

DESCOBERTA DE ROTAS SEGURAS ATRAVÉS DE METAHEURÍSTICAS

LUIS O. FACONI¹, VINICIUS R. ZOQUETI², CRISTIANE Y. IMAMURA³

¹ Graduando em Engenharia de Software, CENTRO UNIVERSITÁRIO DAS FACULDADES ASSOCIADAS DE ENSINO - UNIFAE, Campus São João da Boa Vista, luisfaconi@gmail.com

² Graduando em Engenharia de Software, CENTRO UNIVERSITÁRIO DAS FACULDADES ASSOCIADAS DE ENSINO - UNIFAE, Campus São João da Boa Vista, vinicius.zoqueti@gmail.com

³ Mestre em Ciências pela USP -São Carlos. Professora dos cursos de Engenharia de Computação e Engenharia de Software, CENTRO UNIVERSITÁRIO DAS FACULDADES ASSOCIADAS DE ENSINO - UNIFAE, Campus São João da Boa Vista, cristiane@fae.br.

Área de conhecimento (Tabela CNPq): 1.03.01.01-1 Computabilidade e Modelos de Computação

Apresentado no

10º Congresso de Inovação, Ciência e Tecnologia do IFSP ou no 4º Congresso de Pós-Graduação do IFSP

27 e 28 de novembro de 2019- Sorocaba-SP, Brasil

RESUMO: O problema do Caixeiro Viajante (PCV) é um problema de otimização de rotas com diversas variações, que envolve uma série de soluções dependendo dos aspectos a serem considerados. O presente trabalho visa fornecer o cálculo de trajetos por vias seguras, segundo informações disponíveis sobre os acidentes que levaram a óbito em vias do estado de São Paulo combinadas com a determinação de rotas de menor distância. Para isso, envolve a resolução de uma variante do PCV utilizando a mineração de dados e algoritmos genéticos, além dos aspectos da teoria dos grafos. Assim, são descritas as técnicas utilizadas e os resultados que apontam a colaboração para a construção de um software que prediz os deslocamentos necessários que possam ser considerados mais seguros no cálculo de uma determinada rota.

PALAVRAS-CHAVE: problema do caixeiro viajante; algoritmos genéticos; mineração de dados; planejamento de rotas.

THE DISCOVERY OF SAFE ROUTES USING METAHEURISTICS

ABSTRACT: The Traveling Salesman Problem (TSP) is an optimization problem that involves a number of solutions according many aspects to be considered. In this work, we present techniques that allow planning safe routes based on information about fatal accidents in the state of São Paulo, Brazil, and determination of the minimal cost. This work has been using data mining and genetic algorithms, as well as aspects of graph theory, to solve a variant of TSP. Thus, we present the results that indicate the collaboration for the construction of a software that predicts safe routes.

KEYWORDS: travelling salesman problem; genetic algorithms; data mining; route planning.

INTRODUÇÃO

Dentro do contexto de rotas se destaca o clássico conhecido como Problema do Caixeiro Viajante (PCV) cujo objetivo é encontrar a menor rota entre um conjunto de pontos que devem ser percorridos, sabendo que todos podem ser visitados uma única vez (CARVALHO, 2017).

De acordo com Buriol (2000) o PCV pertence à classe dos problemas NP-difíceis, ou seja, as soluções já conhecidas deste problema podem não chegar a convergir para a solução na busca de uma rota de acordo com as variáveis a serem consideradas em um tempo factível.

O presente trabalho tem os aspectos do PCV uma vez que objetiva encontrar o deslocamento por vias mais seguras com a menor distância que possa incluir o percurso entre múltiplos pontos, sendo que estes pontos só podem ser acessados uma única vez a partir da origem.

Tendo em vista que as informações necessárias para alcançar o objetivo podem ser coletadas com a ajuda de API's ou de bases de dados públicas relacionadas a trânsito, e que o volume de dados a serem analisados pode inviabilizar técnicas comuns de consultas SQL, foi escolhido realizar a mineração de dados para que uma heurística pudesse ser definida para o planejamento das rotas e, assim, foi possível implementar um software capaz de indicar rotas para percorrer trechos mais seguros de acordo com o índice de acidentes que levaram a óbito nas vias do estado de São Paulo (FACONI, ZOQUETI e IMAMURA, 2019).

Entretanto, na solução obtida, as rotas podem considerar o cálculo de um custo mínimo local e não global. Assim, como o objetivo de melhorar o desempenho da busca da solução de custo mínimo está sendo empregada uma metaheurística, uma vez que muitos trabalhos correlatos apresentam resultados interessantes com uso dessa abordagem, como exemplo Barros e Cunha (2019).

MATERIAL E MÉTODOS

Para o desenvolvimento da ferramenta proposta, foi utilizado inicialmente o algoritmo Dijkstra que visa encontrar a rota mínima entre nós de grafos hamiltonianos, com uso de pesos positivos, uma vez que esse algoritmo é empregado na solução de problemas de cálculo de rotas com custo mínimo. O algoritmo de Dijkstra pode ser observado na Figura 1. Seja o grafo $G = (N, A)$, com o conjunto de nós representado por N e o conjunto de arcos entre esses nós A , a cada passo do algoritmo a partir de um nó i atingido é avaliado qual nó será adicionado na rota, avaliando-se a menor distância dos arcos A ligados ao nó i que ainda não estão adicionados na rota. Após a escolha do menor caminho partindo de um N inicial, será continuada a construção do circuito a partir do novo nó alcançado, até que se esgotem todos os nós. Este processo deve garantir então que cada N escolhido só apareça uma vez na rota e que ela passe ao final por todos os nós do grafo (SILVA, 2017).

Algoritmo Dijkstra;

Início

$P := \emptyset$; $T := N$;
 $d(i) := +\infty$ para cada nó $i \in N$;
 $d(s) := 0$ e $\text{pred}(s) := 0$;
 C : caminho mínimo entre o nó inicial e final;

Enquanto $|P| < n$, faça

Início

Permitir $i \in T$ seja o nó para qual $d(i) = \min \{d(j) : j \in T\}$;
 $P := P \cup \{i\}$;
 $T := T - \{i\}$;

Para o arco $(i,j) \in A(i)$, faça

Se $d(j) > d(i) + C_{ij}$, então $d(j) := d(i) + C_{ij}$ e $\text{pred}(j) := i$;

Se $j = n$, recupere C (a partir dos predecessores, $\text{pred}(j)$);

Fim;

Fim.

FIGURA 1. Algoritmo de Dijkstra. Fonte: Silva (2017).

Para a proposta inicial de indicar rotas mais seguras, no estado de São Paulo, o peso considerado em cada aresta leva em consideração a distância entre os trechos e também uma heurística calculada em

função das informações da base de dados pública do Infosiga que contém 5.460 registros e descreve os óbitos ocorridos nas vias do estado no ano 2018 (INFOSIGA, 2018). Essa heurística se mostrou eficiente para determinar as rotas considerando trecho mais seguros e foi baseada na identificação de Clusters, com uso do algoritmo SimpleKMeans (FACONI, ZOQUETI e IMAMURA, 2019). O SimpleKMeans foi escolhido pela natureza das informações contidas na base e foi utilizado a partir da ferramenta WEKA, por ser um software de código aberto e disponibilizar muitos algoritmos de aprendizado de máquina, utilizados para fazer generalizações ou formulação de uma coleção de dados de amostragem (ASRONI, 2015).

Para definir as distâncias entre os locais escolhidos da rota, onde foi utilizada a API do Bing Maps que se chama Distance Matrix e que fornece as distâncias para um conjunto de origens e destinos (BING MAPS, 2019). Além disso, por ser necessário validar coordenadas presentes na base de dados com as do trajeto em questão, foi necessário armazenar latitudes e longitudes das ligações dos N pontos a serem tratados, e para isso foi utilizada a ferramenta Interactive Maps fornecida pela API do Here Maps (HERE MAPS, 2019).

Tendo em vista essa solução já desenvolvida e os objetivos do presente trabalho que visa solucionar o problema do mínimo local, já que pode não conduzir a uma rota ótima, foi utilizado o algoritmo genético que é uma metaheurística e consiste em um método de busca estocástica baseado na Teoria da Evolução Natural das Espécies de Darwin. Para isso, estão sendo considerados os aspectos empregados no trabalho de Vargas, (2000), no qual se usa uma população inicial P de soluções chamadas de cromossomos e esses cromossomos iniciais vão sendo evoluídos através de diversas operações até chegar a uma solução que melhor atenda o critério objetivado.

Conforme demonstrado na Figura 2, a cada geração os cromossomos são avaliados de acordo com uma função que valida seu nível de aptidão, chamada de função fitness. Os cromossomos que obtiverem o melhor fitness são selecionados para gerarem a próxima geração, através de cruzamentos e mutações. O intuito é que em cada nova geração o conjunto de soluções melhore até que um bom resultado seja atingido (BARROS; CUNHA, 2019 apud GOLDBERG, 1989).

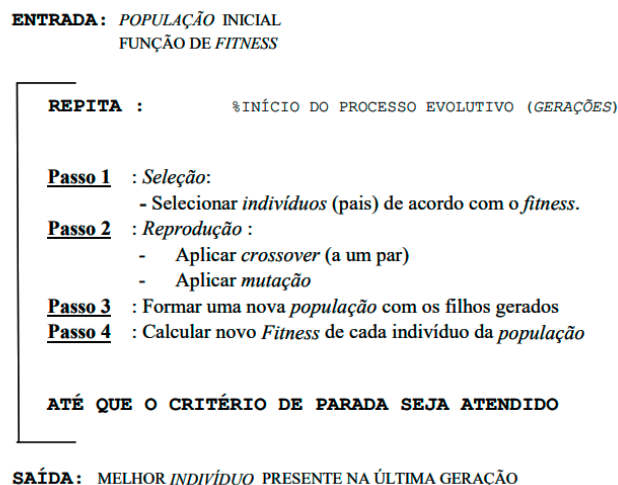


FIGURA 2. Ideia central do algoritmo genético. Fonte: Vargas (2000).

RESULTADOS E DISCUSSÃO

Com base nas técnicas mencionadas, foi factível desenvolver um programa em C# para estimar a rota a ser apontada, percorrendo múltiplos pontos de forma a escolher o trajeto com o menor risco de acidentes e, portanto, mais seguro.

Para auxiliar no entendimento dos resultados adquiridos no presente trabalho com a aplicação da metaheurística, estes são confrontados com os resultados já obtidos sem o uso do algoritmo genético e publicados por Faconi, Zoqueti e Imamura (2019). Para isso, serão usados também os valores da Tabela 1, na qual se observam algumas distâncias, ponto a ponto, entre as cidades de Mogi Mirim (A), Limeira (B), Paulínia (C), Campinas (D). Os valores observados na Tabela1 foram adquiridos através de requisições efetuadas a API do Bing Maps.

TABELA 1. Alocação de custos em um exemplo real de cidades através das suas distâncias.

	A	B	C	D
A	0	52,2	66,1	59
B	53,2	0	47,1	56
C	67,8	47,9	0	23,4
D	58,7	57	21	0

Fonte: Faconi, Zoqueti e Imamura, 2019.

A partir os dados da Tabela1 e no fator de segurança envolvido no peso de cada aresta obtido com uso das informações dos Clusters resultantes da aplicação do SimpleKMeans na base de dados do Infosiga, a ferramenta sem uso da metaheurística fornece a seguinte rota segura a partir de A: $A \rightarrow C \rightarrow D \rightarrow B \rightarrow A$, com o custo total de 320,4. (FACONI, ZOQUETI e IMAMURA ,2019).

Com uso da solução descrita neste trabalho, considerando a distância, os Clusters encontrados na base do Infosiga e a aplicação do algoritmo genético com uma população inicial com 5 indivíduos, sendo um deles o resultado obtido através do algoritmo Dijkstra. O cruzamento que ocasiona uma nova geração é efetuado de forma que ocorra a cópia dos seus genes e ao final o que não foi possível ser copiado é gerado aleatoriamente, para mutação. Com todo o processo efetuado o resultado obtido foi: $A \rightarrow D \rightarrow C \rightarrow B \rightarrow A$ com o custo total de 286,9, sendo encontrado na segunda geração de população dos cromossomos.

Observando os dados da base do Infosiga e dos clusters obtidos, é possível perceber que o trecho de Mogi Mirim a Limeira possui um trecho que as conecta que possui muitos acidentes, embora possua a menor distância. Assim tanto o software que usa a metaheurística como a versão que não usa não optam por escolher esse trecho inicialmente. Entretanto, comparando os custos totais alcançados respectivamente por cada solução, fica evidente a diferença das rotas calculadas.

CONCLUSÕES

De acordo com os resultados apontados neste trabalho é possível validar que a ferramenta desenvolvida consegue calcular um percurso seguro utilizando a distância entre os pontos e identificando ocorrências de acidentes fatais nos trechos a serem trafegados.

É possível notar que o uso do algoritmo de Dijkstra conquistou uma solução aproximada para a problematização, aliado aos dados obtidos com a mineração de dados sobre a base de óbitos em vias do estado de São Paulo ocorridos em 2018.

Embora ainda não se encontre totalmente concluído e sejam necessários mais testes e análises mais criteriosas, foi possível notar com os resultados mostrados que a solução que usa a metaheurística, baseada em algoritmo genético, tende a amenizar o problema evidente do mínimo local, na tentativa de encontrar de fato a rota mais segura e com o menor custo entre os pontos a serem trafegados.

REFERÊNCIAS

ASRONI, R. A. Implementation Method for K-Means Clustering Based Student Value with Weka Interface a Case Study of Department of Information UMM Magelang. Jurnal Ilmiah Semesta Teknika Vol. 18, No. 1, 76-82, Mei 2015. Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, 2015.

BARROS, J.; CUNHA, F. Aplicação de Algoritmos Heurísticos na Otimização de Rotas Comerciais em Tempo Real. In: WORKSHOP BRASILEIRO DE CIDADES INTELIGENTES (WBCI), 2., 2019, Belém. Anais do II Workshop Brasileiro de Cidades Inteligentes. Porto Alegre: Sociedade Brasileira de Computação, 2019.

BING MAPS. Bing Maps Resources. Disponível em <<https://www.microsoft.com/enus/maps/documentation>>. Acesso em 04 de mar de 2019.

BURIOL, Luciana Salete. Algoritmo memético para o problema do caixeiro viajante assimétrico como parte de um framework para algoritmos evolutivos. 2000. 136 p. Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP.

CARVALHO, E. B. de. Análise experimental de variações da heurística de Lin e Kernighan para o problema do caixeiro viajante multiobjectivo - Dissertação de Mestrado, Centro de Ciências Exatas e da Terra, Universidade Federal do Rio Grande do Norte, Natal, 2017.

FACONI, L. O.; ZOQUETI, V. R.; IMAMURA, C. Y. M. Aplicação de Clustering de Dados em Heurística para o Problema do Caixeiro Viajante no Cálculo de Rotas Seguras. To appear VI Encontro Sanjoanense de Engenharia e Tecnologia, Unesp, São João da Boa Vista-SP. 2019.

HERE MAPS. API references, developer guides and quick starts for HERE APIs and SDKs. Disponível em <<https://developer.here.com/documentation>>. Acesso em 10 de ago de 2019.

INFOSIGA. Base de Óbitos, 2018. Disponível em<<http://www.infosiga.sp.gov.br/Home/Relatorio>>. Acesso em 02 de mar de 2019.

SILVA, G. T. de L. e. Aplicação do algoritmo de DIJKSTRA na otimização multiobjetivo de rotas de evacuação em cenários de nuvem tóxica. Programa de Pós-Graduação - Universidade Federal de Pernambuco, 2017.

VARGAS, P. A. Sistemas Classificadores para Redução de Perdas em Redes de Distribuição de Energia Elétrica (98 p). Dissertação (Mestrado) - Universidade Estadual de Campinas (Unicamp), SP, 2000.