A Caching Algorithm for Information Centric Network Using Fuzzy Logic

Mark Lorenze R. Torregoza[1], Elmer P. Dadios [2]
[1]*Department of Electronics and Communications Engineering, De La Salle University, 2401 Taft Avenue Manila, Philippines.*
[2]*Department of Manufacturing Engineering & Management, De La Salle University, 2401 Taft Avenue Manila, Philippines.*
mark.lorenze.torregoza@dlsu.edu.ph

***Abstract*—The internet today has evolved from information superhighway to a household necessity that offers more than just information. Nowadays, the internet serves a lot of purpose. It is a tool for not just information but entertainment that offers music, graphics and videos that is available for downloading or streaming. It has also evolved to be a medium of communication that offers a global link from people around the globe. From emails, short message services and even voice communication, the internet has all of these to offer. The former information superhighway is today a social media platform that is open to all ages to all variety of users. With this development, it is logical to think that the current internet network scheme should also be subjected to evolution. The emerging Information Centric Network is quite a good fit to the future of internet. The idea to be concerned to the content that is to be accessed more than the identity of the one accessing the content is tailor-fit to the current application of internet. In a nutshell, ICN requires node with caching functionality. An effective caching algorithm is a great help to attain the very purpose of ICN which is to come up with an efficient network. Meanwhile, fuzzy logic, which has proven to be effective in control or optimization applications, can also be applied in improving caching functionality of ICN. This paper explores the application of fuzzy logic to the caching algorithm that can be used to further improve current information centric networks. The results were obtained from hypothetical data because this is just to prove that fuzzy logic can be applied in the caching dynamics of Information Centric Network.**

***Index Terms*—Caching Algorithms; Computer Networks; Future Internet; Fuzzy Logic; Information Centric Network.**

I. Introduction

The internet today is very far from what it was decades ago, during its first launch into human awareness. Started as a research project from military applications, then later on morphed as a source of information, the internet today serves a lot of purpose. In its early applications, the internet was used to browse files like pictures, documents, and others hosted by a remote server. The server and the one accessing the information are both connected to the internet and talking to each other across thousands of computer connected on the same vast network. Years later, the simple downloading of information changed to exchange of material. E-mail was developed, and the internet was never the same. More and more people started using the internet because of the need to send electronic mails and the usual access to information. Several years after, the internet becomes a household necessity that can be used not only for researches, not only for exchange of e-mails, not only to download pictures but a medium for exchange of information, a social media platform, and avenue for both entertainment and education. It is clear that the 21st century internet is a vast collection of services that is readily available to its users.

As the use of internet grows, the number of users also multiplies. This is a happy problem though. However, as more people use the internet, the current network turns to be over-crowding and majority of the core services are greatly affected. Like for example, as more people use the internet to make calls using the internet protocol network the quality of these calls will be affected. Though it may not be evident in paid VoIP calls, but for free call services that is very popular today, the quality of service is a function of the number of users and the capability of the network to handle such volume as well.

Though the applications and uses of the internet went under a drastic change, the network architecture behind is still almost identical to its first launch. The current network architecture of the internet still uses the client-server design. Every user needs to log in to the network and be connected to the server in order for the exchange in information to take place. In the event that the server is unavailable, the service will also be unavailable to the user regardless if the whole network is up and running. Faster access to information is also an issue in this current network design. If the network routes leading to the server are full or loaded, then the client will have a hard time reaching the server, increasing chances of slow access time and making the over-all performance of the network inefficient on the perspective of the user. It is evident that there is a need to improve the current internet architecture to make it fir for the current internet application.

There are researches today that are gearing towards a change in the current internet architecture. This is to further improve the internet on top of giving users more favourable experience in browsing data, exchanging information and communicating over the once known as information superhighway. One thing of interest in the future of internet is the currently emerging Information Centric Network. ICN is a promising solution to the problems related to modern-day computer networking.

II. ICN at a Glance

With the inevitable evolution of the internet, the network supporting such service should also be open for a drastic change. So far, information centric network shows promising solutions to current network problems. Numerous efforts through researches such as [1,2,3] are presented to the public explaining the core concepts of ICN.

ICN is a network architecture that is more concern on the 'what' rather than the 'who'. In other words, the focus of ICN is to make sure that the information is available to those who are accessing it all throughout the network. Contrary to the current IP network wherein the first step to establish access to information is to identify the client connecting to the server, ICN is leaning towards the idea that the information is more important than the identity of the client or the server. In the current IP architecture, once the server is down then the information is unavailable also. This is further explained in Figure 1.

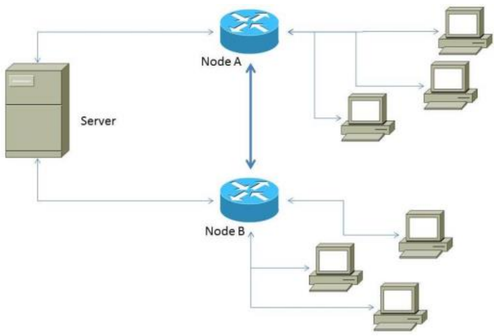

Figure 1: Conventional IP Network Architecture

In the event that the server failed, regardless if the whole network is up, the service is still unavailable. To make matters worse, if the same information is being accessed by all the computers for both node A and node B, then all the computers will not be able to retrieve the information because the server is down.

On the contrary, this can be eliminated using ICN. ICN operates in a very simple exchange of information. Unlike in the conventional IP networks wherein there are a lot of messages needed for nodes to communicate, in ICN only data and interest message are involved. The client that requires certain information sends an interest message to the node it connected to. If the node has the content that the client needs, it will send back a data message containing the information requested. However, if the node does not contain the needed information, then it will take hold of that interest message and the incoming interface that points to the origin of the interest message. Then the same node will send a new interest message to other nodes and the cycle goes on and will only stop once a data message that signals that the node containing the data is already found. Since each node takes hold of the incoming interface, the data will be sent to the origin of the interest message by just tracing back the route.

It is evident that in ICN the concern is the information, one of its strength is the spreading out of information throughout the network. It simply means that not only the server is the one holding the data to be accessed, but also some other parts of the network. This will be explained with the help of Figure 2.

Some nodes along the network is equipped with caching functionality. In the event that a computer of one node accesses content, the nodes will cache this information. If another computer accesses the same content, then there is no need for that computer to reach the server because the information is already present in the nearby node.

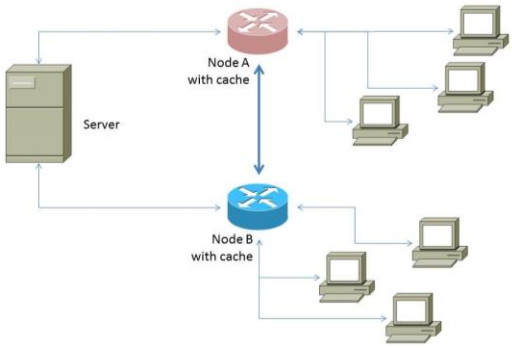

Figure 2: ICN with Nodes with Cache

This scenario is greatly applicable to social media networks as discussed in [4,5]. It is because of the fact that in social media, majority of the users are retrieving the same content. One good example is Twitter. All followers of a certain Twitter account retrieve the same information. If this information is scattered all though out the network then it will be easy to all the followers to access the information, even without direct connection to the Twitter server.

Like with current computer networks, over all network traffic is also put into consideration. The idea is of course to be able to give users ease of access to information. In [6] traffic engineering for information centric network has been discussed and highlighted methods to evenly spread the information across the network.

III. Fuzzy Logic and Caching

It is evident that the placement of nodes with cache is of vital importance to the performance of the network as a whole. If the nodes with caching functionality are not strategically placed in the network, it may not serve the purpose correctly, worse, it may just increase network traffic resulting to poor quality of service or even network congestion.

In [7] the importance of good caching dynamics was highlighted, modelled as well as evaluated. It has been proven that caching plays an important role in the whole purpose of information centric networking.

In [8] autonomic cache management architecture is proposed. It was then proven that it can give significant benefit in terms of the overall network performance. In [9, 10], the performance of cache in ICN was further investigated and evaluated and related to the actual performance of the network.

Caching has also been incorporated in actual routing procedures that also bring in good effects to the network performance. This is discussed and evaluated in [11]. In [12] resource allocation and caching in ICN was presented and evaluated further to highlight the effects of caching in network performance. Another proposal in [13, 14] is the cooperative caching and routing in ICN. In [15] multicast in caching techniques has also been proposed and evaluated to give significant improvement in download time.

This paper presents an algorithm and decision making using fuzzy logic to further develop caching in information centric network. Fuzzy logic has been widely used in many applications and consistently attracting attention to new users to explore additional applications in other fields [16]. It provides solution to control problems. It is also useful in communications; computer networks and congestion control [17, 18, 19] and even in medical decision making [20]

In this paper, fuzzy logic is use to make decision whether a certain information should stay or not in the cache of a certain node. The frequency of request for certain information and the size of the information itself will be taken into consideration by the fuzzy logic system and later on make decisions based on the two parameters.

IV. The Fuzzy Logic System

There are two inputs that the system requires. First is the size of packet to be kept in cache. Second is the frequency or the number of times that the packet has been requested. The fuzzy logic system will decide based on those input and will give a priority number ranging from one to five. This is can be seen in Figure 3.

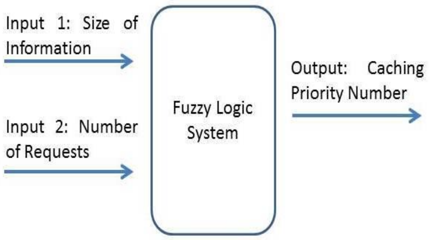

Figure 3: Fuzzy Logic Block Diagram

This is because of the assumption that there are only five slots per cache per node. This will still work however in systems where the cache is not limited to the number of information, as most of the time the amount of cached information is dependent on the size of the storage. In this event, the information with the lowest priority numbers will be retained, the number of information will depend on the size of the storage. So for as long as the information can be saved, it will be cached. Only those with relatively high priority number will not be cached.

The output is a priority number that indicates whether the content will be removed from the cache or will stay there for others to access the information. All information will have its own priority number for a given node.

V. The Membership Function

The membership function is given in the figures below. The first is for size of information. Very Small means that the size is from 0kB to 100kB. The typical size of a pure text tweet as an example is around 5kb to 10kb, this will fall in the category very small.

For further understanding the value for each membership function is tabulated and presented in Table 1 and 2.

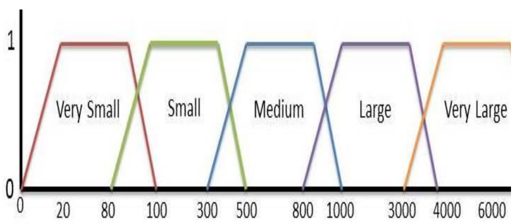

Figure 4: Membership Function for Size

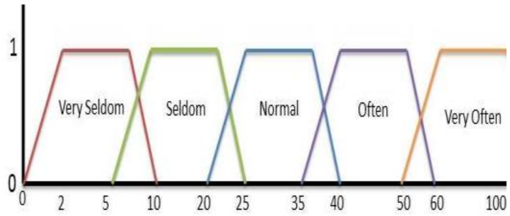

Figure 5: Membership Function for Request Count

Table 1
Membership Function for Size

Membership Function for Size	
Very Small	0-100kB
Small	80kB-500kB
Medium	300kB-1MB
Large	800kB-4MB
Very Large	3MB-Infinity

Table 2
Membership Function for Request Count

Membership Function for Request Count	
Very Seldom	0-10
Seldom	5-25
Normal	20-40
Often	35-60
Very Often	50-Infinity

VI. The Fuzzy Associative Matrix

The fuzzy logic system will decide whether a certain information will be cached, this will happen because the fuzzy system will output a priority number and if the number of allotted. The decision will be based on rules. The important rules are the following.

Rule 1: If the information size is very small and it is requested very often then it should get the highest priority number.

Rule 2: If the information size is very small and it is requested not often, it should get a higher priority number than the case that will satisfy rule number 1.

Rule 3: If the information takes a lot of space in storage, meaning, large in size, then it should not get a low priority number unless it is requested very often.

The fuzzy associative matrix given in table 3 further explained the rules where the decision making process will be based.

Table 3
Fuzzy Associative Matrix

Size	Request Count				
	VS	S	N	O	VO
Very Small	4	3	2	1	1
Small	4	3	3	2	1
Medium	5	3	3	2	2
Large	5	4	3	2	2
Very Large	5	5	4	3	2

VII. Data and Results

The paper only proves that fuzzy logic is applicable to be part of the caching dynamics of ICN. Using hypothetical data, and running the fuzzy logic system and recording the priority number as the output, the following results were obtained. Though the testing was done numerous times, only a portion was shown for the purpose of presentation.. Figure 6 shows the size of information used in the simulation.

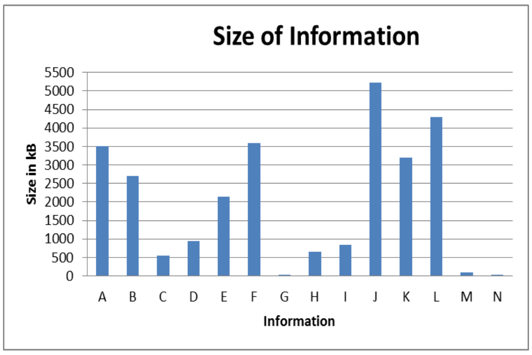

Figure 6: Information Size

Figure 7 shows the number of requests each information has in the whole simulation.

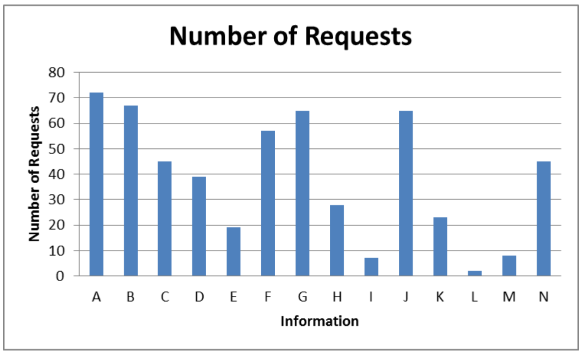

Figure 7: Number of requests.

Table 4 shows the output of the algorithm. The first column is the tag of the particular information. The second column is the size given in kB and the third column is the number of request for that given information in a particular time. It follows that information A occupies 3.5MB of storage space with 72 requests and information N occupies 25kB of storage space and has 45 requests.

Table 4
Output of the Algorithm

Time	Size	Request	Output
A	3500	72	2
B	2700	67	2
C	550	45	2
D	950	39	2.285
E	2150	19	4
F	3600	57	2.1875
G	25	65	1
H	650	28	3
I	840	7	4.285
J	5215	65	2
K	3200	23	3.71
L	4300	2	5
M	92	8	3.44
N	25	45	1

The presented table covers most of the possible combination of size and number of requests. Figure 8 summarizes the output of the fuzzy logic system.

VIII. Analysis of Results

After simulating possible scenarios in an information centric network, the results were gathered and analyse. The algorithm showed that it is not only concerned on the number of request information has, but it takes into consideration the amount of space that the particular information consumes in the storage. This can be seen in the result, particularly in information J where in the number of request is 65 and it falls under 'very often' meaning, there are multiple clients requesting for this content. However, information J consumes a little more than 5MB in storage therefore it is only given a priority number 2.

This is contrary, in terms of number of requests to information N which has only 45 requests tag under its name. But since it consumes only a small amount of storage capacity amounting to 25kB, then it has a priority number of 1.The algorithm was able to demonstrate that the output it gives is not just based on one parameter alone but based on both parameters.

Figure 8 is shows the output priority numbers for each information. The higher the priority number means most likely, the information will not be cached. Like in figure 8, information L, information I and information E are more likely excluded for caching. These information are those with lower number of requests, consumes large storage capacity or a combination of both. However, Information G and Information N will be the priority for caching as they are the most demanded information and consumes less storage space or a combination of both.

The priority number indicates whether information will stay in cached or will be replaced. This can be applied in different scenarios. If a node has a limited number of spaces, say 5 compartments to cache file regardless of its size, then only the five chunks of information with the lowest priority numbers will be cached. This is shown in Figure 9. Only five chunks of information occupied the allotted spaces in cache.

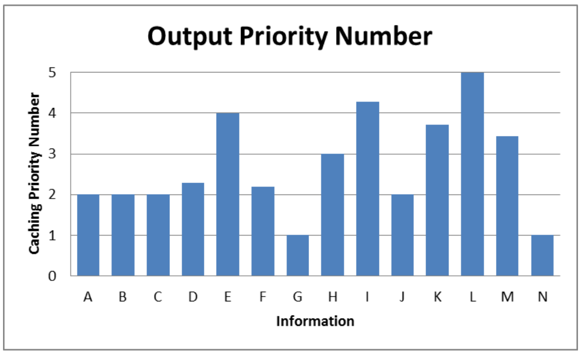

Figure 8: Output of the Fuzzy system

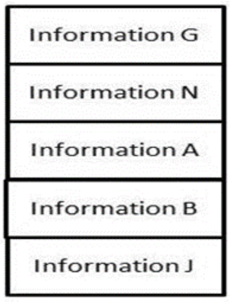

Figure 9: Caching with Limited Number of Compartments

However, if the cache has no limitation in the number of information but only in the total storage capacity, then all information that will be fitted in the storage will be cached, and again the basis is the priority number. So if the total storage is 100MB then all information that will fit in will be cached, the remaining information which do not have a low priority number will not be cached. Figure 10 shows this scenario with a conservative assumption of 16MB caching capacity.

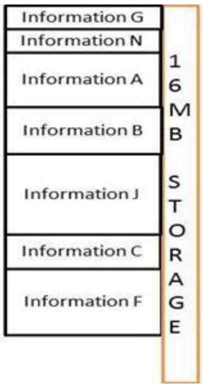

Figure 10: Caching based on storage capacity

In the event that there are same output priority numbers, the one with the greater number of requests will be the priority among those tied. This is to fully realize one objective of ICN which is to spread out the most demanded information all over the network as scattered as possible for all network users to gain easy access.

IX. Conclusion and Recommendation

ICN offers a promising solution to current network problems concerning the internet and its applications. This paper was able to present ICN and the application of fuzzy logic in one of its integral part which is caching.

Fuzzy logic was utilized to create an algorithm to help in the caching part of information centric network. The algorithm successfully demonstrated, using hypothetical data, that fuzzy logic can be used in improving caching dynamics of ICN. It is because of the fact that using fuzzy logic, the decision can be influenced by not only one parameter but both size and number of requests.

Though the simulation was successful, there are some improvements that can be done in the paper. First, if the algorithm will be integrated in actual and running information centric network, then the actual performance of the fuzzy logic system can further be evaluated. Second, since the system is connected to a live network, real time data can be used for further evaluation. In addition, the membership function can further be improved by incorporating other algorithms used in optimization and decision making like genetic algorithm or neural networks. Last, the system can be converted to a smart adaptive system once integrated to live network by taking into account the dynamic nature of the number of requests parameter. Since the paper only focused on the decision making, data gathering and execution should be added in order to make this algorithm adaptive and stand-alone.

References

[1] Y. Zhu and A. Nakao *"A Deployable and Scalable Information-Centric Network Architecture" IEEE ICC 2013 - Next-Generation Networking Symposium*, 2013

[2] Md. Faizul Bari, Shirabur Rahman Chowdhury, Reaz Ahmed "*A Survey of Naming and Routing in Information-Centric Network," IEEE Communications Magazine,vol. 50 no. 12 December 2012, pp44-53*

[3] B. Ahlgren et al., "*A Survey of Information-Centric Networking," IEEE Communications Magazine,vol. 50 no. 7 July 2012, pp26-36*

[4] B. Mathieu et al., "Information-Centric Networking: A Natural Design for Social Network Applications*," IEEE Communications Magazine,vol. 50 no. 7 July 2012, pp44-51*

[5] B. Trossen et al., "*Designing and Realizing an Information-Centric Internet," IEEE Communications Magazine, vol. 50 no. 7 July 2012, pp60-67*

[6] M. Reed, *"Traffic engineering for information-centric networks" IEEE ICC 2012 - Next-Generation Networking Symposium*, 2012

[7] I. Psaras et al., "*Modelling and Evaluation of CCNCachingTrees*," Networking 2011, ser. Lecture Notes inComputer Science, vol. 6640, Springer Berlin/Heidelberg,2011, pp. 78–91.

[8] V. Sourlas et al, "*Distributed Cache Management in Information-Centric Networks" IEEE Transactions on Network and Service Management, vol. 10, no. 3, September 2013, pp286-299*

[9] S. Hassan et. al, "*On the Cache Prformance of the Information Centric Network"* Malaysia, 2013

[10] V. Sourlas and L. Tassiulas *"Effective Cache Management and Performance Limits in Information-Centric Networks"* Greece, 2013

[11] V. Sourlas et. al *"Cache-Aware Routing in Information-Centric Networks"2013*

[12] I. Psaras et. al, "In-Network Cache Management and Resource Allocation for Information-Centric Networks", May 2013

[13] S. Saha et. al, *"Cooperative Caching through Routing Control inInformation-Centric Networks"* 2013

[14] Z.Ming et al, *"Age-based Cooperative Caching in Information-Centric Networks"* IEEE INFOCOM 2012 Workshop on Emerging Design Choices in Name-Oriented Networking, 2012

[15] K. Katsaros et.al, *"MultiCache: An Incrementally Deployable Overlay Architecture for Information-Centric Networking"*2010

[16] Yasuhiko Dote *"Introduction To Fuzzy Logic"* Japan, 1995

[17] Phan Van Su et. al *A New Fuzzy Logic Application to Variable Step Size LMS Algorithm for Adaptive Antennas in CDMA Systems*, Korea

[18] Jungang Liu and Oliver *Using Fuzzy Logic Control to Provide Intelligent Traffic Management Service for High-Speed Networks*

[19] E. Jammeh et alInterval *Type-2 Fuzzy Logic Congestion Control for Video Streaming Across IP Networks.*

[20] D. Hudson and M. Cohen *"Fuzzy Logic in Medical Experts" IEEE Engineering in Medicine and Biology* 1994.