IEEE Access
The journal for rapid open access publishing
IEEE Xplore Digital Library
http://ieeexplore.ieee.org/Xplore

Cooperative Content Caching for Mobile Edge Computing with Network Coding

WEISHENG HE[1], YUHAN SU[1], XUETING XU[1], ZHAOHUI LUO[1], *Student Member, IEEE* LIANFEN HUANG[1], AND XIAOJIANG DU[2], *Senior Member, IEEE*

Abstract—**As an important technology in next-generation networks, mobile edge computing (MEC) can bring high-speed intelligent network services to users. However, planning MEC file transmission represents a problem to be urgently solved. In addition, network coding has been an emerging technology in recent years, and certain forms of cooperation with wireless relays can improve the transmission efficiency of network data. In this paper, we focus on cooperation in MEC intelligent cache with network coding. First, we built a MEC intelligent cache analysis platform that includes multiple algorithms; it can predetermine which streaming media files should be stored in the limited storage space of the MEC device to provide users the most efficient intelligent network experience. The platform's decision algorithm includes an innovative composite recommendation algorithm that is more comprehensive than the traditional recommendation algorithm. The MEC file extracted from the core network is then simultaneously transmitted to the edge with the real-time normal communication file using an innovative network coding and wireless relay cooperation method; with that, files stored in the MEC can be cached to the edge. The normal network communication process does not require additional slot costs. Simulation results show that the proposed cooperation scheme can effectively improve network data transmission, improve users' average access rate, and offer users better intelligent network services.**

Index Terms—**Mobile edge computing; network coding, wireless relay, intelligent cache, cooperative communications**

I. Introduction

As an important applications of next-generation networks (5G) network, fog radio access networks aim to fully leverage the processing and storage functions of users and edge network devices to offer users more intelligent, efficient and humanized network services. As a prominent technology in 5G fog wireless access networks, mobile edge computing (MEC) has recently attracted extensive attention [1]. Although MEC can bring more efficient and humane network services to surrounding users, many problems remain to be addressed in the MEC landing process, such as how to make intelligent decisions about host files the MEC needs to cache in advance or how to reasonably arrange MEC data transmission [2].

Network coding has also become a hotspot within the past decade. Its core idea is to allow intermediate nodes in the network to re-encode different information flows and become one information flow to be transferred. After gathering enough information, the flow can be decoded at the target node according to certain rules [3]. Recently, the question of how to apply network coding technology to enhance collaboration with network devices has become a popular research area.

The authors in [4] and [5] noted that methods to reasonably formulate a MEC cache scheme warrant further study. The authors in [6] found that different MEC caching decisions could lead to distinct optimization effects for ambient users. In [7], the authors innovatively improved the recommendation algorithm, which is of paramount importance in the MEC cache decision-making scheme. Their collaborative filtering algorithm, which was based on user attributes and scores, could analyze user data more objectively than the original recommendation algorithm, but it could not consider the user and project characteristics simultaneously. In [8], a recommendation algorithm that considered both user and project characteristics was proposed, but it did not take data objectivity into account.

In [9], the authors put forward a scheme of applying network coding to data retransmission in satellite communication. The authors in [10] introduced a network coding scheme in data broadcasting and found that network coding could also improve the efficiency of data transmission in wireless communication networks. In [11], the authors sought to combine network coding and wireless relay devices to select more suitable communication lines for users. This scheme inspired us to apply network coding and wireless relay cooperation to communication networks and then optimize data transmission in such networks.

In this paper, we study a cooperation scheme of network coding and wireless relay in 5G MEC intelligent cache. Overall, our work makes three contributions:

- We construct a MEC intelligent cache analysis platform composed of various algorithms to determine which files should be stored in MEC devices.
- We propose an innovative compound recommendation algorithm in the algorithmic flow of the MEC cache decision, which may address shortcomings of traditional recommendation algorithms that cannot consider numerous factors simultaneously.
- An innovative cooperation involving network coding and wireless relay station is recommended to transfer MEC files and real-time normal communication files to the edge at the same time, thereby implementing the intelligent cache of mobile edge computing efficiently.

The remainder of this paper is organized as follows. We describe the overall framework of our proposed MEC intelligent cache analysis platform and cooperative application scheme of network coding and wireless relay station in Section II. Then, we analyze each algorithm step in Section III. In Section IV, we introduce the cooperation scheme of network coding and wireless relay station. Lastly, in Section V, we

Digital Object Identifier: 10.1109/ACCESS.2019.2917977

combine the algorithm framework with streaming media data for simulation and determine that the proposed scheme can substantially improve user access efficiency and reduce user access latency. This scheme can also greatly enhance users' network service experience.

II. MODEL FRAMEWORK

After finalizing the decision of MEC intelligent caching, the efficient transfer of pre-cached files from the core network to a MEC device has become an important problem. According to prior research, the transfer generally occurs when the surrounding network load is relatively low; alternatively, it is possible to allocate certain channel resources for transmission. Although this method can also complete MEC file pre-caching, the following shortcomings persist.

1) For office or residential areas with a high population density, the network load may continue to be high 24 hours a day. Although the network load may decline during certain periods, it is still higher than the normal network load level. In this case, it is difficult to allocate time resources for MEC file pre-caching.
2) The changing speed of network content popularity is quite fast today. Therefore, to guarantee real-time synchronization of users' intelligent experience, streaming media files stored in MEC should maintain a relatively high update frequency. If all streaming media files are waiting for a relatively idle period in the network, a lag effect may follow and result in failure to achieve the desired effect of MEC service.
3) Whether using dedicated channel resources or devoting dedicated slot resources to MEC file caching during idle times, if we wish to ensure that normal communication is unaffected and the MEC file can cache synchronously, then the network equipment workload will inherently grow. Over time, the network equipment will likely become overloaded.

In other words, to realize intelligent MEC services, it is necessary to construct an efficient and useful intelligent cache analysis platform and propose novel and efficient data transmission methods to support advance MEC file transfer.

In the technical introduction, we mentioned that network coding can compress information on nodes and transmit them together. The wireless relay represents a form of smart base station equipment that has been used extensively and extended in recent years. Therefore, we propose a scheme that combines collaborative network coding and wireless relay equipment in a specific way to solve MEC file caching in advance. The schematic diagram is displayed in FIGURE 1.

In this scheme, the innovative supporting equipment includes a new type of relay base station (RS-MEC) with a MEC server function, which operates in half-duplex mode. The macro base station (BS) and the new relay base station can access core network communication. The new relay base station RS with MEC server function possesses the communication and wireless transmission capabilities of the BS along with the ability for MEC content intelligent sensing and data storage. By collecting the potential needs of peripheral users, combined with hot content in the network, we can output several potential needs files that are most suitable for peripheral users. These requirements are fed back to the core network, and corresponding requirement files are generated in the core network and then transmitted to edge MEC via the scheme proposed herein.

The MEC cache intelligent analysis platform is responsible for intelligent decision making regarding which files must be stored in advance based on collected preference information from peripheral users about streaming media files (i.e., the size of streaming media files and the size of the MEC storage module) to offer users an optimally intelligent experience. The algorithm flow of the MEC intelligent cache analysis platform constructed in this paper is illustrated in FIGURE 2.

Although existing streaming media platforms on the network provide a file popularity ranking that can serve as a similar reference, the MEC intelligent caching analysis platform in this paper can offer more comprehensive, practical, and targeted analysis results. Two points in the algorithm process are noteworthy:

1) MEC is aimed at analyzing the preferences of surrounding user groups in different regions. Because the databases of streaming media platforms on the network operate independently, it is impossible to reflect users' preferences for various streaming media content in real time in different regions. The MEC intelligent caching analysis platform in this paper provides more effective analysis results via real-time intelligent analysis for users surrounding the device.
2) In the process of MEC intelligent cache analysis, it is necessary to consider the popularity of streaming media files as well as differences in users' usage frequency and service level, the MEC server capacity, and the size of each file. Therefore, the results of comprehensive decision making will reveal the most appropriate and humanized service for peripheral users. Therefore, the algorithm flow in this paper introduces numerous algorithms to analyze various factors step by step to ultimately obtain the most intelligent and humanized decision-making results. The current streaming media platform on the network cannot do this; our method embodies humanized innovation in the proposed MEC intelligent caching analysis platform.

The application of each algorithm is described in detail in Section III along with an overview of the algorithm flow of the MEC intelligent cache analysis platform.

III. ALGORITHM EXPLANATION

A. *New Compound Recommendation Algorithms*

The traditional recommendation algorithm is limited by the uniformity of factors considered in formulas; as such, the algorithm cannot effectively and accurately analyze users' recommendation preferences. To address this problem, we propose a new compound recommendation algorithm that combines multiple factors to improve the practical recommendation effect of the algorithm. The flow chart of this algorithm is depicted in FIGURE 3. Specific implementation of the

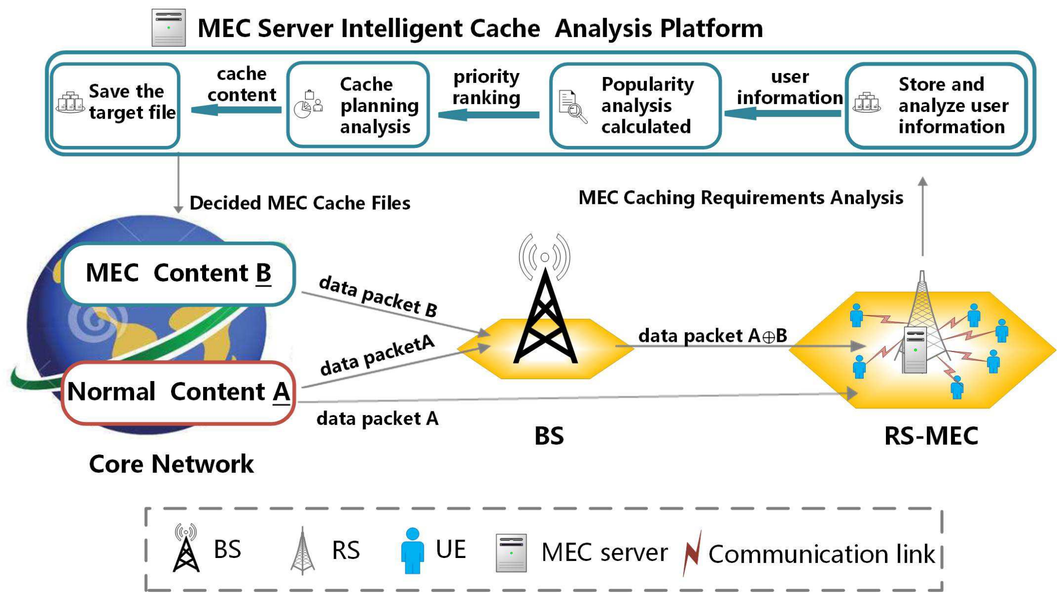


Fig. 1: Architecture of cooperative application of network coding and wireless relay in MEC.

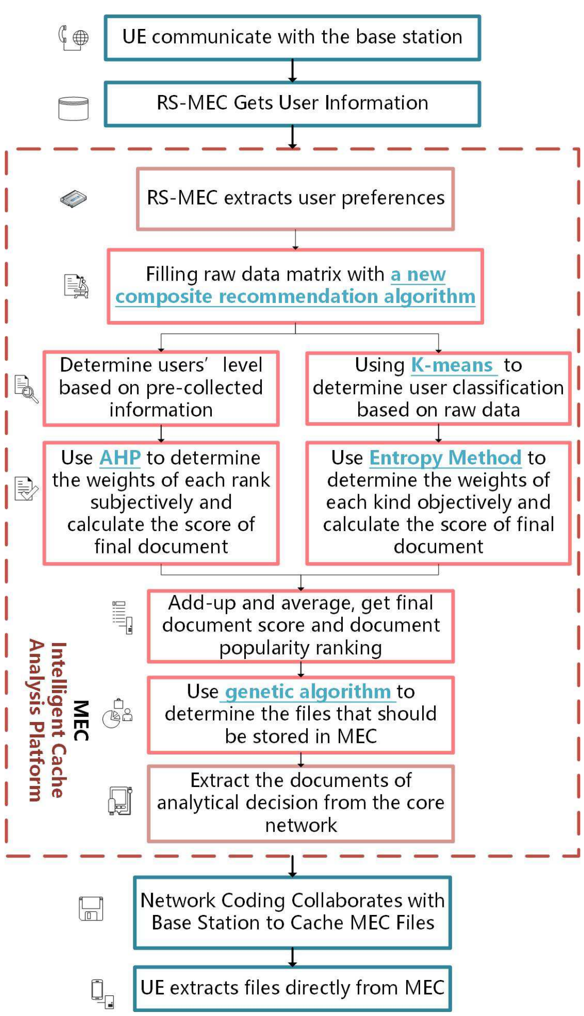


Fig. 2: Algorithmic framework of MEC intelligent cache analysis platform.

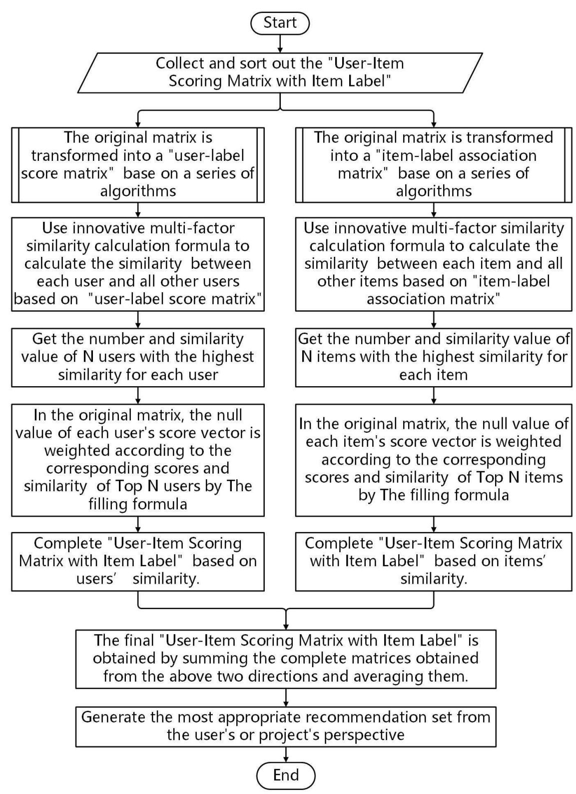


Fig. 3: Flowchart of new compound recommendation algorithms.

algorithm will be introduced in accordance with key points of the process.

1) SUMMARIZE ORIGINAL "USER-ITEM SCORING MATRIX" WITH VACANCY VALUE: First, we collect and organize information for the items and users to be studied and compose the "User-Item Scoring Matrix with Item Label". For items that users have not touched, the score is null.

2) CALCULATE "USER-LABEL SCORE MATRIX": According to the average score of the items, items can be divided into two sets: an item set that is equal to or greater than the average user rating (positive item set); and an item set that is less than the average user rating (negative item set). The positive item set represents the set that user prefers, whereas the negative item set represents the set that user dislikes [8].

We define the user's tag set as $T = \{t_1, t_2, ..., t_v\}$. The weight $w_{u,j}$ of user u on tag t in item i can be obtained by using the following formula:

$$w_{u,j} = \frac{R_{u,i}}{\sqrt{\sum_{j=1}^{|t|} R_{u,j}^2}}, \tag{1}$$

where $R_{u,i}$ represents user u's evaluation of item i.

A many-to-many relationship exists between tags and items. A tag can mark multiple items, and an item can have multiple tags. Therefore, to measure the user's overall preference for a tag, it is necessary to calculate the user's average preference for a tag in the positive item set and the negative item set respectively. Label weights of users on positive item sets and negative item sets are recorded as $\mu_{u,t}^{pos}$ and $\mu_{u,t}^{neg}$ respectively. The specific calculation methods are as follows:

$$\mu_{u,t}^{pos} = \frac{1}{|I_u^{pos}(t)|} \times \sum_{j \in I_u^{pos}(t)} w_{u,j}(t); \tag{2}$$

$$\mu_{u,t}^{neg} = \frac{1}{|I_u^{neg}(t)|} \times \sum_{j \in I_u^{neg}(t)} w_{u,j}(t). \tag{3}$$

Where $I_u^{pos}(t)$ is an active item set marked with label t by user u , and $I_u^{neg}(t)$ is a negative item set marked with label t by user u. Ultimately, the overall preference level of user u for label t can be obtained with $\mu_{u,t}^{pos}$ and $\mu_{u,t}^{neg}$:

$$\mu_{u,t} = \begin{cases} (\mu_{u,t}^{pos} + \mu_{u,t}^{neg})/2 & t \in [I_u^{pos}(t) \bigcap I_u^{neg}(t)] \\ \mu_{u,t}^{pos} & t \in I_u^{pos}(t), t \notin I_u^{neg}(t) \\ \mu_{u,t}^{neg} & t \notin I_u^{pos}(t), t \in I_u^{neg}(t) \end{cases} \tag{4}$$

Thus far, the "User-Label Score Matrix" can be established preliminarily according to the overall preference of user u for label t.

3) CALCULATE "ITEM-LABEL ASSOCIATION MATRIX": Let N_t represent the number of items with label t , N represents the total number of items, $Freq_{t,i}$ represents the frequency of tagging item i with label t, and $\sum\limits_{t \in T} Num_{t,i}$ represents the number of times that item i has been tagged with labels in the label set. Then, the degree of correlation $q_{t,i}$ between label t and item i can be expressed as:

$$q_{t,i} = \frac{Freq_{t,i}}{\sum\limits_{t \in T} Num_{t,i}} \lg(N/N_t). \tag{5}$$

Hence, the "Item-Label Association Matrix" can be established.

4) MULTIVARIATE SIMILARITY CALCULATION FORMULA: After obtaining the two new label matrices, we must calculate the similarity values between each user and other users and between each item and other items based on the tag content. Then, we can record the top N user numbers and similarity values corresponding to the highest similarity for each user and do the same for items [7].

The following formulas can be used to calculate multivariate similarity, where u and v represent two vectors, respectively.

$$sim_{total}(u,v) = \left[\frac{1}{|C_{uv}|} \sum_{i \in C_{uv}} sim_1(u,v,i) \cdot sim_2(u,v,i) \cdot sim_3(u,v) \right]. \tag{6}$$

sim_1 represents the numerical similarity of scores:

$$sim_1(u,v,i) = 2 \left[1 - \frac{1}{1 + \exp(-|p_{ui} - p_{vi}|)} \right], \tag{7}$$

where u denotes the principal vector u, v denotes the vector v need to be analyzed, i denotes current position, and P_{ui} denotes the value of vector u in position i. Parameter 2 is introduced to modify the similarity between (0,1).

sim_2 represents interest tendency similarity:

$$sim_2(u,v,i) = 2 \left\{ 1 - \frac{1}{1 + \exp\left[-(p_{ui} - \overline{p}_u)(p_{vi} - \overline{p}_v) \right]} \right\}, \tag{8}$$

where u denotes the principal vector u, v denotes the vector v to be analyzed, i denotes current position i, $P_(ui)$ denotes the value of vector u in position i, and $\overline{P_u}$ denotes the numerical mean of vector u. Parameter 2 is introduced to modify the similarity between (0,1).

sim_3 represents the overlapping confidence:

$$sim_3(u,v) = \frac{|I_u \cap I_v|}{|I_u \cup I_v|}, \tag{9}$$

where I_u represents the set of valid non-null elements of vector u, and I_v represents the set of valid non-null elements of vector v.

5) FILLING FORMULA WEIGHTED BY SIMILARITY AND MODIFIED BY AVERAGE SCORE: In the original matrix, the null value of each user's score vector is weighted according to the corresponding scores and similarity of the top N users by the filling formula. The same weighting scheme is applied for items.

$$p_{ui} = \overline{p}_u + \frac{\sum\limits_{u' \in U'} (p_{u'i} - \overline{p}_{u'}) \times sim(u,u')}{\sum\limits_{u' \in U'} |sim(u,u')|}. \tag{10}$$

Where u represents the subject object (user or item), i represents the corresponding position of the subject object (i represents the item if the subject is the user; i represents the user if the subject is an item); P represents the score value; and u' represents the collection of the N highest similarity related subjects.

6) OBTAIN COMPLETE "USER-ITEM SCORING MATRIX" AND PROVIDE RECOMMENDATION: The final "User-Item Scoring Matrix with Item Label" is obtained by summing the complete matrices obtained from the above two directions and averaging them. We can then generate the most appropriate recommendation set from the user's or project's perspective.

The "User-Item Scoring Matrix", which is finalized in this step, will be used for subsequent weighting calculations.

B. K-means Clustering Analysis

In this paper, a K-means clustering algorithm is used for clustering analysis based on the original "User-Item Scoring Matrix" which contains null. Users are divided into four types given the objectivity of data information.

The algorithm formalization can be described as follows:

Input: Number of clusters k, and database D containing N data objects.

Output: K clusters meeting minimum variance criteria.

Processing flow:

Step 1: Randomly select K objects from N data objects as initial clustering centers.

Step 2: Assign each object to the most similar cluster according to the average value of objects in the cluster.

Step 3: Update the average value of clusters (i.e., calculate the average value of objects in each cluster).

Step 4: Cycle from Step 2 to Step 3 until each cluster no longer changes [12].

Through K-means data analysis of the original "User-Item Scoring Matrix", users can be divided into four categories according to objective information of the vector data. From there, the average of users' scores by user type can be obtained and applied for weighted calculations.

C. Entropy Method

The entropy method is a common approach for calculating weights. Its most useful feature is the ability to determine the weights of each index parameter according to the amount of information contained in the parameter itself, contrary to the analytic hierarchy process (AHP).

Step 1: Import a complete index matrix:

$$A = \begin{pmatrix} X_{11} & \dots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix}_{n\times m}. \tag{11}$$

Where X_{ij} is the value of the j-th index of the i-th scheme.

Step 2: Complete the matrix non-negative digitalization operation.

Because the algorithm uses the ratio of one index per scheme to the total of the same index, it can avoid the influences of dimensions and eliminate the need for data standardization. However, if a negative number exists in the data matrix, then a non-negative translation calculation is needed; the formula is as follows (ideally, the larger the index value, the better):

$$X'_{ij} = \frac{X_{ij} - \min(X_{1j}, X_{2j}, \cdots, X_{nj})}{\max(X_{1j}, X_{2j}, \cdots, X_{nj}) - \min(X_{1j}, X_{2j}, \cdots, X_{nj})} + 1. \tag{12}$$

When a smaller index value is better, the following formula should be used:

$$X'_{ij} = \frac{\max(X_{1j}, X_{2j}, \cdots, X_{nj}) - X_{ij}}{\max(X_{1j}, X_{2j}, \cdots, X_{nj}) - \min(X_{1j}, X_{2j}, \cdots, X_{nj})} + 1. \tag{13}$$

For the convenience of subsequent formulas, non-negative digitized data are represented by X_{ij}.

Step 3: After completing the non-negative digitization operation, the obtained data matrix fulfills the standard. Then, the proportion of the j-th index to the i-th scheme can be calculated:

$$P_{ij} = \frac{X_{ij}}{\sum\limits_{i=1}^{n} X_{ij}} (j = 1, 2, \cdots, m). \tag{14}$$

Step 4: Calculating the entropy value of the j-th index:

$$e_j = -k * \sum_{i=1}^{n} P_{ij} \log(P_{ij}). \tag{15}$$

When $k > 0$, ln is the natural logarithm, $k = 1/\ln m$.

Step 5: Compute the difference coefficient of the j-th index. For the j-th index, the greater the numerical difference of parameter X_{ij}, the greater the effect of parameter X_{ij} on scheme evaluation. The smaller the entropy value, the larger the corresponding weight parameter should be.

$$g_j = 1 - e_j. \tag{16}$$

The larger the value of g_j , the more important the index.

Step 6: Determine the final weight of each index:

$$W_j = \frac{g_j}{\sum\limits_{j=1}^{m} g_j}, j = 1, 2, \cdots, m. \tag{17}$$

At this point, the calculation process of the entropy method is complete [13].

By using the entropy method to analyze four kinds of representative scoring vectors for each item, we can determine the representative scoring weights of each kind in the final total scoring calculation. Results are listed in TABLE I.

TABLE I: WEIGHTS DETERMINED BY ENTROPY METHOD

Kind	Kind 4	Kind 3	Kind 2	Kind 1
Weight	0.2816	0.2112	0.1894	0.3179

Thus, the final score and popularity ranking of each item can be calculated objectively.

TABLE II: ANALYTIC HIERARCHY PROCESS NUMERICAL RULES TABLE

No.	**Level of importance**	a
1	i, j are equally important	1
2	The i element is slightly more important than the j element	3
3	The i element is significantly more important than the j element	5
4	The i element is more important than the j element	7
5	The i element is more important than the j element greatly	9
6	The i element is slightly less important than the j element	1/3
7	The i element is obviously less important than the j element	1/5
8	The i element is strongly less important than the j element	1/7
9	The i element is much less important than the j element	1/9
	2, 4, 6, 8 and 1/2, 1/4, 1/6, 1/8 are among the above indicators	

D. Analytic Hierarchy Process (AHP)

The AHP is a classic weight determination algorithm. The most important feature of this method is that users' subjective judgments can be converted into weight parameters via matrix operations; then, the priority that individuals assign to indicators can be transformed into respective weights.

The specific calculation steps are as follows:

Step 1: Establish a two-by-two comparison matrix for tectonic indicators:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{bmatrix} \tag{18}$$

where a_{ij} indicates the importance of item i relative to item j. The relative importance of item j to item i is $a_{ji} = 1/a_{ij}, a_{ij} > 0$ and $a_{ii} = 1$. The numerical value of a can be determined by referring to the numerical rules in TABLE II.

Step 2: Normalize indicators' two-by-two comparison matrix by column:

$$b_{ij} = \frac{a_{ij}}{\sum\limits_{i=1}^{n} a_{ij}} (i = 1, 2, \cdots, n). \tag{19}$$

Step 3: Add the normalized matrix by rows:

$$\overline{W}_l = \sum_{j=1}^{n} b_{ij} (i = 1, 2, \cdots, n). \tag{20}$$

Step 4: Normalize the vector $\overline{W} = \left[\overline{W}_1 \overline{W}_2 \overline{W}_3 ... \overline{W}_n\right]^T$:

$$W_i = \frac{\overline{W}_l}{\sum\limits_{j=1}^{n} \overline{W}_l} (i = 1, 2, \cdots, n). \tag{21}$$

The resulting vector $W = [W_1 W_2 W_3 ... W_n]^T$ is a feature vector and is the weight parameter of the decision matrix. However, follow-up verification is required.

Step 5: Calculate the maximum eigenvalue of the two-by-two comparison matrix:

$$\lambda_{\max} = \sum_{i=1}^{n} \frac{(AW)_i}{nW_i}, \tag{22}$$

where $(AW)_i$ is the i-th component of a vector matrix (AW).

TABLE III: VALUE OF ANALYTIC HIERARCHY PROCESS R.I.

Dimension	R.I.
1	0
2	0
3	0.58
4	0.96
5	1.12
6	1.24
7	1.32
8	1.41
9	1.45

TABLE IV: WEIGHTS DETERMINED BY ANALYTIC HIERARCHY PROCESS

Level	Level 4	Level 3	Level 2	Level 1
Weight	0.5477	0.2588	0.1448	0.0517

Next, we must verify the consistency of the matrix. Because the matrix is artificially constructed, we need to ensure there are no significant contradictions between indicators.

$$C.I. = \frac{\lambda_{\max} - n}{n - 1}. \tag{23}$$

When $\lambda_{\max} = n$, $C.I. = 0$, in which case the matrix is fully logically consistent. The higher the value of $C.I.$, the poorer the logical consistency of the structure matrix. When $C.I. \leq 0.1$, the logical consistency of the matrix is acceptable; otherwise, we need to recalculate.

However, because too many indicators comprise a large dimension of the matrix, the judgment consistency will deteriorate. Therefore, it is necessary to introduce a correction factor parameter $R.I.$ to improve the requirements for relaxing consistency test parameters:

$$C.R. = \frac{C.I.}{R.I.}. \tag{24}$$

When $C.R. \leq 0.1$, the logical consistency of the constructed matrix is acceptable; otherwise, it needs to be reconstructed. In addition, the AHP can usually only apply to the calculation of nine indicators. The value of $R.I.$ is as shown in TABLE III.

This step completes the AHP calculation. Then, it is only necessary to input the constructed two-by-two comparison decision matrix for different user levels. Following the above series of operations, the calculated weight vector can be obtained along with the results of the consistency check [14].

In this paper, according to the declining order of importance of Levels 4, 3, 2, and 1, we can construct the following two-pair comparison decision matrix:

$$\begin{bmatrix} 1 & 3 & 5 & 7 \\ 1/3 & 1 & 3 & 5 \\ 1/5 & 1/3 & 1 & 5 \\ 1/7 & 1/5 & 1/3 & 1 \end{bmatrix} \tag{25}$$

The corresponding levels from left to right and top to bottom are 4, 3, 2, and 1. Results are displayed in TABLE IV.

The subjective computing branch is based on the classification criteria of a user's previous grades. We take the average of

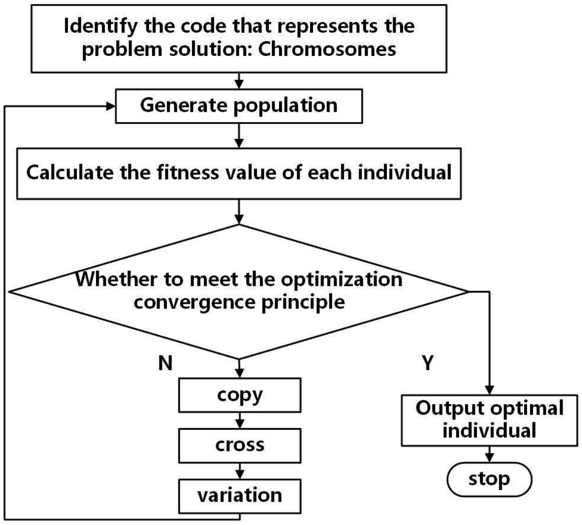


Fig. 4: Genetic algorithm flow chart.

users' scores in each level to obtain the representative scores of each item. Then, we use the AHP results to weigh the scores of each item to determine the final subjective scores per item.

Next, the final scores of each project and corresponding popularity ranking can be obtained by summing and averaging the previous objective and subjective calculation scores. Together with the file size information, the next step can be used to solve the decision-making process, similar to the knapsack problem.

E. Improved Genetic Algorithm

After processing the project scoring data, it is necessary to decide which items to deposit on the edge of the MEC server. The capacity of the MEC server is limited, such that only a certain amount of space is allocated for different types of content (i.e., to store the most suitable content for surrounding users). Therefore, the question of how best to leverage storage space on the MEC server and balance the score ranking and file sizes of streaming media files warrants careful consideration. This problem can be regarded as a backpack problem.

We can use the genetic algorithm to solve the backpack problem here. FIGURE 4 outlines the operation flow of the algorithm.

In addition, in the previous research process, we found that the most basic genetic algorithm will likely fall prey to a local optimal problem. To solve this optimization problem, appropriate improvements should be made to link variation. Here, we introduce an improved genetic algorithm with a large variation effect.

The so-called improved genetic algorithm can elicit the following enhancements for a small probability of random variation:

1) Calculate the population's maximum fitness F_{max} the average fitness F_{ave} of a certain generation to judge $\alpha F_{\max} < F_{ave}$ and $0.5 < \alpha < 1$.
2) If the above conditions are satisfied, set a value close to 1 (e.g., 0.9). If the obtained α reaches the set value of 0.9, one can assume that the population has reached a certain level of fitness and it requires a random shatter.
3) For random dispersion, the mutation probability at this time is multiplied by several multiples. We set the probability to amplify by 5 times to perform a large-variation operation, which causes individuals in the population to exhibit a relatively large variation [15].

Finally, after analyzing the genetic algorithm, we can determine which files are suitable for peripheral users to store in advance at the edge of the MEC server. The MEC intelligent cache analysis platform algorithm framework has thus been described in full.

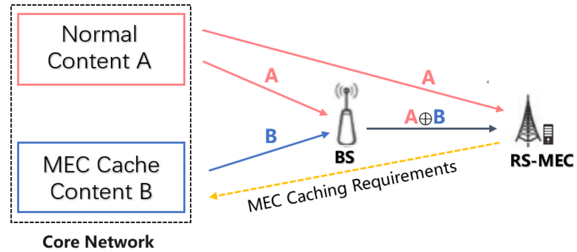


Fig. 5: Scenario of Network Coding and Wireless Relay Cooperation in MEC.

IV. COOPERATIVE APPLICATION OF NETWORK CODING AND WIRELESS RELAY

This paper proposes a cooperation scheme between network coding and wireless relay and its matching equipment to innovatively solve the problem of transmission resource coordination; such coordination will affect normal, real-time communication when potential demand files are cached in advance. This scheme improves the operating efficiency of the network as well as the reliability of information transmission. The scheme combines the flexibility of the relay base station with the high efficiency of network coding and realizes an intelligent and convenient MEC (FIGURE 5).

In this scheme, A is transmitted to BS and RS-MEC; B is transmitted to BS. BS then encodes A and B in XOR network coding and transmits them to RS. RS-MEC obtains encoding packet $A \oplus B$. Together with previous packet A, $(A \oplus B) \oplus A = B$, thus, RS-MEC can get B and keep it in the server storage module.

We can use a dynamic slot graph to describe the transmission dynamics of the proposed scheme. Illustrations are given for tags in the following series of time-slot graphs:

1) Normal communication content A will be divided into several data packets $A_1, A_2, ..., A_n$; MEC caching content is divided into several packets $B_1, B_2, ..., B_m$;
2) The channel capacity of each slot can transmit a data packet (i.e., an A_n or B_m);
3) "$\oplus$" represents XOR network coding, which does not change the size of the packet file; that is, the size of two packets after the encoding operation is the same size as one packet.

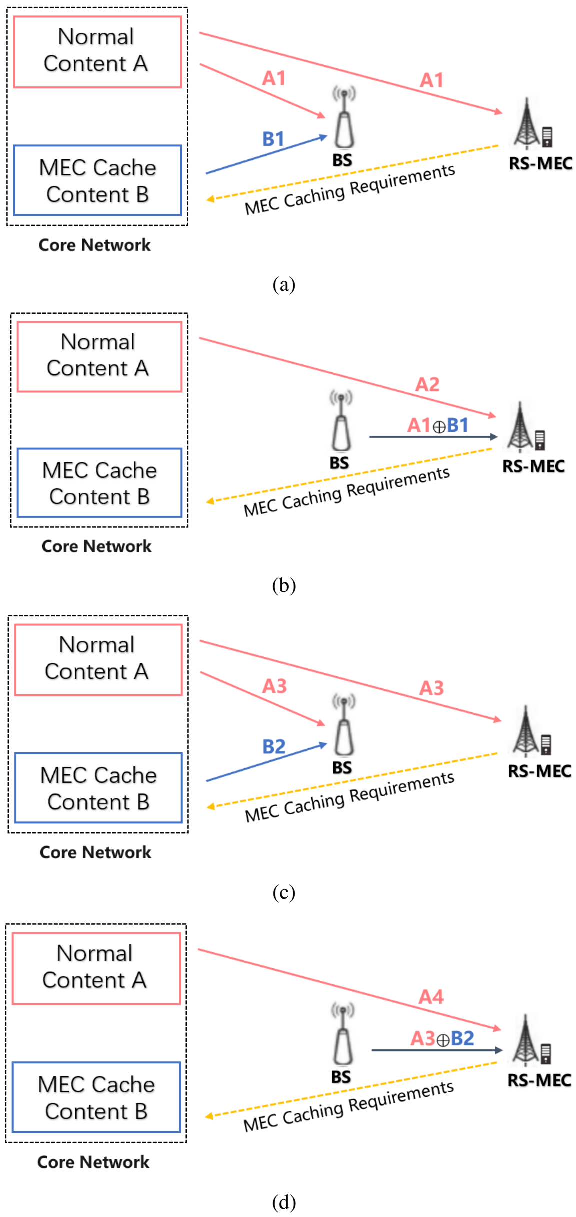


Fig. 6: Schema transmission Schematic. (a) slot 1. (b) slot 2. (c) slot 3. (d) slot 4.

The following describes the operation time-slot graph under normal communication conditions, assuming all transmission contents can be received correctly (see FIGURE 6).

Every two transmission slots can be regarded as a small operation cycle. If A file has n data packets, then after n slots, the A file transmission is completed, and B file's data packet transmission is $n/2$. Through this innovative collaboration scheme, MEC files can be pre-cached smoothly.

V. SIMULATION EXPERIMENT DESCRIPTION

We selected a movie file as an example of a MEC streaming media file and chose 50 movies as the research set on a video website. We recorded the names, scores, and sizes of the format files along with label information for each movie.

TABLE V: FILE LIST FOR OPTIMAL DECISION MAKING

Number	File Size (MB)	Ranking of Intelligent Analysis	Scores of Intelligent Analysis
1	242.16	1	9.38
2	184.99	2	9.31
3	305.85	3	9.28
4	203.84	4	9.19
5	171.59	5	9.19
6	251.98	6	9.13
7	182.20	7	9.11
8	196.48	8	9.09
9	167.12	10	9.03
10	188.34	11	9.03
11	185.78	13	9.01
12	192.76	20	8.93
13	158.31	21	8.89
14	143.99	30	8.67
15	223.77	33	8.64
Total file size	2999.16 MB		

After sorting the data, we investigated 100 users' preference scores on 50 movies on the WeChat platform. We used a web-based questionnaire to collect experimental data for follow-up research, analogous to the data collection process of MEC's intelligent perception function. The following parameters applied to the questionnaire:

1) Users could choose any integer score for each movie according to their preferences in a range from 1 to 10.
2) If a user had not seen a particular movie, he or she can assign the movie either no score or a score of zero.
3) At the end of the questionnaire, users were asked to choose a frequency rating for movie-watching (4 = frequently, 3 = generally, 2 = occasionally, basically never = 1).

In total, 100 scores for the 50 films were obtained. After summarizing, a user-item score matrix with complete information and vacancy values was taken as the data basis for a subsequent experimental simulation.

After investigation, we decided to select approximately 1/3 of the 50 movies for pre-caching; therefore, MEC storage space was set to 3GB in this experiment. After operating the MEC intelligent cache analysis platform, a list of the best cached files in the dataset were compiled in TABLE V.

A. Comparing Simulations of New Compound Recommendation Algorithms

The new compound recommendation algorithm is an improved recommendation algorithm as proposed in this paper. Compared with the original recommendation algorithm, our version considers many factors more carefully to better fill vacancies in the "user-item score matrix" and provide relevant recommendations. In this section, the innovative algorithm is compared with the traditional user CF algorithm and the project CF algorithm to verify the analysis results.

Because pre-collected information included the scores and rankings of each movie on the website, we can use this information as the standard to compare the results of each algorithm. The data deviation value is introduced below for comparative analysis.

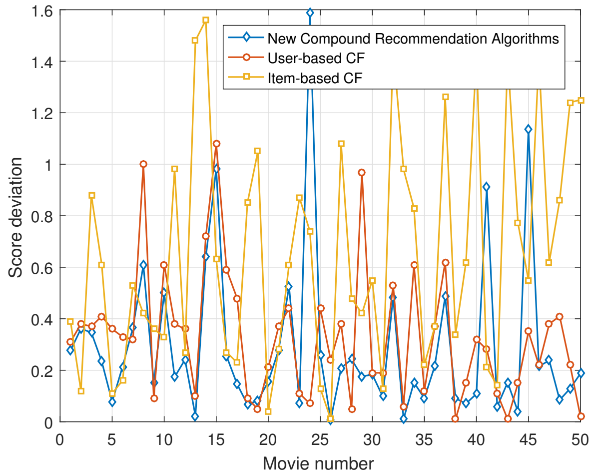


Fig. 7: Comparison of score deviation values obtained by different methods.

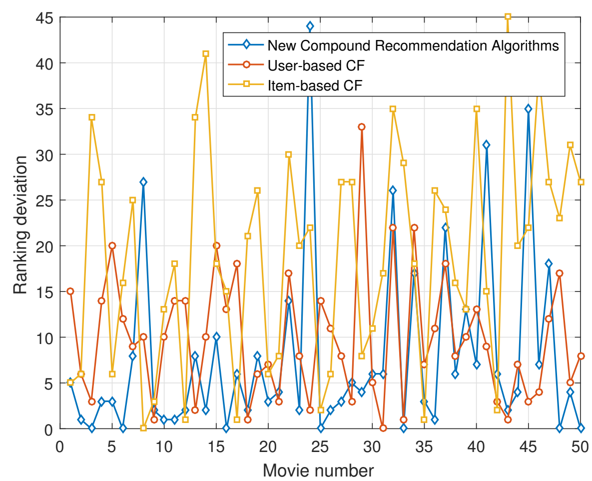


Fig. 8: Comparison of ranking deviation values obtained by different methods.

TABLE VI: DEVIATION COMPARISON

Average Deviation	New Compound Recommendation Algorithms	User-based CF	Item-based CF
score	0.29	0.33	0.65
ranking	7.68	9.80	18.84

Score deviation=|Score from algorithmic analysis - Score on the website|

Ranking deviation=|Ranking from algorithmic analysis - Ranking on the website|

The simulation results as shown in FIGURE 7 and FIGURE 8. The average deviation can be obtained by dividing the total deviation of all movies by the total number of movies.

TABLE VI reflects deviation values from the analysis results of different movies, where the smaller the numerical value, the better the analysis effect. The proposed new recommendation algorithm maps were mostly below those of the other methods.

Although deviations in the prediction scores or rankings of some movies were greater than the other two methods, the average deviation was superior to the original two algorithms. Therefore, the compound recommendation algorithm proposed in this paper can analyze user information more carefully, consider users' potential preferences more effectively, and produce more practical analysis results.

TABLE VII: FILE LIST FOR CONTRAST METHOD 1

Number	File Size(MB)	Scores in The Website
1	203.84	9.4
2	326.87	9.3
3	251.98	9.2
4	242.16	9.2
5	250.53	9.2
6	305.85	9.2
7	188.34	9.1
8	428.95	9.1
9	182.20	9.1
10	291.68	9.1
11	238.78	9.0

TABLE VIII: FILE LIST FOR CONTRAST METHOD 2

Number	File Size (MB)	Ranking	Scores
1	242.16	1	9.38
2	184.99	2	9.31
3	305.85	3	9.28
4	203.84	4	9.19
5	171.59	5	9.19
6	251.98	8	9.13
7	282.20	9	9.11
8	196.48	11	9.09
9	356.87	12	9.08
10	167.12	18	9.03
11	188.34	16	9.03
12	250.53	19	9.02
13	185.78	22	9.01

B. Comparison of Decision-making Results of MEC Intelligent Cache Analysis Platform

The MEC intelligent caching analysis platform developed in this paper can finally obtain an optimal list of cached files after a series of algorithmic decisions. To confirm that the proposed algorithm can produce effective optimization results, we compare our findings with two other decision-making methods.

Contrast Method 1: Film files with a total capacity of less than 3000MB were selected according to the original movie ranking on the website; this is the most common MEC intelligent cache processing method. File information is listed in TABLE VII.

Contrast Method 2: After obtaining the complete file ranking using the original user CF recommendation algorithm, movie files with a total capacity of less than 3000MB were selected using the genetic algorithm [8]. Data are shown in TABLE VIII.

We present the following scenarios for simulation.

Assuming 1000 users around the MEC server, 50 experiments were carried out, and the number of files requested increased gradually from 1 to 50 (i.e., 1 file was requested randomly in the first experiment, and N files were requested

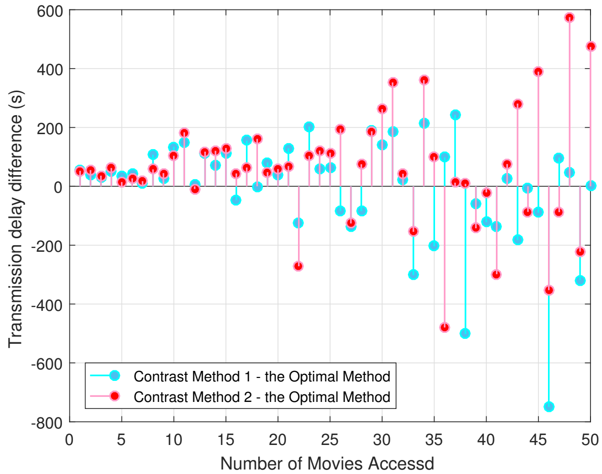


Fig. 9: Comparison of transmission delay differences between comparison method and optimal method.

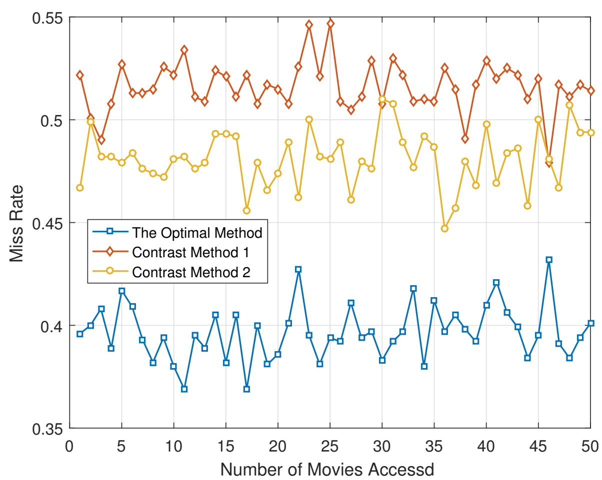


Fig. 10: Comparison of miss rates among decision-making methods.

randomly in the nth experiment). Files were requested randomly according to a certain probability (the movie's score on the website/the total score of 50 movies on the website). If the requested files were stored in the MEC server, we could obtain them at a rate of 5 MB/s; otherwise, we could obtain them through the core network at a rate of 1 MB/s [2]. The average request time of 1000 users in each experiment was calculated, and the delay parameters of 50 experiments were obtained. The following indicators were then compared:

1) Transmission time difference between the two comparison methods and the optimal method presented in this paper;
2) The probability of accessing movie files not in MEC under different decision-making methods (i.e., miss rate).

Simulation results are shown in FIGURE 12.

In the above scatter plot, if points were distributed on the regular axis, then the decision-making method in this paper saved transmission time compared with other methods under the number of movies accessed. After many experiments and simulations, we chose one to showcase the results. In a few cases, the two contrast methods achieved better results; in most cases, however, the optimal method in this paper conserved transmission time compared with other methods, suggesting that our proposed intelligent caching algorithm can optimize transmission delay more effectively.

FIGURE 10 displays a comparison of file miss rates. Although specific values of each simulation deviated, the overall trend was stable. The decision-making result of the MEC caching intelligent analysis platform in this paper was significantly lower than that of the two comparison methods, reducing the miss rate by about 10%. Although the decision-making effect of this scheme has room for improvement and optimization, compared with other methods, it presents a substantial enhancement; thus, this method can offer users more humanized intelligent service.

C. Integrated Scene Simulation of Cooperation of Network Coding and Wireless Relay in Mobile Edge Computing

Previously, the optimization effect of the MEC intelligent caching platform was demonstrated through simulation experiments. Then, the cooperative application scheme of network coding and wireless relay in MEC was simulated, and a comprehensive scenario was constructed to demonstrate the optimization effect of the proposed scheme.

Scenario conditions of the integrated simulation scheme were as follows:

1) Following experimental data from the MEC intelligent cache analysis platform, files in the final decision-making list were cached to the edge MEC from high to low scores in 50 movies; the rest remained in the core network.
2) Using the cooperative scheme of network coding and wireless relay, the transmission data volume of a normal communication file and MEC file was 2:1. The data transmission process was abstracted as a time-travel process.
3) Given access to normal communication files and MEC files with existing edges, a user's access rate was 5 MB/s; for access to MEC files stored in the core network, a user's access rate was 1 MB/s.

Coordinate axes of the simulation drawings were as follows:

1) The abscissa denotes the amount of data transmitted by normal files, ranging from 1 MB to 6000 MB (this can be understood as a time process, reflecting gradual transmission process of MEC files from 0 MB to 3000 MB via communication of normal files);
2) The vertical axis represents the average access rate of 100 peripheral users (MB/s).

The scene description is as follows:

Under the set scheme, normal file transfer occurred with the passage of time, and future MEC files were transferred synchronously. We simulated roughly 100 users, some of whom accessed normal communication files whereas others accessed MEC experimental files (the number of each type of user was generated randomly each time). Files were tested

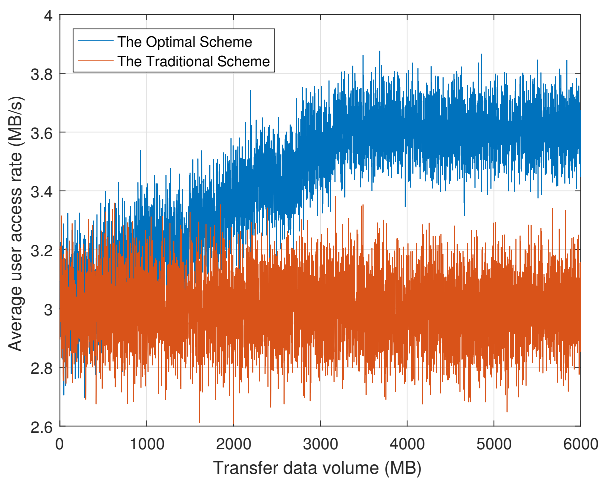


Fig. 11: Average user access rate comparison.

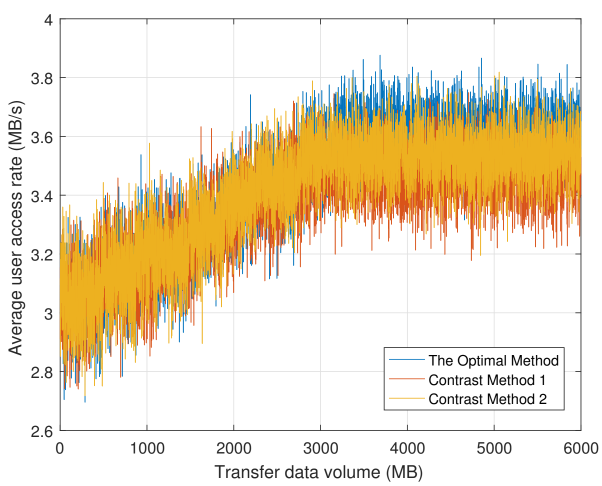


Fig. 12: Average user access rate comparison.

TABLE IX: AVERAGE TRANSMISSION TIME COMPARISON

Average Transmission Time comparison	The Optimal Method	Contrast Method 1	Contrast Method 2
Average User Access Rate for The Overall Process	3.44 MB/s	3.36 MB/s	3.40 MB/s

100 times, and the average access rate of 100 users at each horizontal point was counted.

We first compared the traditional mode without MEC and the network coding scheme. Then, we compared the optimization effect of the MEC decision-making method proposed in this paper with the two comparison methods from the previous section. Simulation results appear in FIGURE 11 and FIGURE 12.

The average value of the results in FIGURE 12 as shown in TABLE IX. Compared with the traditional scheme that does not introduce MEC, network coding, or wireless relay cooperation, the rate improvement from the optimal scheme in this paper became increasingly significant during the process of MEC file caching. In the final case, the optimal scheme in this paper could increase the average access rate of peripheral users from an average of 3 MB/s to about 3.6 MB/s, eliciting an optimization effect. The significant degree of this effect is subject to the setting of experimental parameters, but our experiment verified that the proposed cooperative application scheme can indeed bring more efficient network services to users compared with the current network operating mode.

Although fluctuations occurred in each experiment, the trend of the overall effect was essentially stable. We found that, as long as the MEC and network coding and wireless relay cooperation schemes were introduced, even different decision-making algorithms could elicit significant optimization results compared with the traditional network working mode. In addition, the average access rate of the optimal decision-making method proposed herein was slightly higher than that of the other two methods. These findings verify the practical significance of the cooperative application of network coding and wireless relay in MEC and confirm that the algorithm flow of the MEC intelligent caching analysis platform proposed in this paper can bring more benefits to users compared with the existing decision-making scheme. Our method can also result in a higher access rate and more humanized intelligent network services.

Finally, although the proposed scheme can improve users' network access rate and provide a more user-friendly experience with network services, further in-depth study is needed to ensure the proposed scheme can bring more significant improvements to 5G networks.

VI. CONCLUSION

The era of the 5G intelligent network is imminent. As we prepare for a more intelligent, humane, and efficient network life, we must also provide adequate technical support for the landing of 5G technology. The MEC method in this paper represents an important technology for the 5G fog wireless access network, and the cooperative application scheme of network coding and wireless relay described herein offers a technical scheme that can support 5G to realize high-speed transmission of large amounts of data. Therefore, our proposed cooperative application scheme of network coding and wireless relay in MEC is practical and feasible in today's era. In addition, results of our experimental simulation in MATLAB imply that the new compound recommendation algorithm, the algorithm flow of the MEC intelligent cache analysis platform, and the proposed cooperative application scheme of network coding and wireless relay can effectively improve network data transmission efficiency, reduce user access delay, and improve users' average network speed. This approach provides a novel method for improving users' network service experience. However, more work remains to be done to improve the effects of the methods studied in this paper. Our proposed approach is worthy of further investigation and optimization to contribute more efficient and reliable technical solutions to impending 5G communication networks.

REFERENCES

[1] Mach, Pavel and Zdenek Becvar. ”Mobile Edge Computing: A Survey on Architecture and Computation Offloading.” IEEE Communications Surveys & Tutorials 19 (2017): 1628-1656.

[2] Schiller, Eryk, Navid Nikaein, Eirini Kalogeiton, Mikael Gasparyan and Torsten Braun. ”CDS-MEC: NFV/SDN-based Application Management for MEC in 5G Systems.” Computer Networks 135 (2018): 96-107.

[3] Xiao, Yang, Xiaojiang Du, Jingyuan Zhang, Fei Hu and Sghaier Guizani. ”Internet Protocol Television (IPTV): The Killer Application for the Next-Generation Internet.” IEEE Communications Magazine 45 (2007): n. pag.

[4] Paschos, Georgios S., Ejder Bastug, Ingmar Land, Giuseppe Caire and Mrouane Debbah. ”Wireless caching: technical misconceptions and business barriers.” IEEE Communications Magazine 54 (2016): 16-22.

[5] Zeydan, Engin, Ejder Bastug, Mehdi Bennis, Manhal Abdel Kader, Ilyas Alper Karatepe, Ahmet Salih Er and Mrouane Debbah. ”Big data caching for networking: moving from cloud to edge.” IEEE Communications Magazine 54 (2016): 36-42.

[6] Zheng, Kan, Zhe Yang, Kuan Zhang, Periklis Chatzimisios, Kan Yang and Wei Xiang. ”Big data-driven optimization for mobile networks toward 5G.” IEEE Network 30 (2016): 44-51.

[7] Gao, Zhipeng, Zehui Lu, Nanjie Deng and Kun Niu. ”A novel collaborative filtering recommendation algorithm based on user location.” 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) (2016): 1-2.

[8] Yang, Ling, Ling Tian and Wei Hong. ”An ultra-broadband, high resolution frequency synthesizer.” 2012 International Conference on Microwave and Millimeter Wave Technology (ICMMT) 4 (2012): 1-4.

[9] Tao, Shaoguo, Jiaqing Huang, Zongkai Yang, Wenqing Cheng and Wei Liu. ”An Improved Network Coding-Based Cooperative Content Distribution Scheme.” ICC Workshops - 2008 IEEE International Conference on Communications Workshops (2008): 360-364.

[10] Peng, Cong, Qian Zhang, Ming Zhao and Yan Yao. ”Opportunistic Network-Coded Cooperation in Wireless Networks.” 2007 IEEE Wireless Communications and Networking Conference (2007): 3358-3363.

[11] Kramer, Gerhard, Michael Gastpar and Piyush Gupta. ”Cooperative strategies and capacity theorems for relay networks.” IEEE Transactions on Information Theory 51 (2005): 3037-3063.

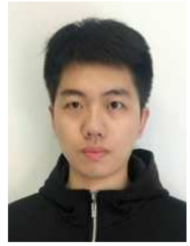
[12] Kanungo, Tapas, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman and Angela Y. Wu. ”An Efficient k-Means Clustering Algorithm: Analysis and Implementation.” IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002): 881-892.

[13] Zou, Zhi-hong, Yi Yun and Jing-nan Sun. ”Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment.” Journal of environmental sciences 18 5 (2006): 1020-3 .

[14] Vaidya, Omkarprasad S. and Sushil Kumar. ”Analytic hierarchy process: An overview of applications.” European Journal of Operational Research 169 (2006): 1-29.

[15] Lee, Zne-Jung, Shun-Feng Su, Chou-Yuan Lee and Yao-Shan Hung. ”A Heuristic Genetic Algorithm for Solving Resource Allocation Problems.” Knowledge and Information Systems 5 (2003): 503-511.

WEISHENG HE received his B.S. degree from Xiamen University, Xiamen, China. He is currently a graduate student in Xiamen University, Xiamen, China. His research interests include wireless communications, network coding, etc.



YUHAN SU received his B.S. degree from Huaqiao University, Xiamen, China. He is currently pursuing the Ph.D. degree in communication and information system at Xiamen University, Xiamen, China. His research interests include wireless communications, congestion control, network coding, etc.

XUETING XU received her B.S. degree from Nanjing University of Science and Technology, Nanjing, China. She is currently a graduate student in Xiamen University, Xiamen, China. Her research interests include wireless communications, massive MIMO, etc.

ZHAOHUI LUO (S’17) received the B.S. degree in Electronic & Information Engineering and the M.S. degree in Pattern Recognition and Intelligent System from Guangxi Normal University, China, in 2011 and 2015, respectively. Currently, he is pursuing the Ph.D. degree with the Department of Communication Engineering, Xiamen University, Xiamen, China. His research interests include Mobile Edge Computing with emphasis on edge caching and wireless communications.

LIANFEN HUANG received the B.S. degree in radio physics and the Ph.D. degree in communication engineering from Xiamen University, Xiamen, China, in 1984 and 2008, respectively.

She was a Visiting Scholar at Tsinghua University, Beijing, China, in 1997, and a Visiting Scholar at Chinese University of Hong Kong, Shatin, Hong Kong, in 2012. She is a Professor of Communication Engineering with Xiamen University. Her research interests include wireless communication, wireless network, and signal process.

XIAOJIANG DU (SM'09) received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park in 2002 and 2003, respectively. He is currently a Professor with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA. He received over U.S. 5 million research grants from the U.S. National Science Foundation, the Army Research Office, Air Force, NASA, the State of Pennsylvania, and Amazon. His research interests include security, wireless networks, and systems. He has authored over 320 journals and conference papers in these areas, as well as a book published by Springer. He is a Life Member of the ACM. He received the Best Paper Award at the IEEE GLOBECOM 2014 and the Best Poster Runner-Up Award at the ACM MobiHoc 2014. He served as the Lead Chair of the Communication and Information Security Symposium of the IEEE International Communication Conference 2015 and a Co-Chair of the Mobile and Wireless Networks Track of the IEEE Wireless Communications and Networking Conference 2015.