Research Article

Efficient Content Delivery for Mobile Communications in Converged Networks

Mahfuzur Rahman Bosunia and Seong-Ho Jeong

1Bangladesh Bank, Dhaka, Bangladesh
2Dept. of Information and Communications Engineering, Hankuk University of Foreign Studies, Seoul, Republic of Korea

Correspondence should be addressed to Seong-Ho Jeong; shjeong@hufs.ac.kr

Received 3 August 2018; Accepted 4 November 2018; Published 14 January 2019

1. Introduction

Recently, as the use of mobile devices such as smart phones and tablets has increased exponentially, the number of services that connect mobile devices to the Internet has also increased. Accordingly, the volume of the mobile content has increased enormously, and it is a big challenge to support fast content delivery as well as seamless mobility in the wireless network. The content-centric networking (CCN) [1, 2] architecture can be useful to resolve this issue.

The current Internet is configured for host-to-host communication, and therefore it is not suitable for the future Internet which will deal with various contents and provide seamless mobility for moving users. The main focus of CCN is to provide more efficient, faster, and secured delivery of a content rather than to establish the communication path to the content source. CCN provides name-based routing without exploiting the content/device address. CCN operates using two simple messages for content transmission: first one is the Interest packet and the other is the Data packet. The Interest packet contains the request for a desired content, and the request consists of different attributes such as content name, content type, and content version. The Data packet contains the original data with the content name, security related information, and several other attributes, e.g., hop distance and content source description.

Figure 1 shows the CCN forwarding module which is equipped with three functional and operative elements for content-based routing: Content Store (CS), Forwarding Information Base (FIB), and Pending Interest Table (PIT). The CS is the physical storage of the content and stores the identical name of the published content. The FIB preserves the content routing information for mapping between the content name and the next hop towards the content source. The PIT keeps track of and records the status of all received Interest packets in order to satisfy later when the content is on hand. Any node that wants to publish a content spreads the content name to the nearby nodes to make it available.
and accessible to the content consumers. A node sends an Interest packet to retrieve a content based on the routing information available in the FIB. When a node receives the Interest packet, it enforces a lookup inside the CS to check the content availability. If the requested content is available then the content is delivered back to the content requester. If the content is not available inside the CS, a pending Interest is recorded in the PIT and the Interest packet is forwarded via the face guided by the FIB.

The LTE-based 4G mobile network establishes a network architecture to provide faster data retrieval and seamless mobility management. Even though LTE promises to be a faster and more efficient data delivery network, its communication architecture is still centralized and host oriented; thus various performance degradation issues, e.g., high bandwidth consumption, cross domain traffic, high delay for the long communication path, and waste of network resources, may need to be resolved. A new paradigm called CCN is considered as an appropriate way of efficient content delivery as mentioned above. The integration of CCN and LTE can be used to resolve the issues by providing timely and fast delivery of the data with efficient resource utilization.

In this paper, we propose a CCN-based efficient content delivery mechanism in the 4G network and also in the upcoming 5G network where various heterogeneous networks are converged. We also propose an efficient mobility management mechanism to address the content diversity and network diversity by leveraging the abundant computational resources in the LTE-based 4G network. Our proposed mobility management scheme introduces a context-aware handover prediction mechanism to deal with the heterogeneity of wireless content providers and consumers. Context awareness provides a significant effect to guarantee data continuity in the mobile environment. We also introduce a content similarity matching mechanism to provide the data continuity when mobility increases the chance of data unavailability. The rest of the paper is organized as follows. Section 2 describes the recent work related to content delivery in CCN. Section 3 presents a novel approach to integrate CCN with LTE and some mechanisms for efficient and seamless content delivery in the LTE network in detail. Section 4 describes implementation details and performance issues, and finally, Section 5 concludes the paper.

2. Related Work

Seamless content delivery and mobility management in wireless networks are important research issues which have been focused a lot under various scenarios. A lot of works have been proposed so far in this area. The LTE network is limited to the measured value of the received signal strength which is a factor that triggers handover in the wireless network. In reality, handover involves content delivery rate requirements for a particular application, packet loss, end-device demand for high/low data rate, and many more.

There are various mobility-related mechanisms and protocols across all the layers in the TCP/IP protocol stack for providing seamless content transmission. For example, Mobile IP [3] provides mobility support at the Network layer; the Stream Control Transmission Protocol (SCTP) [3] and the Datagram Congestion Control Protocol (DCCP) [3] provide seamless mobility support at the transport layer. Dynamic DNS (DDNS) [3] and Session Initiation Protocol (SIP) [3] are examples to provide mobility management at the application layer. The efficiency of these mechanisms is limited due to the acting of IP address as a locator as well as an identifier and also due to the cross layer communication between different layers. The separation of the locator from its identifier is proposed in Host Identity Protocol (HIP) [3] and Locator Identifier Separation Protocol (LISP) [4], but they cannot come out from the host-to-host communication scenario.

![Figure 1: A forwarding module for CCN.](image-url)
The future Internet architecture based on Information-Centric Networking (ICN) such as CCN, Name Data Networking (NDN), Data-Oriented Network Architecture (DONA), and Network of Information (NetInf) uses in-network content caching to improve the efficiency of content transmission, reduce the network traffic and content access latencies, alleviate the present communication bottlenecks, and support ubiquitous access and efficient mobility management. DONA [5] introduces integrated name resolution and content-based routing schemes by replacing the concept of DNS in the traditional TCP/IP based Internet architecture. DONA uses a flat, self-certifying name of a content and registers the content name and the location in a domain server called Resolution Handlers (RHs). RHs are structured into a BGP topology of the network and content lookups are performed by querying a consumer to its local RH. If no reference of the requested content is found, the query is forwarded up the tree until a content source is found; an out-of-band delivery path is then established by the source (over IP). DONA reduces the applicability due to the dependency on RHs like DNS and location-based communication. The amount of delay and overhead may also be a big concern for DONA.

Network of Information (NetInf) [6] follows the similar mechanisms as DONA and also provides content delivery using a name resolution (NR) service. To handle content and device mobility issues and provide content-based routing, NetInf uses Multiple Distributed Hash Table (MDHT) [7] and Late Locator Construction (LLC) [8] schemes. MDHT and LLC try to make the content management and name resolution simple and provide in-network content caching. PERSUIT [9] uses three key components, Rendezvous, Topology, and Routing, to provide seamless content delivery and mobility management in a publish/subscribe architecture. Bloom filter based source routing is used to transfer content through the network [10]. However, NetInf reduces its efficiency due to the dependency on the NR which is responsible for content registration, content updates, and content-based route establishment. It requires a re-binding similar to DONA. The mechanism for handling mobility in NetInf may vary according to the chosen content locator; thus the implementation can be complex. The cost of updating the routing information for PERSUIT is very high. The packet loss may be significant in the case of high mobility.

The CCN architecture [11,12] decouples the content from the location and device and distributes the content using content name-based routing. In CCN, consumer mobility is managed inherently by its receiver driven nature. There is no need to update the routing information due to mobility from a consumer's perspective; the consumer just retransmits Interest packets if the content is not available yet. Even though CCN supports consumer mobility inherently, it faces long delays to re-issue the Interest after rebinding to a new network and cannot provide seamless content transmission. Content provider mobility is still an open issue in CCN; there are several issues to be resolved such as update of routing and location information, repeated transmission of Interest/Data packets, and undesirable content delivery delays due to mobility.

A proxy-based approach [13] proposes a publisher mobility support protocol in CCN and a fast FIB update mechanism. It introduces the mobility entry in FIB for mobile and temporal destination and also defines the Home Router (HR) or proxy to announce the original entry of the publisher to the network and establish a tunnel between the previous point of attachment and a new point of attachment to reduce the packets loss due to mobility. Clustered CCN [14] introduces the cluster concept to support mobility which can be viewed as a hierarchical mobility management scheme to support the extensive mobile domain. It forms a cluster with a cluster head that manages all the responsibilities of its members like Interest processing and mobility tracking. With the assistance of the cluster or proxy, it can reduce the Interest dissemination and content distribution, but the overhead and complexity of this approach are high due to its centralized and hierarchical nature.

A converged network is useful to exploit content diversity and device heterogeneity by disseminating contents through several networks, e.g., Wi-Fi, broadcast networks or cellular networks [15, 16]. Despite a large variability of content requests to several routes and several content sources for efficient transmission of user requests and content storing, content requests are typically satisfied by the nearby devices or networks [17–19]. The separation among content delivery, content storage, and content and device mobility operations may reduce the performance efficiency of the network and increases the operational complexity of the network.

Fetching the content before handover was proposed in [20] to support producer mobility in name-based routing. Software defined controller for CCN [21] proposed a mobility management mechanism for allowing packet forwarding and intermediate routing on the device mobility. Software Defined Mobile Network [22] was proposed to improve the content delivery efficiency by optimizing caching in the LTE network in which Software Defined Networking (SDN) mechanisms were integrated with the Mobility Management Entity (MME). It allows dynamic relocation of contents in any intermediate node. This mechanism was compared and evaluated in [23] by using simulation. The simulation results showed that the in-node content caching reduced traffic load and improved content delivery efficiency. Even though intermediate content storing ensures faster content access, the virtual tunnel-based content redirection in the LTE network increases the overhead and reduces the transmission efficiency.

A mobility direction prediction mechanism was proposed in [24] for reducing the number of handovers and data losses in the LTE network. However the scope of this work is very limited due to the TCP/IP-based communication nature. The future mobile networks demand the mobility and portability of devices and data or contents in an autonomous and adaptive way to provide seamless content delivery, to be connected to several access networks simultaneously, and maintain the high quality of content transmission without any interruption even in the highly mobile environment. The proposed mechanism is an enhancement to our previous work [25], which makes it possible to directly fetch and store contents in any appropriate node before handover to enable
faster content retrieval for reducing content transmission delays and preventing the repeated transmission of Interest/Data packets to avoid the network congestion.

3. A CCN-based Mechanism for Seamless and Efficient Content Delivery in the Mobile Network

Mobility in the Internet means that either the consumer or the provider is moving away from its point of attachment to another or both of them are moving away together. Mobility support is a service such that the mobility of the node should not result in any loss of data or extended periods of disconnection. It is still an issue how CCN will be integrated with the LTE network. We simplify here a seamless data delivery procedure in the content-centric LTE network. The eNodeB, SGW/PGW, and MME can support the CCN function and protocols. The detailed content delivery procedure in the content-centric LTE network is described below and also shown in Figure 2.

(i) Step 1: UE$_1$ establishes a connection to an eNodeB.
(ii) Step 2: the eNodeB sends the UE$_1$ information containing the node identifier and the interface identifier to the SGW/PGW.
(iii) Step 3: UE$_1$ sends the Interest packet via the eNodeB. When the eNodeB receives the Interest packet, it performs according to the basic operation of the CCN node.
(iv) Step 4: the SGW/PGW receives the Interest packet from the eNodeB. After searching its CS and PIT, if no matched content is found, it checks its FIB.
(v) Step 5: the SGW/PGW’s FIB sends the Interest packet via the face of the content provider.
(vi) Step 6: when a content provider receives the Interest packet, it searches its CS. When the matched content is found, it sends the Data packet of the content out via the incoming face.
(vii) Steps 7, 8: when the SGW/PGW receives the Data packet, it forwards it to the eNodeB and UE$_1$ based on their PIT.
(viii) Step 9: UE$_1$ terminates the connection.
(ix) Step 10: UE$_2$ establishes a connection to the eNodeB.
(x) Step 11: the eNodeB sends the UE$_2$ information containing the node identifier and the interface identifier to the SGW/PGW.
(xi) Step 12: UE$_2$ sends the Interest packet to the eNodeB.
(xii) Step 13: when the eNodeB receives the Interest packet, it performs according to the basic operation of the CCN node. Since it is cached in the previous eNodeB, it sends the Data packet to UE$_2$.
(xiii) Step 14: UE$_2$ terminates the existing connection.

3.1. Device Mobility Prediction. Device mobility or consumer mobility allows consumers to change their point of attachment without disrupting the connectivity. We have defined a mathematical model which takes into account the preference of end devices, e.g., UEs, when selecting a base station for seamless and fast content retrieval and content delivery services. Let $x$ be a value for a single criterion and $\alpha$ be the steepness. $x_{\text{min}} \leq x_m \leq x_{\text{max}}$ where $x_m$ is the midpoint of the
variation range. These variations can be defined as [26] using the single criteria utility function as shown in

\[ u(x) = \begin{cases} 
0 & \text{if } x \leq x_{\text{min}} \\
\frac{1}{1 + e^{\beta(x-x_{\text{m}})/(x_{\text{m}}-x_{\text{min}})}} & \text{if } x_{\text{min}} < x \leq x_{\text{m}} \\
m & \text{if } x_{\text{m}} < x \leq x_{\text{max}} \\
1 & \text{if } x \geq x_{\text{max}} 
\end{cases} \tag{1} \]

where

\[ \beta = \frac{\alpha(x_{\text{max}} - x_{\text{m}})}{x_{\text{m}} - x_{\text{min}}} \tag{2} \]

and \( \alpha > 0 \) is the tuned steepness parameter. The proposed utility function satisfies the following properties: \( u(x) = 0 \) \( \forall x \leq x_{\text{min}} \), \( u(x) = 1 \) \( \forall x \geq x_{\text{max}} \), and \( u(x_{\text{m}}) = 0.5 \). The point of attachment selection in the wireless networking environment is based on the aggregation of different utility functions for decision processes. Hence, we define here a multicriteria utility function that is able to integrate the end devices’ different choice metrics to select a best point of attachment. Let \( R_1, \ldots, R_n \) be a set of potential alternatives (e.g., possible different eNodeBs) and each alternative can be described as a different descriptor or attributes (e.g., received signal strength, mobility direction, and load in terms of data transfer rate) \( x = x_1 \ldots x_n \), and each alternative attribute being described as a utility function \( u(x_i) \), the simple weighted average of different alternatives \( A(R_1 \cdots R_n) \) is used to maximize the selection probability of the best eNodeB as follows:

\[ A_R = \sum_{i=1}^{n} w_i u(x_i) \tag{3} \]

where \( w_i \) is a weight that reflects the content receiver’s preference. Weights are assigned according to the UE’s expected criteria.

(i) Received Signal Strength. Received signal strength (RSS) is one of the most popular parameters to take a handover decision. By monitoring the RSS, it is easily determined whether the UE should connect to a new eNodeB or not. The UE reports the received RSS value for all the neighbor eNodeBs to the serving eNodeB. The eNodeB that takes a handover decision uses the RSS values of each eNodeB in a single criteria utility function as shown in (1).

(ii) Mobility Direction. The proposed mechanism also uses the moving direction prediction of each UE to make the decision of the movement towards an eNodeB. We assume that each eNodeB is aware of the position of the 2-hop neighbor eNodeBs and each UE is aware of its own position. Assuming the serving eNodeB position is \( (X_u, Y_u) \), the position of the UE is \( (X_n, Y_n) \), and the candidate eNodeB position is \( (X_m, Y_m) \), using the Pythagorean Theorem, it is possible to estimate distance between a UE and a candidate eNodeB as shown in

\[ d = \sqrt{(X_n - X_u)^2 + (Y_n - Y_u)^2} = \sqrt{(\Delta X)^2 + (\Delta Y)^2} \tag{4} \]

Therefore, the probability of a UE being in a coverage area of an eNodeB is

\[ P_r \{ d \leq R_r \} = P_r \left\{ \sqrt{(\Delta X)^2 + (\Delta Y)^2} \leq R_r \right\} \tag{5} \]

So \( d \) is normalized using the coverage \( R_r \) as follows:

\[ d_R = \frac{R_r - d}{R_r} \tag{6} \]

Since the velocity is a vector and a UE moves to different direction, it is reasonable to predict the direction or angle of the UE towards a candidate SBS using the vector formula. The moving angle of the UE from the associated eNodeB to a new candidate eNodeB can be estimated as shown in

\[ \theta = \cos^{-1} \frac{X_nX_u + Y_nY_u}{\sqrt{X_u^2 + Y_u^2} \sqrt{X_n^2 + Y_n^2}} \tag{7} \]

It is considered that the \( 120' \) angle is the acceptable angle towards a candidate eNodeB as in [24]. So it is considered as an offset of \( \mu = \pm 60' \) to normalize the \( \theta \) value as shown in the following formula

\[ \theta_R = \frac{\mu - \theta_R}{\mu} \tag{8} \]

Then \( \theta \) and \( d \) are used to estimate the movement prediction \( P_m \) as shown in the following formula

\[ P_m = (1 - \alpha) * d_R + \alpha * \theta_R \tag{9} \]

The eNodeB that takes a handover decision uses the movement prediction values of (9) of each eNodeB in a single criteria utility function as shown in (1).

(iii) Load. In order to take the accurate context-based decision for handover, the load of a candidate eNodeB was additionally considered. In some cases, the UE can attach to an eNodeB which has a greater RSS value but might be overloaded in terms of connected UEs. In other words, based only the RSS and the number of associated UEs that are currently associated with an eNodeB, the handover decision may experience the degradation of a performance. Each eNodeB transmits its current workload to its two hop neighbor eNodeBs. The eNodeB that takes a handover decision uses the work load value of each eNodeB in a single criteria utility function as shown in (1).

3.2. Candidate eNodeB Selection. The handover decision is made in the serving eNodeB. The UE reports the measurement of RSS values of all the candidate eNodeBs and its own position information to the associated eNodeB. Each eNodeB sends its load estimation to 2-hop neighbor eNodeBs. Then the associated eNodeB uses (1) to make the utility estimation of each alternative of the eNodeB. For each candidate eNodeB, three different utility values are estimated and combined in the aggregated metric function as shown in (3). The serving eNodeB selects the eNodeB which has the highest aggregated metric value.
3.3. Seamless Content Retrieval Using MME. Consumer mobility allows consumers to change their point of attachment without disrupting connectivity. The point of attachment selection in the wireless networking environment is based on an aggregation of different utility functions for decision processes. Hence, we define here multicriteria that are able to integrate the end devices’ different choice metrics to select the best point of attachment. This paper proposes a soft-handover approach where a new connection is established with the new eNodeB before breaking the current connection. This is almost similar to the follow-me service that is the modern trend of mobile communications. In the proposed approach, UEs send the measurement reports to the serving eNodeB. Then the eNodeB follows the procedure mention in Section 3.1 to decide whether the UE will move to a new eNodeB or not. If the serving eNodeB decides the necessity for a new eNodeB to continue the seamless content retrieval of the UE, it forwards the Interest and related information to the new eNodeB; then the new eNodeB forwards the Interest to the most appropriate content provider to retrieve the content. If the content retrieval is successful, the UE releases the connection with the old eNodeB and continues the content transfer using the new eNodeB. The message flow for seamless content retrieval of the UE is shown in Figure 3.

The detailed operational procedure for seamless content retrieval of the UE in the content-centric LTE network is described below.

(i) Step 1: an UE establishes a new connection to eNodeB1.

(ii) Step 2: eNodeB1 sends details about the UE, e.g., UE identifier, interface identifier to MME.

(iii) Step 3: the UE sends an Interest message via eNodeB1. When eNodeB1 receives the Interest, it follows the same procedure performed by a CCN node.

(iv) Step 4: eNodeB2 receives the Interest message from eNodeB1 and follows the same procedure performed by a CCN node. After doing look-up on its CS and PIT, it forwards the Interest to the mobile content source (UE Provider).

(v) Step 5: after forwarding the Interest message to the content source, it adds PIT entry to forward content in future.

(vi) Step 6: when the content source receives the Interest message, it looks up its CS. When the matching content is found, it replies back with the Data message as a response through the arrival interface of the Interest message.

(vii) Step 7,8: eNodeB2 and eNodeB1 forward Data packets to the UE.

(viii) Step 9: during the ongoing content transfer, the eNodeB1 estimates the mobility prediction and decides whether it will move from eNodeB1 or not using (3), as in Section 3.1. If the serving eNodeB1 finds the best candidate for content transmission, the eNodeB1 sends the chunk Interest to the MME.
(ix) Step 10: eNodeB1 receives the Data packet but cannot transmit it to the UE efficiently. In our proposed mechanism, we maintain a separate buffer as shown in Figure 4 to store and forward the Data packets for later transmission. If an eNodeB tries to buffer the packets in the main queue, buffer overflow may hamper the normal operation of the eNodeB. To solve the problem, we maintain an extra buffer so that all the Data packets for upcoming handover can be stored separately by stamping as buffered packets. To avoid buffer overflow it checks the current queue status using the Exponentially Weighted Moving Average (EWMA) formula as shown in the following formula

\[ Q_{avg} = (1 - \alpha) \times Q_{avg} + Q_{curr} \times \alpha \]  

(10)

where \( \alpha \) is a weight factor and \( Q_{curr} \) is the current queue size. eNodeB1 can buffer packets if the current status falls below or is equal to a minimum threshold called \( Q_{Th} \) calculated as shown in the following formula

\[ Q_{Th} = w \times Q_{size} \]  

(11)

where \( w \) is a weight factor. Based on the extra buffer occupancy, if the value of \( Q_{avg} \) is \( \leq Q_{Th} \), eNodeB1 discards the Data packet. MME sends the Interest message to eNodeB1 to transfer the buffered Data message to the desired eNodeB where the UE wants to move, e.g., eNodeB3.

(x) Step 11: MME forwards the received Interest message to the desired eNodeB where the UE wants to move, e.g., eNodeB3.

(xi) Step 12: eNodeB3 receives the Interest message and forwards the Interest message to the eNodeB2. It also receives the buffered Data messages sent by eNodeB1.

(xii) Step 13: after receiving the Interest message from eNodeB3, eNodeB2 sends the Data message as a response to the Interest message.

(xiii) Step 14: The UE terminates the connection to the eNodeB1 and uses the same physical handover operation like LTE and makes a connection to eNodeB3. It is a UE initiated handover, and the UE establishes a new connection to eNodeB3.

(xiv) Step 15: eNodeB3 sends the UE details, e.g., UE identifier, interface identifier to MME.

(xv) Step 16: since the UE's identifier is already registered in the MME, MME identifies that the UE is moved from eNodeB1 to eNodeB3. MME sends the Interest message to eNodeB1 and eNodeB2 to reconfigure the previous path.

(xvi) Steps 17, 18: the UE sends the Interest to the new eNodeB, e.g., eNodeB3, and continues to receive the content seamlessly using the new eNodeB.

3.4. Seamless Content Delivery Using MME. Provider mobility allows sources to relocate without disrupting content availability. In order to reduce handover latency and the cost of the provider mobility in CCN, we propose a new mechanism that can allow soft-handover approach where a new connection is established to a new eNodeB before breaking the old connection. Once a handover occurs, the producer will update its prefix to match the new location (e.g., when a producer named /prefix moves from eNodeB1 to eNodeB2, the producer's name will change from /eNodeB1/prefix to /eNodeB2/prefix).

If the producer changes its attachment point, i.e., eNodeB, its location name becomes invalid and Interests from the UE and eNodeB will no longer reach the content source. As soon as it is assigned a new location name at the new eNodeB, the eNodeB and MME update the binding information. Consumers exploit CCN’s multipath forwarding to handle handovers. Due to mobility, if the content name is changed and the producer receives the old named Interest message, then it can use similarity matching mechanism to satisfy the Interest. Let a content be denoted by \( d \) which consists of naming components or attributes, e.g., location and type denoted by \( v_i \). Thus the content name is represented by \( d = (v_1^d, v_2^d, \ldots, v_m^d) \). Then for similarity matching, the following formula shown in (12) is used to calculate the similarity, \( S_{1,2} \), between the content item \( d_1 \) and the content item \( d_2 \).

\[ S_{1,2} = \frac{\sum_{i=1}^{m} w_i \times B(v_i^1, v_i^2)}{\sum_{i=1}^{m} w_i} \]  

(12)

where \( m \) is the number of qualitative attributes that present the content, e.g., movie, video, size, and length, \( w_i \) is the weight for each attribute based on its significance, and \( B(i, m) \) is a similarity function returning 1 if \( v_i^1 = v_i^2 \) and 0 otherwise. Using the similarity value obtained from (9) based on the requested content and available content and also based on the significance of the data, the producer can determine whether the Interest was satisfied or not. The detailed operational procedure for seamless data delivery of a provider UE in the content-centric LTE network is described below. The message flow regarding seamless data delivery of the provider UE is shown in Figure 5.

(i) Step 1: UE (provider) establishes a connection to eNodeB1 and registers its content name to eNodeB1.

Figure 4: Separate buffer to store and forward data.
(ii) Step 2: eNodeB1 sends the UE the details, e.g., UE identifier, interface identifier to MME.

(iii) Step 3. UE (consumer) sends Interest message via eNodeB2. When eNodeB2 receives the Interest, it follows the same procedure performed by a CCN node.

(iv) Step 4: eNodeB1 receives the Interest packet from eNodeB2 and follows the same procedure performed by a CCN node. After doing look-up on its CS and PIT, it forwards the Interest to the mobile content source.

(v) Step 5: after forwarding the Interest packet to the content source, it adds PIT entry to forward data in the future.

(vi) Step 6: when the content source receives the Interest packet, it looks up its CS. When the matching content is found, UE replies back with the Data packet as a response through the arrival interface of the Interest packet.

(vii) Step 7, 8: eNodeB1 and eNodeB2 forward the Data packet to the UE.

(viii) Step 9: during the ongoing content transfer, the eNodeB1 estimates the mobility prediction using (3) and decides whether UE will move from eNodeB1 or not. If the serving eNodeB1 finds any best candidate for content transmission, it uses the same physical handover operation like LTE and triggers a connection to eNodeB3. UE establishes connection to eNodeB3.

(ix) Step 10: eNodeB3 sends the UE details, e.g., UE identifier, interface identifier to MME.

(x) Step 11: Since the UE's identifier is already registered in the MME, MME identifies that the UE is moved from eNodeB1 to eNodeB3. MME sends the Interest packet to eNodeB1 and eNodeB2 to reconfigure their path.

(xi) Steps 12, 13: eNodeB1 forwards the Interest packet to eNodeB3. When eNodeB3 receives the Interest, it follows the same procedure performed by a CCN node. For content matching, it can use the similarity matching equation (9). After doing look-up on its CS and PIT, it forwards the Interest to the mobile content source.

(xii) Step 13: UE (Content Producer) receives the old named Interest packet; it uses the similarity matching equation (9) to satisfy the Interest that is matched with any appropriate content.

(xiii) Steps 14, 15: after receiving the Interest from eNodeB3, eNodeB2 sends the Data packet as a response to the Interest packet to eNodeB1.

4. Performance Evaluation

This section presents simulation results in order to demonstrate that the content-centric LTE network is well suited to
today's communication trend. This section also analyzes the performance of the content-centric LTE network with our proposed mobility management scheme and compares it with the prediction-based LTE [24] and mobility management for CCN [21].

The simulations were performed using CCNx, LENA/NS-3, Direct Code Execution (DCE) on VMware, and Ubuntu 12.04 environment. We used a simulation topology as shown in Figure 6. The size of the content which is transferred between the mobile producer and mobile consumer is 1.1 Mbytes. The number of eNodeBs was considered to be three to show the mobility scenario of the UEs. The number of UEs covered by each eNodeB varies from 1 to 10 to show the content delivery efficiency in the low load and high load environment. The UEs may work as a content provider or a content consumer.

We showed the performance for different number of content providers who publish the video files and different number of consumers. Each content provider publishes different content files after 1-minute interval in the whole simulation time, and contents are requested randomly from the different consumers. We selected a half of all UEs to work as consumers and the other half as producers at each eNodeB. To show the efficiency of the proposed mechanisms and also to show the content delivery efficiency in the low load and high load environment. The UEs may work as a content provider or a content consumer.

We showed the performance for different number of content providers who publish the video files and different number of consumers. Each content provider publishes different content files after 1-minute interval in the whole simulation time, and contents are requested randomly from the different consumers. We selected a half of all UEs to work as consumers and the other half as producers at each eNodeB. To show the efficiency of the proposed mechanisms and also to show the content delivery efficiency in the low load and high load environment. The UEs may work as a content provider or a content consumer.

We evaluated the efficiency of the proposed approach and applicability by showing the average content transfer time, the average throughput observed by each consumer over time, and the average content delivery success ratio as performance parameters. We measured the average content transfer time $T$ using the following formula:

$$T = \frac{\sum_{i=1}^{n} T_{i,f} - T_{i,s}}{n}$$

where $n$ is the total number of UEs which are involved in receiving the content, $T_{i,s}$ is the time at which the UE makes a request to retrieve the content, and $T_{i,f}$ is the time at which the UE receives the requested content. Average throughput is measured as the average number of data bytes received by all the consumer UEs per second. Data transmission success ratio is the ratio of the total number of data packets received...
4.1. Content Transfer Time. Figure 7 shows the average content transfer time observed by UEs. The average content transfer time of our proposed CCN-based mobility management mechanism has been evaluated and compared with the prediction-based LTE and CCN-based mobility management mechanism in the LTE network.

The content transfer time of our proposed approach is shorter than the content transfer time of others because of its efficient soft handover based mobility management mechanism for both consumers and producers. The introduction of the extra buffer reduces the chance of long transmission delay, queuing delay, propagation delay, and processing delay at intermediate nodes in case of high mobility scenarios. In case of high mobility cases, our proposed mechanism uses the make-before-break approach when changing the routing path if needed. To avoid the high cost of tunnel setup, we use the same approach in case of consumer mobility. The main reason is that a content is retrieved in the candidate eNodeB before the original handover occurred. In the proposed mechanism, a node acquires the content from the edge network whereas in the other content communication a node acquires the content from a remote network. Therefore, the routing path, cost, and latency in the proposed mechanism are smaller.

4.2. Average Throughput. Figure 8 depicts the average throughput of the prediction-based LTE, CCN-based mobility management mechanism, and our proposed mechanism with the varying number of producers and consumers in random mobility scenarios. The throughput of our proposed mechanism is better and stable as it is able to detect and differentiate losses due to congestion, link failure, and mobility. Our in-network buffering capability does not affect the normal operation of the network in case of high mobility, so the throughput rate is always consistent in case of our approach. However, the content delivery rates are hardly affected by mobility and tend to be stable, as shown in Figure 8.

4.3. Data Transmission Success Ratio. We also measured the performance of reachability and continuity of our mechanism in terms of data transmission success ratio, which implies how much data were received correctly by the consumer in the random mobility scenario. The proposed content similarity approach increases the content availability when mobility changes the content location. Also the buffering capability, fast path switch, and handover prediction reduce the packet loss rate. The simulation result showed a significant improvement in this case as illustrated in Figure 9.

5. Conclusion

In this paper, we proposed a novel content delivery mechanism and a mobility management scheme for the evolved communication architecture such as 4G/5G to make the balance between the content diversity and network diversity. We then analyzed the performance of the proposed schemes with the LTE network in the mobile environment. By presenting different simulation results, we showed that the proposed schemes can be used as a possible solution for faster content transmission and seamless content delivery in the mobile environment. It is possible to provide accelerated, reliable, resource-efficient, and cost-effective communication, which will also be helpful for 5G.
Data Availability

The simulation parameters and results and other relevant data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE) and Korea Institute for Advancement of Technology (KIAT) through the International Cooperative R&D program. This work was supported by Hankuk University of Foreign Studies Research Fund of 2018.

References


