

Análise de consistência de um algoritmo regulador de semáforos inteligentes para mobilidade urbana

Consistency analysis of a smart traffic light regulatory algorithm for urban mobility

A. L. F. da Silva, J. N. A. Silva, J. L. S. Feitoza, L. A. C. de Lima, S. S. E. A. Pereira

*Centro de Informática(CIn) - Universidade Federal de Pernambuco(UFPE)
Av. Jorn. Aníbal Fernandes, s/n - Cidade Universitária, Recife - PE, Brasil*

alfs2@cin.ufpe.br
jnas2@cin.ufpe.br
jlfs@cin.ufpe.br
lacl@cin.ufpe.br
sseap@cin.ufpe.br

Resumo— Esse artigo tem como objetivo analisar os resultados colhidos do tempo de execução de um algoritmo ganancioso implementado em semáforos inteligentes, obtendo, dessa forma, como principais conclusões que o comportamento do algoritmo é consistente e a amostra de telemetria segue uma distribuição normal.

Palavras-chave— Mobilidade urbana, semáforos inteligentes, tecnologia, distribuição normal, algoritmo ganancioso.

Abstract— This article aims to analyze the results gathered from the runtime of a greedy algorithm implemented on intelligent traffic lights, thereby deriving as main conclusions that the algorithm's behavior is consistent and the telemetry sample follows a normal distribution.

Keywords— Urban mobility, smart traffic lights, technology, normal distribution, greedy algorithm.

I. INTRODUÇÃO

Durante a disciplina de análise das organizações, foi proposto uma atividade para abordar os impactos sociais das organizações, a criação de um business model canvas (quadro de modelo de negócios) que impactam diretamente a sociedade. Durante a atividade, um grupo composto por integrantes do curso de sistemas de informação, projetaram uma empresa de mobilidade urbana que tinha como base usar tecnologia para gerar soluções, sendo sua primeira proposta, semáforos

inteligentes. Essa ideia que surgiu em sala de aula, agora visa concorrer na modalidade de pré-incubação da PoloTec UFPE e atrair possíveis investidores, para isso, foi necessário criar um algoritmo que tomasse decisões sobre quando abrir ou fechar o semáforo com base em informações como o fluxo de veículos.

Cumprе salientar que a ideia do grupo tem como escopo um dos problemas centrais das grandes centros, a mobilidade urbana, em especial, o tráfego nas vias públicas, visto que, conforme a pesquisa de mobilidade urbana de 2022, é o maior responsável pela locomoção de pessoas pelo território, bem como o gasto médio de tempo em engarrafamentos girou em torno de 1h por dia.

Sendo assim, a ideia dos semáforos inteligentes, tem o condão de agilizar o fluxo dos veículos a partir do uso de um algoritmo greedy que otimiza a tomada de decisão de trocar o estado dos semáforos de acordo com os dados do tamanho da fila de veículos, bem como de outros dados sobre o tráfego da região.

Esse relatório tem como objetivo analisar a eficiência do semáforo inteligente a partir da análise dos dados coletados. Nesse sentido, o algoritmo foi executado 1000 vezes e foi aferido o tempo de

resposta de cada execução e essa será a amostra a ser analisada.

Quanto à organização, na 2ª Seção do presente relatório é apresentada a metodologia empregada na análise, o funcionamento do algoritmo, a realização da coleta de dados, as bibliotecas usadas para análise e as condições do teste de hipótese. A análise descritiva dos dados e a visualização por meio de gráficos estão na Seção 3. A Seção 4 traz as conclusões do relatório. Por fim, a Seção 5 apresenta os dados e o repositório criado para a análise do algoritmo.

II. METODOLOGIA

A. Funcionamento do algoritmo

O objetivo do presente trabalho é analisar a consistência da implementação de um algoritmo greedy utilizado para otimizar o processo decisório de semáforos inteligentes. Nesse sentido, é importante para fins metodológicos descrever o funcionamento do algoritmo analisado, essa seção, terá essa finalidade.

Os greedy algorithms (algoritmos gananciosos) tem como característica principal o seu funcionamento que faz escolhas locais, visando obter soluções globais ótimas. Nesse sentido, o algoritmo começa com uma solução inicial vazia e em cada passo o algoritmo faz uma escolha selecionando a opção que parece ser a melhor naquela situação em específico, essa opção passa a ocupar o local da solução inicial, tal procedimento se repetiria com o fito de ir tomando decisões ótimas com base nos dados do tráfego de veículos em uma área geográfica específica.

Cumprе salientar que por ser um protótipo, ele faz a análise de dados do fluxo de veículos em determinada via, no intervalo de hora informado e no intervalo de velocidade definido naquele mês, usando a database disponibilizada no website Dados Recife, pela Autarquia de Trânsito e Transporte Urbano do Recife (CTTU). Dessa forma, foram obtidos os dados de telemetria do programa criado pelos alunos, os quais são objetos da análise deste relatório.

B. Coleta de dados

Durante os testes de execução do semáforo, foi realizada a observação do tempo (em segundos) do desempenho do algoritmo, implementado de maneira alternada em semáforos da Av. Caxangá, Av. Agamenon Magalhães, e Av. Norte, as três com trânsito mais intenso da cidade segundo CTTU, e em suas proximidades. Ao final, foram coletadas 1000 respostas como dados de observação. Esses testes foram conduzidos na IDE Visual Studio Code (VSCode).

C. Análise de dados

A análise do projeto foi feita utilizando a linguagem de programação Python. Foram utilizadas as bibliotecas pandas, para a leitura da database, Matplotlib para a plotagem dos gráficos, Numpy e Scipy para medidas de dispersão e testes estatísticos.

D. Teste de hipótese I - Distribuição normal

Para avaliar a normalidade dos dados, foram empregadas duas abordagens distintas. A primeira delas envolveu o uso da função 'scipy.stats.normaltest', que combina medidas baseadas na curtose e na assimetria da distribuição para calcular uma estatística de teste. Esta técnica permite avaliar se a amostra é proveniente de uma distribuição normal, assumindo como hipótese nula (H_0) que ela segue essa distribuição, e hipótese alternativa (H_a) que ela não segue a distribuição normal. Além disso, foi utilizada a função 'scipy.stats.probplot', que gera um gráfico de probabilidade normal, conhecido como Q-Q plot. Nesse gráfico, se os pontos de dados formam uma linha reta próxima à linha de 45 graus, isso indica concordância com uma distribuição normal.

E. Teste de hipótese II - Aplicabilidade do algoritmo

Para avaliar a aplicabilidade do algoritmo, duas métricas foram analisadas: A consistência, observada através do desvio padrão e a eficiência, observada através do tempo médio de execução. As condições para o projeto ser considerado aplicável na prática era atingirem um tempo médio igual ou menor a 1.2 segundos e um desvio padrão inferior a

0.5 segundos. Assim, nossa hipótese nula(H_0) é que o algoritmo é eficiente, e nossa hipótese alternativa(H_a) é que ele é ineficiente.

III. RESULTADOS

Neste relatório, apresentamos as principais medidas estatísticas de um conjunto de dados. Os resultados destacam a tendência central, a dispersão e outras características relevantes da distribuição dos dados.

F. Análise das medidas de centralidade e de dispersão.

Por se tratar de representações de tempo, os dados analisados são quantitativos contínuos. O menor valor registrado foi 0,17379 segundos e o maior 2,28593 segundos.

TABELA I. MEDIDAS DE CENTRALIDADE E DISPERSÃO

Medida	Valor
Média	1.1309
Mediana	1.11973
Moda	0.78624
Variância	0.10068
Desvio Padrão	0.3173
1° Quartil	0.919607
2° Quartil	1.119730
3° Quartil	1.339840
4° Quartil	2.285930
Curtose	0.04833

Uma curtose de aproximadamente 0.04833 sugere que a distribuição dos tempos de resposta do seu algoritmo é relativamente próxima de uma distribuição normal, que é a distribuição comumente considerada como tendo uma curtose de 0. Se a curtose fosse positiva, indicaria que a distribuição dos dados é mais concentrada em torno da média e possui caudas mais pesadas do que uma distribuição normal (distribuição leptocúrtica). Se fosse negativa,

indicaria que a distribuição é mais espalhada e tem caudas mais leves do que uma distribuição normal (distribuição platicúrtica).

A partir das medidas foram elaborados os gráficos de boxplot (Figura 1), histograma (Figura 2) e gráfico de linha (Figura 3).

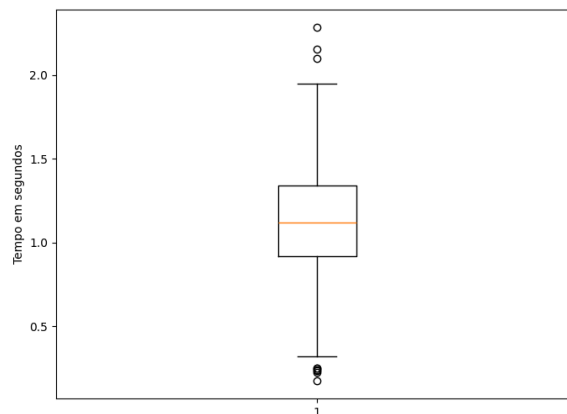


Fig. 1 Boxplot. Foram observados 3 outliers superiores e 6 outliers inferiores dentro dos 1000 valores analisados.

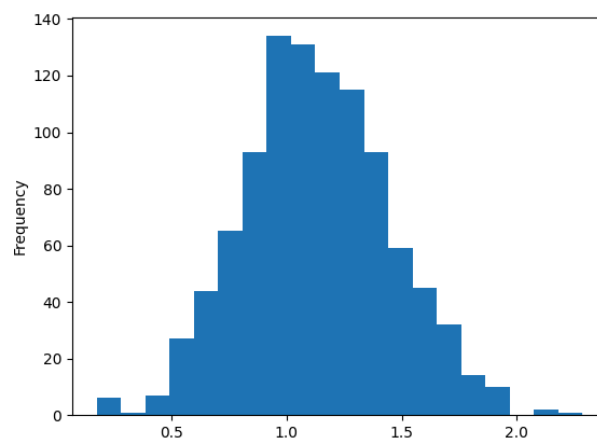


Fig. 2 Histograma (frequência x segundos) .

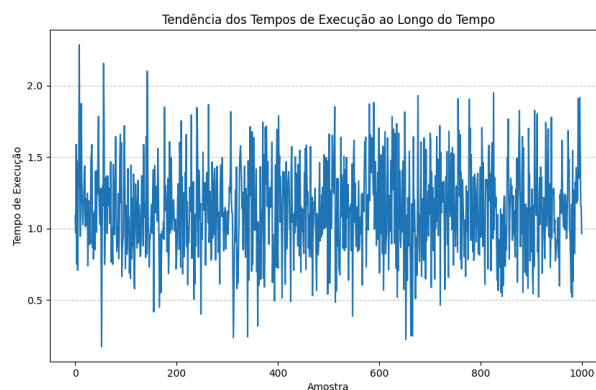


Fig. 3 Gráfico de Linha .

A média do tempo de resposta sugere que, em geral, o algoritmo leva cerca de 1,13 segundos para tomar uma decisão sobre abrir ou fechar o semáforo. A interpretação do gráfico de boxplot indica que, levando em consideração que a mediana é muito próxima da média e pela dispersão no gráfico, não há grande assimetria dos dados. Isso sugere que a maioria dos tempos de resposta são em torno desse valor. É possível notar a presença de poucos outliers, o que sugere que a maioria dos valores é específica dentro do intervalo interquartil, com apenas algumas observações atípicas ou extremas, podendo significar flutuações no tráfego, incluindo acidentes ou bloqueios, isso indica que a maioria das decisões do semáforo é tomada dentro de um intervalo previsível de tempo.

A análise do histograma revela indícios de uma distribuição normal, notáveis pela curva centralizada em torno da média e pela dispersão simétrica dos dados ao redor dela. Além disso, a curtose, que indica o "pico" ou "achatamento" da distribuição, tende a se aproximar de zero em distribuições normais, refletindo uma forma moderadamente aguçada. Para validar essa hipótese, empregamos um teste de normalidade, detalhado no Teste de Hipótese I.

G. Teste de Hipótese I - Distribuição normal

A normalidade dos dados foi verificada utilizando as funções 'scipy.stats.normaltest', para calcular uma estatística de teste, com a hipótese nula de que a amostra é proveniente de uma distribuição normal, e 'scipy.stats.probplot' que gera um gráfico de probabilidade normal (Q-Q plot) (Figura 3).

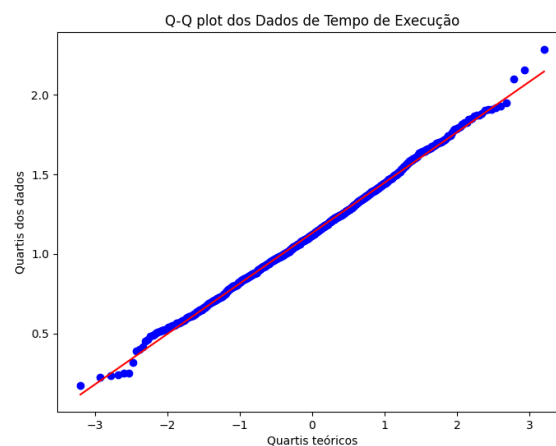


Fig. 3 Gráfico Q-Q

Os pontos em um gráfico de Q-Q plot devem formar uma linha aproximadamente reta, que coincide com a linha de 45°, ou seja, a linha em que os quartis teóricos são iguais aos quantis da amostra. Se os pontos estão todos próximos ou acima dessa linha, isso indica que os quartis observados na amostra estão em concordância com os quartis que seriam esperados em uma distribuição normal. Portanto, ao interpretar um Q-Q plot, se os pontos estiverem principalmente ao longo da linha de 45°, sem desvios significativos, como acontece na Figura 4, isso confirma que os dados estão distribuídos de forma normal.

Além do mais, foi calculada a 'statistic', a estatística do teste de normalidade, onde quanto mais próximo de 1, mais próximo os dados estão de uma distribuição normal, e 'p-value', o valor associado a essa estatística que indica a probabilidade de observar os dados ou algo mais extremo, se a hipótese nula for verdadeira.

O resultado do teste de normalidade revelou que a estatística de teste calculada foi de 2.8628, enquanto o valor-p associado a essa estatística foi de 0.2389. Esses valores indicam que não há evidências suficientes para rejeitar a hipótese nula de que os dados são provenientes de uma distribuição normal. Com um valor-p superior ao nível de significância comum de 0.05, podemos inferir que os dados possivelmente seguem uma distribuição normal, visto que não há evidências suficientes para rejeitar a hipótese nula (H_0). Portanto, baseado neste teste, não há justificativa

estatística para questionar a normalidade dos dados testados.

em: <https://iopscience.iop.org/article/10.1088/1742-6596/974/1/012013/pdf>. Acesso em: 20/04/2024.

H. Teste de Hipótese II - Aplicabilidade do algoritmo

Através das medições, observou-se uma média de aproximadamente 1.1309 segundos, atingindo a exigência de eficiência ($\mu \leq 1.2s$), e um desvio padrão de aproximadamente 0.3173 segundos, atingindo a exigência de consistência ($\sigma \leq 0.5s$). Assim, não há motivos para rejeitar a hipótese nula (H_0): O algoritmo é aplicável.

IV. CONCLUSÃO

Portanto, é possível concluir que o algoritmo regulador de semáforos inteligentes é aplicável de maneira consistente e eficiente, dentro dos padrões da Prefeitura de Recife, visto que atingiu uma distribuição próxima da normal e valores de média e desvio padrão dentro do limite desejado.

V. ANEXOS

Todo o material do projeto pode ser acessado pelo link a seguir:

https://github.com/adnalsia/projeto_estati

RECONHECIMENTOS

O Centro de Informática da UFPE deseja agradecer aos alunos Luiz Miguel, José Leandro e Talisson Mendes pelo desenvolvimento e concessão do algoritmo para estudo e pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Pesquisa de Mobilidade Urbana de 2022. Disponível em: <https://static.poder360.com.br/2022/06/An-lise-Mobilidade-Urbana.pdf>. Acesso em: 20/03/2024.
- [2] Semáforo inteligente melhora o trânsito em até 30%. Disponível em: <https://exame.com/brasil/semaforo-inteligente-melhora-transito-em-ate-30/>. Acesso em: 20/03/2024.
- [3] VECCHIA, Alessandro Dalla. Semáforos adaptativos, uma abordagem baseada em inteligência de enxames para a mobilidade urbana. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/86282/000909818.pdf?sequence=1>. Acesso em: 20/03/2024.
- [4] SISWIPRAPTINI, Puji Catur et al. Reducing a congestion with introduce the greedy algorithm on traffic light control. 2018, Journal of Physics: Conference Series 974 012013. Disponível