New Efficient Caching Strategy Based on Clustering in Named Data Networking

Nour El Houda Fethellah, Hafida Bouziane, Abdallah Chouarfia [✉]
Université des Sciences et de la Technologie d'Oran- Mohamed Boudiaf, Oran, Algérie
`hafida.bouziane@univ-usto.dz`

Abstract—The Named Data Networking NDN is one of the most proposed architecture for the new model of Internet communications based on contents distribution, called Information-Centric Network ICN. It is widely accepted by the research community since it has become dominant in ICN design that resolves TCP-IP based Internet problems such as bandwidth, delay, location dependent and congestion. Based on location host IP addresses, TCP-IP designed for Peer-to-Peer communication P2P. NDN architecture is oriented Content Centric Networking CCN, where the data is stored on routers and distributed to users from the nearest router. Cache capacities of routers are limited compared to forwarded contents. To move from TCP-IP model to CCN model, many papers propose several new contents distribution based architecture ICN. In this paper, we propose a novel strategy to optimize the use of network resources inspired from Network clustering and cluster head selection in MANETs. Specifically, the improved K-medoids cluster algorithm is used to divide the global network in clusters, where for each cluster; three routers are selected as content routers. The first is the main caching router as well as the second and the third are the secondary caching router. The caching router selection process relies on three relevant criteria consisting of the distance between a node and its cluster centroid, the number of neighbors, and the congestion level. Two Multi Attribute Decision–Making methods MADM are applied, namely TOPSIS and AHP. Performance analysis of our proposed strategy with the established criteria showed its effectiveness and strong potential.

Keywords—Named Data Networking, Content Centric Networking, Information-Centric Network, Caching router, Clustering methods, Analytic Hierarchy Process

1 Introduction

Roughly speaking, TCP/IP is used for distribution content instead of host to host communication. In addition, users are interested by getting the content in any time as fast as possible rather than where it come from. Since, the current architecture is not designed for this purpose, researchers have proposed a new network architecture based on in-network caching called Information Centric Networking ICN [1]. Hence, such architecture aims to satisfy the network requirements such as decrease network

traffic, ease the access user latencies and resolution of server bottleneck problem [2]. During the last years, ICN has attracted much interest as the most advanced architecture which has lead to numerous suggestions such as Data Oriented Network Architecture DONA [3], Publish Subscribe Internet Routing Paradigm PURSUIT [4], Scalable and Adaptive Interest Solutions SAIL [5] and Named Data Networking or Content Centric Networking NDN/CCN [6]. The NDN architecture is the most promising architecture among all the other ICN suggestions. It is a caching network in which nodes store contents satisfying the future requests. The content is duplicated in the network which raise it accessibility, reduce the time of reply and alleviate the node charge. Furthermore, it reduces the traffic in the network and optimizes the use of resources.

Several studies focused on routing and forwarding strategies [2] in NDN architecture in order to suggest a full solution to ease its implementation in the real world. However, the majority of them considers the network as one unit, whereas, another possibility consists to divide the network in clusters which is an approach of paramount importance and still rarely explored. Clustering gives more flexibility to the network and allows a highly effective manipulation. In our view, this effectiveness cannot be achieved without an optimal clustering and an optimal selection of caching routers. The cluster head is the main caching router in the cluster, so, its selection is a critical task for that purpose since random selection is ineffective. We have to select also two other routers as first and second caching routers.

Thus, to ensure the best selection of the three caching routers by the TOPSIS decision method; we choose three pertinent criteria to evaluate each node in the cluster. The node with the least congestion level, the maximum of neighbors and the closest to the cluster centroid will be selected as the main caching router of its cluster. We assume that all routers have the same powerful and the same cache memory. Simultaneous consideration of all these criteria in main caching router selection is a hard task that can be resolved by using Multiple Attribute Decision-Making approaches [7]. In this paper, we retain TOPSIS and AHP as MADM approaches in order to select the best nodes in the cluster as caching routers. TOPSIS method has been successfully applied in various fields such as, engineering, economic and industrial problems based on decision-making [8]. This method quantitatively ranks the alternatives based on their multiple criteria from the best alternative to the worst.

Since TOPSIS is used with different criteria weights, thus we evaluate the impact of the affected weights for the main caching router selection. To this end, AHP has been also adopted to calculate different values of the criteria weights which then will be fed into TOPSIS method to classify the nodes. It is worth noticing that the selection of the right criteria weights values is critical since they have a great influence on the nodes rank in the cluster, the right choice lead to the best selection and inversely.

To do so, we have conducted several experiments, using different scenarios and different criteria weights values to ensure a best selection of the nodes and high performance of the cluster. Our main idea consisted on first, affecting the weights values to criteria in descending order and next, placing in the first position the number of neighbors, in the second position the distance from the cluster centroid and in the third position, the congestion level.

The reminder of the paper is organized as follows. Section 2 provides a brief overview of the NDN architecture. Section 3 presents the related work and section 4 describes our proposed approach. Section 5 reports and discusses the experimental results, and section 6 concludes the paper.

2 Brief NDN Overview

The Named Data Networking NDN also known as Content Centric Networking CCN is a novel proposed network architecture by Van Jacobson and his team in 2009 [6]. The NDN architecture is a part of the Information Content Networking ICN paradigm which relies on content-based connectivity rather than the traditional host to host communication. The nodes in NDN store content in the network to increase the data availability, keep data closer to the users, reduce the replay time and economies the use of resources, and relive the burden on the servers.

In the current architecture, IP addresses are used to locate the hosts, however, in NDN architecture, hierarchical names are used to identify contents. The names structure is similar to the URLs structure (for example /usto/mi/course/network- security.pdf).

There are two types of exchange packets in NDN architecture: Interest and Data. The consumer uses Interest packet to ask for content. The interest contains the request and the data producer responses with data packet contain the requested content. The data producer can be a node in the network or a server [9].

Each NDN router has three basic components see Figure 1: Content Store CS: a cache or storage memory holds contents, Pending Interest Table PIT: a table contains the name and the interfaces of requests not satisfied yet, Forwarding Information Base FIB: a table has the prefix and the interface to forward the interest [10].

An Interest is sent when a user asks for content. When, the interest reaches the router from one of its interfaces, the router checks the content store if the requested content exists, then the router sends a copy in a data packet through the same arrival interface. Else, the router looks for the interest in its pending interest table. If the Interest exists in the PIT then, the router adds the arrival interface and rejects the Interest packet. If not, the router creates an entry in the PIT for this Interest and adds the arrival interface to it and then the router resends the interest to the next hope via an interface according to its FIB. Once the data packet comes at the router, it will be transmitted to all the interfaces that requested this Interest in the corresponding PIT entry and removes that entry, and then caches the content in the content store according to certain strategies [6].
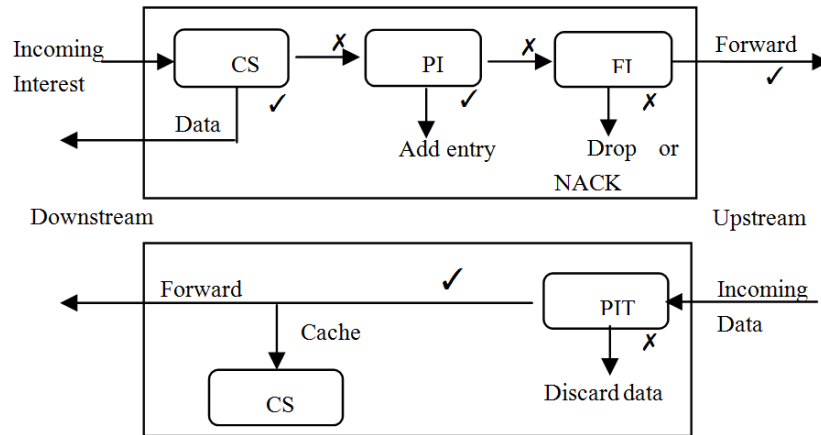
Fig. 1. Interest and data processing in NDN node

3 Related Work

In the beginning, the interest forwarding strategy in NDN architecture was naïve since it was mainly based on flooding the network. When an interest comes at the router, it was retransmitted through all its interfaces to next hop and so on, until the interest packet reaches the requested content. The data packet will take the opposite way that was taken by the Interest packet. Thereafter, the protocol Open Shortest Path First for Named Data Networking OSPFN was proposed [11]. It was the adapted version of Open Shortest Path First OSPF protocol which was designed for the actual architecture. Then, another routing protocol was proposed, called Named-data Link State Routing Protocol NLSR [12]. It was designed especially for the NDN architecture as the default protocol. The NLSR communication is built only on interest and data packet.

There are many propositions of forwarding strategies in NDN architecture, for example: Probability-based Adaptive Forwarding Strategy [13], Entropy-based Probabilistic Forwarding Strategy [14], Stochastic Adaptive Forwarding SAF strategy [15] and Bloom Filter-based Routing approach for information – centric networking BFR [16].

As it has previously mentioned, the majority of the proposed routing protocols and the forwarding strategies in NDN architecture take the entire network as one unit. However, there is a better approach to manage the network in NDN architecture and it is rarely explored. Clustering approach allows scaling a large network into numerous smaller autonomous groups furthermore, the operational functions and message exchanges inside a cluster offer high efficiency and scalability. Clustering is common in network design to manage distributed entities in a network, such as Wireless Sensor Networks WSN [17] and Mobile Ad hoc NETworks MANET [18].

We cannot benefit of the advantages of clustering without the right main caching router selection. To the best of our knowledge, there are no previous studies on main

caching router selection in NDN architecture. Thus, the aim of our paper is to ensure an effective main caching router selection in this architecture.

Before selecting caching routers, we have to divide the network into clusters using the improved K-medoids cluster algorithm [19] which is one of the rare works on clustering in NDN architecture. The Euclidean distance in the classical clustering algorithm is replaced by a new distance based on the delay, the bandwidth, the cache size and the number of hops. This distance allows a reasonable and more real clustering of the network.

4   Proposed Approach

First, the clustering step is performed using the improved K-medoids cluster algorithm [1], then the MADM methods, AHP and TOPSIS are applied. AHP method evaluates the criteria weights and TOPSIS method evaluates the alternatives solutions. These methods solve the problem of caching routers selection.

4.1   Clustering

Before selecting the three caching routers by TOPSIS decision method, we need to cluster the network shown in Figure 2.
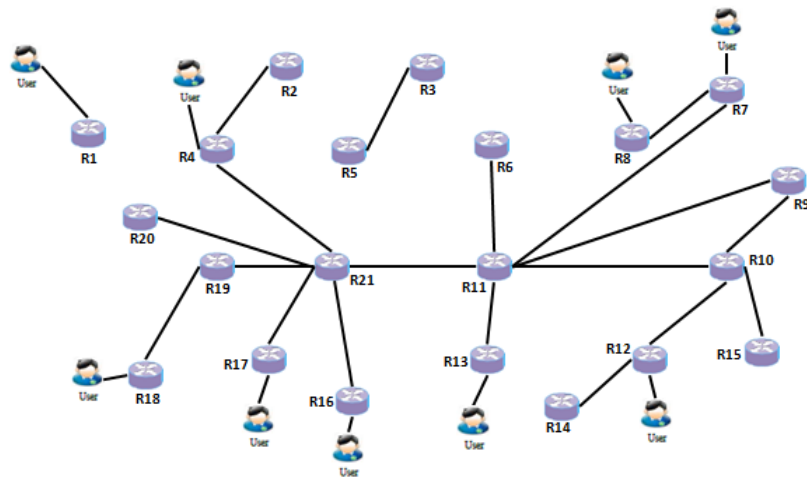

Fig. 2. Initial network

The initial network is represented as a graph G = (N, L), where N denotes the nodes (routers) in initial network G and L denotes the links between the nodes.

The improved K-medoids cluster algorithm is applied to cluster the initial network in K clusters. The initial network has twenty–one nodes. A node or router in cluster is noted Ri where i is the node number. We give as input the following settings:

1. The number of desired clusters K = 2,
2. The size (number of nodes) in the initial network is | N | = 21.

In the first step, the improved K-medoids algorithm deletes systematically the nodes with only one link. So, R1, R2, R3, R5, R6, R14, R15, R20 are deleted. In the second step, K nodes with the maximum of links are selected as initial clusters centroids. One centroid for each cluster (R11 for cluster 01 and R21 for cluster 02). In the third step, the distance between each node in the pruned network and the cluster centroids is calculated, and each node is assigned to the closest centroid. Once the clusters are formed, the fourth step consists in averaging the distance for each node in the given cluster and the node with the minimum distance will be the new centroid. The two latest operations are repeated until the stationary is achieved (the centroid becomes unchanged). Thereby, the cluster centroid is the closest node to all the other nodes in the cluster and the main node or the cluster head if necessary.

The improved K-medoids algorithm gives as output two clusters as it is shown in Figure 3:

1. Cluster 01: (R7, R8, R9, R10, R11, R12, R13)
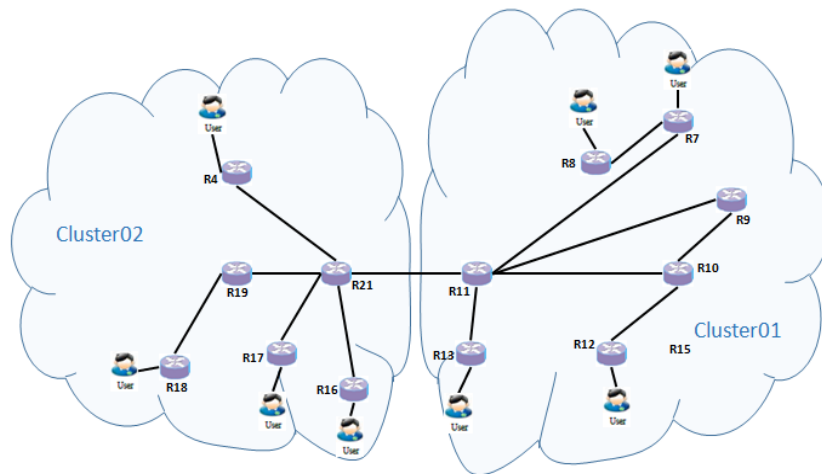2. Cluster 02: (R4, R16, R17, R18, R19, R21)

Fig. 3. Clustered network

4.2 AHP method

Analytic Hierarchy Process method (AHP) [20], [21] is a multi-criteria decision method considering several criteria in order to make the best decision. It is useful to determine the relative importance of a set of criteria in a multi-criteria decision problem. it has been applied in several domains since its development [22].

The procedure for this method is presented in the following steps [23]:

1. Establish hierarchical structure then perform a pair-wise criteria comparison using a scale of relative importance. The relative scale of importance is described in Table 1. The matrix A is the result of a pair-wise comparison of n criteria. Each element a_{ij} indicates the importance of a criterion i compared to criteria j. The value 1 is always assigned to a criterion compared to it.

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix}$$
$$a_{ji} = 1 / a_{ij}, a_{ij} \neq 0 \tag{1}$$

2. Calculate the relative normalized weight (Wi) of each criterion as defined in (2) and (3), and then calculate the vector X as defined in (4).

$$W_i = \frac{GM_i}{\sum_{j=1}^{j=n} GM_i} \tag{2}$$

$$\text{where } GM_i = \left(\prod_{J=1}^{N} a_{ij} \right)^{1/n} \tag{3}$$

$$X = A*W = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \cdots \\ W_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{bmatrix} \tag{4}$$

3. Calculate the coherence values (CVi) as defined in (5), and then average the coherence value that is the maximum Eigenvalue λ_{max} as defined in (6).

$$CV_i = \frac{c_i}{Wi} \tag{5}$$

$$\lambda_{max} = \frac{\sum_{i=1}^{i=n} CV_i}{n} \tag{6}$$

4. Calculate the consistency index (CI) as defined in (7).

$$(CI) = (\lambda_{\max} - n)/(n-1) \tag{7}$$

5. Obtain the random number of criteria (RI) from Table 2. Calculate the consistency ratio (CR) as defined in (8). If final consistency rate exceeds the upper value 0.1, the evaluation procedure must be repeated.

$$CR = CI/RI \tag{8}$$

Table 1. Scales of relative importance

Relative importance (a_{ij})	Description
1	Equal importance of i and j
3	Moderate importance of i over j
5	High importance of i over j
7	Very important of i over j
9	Absolute importance of i over j
2,4,6,8	Intermediate values
1/a_{ij}	If a criterion i have any of the preceding numbers when compared to criterion j. So, the criterion j has the reciprocal value when it is compared to criterion i.

Table 2. Random index values

Criterion	RI
2	0
3	0.58
4	0.90
5	1.12
6	1.24
7	1.32
8	1.41
9	1.45
10	1.49
11	1.51

4.3 TOPSIS method

Technique for Order of Preference by Similarity to Ideal Solution TOPSIS [24],[25] is one of multiple attribute decision-making methods used in many real-world problems. This technique allows to classify alternatives in decreasing order from best to worst, so the best alternative would be the one that is closest to the ideal positive solution and the farthest from the ideal negative solution. The ideal positive solution is a solution that maximizes the benefit criteria and minimizes the cost criteria in the opposite sense for the ideal negative solution.

The TOPSIS procedure is presented in the following steps:

1. Establish a decision matrix D in which the criteria or attributes $(C_1, C_2, ..., C_n)$ are indicated by the columns, the competing alternatives $(A_1, A_2, ..., A_m)$ are indicated by the rows and the weights $(W_1, W_2, ..., W_n)$ are the values assigned to the criteria. An element x_{ij} of the decision matrix D indicates the evaluation of the performance of the ith alternative A_i, with respect to the jth criterion C_j, as indicated in the matrix (9).

$$
\begin{array}{cccccc}
 & C_1 & C_2 & \ldots & C_j & \ldots & C_n \\
A_1 & \left[\begin{array}{cccccc} x_{11} & x_{12} & \ldots & x_{1j} & \ldots & x_n \\ \right. \\
A_2 & x_{21} & x_{22} & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
A_j & x_{i1} & \ldots & \ldots & x_{ij} & \ldots & x_{in} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
A_m & x_{m1} & x_{m2} & \ldots & \ldots & \ldots & x_{mn} \left. \right] \\
\end{array}
\tag{9}
$$

$$
\begin{array}{cccccc}
W_1 & W_2 & \ldots & W_j & \ldots & W_n
\end{array}
$$

2. Calculate the normalized value r_{ij} for each element x_{ij} of the decision matrix D. The normalized value r_{ij} of x_{ij} is calculated as defined in the equation (10):

$$
r_{ij} = \frac{x_{ij}}{\sqrt{\displaystyle\sum_{i=1}^{i=m} x_{ij}^{2}}}, i = 1, 2, ..., m; j = 1, 2, ..., n
\tag{10}
$$

3. Calculate the normalized weighted value v_{ij} by multiplying the normalized decision matrix by its associated weights that are obtained by applying the AHP method as defined in (11).

$$
v_{ij} = W_j r_{ij}
\tag{11}
$$

4. Determine the ideal positive solutions (v^+) as shown in (12) and the ideal negative solutions (v^-) as shown in (13).

$$
v^{+} = \left\{ \left(\overset{max}{\sum_{i}} v_{ij} \middle/ j \in J \right), \left(\overset{min}{\sum_{i}} v_{ij} \middle/ j \in J' \right) \middle/ i = 1, 2, ..., m \right\}
\tag{12}
$$

$$
v^{-} = \left\{ \left(\overset{min}{\sum_{i}} v_{ij} \middle/ j \in J \right), \left(\overset{max}{\sum_{i}} v_{ij} \middle/ j \in J' \right) \middle/ i = 1, 2, ..., m \right\}
\tag{13}
$$

Where, J = (j = 1, 2... N) / j is a set of the biggest criteria

and J '= (j = 1, 2... N) / j is a set of the smallest criteria.

5. Calculate the separation measures for each alternative Ai. The separation of each alternative from the ideal positive solution (S+) as shown in (14) and the ideal negative solution (S-) as shown in (15).

$$S^+_i = \sqrt{\sum_{J=1}^{n} {v_{ij} - v_j^{+}}^2} \quad i = 1, 2, ..., m \tag{14}$$

$$S^-_i = \sqrt{\sum_{j=1}^{n} {v_{ij} - v_j^{-}}^2} \quad i = 1, 2, ..., m \tag{15}$$

6. Calculate the value of the relative proximity of the ideal solution (Ci) as shown in (16) for each alternative Ai. Then rank the alternatives in descending order of the values of the relative proximity coefficient (Ci).

$$C_i = \frac{S_i^-}{S_i^+ + S_i^-} \tag{16}$$

4.4 Experimental evaluation

In this part, we present our experiments conducted to ensure that the best node in a cluster is selected as main caching router and this is due to the huge responsibly and the critical role that it plays as cluster manager. In addition, to prevent cluster paralysis in case of failure of the main caching router, two other nodes in the same cluster are selected to be respectively second and third caching routers. The second caching router takes the control if the main caching router breakdown or if its congestion level up to a critic level or its caching memory up to a critic size. If the second caching router is also failing, the third caching router becomes the main caching router.

In order to select the best nodes as first, second and third caching routers, we choose relevant criteria consisting in the congestion level (cl), the number of neighbors (nbrn) and the distance (dis) between a node and its cluster centroid. So, the node with the minimum congestion level and the maximum neighbors, and the closest to the cluster centroid will be the main caching router.

To ensure this simultaneously, we are facing a multiple attribute decision-making problem, and to try to solve it we have chosen the most effective and the most adequate MADM methods explained in the previous section. AHP method is used to calculate the criteria weights and the TOPSIS method ranks the nodes from the best node to the worst. In our proposition, the three first best nodes are selected.

The decision matrix used to apply TOPSIS method is presented in Figure 4. We present the decision matrix only for the cluster 01. The rows represent the nodes in the cluster 01 and the columns represent the three criteria, namely, congestion level, number of neighbors and distance from the cluster centroid. The criteria weights are presented in the follow form W (w1, w2, w3) where w1 is the first criterion (conges-

tion level), w2 is the second criterion number of neighbors) and w3 is the third criterion (distance to the cluster centroid).

$$
\begin{array}{c}
\begin{array}{ccc} cl & nbrn & dis \end{array} \\
\begin{array}{c} R07 \\ R08 \\ R09 \\ R10 \\ R11 \\ R12 \\ R13 \end{array}
\left[
\begin{array}{ccc}
0.3 & 2 & 1.9 \\
0.55 & 1 & 5.076 \\
0.47 & 2 & 1.9 \\
0.45 & 3 & 1.9 \\
0.65 & 5 & 0 \\
0.24 & 1 & 5.076 \\
0.4 & 1 & 1.9
\end{array}
\right]
\end{array}
$$

Fig. 4. Cluster 01: normalized decision matrix

Furthermore, we have not only applied the TOPSIS method, but also we have study the ranking of the TOPSIS method with all the possible weights. The weights selection is very important and cannot be random. The criteria weights have great impact and influence on the rank of nodes as it shown in the results reported in Table 3 that shows all cases of cluster 01.

- Case 01: all the criteria have the same importance W (0.3333, 0.3333, 0.3333)
- From case 02 to case 04: two criteria have the same value (0.4545) and the third criterion has less importance (0.0909)
- From case 05 to case 10: three values are distinguished, high importance (0.7370), medium importance (0.1863) and less importance (0.0768)

Table 3. Criteria weights

Rank criteria	case 01	case 02	case 03	case 04	case 05	case 06	case 07	case 08	case 09	case 10
Cl	0.3333	0.4545	0.4545	0.0909	0.7370	0.7370	0.1863	0.0768	0.1863	0.0768
Nbrn	0.3333	0.4545	0.0909	0.4545	0.1863	0.0768	0.7370	0.7370	0.0768	0.1863
Dis	0.3333	0.0909	0.4545	0.4545	0.0768	0.1863	0.0768	0.1863	0.7370	0.7370

5 Results and Discussion

The results obtained for cluster 01 reported in Table 4, which represent the rank of the nodes in descending order from the best node to the worst node and the criteria weights (W) in each case, calculated by the AHP method. The table structured as follow: the first column presents the rank, the second to eleventh columns present the ranked nodes in all the cases and the last column shows the node's role selection. The node in the first position is the main caching router, while the nodes in the second and third positions are the second and the third caching routers, respectively.

For example, the seventh and eighth columns show the result of the cases 05 and 06, they give us as a classification (R7> R12> R13> R10> R9> R11> R8). In this case, R7 is the main caching router, while R12 and R13 nodes are designed as second and third caching routers, respectively.

The case 03 gives as rank of nodes (R7> R11> R13> R10> R9> R12> R8). R7 is thus selected as main caching router, it has the minimum congestion level but it is not the closest node to the centroid neither the node with the maximum of neighbors. R11 is selected as the second caching router; it has the maximum number of neighbors and the smallest distance from the centroid. R13 is selected as the third caching router, although there are other nodes better than this node in congestion level, number of neighbors and distance from the centroid.

The cases 05 and 06 give as rank of nodes (R7> R12> R13> R10> R9> R11> R8). In these cases, the main caching router is the node R7. It has the minimum congestion level but it has not the smallest distance to the centroid neither the node with the maximum of neighbors. R12 is one of the worst nodes in the cluster. It has the furthest distance from the centroid and the minimum of neighbors. R12 and R13 are respectively selected as the second and third caching router. However, there are other nodes better than R13 in congestion level, number of neighbors or distance to the centroid.

The case 02 gives as rank of nodes (R11> R10> R7> R12> R9> R13> R8). R12 is ranked well than the nodes R9 and R13. R9 has a smaller distance to the centroid than R12 and has a larger number of neighbors than R12. R13 has a smaller distance to the centroid than R12.

The case 07 gives as rank of nodes (R11> R10> R7> R9> R12> R13> R8). R12 is ranked well than R13, although R12 has a larger distance to the centroid than R13.

The case 09 gives as rank of nodes (R11> R7> R10> R9> R13> R12> R8). R11 is thus selected as a main caching router; it has the maximum of neighbors and the minimum distance to the centroid. R7 is selected as the second caching router; it has fewer neighbors than R10. The node R10 is selected as the third caching router, it is more congested than R7, whatever R7 and R10 have the same distance to the centroid.

The cases 01, 04, 08 and 10 give the same rank of nodes (R11> R10> R7> R9> R13> R12> R8). This four cases rank all the nodes from the first to the last node correctly. These cases rank the nodes in the order with the minimum congestion level, the maximum number of neighbors and the minimum distance to the centroid.

As it is illustrated above, applying the TOPSIS method is not sufficient to have the best rank. The different cases lead to different ranking. However, we need to choose the right case to benefit from the advantages of TOSPIS method and obtain the best rank of nodes from the first node to the last. Thus, the criteria weights may influence the rank of nodes. The best choice of criteria weight leads to a best selection and the worst choice to a worst selection. Therefore, the choice of criteria weights is also very important to have the best ranking. Accordingly, the criteria weights cannot be chosen randomly.

The experimental results revealed that, the best choice of criteria weights consists in giving high importance to the number of neighbors followed by the medium importance to the distance between a node and its centroid and at last the small importance to the congestion level. We found that this situation is the most appropriate

among all the studied scenarios. This is well observed in case 08, where we have low importance congestion level, high importance number of neighbors and medium importance distance. The nodes are thus ranked nodes as follows (R11> R10> R7> R9> R13> R12> R8). All the nodes are ranked perfectly in the order by respecting the conditions on the criteria. The least congested node with the maximum number of neighbors and the minimum distance to the centroid is selected as main caching router. The second and third better nodes are selected as second and third caching routers. Finally, the analysis of our proposed strategy optimizing the use of network resources confirms its effectiveness that relies mainly on the established criteria and the best selection of the criteria weights.

Table 4. Rank of nodes

Rank	case 01	case 02	case 03	case 04	case 05	case 06	case 07	case 08	case 09	case 10	Role selection
01	R11	R11	R7	R11	R7	R7	R11	R11	R11	R11	Main caching router
02	R10	R10	R11	R10	R12	R12	R10	R10	R7	R10	Second caching router
03	R7	R7	R13	R7	R13	R13	R7	R7	R10	R7	Third caching router
04	R9	R12	R10	R9	R10	R10	R9	R9	R9	R9	/
05	R13	R9	R9	R13	R9	R9	R12	R13	R13	R13	/
06	R12	R13	R12	R12	R11	R11	R13	R12	R12	R12	/
07	R8	R8	R8	R8	R8	R8	R8	R8	R8	R8	/

6   Conclusion

This paper proposes a novel strategy to optimize the use of network resources in NDN architecture, inspired from Network clustering and cluster head selection in MANETs. For this purpose, two Multi Attribute Decision–Making (MADM) methods are used, namely, Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and Analytic Hierarchy Process (AHP). The first method is applied for the selection of the three best nodes as caching routers of the cluster, using three criteria consisting in congestion level, number of neighbors and distance to the centroid. The second multi-criteria method is applied for studying the impact of the criteria weights on the result. From a practical point of view, our proposition allows an effective optimization of the network resources. The content can be retrieved directly from the three selected caching routers. Then, the retrieval delay can be significantly reduced. This paper has also discussed the issues and challenges of caching memory routers in NDN architecture. The latest is still regarded as a new paradigm which can bring a revolution in the internet usage; however, the subject requires a cautious handling. As future research directions, we will:

- Add others criteria in the clustering step and in the selection of the caching routers such as the rooter's power and the cache size
- Take into account how to manage contents in ICN network caches and coordinate between caching memory routers
- What happens if the original content changes on the source node?

7 Acknowledgement

8 References

[1] G. Xylomenos et al, "A Survey of Information-Centric Networking Research", *IEEE Communications Surveys and Tutorials*, vol. 16, no. 2, pp. 1024–1049, 2014.

[2] M. Zhang, H. Luo, and H. Zhang, "A Survey of Caching Mechanisms in Information-Centric Networking", *IEEE Communications Surveys and Tutorials*, vol 17, no 3, pp. 1473–1499, 2015. https://doi.org/10.1109/comst.2015.2420097

[3] T. Koponen et al., "A Data-oriented (and Beyond) Network Architecture", in *Proceedings of the SIGCOMM'07 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pp. 181-192, 2007, New York, NY, USA.

[4] N. Fotiou, P. Nikander, D. Trossen, and G. C. Polyzos, "Developing Information Networking Further: From PSIRP to PURSUIT", in *Broadband Communications, Networks, and Systems*, vol. 66, pp. 1-13, 2012, I. Tomkos, C. J. Bouras, G. Ellinas, P. Demestichas, and P. Sinha, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30376-0_1

[5] Scalable Adaptive Internet Solutions : SAIL'. Available: https://sail-project.eu/wp-content/uploads/2013/02/SAIL_EFRAIM_PressRelease_130213-b-.pdf, Accessed on: May 19, 2018

[6] V. Jacobson et al, "Networking Named Content", in *Proceedings of theCoNEXT'2009: 5th International Conference on Emerging Networking Experiments and Technologies*, pp. 1-12, 2009, New York, NY, USA.

[7] C.-L. Hwang and K. Yoon, "Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey". *Springer Science & Business Media*, 2012.

[8] M. Behzadian, S. Khanmohammadi Otaghsara, M. Yazdani, and J. Ignatius, "A state-of the-art survey of TOPSIS applications", *Expert Systems with Appications*, vol. 39, no. 17, pp. 13051–13069, Dec. 2012. https://doi.org/10.1016/j.eswa.2012.05.056

[9] L. Zhang et al., "Named data networking", *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66-73, 2014.

[10] V. Jacobson et al., "Named Data Networking (NDN) Project 2012 - 2013 Annual Report", CAIDA, 2013. [Online]. Available: http://www.caida.org/publications/papers/2013/named_data_networking_2012-2013/index.xml. [Accessed: 17-Sep-2018].

[11] L. Wang, A. Alyyan and B. Zhang, "An OSPF Based Routing Protocol for Named Data Networking", *NDN Technical Report NDN-0003*, July 2012, https://www.named-data.net/techreport/TR003-OSPFN.pdf

[12] A. K. M. M. Hoque et al, 'NLSR: "Named-data Link State Routing Protocol", in *Proceedings of the 3rd ACM SIGCOMM Workshop on Information-centric Networking*, New York, NY, USA, pp. 15-20, 2013. https://doi.org/10.1145/2491224.2491231

[13] H. Qian, R. Ravindran, G. Wang, and D. Medhi, "Probability-based adaptive forwarding strategy in named data networking", in *2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, pp. 1094-1101, 2013.

[14] K. Lei, J. Wang, and J. Yuan, "An entropy-based probabilistic forwarding strategy in Named Data Networking", in *2015 IEEE International Conference on Communications (ICC)*, London, pp. 5665-5671, 2015. https://doi.org/10.1109/icc.2015.7249225

[15] D. Posch, B. Rainer, and H. Hellwagner, "SAF: Stochastic Adaptive Forwarding in Named Data Networking", *IEEE-ACM Transactions on Networking.* vol. 25, no. 2, pp. 1089–1102, Apr. 2017. https://doi.org/10.1109/tnet.2016.2614710

[16] A. Marandi, T. Braun, K. Salamatian, and N. Thomos, "BFR: A bloom filter-based routing approach for information-centric networks", in *2017 IFIP Networking Conference (IFIP Networking) and Workshops*, pp. 1-9, 2017. https://doi.org/10.23919/ifipnetworking.2017.8264842

[17] K. Akkaya and M. Younis, "An energy-aware QoS routing protocol for wireless sensor networks", in Proceedings *23rd International Conference on Distributed Computing Systems Workshops*, *Providence*, RI, USA, 2003. pp. 710-715, 2003. https://doi.org/10.1109/icdcsw.2003.1203636

[18] J. Y. Yu and P. H. J. Chong, "A survey of clustering schemes for mobile ad hoc networks", *IEEE Communications Surveys and Tutorials*, vol. 7, no. 1, pp. 32–48, 2005. https://doi.org/10.1109/comst.2005.1423333

[19] C. Li and K. Okamura, "Cluster-based In-networking Caching for Content-Centric Networking", *International Journal IJCSNS*, vol. 14, no. 11, pp. 1-9, 2014. http://paper.ijcsns.org/07_book/201411/20141101.pdf

[20] R. W. Saaty, "The analytic hierarchy process—what it is and how it is used", *Mathematical. Modeling.* vol. 9, no. 3, pp. 161–176, Jan. 1987. https://doi.org/10.1016/0270-0255(87)90473-8

[21] T. L. Saaty, "Decision making with the analytic hierarchy process", *Int. J. Services. Sciences*, vol. 1, no. 1, p. 83, 2008.

[22] W. Ho and X. Ma, "The state-of-the-art integrations and applications of the analytic hierarchy process", *European. Journal of Operational Research*, vol. 267, no. 2, pp. 399–414, Jun. 2018. https://doi.org/10.1016/j.ejor.2017.09.007

[23] L. Socaciu, O. Giurgiu, D. Banyai, and M. Simion, "PCM Selection Using AHP Method to Maintain Thermal Comfort of the Vehicle Occupants", *Energy Procedia*, vol. 85, pp. 489–497, Jan. 2016. https://doi.org/10.1016/j.egypro.2015.12.232

[24] C.-L. Hwang and K. Yoon, "Multiple Attribute Decision Making", vol. 186. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981.

[25] V. B.Vommi and S. R. Kakollu, "A simple approach to multiple attribute decision making using loss functions", Journal of Industrial Engineering International , vol 13, n° 1, pp.107-116, March, 2017. https://doi.org/10.1007/s40092-016-0174-6

9 Authors

Nour El Houda Fethellah Currently is PHD computer science student in the USTO-MB University of Oran in Algeria. Having graduated from the same university with a Master degree in Computer Science Engineering in. She is a member of the research team that focused on networking. nourelhouda.fethellah@univ-usto.dz

Hafida Bouziane is an Associate Professor lecturer at Computer Science Department Faculty of Mathematics and Computer Science of USTO-MB University Oran – Algeria. She received the PhD degree in Computer Science from USTO-MB in 2014. She is Member of LIM (Computer science and mathematics) Laboratory – University

Ibn Khaldoun of Tiaret. Her research interests include Intelligent System, networking, bioinformatics, m-learning. abdallah.chouarfia@univ-usto.dz

Abdallah Chouarfia was born in Algeria, graduated with Engineer Degree in Computer Science from National Institute of Computer Science, Algiers, in 1981. He received the PhD degree in Computer Science from the Paul Sabatier University in Toulouse, France in 1983. He was a Senior Lecturer in the Computer Science Department at Sciences and Technology University, Oran Algeria from 1987-2019. He is currently a full Professor at Science and Technology University, Oran Algeria. His research interests include networking, software engineering, information systems interoperability, web engineering, pervasive systems.