SKCFS: Stochastic Optimal Forwarding Strategy for Knowledge Centric Network

SHUJIE YANG1, HAIYONG XIE2, YIFENG LIU2, ZHONGBAI JIANG1, LUJIE ZHONG3
1State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, 100876, China
2Innovation Center
China Academy of Electronics and Information Technology
Beijing, 100041, China
3Information Engineering College
Capital Normal University
Beijing, 100048, China
E-mail: zongshi@bupt.edu.cn; haiyong.xie@ieee.org; yliu@csdslab.net; zbjiang@bupt.edu.cn; zhonglj@cnu.edu.cn

The popularization of intelligent equipment, following the rapid development of wireless communication technology, has also promoted the demand for bandwidth occupancy services such as video streaming. The technology combining KCN technology and mmWave technology can meet the demand of rich bandwidth and large rate transmission in the future. However, the combination of these two technologies will bring novel challenges in forwarding strategies, routing algorithms and so on. In this paper, according to the characteristics of the KCN network with mmWave communication, a stochastic optimization algorithm is proposed for the multi-path content forwarding problem aiming at the base station with the function of sending mmWave and in-network caching. Then the Lyapunov stability equation is adopted to further analyze this optimal problem. Through the optimization of the queue problem, the problem is further decomposed into three optimal sub-problems. At the same time, each physical meaning of these three decomposed problems is shown. At last, extensive simulations are performed in terms of provided stochastic optimal forwarding algorithms.

Keywords: KCN, forwarding strategies, Millimeter-Wave Communication, Lyapunov queue stability

1. INTRODUCTION

The popularization of intelligent equipment, following the rapid development of wireless communication technology, has also promoted the demand for bandwidth occupancy services such as video streaming. It is very predictable that when 5G technology arrives, mobile data will enter an explosive growth era, which implies the main function of mobile Internet be from the host communication to the content distribution. Content delivery networks (CDN) as a current popular method of nearby content fetching have been widely used by replicating the content copies at network edges as well as the peer-to-peer networks [1-4]. However, these solutions are still based on IP network which focus on host-centric and results in a low efficiency of large scale content distribution[5-7].

In this context, knowledge centric networking (KCN) becomes a promising solution that content takes the place of host for routing and in-network was employed which can increase the efficiency of content lookup and the speed of content distribution. Different with traditional IP network, users obtain the knowledge by sending the interested packet

Communicated by the editor.
for content lookup and content providers will return data packet if they receives the interested packet. In-network caching is another feature since each node caches the content which can reduce the latency of content access and improve the rate of content distribution as well. Users will find out the nearby content copies through the name-based routing technology in KCN, which improves caching utilization than normal host-centric methods.

Then, as a novel architecture, KCN has been provided recently [8,9], which possesses the two feathers of name routing and in-networking caching. Besides, in KCN, it need to meet four scenario requirement, the continuous wide-range coverage scenario, the hot spot high volume scenario, the low power consumption scenario, and the low time-delay but reliability scenario, respectively. Therefore, lager caching capacity and higher data transmission rate are desirable. All of these bring enormous challenge for KCN to overcome the shortage of limited wireless bandwidth.

Millimeter-Wave (mmWave) communication provides a solution for these challenges. The traditional wireless band is 300MHZ to 3GHz, while millimeter-wave communication can be carried out between 3G-300GHz, which increased more than 200 times available bandwidth [10]. Therefore, mmWave possesses extremely high transmission rate and rich available bandwidth resources [11]. However, mmWave communication has three major problems. First, the attenuation is large and the packet loss rate is very high, because the spatial range of free transmission is proportional to the carrier frequency, namely, the higher the carrier frequency is, the smaller the spatial range [12]. Secondly, because of the wave length of mmWave is very short (generally, 10mm-100mm), it is difficult to diffract around obstacles [13]. Finally, mmWave signals are difficult to penetrate solid materials [14]. Based on the analysis above, in order to make good use of mmWave characteristics in 5G-KCN, a high-gain directional antenna is in need to compensate for the tremendous propagation loss and reduce the shadowing effect [15].

Through the analysis above, the technology combining KCN technology and mmWave technology can meet the demand of rich bandwidth and large rate transmission in the future. However, the combination of these two technologies will bring novel challenges in forwarding strategies, routing algorithms and so on because of these new characteristics, such as, routing by name, in-networking caching and directed forwarding content and so on. At the same time, how to balance the tradeoff between stringent quality of service (QoS) requirements and energy efficiently consumption in the above-mentioned architecture is an urgent problem to be solved. In this paper, according to the characteristics of the KCN network, a stochastic optimization algorithm is proposed for the multi-path content forwarding problem aiming at the base station with the function of in-network caching. Then the Lyapunov stability equation is adopted to further analyze this optimal problem. The main contributions to this article are shown below.

- Considering a large number of multimedia streaming applications, as well as the demand of available bandwidth in future, we first propose a novel network architecture through combing the KCN technology and mmWave technology. Then the novel architecture features are analyzed.
- Based on the characteristics of the two technologies, this paper studies the content forwarding problem of KNC network, and provide a stochastic optimization model to realize the multi-path content distribution on the base station.
- By using Lyapunov stability equation, the above optimization problem is transformed into queue stability problems. Through the optimization of the queue problem, the problem is further decomposed into three optimal sub-problems. At the
same time, each physical meaning of these three decomposed problem is shown.

The remaining sections are laid out as follows. In Section 2, an overview of the KCN is shown in Section 3. In Section 4, several models have been setup which provide essential information for our strategy. Based on this, in Section 5, we analyze the implementation of the SKCFS in detail. And Simulations are shown in Section 6. At last, Section 7 concludes the paper.

2. OVERVIEW OF KCN

In this section, according to the related technology features, KCN network based on millimeter wave is introduced in detail shown in Fig. 1.

The routing process of Knowledge Centric Networking (KCN) mainly involves three kinds of data structure: Knowledge Forwarding Table (KFT), which stores the mapping relationship between the knowledge network names and the knowledge source ports. Pending Requested Table (PRT) stores the mapping of a knowledge network name that has not yet received a response from the request packet source port, and the Knowledge Storage Table (KST) records the local knowledge cache information for the Knowledge interconnection node.

The knowledge consumer broadcasts the request package after getting the requested knowledge network name. The node which the request package arrives on first queries the KST and returns the response packet directly if the corresponding knowledge is found in the local cache; otherwise, the knowledge network name and the requested package source port are recorded in PRT. Then looking up in KFT, and if find the source port for the corresponding knowledge, the node will forward the request package to that port, otherwise the node continues to broadcast the request packet until the requested package arrives at the producer or cache node that owns the request knowledge. The producer or cache node encapsulates the response packet and returns to the consumer along the path that the requested package arrives. The Knowledge Network name and source port of the Knowledge Interconnect node that receives the response information will be recorded in the KFT, and update local knowledge base and KST according to cache policy.

The routing mechanism of Knowledge Centric Networking (KCN) supports the aggregation and integration of the requested knowledge. The routing node merges the requested package for the same knowledge, and only sends the first requested packet to the
upstream knowledge interconnect node. Each entry in PRT records the requested knowledge network name and a set of ports that receive the same requested knowledge. When the response packet arrives at the router node, it matches all matching request entries, and then forwards the data according to the port list in the entry.

3. SYSTEM MODEL

3.1 Problem Description

As shown in Fig. 2, the problem we care about can be described as following. We assume in KCN networks with mmWave, many mobile users send Interest packets (assuming a request to obtain HDTV video service) to a KCN base station (KCN-BS) with in-networking caching capabilities to obtain the request content. The Interest packets will be forwarded in accordance with the FIB table if the KCN-BS not meet the request. At the same time, the KCN BS will record the path in its own PIT table. When the forwarded Interest packets obtain the corresponding requested content, a Data packet containing the request content will be sent back to the KCN BS along the reverse path that Interest packets transmit. Considering the two practical factors. The first is that there are large amount of content will be sent back to this KCN BS every time. The other is that this KCN BS leverage mmWave communication to forward Data packet (namely, directional straight line transmission mode), which need to communicate with a user after completing the communication of previous user. Based on these two reasons above, packets that need to be forwarded in KCN BS will form queues. Therefore, in this paper we will study the problem of multi-path content distribution based on KCN BS with mmWave communication. Through the research of this problem, a stochastic optimal forwarding strategies is provided, which can not only satisfies the user QoS, but also realize energy efficient forwarding.

Fig. 2. Problem Description.

3.2 Model Setup

In order to be able to analyze problem described above better, in this part, we first abstract the problem to mathematical models. We briefly summarize some key parameters in Table I in order to improve the readability. Fig. 2 can be divided into three parts, the content source $O$, the KCN-BS, and the destination user node set $U$, respectively. We let
Table 1. Some Key symbols.

<table>
<thead>
<tr>
<th>parameter</th>
<th>definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>mobile user node set</td>
</tr>
<tr>
<td>$O$</td>
<td>the multimedia content source set</td>
</tr>
<tr>
<td>$t$</td>
<td>the time slot and each slot have the same size</td>
</tr>
<tr>
<td>$a_i(t)$</td>
<td>the number of Data packets requested by user $i$</td>
</tr>
<tr>
<td>$P(x)$</td>
<td>random variable $x$ obeys Poisson distribution</td>
</tr>
<tr>
<td>$d_i(t)$</td>
<td>the distance between users $i$ and KCN BS</td>
</tr>
<tr>
<td>$PL(x)$</td>
<td>the pass loss on wireless link with distance $x$</td>
</tr>
<tr>
<td>$H_i(t)$</td>
<td>the transfer rate of each mobile user $i$ at time slot $t$</td>
</tr>
<tr>
<td>$r_i(t)$</td>
<td>the number of Data packets obtained by user $i$ at time $t$</td>
</tr>
<tr>
<td>$P_i(t)$</td>
<td>the packets loss rate on the wireless link</td>
</tr>
<tr>
<td>$Q_i(t)$</td>
<td>the dynamic updating queue backlog produced by user $i$ at time slot $t$</td>
</tr>
<tr>
<td>$\phi(x)$</td>
<td>the utility function of variable $x$</td>
</tr>
<tr>
<td>$R_i(t)$</td>
<td>the auxiliary queue backlog produced by mobile user $i$ at time slot $t$</td>
</tr>
<tr>
<td>$e_i(t)$</td>
<td>the auxiliary variable meeting $e_i(t) &lt; a_i(t)$</td>
</tr>
<tr>
<td>$L(t)$</td>
<td>the Lyapunov queue function at time slot $t$</td>
</tr>
</tbody>
</table>

$U = \{1, 2, \cdots, m\}$ to denote there are $m$ users send Interest packet for the same content at time $t$. For the sake of analysis, time are divided into slots with the same size. The time slot is denoted as $t = 0, 1, \cdots, \tau, \cdots$, where $\tau$ is a positive integer number. Now, we assume the KCN BS is providing service for $m$ users at time slot $t$ and the queue backlog of Data packet is $Q(t)$. Assume source node $O$ return Data packet continuously, and the Data packet can only forwarded by KCN BS to one user at each time slot $t$. Let $a_i(t), (1 < i < m)$ denote the number of Data packets requested by user $i$. Assume $a_i(t)$ following Poisson distribution of intensity $\lambda_i$ and can be denoted as follow:

$$P[a_i(t) = k] = \frac{e^{\lambda_i} (\lambda_i)^k}{k!}, k = 1, 2, \cdots, A_{max},$$  \hspace{1cm} (1)

where $A_{max}$ satisfies $P[a_i(t) = A_{max}] \to 0$.

Besides, assume the distance between users $i$ and KCN BS can be denoted as $d_i(t)$. Then the path loss on wireless link can be denoted as follow:

$$PL(d_i(t)) = 44.2 + 21.5 \log (d_i(t)) + 20 \log \frac{f_c}{5 \times 10^9},$$  \hspace{1cm} (2)

where $f_c$ is the carrier frequency. Therefore, the signal-to-noise ratio (SNR) between user $i$ and KCN BS can be denoted as

$$SNR = P_t - PL(d_i(t)) - N_0,$$  \hspace{1cm} (3)

in which, $P_t$ is the transmission power of KCN BS. Then the transfer rate of each mobile user $i$ can be denoted as follow:

$$H_i(t) = W \log_2 (1 + SNR),$$  \hspace{1cm} (4)

where, $W$ denotes system bandwidth.

Since we adopt discrete time, we need to assume that the length of each discrete moment is the same given by $t_s$ second. At each time slot $t$, the transfer rate of every user
is constant. Therefore, the number of Data packets obtained by user $i$ at time $t$ can be given by equation 5

$$r_i(t) = \frac{H_i(t) \cdot t_s}{G} \cdot P_i(t),$$

where $G$ is the Data packet size, and $P_i(t)$ represents the packet loss rate on the wireless link (generally produced by random number).

Therefore, for every user $i$, the dynamic updating queue backlog can be defined as follow equation:

$$Q_i(t+1) = \max [Q_i(t) - r_i(t), 0] + a_i(t).$$

Then we can adjust the length of the caching queue backlog in the KCN-BS by controlling the size of the $a_i(t)$. Considering every mobile user request HDTV video service in our scenario, we use the number of receiving Data packet at every time slot $t$ to reflect user satisfaction with the service, namely the QoS for every user given as follow:

$$\phi(\bar{a}_i) = \ln(1 + \beta \bar{a}_i)$$

where $\bar{a}_i = \lim_{T \to \infty} \frac{1}{T} \times \sum_{t=0}^{T} a_i(t)$ and $\beta$ is a positive constant.

Therefore, problem described in 3.1 can be treated as the follow optimal problems, namely,

$$\max \sum_{i=1}^{m} \phi(\bar{a}_i)$$

subject to $$\bar{Q} < \infty$$ (8)

Solving the problem (8) has many difficulties, because the function of time average is involved in the optimization objective, and it is nonlinear. thus, we will try to solve the problem by exploiting the theory of optimizing functions of time averages in the next stage.

4. ANALYSIS OF PROBLEM

In this section, a stochastic optimal algorithm is proposed based on Lyapunov Optimization to solve the forwarding strategies in the previous section.

4.1 Transformation Process

Since $\phi(\bar{a}_i)$ in (8) is a concave function, we can transfer this problem to another problem for analysis easily. At first, we construct an auxiliary variable $e_i(t)$, meeting $e_i(t) < a_i(t)$. Based on this, an auxiliary function is constructed:

$$R_i(t + 1) = \max [R_i(t) - a_i(t), 0] + e_i(t).$$

As a result, the optimal problem (8) can be transformed as follows:

$$\maximize \sum_{i=1}^{m} \phi(e_i)$$

subject to $$\bar{Q} < \infty$$

$$\bar{R} < \infty$$ (10)
To solve problem (10), we can construct a Lyapunov function and is defined as:

\[ L(t) = \frac{1}{2} \sum_{i=1}^{m} [Q_i^2(t) + R_i^2(t)] \]  

(11)

Then the Lyapunov drift can be defined as

\[ \Delta(t) = L(t+1) - L(t) \]

\[ = \frac{1}{2} \sum_{i=1}^{m} [Q_i^2(t+1) - Q_i^2(t)] + \frac{1}{2} \sum_{i=1}^{m} [R_i^2(t+1) - R_i^2(t)] \]  

(12)

in which,

\[ Q_i^2(t+1) - Q_i^2(t) \leq r_i^2(t) + a_i^2(t) - 2Q_i(t) \{r_i(t) - a_i(t)\} \]  

(13)

In the similar way, we have

\[ R_i^2(t+1) - R_i^2(t) \leq a_i^2(t) + e_i^2(t) - 2Q_i(t) \{a_i(t) - e_i(t)\}. \]  

(14)

Then equation (12) can be further rewritten as follow

\[ \Delta(t) \leq \frac{1}{2} \left[ r_i^2(t) + 2a_i^2(t) + e_i^2(t) \right] - \sum_{i=1}^{m} \left\{ Q_i(t) \{r_i(t) - a_i(t)\} + R_i(t) \{a_i(t) - e_i(t)\} \right\} \]  

(15)

The first term of (15) can be proven to have a upper limit value, denoted by \( B \). Considering optimizing goals at the same time, we have

\[ \Delta(t) - V \sum_{i=1}^{m} \phi(e_i(t)) \leq \]

\[ B - \sum_{i=1}^{m} \left[ Q_i(t)r_i(t) - Q_i(t)a_i(t) + R_i(t)a_i(t) \right. \]

\[ - R_i(t)e_i(t) + V \phi(e_i(t)) \]  

(16)

where \( V \) is a constant to get the data transmission utility.

4.2 Subproblem Optimization

Based on analysis above, to minimize \( \Delta(t) \), we need to maximize the follow value

\[ \sum_{i=1}^{m} \left[ Q_i(t)r_i(t) \right. \]

\[ + [R_i(t) - Q_i(t)]a_i(t) \]

\[ - R_i(t)e_i(t) + V \phi(e_i(t)) \]  

(17)

The problem then can be divided into three suboptimal problems (each line term stand for a suboptimal problem). Through a further solution to these three problems, we discovery they exactly correspond to three practical process, including transmission control problem, accept control problem and utility maximization problem, respectively. By solving these three problems, we can obtain the optimal solution of problem (8).
1) Transmission Control Problem: Through solving the first line term of equation (17), we can obtain the transmission control problem, given by
\[
\max \sum_{i=1}^{m} Q_i(t)r_i(t) \tag{18}
\]
To obtain the solution of this problem, we just need to pick the maximum value of \(r_i(t)\) at each time \(t\). Then through the solving process of the subproblem (18), the KCN BS will determine which user should be transfer the Data packets at each time slot \(t\). Let \(i^*\) denote the user that get the transmission data at time \(t\). We can get the algorithm 1.

\begin{algorithm}
\caption{Transmission Control Problem}
\begin{algorithmic}[1]
\State \textbf{Input}: The number of Data packets obtained by all users at time \(t\), namely, \(r_i(t), i = 1, 2, \ldots, m\). \textbf{Output}: Forwarding Data packets to user with index \(i^*\).
\State Initialize \(\text{argmax} = 0, i^* = 0\); \State for \(i = 1\) to \(m\) do \State \text{calculate} \(r_i(t)\); \State \text{if} \(r_i(t) > r_{i-1}(t)\) then \State \(i^* = i;\) \State \text{end if} \State end for \State for \(i = 1\) to \(m\) do \State \text{if} \(i \neq i^*\) then \State \(r_i(t) = 0;\) \State \text{end if} \State end for \State return \(r_{i^*}(t)\);
\end{algorithmic}
\end{algorithm}

2) Accept Control Problem: Through solving the second line term of equation (17), we can obtain the admission control problem as follow:
\[
\max \sum_{i=1}^{m} [R_i(t) - Q_i(t)]a_i(t) \tag{19}
\]
This suboptimal problem response for KCN BS to decide if accept Data packets from content service \(O\) according to its congestion situation at time slot \(t\). From equation (19), we can know, when \(R_i(t) - Q_i(t) \geq 0\), \(a_i(t)\) is just right the number of Data packet coming. Then equation (19) can be solved according Algorithm 2.

3) Utility Maximization Problem: Through solving the third line term of equation (17), we can obtain the utility maximization problem can be write as:
\[
\max \sum_{i=1}^{m} -R_i(t)e_i(t) + V\phi(e_i(t)) \tag{20}
\]
This suboptimal problem can be describe as a utility maximization problem. That is according to congestion situation of KCN BS, we can get the auxiliary variable value \(e_i(t)\) through solving the equation (20). The solution steps are listed as follows. Firstly, let \(e_i(t) = x\), then we have
\[
f(x) = V\ln(1 + \beta x) - R_i(t)x \tag{21}
\]
Algorithm 2 Accept Control Problem

Input: Queue $Q_i(t)$ and $R_i(t)$ in KCN BS at time slot $t$
Output: Accept Data packets form service with index $i^*$
1: Initialize $\text{argmax} = 0$, $i^* = 0$;
2: for $i = 1$ to $m$ do
3: if $R_i(t) - Q_i(t) > 0$ then
4: $a_i^*(t) = a_i(t)$;
5: else $a_i^*(t) = 0$;
6: end if
7: end for
8: return $a_i^*(t)$;

Then by the derivative of equation (21), we have

$$\frac{df(x)}{dx} = \frac{Vx}{1+\beta x} - R_i(t)V \ln(1+\beta x) - R_i(t)x$$

(22)

Let $\frac{df(x)}{dx} = 0$, we have

$$x = \frac{R_i(t)}{V - R_i(t)\beta}$$

(23)

Based on analysis above, the solution of suboptimal problem (20) is detailed in Algorithm 3.

Algorithm 3 Utility Maximization Problem

Input: Queue backlogs of $R_i(t)$ at time slot $t$
Output: the auxiliary variable value $e_i(t)$ of every user.
1: Initialize $\text{argmax} = 0$;
2: for $i = 1$ to $m$ do
3: calculate $e_i(t)$ of user $i$ according to equation (23);
4: end for
5: return $e_i(t)$;

Through Algorithm 1, Algorithm 2, and Algorithm 3, we can solving these three suboptimal problems well. Then we can obtain the value of $r_i(t), a_i(t)$ and $e_i(t)$ of every user. Then the queue backlog of $Q_i(t)$ and $R_i(t)$ can be update by equation (6) and (9).

5. PERFORMANCE ANALYSIS

In this section, we first introduce the comparison algorithm and simulation environment. Then we provide a detailed performance analysis of the proposed SKcSA.

To demonstrate the performance of out proposed algorithm, we select the Round Robin Algorithm (RRA) as the comparison algorithm. RRA transmit knowledge the mobile users sequentially. Even though it is a simple algorithm, it is utilized widely in network communications due to its ease of implementation. Therefore, we select it as the comparison algorithm.

During the simulation, we assume there are 10 mobile users are communicating with BS. And the average number of arrived packets is $10^6$. Since the mm-WAVE technology is adopted, the system bandwidth is set as 60GHz.
We first compare the queue length of BS. As shown in Fig. 3, the queue length of SKCFS vibrates in the range of $[4 \times 10^4, 1 \times 10^5]$. However, the queue length of RRA is out of control. This is because SKCFS employs effective admission control and transmission scheduling algorithms. It is able to control the queue length efficiently. But these algorithms are absent in RRA. When the users are using the data-consumption service, it can not perform necessary control of the arrived data packets. It is worth noting that the longer queue length at BS means the packets will experience longer transmission delay. In this way, it also demonstrates that SKCFS will reduce the transmission delay of data packets.

Fig. 3. Comparison of Queue Length

Fig. 4 shows the impact of parameter $V$ on queue length at BS. It can be observed that with the increase of $V$, the average queue length is also increasing. Recalling that $V$ means the significance of QoE of users, the increase of $V$ denotes that the SKCFS place more attention on the QoE. It would result in more data packets needed to be transmitted and forwarded. Then the average queue length will increase. It is noteworthy that when the $V$ increase from 1 to 20, the average queue length increases obviously. After $V$ reaches to 20, the average queue length does not increase anymore.

Fig. 4. Impact of $V$ on Queue Length
Fig. 5 shows the impact of parameter $V$ on the number of accepted packets. At first, it can be observed that both the accepted packets and queue length shows the same increasing tendency. The number of accepted packets increases when $V$ increase from 1 to 20. After $V$ arrives at 20, the number of accepted packet nearly remains the same. This is because with the increase of $V$, the SKCFC put more attention on the QoE of mobile users. Therefore, the number of accepted packets should be increased to improve the quality of knowledge.

From Fig. 4 and Fig. 5, it also demonstrates that there exists a tradeoff between the service quality and transmission delay. This is because the improved service quality would incur more data packets to be transmitted. However, with the increase of packet number, the queue length of BS will also increase as shown in Fig. 4 and the packets will experience longer transmission delay. Therefore, it can be concluded that there exists a tradeoff between the service quality and transmission delay.

6. CONCLUDING

In this paper, according to the characteristics of the KCN network, a stochastic optimization algorithm is proposed for the multi-path content forwarding problem aiming at the base station with the function of in-network caching. Then the Lyapunov stability equation is adopted to further analyze this optimal problem. Through the optimization of the queue problem, the problem is further decomposed into three optimal sub-problems.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China (NSFC) under Grant Nos. 61522103, 61372112, in part by the joint advanced research foundation of CETC (No.20166141B08010102).
REFERENCES

Author one

Author two