Computers, Materials & Continua CMC, vol.65, no.1, pp.321-336, 2020

A Cache Replacement Policy Based on Multi-Factors for Named Data Networking

Meiju Yu[1], Ru Li[1, *] and Yuwen Chen[2]

Abstract: Named Data Networking (NDN) is one of the most excellent future Internet architectures and every router in NDN has the capacity of caching contents passing by. It greatly reduces network traffic and improves the speed of content distribution and retrieval. In order to make full use of the limited caching space in routers, it is an urgent challenge to make an efficient cache replacement policy. However, the existing cache replacement policies only consider very few factors that affect the cache performance. In this paper, we present a cache replacement policy based on multi-factors for NDN (CRPM), in which the content with the least cache value is evicted from the caching space. CRPM fully analyzes multi-factors that affect the caching performance, puts forward the corresponding calculation methods, and utilize the multi-factors to measure the cache value of contents. Furthermore, a new cache value function is constructed, which makes the content with high value be stored in the router as long as possible, so as to ensure the efficient use of cache resources. The simulation results show that CPRM can effectively improve cache hit ratio, enhance cache resource utilization, reduce energy consumption and decrease hit distance of content acquisition.

Keywords: Cache replacement policy, named data networking, content popularity, freshness, energy consumption.

1 Introduction

In recent years, with the rapid development of the Internet and the explosive growth of Internet users, data traffic in the network has increased dramatically. The main function and objective of the Internet is also changing from the initial pursuit of network interconnection to information sharing and efficient access. The users are more interested in content itself rather than the address of the content [Wang, Kong, Li et al. (2019)]. Therefore, the existing IP network architecture is facing unprecedented challenges. With the emergence of these problems, many research institutes have been devoting themselves to the research and design of the future network. Among these design

[1] College of Computer Science, Inner Mongolia University, Hohhot, 020021, China.

[2] Department of Computer Systems Technology, New York City College of Technology-The City University of New York, Brooklyn, NY 11201, USA.

[*] Corresponding Author: Ru Li. Email: csliru@imu.edu.cn.

Received: 31 March 2020; Accepted: 08 May 2020.

schemes, Named Data Networking (NDN) has become a promising architecture based on contents for future network [Xylomenos, Ververidis, Siris et al. (2014)].

NDN is a content-centric and user-driven network architecture, in which routers has the capability of caching [Jacobson, Smetters, Thornton et al. (2009); Zhang, Afanasyev, Burke et al. (2014)]. Namely, NDN is more concerned with the content itself than the actual physical location of content. By caching the content passing by, routers can provide data services for users. So, users can obtain the content they need from a nearby router instead of going to the original content server. This can greatly alleviate the bandwidth pressure and reduce the content acquisition time.

In NDN, a user requests a content using interest packet with content name. When routers receive an interest packet, they will check their Content Store (CS) for the matched content. If there is a matching content, routers will send the content to user who requests it. Otherwise, they will forward the interest package to the next router according to the Forwarding Information Base (FIB). When a data packet is passing by, routers will cache the content for severing the incoming requests instead of original producer.

Caching strategy is the key research area of NDN, which greatly affect the performance of network [Aggarwal, Nilay and Yadav (2017)]. And there also many researches on cache strategies recently [Liu, Zhang, Ge et al. (2019)]. When the cache space of router is full and new content arrives, it is important to decide which content should be discarded from CS. Hence, considering the limited caching resource of routers, it is crucial to make effective cache replacement policy to improve the utilization of caching resources.

Based upon the multiple factors affecting cache performance, in this paper, we propose a cache replacement policy based on multi-factors which is named CRPM for Named Data Networking to improve the overall performance of NDN in this paper. In CRPM, the content with the least cache value will be evicted from the caching space. The contributions in this paper are summarized as follows: 1. We fully analyze multi-factors that affect the caching performance from different perspectives and puts forward the corresponding calculation methods. 2. A new cache value function is constructed to measure the cache value of contents. 3. A cache replacement policy is proposed which makes the content with high value be stored in the router as long as possible, so as to ensure the efficient use of cache resources.

The reminder of this paper is organized as follows. In Section 2, we present the related work. Then we illustrate the CRPM strategy in detail in Section 3 and we present the performances of our cache replacement policy in Section 4 subsequently. Finally, we conclude the paper in Section 5.

2 Related work

Many researchers have proposed various cache replacement policies to manage the in-networking caches in order to improve the performance of network [Zhang, Luo and Zhang (2019)]. The default cache replacement policies are: First in First out (FIFO), Least Frequently Used (LFU) and Least Recently Used (LRU). In FIFO, the content that first be cached in CS is replaced first. LFU regards the content request frequency as the most important reference factor. It believes that if data has been accessed many times in the past, it will be accessed more frequently in the future. By using the request count of content, the content with the smallest count will be moved out of CS. LFU improves the cache hit rate to some extent. However, this method may lead to the situation that the content with high frequency of access in history has not been accessed recently. Hence, the caching resource has not been effectively utilized because of long-term occupation of caching resources. Considering that the content that has not been accessed for a long time in CS is likely not to be accessed again, LRU regards the time interval of content being accessed as an important reference factor and removes the data that has not been accessed for the longest time. LRU improves the cache hit rate to a certain extent and is relatively easy to implement. It is widely used in current research.

Some researchers argue that in the existing cache replacement policies, such as LFU and LRU, the influence factors of content popularity are not taken into account, resulting in low cache efficiency in dynamic networks [Ran, Lv, Zhang et al. (2013); Xin, Li, Wang et al. (2016)]. Therefore, cache replacement strategies based on content popularity are proposed to remove the content with the smallest popularity from CS. These strategies not only improve the cache hit rate, but also significantly reduce server load and increase network capacity. In order to balance the distribution of content with different popularity in the network, Zhu et al. [Zhu, Mi and Wang (2013)] propose a cache probability replacement strategy based on content popularity, and selected the replacement content according to the characteristics of content popularity distribution. Fan et al. [Fan, Wu, Zhang et al. (2017)] consider that the cache value of contents in the CS are dynamic and increase exponentially with the number of requests. So, the value of content with high popularity is much higher than that of content with low popularity. When the value of content in CS are lower than the threshold value, the space occupied by them are marked as “idle”, and they can still provide services for users before being replaced.

Li et al. [Li, Nakazato, Detti et al. (2015)] analyze the relationship between different blocks of the same video segment. When the video requester requests a video block of a sequence number, the block after the sequence number is more likely to be requested, so it has a higher request probability. In view of the above analysis, Li et al. [Li, Nakazato, Detti et al. (2015)] propose a cache replacement strategy based on future request probability for video content transmission, which removes the content with the minimum request probability in the future. The proposed method increases the average cache hit rate and shortens the average content transmission distance. There are cache replacement strategies based on popularity prediction which can predicts content popularity in CS [Zhang, Tan and Li (2017); Ren, Zhao, Sun et al. (2018)]. For efficient caching in CCN, Ren et al. [Ren, Zhao, Sun et al. (2018)] present a caching framework of Prefix-based Popularity Prediction (PPP), which assigns a lifetime to the prefix of a content name based on its access history or popularity. Zhang et al. [Zhang, Tan and Li (2017)] propose a blocklevel cache replacement method called Predictive Popularity Caching (PPC) to discover the relationship between video blocks from the perspective of user’s watching behavior. PPC predicts and caches the most popular blocks in the future, compares the future popularity of the arrived new content with the minimum future popularity in the CS, and keeps the content with higher value. In order to prevent popular content from being replaced in a short time, a limited-LRU strategy is proposed [Xin, Li, Wang et al. (2016)]. Based on the results of content popularity statistics, this strategy tries to cache popular content copies as long as possible and remove unpopular content copies in CS.

There are also many replacement strategies that take into account a variety of factors and try to improve the cache performance of routers. Some researchers start with the frequency of content being accessed, combining other factors that affect the performance of the replacement strategy. Considering the frequency and size of content access, Ma et al. [Ma, Chen and Zhao (2013).] propose a cache replacement strategy based on priority, which decides which content to replace according to priority. The filtering effect and aggregation of requests in NDN network are discussed in Hu et al. [Hu, Gong, Cheng et al. (2015)]. The cache replacement strategy is formulated considering the frequency of requests and the distance between users. There is a weighted frequency based real-time data replacement policy (WFRRP) is proposed to predict the real-time popularity of content by using the weighted frequency in different time periods and the cost of data requests [Liao, Hu, Wu et al. (2016)]. WFRRP enables real popular content far away from the source server to achieve high real-time popularity and high cache priority. This strategy improves the hit rate of data and reduces the average hop count and transmission delay.

Some of the above policies only consider the improvement of a single factor to a certain performance of the network, and cannot meet the diverse network performance requirements. Therefore, considering the cache content in router from various perspectives, proposing cache replacement function and measuring the cache value of the content are the main ideas of cache replacement policies based on cache value. Currently, there are various methods based on cache value, and the algorithm performance and complexity are also different.

3 Proposed policy

CRPM is a cache replacement policy that utilizes multi-factors to measure the caching value of contents in CS. It can effectively utilize caching resources. In CRPM, we fully analyze content popularity, content acquisition energy consumption, content freshness and the last access time interval and determine the calculation method of these four factors. Then we utilize analytic hierarchy process (AHP) to determine the weight of the important factors affecting cache replacement and construct a new cache value function to measure the cache value of contents. At last, we present a cache replacement policy based on content value to decide which content should be evicted form CS and describe the working details of CRPM. CRPM can achieve a better balance between cache performance and acquisition cost.

In this section, we sketch out the environment and assumption under which CRPM is designed. In addition, the details of our proposed CRPM policy are described. CRPM consists of three main parts, including analysis of the important factors, using AHP to determine weight and CRPM cache replacement policy.

3.1 Analysis of the important factors

In the process of designing cache replacement policy, we consider the attribute characteristics of content itself. The content attribute factors affecting the replacement policy mainly include content popularity, content acquisition energy consumption, content freshness and content hit interval. Before introducing the replacement policy proposed in this paper, we analyze the factors affecting cache replacement strategy from different perspectives and determine the calculation method of these four factors.

3.1.1 Content popularity

As we all know, there are huge amount of contents in the network, but only a small part of contents will frequently be accessed by a large number of users, most of the contents are rarely requested. Consequently, caching popular contents on routers will greatly improve the retrieval efficiency of users. In our strategy, we propose a dynamic content popularity algorithm, in which the content popularity includes three parts: the historical popularity information of content, the number of requests of content in the current period and the future trend of requests changing over time. The proposed algorithm can dynamically reflect the changes in content popularity and provide a more accurate cache basis for content objects arriving at routers.

In this paper, the time period is set to τ and the popularity of content c_i in period k is defined as follows:

$$\mathrm{P}_k(\mathrm{c}_i)=\alpha\times\mathrm{P}_{k-1}(\mathrm{c}_i)+\beta\times\mathrm{LP}_k(\mathrm{c}_i)+\gamma\times T_k(\mathrm{c}_i),\quad k=2,3,\cdots \tag{1}$$

where $\mathrm{P}_{k-1}(\mathrm{c}_i)$ is the popularity of content c_i in period k-1, $\mathrm{LP}_k(\mathrm{c}_i)$ is the relative popularity of content c_i and $T_k(\mathrm{c}_i)$ is the future trend of requests changing about content c_i. α , β , γ are the weight coefficients and $\alpha+\beta+\gamma=1$.

$\mathrm{LP}_k(\mathrm{c}_i)$ and $T_k(\mathrm{c}_i)$ are denoted in Eqs. (2) and (3).

$$\mathrm{LP}_k(\mathrm{c}_i)=\frac{r_k(c_i)\text{-}\min\limits_{c_j\in C}\{r_k(c_j)\}}{\max\limits_{c_j\in C}\{r_k(c_j)\}-\min\limits_{c_j\in C}\{r_k(c_j)\}},\quad k=2,3,\cdots \tag{2}$$

where $r_k(c_i)$ is the number of requests of content c_i in the kth period, $\min\limits_{c_j\in C}\{r_k(c_j)\}$ represents the minimum number of requests of content c_i in the kth period, and $\max\limits_{c_j\in C}\{r_k(c_j)\}$ represents the maximum number of requests of content c_i in the kth period.

$$T_k(\mathrm{c}_i)=\frac{1}{1+\exp\{-\dfrac{r_k(c_i)-r_{k\text{-}1}(c_i)}{r_{k\text{-}1}(c_i)}\}},\quad k=2,3,\cdots \tag{3}$$

where $r_{k-1}(c_i)$ is the number of requests of content c_i in period k-1.

When k=1, the initial values of content popularity for content c_i in the first period are given as follows:

$$\mathrm{P}_1(\mathrm{c}_i)=\frac{r_1(c_i)\text{-}\min\limits_{c_j\in C}\{r_1(c_j)\}}{\max\limits_{c_j\in C}\{r_1(c_j)\}-\min\limits_{c_j\in C}\{r_1(c_j)\}} \tag{4}$$

3.1.2 Content acquisition energy consumption

In this paper, the energy consumption of content acquisition in CS is mainly the energy consumption of content transmission, that is, the transmission cost from the router or server who provide content to the user. It is related to the size of content and the transmission distance. Assume $hop_j^m(c_i)$ represents the number of hops from the content acquisition node to the local node, s_i (bit) is denoted the size of the content c_i and P_l is the transmission energy consumption per bit per hop. The content acquisition energy consumption is defined as follows:

$$E_j(c_i) = Pl \times s_i \times hop_j^m(c_i) \tag{5}$$

Therefore, the higher the energy consumption is, the higher the cache value is, namely, caching the content can consume less energy. The larger the number of hops and the larger the content size, the higher the cost of data transmission. Therefore, our policy avoids replacing the cached content far from the local node.

3.1.3 Content freshness

Because the content replica in CS may be outdated, it is necessary to use the freshness mechanism to check the validity of hitting content. This problem is called cache freshness validity [Hail, Amadeo, Molinaro et al. (2015)]. In this paper, let t is the current time, $t_{ini}(c_i)$ is the time when content is generated by the content source server, and $t_l(c_i)$ is the life cycle of content c_i. Therefore, the freshness of content c_i in r_j is defined as follows:

$$F_j(c_i) = t_l(c_i)\text{-}(t - t_{ini}(c_i)) \tag{6}$$

The fresher the content in the cache, the later it is generated and it is more likely to provide services. Therefore, this paper argues that the content with high freshness has more caching value.

3.1.4 Content hit interval

When the content in CS is not accessed for a long time and still occupies cache resource, it will cause cache pollution [Xu, Wang and Wang (2017)]. Therefore, CRPM considers the interval between current time and the last access time of the content. The longer the interval, the smaller the cache value of the content. Let t denote the current time and $t_{Hit}(c_i)$ denote the last time the content was hit. So, the content hit interval is defined as follows:

$$\mathrm{T}_{\mathrm{inter}}(c_i) = t - t_{Hit}(c_i) \tag{7}$$

3.2 Using AHP to determine weight

In this section, the important factors will be quantitatively analyzed, and each factor will be given the corresponding weight. Because these factors have different effects on the performance of cache replacement policy, and each factor considers different purposes, the importance of these factors are different. Considering the linear requirement of routers for data processing in NDN network, the algorithm of cache replacement policy should not be too complex. We utilize analytic hierarchy process (AHP) to determine the weight of each influencing factor, and establishes the cache value function on this basis.

We will introduce the specific steps of the implementation of the AHP method in this section, including constructing hierarchy structure model, establishing judgement matrixes, calculating weight vectors and consistency check.

3.2.1 Constructing hierarchy structure model

We establish the hierarchical structure model according to AHP, as shown in Fig. 1. The problems to be solved are as follows: selecting the content with the lowest cache value in the CS as the goal layer; four important factors affecting cache replacement are taken as the criterion layer for decision-making; and the lowest level is each content in the CS which serve as the alternative layer.

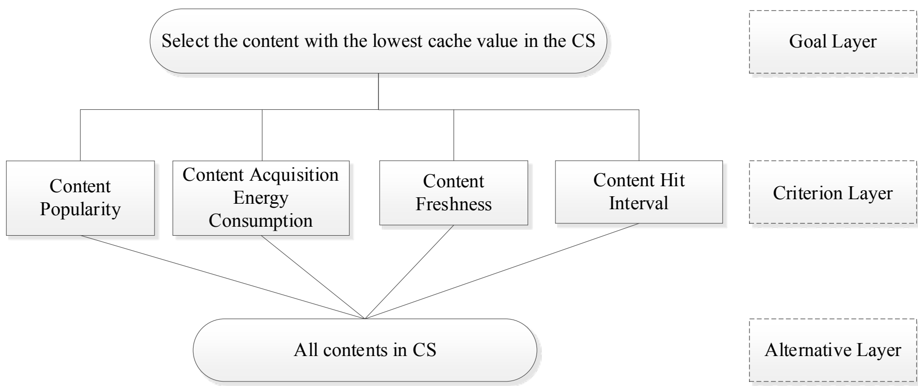

Figure 1: AHP hierarchical structure diagram

3.2.2 Establishing judgement matrixes

Let $G = (G_1, G_2, G_3, G_4)$ is denoted the four factors in criterion layer and compare the four factors in pairs. Let a_{ij} denotes the result of important comparison. The judgment matrix is as follows:

$$A = \begin{bmatrix} 1 & 7 & 3 & 5 \\ 1/7 & 1 & 1/5 & 1/3 \\ 1/3 & 5 & 1 & 3 \\ 1/5 & 3 & 1/3 & 1 \end{bmatrix} \tag{8}$$

The maximum eigenvalues and eigenvectors of the judgment matrix are calculated and the consistency check is carried out in next section.

3.2.3 Calculating weight vectors and consistency check

We calculate the maximum eigenvalue of judgment matrix A and its corresponding eigenvectors utilizing matlab. The maximum eigenvalue is $\lambda_{\max}$ =4.117 , and its corresponding eigenvector is W=(0.8880,0.0869,0.4121,0.1847). The normalized eigenvector is obtained by normalizing the eigenvectors and the value is W'=(0.5650,0.0553,0.2622,0.1175).

Then the consistency check is carried out. The consistency index to be tested is as follows:

$$CI=\frac{\lambda_{\max}\text{-}n}{n-1} \quad (9)$$

where n denotes the order of the matrix A. Consequently, the consistency check value for the problem in this paper is $CI=\frac{4.117\text{-}4}{4-1}=0.039<0.1$. So, there is satisfactory consistency.

In order to the value of CI, we look up the average random consistency index table and find that the corresponding random consistency index is 0.9. By comparing CI with random consistency index, the test coefficient is $CR=\frac{CI}{RI}=\frac{0.039}{0.9}=0.043<0.1$. It is considered that the judgment matrix passes the consistency check.

By using the AHP method, we obtain the weight parameters of content popularity, content acquisition energy consumption, content freshness and content hit interval. They are 0.5650, 0.0553, 0.2622 and 0.1175.

3.3 CRPM cache replacement policy

This section constructs a cache value function to calculate the cache value of the contents in CS, removes the content with the least cache value from the CS, and leaves the cache space for more valuable content. Then we will introduce the description of CRPM cache replacement policy model and its working details. The CRPM cache replacement model mainly consists of the following modules: influencing factor calculation module, influencing factor weight assignment module and cache replacement module. The overall design idea of the cache replacement policy model is shown in Fig. 2.

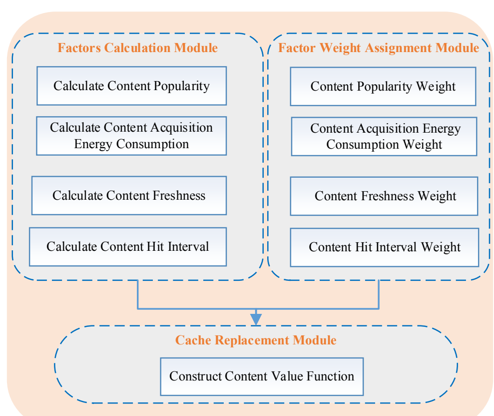

Figure 2: The modules of CRPM

When the content reaches the router and the cache space is full, the replacement policy occurs. It will decide which content of CS will be moved out of CS. CRPM cache replacement policy proposes a cache value function according to the calculation formula of four factors affecting the content in Section 3.2 and the cache weight determined by AHP method. The cache value of each content in CS is calculated, and the content with the lowest cache value is evicted from CS. The definition of the cache value function is defined as follows:

$$\mathrm{CV}(c_i) = \alpha_1 \mathrm{P}_j(\mathrm{c}_i) + \alpha_2 E_j(c_i) + \alpha_3 F_j(c_i) + \alpha_4 \mathrm{T}_{\mathrm{inter}}(c_i) \quad (10)$$

where $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = W' = (0.5650, 0.0553, 0.2622, 0.1175)$.

4 Performance evaluation

In this section, the simulation results of CRPM caching replacement policy are presented, compared and analyzed using ndnSIM [Mastorakis, Afanasyev and Zhang (2017)]. We detail the simulation evaluation, including simulation parameter setting, evaluation metrics and experiment results analysis.

4.1 Simulation parameter setting

In all our simulations, we choose LRU, FIFO and a cache replacement policy based on content popularity (CCP) [Ran, Lv, Zhang et al. (2013)] as comparison strategies with our CRPM policy and record experimental results in 1000 s.

Most of the current networks are complex network topologies. This paper mainly studies the cache replacement policy on routers. Therefore, in the simulation, we use a simple linear topology which includes one user and one server distributed at both ends of the linear topology. There are three routers in the middle as cache nodes. The network topology is shown in Fig. 3.

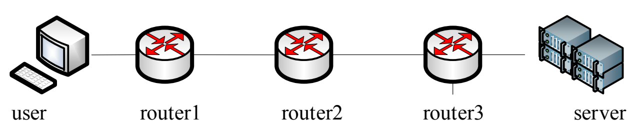

Figure 3: Network topology

The distribution of content requests is very important for our research. In different fields, there are various professional methods to study the distribution of random variables [Mallouli (2019)]. In our simulation, the total content requests in the network have been modeled following a Zipf distribution function [Yang and Zhu (2016)]. In our experiments, there are 1000 different content in the network. We observe the impact of cache capability on cache performance and cache capability refers to the number of contents that can be stored in CS. The cache size of each node ranges from 5 to 30 and the default value is 5. This paper also observes the impact of different Zipf exponents on cache performance through experiments. The frequency of interest packets is 100/s. The request count period is 4 seconds. The simulation configurable parameters are depicted as Tab. 1.

Table 1: Simulation parameter setting

Parameter	Description	Value
n	Number of contents	1000
Request Rate	Number of requests of user	100
Cache Size	Number of contents stored per router	5, 10, 15, 20, 25, 30
τ	Count period (s)	4
s	Zipf exponents parameter	0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5

4.2 Evaluation metrics

In order to effectively improve cache-hit ratio, enhance cache resource utilization, reduce energy consumption and decrease hit distance of content acquisition, we use the following metrics to compare CRPM with LRU, FIFO and CCP Cache replacement policy.

(1) Cache-Hit Rate

The ratio of the total number of interest packets satisfied at all routers to the total number of interest packets arriving at all routers, which indicates the utilization rate of the router's cache resources and the average cache efficiency of the routers. It is calculated as follows:

$$cache-hit\ rate=\frac{\sum_{router} Hit}{\sum_{router} Hit+\sum_{router} Miss} \tag{11}$$

(2) Replacement Frequency

The average number of replacements on a router per second. It is used to define the load degree of router in network. If the number of replacement operations occurs in routers is small, it means that contents stay in the CS for a long time and the cache space is effectively utilized.

(3) Average Cache Hit Distance

The average number of hops that contents have travelled in the network between requesters to the corresponding cache nodes. It is used to define the responding speed of user requests.

4.3 Experiment results and analysis

4.3.1 Cache-Hit Rate

In this experiment, we observe the effect of cache size and Zipf exponents on cache hit rate. The results will be introduced separately.

Fig. 4 is a comparison of cache-hit rates of CRPM, FIFO, LRU and CCP under the change of cache size. The cache-hit rates of the four cache replacement policies are increasing with the increase of cache capacity. Therefore, the more contents the router caches, the stronger the caching capability, the greater the possibility of providing services for users, and the higher the cache hit rate. The cache-hit rate of CRPM policy is always higher than that of the other three policies. The cache hit rate of FIFO is the lowest, always under the other three policies. Compared with FIFO, LRU, CCP and CRPM have significantly improvement in cache-hit rate. When the cache size is small, the cache-hit rate of CRPM is significantly higher than that of LRU and CCP, especially when the cache size is less than 8, the contrast is obvious; when the cache size is increasing, the effect of the cache replacement policy is no longer obvious, so the gap between the three policies is gradually smaller. The result shows that CRPM plays a very good role in improving the hit-rate of NDN, which meets the original intention of designing this policy.

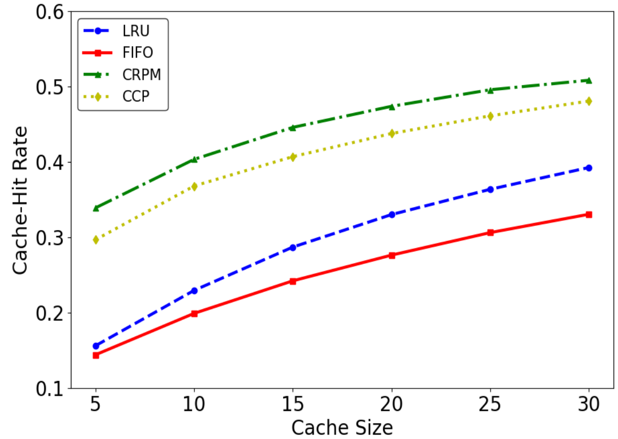

Figure 4: Cache size-cache-hit rate

Fig. 5 is a comparison of cache-hit rates of CRPM, FIFO, LRU and CCP under the change of Zipf exponents when the default cache size is 3. The cache-hit rates of the four cache replacement policies are increasing with the increase of Zipf exponents and FIFO has always been the lowest one. FIFO always replaces the content that first enters the CS, so it is insensitive to popularity and the hit rate increases slowly. While LRU replaces the content that has not been accessed for the longest time and the cache-hit rate is higher than FIFO. Both CRPM and CCP are based on content popularity, so they are more sensitive to the increase of Zipf exponents, and CRPM policy is better than CCP.

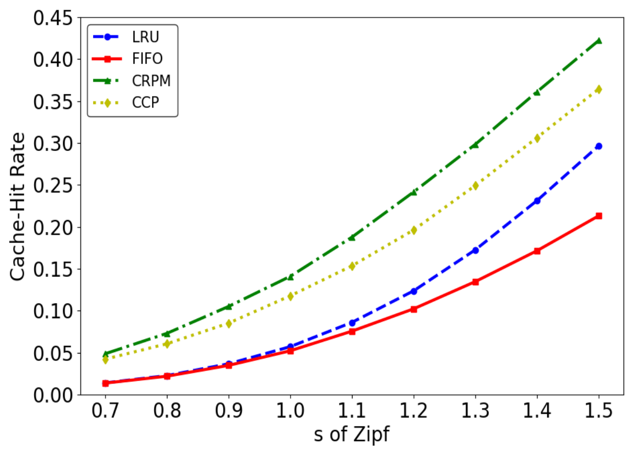

Figure 5: s of Zipf-cache-hit rate

4.3.2 Replacement frequency

In this experiment, we observe the effect of cache size and Zipf exponents on replacement frequency. The results are introduced separately.

Fig. 6 is a comparison of replacement frequency of CRPM, FIFO, LRU and CCP under the change of cache size. The replacement frequency of the four cache replacement policies is reducing with the increase of cache capacity. Therefore, the more contents the router caches, the stronger the caching capability, the greater the possibility of providing services for users, the fewer replacements frequency, and the smaller the processing load of the router. When the router's cache capacity is unlimited, all contents can be cached without replacing. As shown in Fig. 6, FIFO always has the highest cache replacement frequency. FIFO always replaces the content that first enters the CS, without considering the characteristics of the content itself, so the cache replacement frequency is the highest. LRU replaces the content that has not been accessed for the longest time, so the cache replacement frequency is lower than FIFO. Both CRPM and CCP policies take content popularity into account, so the more likely the content in CS is to provide services, the fewer cache replacements. When the cache space is small, CRPM has greater advantages, but when the cache space increases, the difference between CRPM and CCP is no longer obvious.

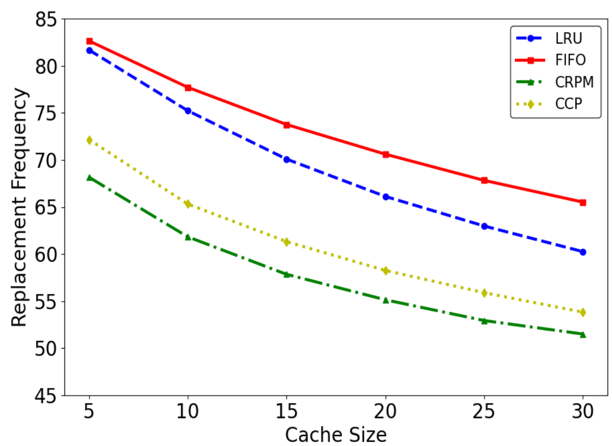

Figure 6: Cache size-replacement frequency

Fig. 7 is a comparison of replacement frequency of CRPM, FIFO, LRU and CCP under the change of Zipf exponents when the default cache size is 3. The replacement frequency of the four cache replacement policies is reducing with the increase of Zipf exponents. This is because the larger the Zipf exponents, the more requests users request for popular content, the greater the difference in content popularity. FIFO always replaces the content that first enters the CS, so it is insensitive to popularity, and FIFO always has the highest cache replacement frequency. LRU replaces the content that has not been accessed for the longest time, so the cache replacement frequency is lower than FIFO. CRPM and CCP are based on content popularity, so they are more sensitive to the increase of Zipf exponents, and CRPM is better than CCP strategy.

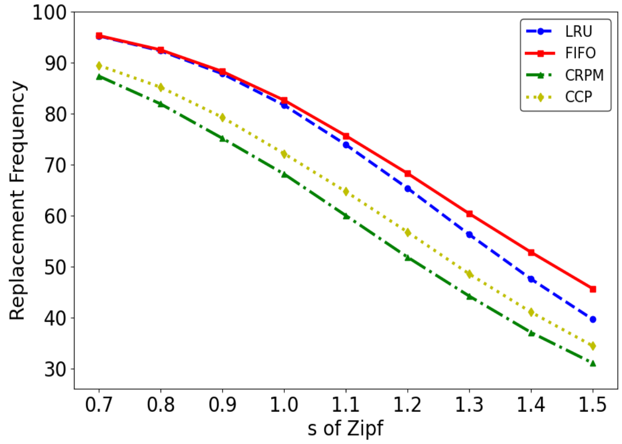

Figure 7: s of Zipf-replacement frequency

4.3.3 Average cache hit distance

In this experiment, we observe the effect of cache size and Zipf exponents on average cache hit distance. The results are introduced separately.

Fig. 8 is a comparison of average cache hit distance of CRPM, FIFO, LRU and CCP under the change of cache size. The average cache hit distance of the four cache replacement policies are reducing with the increase of cache capacity. Therefore, the more contents the router caches, the stronger the caching capability, the greater the possibility of providing services for users, and the less average cache hit distance. Therefore, the more contents can be cached on the router close to the user, the user's requests can be satisfied on these routers, so the average cache hit distance decreases continuously. As shown in Fig. 8, FIFO always has the highest average cache hit distance. FIFO always replaces the content that first enters the CS, without considering the characteristics of the content itself, so the average cache hit distance is the highest. LRU replaces the content that has not been accessed for the longest time, so the average cache hit distance is lower than FIFO. Both CRPM and CCP policies take content popularity into account, so the more likely the content in CS is to provide services, the fewer average cache hit distance. When the cache space is small, CRPM has greater advantages, but when the cache space increases, the difference between CRPM and CCP is no longer obvious.

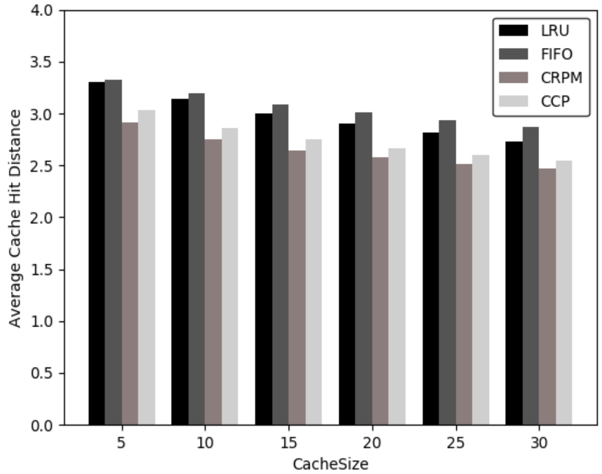

Figure 8: Cache size-average cache-hit distance

Fig. 9 is a comparison of average cache hit distance of CRPM, FIFO, LRU and CCP under the change of Zipf exponents when the default cache size is 3. The average cache hit distance of the four cache replacement policies are reducing with the increase of Zipf exponents. This is because the larger the Zipf exponents, the more requests users' requests for popular content, the greater the difference in content popularity. FIFO always replaces the content that first enters the CS, so it is insensitive to popularity, and FIFO always has the highest average cache hit distance. LRU replaces the content that has not been accessed for the longest time, so the average cache hit distance is lower than FIFO. CRPM and CCP are based on content popularity, so they are more sensitive to the increase of Zipf exponents, and CRPM is better than CCP strategy.

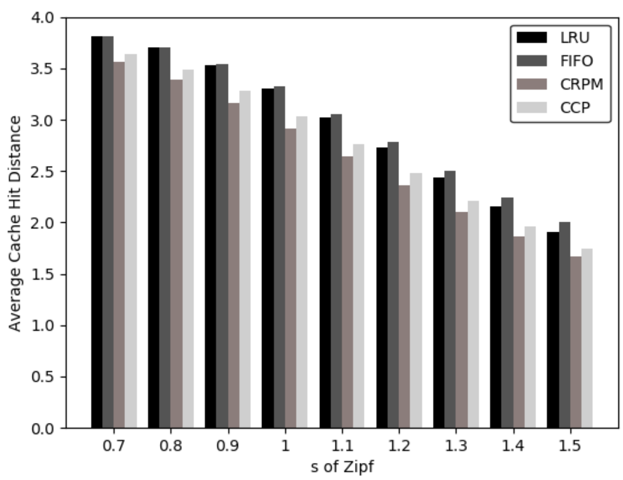

Figure 9: s of Zipf-average cache-hit distance

5 Conclusions

In this paper, we proposed the CRPM cache replacement policy for named data networking. The policy fully analyzes multi-factors that affect the caching performance, puts forward the corresponding calculation methods, and utilize the multi-factors to measure the cache value of contents. Furthermore, a new cache value function is constructed, which makes the content with high value be stored in the router as long as possible, so as to ensure the efficient use of cache resources. The simulation results show that CPRM can effectively improve cache hit ratio, enhance cache resource utilization, reduce energy consumption and decrease hit distance of content acquisition.

Funding Statement: This research was funded by the National Natural Science Foundation of China (No. 61862046), the Inner Mongolia Natural Science Foundation of China under Grant No. 2018MS06024, the Research Project of Higher Education School of Inner Mongolia Autonomous Region under Grant NJZY18010, the Inner Mongolia Autonomous Region Science and Technology Achievements Transformation Project (No. CGZH2018124) and the CERNET Innovation Project under Grant No. NGII20180626.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Aggarwal, M.; Nilay, K.; Yadav, K. (2017): Survey of named data networks: future of internet. *International Journal of Information Technology*, vol. 9, no. 2, pp. 197-207.

Fan, Z.; Wu, Q.; Zhang, M.; Zheng, R. (2017): Popularity and gain based caching scheme for information-centric networks. *International Journal of Advanced Computer Research*, vol. 7, no. 30, pp. 71-80.

Hail, M.; Amadeo, M.; Molinaro, A.; Fischer, S. (2015): Caching in named data networking for the wireless internet of things. *Proceedings of the International Conference on Recent Advances in Internet of Things*, Singapore.

Hu, X.; Gong, J.; Cheng, G.; Fan, C. (2015): Enhancing in-network caching by coupling cache placement, replacement and location. *Proceedings of the IEEE International Conference on Communications*, London, United Kingdom.

Jacobson, V.; Smetters, D.; Thornton, J.; Plass, M.; Briggs, N. et al. (2009): Networking named content. *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, Rome, Italy.

Li, H.; Nakazato, H.; Detti, A.; Melazzi, N. (2015): Popularity proportional cache size allocation policy for video delivery on CCN. *Proceedings of the European Conference on Networks and Communications*, Paris, France.

Liao, Y.; Hu, Y.; Wu, L.; Qin, Z. (2016): A weighted frequency-based cache memory replacement policy for named data networking. *Proceedings of the International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, Zhangjiajie, China.

Liu, W.; Zhang, J.; Liang, Z.; Peng, L.; Cai, J. (2017): Content popularity prediction and caching for ICN: a deep learning approach with SDN. *IEEE Access*, vol. 6, pp. 5075-5089.

Ma, G.; Chen, Z.; Zhao, K. (2013): A cache management strategy for content store in content centric network. *Proceedings of the Fourth International Conference on Networking and Distributed Computing*, Hong Kong, China.

Mallouli, F. (2019): Robust EM algorithm for iris segmentation based on mixture of gaussian distribution. *Intelligent Automation and Soft Computing*, vol. 25, no. 2, pp. 243-248.

Mastorakis, S.; Afanasyev, A.; Zhang, L. (2017): On the evolution of ndnSIM: an open-source simulator for NDN experimentation. *ACM SIGCOMM Computer Communication Review*, vol. 47, no. 3, pp. 19-33.

Ran, J.; Lv, N.; Zhang, D.; Ma, Y.; Xie, Z. (2013): On performance of cache policies in named data networking. *Proceedings of the International Conference on Advanced Computer Science and Electronics Information*, Beijing, China.

Ren, J. J.; Zhao, S.; Sun, J. D.; Li, D.; Wang, S. et al. (2018): PPP: prefix-based popularity prediction for efficient content caching in content-centric networks. *Computer Systems Science and Engineering*, vol. 33, no. 4, pp. 259-265.

Wang, B.; Kong, W.; Li, W.; Xiong, N. (2019): A dual chaining watermark scheme for data integrity protection in Internet of Things. *Computers, Materials & Continua*, vol. 58, no. 3, pp. 679-695.

Xin, Y.; Li, Y.; Wang, W.; Li, W.; Chen, X. (2016): Content aware multi-path forwarding strategy in Information Centric Networking. *Proceedings of the IEEE Symposium on Computers and Communication*, Messina, Italy.

Xu, C.; Wang, H.; Wang, H. (2017): A cache replacement scheme based on contribution to hit ratio of node in content-centric networking. *Telecommunication Engineering*, vol. 57, no. 3, pp. 311-315.

Xylomenos, G.; Ververidis, C.; Siris, V.; Fotiou, N.; Tsilopoulos, C. et al. (2014): A survey of information-centric networking research. *IEEE Communications Surveys and Tutorials*, vol. 16, no. 2, pp. 1024-1029.

Yang, Y.; Zhu, J. (2016): Write skew and Zipf distribution: evidence and implications. *ACM Transactions on Storage*, vol. 12, no. 4, pp. 1-19.

Zhang, L.; Afanasyev, A.; Burke, J.; Jacobson, V.; Crowley, P. et al. (2014): Named data networking. *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66-73.

Zhang, M.; Luo, H.; Zhang, H. (2015): A survey of caching mechanisms in information-centric networking. *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1473-1499.

Zhang, Y.; Tan, X.; Li, W. (2017): PPC: popularity prediction caching in ICN. *IEEE Communications Letters*, vol. 22, no. 1, pp. 5-8.

Zhu, Y.; Mi, Z.; Wang, W. (2013): A cache probability replacement policy based on content popularity in content centric networks. *Journal of Electronics & Information Technology*, vol. 35, no. 6, pp. 1305-1310.