Content Searching Scheme Using Interesting Keyword Based Overlay Network

Rim Haw and Choong Seon Hong*
Department of Computer Engineering
Kyung Hee University
Yong In, Korea
{rhaw, cshong}@khu.ac.kr

Byeongsik Kim and Ho Young Song
Internet Research Laboratory
Electronics and Telecommunications Research Institute
Daejeon, Korea
{bskim25, hsong}@etri.re.kr

Abstract—**Content-Centric Network (CCN) sends data to avoid duplicated transmission with content name in the network. However, CCN may cause increase in overhead while maintaining contents management table and create storage problem while storing contents information in the node. In this paper, we proposed a new searching scheme using interesting keyword and manage contents table via overlay network. The content providers with high frequency are chosen in the network. Then an overlay network is constructed using content providers. Via overlay network, it is possible to reduce the number of message transmissions and overhead to manage content management table in the network.**

Keywords-Content Based Network, Overlay Network, Content Centric Network

I. INTRODUCTION

Internet is started by ARPANET in 1970 and it has rapidly grown up via the appearance of World Wide Web (www) in the early 1990. As the result, Internet has become very important part of human life. Many people use the Internet every day and everywhere using smart phone and other smart devices.

However as recent researches show, Internet traffic is not one-to-one communication traffic but most of it is download traffic to receive various contents which are spread in the Internet like music files, movie files and photographs. It means that download traffic is taken up recent Internet traffic [1]. Today, IP based data transmission resends same data at the number of people who want to use data in the network, because it provides services regardless of kinds of data.

To discard this inefficient data transmission, many researchers start studying of the way which does not use IP address in recent Internet but is uses data name for data communication in the network. Among these researches, Content Centric Network (CCN) is one of the content based networks for future Internet [2, 3, 4]. In CCN, routing scheme is content-aware, so it can search close contents efficiently. This routing scheme can solve retransmission problem in recent Internet. Some researchers predict that the number of contents would increase rapidly in the near future. It might create overhead for maintaining routing table and storages and it is a big problem in the Content Centric network [5]. To solve overhead problems, in this paper we discuss about content searching scheme for increasing efficiency of content delivery in the content based network. We consider overlay network which consists of content providers. Proposed scheme can search content using interesting keyword in the overlay network.

The remainder of this paper is organized as follow. In section II, we discuss about what is CCN, CCN routing scheme [6]. Section III introduces our proposed scheme using interesting keyword based overlay network in the content based network. For the verification of our scheme, we present the result of simulation in section IV and we explain conclusion and our next step in section V.

II. RELATED WORKS

A. Content-Centric Network

Recently, to improve the current Internet, many researchers have tried to keep recent Internet architecture and change some parts of function in the network. However this partial improvement did not solve structural problem of recent Internet. As a result, Future Internet research is started via clean slate approach. Content-Centric Network (CCN) is proposed as a part of future Internet scheme for research which concern with sharing data.

Present research related content-centric network has just started. Many research organization and researchers are interested in CCN. PARC proposed Content Centric Network and FIND project researches Data Oriented Network Architecture (DONA). In this content centric network, proposed scheme provides advantages like fast services and protect data sender's repetitive data transmission, etc. to supplement disadvantage which resends data via one-to-one communication between sender and receiver [7]. Figure 1 is shown a sandglass of CCN architecture and the current Internet.

This research was supported by the KCC(Korea Communications Commission), Korea, under the R&D on Smart Node Technology for Cloud Networking and Contents Centric Networking support program supervised by the KCA(Korea Communications Agency)"(KCA-2011-11913-05003). Dr. CS Hong is the corresponding author

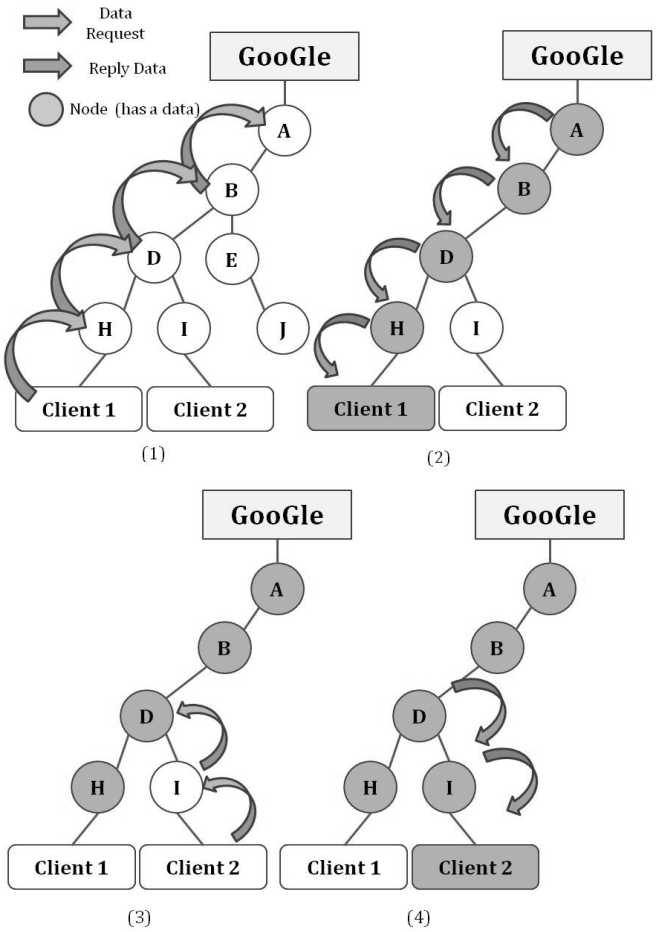

Figure 1. Basic Routing Scheme for CCN

B. Basic routing scheme for CCN

Figure 2 shows a basic routing scheme [8] for content centric network. In figure 2 (1), client 1 requests content to its parent node. When parent node receives client 1's request message, it checks its content management table to be request content or not. If request content is found from parent node, parent node sends a request content to client 1. However if content is not in parent content management table, parent node sends a request message to upper node. Such as the way, request message is sent to the highest parent node. Figure 2 (2) shows the procedure of sending request content. Top node (the highest parent node) sends a content using reverse way and when each node receives contents, each node stores the content into content management table. Finally, client 1 receives requested content from node H. Figure 2 (3) and (4) show that client 2 requests same content which is requested from client 1. In this case, client 2's request message is sent to node D. Then node D sends requested content to client 2.

C. Problem statement

In a basic routing scheme for CCN, network traffic is reduced in transmitting same data. However, it is possible to increase network overhead in maintaining content storage and content management tables, because CCN stores all requested content in the node storage. When data searching fails, CCN also sends messages using flooding to find the requested content.

III. PROPOSED SCHEME

In this paper, we try to solve problems caused by CCN routing scheme. We make routing scheme called interesting keyword based overlay network in CCN. Overlay network via content provider who has a high frequency rate and use interesting keyword to find content.

A. The Architecture of Interesting Keyword based Overlay Network

To construct a proposed overlay network in this paper, each node makes a content management table like table I. Content management table includes address of content provider, interesting keyword which is found by users, information of neighbor nodes who have same kind of contents, registered time and the number of frequency rate.

Each node constructs overlay network like a process shown in figure 4. After deploying overlay network with content providers, each node does not use neighbor node but they use overlay network to find contents.

Content providers, which are consisted of overlay network, make a table which has a hash value such as table II. In overlay network, each node checks which content provider has an aim content (user wants to find content) via interesting keyword or not. Figure 4 shows an overall architecture of the interesting keyword based overlay network.

B. Maintaining of Content Management Table

To solve the problem of increasing size of content management table in CCN, each table gives a high priority to content which has a high frequency rate within a certain time period and the number of searches in the content searching.

TABLE I. CONTENT MANAGEMENT TABLE

Content-Provider	Interesting Keyword	Neighbors	The Number of Frequency	The Time of Entrance
Google	Paper	B, C	110325140	42
YouTube	Movie	D	110325135	200
Naver	New Paper	E,F	110325130	60
MNet	Music	A	100407830	70
-	-	-	-	-
-	-	-	-	-

TABLE II. INTERESTING MANAGEMENT TBALE IN THE OVERLAY NETWORK

Hash key	Interesting Keyword	The Number of Frequency	The Time of Entrance
10352a1	Movie	1103251400	200
10452a1	Music	100407830	70
-	-	-	-

(1) Each nodes check frequency rate of content providers for a period in the table.
(2) Nodes choose content provider which has the highest frequency rate.
(3) Nodes makes a hash key of content provider via hash function
(4) Overlay network make these hash key and their interest keywords

Figure 2. How to construct the Overlay Network

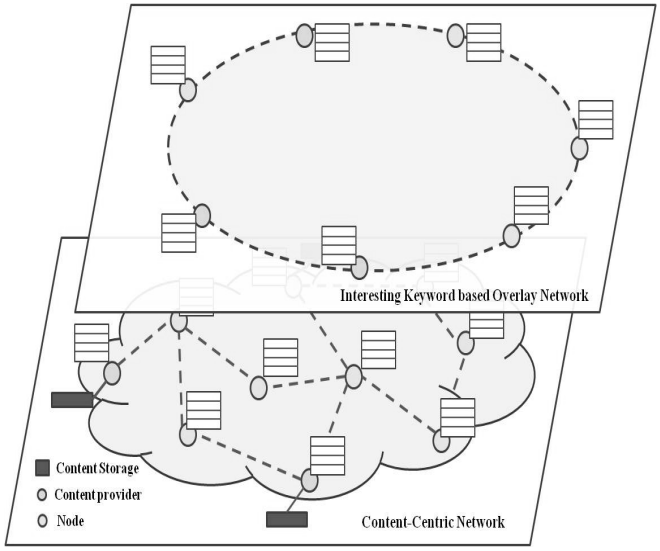

Figure 3. The Architecture of Interesting Keyword based Overlay Network

And management table deletes contents which have low priorities. However, to protect new registered contents which has a low selecting rate are deleted, table compares registered time with the number of searching contents. If registered time is passed with some threshold value (time), content management table delete the content information with low priority.

C. *A Scenario of Searching Content*

In this paper, Interest Keyword Based Overlay Network searches for contents as the following sequence.

- When the network is deployed, each node registers content information in the table and checks the frequency rate of contents for certain time period.
- After a certain period of time, each node chooses a content provider which has a high searching rate and constructs an overlay network.
- Each content provider makes a hash key to reduce the size of interesting management keywords.
- Each node can find contents using interesting key word through content providers via overlay network which has a ring topology.

provider 3 sends user A's request message to its neighbor, as shown in figure 6 (b).

3) Content provider 2 receives user A's request message. Content provider checks its table again. If content provider 2 has a same interesting keyword, it checks its content storage to find a required content and sends content to user A, as shown in the figure 6 (c).

Figure 6 shows a scenario that how user A finds a required content, "music" in the proposed network.

1) User A sends a request message including user A's interesting keyword to content providers in overlay networks as shown in figure 6 (a).

2) When content provider 3 receives user A's request message, content provider 3 checks its content management table to compare its interesting keyword and user A's interesting keyword. However it does not match with user A's interesting keyword. So, content

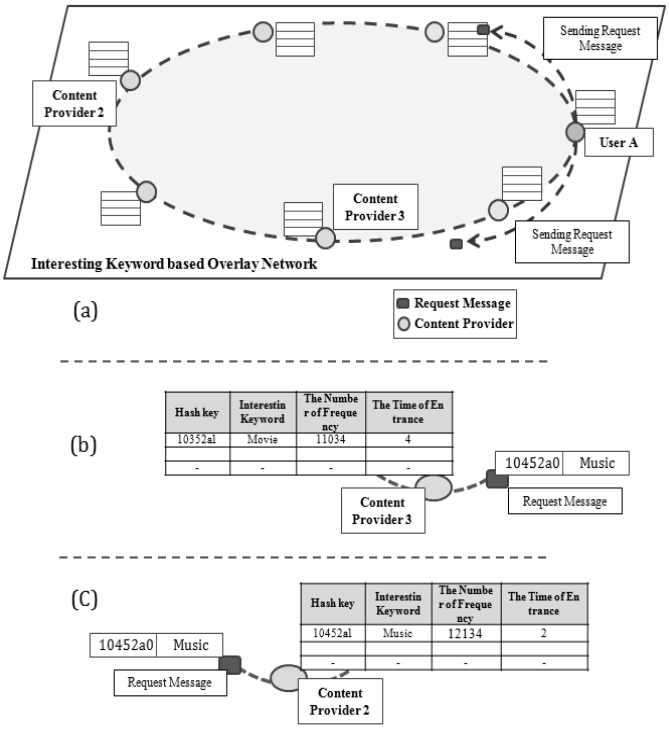

Figure 4. A Scenario of Searching Content in the Overlay Network

IV. PERFORMANCE EVALUATION

In this paper, we have introduced our proposed scheme which constructs overlay network and content searching using interesting keyword. To evaluate the performance of our scheme we have performed simulation using C++. We increased the number of nodes and compared with the number of packets for both recent CCN and our proposed overlay based CCN. We consider 10 interesting keywords such as Movie, Music News paper, Book, electronic device, and car etc., 10 content providers and 200 nodes. The experimental environment is illustrated in table III.

TABLE III. THE EXPERIMENTAL ENVIRONMENT

Parameter	Value
Total Number of Nodes	200
Total Number of Interesting Keywords	10
Total Number of Content Providers	10

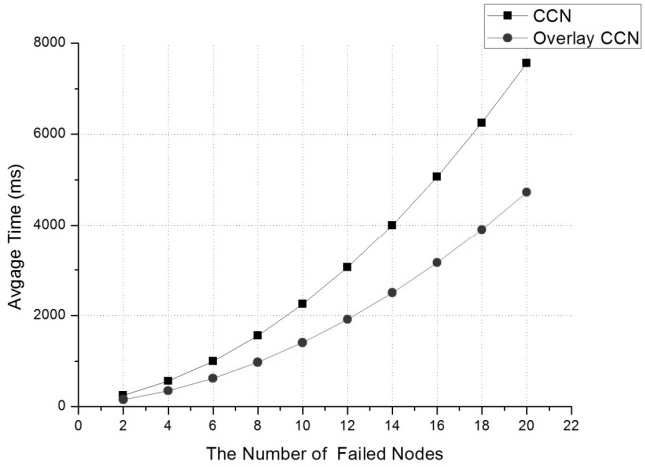

Figure 5. Comparing Average Delivery Time as a Fail of Searching Content

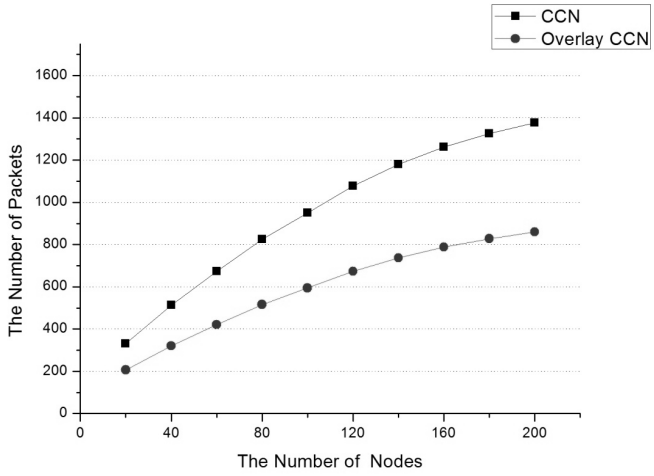

Figure 6. Comparing increasing the number of packets as the number of nodes for finding content

Figure 7 shows the number of increasing packets between recent CCN and proposed overlay based CCN when user searches contents in the network. As a result of simulation, our proposed routing scheme can communicate using lesser number of packets than CBCB routing.

As shown in figure 8, we have compared average delivery time between our overlay based CCN and CCN. In this case, we can know request time is increased continuously as CCN sends request message to neighbor nodes with flooding. And our proposed scheme requires less time but we can see our scheme's decreasing rate is not high. That reason is it spent time to deploy overlay network. We try to solve this time problem, which is increased content searching time when searching content is failed, in the near future.

V. CONCLUSION AND FUTURE WORKS

In this paper, we try to solve problems of increasing the number of messages, storage of each node and maintaining routing table when content is searched in a Content Centric Network. So we proposed content searching scheme using interesting keyword in the overlay network. We choose content providers which has high frequency rate in the network after constructing CCN. Then, overlay network, which consists of chosen content providers, is deployed to reduce the number of message transmissions. Nodes can find requested contents with interesting keywords in the overlay network. As the simulation result shows, our proposed scheme is better in performance than recent CCN.

Our scheme has a few problems. One is that our scheme uses the way to select content providers which has high selecting rate in the network. This choosing of providers causes problem of central traffic at the providers. The other one is that when overlay network is constructed, there are management table in the overlay network. It causes overhead to maintain table. So, our future works is to solve these problems. Also we try to find optimal way to select content provider with high frequency rate.

REFERENCES

[1] V. Jacobson, D.K. Smetters, J.D. Thornton, M.F. Plass, N.H. Briggs, and R.L. Braynard, "Networking Named Content," In CoNEXT' '09, Rome, Italy, Dec. 2009.

[2] Diana Smetters and Van Jacobson, "Securing Network Content," PARC Technical Report, Oct. 2009.

[3] Ipoque - Internet study 2008/2009 The Impact of P2P File Sharing, Voice over IP, Instant Messaging, One-Click Hosting and Media Streaming on the Internet. April, 2009.
http://www.ipoque.com/resources/Internet-studies/Internetstudy-2008_2009

[4] M. Meisel, V. Pappas, L. Zhang. "Ad Hoc Networking via Named Data," MobiArch '10, Chicago, USA Sept. 2010.

[5] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," Proc. ACM SIGCOMM, Aug. 2007.

[6] Content-Based Networking in University of Colorado
http://serl.cs.colorado.edu/serl/cbn/

[7] Kideok Cho, Jaeyoung Choi, Dong-il Diko Ko, Taekyoung Kwon, and Yanghee Choi, " Content-Oriented Networking as a Future Internet Infrastructure: Concepts, Strengths, and Application Scenarios," in Proc. of International Conference on Future Internet Technologies (CFI) 2008, Seoul, Korea, June 2008

[8] A. Carzaniga, M.J. Rutherford, and A.L. Wolf, "A Routing Scheme for Content-Based Networking," Proceedings of IEEE INFOCOM 2004. Hong Kong, China. March, 2004