Towards Objective Cloud Computing Services Selection - Multi-Criteria Based Approach

Aleksandra Bączkiewicz
University of Szczecin
Szczecin, Poland *aleksandra.baczkiewicz@usz.edu.pl*

Jarosław Wątróbski
University of Szczecin;
National Institute of Telecommunications
Szczecin; Warsaw, Poland *jaroslaw.watrobski@usz.edu.pl*

Ewa Ziemba
University of Economics in Katowice
Katowice, Poland *ewa.ziemba@ue.katowice.pl*

Abstract

Cloud servers are becoming more widely used due to the growing number of mobile devices with limited capabilities for complex computing, processing, and storage. Choosing the optimal cloud server is challenging due to the continuous technological development of the growing number of cloud service providers. The need to evaluate cloud services according to multiple attributes suggests that multi-criteria decision analysis (MCDA) methods are appropriate. This paper proposes an approach to multi-criteria assessment of cloud services using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method employing various distance metrics with different approaches to prefer sustainable solutions and criteria compensation. A sensitivity analysis that considered changing criteria weights was applied to assess the robustness of evaluated solutions. The demonstrated approach proved its applicability for multi-criteria cloud service assessment considering sustainability and robustness.

Keywords: Cloud computing, Cloud services, Cloud servers, Multi-criteria assessment, Sustainability

1. Introduction

Cloud computing is a term for a distributed paradigm running on the Internet. Its resources ensure computing, data storage, and service solutions [9]. Cloud computing solves challenges such as limited battery power and processing capabilities encountered by mobile devices [7]. Mobile cloud computing (MCC) specifies a distributed paradigm that includes mobile computing, cloud computing, and networking [19]. MCC enables the transfer of tasks requiring complex computing to cloud servers with extensive resources [8]. Cloud computing enables the usage of virtualized resources available on the Internet instead of maintaining one's own computing infrastructure [14]. This technology allows users to avoid the administrative and technical problems associated with IT infrastructure, such as the cost of development, maintenance, and providing security [5, 18]. Cloud computing enables users to pay for computing and storage services on a pay-per-use model [11].

Selection of the optimal server among those available in the cloud is a challenge and is considered a research area that needs exploration. This is a problematic task due to the continuous development of cloud computing technology and the growing number of cloud service providers [13]. Multi-criteria decision analysis (MCDA) is among the techniques suitable for selecting the optimal cloud server [11]. This method makes it possible to consider multiple conflicting criteria, such as response time, cost, and speed during selection. MCDA methods show applicability in selecting the optimal cloud from among cloud services providing similar services [9].

This paper presents a multi-criteria approach to optimal cloud server selection considering multiple attributes, sustainability, and robustness of alternatives. The proposed approach is based on the well-known and widely used the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method applied with various distance metrics representing different considerations of sustainability and criteria compensation. Additionally, sensitivity analysis considering changes in decision-makers preferences reflected in criteria weight modification was performed. The presented approach involving modified parameters of multi-criteria analysis enables the identification of an optimal alternative considering robustness.

2. Methodology

The aim of this paper is to evaluate four cloud servers located in different regions of the world concerning six Quality of Service (QoS) [9] attributes using the TOPSIS method applying different distance metrics for one of the steps, which is to determine the distance of the considered alternatives from the reference solutions. Research was carried out based on a generalized form of the Minkowski metric [12], in a form adapted to the specifics of determining distances from reference points. The QoS parameters representing the criteria assessment considered in this research are given in Table 1.

Table 1. Description of QoS parameters considered for cloud servers multi-criteria assessment.

QoS Attribute	Description	References
C_1 - Speed	The time required for the cloud server to perform a given task.	[3, 19]
C_2 - Response Time	The time required to receive a response from the cloud server to the mobile device from the time the request is sent from the mobile device to the cloud servers without considering computation time.	[2, 6, 7]
C_3 - Proximity	The distance specified in kilometers from the user's location to the cloud servers.	[1, 10, 13]
C_4 - Cost per hour	The price of a virtual machine per hour of use. The cost depends on performance and location.	[8, 15]
C_5 - Availability	The availability of virtual machine resources.	[14, 15]
C_6 - Security	Security setting dependent on cloud providers.	[5, 18]

Table 2. Criteria values assigned for considered alternatives.

Cloud	Location	Speed	Response Time	Proximity	Cost	Availability	Security
A_1	Mumbai	3.64	19.1	842	0.0496	5	5
A_2	Paris	3.74	21.43	7833	0.0528	4	4
A_3	Sydney	3.75	22.56	9458	0.0588	4	4
A_4	North Virginia	3.85	29.19	14001	0.0464	4	4
Weight		0.27	0.34	0.22	0.02	0.1	0.04
Type		-1	-1	-1	-1	1	1

The values of performance regarding the parameters considered in the evaluation of each cloud server are presented in Table 2. A set of criteria assessment and performance values were acquired from the research conducted by Kurup S. and Guruprasad H. S. in 2022 [9]. The data was determined based on an experiment conducted for four real-time cloud servers located in different world regions. Criteria weights were determined by the authors of the mentioned paper using the Analytical Hierarchical Process (AHP) method [16]. Profit criteria type is represented by 1, and cost criteria are denoted by -1. Each considered device is Amazon EC2 instance t2.medium model with 2 vCPUs, 2.3 GHz, and 4 GiB. The mobile device used for the performed experiment is the Sony Xperia M C1904 model with CPU Octa-Core 1.6 GHz and 3 GB of RAM.

2.1. The TOPSIS Method

The TOPSIS method evaluates alternatives based on the calculation of alternatives' distances to ideal and anti-ideal reference solutions. In the original algorithm, Euclidean distance is used as the distance metric, although there are other metrics whose use in multi-criteria analysis requires exploration. The following stages of the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) method are provided below, based on [16].

Step 1. Normalize the decision matrix. The Minimum-Maximum normalization method for profit criteria r_{ij}^{+} and for cost criteria r_{ij}^{-} can be applied using Equation (1). another normalization method can be also applied in this aim.

$$r_{ij}^{+} = \frac{x_{ij} - min_j(x_{ij})}{max_j(x_{ij}) - min_j(x_{ij})},\ r_{ij}^{-} = \frac{max_j(x_{ij}) - x_{ij}}{max_j(x_{ij}) - min_j(x_{ij})} \tag{1}$$

Step 2. Calculate weighted normalized decision matrix with Equation (2).

$$v_{ij} = w_j r_{ij} \tag{2}$$

Step 3. Determine Positive Ideal Solution (PIS), namely v_j^{+} and Negative Ideal Solution (NIS) v_j^{-} using Equation (3).

$$v_j^{+} = \{v_1^{+}, v_2^{+}, \ldots, v_n^{+}\} = \{max_j(v_{ij})\},\ v_j^{-} = \{v_1^{-}, v_2^{-}, \ldots, v_n^{-}\} = \{min_j(v_{ij})\} \tag{3}$$

Step 4. Calculate distance from PIS D_i^{+} and NIS D_i^{-} for each alternative as Equation (4) shows. The original metric used for distance computation in TOPSIS is Euclidean distance.

$$D_i^{+} = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^{+})^2},\ D_i^{-} = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^{-})^2} \tag{4}$$

Step 5. Calculate the score for each considered alternative using Equation (5). The C_i value is in range from 0 to 1, and the alternative with the highest C_i value is the ranking leader.

$$C_i = \frac{D_i^{-}}{D_i^{-} + D_i^{+}} \tag{5}$$

2.2. Distance metrics

The Euclidean distance measured between two sets of points is received by calculating the square root of the sum of the squares of the differences between the respective points in compared sets a and b [4]. Equation (6) is applied for computation of the Euclidean distance, where m represents the size of sets.

$$d(a, b) = \sqrt{\sum_{i=1}^{m}(a_i - b_i)^2} \tag{6}$$

Manhattan (Taxicab) distance is used for the determination of the distance between two sets a and b by calculating the sum of absolute differences between corresponding particular points contained in compared sets, as Equation (7) shows [20].

$$d(a,b) = \sum_{i=1}^{m} |a_i - b_i| \tag{7}$$

The Chebyshev distance between two sets of points is computed with Equation (8). This metric is also called the chessboard distance [17].

$$d(a,b) = \max_{i=1,\dots,m} \{|a_i - b_i|\} \tag{8}$$

The Minkowski distance metric defined by Equation (9) provides a wide range of analytical possibilities for finding compromise solutions with different degrees of substitution and sustainability. The Minkowski metric allows modification of the degree of substitution in the solutions under consideration. The form of this function has useful decision-making properties. For p=1, the function takes the form of a Manhattan-type urban distance. For p=2, the function is a Euclidean distance, while for higher values of p reaching ∞, it becomes a Chebyshev distance.

$$d(a,b) = (\sum_{i=1}^{m} |a_i - b_i|^p)^{\frac{1}{p}} \tag{9}$$

3. Results

The research involves an experiment of conducting a multi-criteria evaluation of the considered cloud servers using the TOPSIS method, applying the Minkowski metric to measure the distance of the alternatives from the reference solutions. During the experiment, the value of the p parameter was increased in the range of 1 to 40 with a step equal to 1. The experiment was conducted to determine the range of values of the p parameter of the Minkowski metric for which there are changes in TOPSIS results for the considered set of alternatives. The TOPSIS scores obtained in each step of the experiment were compared with those obtained using the TOPSIS method using three other metrics: Manhattan, Euclidean, and Chebyshev. The Pearson correlation coefficient was used to compare results received with different distance metrics. Correlation of compared results is visualized in Figure 1.

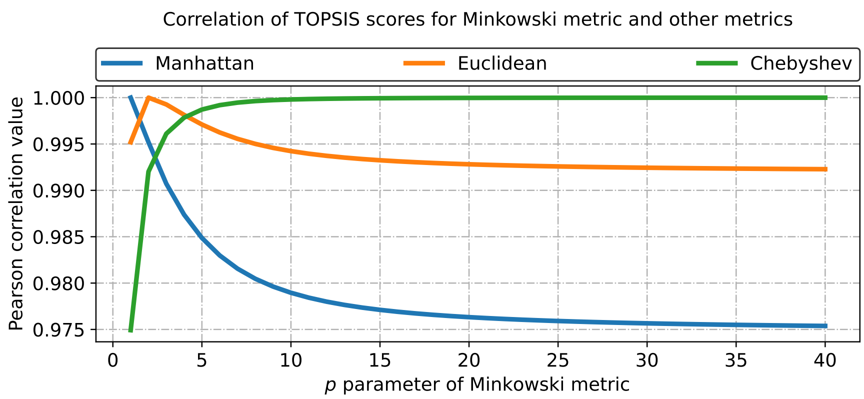

Fig. 1. Pearson correlation of TOPSIS scores for Minkowski metric and other distance metrics.

It can be noted that the TOPSIS scores obtained with the application of the Minkowski metric with increasing the p parameter from $p = 5$ and more are consistent with TOPSIS scores received using the Chebyshev metric. Increasing the p parameter in the Minkowski metric causes the obtained results to be increasingly less convergent with TOPSIS scores achieved using the Manhattan metric. The TOPSIS scores obtained using the Minkowski metric with $p = 2$ are identical to the TOPSIS scores achieved with the Euclidean metric. Then, an increment of the p parameter value causes a decrease in correlation with TOPSIS results received with the Euclidean metric. However, this decrease is less significant than observed in the case of Manhattan distance.

The experiment allowed us to determine the range of values of the p parameter from a value of 1 to a value of 10 in the Minkowski metric, in which TOPSIS scores and rankings stabilize for the considered alternatives. Table 3 shows the TOPSIS scores received using the mentioned range of the p parameter for the Minkowski metric. The calculated TOPSIS scores allowed to generate final positional ranking presented as $A_1 \succ A_2 \succ A_3 \succ A_4$ for all considered scenarios.

Table 3. Scores of the TOPSIS assessment received for considered distance metrics.

Clouds	p = 1	p = 2	p = 3	p = 4	p = 5	p = 6	p = 7	p = 8	p = 9	p = 10
A_1	0.9948	0.9898	0.9877	0.9866	0.9861	0.9857	0.9855	0.9854	0.9853	0.9852
A_2	0.5209	0.5895	0.6162	0.6316	0.6417	0.6487	0.6536	0.6571	0.6598	0.6618
A_3	0.4323	0.5106	0.5400	0.5571	0.5683	0.5762	0.5818	0.5860	0.5892	0.5917
A_4	0.0202	0.0386	0.0461	0.0498	0.0519	0.0531	0.0539	0.0544	0.0548	0.0550

It can be noticed that the TOPSIS rankings of evaluated cloud servers, regardless of the p parameter, are identical. The leader of these rankings is cloud A_1, located in Mumbai. This cloud server shows the best performance in terms of response time, which is the criterion with the most significant weight. Cloud A_1 also has the best performance regarding speed, proximity, availability, and security. In second place is cloud server A_2 located in Paris, and in third place is cloud server A_3 located in Sydney. The results obtained for these alternatives are in line with their performance values collected for each criterion. Cloud server A_4 located in North Virginia received the last place in all rankings despite having the best performance value in terms of the cost criterion. This shows that getting the best performance value for a single criterion by an alternative is insufficient to achieve a good score in a multi-criteria evaluation, especially when the relevance of this criterion is low, and the performance values for the other, more relevant criteria are weaker compared to the other alternatives.

Obtained results proved that cloud server A_1 is a stable leader among considered alternatives regarding provided criteria weights. Cloud server A_1 scored significantly superior TOPSIS scores compared to the alternative cloud servers included in this research for all p parameter values considered, confirming the competitiveness of this alternative. Cloud servers A_2 and A_3 scored very close to TOPSIS scores, which allows us to conclude that these alternatives are similar in terms of advantages like speed, availability, and cost aspects. Cloud server A_4 scored significantly low compared to the other alternatives, allowing this alternative to be considered the least advantageous.

The following stage was to conduct a sensitivity analysis to estimate the robustness of the considered cloud servers. For this purpose, different scenarios were generated with a different distribution of criteria weights. The weight of successive criteria was increased in steps, while the weights of the other criteria were changed equally accordingly so that the sum of the weights of all criteria was 1. Simulations were carried out for the three different distance metrics used in the TOPSIS method: Manhattan, Euclidean, and Chebyshev. Figure 2 visualizes the sensitivity analysis results performed for the Manhattan distance metric. It can be observed that evaluated alternatives are not susceptible to weight modification of speed (C_1), proximity (C_3), availability (C_5), and security (C_6). When the significance of response time is low, cloud server A_4 is better scored than A_3. In this situation, the weights of another criterion, namely cost (C_4), are higher, and A_4 shows good performance in terms of cost per hour. Thus, it causes its advance in the described situation. When the weight of cost is increased to 25%, A_4 jumps to the third rank. Increment of cost criterion to 45% affects that A_4 climbs to the second rank, and when the significance of cost is increased to 85%, A_4 becomes the ranking leader.

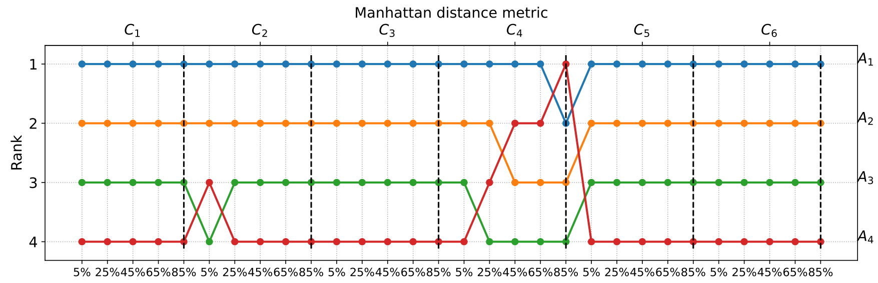

Fig. 2. Impact of changes in criteria weights on TOPSIS rank shifts for Manhattan metric.

For sensitivity analysis performed for Euclidean distance presented in Figure 3, increment of cost weight to 25% impacts on advance of A_4 to the second position. Further increment of cost weight causes a jump of A_4 to the leader position. Additionally, modifications of other criteria weights affect shifts of A_3 and A_4. It proves that this dataset evaluated using TOPSIS with Euclidean distance is more susceptible to individual criteria value change.

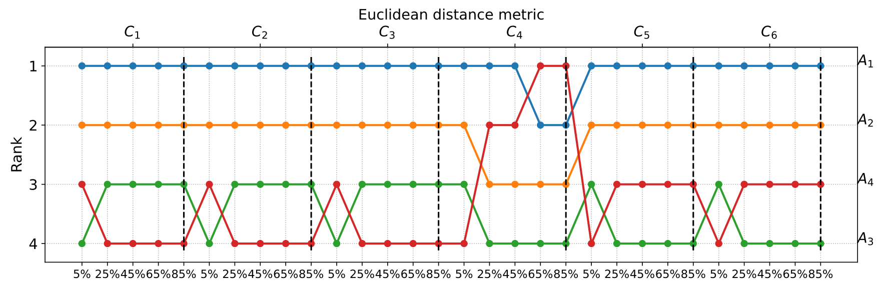

Fig. 3. Impact of changes in criteria weights on TOPSIS rank shifts for Euclidean metric.

Figure 4 displays the results of sensitivity analysis conducted for the Chebyshev metric.

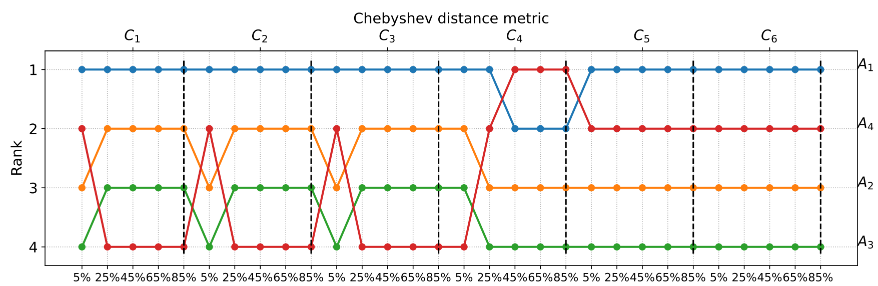

Fig. 4. Impact of changes in criteria weights on TOPSIS rank shifts for Chebyshev metric.

Similarly to previous simulations carried out for other distance metrics, it shows the advance of A_4 to the leader position when cost weight is increased. There are also observed shifts among A_2, A_4, and A_4 when other criteria weights are changed. When weights of criteria in terms of which A_4 shows poor performances are reduced, it advances ahead of A_2 and A_3, as demonstrated for C_1, C_2, and C_3 for the Chebyshev metric. The results of the performed sensitivity analysis proved that cloud server A_1 demonstrates a robust position in terms of performance values regarding considered criteria assessment. It maintains a stable leader position except for an enormous increment of cost criterion weight, causing its overtaking by A_4 with the best performance in terms of cost attribute.

4. Conclusions

During the selection of the optimal cloud server, it is essential to consider several attributes, both those whose values should be as low as possible, i.e., response time, computation time, proximity, and cost of use, as well as profit criteria such as availability and security. Multi-criteria methods such as TOPSIS offer a manner of selecting the optimal solution with limited objectivity. Analysis that considers parameters with different properties and approaches to prefer sustainable solutions, criteria compensation, and sensitivity analysis that considers decision makers' varying preferences for criteria weights extend the capabilities of multi-criteria methods with the ability to identify the most sustainable and robust alternatives.

Since this paper has a preliminary character and presents work in progress, it has several limitations. Among them is the incorporation of only four cloud servers in the investigation. Broadening the dataset to include more servers and their attributes is included in future work directions. Directions for future work also include a more dynamic approach to setting criteria weights and more sophisticated sensitivity analysis considering criteria weights, such as probabilistic sensitivity analysis or scenario-based analysis, and taking into account the different priorities and preferences of stakeholders. Another shortcoming is the inclusion of only one MCDA method in the analysis. Among future work, a comparative analysis considering the exploration of other MCDA methods is planned, including methods with distance metric-based algorithms such as CODAS, exploration of other distance metrics with applicability in MCDA methods, and their influence on assessment results.

Acknowledgements

Publication funded by the state budget under the program of the Minister of Education and Science named Perły Nauki, Poland, project number: PN/01/0022/2022, total project value: PLN 165 000,00 (A.B.) and Co-financed by the Minister of Science under the "Regional Excellence Initiative" Program RID/SP/0046/2024/01 (J.W.).

References

1. Al-Qerem, A., Alauthman, M., Almomani, A., Gupta, B.B.: IoT transaction processing through cooperative concurrency control on fog–cloud computing environment. Soft Computing 24, pp. 5695–5711 (2020)
2. Ding, D., Fan, X., Zhao, Y., Kang, K., Yin, Q., Zeng, J.: Q-learning based dynamic task scheduling for energy-efficient cloud computing. Future Generation Computer Systems 108, pp. 361–371 (2020)
3. Huang, L., Zhou, H., Feng, K., Xie, C.: Quantum random number cloud platform. npj Quantum Information 7(1), pp. 1–7 (2021)
4. Jain, G., Mahara, T., Tripathi, K.N.: A survey of similarity measures for collaborative filtering-based recommender system. In: Soft Computing: Theories and Applications: Proceedings of SoCTA 2018. pp. 343–352. Springer (2020)
5. Jangjou, M., Sohrabi, M.K.: A comprehensive survey on security challenges in different network layers in cloud computing. Archives of Computational Methods in Engineering 29(6), pp. 3587–3608 (2022)

6. Kazeem, A.M., Bamidele, J.A., Oluwaseun, R.O., Misra, S., Emmanuel, A.A.: Applicability of MMRR load balancing algorithm in cloud computing. International Journal of Computer Mathematics: Computer Systems Theory 6(1), pp. 7–20 (2021)
7. Khare, S., Chourasia, U., Deen, A.J.: Load balancing in cloud computing. In: Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 1. pp. 601–608. Springer (2022)
8. Kim, M.H., Lee, J.Y., Raza Shah, S.A., Kim, T.H., Noh, S.Y.: Min-max exclusive virtual machine placement in cloud computing for scientific data environment. Journal of Cloud Computing 10, pp. 1–17 (2021)
9. Kurup, S., Guruprasad, H.S.: Hybrid multi criteria decision methods for optimal cloud selection in mobile cloud computing. Indonesian Journal of Electrical Engineering and Computer Science 27(1), pp. 404–412 (2022)
10. Lai, P., He, Q., Cui, G., Xia, X., Abdelrazek, M., Chen, F., Hosking, J., Grundy, J., Yang, Y.: QoE-aware user allocation in edge computing systems with dynamic QoS. Future Generation Computer Systems 112, pp. 684–694 (2020)
11. Mostafa, A.M.: An MCDM approach for cloud computing service selection based on best-only method. IEEE Access 9, pp. 155072–155086 (2021)
12. Mukherjee, A., Li, Q., Song, Z.: An assessment of the effect of using different mappings and Minkowski distances in joint monitoring of the time-between-event processes. Journal of Computational and Applied Mathematics 404, pp. 113776 (2022)
13. Polyviou, A., Venters, W., Pouloudi, N.: Distant but close: Locational, relational and temporal proximity in cloud computing adoption. Journal of Information Technology 39(1), pp. 71–93 (2024)
14. Shahapure, N.H., Jayarekha, P.: Virtual machine migration based load balancing for resource management and scalability in cloud environment. International Journal of Information Technology 12, pp. 1331–1342 (2020)
15. Shirvani, M.H., Rahmani, A.M., Sahafi, A.: A survey study on virtual machine migration and server consolidation techniques in DVFS-enabled cloud datacenter: taxonomy and challenges. Journal of King Saud University-Computer and Information Sciences 32(3), pp. 267–286 (2020)
16. Sidhu, S.S., Singh, K., Ahuja, I.S.: Ranking of implementation dimensions for maintenance practices in Northern Indian SMEs using integrated AHP-TOPSIS approach. Journal of Small Business & Entrepreneurship 34(2), pp. 175–194 (2022)
17. Sun, Y., Li, S., Wang, X.: Bearing fault diagnosis based on EMD and improved Chebyshev distance in SDP image. Measurement 176, pp. 109100 (2021)
18. Swain, S., Tiwari, R.K.: Cloud security research-a comprehensive survey. Int. J. of Electronics Engineering and Applications 8(2), pp. 29–39 (2020)
19. Wu, H., Zhang, Z., Guan, C., Wolter, K., Xu, M.: Collaborate edge and cloud computing with distributed deep learning for smart city internet of things. IEEE Internet of Things Journal 7(9), pp. 8099–8110 (2020)
20. Zhang, X., Li, R., Hou, W., Shi, J.: Research on Manhattan distance based trust management in vehicular ad hoc network. Security and Communication Networks 2021(1), pp. 9967829 (2021)