity distribution companies in Pakistan. Pakistan Dev. Rev. 53 (3), 239–254.
Jamil, F., Ahmad, E., 2019. Policy considerations for limiting electricity theft in the developing countries. Energy Pol. 129, 452–458.
Kazmi, H., Mehmood, F., Tao, Z., Riaz, Z., Driesen, J., 2019. Electricity load-shedding in Pakistan: unintended consequences, opportunities and policy recommendations. Energy Pol. 128, 411–417.
Khalid, R., Sunikka-Blank, M., 2017. Homely social practices, uncanny electricity demands: class, culture and material dynamics in Pakistan. Energy research & social science 34, 122–131.
Khan, Y.H., Riaz, M., Khan, I.J., Khan, J.A., Khan, S., others, 2016. Comparison OF power theft techniques and their identification: an overview concerning Pakistan. Sci. Int. 28 (3).
Krishna, V.B., Gunter, C.A., Sanders, W.H., 2018. Evaluating detectors on optimal attack vectors that enable electricity theft and DER fraud. IEEE J. Selected Topics Signal Process. 12 (4), 790–805.
LESCO, 2019. Schedule of Approved Shut Down. LESCO [Online]. Available: http://lesco-bills.gov.pk.iis4001.shared-servers.com/lescoadmin/ShutdownFiles/Shutdown%20Schedule.pdf. (Accessed 25 April 2019). Accessed.
Mwaura, F.M., 2012. Adopting electricity prepayment billing system to reduce non-technical energy losses in Uganda: lesson from Rwanda. Util. Pol. 23, 72–79.
LESCO, 2019. Load Management Schedule. LESCO [Online]. Available: http://www.lesco.gov.pk/LoadManagement. (Accessed 25 April 2019). Accessed.
Phoneworld, 2015. IESCO to Use Android Smartphones for Electricity Meter Readings [Online]. Available: https://www.phoneworld.com.pk/iot-initiative-iesco-to-use-android-smartphones-for-electricity-meter-readings/. (Accessed 25 April 2019). Accessed.
Punmiya, R., Choe, S., 2019. Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. IEEE Trans. Smart Grid 10 (2), 2326–2329.
Razavi, R., Gharipour, A., Fleury, M., Akpan, I.J., 2019. A practical feature-engineering framework for electricity theft detection in smart grids. Appl. Energy 238, 481–494.
Recorder, Business, 2019a. AMI in IESCO & LESCO [Online]. Available:, 2019 4 2019. https://fp.brecorder.com/2019/02/20190209445838/.
Recorder, Business, 2019b. Electricity Act Amended: Power Theft Declared Serious Crime [Online]. Available: https://fp.brecorder.com/2016/02/2016021817477/. (Accessed 25 April 2019). Accessed.
Singh, S.K., Bose, R., Joshi, A., 2018. Energy Theft Detection in Advanced Metering Infrastructure. IEEE 4th World Forum on Internet of Things (WF-IoT), 2018.
Smart Energy International. Germany moving ahead with smart meter rollout plans,. https://www.smart-energy.com/magazine-article/germany-moving-ahead-smart-meter-rollout-plans/.
The Nation, 2019. Lesco Tops in Line Losses in Punjab Discos [Online]. Available: https://nation.com.pk/09-Jan-2018/lesco-tops-in-line-losses-in-punjab-discos. (Accessed 25 April 2019). Accessed.
Trading Economics, 2019. Pakistan average temperature [Online]. Available: https://tradingeconomics.com/pakistan/temperature. (Accessed 25 April 2020). Accessed.
Ullah, K., 2013. Electricity infrastructure in Pakistan: an overview. Int. J. Energy Inform. Commun. 4 (3), 11–26.
Yip, S.-C., Tan, C., Tan, W.-N., Gan, M.-T., Wong, K., Phan, R.C.-W., 2018. Detection of energy theft and metering defects in advanced metering infrastructure using analytics. In: International Conference on Smart Grid and Clean Energy Technologies. ICSGCE), 2018.
Yu, N., Shah, S., Johnson, R., Sherick, R., Hong, M., Loparo, K., 2015. Big Data Analytics in Power Distribution Systems. IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2015.
Yurtseven, Ç., 2015. The causes of electricity theft: an econometric analysis of the case of Turkey. Util. Pol. 37, 70–78.
Zehng, K., Chen, Q., Wang, Y., Kang, C., Xia, Q., 2019. A novel combined data-driven approach for electricity theft detection. IEEE Trans. Industrial Informatics 15 (3), 1809–1819.
Zhang, Q., Zhang, M., Chen, T., Fan, J., Yang, Z., Li, G., 2018. Electricity theft detection using generative models, 2018. In: IEEE 30th International Conference on Tools for Artificial Intelligence. ICTAI).
Zheng, Z., Yang, Y., Niu, X., Dai, H.-N., Zhou, Y., 2018. Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids. IEEE Trans. Industrial Informatics 14 (4), 1606–1615.
Zhou, K., Fu, C., Yang, S., 2016. Big data driven smart energy management: from big data to big insights. Renew. Sustain. Energy Rev. 56, 215–225.
Zorn, B.. Strategies to consider when implementing AMI [Online]. Available: https://electricenergyonline.com/energy/magazine/260/article/Strategies-to-Consider-When-Implementing-AMI.htm.