International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Fuzzy Approach to Evaluate Data Colocation Centers

Sadia Husain

Submitted: 08/12/2023 **Revised**: 15/01/2024 **Accepted**: 29/01/2024

Abstract: We are living in the era of information technology where data is generated at enormous rate and every modern business needs their data to be stored and managed efficiently. Big corporations and government organization have enough resources to create and manage in-house data center while others may choose to rent the services from a data colocation center. A data colocation center basically provides IT related rental services to companies that required services like bandwidth, technologies, spaces etc. Colocation centers provide business with an efficient way to expand processing capabilities and grow their facilities without building everything from the ground up. Selecting data colocation provider not only increases prospect of additional flexibility and business value but also some potential risk. There are many data colocation service provider available and to choose the best one is not an easy task. In this research I have proposed an approach using fuzzy set theory to evaluate the data colocation centers. I have judged the colocation center on different criteria and used fuzzy sets to reach the decision. The fuzzy set theory gives the flexibility to express the expert opinion in linguistic terms, these linguistic terms has been assigned to some degree which can be evaluated by fuzzy set theory to conclude decision.

Keywords: *Data colocation, Data center, Fuzzy set, Fuzzy operations, implication operation*

1. Introduction

In today's world for any company who is using technology required a data center as a vital asset. The sizes of datasets are getting so enormous that businesses are finding it increasingly difficult to updates their data centers with technology, security and maintenance [1-3].

Traditionally, IT companies used to create their own data center and invest in upgrading technology time to time but now they are realizing that it's not sensible to invest in data center as there are better options available to accomplish their goals. One of such option is data colocation center. There could be two main reasons for shifting to data colocation center [4-7]. The first reason is that management of a data center is a tedious work, and the companies need to investment in multiple skills and technologies. Second, there is a growing demand from business functions for IT to participate in business transformation.

Data Colocation facilities may also provide physical security for the server, space, power, cooling, networking equipment, storage etc. It can also allow the companies to choose from different telecommunications and network provider with reasonable of cost and complexity.

Finding a good data colocation may include exploiting expertise throughout the organization, establishing formal requirements and collecting the unbiased reviews to evaluate colocation options. In this research I have introduce a fuzzy method to evaluate the colocation centers; this evaluation will help a business to find an efficient colocation center. Data center evaluation is done on criteria's which are based on expert's knowledge and by assigning some weight to each of them. A thorough search has been done to identify and analyze research articles based on the related topic. [8-12]. There are many criteria on which a data center can be evaluated but after reviewing main six qualitative criterions were identified as shown in Table 2. The proposed approach is defined by a framework as shown in figure 1.

College of Computer Science and Information Technology, Jazan University, Jazan, Saudi Arabia
ORCID ID : 0000-0002-9361-8701
Author Email: husain.sadi@gmail.com

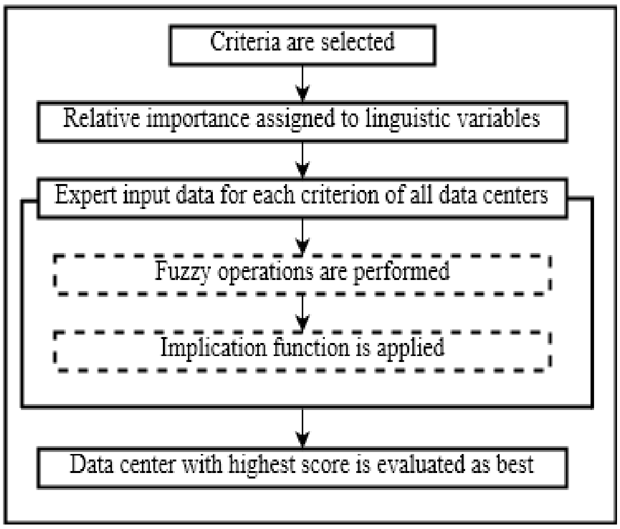

Fig. 1. Framework of proposed approach

1.1. Organisation of the paper

This paper is organized in three sections. Preliminaries is the first section which explained a brief review on fundamentals of fuzzy set theory and basic specification of data colocation centers. In the second section I explained different operation of fuzzy sets have used in this research. In the third section I presented a fuzzy approach to evaluate the data center. I have justified our approach with a numerical sample. Finally, after section III section of conclusion and future scope is presented.

2. Preliminaries

In the following section a brief introduction regarding the Fuzzy set theory and Data colocation center is discussed.

2.1. Fuzzy Set Theory

Lotfi Zadeh, who is also known as father and master of fuzzy logic introduce the concept of Fuzzy sets in year 1965. He basically worked to extend possibility theory to design a mathematical logic that can be used to express natural language into mathematical terms. This new logic is used to cater the fuzziness and termed as fuzzy logic [13-18]. Fuzzy set theory is used to define the fuzziness and Fuzzy logic is a part of Fuzzy set theory which can be used with different and boarder perspective.

Fuzzy logic worked with the concept that all things admit of some degrees of truth rather than true or false as we see in Boolean logic. Weather conditions, distance, speed, height, etc. all these data can come on a sliding scale. These days Fuzzy logic is extensively used to describe fuzziness. Fuzzy sets are mathematical means of representing information that is fuzzy in nature i.e. the information which is not precise. In fuzzy set theory, certain degree of membership is defined with each element. According to the theory, an element's value is not always true or false, but it can be partially true or partially false and can be represented with any degree. A degree can be given in any real number which lies in the interval [0, 1].

Table 1: Sample of A Fuzzy Set 'Near" With Its Membership

Name	Distance	Membership Degree (Crisp set)	Membership Degree (Fuzzy)
Car1	2.1	0	1.00
Car2	1.9	1	1.00
Car 3	2.4	0	0.76
Car 4	2.5	0	0.70
Car 5	1.5	1	0.80

An easy-to-understand example of fuzzy sets is to evaluate distance between two cars on a street. We assume cars are having distance less than two meters between them are considered near and above 2 meters are considered far or not near. When we evaluate this scenario in a crisp set a question is raised like: whether the cars are near to each other or not? The possible answer will be either yes or no depending on the value of the distance between them. If distance is 2.2 meter it will be considered not near. In Table 1 we can see the representation of this example in fuzzy sets numbers. The fuzzy set "near" are defined by degrees of membership depend on their distance provided for each element in meter. When we define a Fuzzy set, a question is raised like as follows; How near the cars are? The distance variable can have a partial membership in the fuzzy set, Car 5 is 0.85 near as shown in Table 1. In a crisp set car can be either near or not and can have membership degree 0 or 1.

According Zadeh's fuzzy set theory [13] A Fuzzy set R of universe X is defined by function $\mu_R(x)$ called the membership function of set R".

R= {x, $\mu_R(x)$ x | ∈X}

μ_R (x): X → [0, 1]

"Where μ_R (x) (which is also called the membership value) is equal to 1 if x is completely lies in R; $\mu_R(x) = 0$ if x is not in R; $0 < \mu_R(x) < 1$ if x is partially lies in R The non membership value is calculated by subtracting membership value from 1. The Fuzzy set gives the possibility to express the values of membership function in a range of possible choices. For any element x of universe X, membership function $\mu_R(x)$ equals the degree to which x is an element of set R. The degree of membership can also be expressed as membership value, of any element x in set R.

Take another example where communication skills are a fuzzy linguistic variable and represented with terms: excellent, poor, good, very poor and average etc. This linguistic term can be represented with linguistic modifier. By Combining the linguistic variable and a linguistic modifier a fuzzy set can be formed. To represent the level of communication a scale of 1 to 5 is defined and assign a value to each linguistic modifier. Thus, a fuzzy performance set Q and its modifiers can be represented as:

Q= {1.0 | 0, 2 | 0.4, 3| 0.6, 4| 0.8, 5| 1}

According to Zadeh "An element of fuzzy set, in a universe of discourse should have values with a degree of membership or non-membership". In our example, value 5 has a membership value of 1 respective to very excellent communication-skills, and a non-membership value of 0 (which is calculated by subtracting membership value from 1) indicating poor communication-skills. The value 3 with a partial membership value of 0.6 having average communication-skills.

2.2. Data colocation center basic specification

When a business chooses a colocation center over building company's own data center it must go through lots of considerations, but as per analysis the major reason is associated with bearing the cost of building, updating, or maintaining a large computing facility etc. After the review of several articles on data colocation following qualitative specification of a data center is identified by experts [19].

Location: The location plays an important role in data colocation as the company that is outsourcing its colocation center may want to access the facility remotely using a dedicated connection to one of the data hubs in the facility. A dedicated connection shall provide added security and speed which may be in the form of a fibre optics line. Hence the cost of setting up a dedicated line shall be proportional to the distance between the facility and the company and its data colocation facility. If speed and security is not the concerned then the company may use Internet connection, in such cases the location of the facility may not be an important issue as far as it guarantees reliable high-speed internet.

Power: An important concerned for choosing the data colocation center is the availability of the power backup as and when it may be required. As the power is the essential requirement of the data center, the facility must guarantee the availability of the power through 365x7x24.

Space: It is a concern because before choosing a data center facility one must be assured that the data center has enough space to accommodate the primary data storage devices and its backup devices. Besides storage devices the facility must also have adequate space for power and power backup equipment's along with other amenities to run the facility smoothly like cooling equipment to control the temperature of the facilities and the equipment within it.

Security: It is of great importance because the company might be sharing very sensitive data. Hence, before choosing a data colocation facility, it must be ensured that proper state of art features for guarantee the security of the data are incorporated. This must include protection not only from internal threats like the loss of power but also from external threats like unauthorized access of the data.

Connectivity: The bandwidth for the internet and network connection available in the data center facility must be adequately high to provide seamless and lag free experience to the users who are accessing the data from the facility [20]. Moreover, multiple connectivity options shall be available to ensure the availability of the data whenever required.

Industry Compliance: A colocation center that follows the industry compliance standard is assumed to be legitimate. Big business before hiring a colocation center assures that the center is industry standard complaint and having required certification, example: There are international service organization assurance standard like ISAE 3204, PCI DSS etc. ISAE 3204, provided by accountants to help customers make better decisions. A set of security standards like payment card industry (PCS) and data security standard (DSS), are secure environment meant to cater that all companies that work with credit card information must operate in a secure environment.

Customer service: A colocation center must provide an efficient customer service. Customers of any business remain loyal to the company if they are satisfied with the customer care service. Most of the customers of a data colocation center look for 24/7/365 remote hands services, they need dedicated and expert support staff on-site that can handle any queries any time and make any service changes they need or take care of any potential issues immediately.

Service Cost: A colocation center must offer a competitive price for its services to remain in market. A colocation center service provider should assure a customer that it is offering the best price for the services, facilities and flexibilities required by customer. Most of the time if company sign up to a long-term deal, it likely pays rather less per month than a flexible month-by-month contract. Ultimately, a company must try to find out a colocation provider with flexible pricing options also the company expect that the provider should structure a deal in a way that a company is comfortable with and fully understand.

Services: It's not necessary that all data colocation service provider offers all services as per industry demands. Services available in a data colocation center depend on its policy or how much a provider has invested in the business. Some colocation centers don't have customer care services, some colocation center doesn't allow a company to install its own physical server at their location instead they offer their own sever to do the task. Some colocation centers are not industries compliant they don't have required certification. So, it's advisable to a business before finalizing a data colocation center make sure it is offering you all required services along with best security.

3. Fuzzy Set Operations

There are many fuzzy set operations, but four fuzzy set operations are fundamental: Intersection, union, complementation, and implication operations [20-24]. To understand these basic fuzzy operations, assume P and Q are fuzzy sets with membership functions:

$\mu_P(x) = \{1,2,3,4,5,6\}$ and

$\mu_Q(x) = \{1,2,3,4,5,7,8,9,10\}$

The union of P and Q is a fuzzy set.

$C = P \cup Q$, where

$\mu_C(z) = \mu_{P\cup Q}(z) = \mu_P(x) \cup \mu_Q(y)$

Union operation is same to logical OR operation and fuzzy set union is performed by taking maximum elements of each fuzzy set's elements. Here maximum is calculated by applying a MAX() function.

So, $\mu_{P\cup Q}(z) = \{1,2,3,4,5,6,7,8,9,10\}$

The intersection of P and Q is a fuzzy set $D= P\cap Q$, where

$\mu_D(w) = \mu_{P\cap Q}(w) = \mu_P(x)\cap\mu_Q(y)$

Intersection operation is same to logical AND operation and fuzzy set intersection is performed by taking minimum elements of each fuzzy set's elements. Here minimum is calculated by applying a MIN() function.

$\mu_{A\cap B}(w)= \{1,2,3,4,5\}$

The complement operation on a fuzzy set is done by subtracting every member of fuzzy set from its highest possible value, in our example the highest value is 10.

So, $\mu\ (x)= 10-\mu_{\sim p}\ (x)= \{9,8,7,6,5,4\}$

$\mu_{\sim Q}\ (x)= 10-\mu_{\sim Q}\ (x)= \{9,8,7,6,5,3,3,1,0\}$

The fuzzy implication function is used to evaluate the degree of fulfilment of a rule represented as IF X THEN Y, where both the antecedent and the consequent are fuzzy in nature. These functions must adhere to fundamental properties, with the most common being the Kleene-Dienes implication, which is derived on the classical implication definition expressed as: $(x\rightarrow y=\sim x\cup y)$ which utilizes the Zadeh fuzzy operation properties of negation and the maximum S-norm, but there are many other fuzzy implication functions also exist which can be used in fuzzy applications like S, N- implications [25]. A Fuzzy implication function is derived from classical implication operator or we can say it's an extended version of a classical implication operator in which the two values can belong to the set [0,1] as oppose to classical implication where implicatory operator can have value either true or false (0 or 1). Thus, it is a function $f:[0,1]\rightarrow [0,1]$. So, a fuzzy implication function can be utilized to conclude if P is true, to what degree that indicates that Q is true? The implication operation is performed by evaluating $\mu_{\sim pUQ}(u)$, recognized as Kleene Dienes implication.

4. Data Colocation Center Evaluation Using Fuzzy Approach: A Numerical Example

Evaluating a data center is not easy. Lots of analysis has been done for deciding common criteria for evaluation of colocation center. I have finalized some criteria after getting good insights on data colocation center specification from online research articles and library. However, for illustration purposes and keep the paper numerical calculation simple our paper includes six criteria namely Services offers, Security Measures, Cost of Services, Connectivity/Speed of accessibility, Customer Services, and Industry Compliance. The criteria are presented in Table 1I.

Table 2: Sample list of criteria and its relative importance

Criteria	Symbol	Relative Importance
Services offers	C1	1
Customer Services	C2	0.6
Industry Compliance	C3	0.4
Cost of Services	C4	0.8
Security Measures	C5	0.9
Connectivity or Speed of accessibility	C6	0.7

I have formed a fuzzy set C in a universe of discourses U having unit interval [0, 1] given as : C={u|μ_C(u), u∈U}

μ_C(u)= {1. 0, 0.60, 0.40.0, 0.80, 0.90, 0.70}

Each element in this set is scored between 0 and 1. These scores also signify the relative importance of each criterion to the decision maker. For each criterion a qualitative judgment will be given to assess the degree of data center performance for those criteria. Based on criteria qualitative judgments can be mentioned or represented in many options like poor, above average, average, outstanding, below average, slightly above average etc. Thus, a new fuzzy set P is formed in the universe V with unit interval [0,1] and a fuzzy membership function. These linguistic variables are shown in Table 3.

P= {v | μ_P(v), v∈U}

μ_P(v)= {0.10, 0.20, 0.30.0, 0.40, 0.60, 0.70, 0.80, 1. 0}

Once the linguistics terms are finalized and relative importance is mentioned, I will exemplify how the evaluation of fuzzy sets can help in decision making. There are various steps involved to reach the decision. The very first step is to measure the performance of colocation centers against the criteria considered for evaluation. In our sample I have assumed that five data colocation centers are to be evaluated. The decision maker who asses the colocation center is assumed to be a colocation centers auditor who has sufficient knowledge about all the five data colocation centers named as CC1, CC2, CC3, CC4 and CC5. The decision maker expresses the opinion in form of linguistic terms available with assigned symbol as depicted in Table 4.

Table 3: Importance Of Linguistic Variables in Sample

Relative Importance	Symbol	Linguistic Variable
0.1	NA	Not acceptable
0.2	P	Poor
0.3	BA	Below average
0.4	SBA	Slightly Below average
0.6	AVG	Average
0.7	SAA	Slightly above average
0.8	AA	Above average
1	OS	Outstanding

Table 4: Colocation centers grading against the criteria.

	CC 1	CC2	CC3	CC4	CC5
C1	BA	SAA	SAA	AVG	AA
C2	P	AVG	AVG	BA	SAA
C3	SAA	AA	SBA	SBA	SAA
C4	AA	SBA	AVG	AA	AVG
C5	AVG	AVG	AA	SBA	SAA
C6	AVG	P	OS	SAA	AA

Table 5 contains five fuzzy sets P1, P2, P3, P4, P5 with belonginess/membership grades of each element of the set:

μ_{p1}(v), μ_{p2}(v), …….μ_{p5} (v)

For example, I can represent the fuzzy set along with its membership function for data center CC1 as:

P1= {1.0| 0.30, 2.0|0.2, 3.0|0.7, 4.0|0.8, 5.0|0.6, 6. 0|0.6}

μ_{p1} (v)= {0.30, 0.2, 0.7, 0.8, 0.6, 0.6}

Table 5: Membership Grades Assigned to Criteria

	CC 1	CC2	CC3	CC4	CC5
C1	0.3	0.7	0.7	0.6	0.8
C2	0.2	0.6	0.6	0.3	0.7
C3	0.7	0.8	0.4	0.4	0.7
C4	0.8	0.4	0.6	0.8	0.8
C5	0.6	0.6	0.8	0.4	0.7

Now comes the important stage of the decision-making process where I applied a fuzzy implication between a particular criterion with each data centers performance. This step is done to understand, each criterion along with its given relative importance, we can understand this in terms of question: does the data center imply a good performance for that criterion?

Now I apply the Kleene & Dinies [23] implication function as:

$\mu_{\sim C \cup p1}(r) = \mu_{\sim C(u) \cup} \mu_{P1}(v)$

= {0.0, 0.4, 0.6, 0.2, 0.1, 0.3} ∪ {0.30, 0.2, 0.7, 0.8, 0.6, 0.6}

= {0.3, 0.4, 0.7, 0.8, 0.6, 0.6}

Similarly, I will calculate for all fuzzy set P2…P5. The final step is to get the result by combining all the performances of colocation centers against criteria. The is achieved by utilizing min () function to the set derived by implication function for all fuzzy set P. This result is shown in Table 6. From the table no. 6 it can be deduced that none of the data colocation center is outstanding, but colocation center no. 5 has achieved the highest score indicating slightly above performance than other data colocation centers. Whereas colocation centers 2 and 1 are below average performers with total score .3.

Table 6: Overall Rating of Data Colocation Centers

Data colocation centers	Score
CC1	0.3
CC2	0.3
CC3	0.6
CC4	0.4
CC5	0.7

5. Conclusion And Future Scope

For some organizations data colocation centers are proved to be an ideal solution to expand their business with less investment in managing their data. But the selection of colocation center is a tedious task as there are many colocation center providers available in market. A wrong choice can lead to business downfall because for many companies' data is among the company's most valuable assets. Now days most of the colocation centers are industry compliant and go through regular audit to maintain their standard.

In this paper I have presented a novel fuzzy approach to evaluate the colocation centers by utilizing the expertise of colocation centers auditor and allow the knowledge of auditor to be expressed in linguistic terms. I have used various fuzzy set operations to deduce our result. I have use simple calculation / math functions, but MATLAB software can be used to computerize the calculation and result. For the easy understanding, I have used six criteria and five colocation centers, but our approach is flexible to be applicable on a greater number of criteria and a greater number of colocation centers.

This paper introduces an efficient and easy fuzzy model for founding a more competent data colocation center evaluation system. This approach evaluates all data centers against predefined criteria and gives the best performer based on scored achieved but this approach could not work if the company wants to evaluate a data colocation center as per company's requirement for some specific criteria. So, this paper can be extended in future to help the businesses to decide a data colocation center by evaluating their requirement criteria not by simply judging against some expert's predefined criteria.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Wierman, A., Liu, Z., Liu, I., & Mohsenian-Rad, H. (2014). Opportunities and challenges for data center demand response. Paper presented at the International Green Computing Conference.

[2] Shuja, J., Bilal, K., Madani, S. A., & Khan, S. U

[3] (2014). Data center energy efficient resource scheduling. Cluster Computing, 17(4), 1265-1277.

[4] Greenberg, A., Hamilton, J., Maltz, D. A., & Patel, P. (2008). The cost of a cloud: research problems in data center networks. ACM SIGCOMM computer communication review, 39(1), 68-73.

[5] G. Ghatikar, V. Ganti, N. E. Matson, M. A. Piette, "Demand Response Opportunities and Enabling Technologies for Data Centers: Findings from Field Studies", 2012.

[6] Filani, D., He, J., Gao, S., Rajappa, M., Kumar, A., Shah, P., & Nagappan, R. (2008). Dynamic data center power management: Trends, issues, and solutions. Intel Technology Journal, 12(1).

[7] Greenberg, A., Hamilton, J., Maltz, D. A., & Patel, P. (2008). The cost of a cloud: research problems in data center networks. ACM SIGCOMM computer communication review, 39(1), 68-73.

[8] Islam, M. A., Mahmud, H., Ren, S., & Wang, X. (2015). Paying to save: Reducing cost of colocation data center via rewards. Paper presented at the 2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA).

[9] Sun, Q., Wu, C., Ren, S., & Li, Z. (2015). Fair rewarding in colocation data centers: Truthful mechanism for emergency demand response. Paper presented at the 2015 IEEE 23rd International Symposium on Quality of Service (IWQoS).

[10] Zhang, L., Ren, S., Wu, C., & Li, Z. (2015). A truthful incentive mechanism for emergency demand response in colocation data centers. Paper presented at the 2015 IEEE Conference on Computer Communications (INFOCOM).

[11] "Colocation Market - Worldwide Market Forecast and Analysis (2013–2018)", [online] Available: http://www.marketsandmarkets.com/Market-Reports/colocation-market-1252.html.

[12] "Pricing Data Center Co-location Services", 2009, [online] Available: http://enaxisconsulting.com.

[13] J. Novet, Colocation providers customers trade tips on energy savings, Nov.2013, [online] Available: http://www.datacenterknowledge.com/.

[14] L. A. Zadeh, Fuzzy Sets, Information and Control 8 (1965) 338-353.

[15] Kaufmann and M.M. Gupta, Introduction to Fuzzy Arithmetic Theory and Applications (Van Nostrand Reinhold, New York, 1991).

[16] Kaufmann: Introduction to the Theory of Fuzzy Subsets, Vol. I, Academic Press, New York, 1974.

[17] H.-J. Zimmermann, Fuzzy Set Theory and It's Applications, Second Revised Edition.

[18] Zadeh, L.A., "Fuzzy Logic," Computer, Vol. 1, No. 4, pp. 83-93, 1988.

[19] Zadeh, L.A., "Knowledge representation in fuzzy logic," IEEE Transactions on Knowledge and Data Engineering, Vol. 1, pp. 89-100, 1989.

[20] "Colocation Data center ABC", HubSpot, [online] Available: https://cdn2.hubspot.net › hubfs › PP_Colocation Data Center ABC.

[21] Akinola, Ayotuyi, "Performance Evaluation of Data Center Network Setup Architectures", (2017).

[22] Dubois, D., & Prade, H. (1978). Operations on fuzzy numbers. International Journal of systems science, 9(6), 613-626.

[23] Karnik, N. N., & Mendel, J. M. (2001). Operations on type-2 fuzzy sets. Fuzzy sets and systems, 122(2), 327-348.

[24] Mizumoto, M. (1981). Fuzzy sets and their operations, II. Information and control, 50(2), 160-174.

[25] Mas M, Monserrat M, Torrens J, Trillas E. A survey on fuzzy implication functions. IEEE Transactions on fuzzy systems. 2007 Dec 6;15(6):1107-21.

[26] M. Baczyński, B. Jayaram (2007) On the characterization of (S,N)-implications Fuzzy Sets Syst., 158 (2007), pp. 1713-1727