Computer Networks 123 (2017) 88–103

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

RISC: ICN routing mechanism incorporating SDN and community division

Jianhui Lv[a], Xingwei Wang[b,*], Min Huang[c], Junling Shi[a], Keqin Li[d], Jie Li[e]

[a] College of Computer Science and Engineering, Northeastern University, Shenyang 110169, China
[b] College of Software, Northeastern University, Shenyang 110169, China
[c] College of Information Science and Engineering, Northeastern University, Shenyang 110819, China
[d] Department of Computer Science, State University of New York, New York 12561, USA
[e] Department of Computer Science, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan

ARTICLE INFO

Article history:
Received 28 November 2016
Revised 26 March 2017
Accepted 9 May 2017
Available online 10 May 2017

Keywords:
Information-centric networking
Software-defined networking
Community division
Routing mechanism
Clustering analysis

ABSTRACT

Information-Centric Networking (ICN) is one promising architecture paradigm which is a profound shift from address-centric communication model to information-centric one. Although ICN routing has attracted much attention from researchers, there are few researches on improving it inspired by other fields. In this paper, we propose a Routing mechanism for ICN incorporating Software-Defined Networking (SDN) and Community division (RISC), by decoupling control plane from data plane and dividing ICN topology into different communities. Firstly, we propose a community division scheme based on maximal tree in order to help retrieve the content conveniently and effectively. Secondly, we place all information about contents and forwarding into the corresponding information center for the centralized management. Thirdly, we design a routing mechanism which consists of intra-community routing based on same community information and inter-community routing based on social relationship among communities. Finally, the experimental results show that the proposed RISC not only speeds up content retrieval but also outperforms existent methods.

1. Introduction

The current Internet has been suffering from some significant limitations such as those brought by mobility and content distribution, which are inherent and inevitable due to its address-centric communication model [1]. At present, users widely use mobile devices to generate and consume digital media contents, producing large volumes of traffic, which is hard to be undertaken by the current Internet efficiently. For example, Cisco Visual Networking Index (CVNI) predicts that the annual global mobile data traffic is expected to reach 366.8 EB by 2020 from 44.2 EB in 2015 [2], and that the could traffic in Middle East and Africa region is expected to have the highest compound average growth rate (43%) by 2020 [3]. Furthermore, most users are mainly interested in accessing the content irrespective of its physical location with the development of mobile, cloud and big data technologies. For example, CVNI predicts that mobile users are expected to reach 5.5 billion, that is, 70% of the global population, by 2020, and personal mobile devices are expected to reach 8.5 billion by 2020 [2]. Therefore, the current Internet is becoming inadequate to the newly emerging usage patterns [4].

In order to facilitate the development of Internet continuously, many countries (e.g., USA, Germany, Japan and China) have paid high attention to the design of the future Internet [5], and Information-Centric Networking (ICN) is one of the promising networking paradigms. As we know, the research of ICN has been supported by some projects, such as Data-Oriented Network Architecture (DONA) [6], Content-Centric Networking (CCN) [7], Publish/Subscribe Internet Technology (PURSUIT) [8], Named Data Networking (NDN) and its Next Phase (NDN-NP) [9]. Especially, the first NDN community meeting (2014, Los Angeles) [10], the second one (2015, California) [11] and the third one (2016, Kyoto) [12] have highlighted the importance of ICN. As a clean-state networking paradigm, ICN is a profound shift from IP address-centric communication model to information-centric one [13], where content-aware routing and in-network caching are two notable features currently considered [14]. The content-aware routing delivers Interest packets to the content provider which returns the content to the interest requester by Data packets. The main goal

* Corresponding author.
E-mail addresses: lvjianhui2012@163.com (J. Lv), wangxw@mail.neu.edu.cn (X. Wang), mhuang@mail.neu.edu.cn (M. Huang), shijunlingdhr@163.com (J. Shi), lik@newpaltz.edu (K. Li), lijie@cs.tsukuba.ac.jp (J. Li).

http://dx.doi.org/10.1016/j.comnet.2017.05.010
1389-1286/

of in-network caching is to cache the frequently used contents as long as possible since they are expected to be requested by other users. The cached content replicas are capable of quick response to interest requests, which reduces bandwidth usage and latency, thus improving routing efficiency. However, ICN does not properly have an efficient routing scheme for the large-scale production network, which remains one of its major issues to be solved for the future deployment. In fact, the lack of scalable routing scheme is one of the main obstacles which slows down the large-scale deployment of ICN. Therefore, the design of efficient ICN routing mechanism has attracted much attention from a lot of research communities.

Some ICN routing mechanisms have been proposed, for example, flooding [15,16] and forwarding Interest packets according to the "best" outgoing interface(s) [17–23]. However, the flooding usually overloads network, wastes network resources and reduces network performance. It is usually unfeasible in ICN to find the "best" outgoing interface(s) to forward Interest packets with network scale, time scale and Forwarding Information Base (FIB) size considered [22]. Besides, the above mentioned two kinds of routing schemes are hard to realize the global control of contents and retrieve the closest content replica. In recent years, Software-Defined Networking (SDN) has been treated as enabling the new routing scheme development [24], and it can realize the global awareness of contents effectively in the centralized management manner [25]. Furthermore, the scalability of ICN routing is also an important issue which is very difficult to be overcome by the above mentioned two kinds of routing schemes. In [20,26,27], the community division has been proposed to address routing scalability. Moreover, ICN pays attention to user interests, thus the community division based on interest similarity relation can speed up content retrieval. Based on the above statements, in this paper, we adopt the basic ideas of SDN and community division to devise an efficient ICN routing mechanism.

As we know, SDN is considered as one promising networking paradigm to overcome the limitations of the current network architectures, and it was originally conceived at Stanford University [28]. In SDN, the control is decoupled from forwarding, and the network is directly programmable according to Open Networking Foundation (ONF) [29]. On one hand, SDN breaks the vertical integration by separating control plane from data plane; on the other hand, network switches just have the forwarding function whilst controllers are responsible for managing the whole network in the centralized manner [28]. The combination of the above SDN features can improve network management, control and data handling effectively [30–33]. Especially in the industry, Yahoo, Google, Facebook, Cisco and Microsoft have paid attention to SDN development. As two networking paradigms, ICN and SDN have profound impact on the future Internet, thus introducing the ideas of SDN into ICN is becoming a hot research topic, which is also done in this paper.

The community division [34,35] can be used to improve packet forwarding efficiency in ICN routing and help retrieve the content for interest requests conveniently and effectively. In fact, Content Routers (CRs) in the same CommUnity (CoM) usually have the same or similar features (e.g., interest categories). In this case, if the content is retrieved from any CR in its CoM, there is no need to forward Interest packets to CRs in another CoM, and thus avoiding redundant interest forwarding. When and only when the content cannot be retrieved from all CRs in one CoM, Interest packets will be forwarded to CRs in another CoM. It is obvious that the community division can improve routing efficiency and reduce routing overhead. Thus, in this paper, we divide ICN topology into CoMs according to the cached contents in CRs.

This paper proposes a novel Routing mechanism for ICN incorporating SDN and Community division (RISC), which decouples control plane from data plane by introducing SDN at the same time divides ICN topology into CoMs by introducing community, and the main contributions are summarized as follows.

(i) We propose a system framework of RISC which consists of SDN controllers, CoMs and Information Centers (ICs). RISC has the control plane which consists of SDN controllers to make routing decision, and the data plane which consists of CoMs and ICs to forward packets.
(ii) In order to help retrieve the content conveniently and effectively, we propose a novel community division scheme based on maximal tree, which is one clustering analysis method, to divide ICN topology into CoMs.
(iii) In terms of each CoM, we place all information about contents and forwarding into the corresponding IC for the centralized management, which is controlled and managed by SDN controller.
(iv) We propose a new routing mechanism to forward Interest packets and Data packets, which consists of intra-community routing based on same CoM information and inter-community routing based on social relationship among CoMs.

The rest of this paper is organized as follows. Section 2 reviews the related work. Section 3 proposes the system framework of RISC. In Section 4, the clustering analysis based community division is presented. Section 5 introduces the centralized information management. The routing decision of RISC is done in Section 6. Section 7 shows simulation results. Finally, this paper is concluded in Section 8.

2. Related work

2.1. ICN routing

The routing plays an important role in ICN since it determines packet forwarding and content retrieval, thus a lot of ICN routing mechanisms have been proposed. According to what we have learnt, in [13], the flooding algorithm was firstly proposed to address ICN routing issue. Although it could guarantee retrieving the content by sending Interest packets via all available outgoing interfaces, it would cause the heavy message traffic. Therefore, in [15], a cluster-based routing algorithm was proposed to reduce interest flooding. In [16], a flooding-based route discovery mechanism was proposed by considering the relationship among chunks, where the content was divided into a number of chunks. Although [15,16] reduced overhead brought by the successive flooding to some extent, it was hard to manage the degree of flooding.

A number of routing proposals about forwarding Interest packets according to the "best" outgoing interface(s) have been studied. In [17], an adaptive interest forwarding strategy was proposed by ranking different outgoing interfaces and selecting an outgoing interface to forward Interest packets. In [18], a named-data link-state routing protocol was proposed to announce the name prefixes. It used OSPF to propagate link-state advertisements throughout the entire network, which leaded to that the corresponding route computation no longer produced just one shortest path. In [19], a greedy ant colony forwarding algorithm was proposed to provide the intelligent interest forwarding, which improved cost-efficiency and scalability for CCN. In [20], an adaptive forwarding strategy based on probability was presented. It used Ant Colony Optimization (ACO) to compute the probability of path selection and used the statistical mode to compute the timeout for retransmission. In [21], a priority-based interest forwarding strategy was designed, in which Interest packets with higher priority were forwarded prior to those with lower priority. In [22], an exploration and exploitation based interest forwarding strategy was introduced. It spanned over two extremes, i.e., the deterministic exploitation towards a known replica and the random exploration towards an unknown

Table 1
Abbreviations in alphabetical order.

Abbreviation	Full name	Abbreviation	Full name
AD	Average delay	ADSR	Average delivery success rate
ARCC	Average routing community count	AT	Average throughout
CM	Content-related memory	CoM	Community
CR	Content router	CS	Content store
FIB	Forwarding information base	FM	Forwarding-related memory
IC	Information center	ICN	Information-centric networking
PIT	Pending interest table	SDN	Software-defined networking

replica. In [23], a dynamic Q-routing based interest forwarding scheme was proposed to discover routes towards the closest content replica without generating the additional signaling overhead. However, the purposes of [17–23] were to try their best to forward Interest packets according to the "best" outgoing interface(s). It was difficult to determine which outgoing interface(s) in FIB was (were) "best" when considering network scale, time scale and FIB size.

In addition, different from [13,15–23], SDN and community division are introduced in this paper. ICN topology is divided into CoMs, and each IC with the knowledge about CoM is further controlled in the centralized way, which can accelerate content retrieval.

2.2. SDN based ICN routing

There are also some researches on combining ICN and SDN to produce the newly promising routing schemes. In [36], a possible architecture was proposed to provide CCN functionalities over OpenFlow switches without changing OpenFlow, which enabled the name-based routing over OpenFlow switches. In [37], a routing mechanism was proposed to deploy ICN protocols in IP networks with the assistance of SDN, which decoupled control plane from data plane and enabled ICN routing mechanism by the centralized knowledge. In [38], an effective integration mechanism of SDN and ICN was proposed, which extended ICN from the following aspects, i.e., integration with IP, routing scalability, interdomain routing and transport mechanism. However, the studies of [36–38] relied on TCP/IP, and ICN was only used as an overlay at application level. In [25], a new SDN-based routing approach was proposed, which only relied on CCN architecture without TCP/IP and did not adopt the existent communication protocols of SDN (e.g., OpenFlow). However, [25] as well as [36–38] didn't consider community division by extracting user interests, from which our work was different.

2.3. Community based ICN routing

There are a few studies on community division in ICN. In fact, one community is a domain to some extent. In [19], a domain-based greedy ant colony forwarding algorithm was presented. It divided ICN topology into multiple domains, and each one was managed by the specific Internet Service Provider (ISP). In [26], a routing scheme was proposed with SDN controller and various servers considered. A domain could deploy the additional servers and spread them throughout the domain, so that the forwarding interface at an edge router pointed to the closest replica within a domain. In [27], a scalable domain-based routing scheme was proposed. Each domain had one unique identifier and one or more gateways. Packets from (or to) the domain should pass through one of its gateways, and the routing information based on the locator was exchanged among gateways. Different from [19,26,27], instead of the ISP in [19], the edge router in [26] or the locator in [27], we leverage one IC to store all information about contents and forwarding of a domain, and the information is decoupled from its domain. Furthermore, [19,26,27] didn't provide the concrete methods for dividing domain; in contrast, we adopt the clustering analysis method to divide ICN topology into domains, i.e., CoMs. Moreover, [19,26,27] didn't consider social relationship, instead, we adopt social relationship based on interest similarity to divide community for helping retrieve the content for Interest packets conveniently and effectively.

In summary, this paper not only adopts the basic ideas of SDN and community division but also leverages the unique IC of each CoM for packet forwarding and content retrieval instead of directly accessing Content Store (CS), Pending Interest Table (PIT) and FIB of each CR, which is different from the above mentioned related work.

3. System framework

The abbreviations frequently used in this paper are listed in Table 1. ICN usually has two kinds of packets, that is, Interest packets are used to request the content and Data packets are used to carry the corresponding content back to the interest requester. In this paper, we devise a novel routing mechanism called RISC, which is different from existent ICN routing schemes. It is assumed that the proposed RISC has Interest packets and Data packets, but no CS, FIB and PIT. As depicted in Fig. 1, the proposed RISC consists of control plane and data plane by adopting the basic ideas of SDN, which allows decoupling control plane (routing decision) from data plane (packet forwarding). The data plane has ICs and CoMs while the control plane has SDN controllers, where one SDN controller only manages its own CoM and IC, that is, An SDN $controller_i$ corresponds to one CoM_i and one IC_i. An SDN controller uses four kinds of messages: *lookup, forwarding, success* and *failure*. Among them, *lookup* is used to find the content, *forwarding* is used to forward packets, *success* is used when the content is found, and *failure* is used when the content cannot be found. In fact, SDN controller, CoM and IC are three new roles in RISC, and their functions are described as follows.

(i) SDN $controller_i$ is used to control packet forwarding. On one hand, it controls IC_i to provide the content for interest requests in CoM_i (intra-community routing); on the other hand, it controls CRs in CoM_i to forward Interest packets to CRs in another CoM until the content is found (if exist) (inter-community routing).

(ii) A CoM is one community and its members have similar or even same attributes and features (e.g., life habits and interests). In this paper, the content names stored at CRs are regarded as the so-called attributes and features. Then, ICN topology is divided into some sub-ICNs, and each one is considered as one CoM or domain (just as an autonomous system in OSPF) which consists of some CRs. Especially, two CoMs may be nonadjacent (e.g., CoM_3 and CoM_4 in Fig. 1) or adjacent (e.g., CoM_1 and CoM_3 in Fig. 1), and the definition about "adjacent" is as follows.

Definition 1 (adjacent)**.** $\forall$ CR_i, CR_j, $i \neq j$, if there exists an edge between CR_i and CR_j, they are adjacent. $\forall$ $CR_i \notin CoM_j$, if CR_i is adjacent to one CR of CoM_j, CR_i is adjacent to CoM_j.

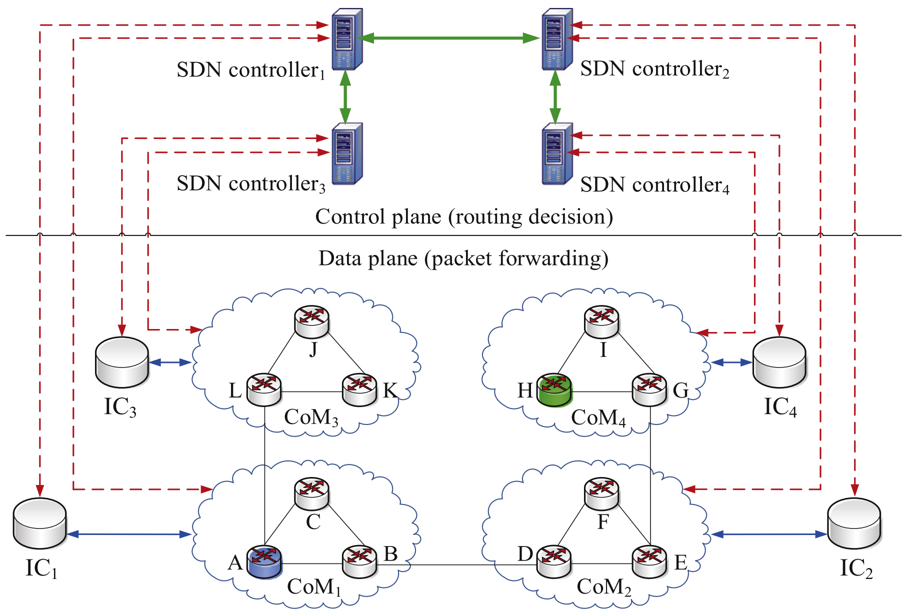


Fig. 1. Architecture of ICN incorporating SDN and community division.

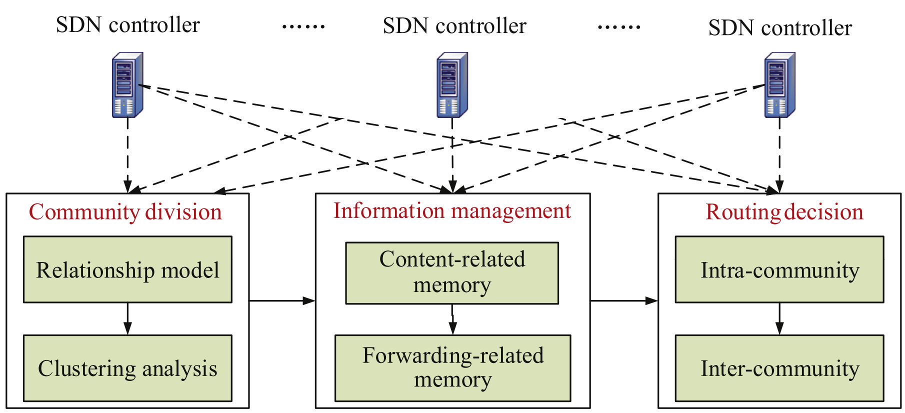


Fig. 2. System framework of RISC.

(iii) IC knows all information about contents and forwarding in its corresponding CoM, and it is used to store and manage the information which refers to (i) content names and (ii) contents stored at CRs of CoM, and (iii) adjacency conditions of CRs with other CoMs, which are used to compute social relationship values among CoMs. Then, the definition about “adjacency condition” is as follows.

Definition 2 (adjacency condition)**.** $\forall\ CR_i \notin CoM_j$, if CR_i is adjacent to CoM_j, CoM_j is the adjacency condition of CR_i; if CR_i is nonadjacent to any CoM, the adjacency condition of CR_i is Null. For example, in Fig. 1, B is adjacent to D of CoM_2, thus the adjacency condition of B is CoM_2.

According to the above statements, we present the system framework of RISC in Fig. 2. RISC consists of the following three modules: community division, information management and routing decision. Among them, the first one is responsible for dividing ICN topology into CoMs by the clustering analysis method; the second one is responsible for placing the information about CoM into its corresponding IC for the centralized management; the last one is responsible for retrieving the content by intra-community and inter-community routing. All operations of three modules are controlled by SDN controller.

4. Community division

We use the maximal tree scheme which is a clustering analysis method to divide ICN topology into some CoMs according to the content name prefixes. At first, we establish a social relationship model among CRs, and then perform the clustering analysis.

4.1. Relationship model based on content name prefixes

As we know, ICN pays attention to user interest similarity which play an important role in social relationship and can be extracted from Interest packets, Therefore, this paper divides community based on user interests, in other words, the interest similarity relation is used to express social relationship. In fact, each CR stores different types of contents (e.g., game, travel, sport, etc.), which reflect user interests and have different content name prefixes. Thus, these prefixes can be used to exploit user interests.

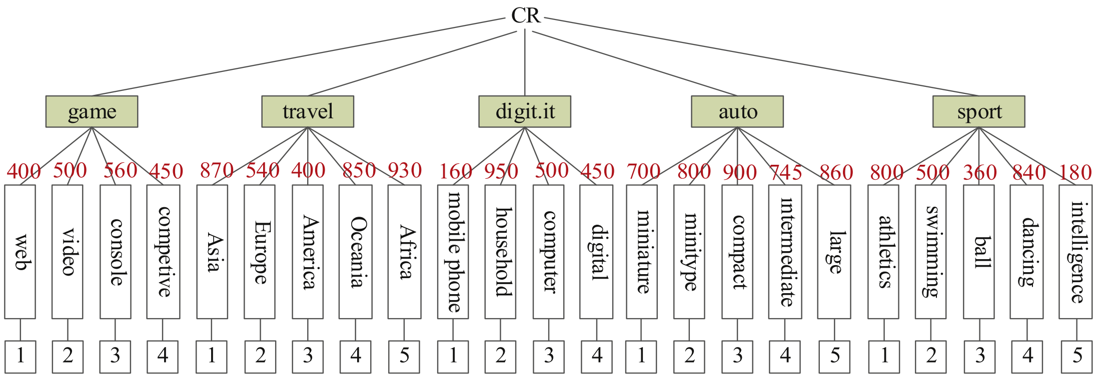


Fig. 3. CR description and sub-attribute value settings.

Definition 3 (interest attribute)**.** Assume that each CR consists of h types of contents, each one is regarded as an interest attribute.

Based on Definition 3, CR_i is expressed as follows:

$$CR_i \sim \left(a_{i,1}, a_{i,2}, \ldots, a_{i,h}\right), \tag{1}$$

where $a_{i,\,k}$ denotes a type of content (an attribute) of CR_i, here $1 \le k \le h$ and $1 \le i \le n$, and n is the number of CRs.

For example, in Fig. 3, CR is expressed by game, travel, digit.it, auto and sport. In other words, according to Eq. (1), $CR \sim$ (*game, travel, digit.it, auto, sport*). It is obvious that CRs with similar or same content name prefixes are likely to be in the same CoM.

However, the interest attribute (e.g., game) is somewhat abstract and hard to be depicted. In order to establish social relationship among CRs effectively, each attribute should be expressed by a definite value, that is, these attributes are quantified rather than just being generally qualified.

Furthermore, an attribute can be divided into some sub-attributes. Let q_k represent the number of sub-attributes in terms of $a_{i,\,k}$, denoted by $a_{i,k,1}, a_{i,k,2}, \ldots, a_{i,k,q_k}$, and these sub-attributes are set as different values from 1 to q_k, denoted by $a'_{i,k,1}, a'_{i,k,2}, \ldots, a'_{i,k,q_k}$, here $a'_{i,k,q_t} = q_t$, $1 \le t \le k$. For example, in Fig. 3, game is classified into web game, video game, console game and competitive game, which are set as 1, 2, 3 and 4, respectively. Suppose that a_{i,k,q_t} corresponds to β_{i,k,q_t} content items, then CR_i is defined as follows:

$$CR_i = \left(\overrightarrow{A_{i,k}}\right)_{1\times h}, \tag{2}$$

$$\overrightarrow{A_{i,k}} = \left(\beta_{i,k,q_t} \cdot a'_{i,k,q_t}\right)_{q_k\times 1} = \left(q_t \cdot \beta_{i,k,q_t}\right)_{q_k\times 1}, \tag{3}$$

where $\overrightarrow{A_{i,k}}$ is a set of q_k sub-attribute values. For example, in Fig. 3, $CR_i = \left(\overrightarrow{A_{i,k}}\right)_{1\times 5}$. Among them,

$\vec{A}_{i,1} = (400 \cdot 1, 500 \cdot 2, 560 \cdot 3, 450 \cdot 4)^T = (400, 1000, 1680, 1800)^T$,

$\vec{A}_{i,2} = (870 \cdot 1, 540 \cdot 2, 400 \cdot 3, 850 \cdot 4, 930 \cdot 5)^T = (870, 1080, 1200, 3400, 4650)^T$,

$\vec{A}_{i,3} = (160 \cdot 1, 950 \cdot 2, 500 \cdot 3, 450 \cdot 4)^T = (160, 1900, 1500, 1800)^T$,

$\vec{A}_{i,4} = (700 \cdot 1, 800 \cdot 2, 900 \cdot 3, 745 \cdot 4, 860 \cdot 5)^T = (700, 1600, 2700, 2980, 4300)^T$,

$\vec{A}_{i,5} = (800 \cdot 1, 500 \cdot 2, 360 \cdot 3, 840 \cdot 4, 180 \cdot 5)^T = (800, 1000, 1080, 3360, 900)^T$.

Let $r_{i,\,j}$ represent the interest similarity value between CR_i and CR_j. A lot of methods, such as dot product, correlation coefficient and absolute subtrahend, have been proposed to compute $r_{i,\,j}$. Among them, the dot product method relies on a variable, which is against the unique determination of $r_{i,\,j}$; the computation of correlation coefficient method is complex due to the square root operation. Thus, in this paper, when CR_i and CR_j are adjacent, we use the absolute subtrahend method to compute $r_{i,\,j}$ as follows:

$$r_{i,j} = 1 - c\sum_{k=1}^{h}\min_{t=1}^{k}\left|q_t \cdot \beta_{i,k,q_t} - q_t \cdot \beta_{j,k,q_t}\right| \tag{4}$$

$$c = 10^{-\psi}, \tag{5}$$

$$10^{\psi} = \sum_{k=1}^{h}\min_{t=1}^{k}\left|q_t \cdot \beta_{i,k,q_t} - q_t \cdot \beta_{j,k,q_t}\right| + \Omega, \tag{6}$$

where $\psi \in \mathbb{N}$, $\Omega \ge 0$, and Ω is set as small as possible. It is obvious that Eq. (4) is capable of quantifying interest similarity relation among CRs. If two CRs are not adjacent, their interest similarity relation does not exist and the corresponding value is 0. If $r_{i,j} = 1$, CR_i and CR_j have the strongest interest similarity relation. Assume that CR_1 is adjacent to CR_2 and CR_3, and that CR_2 and CR_3 are not adjacent, here $CR_1 = (1, 2, 1, 4, 9)$, $CR_2 = (1, 3, 3, 1, 7)$ and $CR_3 = (1, 2, 1, 4, 9)$. We have $r_{1,2} = 0.2$, $r_{1,3} = 1$ and $r_{2,3} = 0$ according to Eq. (4).

4.2. Clustering analysis based on maximal tree

The clustering analysis method can be used for dividing community, and it usually includes transitive closure method, direct cluster method, netting method and maximal tree method. In this paper, we adopt the maximal tree method to divide ICN topology into CoMs because the maximal tree is an undirected graph with weight values. In this way, it does not need the complex conversion process from ICN topology to an undirected graph with weight values. The maximal tree method is described as follows. Firstly, it connects the elements according to $r_{i,\,j}$ in descending order and obtains the maximal tree after marking weight values. Secondly, given a threshold $\lambda \in [0, 1]$, it deletes the edges of which weight values are smaller than λ along the maximal tree. Finally, the interconnected elements in the maximal tree are regarded as one category.

At first, we build a maximal tree according to the given ICN topology, denoted by G. In this paper, these weight values in the maximal tree are interest similarity values among CRs, and they

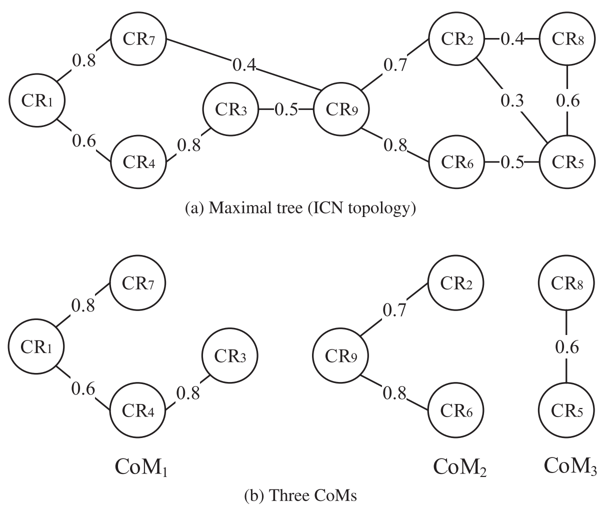


Fig. 4. A simple example for illustrating clustering analysis based on maximal tree with eight CRs and $\lambda = 0.6$.

can be recorded by a matrix, denoted by R, as follows:

$$R = \left(r_{i,j}\right)_{n\times n} = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n,1} & r_{n,2} & \cdots & r_{n,n} \end{bmatrix}. \quad (7)$$

According to the above statements, the maximal tree is built according to G and R (see Fig. 4(a)).

Then, we divide the maximal tree into CoMs according to the cut matrix, and the definition about "λ cut matrix" is as follows:

Definition 4 λ cut matrix. Let $R = (r_{i,j})_{n\times n}$, $\forall\lambda \in [0, 1]$,

$$R_\lambda = \left(r_{i,j}(\lambda)\right)_{n\times n}, \quad (8)$$

where

$$r_{i,j}(\lambda) = \begin{cases} 1, & r_{i,j} \geq \lambda \\ 0, & otherwise, \end{cases} \quad (9)$$

and R_λ is the λ cut matrix of R.

If $r_{i,j} = 1$, CR_i and CR_j are in the same CoM; however, $r_{i,j} = 0$ doesn't mean that CR_i and CR_j are not in the same CoM. For example, in Fig. 4, CR_4 and CR_7 are not adjacent ($r_{4,7} = 0$), and they are in the same CoM; CR_2 and CR_7 are not adjacent ($r_{2,7} = 0$), and they are not in the same CoM. In this way, we can divide ICN topology into CoMs. There are two extreme situations: (i) if λ is larger than the minimum value of R, ICN topology is divided into n CoMs (i.e., each CR is considered as one CoM); and (ii) if λ is smaller than or equal to the maximum value of R, ICN topology is divided into only one CoM (i.e., ICN topology is one CoM). As a matter of fact, ICN topology can be divided into different CoMs with different λ.

For the generated maximal tree, given any λ, if $r_{i,j}(\lambda) = 1$, CR_i and CR_j remain connected; if $r_{i,j}(\lambda) = 0$, the edge between CR_i and CR_j is deleted. For example, in Fig. 4(a), $r_{7,9}(0.6) = 0$, the edge between CR_7 and CR_9 is deleted in Fig. 4(b). Assume that there are m CoMs, denoted by $CoM_1, CoM_2, \ldots, CoM_m$. Let CoM_i be composed of θ_i CRs, denoted by $CR_{i,1}, CR_{i,2}, \ldots, CR_{i,\theta_i}$, and we have

$$\theta_1 + \theta_2 + \cdots + \theta_m = \sum_{i=1}^{m} \theta_i = n. \quad (10)$$

According to the above statements, an algorithm is devised to divide ICN topology into CoMs, and it is described in Algorithm 1.

Algorithm 1 Community division based on maximal tree.

```
Input: G, λ
Output: m CoMs
for i = 1 to n, do
   Obtain CR_i according to Eq. (2);
end for
for i, j = 1 to n, do
   Compute r_{i,j} according to Eq. (4);
end for
Obtain R according to Eq. (7);
Compute R_λ according to Eq. (8);
for i, j = 1 to n, do
   if r_{i,j}(λ) = 0
      Delete the edge between CR_i and CR_j;
   end if
end for
return m CoMs
```

5. Information management

In the classical ICN, the information stored at CRs is composed of contents, forwarding interfaces and pending interests, which are stored at CS, FIB and PIT respectively. However, RISC uses an unique IC to store all information about contents and forwarding in CoM for the centralized management. As shown in Fig. 5, IC consists of Content-related Memory (CM) and Forwarding-related Memory (FM). Among them, CM has the similar functions to those of CS; FM has the similar functions to those of FIB and PIT. If an Interest packet cannot retrieve the content from the corresponding CM, not only the Interest packet is forwarded to another CoM but also its forwarding information is placed into the corresponding FM for conducting the subsequent forwarding.

5.1. Content-related memory

CM is responsible for storing the information about contents in its CoM. When an Interest packet arrives, CR checks its CM instead of CS. As shown in Table 2, the structure of CM is composed of the following six *fields*: CR number, content number, content name prefix (ContPref), content name, content and adjacency condition (AdjaCond). Among them, ContPref reflects the essential attribute of CoM (for example, in Table 2, CRs of CoM prefer to forward these Interest packets about sports and travel); AdjaCond reflects the adjacency condition between CR and another CoM, and it can be a CoM or Null.

The information in CoM is placed into its CM by using ContPref index rather than CR number, which can speed up content lookup. For example, in Table 2, the information about sports is stored at CM firstly, and then the information about travel is stored at CM. If content name of an interest request is /travel.sohu.com/around.shtml, because its prefix is /travel.sohu.com, $Cont_5$ is found at first rather than looking up all information from CR_1.

Each CR of CoM submits the information of CS to CM for the centralized management, thus reducing the burden of CR brought by the frequent content match in CS and improving content lookup speed. For example, in Table 2, assume that (i) a CoM is composed of CR_1, CR_2 and CR_3; and (ii) content name of an interest request is /travel.sohu.com/around.shtml ($Cont_7$), which can be found from CR_3 rather than CR_1 or CR_2. The number of lookup is 7 with the

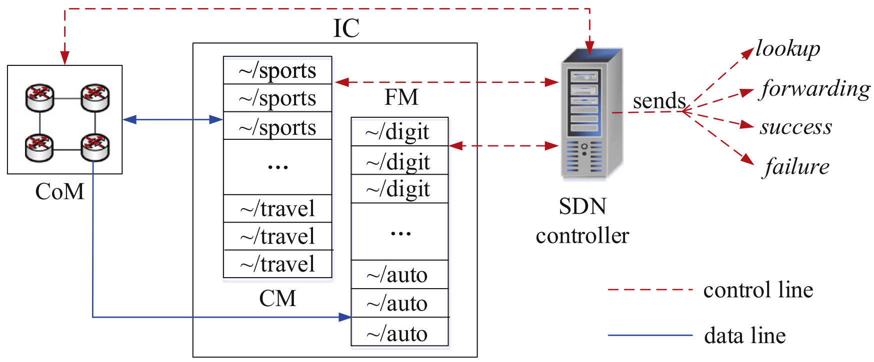


Fig. 5. Centralized management of information from CoM to IC.

Table 2
CM structure used to store the content-related information.

CR number	Content number	ContPref	Content name	Content	AdjaCond
CR_1	$Cont_1$	/sports.sohu.com	~ /basketball.shtml	video1.mp4	CoM_3
CR_1	$Cont_2$	/sports.sohu.com	~ /football.shtml	webpage1	CoM_3
CR_2	$Cont_3$	/sports.sohu.com	~ /tennis.shtml	webpage2	CoM_2
CR_2	$Cont_4$	/sports.sohu.com	~ /pingpong.shtml	video.mp4	CoM_2
CR_1	$Cont_5$	/travel.sohu.com	~ /china.shtml	webpage3	CoM_3
CR_2	$Cont_6$	/travel.sohu.com	~ /world.shtml	video3.mp4	CoM_2
CR_3	$Cont_7$	/travel.sohu.com	~ /around.shtml	webpage4	Null
...	...	...	...	...	...

Table 3
FM structure used to store the forwarding-related information.

Item number	ContPref	Content name	OutgInte
$Item_1$	/digit.it.sohu.com	~ /mobile.shtml	Null
$Item_2$	/digit.it.sohu.com	~ /home.shtml	CoM_2
$Item_3$	/auto.sohu.com	~ /luxury.shtml	CoM_3
$Item_4$	/auto.sohu.com	~ /mpv.shtml	CoM_4
...	...	...	...

classical ICN because the lookup order is $Cont_1$, $Cont_2$, $Cont_5$, $Cont_3$, $Cont_4$, $Cont_6$ and $Cont_7$; however, it is 4 with RISC because $Cont_1 - Cont_4$ only needs one lookup, followed by $Cont_5$, $Cont_6$ and $Cont_7$.

5.2. Forwarding-related memory

If there is no matched content with the interest request in CM, the Interest packet is forwarded to another CoM. The related information used to conduct its forwarding is stored at FM. As shown in Table 3, the structure of FM is composed of the following four *fields*: item number, ContPref, content name and outgoing interface (OutgInte). Among them, OutgInte is the next hop CoM to which the Interest packet will be forwarded; ContPref is used to match with the interest request which cannot be matched at CM, and specify the corresponding forwarding community of Interest packet. Similar to CM, the information is stored at FM according to ContPref index.

Similar to AdjaCond of CM, OutgInte can be a CoM or Null. If OutgInte is a CoM, it means that the Interest packet has been forwarded to the CoM and the content has also been retrieved; in this case, FM serves as the role of FIB. If OutgInte is Null, it means that the Interest packet has been forwarded to the CoM while the content has not been retrieved, that is, the Interest packet is waiting for the content. If the content can be retrieved by waiting from the CoM, Null is changed into the CoM; otherwise, the Interest packet is discarded; in this case, FM serves as the role of PIT.

Examples for illustrating the information storage from CoM to FM are presented as follows. If (i) content name of an interest request is /digit.it.sohu.com/mobile.shtml, (ii) it has been forwarded to CoM_1, and (iii) the content has not been retrieved, the corresponding OutgInte is Null and then the related information is stored at FM, as shown in $Item_1$ of Table 3. If (i) content name of an interest request is /digit.it.sohu.com/home.shtml and (ii) the content has been retrieved from $CoM(2)$, the related information is stored at FM, as shown in $Item_2$ of Table 3.

6. Routing decision

6.1. Intra-community routing

When one CR (CR_x) of CoM_i receives an Interest packet (ix), it first submits the name of ix to $controller_i$; and then $controller_i$ sends *lookup* message to seek CM_i by ContPref.

If the content can be found, CM_i sends *success* message to $controller_i$ at the same time it provides the content to CR_x. In this process, it does not need $controller_i$ to conduct CM_i to provide the content after receiving *success* message, thus saves the time of sending message. Otherwise, CM_i sends *failure* message to $controller_i$ and then *lookup* message is sent to FM_i. Under such condition, ContPref is checked whether it can be matched.

If the matched ContPref cannot be found, FM_i sends *failure* message to $controller_i$ and then *forwarding* message is sent to CoM_i to forward ix to another CoM (see Section 6.2). Otherwise, the content name is checked whether it can be matched.

If the matched content name cannot be found, ix will be forwarded to another CoM (see Section 6.2). Otherwise, it means that CoM_i has forwarded ix while it is uncertain whether the content has been retrieved, thus the following is to check whether OutgInte is Null.

If the matched OutgInte is CoM_j, it means that the previous interest requests have retrieved the content while ix has not retrieved the content, thus forwarding ix to CoM_j. If the matched OutgInte is Null, it means that the content has not been retrieved, thus waiting for the content.

The process of intra-community routing is shown in Fig. 6, where "$\surd$" means that the matched item can be found and "$\times$"

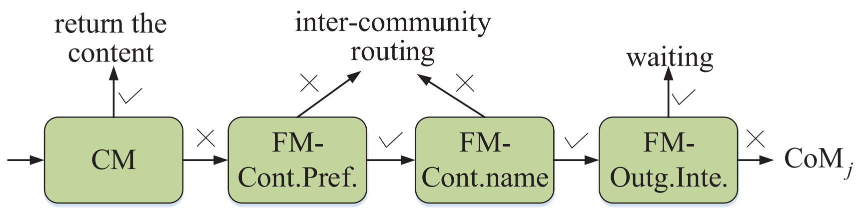


Fig. 6. The process of intra-community routing.

means no. The intra-community routing based on same CoM information is described in Algorithm 2.

Algorithm 2 Intra-community routing algorithm.

Input: CoM_i, CR_x, ix
Output: The content or failure
if the content is found from CM_i
  return the content;
else if (i) no matched ContPref or content name found from FM_i or (ii) the matched OutgInte is CoM_j
  return failure;
else
  Wait for the content;
  return the content;
end if

6.2. Inter-community routing

According to Algorithm 2, in case of no matched content in CM_i, if (i) FM_i has no matched ContPref, or (ii) FM_i has matched ContPref but no matched content name, ix is forwarded to another CoM to retrieve the content, which will be done by inter-community routing and controlled by SDN $controller_i$.

The goal of inter-community routing is to select a proper CoM as the forwarding community. The strong interest similarity relation among CoMs means the large probability to forward same or similar Interest packets. Thus, we use social relationship based on interest similarity to determine which CoM can be as the forwarding community.

The interest similarity relation model between two CRs has been shown in Eq. (4). In this section, we establish that between two CoMs. Suppose that $mr_{i,\,j}$ represents the interest similarity value between CoM_i and CoM_j, and we have $R(CoM_i, CoM_j) = mr_{i,j}$. Let $r_{i,\,o;\,j,\,l}$ represent the interest similarity value between $CR_{i,\,o}$ (one CR of CoM_i) and $CR_{j,\,l}$ (one CR of CoM_j), here $1 \leq o \leq \theta_i$ and $1 \leq l \leq \theta_j$, and $mr_{i,\,j}$ is defined as follows:

$$mr_{i,j} = max\left\{ r_{i,o;j,l} \middle| 1 \leq o \leq \theta_i;\, 1 \leq l \leq \theta_j;\, o, l \in \mathbb{N} \right\}, \tag{11}$$

which indicates that the interest similarity value between CoMs depends on all interest similarity values between CRs in two different CoMs, thus it is different from that between CRs. For example, in Fig. 4, $mr_{1,2} = max\{r_{3,9}, r_{7,9}\} = 0.5$. Then, in terms of the two inter-community routing situations in Fig. 6, we design the forwarding solutions in Sections 6.2.1 and 6.2.2, respectively.

6.2.1. No match with ContPref

If no matched ContPref found, it means that the Interest packet which is similar to ix (i) was not forwarded, or (ii) was forwarded while the related forwarding information had been deleted.

In such situation, suppose that CoM_i is adjacent to p CoMs, $1 \leq p \leq m-1$, we select one CoM which has the highest interest similarity value with CoM_i as the forwarding community. In other words, CoM_k is the forwarding community, $1 \leq k \leq p$, if and only if

$$mr_{i,k} = max\left\{ mr_{i,1}, mr_{i,2}, \ldots, mr_{i,p} \right\}. \tag{12}$$

For example, in Fig. 1, suppose that ix is in CoM_1, and it will be forwarded to CoM_2 or CoM_3, here $p = 2$. Suppose that $mr_{1,2} = 0.5$ and $mr_{1,3} = 0.4$, then CoM_2 is selected as the forwarding community.

6.2.2. No match with content name

If the matched ContPref found while no matched content name found, it means that the Interest packet which is similar to ix was forwarded but (i) the Interest packet which is the same as ix was not forwarded, or (ii) was forwarded while the related forwarding information had been deleted.

In such situation, ix will be forwarded to one CoM which has received the similar Interest packet, where ContPref is the same as that of the Interest packet. Suppose that ω CoMs have received the forwarded Interest packet from CoM_i, $\omega \leq p$, similar to Section 6.2.1, if and only if

$$mr_{i,k} = max\left\{ mr_{i,1}, mr_{i,2}, \ldots, mr_{i,\omega} \right\}, \tag{13}$$

CoM_k is the forwarding community, $1 \leq k \leq \omega$.

For example, in Table 3, suppose that content name of ix in CoM_1 is /auto.sohu.com/suv.shtml and it cannot be found from FM_i, here $\omega = 2$. CoM_3 and CoM_4 can be regarded as the candidate forwarding communities rather than CoM_2, because ContPrefs of $Item_3$ and $Item_4$ are /auto.sohu.com while ContPref of $Item_2$ is /digit.it.sohu.com. Suppose that $mr_{1,3} = 0.3$ and $mr_{1,4} = 0.4$, then CoM_4 is selected as the forwarding community.

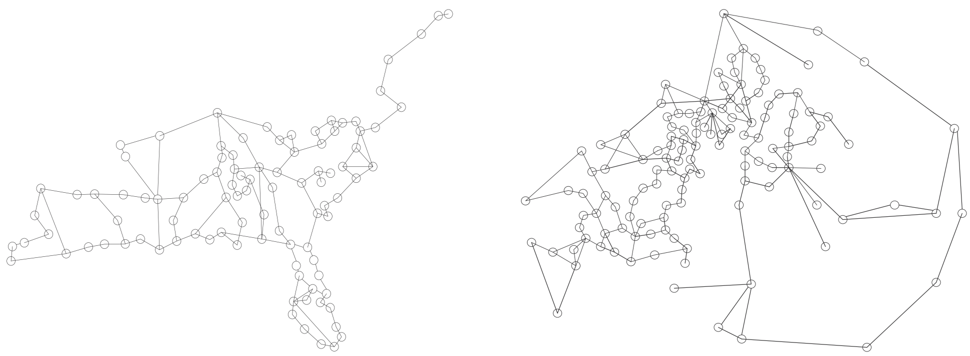
The inter-community routing based on social relationship among CoMs is described in Algorithm 3.

Algorithm 3 Inter-community routing algorithm.

Input: A divided ICN topology, CoM_i, ix
Output: A forwarding community
if no matched ContPref found from FM_i
  $\tau = p$;
else
  $\tau = \omega$;
end if
for $k = 1$ to τ, **do**
  Compute $mr_{i,k}$ according to Eq. (11);
end for
Select the maximum $mr_{i,k}$ from $mr_{i,1}, mr_{i,2}, \ldots, mr_{i,\tau}$;
Forward ix to CoM_k;
return CoM_k;

According to the above statements, we present a theorem about the time complexity of RISC as follows.

Theorem. *The time complexity of RISC is $O(e + \xi \cdot n)$, where* e *is the number of edges and ξ is constant.*



(a) USA-Deltacom with 97 nodes and 124 edges (b) Europe-GTS with 130 nodes and 168 edges

Fig. 7. Two practical topologies used for performance evaluation.

Proof. RISC is composed of three major modules, i.e., community division, information management and routing decision. Among them, the first one depends on extracting the content name prefixes, computing $r_{i,\,j}$ and building the maximal tree; the second one depends on storing the information about contents from CS to CM and that about forwarding from PIT as well as FIB to FM; the last one depends on intra-community routing based on CM lookup and inter-community routing based on FM lookup and computation of $mr_{i,\,j}$.

(i) Community division.

(i.1) Suppose that CS_i has c_i content items, and the extraction times of contents is denoted by E_t as follows:

$$E_t = c_1 + c_2 + \cdots + c_n = \sum_{i=1}^{n} c_i. \tag{14}$$

(i.2) Suppose that G has e edges, and the computation times of $r_{i,\,j}$ is denoted by C_t as follows:

$$C_t = e. \tag{15}$$

(i.3) The maximal tree is obtained by λ cut matrix, that is, if $r_{i,\,j} < \lambda$, the edge between CR_i and CR_j is deleted. In fact, all edges need to be traversed. Let T_t denoted the traversal times, and we have

$$T_t = e. \tag{16}$$

(ii) Information management.

It needs to place all information about contents and forwarding in CoM into its corresponding IC, which means to store them from CRs to ICs. Suppose that PIT_i and FIB_i have p_i and f_i items respectively, and the information storage times is denoted by S_t as follows:

$$S_t = \sum_{i=1}^{n} c_i + \sum_{i=1}^{n} p_i + \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} \left(c_i + p_i + f_i \right). \tag{17}$$

Since the number of information is linear, and we have

$$S_t = \sum_{i=1}^{n} \left(c_i + p_i + f_i \right) \approx \xi \cdot n. \tag{18}$$

(iii) Routing decision.

(iii.1) In the worst case, an interest request is sent from CoM_i to the last CoM, which traverses m CoMs. During this process, all content items in m CMs and all forwarding information items in m FM are looked up. Let L_t denote the lookup times, and we have

$$L_t = S_t. \tag{19}$$

(iii.2) In terms of the computation of $mr_{i,\,j}$, CoM_i is adjacent to p or ω CoMs (see Section 6.2). In the worst case, let $C_t^{'}$ denote the computation times of $mr_{i,\,j}$, and we have

$$C_t^{'} = m - 1. \tag{20}$$

Especially when G is divided into n CoMs (see Section 4.2), and we have

$$C_t^{'} = n - 1. \tag{21}$$

Since three major modules are performed in sequential order, the time complexity of RISC depends on Eqs. (14–16), (18), (19) and (21), that is

$$E_t + C_t + T_t + S_t + L_t + C_t^{'} \approx 2e + 3\xi \cdot n + n - 1 \sim O(e + \xi \cdot n). \tag{22}$$

This completes the proof. □

7. Performance evaluation

7.1. Simulation setup

We have implemented the proposed RISC by C++ programming language, and the test environment has been set up on a personal computer with the Intel Q8400, 2.66 GHZ CPU, 4G RAM over Windows 7 system. We compare RISC with two well-known schemes, i.e., classical NDN forwarding mechanism (denoted by NDNF [17]) and dynamic interest forwarding mechanism (denoted by INFORM [23]). NDNF provides a forwarding strategy to determine which interface is used to forward an interest request by interface ranking, making forwarding decision adapt to the network conditions. INFORM uses Q-routing approach to discover paths to content replicas and forward interest requests by exploration and exploitation, where the exploration phase is to compute Q values for different interfaces and the exploitation phase is to exploit the information. Meanwhile, ADSR, distribution of success times over community count, ARCC, AD, AT and stability are considered as six metrics of performance evaluation.

In this paper, to investigate the performance of RISC, we use two practical network topologies from the Internet topology zoo, i.e., Deltacom (2010-USA) and GTS (2010-Europe), as shown in Fig. 7. Furthermore, we capture data from Sohu website every day for one hour during one week. For the captured data, we extract the names of interest requests from the HTTP requests to divide Deltacom and GTS. Moreover, we generate 600, 800, 1000, 1200, 1400 and 1600 interest requests respectively over two topologies with $\lambda = 0.45$, and the simulation settings are shown in Table 4.

7.2. Average delivery success rate

As illustrated in Fig. 8, ADSRs over Deltacom and GTS with respect to RISC, INFORM and NDNF are reported.

In Fig. 8, over the same topology, we observe that RISC has the highest ADSR, followed by INFORM and NDNF. RISC uses one IC to

Table 4
Simulation settings.

Network topologies	Deltacom (USA) [39]	GTS (Europe) [40]
Number of CRs	97	130
Number of edges	124	168
Size of packet (KB)	10	10
Size of data (GB)	8	8
Total hours	7	7
Number of interest requests	600/800/1000/1200/1400/1600	600/800/1000/1200/1400/1600
λ	0.45	0.45
Simulation times	60	60

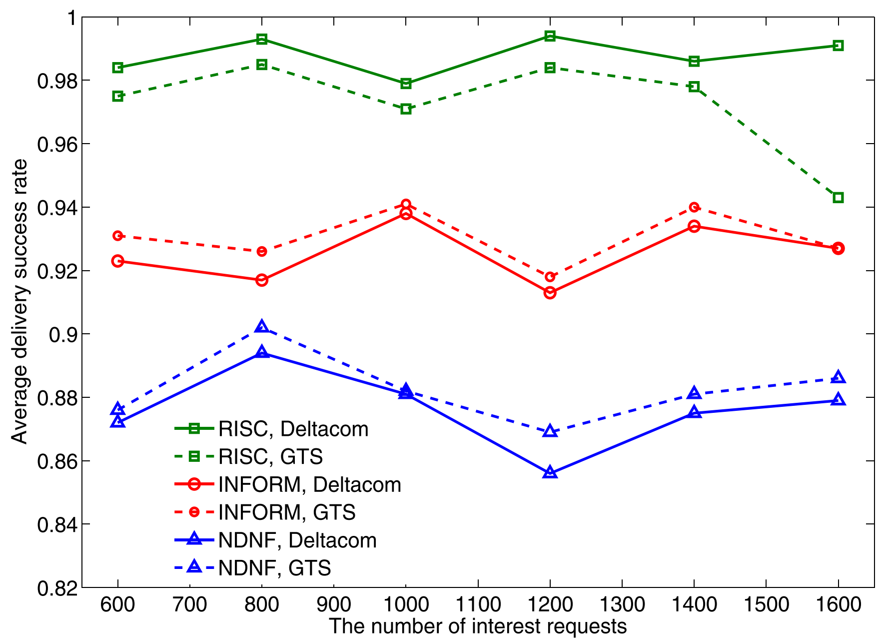


Fig. 8. Average delivery success rate among RISC, INFORM and NDNF.

store the contents of one CoM, and the content replacement frequency is considerably low. As a result, the hitting probability of content is considerably high. Although INFORM does not introduce the centralized management idea, it uses Q-routing based exploration to discover the content replicas, which can be exploited for the same subsequent interest requests. For NDNF, each CR has its CS to store the contents but no exploration and exploitation strategies to discover the content replicas, which causes the considerably high content replacement frequency, thus many interest requests are hard to find the contents.

Furthermore, over the different topologies, we observe that RISC has higher ADSR over Deltacom than that over GTS. In fact, the edges of GTS are more intensive than those of Deltacom, thus CoM is more easily to forward interest requests over GTS than that over Deltacom. In this way, the increasing speed of contents stored at IC over GTS is faster than that over Deltacom, thus the content replacement frequency over GTS is higher than that over Deltacom. We observe that INFORM has lower ADSR over Deltacom than that over GTS because there are more content replicas that respond to interest requests over GTS than over Deltacom. We observe that NDNF has lower ADSR over Deltacom than that over GTS, because GTS has a larger number of edges and thus provides better support for multipath forwarding of interest requests than Deltacom.

7.3. Distribution of success times over community count

As illustrated in Fig. 9, the distribution of success times over routing community count for Deltacom and GTS with respect to RISC, INFORM and NDNF on condition of 1000 interest requests is reported. Here, zero community means that an interest request can retrieve the content within the same CoM; five and five plus communities mean that an interest request cannot retrieve the content within four communities.

In Fig. 9, for RISC, we observe that the average success times of interest requests at zero community is the highest and it decreases with the increasing of routing community count, in other words, the operation of community division can speed up content retrieval. Furthermore, RISC has larger average success times at zero community over GTS than that over Deltacom because some CRs of one CoM over GTS have much more interest similarity relation with the other CRs due to the larger number of edges than that over Deltacom.

We observe that the majority of successful interest requests of INFORM are at three and two communities, while those of NDNF are at three and four communities. Furthermore, in terms of the number of interest requests that can retrieve contents at zero community, RISC is the largest, followed by INFORM and NDNF. In terms of that at three communities, NDNF is the largest, followed by INFORM and RISC. The above analysis suggest that RISC has better performance than INFORM and NDNF.

7.4. Average routing community count

ARCCs over Deltacom and GTS with respect to RISC, INFORM and NDNF are reported in Fig. 10.

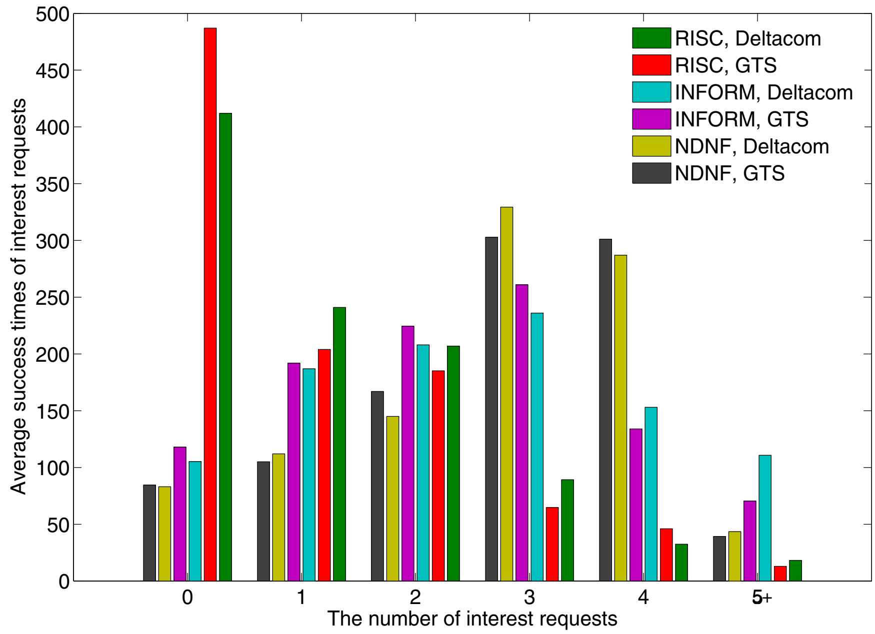


Fig. 9. Distribution of success times over community count among RISC, INFORM and NDNF under 1000 interest requests.

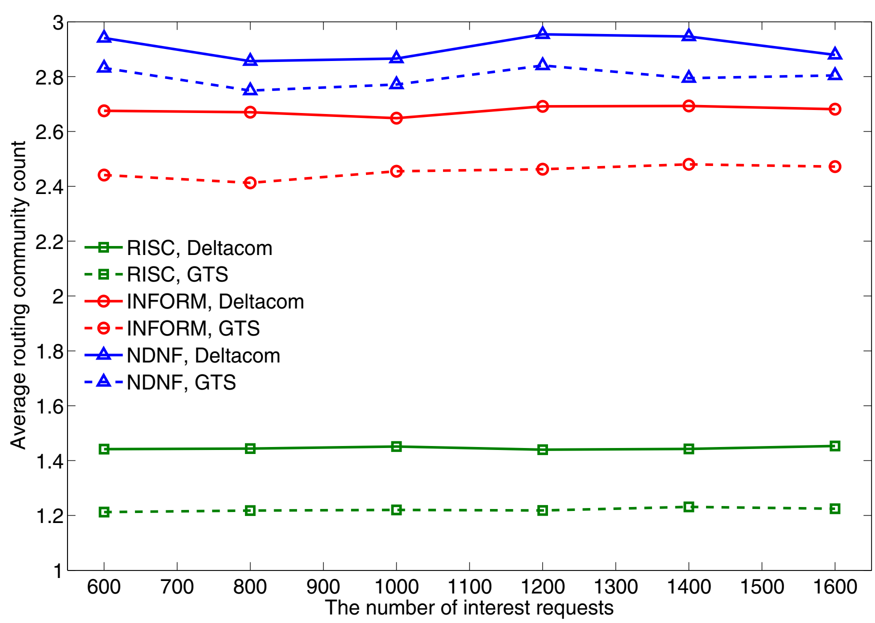


Fig. 10. Average routing community count among RISC, INFORM and NDNF.

In Fig. 10, over the same topology, we observe that RISC has the smallest ARCC, followed by INFORM and NDNF. The reasons are as follows. RISC has the largest average success times at zero, one and two communities at the same time has the smallest average success times at three, four, five and five plus communities. In contrast, NDNF has the smallest average success times at zero, one and two communities at the same time has the largest average success times at three, four, five and five plus communities. Furthermore, over the different topologies, we observe that RISC has larger ARCC over Deltacom than that over GTS. As we know, the edges over Deltacom are sparser than those over GTS, thus the number of CoMs over Deltacom is smaller than that over GTS. We also observe that INFORM and NDNF have larger ARCC over Deltacom than that over GTS, the reasons can be found from Section 7.2.

7.5. Average delay

The delay is defined as the difference between the timepoint when an interest request is generated and that when the corresponding content is returned or failed.

7.5.1. Delay factors

In RISC, the delay mainly consists of community division (division), information storage (storing), IC lookup (lookup), selecting CoM for inter-community routing (selecting), processing the messages of SDN controllers (processing) and forwarding Interest packets (forwarding). As shown in Table 5, the factors are divided into four categories, i.e., preprocessing, intra-community routing, inter-community routing and SDN controller operation. As illustrated in Fig. 11, ADs over Deltacom and GTS with respect to the

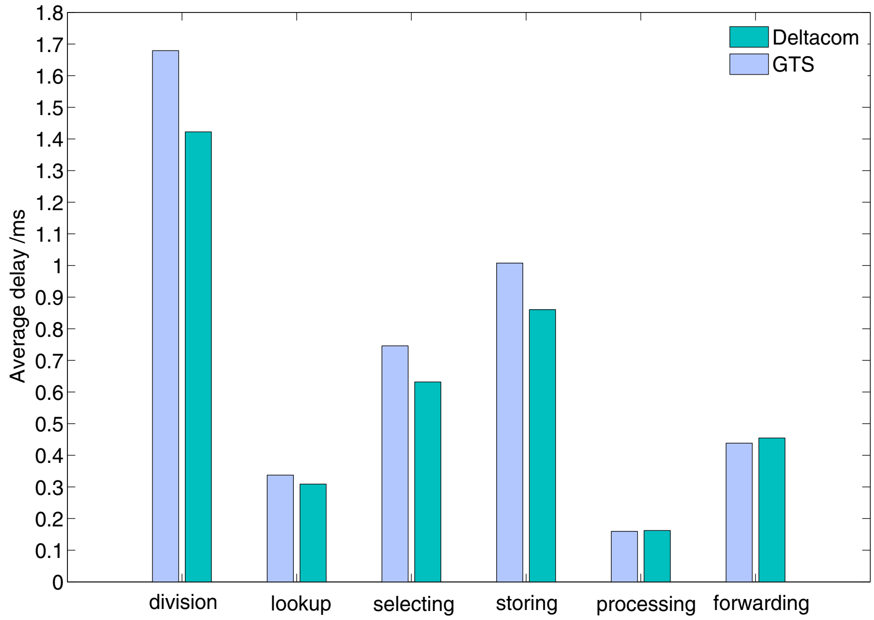


Fig. 11. Multiple delay factors for RISC under 1000 interest requests.

Table 5
Classification on routing delay factors.

Category	Factor
Preprocessing	Division & storing
Intra-community routing	Lookup
Inter-community routing	Selecting & forwarding
SDN controller operation	Processing

six main factors of RISC on condition of 1000 interest requests are reported.

In Fig. 11, we observe that the division takes the largest portion of AD, followed by storing, selecting, forwarding, lookup and processing. For division and selecting, the interest similarity values between two CRs or CoMs need to be computed. When the community division is accomplished, the delay mainly depends on storing the information to IC. Thus, the factors of division, selecting and storing take larger time than the others.

Furthermore, we observe that ADs of division and selecting over GTS are larger than those over Deltacom because GTS has a larger number of CRs and edges and thus needs much more time to compute interest similarity values than Deltacom. We observe that ADs of lookup and storing over GTS are larger than those over Deltacom because GTS has much more contents to be stored to IC and thus needs much more time to look them up than Deltacom. However, ADs of processing and forwarding have no significant change between Deltacom and GTS, which depends on the content types of interest requests.

7.5.2. Delay

ADs over Deltacom and GTS with respect to RISC, INFORM and NDNF are reported in Fig. 12.

In Fig. 12, over the same topology, we observe that RISC takes the smallest average time to accomplish interest requests, followed by INFORM and NDNF. On one hand, RISC has the smallest ARCC and thus has the smallest AD to retrieve the content. In similar way, NDNF has the largest AD to retrieve the content. On the other hand, lookup delay of CM and FM in RISC is smaller than that of CS, PIT and FIB in INFORM and NDNF. INFORM can more easily retrieve the content due to the adoption of exploration and exploitation compared with NDNF and thus gets the smaller AD. Furthermore, the AD increases with the increasing of interest requests. Meanwhile, the AD of RISC shows no significant change; however, ADs of INFORM and NDNF show significant changes. For RISC, the reasons can be found from Section 7.5.1. For INFORM and NDNF, when the number of interest requests increases, the network load becomes heavy and thus much processing time is needed.

Furthermore, over the different topologies, we observe that RISC and NDNF have larger AD to accomplish interest requests over GTS than those over Deltacom. For RISC, the reasons can be found from Section 7.5.1 (see Fig. 11). For NDNF, the edges over GTS are more intensive than those of Deltacom, thus the interest requests are more difficult to retrieve the content over GTS than those over Deltacom. However, INFORM has smaller AD to accomplish interest requests over GTS than that over Deltacom because GTS can provide more content replicas than Deltacom.

7.6. Throughput

The throughput is defined as the processed number of interest requests per unit time (μs). ATs over Deltacom and GTS with respect to RISC, INFORM and NDNF are reported in Fig. 13.

In Fig. 13, over the same topology, we observe that RISC has the largest AT, followed by INFORM and NDNF. RISC uses IC to centrally manage the information of one CoM, which increases the transmission number of interest requests per unit time. Thus, it has larger AT than INFORM and NDNF. In addition, INFORM uses exploration and exploitation to accelerate content retrieval, which makes more interest requests have chance to be sent into the network, thus it has larger AT than NDNF. We also observe that ATs of RISC, INFORM and NDNF decrease with the increasing of interest requests, because the processing capability of network declines due to the congestion. Furthermore, over the different topologies, we observe that RISC and NDNF have smaller AT over GTS than that over Deltacom, while INFORM is just the reverse. The related reasons are similar to the above sections.

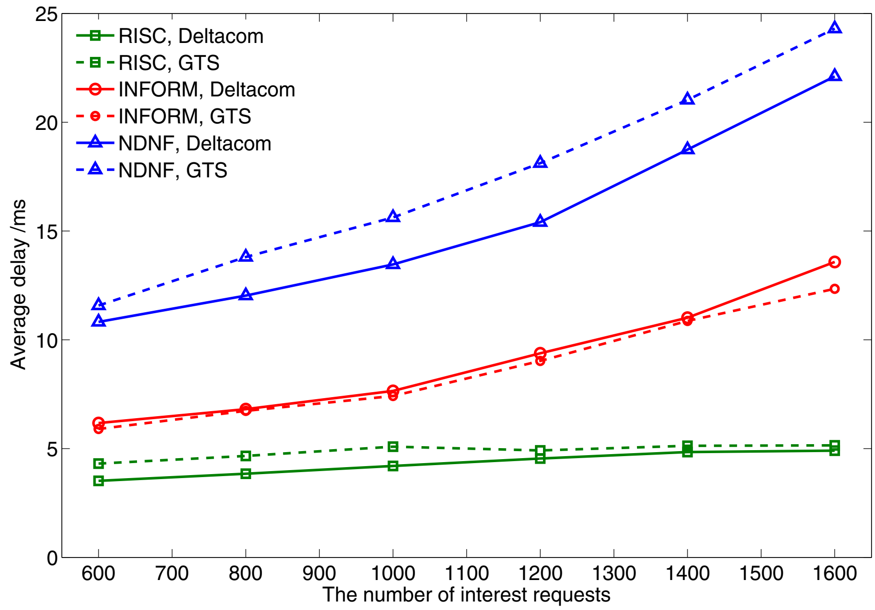


Fig. 12. Average delay among RISC, INFORM and NDNF.

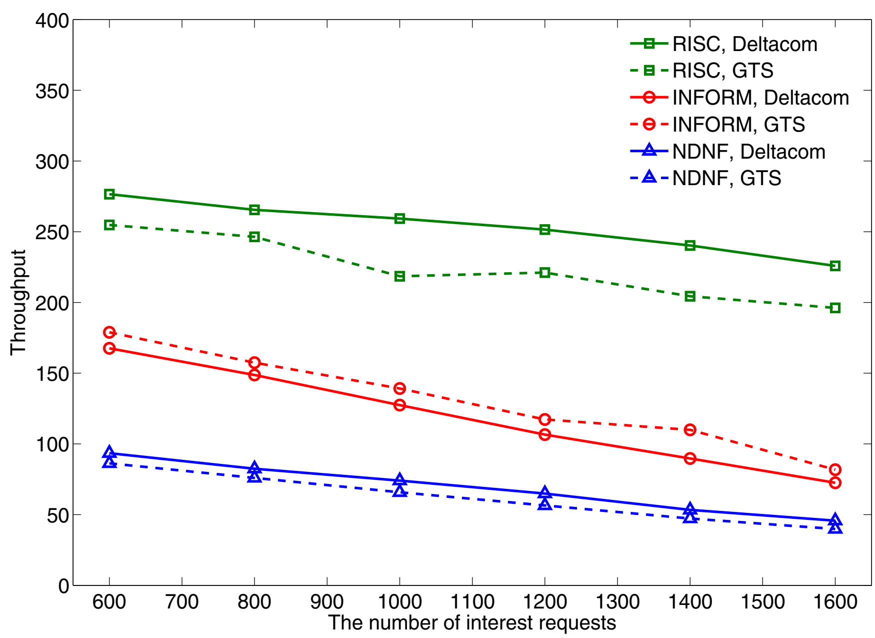


Fig. 13. Average throughput among RISC, INFORM and NDNF.

7.7. Stability

The stability is a metric which reflects the effectiveness of a routing algorithm. In this paper, it is quantified as the fluctuation coefficient [41] among RISC, INFORM and NDNF with respect to ADSR, ARCC, AD and AT. Furthermore, the small fluctuation coefficient means that the algorithm has good performance. As illustrated in Fig. 14, the stability over Deltacom and GTS with respect to RISC, INFORM and NDNF is reported.

In Fig. 14, over the same topology, in terms of the delivery success rate, community count, delay and throughput, we observe that RISC has the strongest stability, followed by INFORM and NDNF. In summary, RISC is more stable than INFORM and NDNF. The reasons are as follows. RISC does community division based on interest similarity relation, and thus provides convenience for the subsequent interest requests with strong similarity. Furthermore, RISC adopts the centralized management manner to place the information into IC, and thus supports the effective and efficient lookup for interest requests. Moreover, in RISC, control plane and forwarding plane are separated, thus forwarding plane does not become heavy loaded with the increasing of interest requests.

7.8. Discussion

Although RISC has good performance on ADSR, ARCC, AD, AT and stability, it encounters a negligible virus propagation issue among CoMs, that is, the community division based on social relationship accelerates content retrieval but may help virus propagation [42]. Therefore, virus defense has become increasingly complex and significant with the increasing of virus types, and the immune strategies have been regarded as a major part of de-

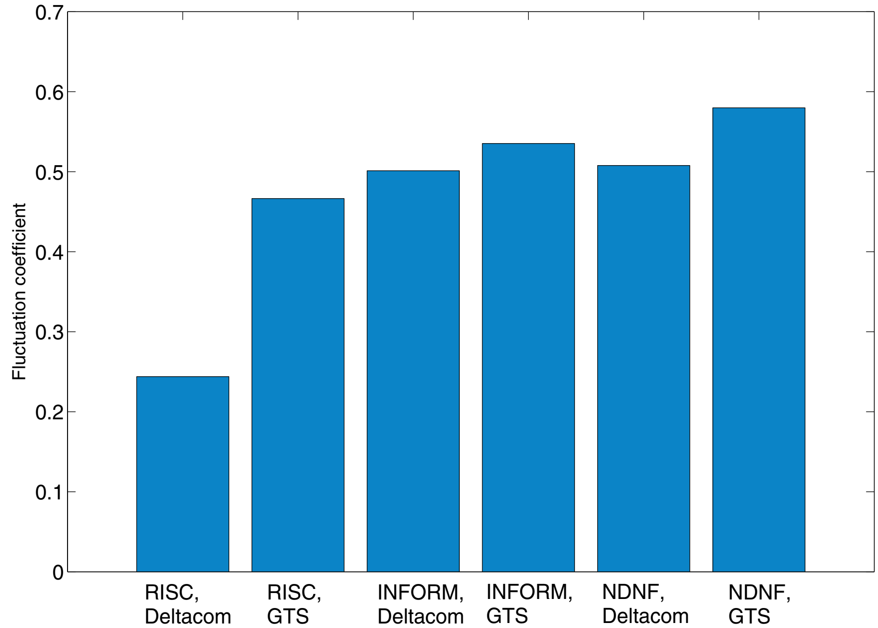


Fig. 14. Stability among RISC, INFORM and NDNF.

fense methods. As we know, some traditional immune strategies (e.g., random immune, targeted immune and acquaintance immune) have been employed to constrain virus propagation.

In particular, on one hand, ICN is to access the named content object rather than its IP address, which increases the amount of information in the network. On the other hand, ICN supports the mobility inherently, in which mobile devices join or leave the network easily. The two aspects adequately indicate that the ICN network changes dynamically. However, these traditional immune strategies cannot effectively adapt to dynamical networks due to the change of network structure over time. Given this, there are some schemes for both accelerating content retrieval and constraining virus propagation in the dynamic network environment. For example, [42] has proposed the evolutionary virus immune strategy based on community vitality to address virus defense. Similarly, when RISC encounters virus propagation issue, we can introduce the method of [42] as an efficient solution.

8. Conclusions

ICN is a clear-state networking paradigm which has attracted the increasing research interests. In this paper, we propose a novel ICN routing mechanism which is consisted of intra-community routing and inter-community routing by introducing the ideas of SDN and community division. We divide ICN topology into CoMs according to the content name prefixes, and the information in each CoM is stored at IC for the centralized management by SDN controller. The performance of RISC is evaluated over two practical topologies by considering success rate, success times, community count, delay, throughput and stability, and the simulation results have demonstrated that RISC has better performance than INFORM and NDNF.

However, as a novel ICN routing mechanism, the design of RISC has some limitations to be overcome. On one hand, the community division is based on the static network deployment without considering the mobility of CRs (i.e., dynamic network deployment). On the other hand, when the amount of information in each IC reaches a certain level, it becomes much difficult to deal with them efficiently, which reduces IC sensitivity and further decreases routing efficiency. In future, we will improve RISC to overcome the mentioned limitations. Furthermore, we will study bio-inspired ICN routing, that is, detecting communities based on the behaviors of biology to find the multipath for interest requests.

Acknowledgments

We would like to thank the editors and all anonymous reviewers for helpful comments and suggestions, which have considerably improved and enhanced the quality of this paper. This work is supported by the National Natural Science Foundation of China (Grant No. 61572123) and the National Science Foundation for Distinguished Young Scholars of China (Grant No. 71325002).

References

[1] A.V. Vasilakos, Z. Li, G. Simon, W. You, Information centric network: research challenges and opportunities, J. Netw. Comput. Appl. 52 (2015) 1–10.
[2] C.I.S.C.O., Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2015-2020, San Jose, CA, USA, White Paper, 2016.
[3] C.I.S.C.O., 2017, http://www.cisco.com/c/en/us/solutions/index.html [accessed 26.03.17].
[4] G. Xylomenos, C.N. Ververidis, V.A. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K.V. Katsaros, G.C. Polyzos, A survey of information-centric networking research, IEEE Commun. Surv. Tutor. 16 (2) (2014) 1024–1049.
[5] J. Pan, S. Paul, R. Jain, A survey of the research on future internet architectures, IEEE Commun. Mag. 49 (7) (2011) 26–36.
[6] T. Koponen, M. Chawla, B.G. Chun, A. Ermolinskiy, K.H. Kim, C. Shenker, I. Stoica, A data-oriented (and beyond) network architecture, ACM SIGCOMM Comput. Commun. Rev. 37 (4) (2007) 181–192.
[7] V. Jacobson, D.K. Smetters, J.D. Thornton, M. Plass, N. Briggs, R. Braynard, Networking named content, Commun. ACM 55 (1) (2012) 117–124.
[8] A. Carzaniga, M. Papalini, A.L. Wolf, Content-based publish/subscribe networking and information-centric networking, in: Proceedings of the ACM SIGCOMM Workshop on Information-Centric Networking, 2011, pp. 56–61.
[9] NDN-NP, 2017 http://www.nsf.gov/ [accessed 26.03.17].
[10] K. Claffy, J. Polterock, A. Afanasyev, J. Burke, L. Zhang, The first named data networking community meeting (NDNcomm), ACM SIGCOMM Comput. Commun. Rev. 45 (2) (2015) 32–37.
[11] A. Afanasyev, Y. Yu, L. Zhang, J. Burke, K. Claffy, J. Polterock, The second named data networking community meeting (NDNcomm 2015), ACM SIGCOMM Comput. Commun. Rev. 46 (1) (2016) 58–63.
[12] ACM-ICN 2016, 2017 http://conferences2.sigcomm.org/acm-icn/2016/ [accessed 26.03.17].
[13] M.F. Bari, S.R. Chowdhury, R. Ahmed, R. Boutaba, A survey of naming and routing in information-centric networks, IEEE Commun. Mag. 50 (12) (2012) 44–53.
[14] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman, A survey of information-centric networking, IEEE Commun. Mag. 50 (7) (2012) 26–36.

[15] V. Ekambaram, K.M. Sivalingam, Interest flooding reduction in content centric networks, in: Proceedings of the IEEE International Conference on High Performance Switching and Routing, 2013, pp. 205–210.
[16] P. Chen, D. Li, NCAF: a reduced flooding mechanism for route discovery in ICN, in: Proceedings of the International ICST Conference on Communications and Networking in China, 2014, pp. 160–165.
[17] C. Yi, A. Afanasyev, L. Wang, B. Zhang, L. Zhang, Adaptive forwarding in named data networking, ACM SIGCOMM Comput. Commun. Rev. 42 (3) (2012) 62–67.
[18] A.K.M. Hoque, S.O. Amin, A. Alyyan, B. Zhang, L. Zhang, L. Wang, NLSR: named–data link state routing protocol, in: ACM SIGCOMM Workshop on Information–Centric Networking, 2013, pp. 15–20.
[19] C. Li, K. Okamura, W. Liu, Ant colony based forwarding method for content-centric networking, in: Proceedings of the International Conference on Advanced Information Networking and Applications Workshops, 2013, pp. 306–311.
[20] H. Qian, R. Ravindran, G. Wang, D. Medhi, Probability based adaptive forwarding strategy in named data networking, in: Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management, 2013, pp. 1094–1101.
[21] M. Aamir, Content-priority based interest forwarding in content centric networks, in: Arxiv, 2014, pp. 1–6.
[22] R. Chiocchetti, D. Rossi, G. Rossini, G. Carofiglio, D. Perino, Exploit the known or explore the unknown? hamlet-like doubts in ICN, in: Proceedings of the ACM SIGCOMM Workshop on Information-Centric Networking, 2012, pp. 7–12.
[23] R. Chiocchetti, D. Perino, G. Carofiglio, D. Rossi, G. Rossini, INFORM: a dynamic interest forwarding mechanism for information centric networking, in: Proceedings of the ACM SIGCOMM Workshop on Information-Centric Networking, 2013, pp. 9–14.
[24] D. Syrivelis, G. Parisisy, D. Trossen, P. Flegkas, V. Sourlas, T. Korakis, L. Tassiulas, Pursuing a software defined information-centric network, in: Proceedings of the European Workshop on Software Defined Networking, 2012, pp. 103–108.
[25] E. Aubry, T. Silverston, I. Chrisment, SRSC: SDN-based routing scheme for CCN, in: Proceedings of the IEEE Conference on Network Softwarization, 2015, pp. 1–5.
[26] B. Raghavan, T. Koponen, A. Ghodsi, M. Casado, S. Ratnasamy, S. Shenker, Software-defined internet architecture: decoupling architecture from infrastructure, in: Proceedings of the ACM Workshop on Hot Topics in Networks, 2012, pp. 43–48.
[27] J.C. Lee, W.S. Lim, H.Y. Jung, Scalable domain-based routing scheme for ICN, in: Proceedings of the International Conference on Information and Communication Technology Convergence, 2014, pp. 770–774.
[28] D. Kreutz, F.M.V. Ramos, P.E. Verissimo, C.E. Rothenberg, S. Azodolmolky, S. Uhlig, Software-defined networking: a comprehensive survey, Proc. IEEE 103 (1) (2015) 14–76.
[29] Open Networking Foundation, Software-defined networking: the new norm for networks, Palo Alto, CA, USA, White Paper, 2012.
[30] W. Xia, Y. Wen, C.H. Foh, D. Niyato, H. Xie, A survey on software-defined networking, IEEE Commun. Surv. Tutor. 17 (1) (2015) 27–51.
[31] C. Bu, X. Wang, F. Li, M. Huang, A big data driven adaptive routing service customization scheme, Chin. J. Comput. 39 (1) (2016) 65–79 (In Chinese).
[32] X. Wang, J. Li, Z. Tan, L. Ma, F. Li, M. Huang, The state of the art and future tendency of "Internet+" oriented network technology, J. Comput. Res. Dev. 53 (4) (2016) 729–741 (In Chinese).
[33] C. Bu, X. Wang, M. Huang, K. Li, SDNFV-based dynamic network function deployment: model and mechanism, IEEE Commun. Lett. (2016) in press, doi:10.1109/LCOMM.2017.2654443.
[34] Y. Zhu, D. Li, W. Xu, W. Wu, F. Fan, J. Willson, Mutual relationship based community partitioning for social networks, IEEE Trans. Emerg. Top. Comput. 2 (4) (2014) 436–447.
[35] J. Shi, X. Wang, M. Huang, K. Li, S.K. Das, Social-based routing scheme for fixed-line VANET, Comput. Netw. 113 (2017) 230–243.
[36] X. Nguyen, D. Saucez, T. Turletti, Providing CCN Functionalities Over Openflow Switches, 2013, Research Report, INRIA.
[37] M. Vahlenkamp, F. Schneider, D. Kutscher, J. Seedorf, Enabling ICN in IP networks using SDN, in: Proceedings of the IEEE International Conference on Network Protocols, 2013, pp. 1–2.
[38] S. Salsano, N. Blefari-Melazzi, A. Detti, G. Morabito, L. Veltri, Information centric networking over SDN and openflow: architectural aspects and experiments on the OFELIA testbed, Comput. Netw. 57 (16) (2013) 3207–3221.
[39] Deltacom, 2017 http://www.topology-zoo.org/ [accessed 26.03.17].
[40] GTS, 2017 http://www.topology-zoo.org/ [accessed 26.03.17].
[41] A.J.X. Lim, M.B.C. Khoo, W.L. Teoh, A. Haq, Run sum chart for monitoring multivariate coefficient of variation, Comput. Ind. Eng. 109 (2017) 84–95.
[42] M. Li, C. Fu, X. Liu, J. Yang, T. Zhu, L. Han, Evolutionary virus immune strategy for temporal networks based on community vitality, Future Generation Computer Systems, 2016. In press, http://dx.doi.org/10.1016/j.future.2016.05.015.

Jianhui Lv received the B.S. degree in mathematics and applied mathematics from the Jilin Institute of Chemical Technology, Jilin, China in 2012, and the M.S. degree in computer science from the Northeastern University, Shenyang, China in 2014. He is currently working toward the Ph.D. degree in the Northeastern University, Shenyang, China. His research interests include routing in ICN and operational research, etc. He has published more than 15 journal and conference papers.

Xingwei Wang received the B.S., M.S., and Ph.D. degrees in computer science from the Northeastern University, Shenyang, China in 1989, 1992, and 1998 respectively. He is currently a Professor at the College of Software, Northeastern University, Shenyang, China. His research interests include future Internet and cloud computing, etc. He has published more than 100 journal papers, books and book chapters, and refereed conference papers. He has received several best paper awards.

Min Huang received the B.S. degree in automatic instrument, the M.S. degree in systems engineering, and Ph.D. degree in control theory from the Northeastern University, Shenyang, China in 1990, 1993, and 1999 respectively. She is currently a Professor at the College of Information Science and Engineering, Northeastern University, Shenyang, China. Her research interests include modeling and optimization for logistics and supply chain system, etc. She has published more than 100 journal papers, books, and refereed conference papers.

Junling Shi received the B.S. degree in networking engineering from the Liaoning Technology University, Huludao, China in 2012, and the M.S. degree in computer science from the Northeastern University, Shenyang, China in 2014. Shi is currently working toward the Ph.D degree in the Northeastern University, Shenyang, China. Her research interests include routing in MSN and ICN, etc.

Keqin Li is a SUNY Distinguished Professor of computer science. His research interests include parallel computing and high-performance computing, distributed computing, energy-efficient computing and communication, heterogeneous computing systems, cloud computing, big data computing, CPU-GPU hybrid and cooperative computing, multicore computing, storage and file systems, wireless communication networks, sensor networks, peer-to-peer file sharing systems, mobile computing, service computing, Internet of things and cyber-physical systems. He has published over 390 journal articles, book chapters, and refereed conference papers, and has received several best paper awards. He is currently or has served on the editorial boards of IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Computers, IEEE Transactions on Cloud Computing, Journal of Parallel and Distributed Computing. He is an IEEE Fellow.

Jie Li is a Professor in computer science in Division of Information Engineering, Faculty of Engineering, Information and Systems, University of Tsukuba, Japan. His research interests are in Intelligent Mobile and Ubiquitous Computing, Networking, and Security, Internet Computing and Networking, Big Data, Cloud, OS, modeling and performance evaluation of information systems. He is a senior member of IEEE and ACM, and a member of IPSJ (The Information Processing Society of Japan). He is the founding chair of the IEEE Technical sub-Committee on Big Data (TSCBD), ComSoc. He received the B.E. degree in computer science from Zhejiang University, Hangzhou, China, the M.E. degree in electronic engineering and communication systems from China Academy of Posts and Telecommunications, Beijing, China. He received the Dr. Eng. degree from the University of Electro-Communications, Tokyo, Japan. He had served as a secretary for Study Group on System Evaluation of IPSJ and on Steering Committees of the SIG of System EVAluation (EVA) of IPSJ, the SIG of DataBase System (DBS) of IPSJ, and the SIG of MoBiLe computing and ubiquitous communications of IPSJ. He has served on several editorial boards for the IPSJ Journal and many international professional journals. He has also served on the program committees for several international professional conferences.