

Proposta de Análise a Potenciais Perdas Econômicas Relativas à Gatos Elétricos

Daniel Carlos Souza Santos
Instituto de Física (IFUSP)

July 2024

Sumário

1	Introdução	1
2	Impactos da Perda de Energia na Sociedade	1
3	Metodologia	2
4	Algoritmo	2
4.1	Considerações iniciais	2
4.2	Código	3
4.3	Padrões de Consumo	4
4.4	Fitting de Dados	5
5	Resultados	5
6	Conclusão	7
7	Referências	7

1 Introdução

Esse artigo visa elaborar uma metodologia simplificada para a identificação de roubos de eletricidade em redes elétricas. A metodologia consiste em um programa em *Python* que ao ser colocado para rodar na rede retornará dados relevantes sobre as perdas de energia na rede. Além disso, o programa irá aprender o padrão de consumo em cada medidor, assim gerando um indicador de quão provável é a presença de roubo de energia em um medidor, o que facilita sua identificação pelas autoridades competentes. Também apresentaremos uma discussão de como essas perdas impactam a sociedade, propondo soluções alternativas ao corte dessa fonte de suprimentos.

2 Impactos da Perda de Energia na Sociedade

O mercado de energia hoje no Brasil é extremamente regulado. Empresas tanto de transmissão quanto de distribuição nesse setor operam concessões fixas nas quais são obrigadas a manter ou ofertar um dado serviço, sendo restituídas pelo estado por um valor previamente

fixado. No caso específico das empresas de distribuição elas além de operarem esse tipo de contrato para o mercado livre de energia, operam também concessões de exclusividade na distribuição sobre subconjuntos do mercado cativo. Então, uma vez que o cenário impõe a empresa altos e necessários custos fixos, ele também a leva a minimizar perdas técnicas e não técnicas uma vez que sua receita está previamente fixada. Ao que se introduzir um roubo de energia na rede, sendo a empresa é incapaz de localiza-lo, ela para manter fixa sua receita rateia o prejuízo referente ao roubo para todos os participantes da rede oque prejudica esses participantes, os estimulando a reduzir o consumo de energia e por consequência prejudicando a empresa e estimulando novos roubos. Esse ciclo representa uma real ameaça à saúde financeira da empresa, tendo em casos como o da Light [1] já levado distribuidoras a falir e assim favorecendo a precarização dos serviços destas. Disso é possível perceber o quão desfavorável é o roubo de energia para a sociedade, oque motiva o estudos de formas de minimizar sua ocorrência.

Inclusive, vale salientar que os impactos dessas práticas não são nada desprezíveis ao em relação aos seus efeitos na economia dos países. Segundo a referência [5] de 13 a 14/100 do consumo de eletricidade no Paquistão é sustentado por práticas de roubo elétrico. Além disso, segundo [6] o Brasil chega a ter perdas anuais de até 3 Bilhões de Dolares em sua economia devido a essas práticas.

3 Metodologia

A metodologia de estudo proposta nesse trabalho consiste no desenvolvimento de um algoritmo que ao rodar sobre uma base de dados contendo informações referentes a leitura no tempo, localização, e temperatura ambiente dos medidores possa apontar para a possibilidade de haver ou não roubo de energia nesse medidor. Para realizar esse feito o algoritmo compara a curva de consumo presente com a curva de consumo padrão desse medidor, compara também o padrão de consumo com o esperado pelas condições climáticas locais, levando em conta também as perdas técnicas sob o fluxo elétrico através da rede na região do medidor. Por fim faremos uma análise social a respeito dos impactos da implementação do algoritmo em um cenário real, ponderando seus benefícios e malefícios para a sociedade.

4 Algoritmo

4.1 Considerações iniciais

Os maiores problemas presentes na identificação de roubo de energia estão na ausência de medidas em tempo real das leituras dos medidores. Conjuntos dessa natureza podem hoje ser gerados e atualizados por medidores modernos, mas tendo em vista que esses são minoria no total hoje presente conclui-se que esse fator importa para a detecção de roubo de energia. Outro fator relevante são as dificuldades geográficas, pois alguns dos medidores estando em regiões de difícil acesso, ou em locais nos quais a presença da equipe responsável pela coleta dos dados é mal vista podendo esses fatores colocarem em risco a equipe durante a coleta. Para alimentar o código dados foram gerados como uma tendência linear acrescida de resíduos aleatórios para representar o crescimento do consumo total ao longo do tempo até a quitação da respectiva conta de luz evento esse no qual consideramos para fins de simplicidade zerado o consumo, sendo indicado na base de dados como um fator que representa a fronteira entre o consumo presente e o consumo passado.

Para construir o algoritmo vamos considerar um cenário ideal no qual todos os medidores atualizam sincronizadamente a base de dados, e que suas atualizações representam leituras totalmente precisas. Também levamos em conta que as dificuldades geográficas são irrelevantes.

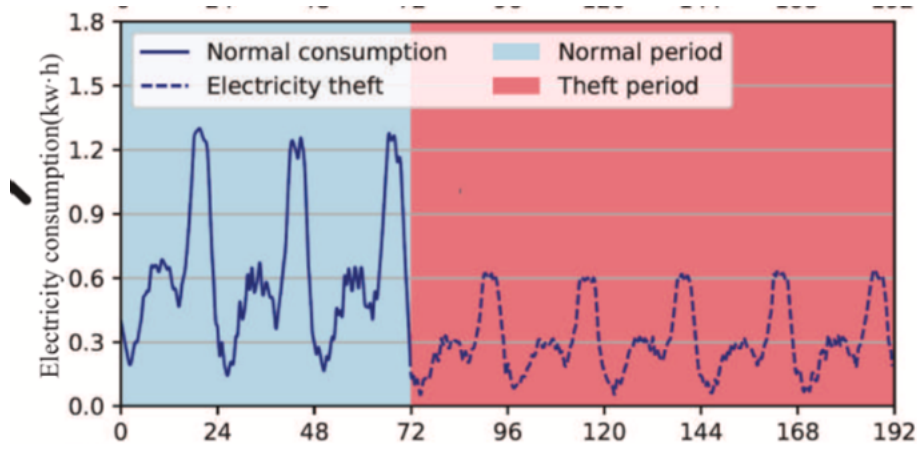


Figura 1: Ilustração do modus operandi de parte do algoritmo [8]

tes, com nosso cenário tendo seus padrões de consumo apenas influenciados pela temperatura ambiente.

Para conseguir uma abordagem efetiva será tomado o seguinte caminho. Inicialmente será feita uma análise por medidor na base de dados, nessa análise inicial o programa irá quebrar o dataset de 1 medidor nos 4 períodos do ano, e então irá aprender os padrões de consumo sazonais nesse medidor. Uma vez aprendidos esses padrões o modelo regressor irá estimar o quanto padrões presentes e passados relativamente próximos divergem dos padrões aprendidos e com isso irá ter uma noção do quão possível é de se ter um gato nesse medidor. Essa abordagem é pensada para conseguir identificar roubos de energia através da forma característica que eles revelam no gráfico de consumo por tempo. Essa abordagem pode ser ilustrada pela figura 1 acima a qual foi confeccionada no artigo da referência [8].

A outra parte do código deve por meio da análise do fluxo de potência obter uma métrica do quão provável é de se ter roubo de energia em uma região. Segundo a referência [2] ao se calcular as perdas técnicas e se comparar o fluxo de entrada - fluxo de saída - perdas técnicas com a diferença real observada, as distribuidoras de energia conseguem ter uma noção das perdas não técnicas em uma dada região. Caso exista uma grande discrepância nesses valores, a empresa consegue induzir a existência de um gato elétrico, sendo essa a maneira hoje usual da detecção desse tipo de fraude elétrica. A segunda parte do algoritmo visa automatizar isso, usando os dados em tempo real do fluxo em uma região da rede, essa parte calcula a diferença no tempo o que alimenta uma métrica relativa à possibilidade de roubo de energia na região dada. Fazendo isso recursivamente encontramos uma medida do quão provável é ali estar ocorrendo roubo de energia. Um estudo mais minucioso sobre como considerar perdas não técnicas está presente na referência [9]. E uma abordagem mais ampla do tema é coberta na referência [10].

4.2 Código

A entrada é uma base de dados com informações de leitura no tempo, localização, e temperatura ambiente dos medidores. Após receber a base de dados ele cria um objeto referente a cada medidor e plota esses medidores com suas leituras nas respectivas localizações. Após isso é considerado para cada medidor o conjunto das leituras registradas o qual é quebrado em um conjunto de leituras antigas e presentes. Feito isso são plotadas curvas de energia por tempo sendo elas comparadas uma a outra e no caso da presença valores substancialmente menores na curva presente com relação a antiga, é associado um índice real a essa diferença.

As temperaturas são tratadas de maneira similar. Os dados de consumo e temperatura no tempo são comparados em termos de formato das curvas, uma vez que temperaturas extremas servem como um incentivo ao consumo elétrico. Para fazer essa comparação nos baseamos no

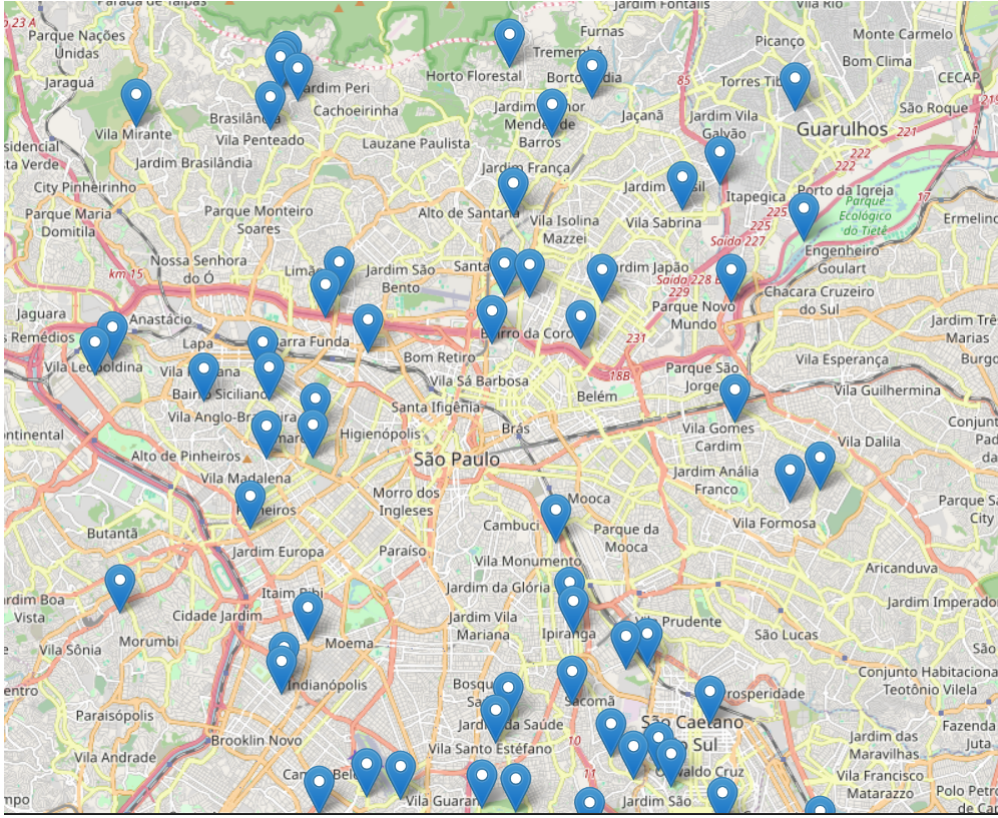


Figura 2: Plot dos Medidores

estudo feito na referência [2]. O resultado dessa comparação também nos dá um número real.

Visando considerar o quanto de energia chega nesse medidor, e se essa energia está ou não sendo consumida será plotada toda a rede elétrica e seus ramos nos respectivos medidores. A esses a energia dada de entrada será contabilizada juntamente com o excedente, tendo essa diferença que representar o acréscimo na leitura. Caso isso não aconteça a diferença será indexada em um número real.

Esses pontos serão plotados no espaço abstrato dos dados sendo esses classificados pelo algoritmo KNN, o qual usará um dataset de dados nos quais já foram inseridos marcadores com roubos de energia. O dataset será quebrado em 0,2 dos dados para treino e 0,8 dos dados para teste. O KNN irá usar as distâncias nesse espaço abstrato para classificar os marcadores como com ou sem roubo.

4.3 Padrões de Consumo

O conhecimento do comportamento de padrões de consumo ocorridos em um medidor é essencial para que se consiga ponderar a existência ou não de roubo de energia nesse local. Nos dias de hoje as pessoas que realizam roubo de energia estão extremamente munidas de ferramentas que viabilizam mascarar esse roubo. Entretanto todas essas formas acabam por alterar permanentemente o padrão de consumo sazonalmente. É importante ressaltar que os padrões de consumo aqui considerados são os padrões de consumo sazonais, os quais são obtidos ao se fixar uma época do ano. Acontece que épocas do ano similares levam a condições climáticas similares, as quais induzem nas pessoas uma forma característica de consumir eletricidade. Isso pode ser ilustrado pela figura 3 acima, em que foi fixo um medidor no dataset, referência [3], e foi acompanhado o padrão de consumo ao longo do ano para um dos medidores do conjunto. Nesse gráfico também foi colocada a média móvel de período 5 para ajudar a visualizar o padrão presente no gráfico.

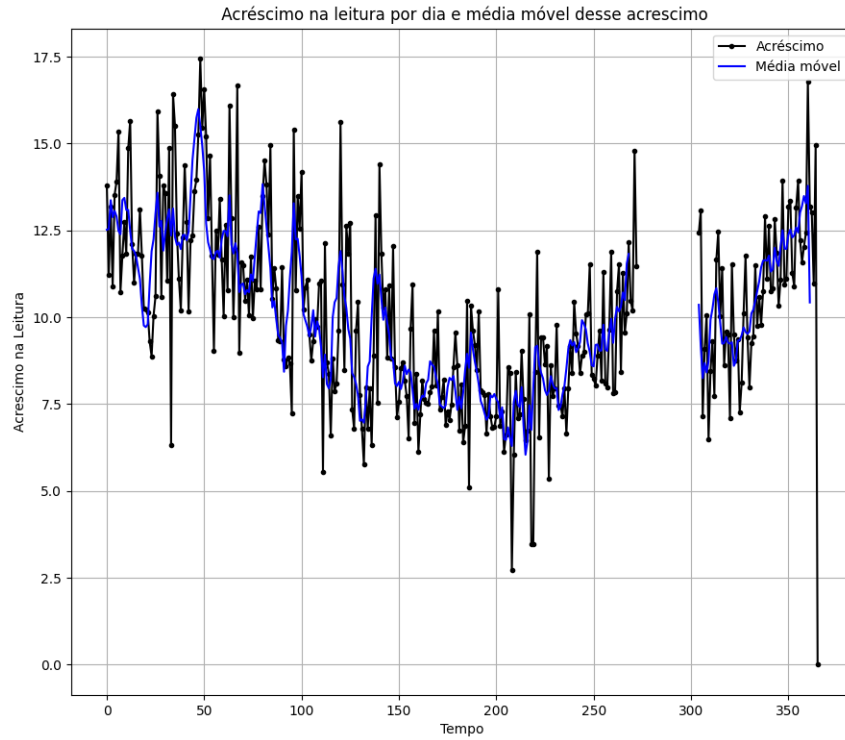


Figura 3: Acréscimos no consumo ao longo do tempo para o marcador 9956 em SI

O padrão ilustra bem o consumo de eletricidade em um país na região zona temperada ao norte, é notável perceber que as baixas temperaturas do inverno induzem a um consumo elétrico maior e mais regular, o que pode ser observado nas bordas do gráfico, enquanto na primavera esse consumo vai lentamente diminuindo, até registrar seu patamar mínimo no verão onde as temperaturas amenas desmotivam o consumo de energia elétrica.

4.4 Fitting de Dados

A não disposição de dados pre-classificados inviabilizou a construção de um classificador KNN para identificar roubos de energia. Além disso, pela impossibilidade de obter dados que contenham informações precisas de temperatura e localização, foi-se escolhida a metodologia de trabalho seguinte.

A metodologia tomada para essa parte do trabalho consistiu na construção de um regressor linear o qual visava aprender os padrões sazonais dos acréscimos e depois treinar nos respectivos dados restantes. O programa usado para isso está na referência [4]. Após isso o modelo foi treinado com o dataset presente na referência [3] separado em frações de 70/30, e então aplicado aos 30/100 de dados restantes.

5 Resultados

Os resultados obtidos pelo modelo são mostrados na figura 3 para o caso do inverno. Com suas respectivas previsões para os dados de teste sendo mostrados em seguida.

Visualmente olhando para as saídas do modelo feito [4] e as comparando a com os dados de teste podemos observar que eles estão relativamente próximos. O que se pode confirmar olhando para as MAEs que constam no intervalo (1.46 e 1.617), sendo esse um indicativo de que as saídas do modelo estão relativamente próximas dos dados ajustados. Quanto a métrica R2, um modelo em que R2 é próximo de 1 está próximo do ideal. Já um modelo de R2 próximo

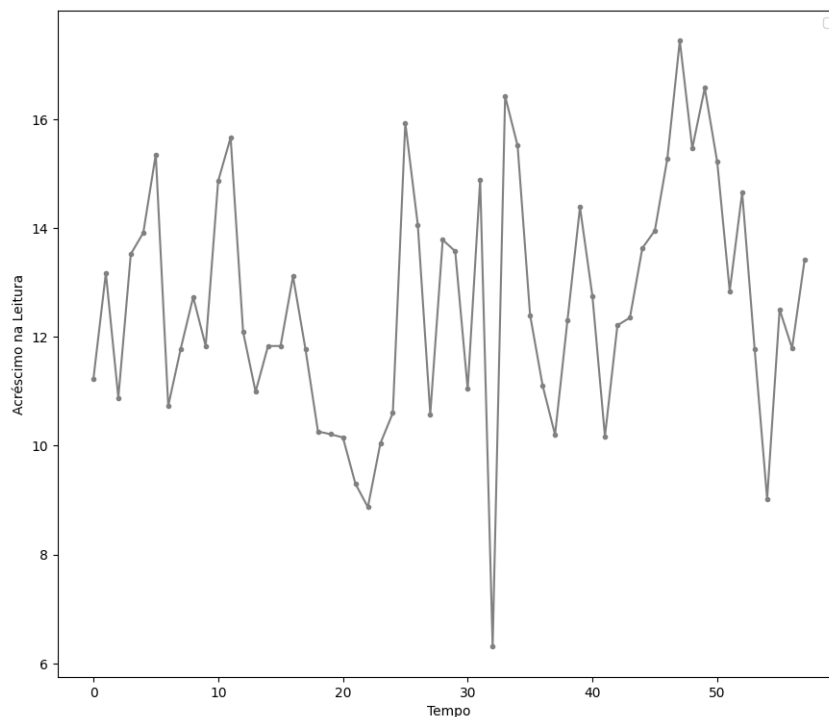


Figura 4: Consumo no inverno para o marcador 9956

de 0 é um mal modelo e um em que R^2 é menor que zero é ainda pior. As métricas R^2 advindas do modelo são mais próximas de 0 do que 1. Além disso R^2 para o modelo de inverno é menor que 0, logo esse é um péssimo modelo para o sistema. Então para concluir exatamente sobre a qualidade dos modelos é necessário o uso de mais métricas como o STD por exemplo, já que MAE e R^2 divergem bastante em suas conclusões em relação aos modelos apresentados.

	Valores Teste(inverno)	Valores Preditos
0	13.91	11.787278872232879
1	10.04	12.491386035916435
2	11.77	12.269036405279522
3	12.50	13.677250732646634
4	10.20	13.010201840735897
5	12.09	12.083745046415428
6	12.84	13.52901764555536
7	15.35	11.824337144005696
8	10.60	12.528444307689252
9	11.11	12.973143568963078
10	12.35	13.232551471372808
11	10.88	11.71316232868724
12	15.46	13.417842830236903
13	10.26	12.30609467705234
14	12.73	11.935511959324153
15	13.17	11.676104056914422
16	13.57	12.713735666553347
17	11.77	13.603134189100997

- **MAE:** 1.5474926459620408
- **R^2 :** -0.2555772900106752

6 Conclusão

Este artigo propôs uma metodologia simplificada para identificar e combater o roubo de eletricidade em redes elétricas. Utilizando um programa em Python, desenvolvemos um algoritmo que analisa dados de consumo de energia, padrões climáticos locais e fluxo elétrico na rede para detectar anomalias que possam indicar roubo de energia.

Os impactos sociais e econômicos do roubo de energia foram discutidos detalhadamente. No Brasil, as perdas associadas a práticas ilegais de energia podem chegar a bilhões de dólares anualmente, afetando negativamente não apenas as distribuidoras, mas também a economia como um todo. Essas perdas impactam a sustentabilidade financeira das empresas, levando até mesmo à falência de algumas, como observado em casos notáveis no setor.

A metodologia proposta no estudo envolveu a criação de um algoritmo que aprende padrões de consumo sazonais e utiliza modelos de regressão para identificar desvios significativos que possam indicar roubo de energia. A análise considerou tanto os dados históricos de consumo quanto fatores externos como temperatura e fluxo de potência na rede.

Os resultados do modelo mostraram uma boa aproximação com os dados de teste, embora métricas como MAE (Erro Médio Absoluto) e R² (Coeficiente de Determinação) indicassem que ajustes adicionais são necessários para melhorar a precisão do modelo, especialmente em diferentes estações do ano.

Em resumo, este estudo destaca a importância de métodos avançados de detecção de roubo de energia para mitigar seus impactos negativos na sociedade e na economia. A implementação de algoritmos eficazes pode não apenas reduzir as perdas econômicas, mas também promover a sustentabilidade financeira das empresas de distribuição de energia, contribuindo para um fornecimento mais estável e confiável de eletricidade para todos os consumidores.

7 Referências

Referências

- [1] Suno. Light (LIGT3): Entenda a recuperação judicial. Disponível em: <https://www.suno.com.br/noticias/light-ligt3-entenda-recuperacao-judicial/>
- [2] MDPI. Sustainability — Free Full-Text — A Comparative Study on Machine Learning Techniques for Predicting Energy Consumption in Buildings. Disponível em: <https://www.mdpi.com/1996-1073/15/23/8794>
- [3] Kaggle. Theft Data. Disponível em: <https://www.kaggle.com/datasets/avinemmatty/theft-data/data>
- [4] GitHub. REnergy. Disponível em: <https://github.com/DanielDCSS/REnergy>
- [5] ScienceDirect. Consumer's Perception Towards Electricity Theft: A Case Study of Islamabad and Rawalpindi using a Path Analysis. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0301421522004098>
- [6] Towards Efficient Energy Utilization Using Big Data Analytics in Smart Cities for Electricity Theft Detection. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2214579621001027>
- [7] Systematic Review of Energy Theft Practices and Autonomous Detection Through Artificial Intelligence Methods. Disponível em: <https://www.sciencedirect.com/science/article/pii/S136403212300401X>

- [8] Machine Learning-Based Detection Framework against intermittent electricity theft attack. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0142061523001321>
- [9] Northeast Group LLC. Electricity theft & non-technical losses: Quantifying the non-technical loss problem. Disponível em: <https://northeast-group.com/research-catalogue/>, accessed October 2021 (2021).
- [10] Xia X, Xiao Y, Liang W, Cui J. Detection methods in smart meters for electricity thefts: A survey. Proceedings of the IEEE 2022;110(2):273-319.