Decentralized and Autonomous Content Overlay Networking (DACON) with WiFi Access Points

Diko Ko  
Computer Science and Engineering  
Seoul National University  
diko@mmlab.snu.ac.kr

Kideok Cho  
Computer Science and Engineering  
Seoul National University  
kdcho@mmlab.snu.ac.kr

Munyoung Lee  
Computer Science and Engineering  
Seoul National University  
mylee@mmlab.snu.ac.kr

Hyunchul Kim  
Computer Science and Engineering  
Seoul National University  
hyunchulk@gmail.com

Ted “Taekyoung” Kwon  
Computer Science and Engineering  
Seoul National University  
tkkwon@snu.ac.kr

Yanghee Choi  
Computer Science and Engineering  
Seoul National University  
yhchoi@snu.ac.kr

ABSTRACT
Accessing contents from mobile devices becomes more and more proliferated and hence the need for the content distribution in the pervasive environment is growing. However, distributing contents in such environments taxes wireless network operators substantially. To provide the content distribution service in a reasonable cost, we pay attention to user-deployable WiFi access points (APs). In this paper, we propose a decentralized and autonomous content overlay networking (DACON) architecture for the pervasive content distribution services, which is the overlay network architecture comprised of public WiFi APs. We identify and answer major challenges in realizing the content distribution service in the pervasive environment by exploiting the overlay network of public WiFi APs.

Categories and Subject Descriptors
C.2.1 [Computer-Communication Networks]: Network Architecture and Design

General Terms
Design

Keywords
content-centric networks, pervasive computing, peer-to-peer

1. INTRODUCTION
Recently, several properties of user-deployable IEEE 802.11 WiFi access points (APs) poise themselves as a viable component of the network infrastructure. First, WiFi-capable user devices have been proliferating, and the coverage of WiFi APs has become so pervasive that APs could form an alternative urban network infrastructure [2]. Second, APs with ample storage (e.g. Apple’s Time Capsule1) are already available in the market, typically with a few to several hundred gigabytes of disk space. Third, unlike usual end-user devices (e.g. PC, mobile device), APs are hardly powered off and normally operate at a fixed location, which means that it has the potential to be a stable2 component of the network infrastructure; for example, some location services are already using the location information of WiFi APs to provide accurate users’ location information both in the indoor and outdoor environments [19]. Finally, the WiFi AP can be the last-mile middlebox [6] between the user and the network service provider. That is, APs can be controlled by the network operator while it is in a user’s domain, which means that it can be the point where the interest, policy and business between the user and the network operator can be established.

In the same context, Yochai Benkler argued that “unlicensed wireless approaches now offer the greatest promise to deliver a common physical infrastructure of first and last resort, owned by its users, shared as a commons, and offering no entity a bottleneck from which to control who gets to say what to whom. [4]” There have also been some approaches that use WiFi technologies to construct a network infrastructure. Google WiFi networks is the freely available outdoor wireless Internet service deployed in Mountain View, California, consisting of 500 Tropos MetroMesh pole-top access points [2]. Fon3 and Whisher4 are two representative examples of allied WiFi networks to provide the registered users with the free pervasive WiFi Internet access.

In this paper, we propose a novel service and network architecture, Decentralized and Autonomous Content Overlay Networking (DACON). This is an experimental approach to explore a new way of network organization and operation, in which we exploit the potential of WiFi APs. DACON

1http://www.apple.com/timecapsule/
2‘stable’ means that we can always use its functionalities at the expected geographical location.
3http://www.fon.com
4http://www.whisher.com
is a self-organized overlay network for the pervasive content distribution service. We define the pervasive content distribution service as the service that distributes contents (e.g., audio, video, etc) to users' mobile devices at anytime, anywhere. While the demand for the pervasive content distribution service is growing [10] [11] [7], the increasing deployment cost is the practical obstacle to wireless network operators. The most straightforward way to distribute contents in the pervasive environment would be via the pervasive wireless Internet access; however, it means that a large amount of contents should be delivered through the provider’s own wireless network. Hence, the provider should make a huge investment to deploy its infrastructure to keep trace with ever increasing user demands. Due to this reason, sometimes wireless network providers simply prohibit users from download bandwidth-intensive contents through their wireless network [8] [14].

We propose DACON as an alternative solution to provide the pervasive content distribution service by leveraging public storage-equipped WiFi APs of voluntary participants. We anticipate that DACONs will be most attractive in crowded commercial areas such as shopping malls, airports and train terminals, where many store managers already open their WiFi APs for their customers' convenience. Note that these areas also would be the places where pervasive services are highly needed.

While designing and implementing DACON, we confront several challenges to distribute contents in the pervasive environment by leveraging public WiFi APs. To address these challenges, we first define the service scenarios; we then design the service architecture and its components in Section 2. Section 3 answers the question of “how to form the local overlay network of APs?” Because APs in a DACON are personal properties of participants, we cannot expect any dominant owner or service provider that establishes and manages scattered DACONs, so that we cannot rely on any central or pre-deployed infrastructure component. Hence, the system should be designed in a fully decentralized and autonomous manner. The last challenge is the motivation of participants. Because a DACON is constructed by participants' APs, the incentives are very important in the success of DACON. We will discuss possible incentives for an AP owner to be a DACON participant and the related security issue about the content license in Section 4. Section 5 shows some preliminary simulation results. Lastly, the conclusion and the future plan will be given in Section 6.

2. DACON SERVICE DESIGN

2.1 DACON Service Architecture

DACON is an overlay networking architecture to provide the content distribution service in the pervasive environment. Figure 1 illustrates the overall DACON architecture. DACON comprises one or more participating users' APs (or DACON APs) and we assume that APs have some sufficient amount of storage; thus contents can be cached in each AP’s storage. If a user wishes to search and retrieve contents in the pervasive environment, the user’s end-user device (e.g., smartphones, netbooks) first makes an association with one of DACON APs in proximity; we call the AP as the host gateway (of the user). Then the content searching and distribution is performed within the DACON overlay network. The user requests and retrieves contents only through his/her host gateway.

DACON is constructed as having the multiple and hierarchical structure for the scalability and flexibility in the content searching process, which will be detailed in Section 2.2. Overlay links between DACON APs are called peer links; the peer link is only used in content searching, while other types of overlay links (i.e. internal link and external link) are used in overlay formation, metadata management and content searching. Some of DACON APs may also serve as DACON portals. A DACON portal (or portal, for short) is a root of a hierarchical tree whose members are DACON APs and one DACON AP can belong to only one portal. We call the overlay links between tree members (and between portals in the same DACON) as internal links. A DACON portal is similar to the supernode in KaZaa [20] in the sense that the portal maintains the information (or metadata) about contents stored in all the APs in its subtree. Hence, all metadata of a DACON AP should be reported to its portal through the internal links. Accordingly, there is an upper limit number of APs which can belong to a portal due to the limitation of the portal’s storage and processing power; therefore, there can be multiple portals in a single DACON. The portals in the same DACON have the full mesh connectivity among themselves via internal links. Some of DACON portals may also serve as DACON bor-

---

5DACON has two meanings: (1) the proposed service/network architecture, or (2) an actual overlay network made up of multiple participating APs. In the latter usage, there will be multiple DACONs, which will be explained later.

8So, the topology among the portals is similar to that of internal BGP.
der portals. A DACON border portal (or a border portal, for short) provides the external connectivity between multiple DACONs and we call the links between border portals as external links. The role of providing inter-DACON connectivity is similar to that of external BGP but the border portal is used for content searching and distribution.

2.2 Search in DACON

The search process locates the requested content in the DACON in a flexible fashion. We adopt the similar search process to Minerva [3] [13]. To trigger a search, an end-user device gives query terms (or keywords) to its host gateway. The host gateway then performs search in collaboration with other APs in the DACON. DACON performs the search process in the incremental manner, as opposed to many other distributed search engines which always inquire the same set of peer nodes. That is, the user in DACON flexibly decides how broadly the query will be distributed in a single DACON or across multiple DACONs. We name this hierarchical search strategy Incremental Content Search (ICS).

In ICS, an end-user device incrementally selects one of 4 search types: from Type 1 to Type 4. Figure 2 shows how broadly a query can be propagated depending on query types. Usually a user starts searching by sending a Type 1 query. The Type 1 query is performed at the host gateway and its peers. If the user is not satisfied with the query results, he/she can choose the higher query type. When the query type is Type 2, the query is forwarded from the host gateway to its portal, so that the portal searches contents of all its sub-tree. The Type 3 query is performed by all the portals in the same DACON and the Type 4 query is done by all the portals in the connected external DACONs. Overall, as the query type increases, the search quality will be enhanced at the cost of time and traffic overhead. Receiving the query results, the user will select one of contents in the results. Then, the user’s host gateway will request the corresponding AP to transfer the content and it is transferred to the end-user device via its host gateway. Optionally, APs can cache the content into its local storage for future requests.

In DACON, we assume that we cannot use the wireless ad-hoc networking to figure out the information of other APs because the wireless coverage of the consumer-level WiFi AP cannot exceed tens of meters in indoor environment [16]. Thus the messaging between DACON APs should be done through the backhaul (Internet) overlay link. When a participant initiates (turns on the power of) a DACON AP, it has no knowledge about other DACON APs. The first task is the bootstrapping: to find out other DACON APs to construct the overlay network. In Gnutella [17] or Chord [21], a peer asks to a central server or other pre-known peers about the rest of peers. However, in DACON, the new-coming AP does not have any knowledge of other peers and we cannot assume any global component as we mentioned in Section 1. Another alternative solution is to exploit flooding through the backhaul link with a time-to-live (TTL) limit. However, it might be impractical to figure out a suitable TTL value; if two geographically close APs are connected to Internet through different ISPs, then the required TTL should be large. Setting a high TTL value will incur the substantial traffic overhead to locate other APs. Furthermore, flooding packets may be filtered out in crossing ISP boundary.

Thus, we designed the bootstrapping of DACON in that the DACON AP exploits end-users’ mobility in obtaining the information about other APs. As already stated, APs are usually stable and have fixed locations while users are walking around especially in crowded commercial areas. Thus we expect the user mobility is suitable as a means of propagating the information of DACON APs in the pervasive environment. If a user wishes to use the DACON service, he/she should first be associated with a host gateway. Suppose an AP, AP1, initiates a DACON by itself; AP1 also serves as a portal of the sole DACON. Assume that a user who was in the coverage of another AP (say, AP2) now comes into the coverage of AP1; then the user learns the existence of AP1 and makes new association with AP1. During that process, the user informs AP1 of AP2’s contact point (e.g. IP address). Then AP1 can contact AP2 over its backhaul link to expand its DACON. If AP2 is behind a NAT, AP2’s STUN [12] server information should be reported.

The above scenario illustrates the beginning of DACON formation/expansion process. We designed DACON in that DACON APs or portals exchange messages to decide relations among them as the DACON grows, which we call the User-assisted DACON Formation (UDF) process. To accommodate APs to structure multiple DACONs hierarchically, we introduce four relations between two APs: peering, joining, uniting and allying. Every decision about the relationship is made by each AP autonomously according to some metrics (e.g. RTT, maximum number of peers/portals, geo/network distance between APs) when an AP knows of another AP.

3.2 Peering Procedure

Figure 3(a) shows the peering procedure when an AP tries to connect to another AP as a peer. Here, a peer is a neighboring AP in the DACON overlay; so there is a one hop peer link between peers. Peers cooperate in the Type 1 search

Figure 2: Search types in Incremental Content Search in DACON
Figure 3: Procedures of Four Relationships in the UDF Process

process. The peering relation is the most basic relation, and each AP wishes to make the peering relation by default when it learns of other APs. So, when AP A learns of AP B, it sends a Peering_Query message, and then AP B replies with a Peering_Reply message. When a peering relation is set between two APs, we make the portals of both APs also know each other; because the portals might make a uniting or allying relation afterward, which will be explained in following sections. For that reason, AP A should forward a Peer_Report message to inform its portal of AP B after processing the Peering_Reply message. Then Portal A sends AP B a Portal_Report message, which is in turn forwarded to Portal B.

The peering relation is not necessarily symmetric, which means that AP A sets up a peer relation to AP B, while AP B may have no relation with AP A. One of the reason is that a DACON AP has the limit on the maximum number of peers (MNP) due to the limitation of the AP’s storage and processing power. If the peer list is full (i.e. the number of peers is equal to the MNP) and the AP wants to have another peer, it has to drop one of existing peers. In doing so, it is necessary for the AP to evaluate existing peers and the candidate by some measures. Parreira et al. have proposed a peer selection mechanism called p2pDating [15] to establish a semantic overlay network [9] [1] by considering three measures: each peer’s behavioral history, metadata similarity between peers, and the overlap of the content inventories between peers. We choose the last criteria for simplicity.

3.3 Joining Procedure
Joining is a procedure that an AP detaches from the current portal and attaches to another portal as shown in Figure 3(b). For example, an AP (here, AP A) that has not been satisfied with the search results with the current portal may wish to move to a new portal9; here, Portal B. Thus, a soliciting AP (AP A) wishes to become a child of the responding AP (AP B) to belong to the new portal (Portal B). In this case, the responder (AP B) becomes the next hop to portal (NHP) of the solicitor (AP A). Eventually, Portal B may grant the joining request from AP A (or not).

3.4 Uniting Procedure
Unlike the two above procedures, the uniting relation is made only between two portals (not between APs). When two portals reach a conclusion that they can merge into a single DACON while they still want to serve as portals (not to be a child of another), the uniting procedure starts; otherwise, the joining procedure is triggered. Assume that there are two portals affiliated with different DACONs. As shown in Figure 3(c), Portal A sends a Uniting_Request message to Portal B. The message contains a local portal list (LPL) of the DACON to which Portal A belongs. Since there is also the maximum number of allowable portals in a DACON, the uniting request will not be granted if the LPL size exceeds the threshold. If the uniting request is granted, the LPLs of both portals will be exchanged and added to each, so that full mesh connectivity among portals is still maintained.

3.5 Allying Procedure
Allying is the procedure by which a portal in a DACON makes an external link with another portal in a different (or external) DACON. If any of these portals has not been a border portal, then it will be a border portal by setting up the allying relation. Suppose Portal A in DACON D110 learns

9 This can be done by the owner of the AP or by some automatic algorithms. Each owner of the AP can establish his/her own policy and it is not the part of the UDF process.

10 There are no unique identifiers for DACONs; D1 and D2 are used for explanation purpose only.
of another portal (say, Portal B in DACON D2) and Portal A decides to make an external link with Portal B by its own policy; e.g. round trip time (RTT) is too long to be the same DACON. In Figure 3(d), Portal A first checks if there is any other portal in DACON D1 that has already made an external link with Portal B by distributing Query_Alliance messages to all the other portals in its LPL. Portals reply with a Reply_Alliance message if they already have an external link with Portal B. If there is no reply, the solicitor (Portal A) sends an Alllying_Request message to the responder (Portal B). Then the same process is performed at Portal B.

4. CONTENT-BASED INCENTIVE SYSTEM

We expect there are at least three kinds of financial incentives to DACON participants: (1) isolated marketing, (2) local networked marketing, and (3) distribution commission. In other words, three kinds of profitable contents can be delivered to DACON users from the standpoint of the DACON AP (or its owner). The first incentive is accrued to the AP owner by advertising his/her business information to users in the wireless coverage. The second one is realized by syndicating the advertising information from retailers in proximity; e.g. Pizza Hut and McDonald stores in the same shopping mall can merge their advertisement information. The third incentive is location-independent and is obtained when APs mediate copyrighted contents to users. While first two incentives will be attractive to retailers or storekeepers, the third financial advantage can accrue to any participant of DACON. The last kind of incentive will leverage the "affiliate marketing" business model\[11\]. There are many variations in affiliate marketing; we will take the "pay-per-sale" example for simplicity. In order to motivate the DACON participant to distribute/mediate commercial contents to users, the original publisher (or owner) of contents might share the revenue with assisting DACON participants. In some sense, this is similar to Amazon's Kindle business. Amazon (the content provider) pays some lump sum cost to Sprint (the wireless service provider) for contents (i.e. e-book) distribution over the Sprint's cellular network.

To date, almost every content provider uses its own, often proprietary, digital right management (DRM) technologies to prevent people from illegally accessing digital contents. We assume that the data-centric security [22] is inherent in DACON, which means that the DRM-related data should be embedded in the metadata of the corresponding content: not in the content itself. In this paper we focus on the aspects of searching and transferring contents by skipping security issues except the following one; how can we securely retrieve and manage users' license information in DACON environments? Because DACON consists of public APs, it is very susceptible to security attacks. To cope with the security issues about users' licenses, we adopt a tunnel-based scheme, which was originally proposed by Sastry et al [18].

In DACON, a user's license information is managed by his/her home gateway, which is the user's credible and representative network component. We expect the home gateway is usually the DACON AP whose owner is the user and the contact point of the home gateway can be pre-written to the end-user device. When the user needs to retrieve content licenses in the DACON environment, he/she set up a virtual private network (VPN) tunnel from the (currently serving) host gateway to his/her home gateway. License purchase and payment also can be done through the VPN tunnel. In this way, most of security and legal issues could be handled because every confidential information of the user is dealt with as if it is in his/her own home gateway.

Figure 4 illustrates entities for the simplified “pay-per-sale” scenario. For example, end-user U locates the requested content whose metadata include the contact points (e.g. Web URI) of the payment and the license servers of the content (P and L, respectively). The payment is made between U and P via U’s home gateway. The channel between U and P will be allowed by the host gateway. Note that during the transaction, U should inform P that the intermediary is the host gateway, like HTTP Referrer. P informs L that the license of the purchased content should be given to U. U now requests the license from L via the home gateway. U asks the host gateway to forward the license of the purchased content from the DACON. Finally, U can access the content with the key in the license from L.

5. SIMULATION RESULTS

In this section, we introduce some preliminary simulation results. We use QualNet 3.9.5\[12\] to simulate the user mobility and the overlay connectivity among DACON APs and portals. Note that each DACON AP has an IEEE 802.11b interface for passersby and a fixed Internet link for the overlay connectivity with other APs. According to the 802.11b experiment in indoor multipath environments [16], the transmission range of a DACON host is set to 29 m at 11 Mbps bit rate for conservative simulation settings. As for the DACON AP placement, we use the map of Gilroy outlet in Gilroy, CA 95020. The outlet spans 200 x 200 m$^2$. The same outlet is replicated at four corners of an 800 x 800 m$^2$ square. We place 50 DACON APs in the four outlets in an almost equi-distant fashion along the store layout.

In this simulation, we check the practicality of leveraging users’ mobility for the DACON formation. We used the random waypoint mobility [5] model for users’ mobility. In the DACON formation process, we observe how DACONs have formed over time. The performance metric is the ratio of the number of the APs attached to the largest DACON (over both internal and external links) to the total number of the deployed APs. Figure 5 shows how this ratio increases as time goes by. As the number of moving DACON users increases, the number of the APs attached to the largest DACON increases sharply at the beginning. We could find that more than 40 DACON users are enough to construct the overlay network of 98% of APs within 30 minutes. We also increase the number of the previous APs whose information is handed over to a new AP when the end-user device is associated with the new AP. Figure 6 shows the DACON formation with various number of previous hop information. Although we suppose there are only 5 DACON users, we find that the knowledge of the two previous APs is sufficiently effective because it can form a DACON consisting of 95% of APs in 30 minutes. Delivering the information of four or

\[11\]http://en.wikipedia.org/wiki/Affiliate_marketing

\[12\]http://www.qualnet.com/
eight APs to a new AP is not so efficient as the case of two previous APs.

6. CONCLUSIONS
In this paper, we propose DACON as an experimental trial for a new kind of infrastructure leveraging user-deployable WiFi APs. DACON is designed to provide contents in pervasive environments. The main idea of DACON is to make individual AP owners participate in establishing a novel local overlay network for the content distribution. DACON is fully decentralized in the sense that no central authority or global component is required for network initialization and operation. Users can search for contents over DACON without incurring the wireless Internet access. A DACON participant’s AP can easily join a DACON just by powering up and it learns of the existence of other APs by exploiting DACON users’ mobility; then the AP can autonomously decide how it will network with other APs. For autonomous network operations, we devise four relations: peering, joining, uniting, and allying. Using these relations between APs, multiple DACONs will grow in a flexible fashion, each of which may have hierarchical structure. They allow users to search for contents at various scales.

We implemented UDF and the content distribution mechanism of DACON with Python 2.5.2 on Linux and Windows XP. Our prototype DACON AP is equipped with an IEEE 802.11b/g interface for end-user connectivity and an Ethernet interface for the overlay connectivity. We plan to deploy 10 DACON APs at a crowded campus building to observe users’ behaviors, and analyze the relation between content locality and content search requests. We believe that it will give us more insights about new content-oriented services in the pervasive environment.

7. ACKNOWLEDGMENTS
This work was supported by NAP of Korea Research Council of Fundamental Science and Technology. This publication is based on work performed in the framework of the Project COAST-ICT-248036, which is partially funded by the European Community. The ICT at Seoul National University provided research facilities.
8. REFERENCES


