1

Information Dissemination in Mobile CDNs

Nicholas Loulloudes, George Pallis, and Marios D. Dikaiakos

1 Introduction

With the recent development of technologies in wireless access and mobile devices, the mobile network has become a key component of today's Internet vision [1, 46]. Current mobile networks, which are being deployed worldwide, enable mobility features to new applications and also extend existing wired Web applications to mobile terminals. The mobile wireless network offers a rich assortment of dynamic and interactive services, such as GPS navigation information, mobile TV, vehicular traffic information, and location-oriented services. The provision of such services requires techniques to disseminate data as efficiently as possible in order to minimize the total network traffic and to improve the mean response time to mobile users.

In the wired Web, network performance can be substantially enhanced by using additional bandwidth, which is often available at low cost. However, this approach is impractical for mobile wireless network infrastructures. Most of these networks have fixed spectrum and achievable data rate is fundamentally limited by interference [46]. This problem is likely to get more serious when more mobile users start using bandwidth-intensive services such as streaming media. In this context, caching and prefetching might be a solution. Specifically, these approaches have been extensively used in the wired Web to optimize the amount of bandwidth consumption by shifting the traffic away from overloaded content providers and closer to the content customers [43]. Although these methods offer several benefits (i.e. conservation of network resources and reduced latency), the dissemination of dynamic content and resource-hungry applications (e.g. multimedia applications) remain a challenge.

Nicholas Loulloudes, George Pallis, and Marios D. Dikaiakos
Department of Computer Science, University of Cyprus, Cyprus,
e-mail: {loulloudes.n, gpallis, mdd}@cs.ucy.ac.cy

Content Delivery Networks (CDNs) promise to address these challenges by moving the content to the "edge" of the Internet, and thus closer to the end-user [45]. An introduction to CDNs can be found in the first chapters of this book. Although there has been much work on wired CDNs [27, 32, 40], content dissemination on mobile environments has received little attention so far [1, 49]. This is due to the limited Internet access capabilities of most mobile terminals in the recent past. However, this situation seems to be changing with the advent of innovative cellular (e.g. 3G) and wireless (e.g. WiFi) services which allow mobile terminals to access Internet and other data services at speeds comparable to traditional wired access [46]. Previous research [6, 48] shows that cooperative caching in mobile environment improves the network performance and information dissemination. In this context, we believe that the infrastructure of CDNs may provide a scalable and cost-effective mechanism for accelerating the information dissemination in the mobile wireless environment [45]. However, the mobile wireless network infrastructure represents a fundamentally different information medium from the traditional Web in terms of access devices used, content availability, bandwidth, and cost to the end-user. Thus, the typical CDNs cannot be enhanced by mobile wireless networks, since CDN architecture does not take the distinguished characteristics of these networks into account. In this context, we define *mobile CDNs as overlay networks of surrogate servers which deliver content in the mobile wireless network infrastructures*. Specifically, CDNs may offer an exciting playground for exploiting the emerging technological advances of mobile computing.

The purpose of this chapter is to present the challenges and the current status of mobile CDNs, discuss the recent evolution of the mobile wireless networking infrastructure, as well as to investigate how information dissemination can be improved by the emerging mobile CDN practices.

The rest of this chapter is structured as follows. Section 2 presents the need for mobile CDNs. Section 3 introduces the mobile CDNs. Section 4 presents the wireless network infrastructure of mobile CDNs. Section 5 presents our vision about how the existing intermediaries in mobile environments can be adopted in mobile CDNs. Section 6 provides some implementation and experimentation perspectives for mobile CDNs. Section 7 presents the future research directions over these networks. Finally, Section 8 concludes the chapter.

2 Motivation

The mobile Internet, defined as wireless access to the digitized contents of the Internet via mobile devices, has significantly advanced in terms of user population. Recent studies have shown that in Japan the number of people using the mobile Internet already exceed those using the stationary Internet [4].

In general, the mobile Internet goes where the users go; users demand Web access when and where they need it, using their mobile devices.

Nowadays, an increasing number of content providers is investing in mobile Internet. The automotive industry has already introduced such mobile technologies in vehicles that provide accurate navigational and traffic aids (traffic conditions such as accidents, congestion, road constructions, diversions) to their drivers. Also, vehicles can be equipped with devices that alert their drivers for emergency situations (e.g. fire, earthquake damages, terrorist attack damages, etc) using multimedia data. Although in most occasions, a simple text message is sufficient, multimedia data, such as images and videos of an accident (or a dangerous situation further ahead), provide drivers with more precise and convincing information in order to take any necessary actions. Furthermore, the banking industry has identified business opportunities in mobile Internet including automated banking services. Nowadays, many mobile phone users readily access such services from their handsets. In addition, mobile Internet has opened new opportunities to the entertainment industry by selling online music, books, films, and games. For instance, travelers/commuters, who are waiting at terminals, can use their mobile devices (e.g. PSP) to play games or interact with their friends who are in other geographic locations.

Implementation of the above examples requires advances both in wireless network technologies and supporting devices, as well as the development of a scalable and resilient internetwork infrastructure that would support efficient information dissemination techniques to mobile users. From a technological viewpoint, the recent advances in wireless networking (e.g. 3G, GPRS, Dedicated Short Range Communications - DSRC[1]) guarantee that mobile devices can access Internet and other data services at speeds comparable to traditional wired access.

The infrastructure of CDNs [45] provides a scalable and cost-effective mechanism for accelerating information dissemination in the wired Web. However, the architecture of typical CDNs is inadequate to enhance the features of mobile Internet. since it does not take mobility into account. The variations in mobile user requests are caused not only by changes in content popularity but also by user mobility. Each user request is characterized by the requested content, the time of the request, and the location of the user. In order to support mobile users, surrogate servers should be located "close" to the base stations of the wireless network providers.

Furthermore, the CDNs architecture should be reconsidered in order to meet the new challenges of mobile users needs. One characteristic of mobile networks is the *scarcity of resources*; due to the small sizes of portable devices, there are implicit restrictions with respect to the availability of storage, computation capacity and energy. For instance, consider a father who shoots with a Wi-Fi camera a digital video of his three-year-old child while playing

[1] Dedicated Short Range Communications:
http://www.leearmstrong.com/DSRC/DSRCHomeset.htm

at the beach. His thought is to upload the video to a server in order to free up his camera memory and thus increase its capacity for more pictures/videos. So far, typical CDNs do not support the uploading of user content to the surrogate servers. In practice, the CDN distributor module is responsible to decide which content would be replicated by surrogate servers. In light of the above, surrogate servers in mobile CDNs should provide user-oriented services. This can be implemented by allocating a portion of their cache to be used by mobile users for direct content uploads.

Another parameter that mobile CDNs should take into account is the *navigational behavior of mobile users*. A significant factor which affects the users navigational behavior is the actual devices being employed. Mobile devices have limited input capabilities; for instance, the numeric keypads of mobile phones allow only "minimal" text entry compared to the keyboard entry on PCs. Moreover, mobile users must also contend with slow download times and incremental billing costs. Consequently, these characteristics have led to differences between the way users access information on the mobile Internet and the way they access information on the wired Web.

The mobility of users urges the development of state-of-the-art *geo-location oriented services* in mobile CDNs. Consider a mobile user who uses a CDN-supported application. While the mobile user is moving, the application should be replicated by another surrogate server so as to be always "close" to the mobile user. Moreover, geo-location services may also be used to detect mobile user Internet connection speed. This is crucial for Web site owners that would like to demonstrate multimedia applications (e.g. ads) to prospective customers.

In addition, typical CDNs do not provide any mechanism that monitors in real time the status of users (who interact with the CDN) and the underlying network infrastructure. However, such a *monitoring mechanism* can be considered as a key component in the support of content dissemination in mobile CDNs considering the inherent limitations of the underlying mobile wireless network infrastructure. These limitations are briefly explained below:

- *Frequent network disconnections*: The random or even organized mobility of users can severely influence their connectivity with the rest of the network. This is mainly due to: 1) small connection periods with base stations or other nearby mobile devices, 2) the presence of obstacles such as high buildings, trees or cars, which significantly degrade the wireless signal quality and 3) the possibility that users have temporarily gone out of radio coverage. The above factors can lead to bandwidth degradation or even total loss of connectivity, which can ultimately cause loss of information.
- *Network fragmentation(s)*: As a result of the frequent disconnections mentioned above, the mobile wireless network is vulnerable to fragmentation(s). These prevent end-to-end connectivity between mobile nodes and consequently minimize the availability of information.
- *Mobile nodes constraints*: The majority of mobile devices face significant constraints in terms of processing power, storage capacity and most im-

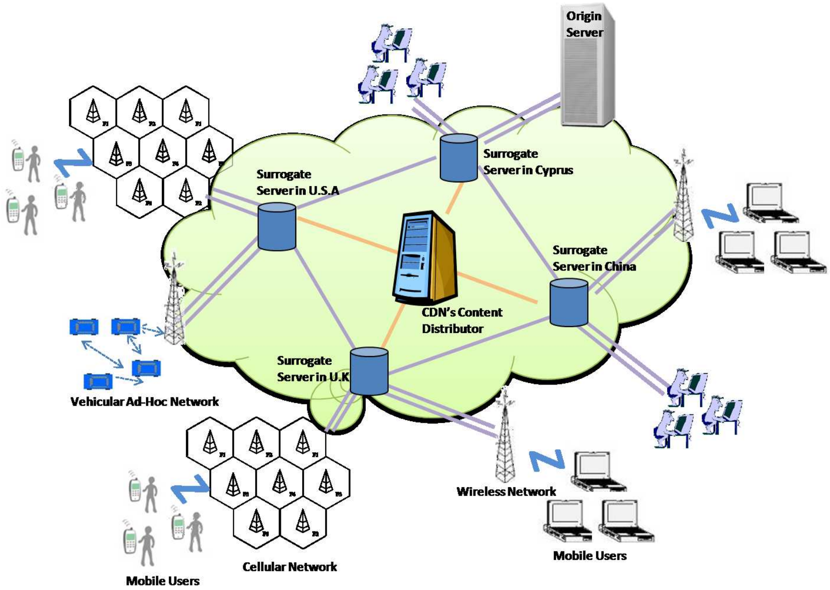

Fig. 1 A typical Mobile Content Delivery Network.

portantly uptime duration. These limitations are imposed primarily from the fact that such devices run on battery power and secondary due to their small size which prevents increased processing and storage capacity.

Due to the above limitations of mobile wireless networks, it is crucial for the CDN to know the status of mobile users so as to minimize the overall traffic in the network. For instance, consider a mobile user who requests to download a podcast to his/her MP3 player, but during the downloading process the device goes offline due to battery drain. In such a case, the CDN should be aware of the user status and block any further data transmission destined to him/her.

All the above issues conclude that disseminating information to mobile users in an efficient and cost-effective manner is a challenging problem, especially, under the increasing requirements emerging nowadays from a variety of applications (e.g. streaming media, dynamic content) and the inherent limitations of the mobile wireless environment. *Mobile CDN* infrastructure may meet these challenges. The next section presents our insight for mobile CDNs.

3 Mobile Content Delivery Networks

Contrary to wired CDNs, mobile CDNs are deployed within the range of a wireless network (e.g. cellular network, WiFi) and offer high quality services for delivering dynamic data and rich multimedia content to mobile devices. Specifically, the network infrastructure in mobile CDNs is de-composed in the two following components: a) the wired network infrastructure and b) the wireless network infrastructure. The former is an infrastructure responsible for the wired environment of the CDN; it provides the communication links which connect origin servers with surrogate servers and surrogate servers with network elements (e.g. switches, routers, 3G/GSM enabled base stations (BS), Wi-Fi enabled access points (AP)). On the other hand, the wireless network infrastructure is responsible for enabling communication and information dissemination among static and mobile users in the wireless environment of the mobile CDN. Therefore, the client-server communication is replaced by three communication flows: 1) between the client and the edge of the wireless network (AP or BS), 2) between the edge of the wireless network (BS or AP) and the surrogate server, and 3) between the surrogate server and the origin server. A typical mobile CDN is depicted in Fig. 1.

Considering that the surrogate servers should be placed “close” to BSs or APs of wireless networks, the network topology of mobile CDN should be redeployed so as to address this issue. Therefore, the placement of surrogate servers should be reconsidered so as to provide a sufficient coverage to the mobile wireless infrastructure. Due to the large amount of base stations, mobile CDNs should incorporate in their infrastructures more surrogate servers than a typical CDN.

Regarding the architecture of surrogate servers, their cache should be segmented into two parts in order to support user-oriented services to mobile users. The one part is devoted for replicating the content of origin servers (which have been contracted by a CDN provider), and the second part is dedicated to mobile users for content upload. Furthermore, surrogate servers of a mobile CDN should also provide geo-location oriented services to mobile users. The deployment of such services requires the definition of a service discovery protocol to select a surrogate server based on the location context of the mobile user (the location context refers to the current geographical area or position of a mobile user), in order to migrate applications to the surrogate server closest to the mobile user.

Apart from the networking issues involved in the establishment of the mobile CDN infrastructure, it is also important to determine which content outsourcing policy to follow (the content which would be replicated). Most CDN providers use either uncooperative or cooperative pull-based approaches [27]. The main characteristic of both approaches is that they are reactive; a data object is cached only when the user requests it and, consequently, these schemes impose large communication overhead (in terms of the number of messages exchanged) when the number of users is large. In addition, this mechanism does not offer high fidelity when the content changes rapidly or when the coherency requirements are stringent. Due to these limitations, the pull-based schemes are prohibitive in a mobile wireless environment. Mobile CDNs should enhance a cooperative push-based scheme. In contrast to the pull-based approaches (uncooperative and cooperative) which wait for the users to request information, the cooperative push-based approach lets the origin servers to proactively push the information into caches close to the mobile user, expecting a further reduction to the access latency. Indeed, Chen et al. [9] concluded that the cooperative push-based policy provides the best overall results, when compared to the other approaches. In such a scheme, the content is pushed (proactively) from the origin server to the surrogate servers. Upon a request, if the surrogate server has an object replica, it serves the request locally, otherwise, it forwards the request to the “closest” server that has the object replica and relays the response to the end-user. In case the requested object has not been replicated by some surrogate server, the request is served by the origin server. This scheme requires cooperation among surrogate servers which incurs extra communication and management costs for its implementation. However, these costs are amortized by the fact that surrogate servers efficiently share the available bandwidth among them and also by the reduction of replication redundancy. In turn, the latter diminishes cache consistency maintenance costs.

Table 1 presents the main differences between a typical CDN and a mobile-specific CDN.

Table 1 Typical CDN vs. Mobile CDN.

Features	**Typical CDN**	**Mobile CDN**
Content type	static; dynamic; streaming	static; dynamic; streaming
Users location	fixed	mobile
Surrogate servers location	fixed	fixed
Surrogate servers topology	“close” to Internet Service Providers”	“close” to Base Stations
Replicas maintenance cost	medium	high
Services	application services	geo-location oriented application services; user-oriented services
Content outsourcing policy	cooperative/uncooperative pull-based scheme	cooperative push-based scheme

Therefore, the architecture of a mobile CDN, should consist of the following components:

- A set of surrogate servers (distributed around the world) which cache the origin servers content; the surrogate servers are not mobile and are located close to mobile base stations,

- A network infrastructure (wired and mobile wireless) which is responsible to deliver content requests to the optimal location and optimal surrogate server,
- A mechanism which monitors the network infrastructure in real-time for available bandwidth, latency and other sources of congestion,
- A cache manager, which manages efficiently the content that has been replicated to surrogate servers,
- A content location manager, which manages the content locations and schedules data prefetching,
- An accounting mechanism which provides logs and information to origin servers and CDN providers.

The above components interact with one another in order to deliver the requested content to mobile users. The content represents a collection of objects, which may change over time. For instance, content may represent a set of objects of a particular Web site, a streaming service, a Web service or any distributed application. When a mobile user requests an object, the user request is sent to the currently "closest" surrogate server. If there is a cache hit (the surrogate server has an updated replica of the requested object), the request is served locally. Otherwise, the content location manager forwards the request to the closest surrogate server that has the requested object.

In response to the high variability in user demands and the dynamic nature of the requested content, mobile CDNs should integrate content management policies in surrogate server caches. There are different approaches related to which content to outsource [40], and which practice to use for outsourcing [19, 28, 44, 51]. The cache manager is responsible for the content which is replicated to surrogate servers as well as for keeping the surrogate server replicas up-to-date. Specifically, several issues related to cache consistency should also be considered [30]. Thus, the cache manager performs periodically checks in order to incorporate the spatial and temporal changes in user demand. On the other hand, the content location manager is responsible to replicate the outsourced content to surrogate servers. Finally, a mechanism monitors in real-time the user status since a user may be unavailable to receive the requested content due to its mobility.

Regarding the CDN market, although several CDN providers have been emerging in the market[2], there is a lack of mobile CDNs. According to the authors knowledge, Ortiva Wireless[3] is the only mobile CDN provider, which is dedicated to delivering video to mobile users under highly variable wireless network conditions. By optimizing the delivery of mobile TV, video, audio, and Web content, Ortiva disseminates high quality information across any wireless network, expanding revenue opportunities, coverage, and network capacity without costly infrastructure modifications.

---

[2] A complete record of the existing CDNs' providers can be found in [29]

[3] Ortiva Wireless: http://www.ortivawireless.com/

4 Wireless Network Infrastructures of Mobile CDNs

As we mentioned above, the network infrastructure in mobile CDNs is decomposed in the wired and the wireless network infrastructure. The wired infrastructure provides the resilience and fault-tolerance that CDNs require since a substantial part of it belongs to the Internet backbone which is provisioned with a high level of redundancy. In this section, we focus on the two variations of wireless infrastructure that are currently available [37] and discuss their suitability in mobile CDNs.

4.1 Mobile CDNs under Centralized Wireless Infrastructures

Cellular and Wi-Fi networks are two indicative network types of centralized wireless infrastructures. From Fig. 2 it can be observed that in such infrastructures all users communicate with a central authority. In the case of cellular networks this authority is usually a 3G and/or GSM enabled base station (BS), whereas in Wi-Fi networks this is an IEEE 802.11 enabled access point (AP). The BS/AP operates as a hardware bridge for all the wireless users [37] and is responsible for the allocation and control of the wireless channel. In addition, BSs and APs keep track of the changes in user mobility in order to prevent content dissemination to them when they are out of the radio coverage or offline.

The centralized wireless network infrastructure provides a good framework for mobile CDNs. For instance, consider user A in Fig. 2, who owns a cellular phone (registered with a cellular network) with an embedded video camera. This user is currently on vacation and makes a video recording of a sightseeing s/he is visiting. At the same time, s/he wants to upload this video to a personal, public online blog, so that users B and C belonging to a Wi-Fi network have access to it. As long as the user is within the radio coverage of the cellular network, the phone transmits the recorded video to the responsible BS of the area. In turn, the BS forwards the video through the cellular network to the Web server hosting the blog. The mobile CDN provider is then responsible to push this content as close as possible to the WiFi network in which users B and C belong to.

However, this centralization in the wireless network can pose several problems. The existence of a central authority, which orchestrates the communication in the wireless environment, limits the scalability and resilience in mobile environments [10]. Therefore, careful and systematic provision of such infrastructures is necessary so as to provide the required communications coverage and Quality of Service (QoS) of mobile users.

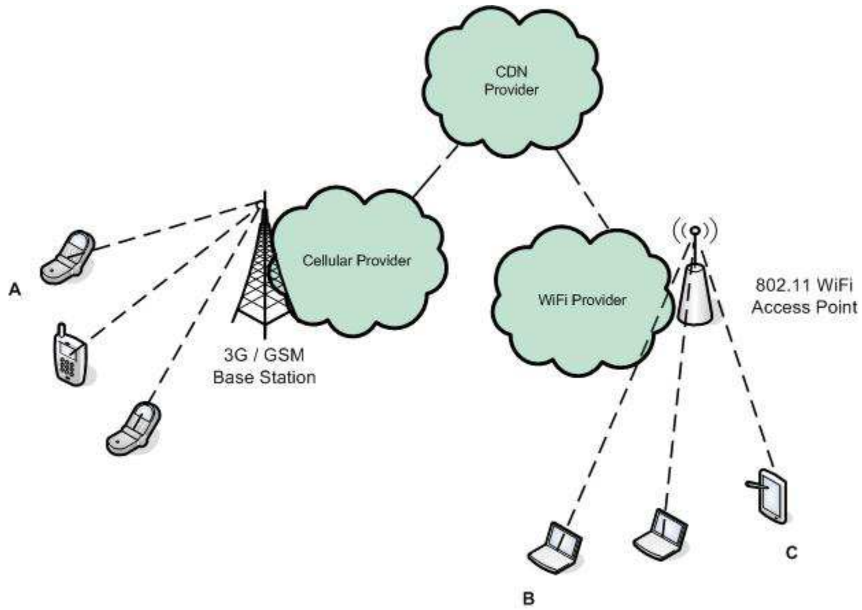

Fig. 2 A Centralized Wireless Network Infrastructure.

4.2 Mobile CDNs under Ad-Hoc Wireless Infrastructures

Ad-hoc wireless infrastructures inherently provide several advantages over centralized ones for mobile environments. From a technical point of view, ad-hoc wireless networks are comprised from autonomous nodes that intercommunicate through wireless ad-hoc links. Communication takes place in a multi-hop fashion without relying on any sort of authority that overlooks channel allocation, topology formation or content dissemination. In contrast to centralized infrastructure networks, they provide cost-effective solutions in terms of network extension when the number of users increases [38, 50]. In summary, the main differences between centralized wireless infrastructures and ad-hoc wireless is depicted in Table 2.

A Mobile Ad-Hoc Network (MANET) is a wireless network with a dynamic arbitrary topology constructed in an ad-hoc, peer-to-peer (P2P) fashion from mobile nodes for information retrieval and dissemination. The basic characteristics of MANETs are summarized as follows:

- *Ad-hoc and Peer-to-Peer (P2P) connectivity*: such networks have the ability to be self-organized and self-configurable. The ad-hoc infrastructure is decentralized and independent of any fixed infrastructure. This inherent characteristic of decentralization highlights the fault tolerance of such networks, as no single point of failure exists. Furthermore, ad-hoc connectivity

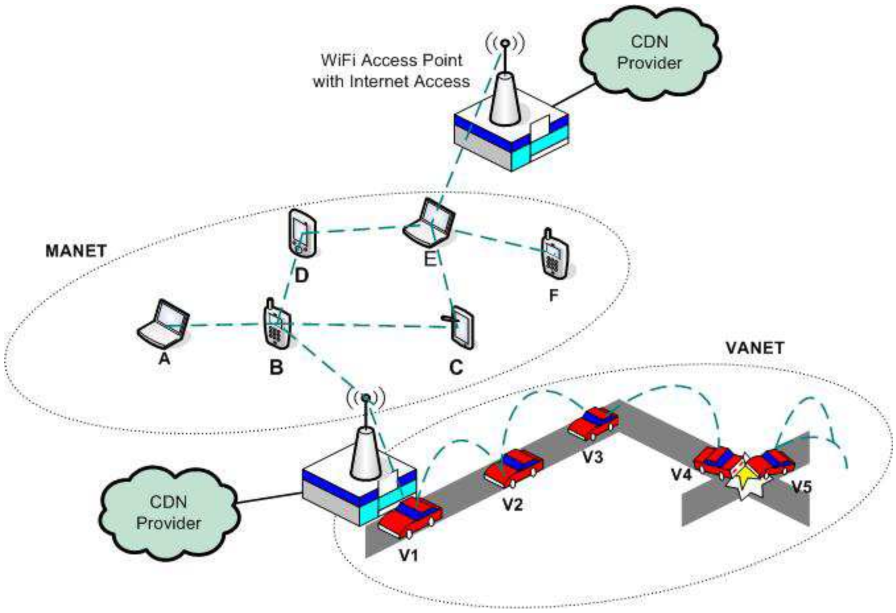

Fig. 3 A Typical Ad-Hoc Wireless Network Infrastructure.

provides network scalability since there is no need of extra infrastructure setup. Their P2P connectivity emerges from the ad-hoc connectivity of the infrastructure and dictates that each participating node in the network is treated as an equal peer. This results from the fact that at any given time, a peer can function both as host and router [21].

- *Network topology*: dynamic due to the unpredictable random mobility patterns of participating nodes. These conditions are supplemented by the ability of mobile nodes to join and leave the wireless network at any given point in time without prior notification. Fixed nodes can peer with mobile nodes, providing gateway functionality to other fixed infrastructures or the Internet.
- *Mobile node constraints*: mobile nodes rely on rechargeable batteries for their operation meaning that uptime duration is substantially limited. In addition, their processing and storage capacities are constraint from their relative small size.

MANETs gained significant attention from the research and industry communities as prominent wireless ad-hoc infrastructures for information dissemination in fast-changing and uncertain mobile environments [15, 42]. For instance, consider the following example which is depicted in Fig. 3. User A wants to download a video from an online multimedia store (registered with a mobile CDN) to his/her laptop. However, in his/her existing location there is no wireless coverage from a WiFi or cellular provider. Thus, s/he switches

Table 2 Centralized Wireless Infrastructures vs. Ad-Hoc Wireless Infrastructures.

Features	**Centralized Wireless Infrastructures**	**Ad-Hoc Wireless Infrastructures**
Fault-Tolerance	No - Access Point is a single point of failure.	Yes
Scalability when number of node increases	No	Yes
Self-Organization and Self-Configuration of Nodes	No	Yes
Maintenance and expansion costs	High; Extra Access Points are needed	Low
Central Control Authority	Access Point/Base Station	Does not exist
Bridging	Access Point or Base Station (hardware bridge)	Each node acts as a bridge for other nodes (software bridge)

to an ad-hoc mode and joins[4] an existing MANET comprised by users B,C,D and F. User E who is also part of the MANET provides Internet connectivity to the rest of the group. Hence, A's request will be forwarded through the MANET and E to the online multimedia store. The responsible mobile CDN will receive this request and will redirect it to a surrogate server which is "close" to E. Then, the aforementioned mobile CDN provides this video in a different encoding and compression so as to compensate for the inherent limitations of mobile nodes [2].

A subclass of MANETs are Vehicular Ad-hoc Networks (VANETs). These are formed among vehicles driven within road constraints and fixed-road side nodes that provide location services and Internet gateway functionality. Specifically VANETs were propelled by the need of implementing a suitable infrastructure for Intelligent Transportation Systems [2]. In general, VANETs are different from MANETs in terms of:

- *Network Topology*: even more dynamic than MANETs, due to the extremely high mobility of vehicles. As opposed to the latter, the movement of vehicular nodes in VANETs is quite predictable since most of the time vehicles are driven in a well organized fashion within the constraints of roads. As in MANETs, road-side nodes peer with vehicular nodes and act as gateway to other fixed infrastructures and the Internet. Also these road-side nodes broadcast location-oriented information to vehicles.
- *Node Constraints*: no constrains exist in terms of processing power and uptime duration since vehicular nodes are constantly fed with power from the vehicle's electrical installation.

[4] Security and how users trust each other are very crucial issues in wireless ad-hoc infrastructures, but these are out of scope in this chapter.

The following scenario depicts the benefits if mobile CDNs are utilized in VANETs. Consider that vehicles V1...V5 have informally agreed to take part in a VANET. Suppose that a traffic accident occurs between V4 and V5 at the junction. Instantly, V4 and V5 broadcast an emergency alert and possibly some cabin photos or short videos to all their neighbor vehicles informing them about the accident. This alert is marked with high priority and is propagated throughout the VANET. As soon as V1 receives the alert, it utilizes its Internet connection from the nearby base station to post this alert to an online traffic monitoring system. The mobile CDN will place this alert and any accompanied photos or videos to the most appropriate surrogate server in terms of the accident location. This enables emergency response teams to receive an overall view of how the accidents looks like before going on site with the minimum latency. Furthermore, far-away vehicles heading towards the accident location will be informed in advance through the aforementioned traffic monitoring system such that necessary adjustments in their course are made so as to avoid congestion build-up. In general, mobile CDNs can be utilized in vehicular environments for:

- the provision of mobile Internet to vehicles with extensive support in dynamic and streaming content.
- dissemination of road traffic-related information to interested vehicles and authorities such as congestion, accidents and diversions.
- location-oriented services such as location and availability of various facilities in an area of interest (i.e. hotel, restaurants, gas stations, parking places).
- dissemination of environmental measurements gathered from vehicular sensors.
- support for distributed gaming.

Finally, Table 3 represents the key differences between MANETs and VANETs.

Table 3 Mobile Ad-Hoc Networks vs. Vehicular Ad-Hoc Networks.

Features	**MANET**	**VANET**
Mobile Nodes	Laptops, Smartphones, PDAs	Vehicles
Node Movement	Random - Unpredictable	Organized - Predictable (within road constraints)
Node Constraints	Limited uptime due to battery constraints	Power is not an issue
Mobility	High	Very High
Network Topology	Dynamic	Dynamic
Supported Network Technologies	UMTS, GSM, WiFi, Bluetooth	DSRC, UMTS, GSM, WiFi

5 Visionary Thoughts for Practitioners

Several projects from the research and industry communities exist in the literature which address the development and deployment of intermediary components between origin servers and mobiles nodes for seamless information dissemination [12]. Such intermediaries are deployed in wireless infrastructures optimizing intermittent network connections and bandwidth utilization. This section presents the existing work in the utilization of intermediaries in mobile environments and our vision of how these can be adopted in mobile CDNs.

IBM's Web Express system follows the client/intercept/server wireless computational model [18] for the optimization of Web access from mobile resource-limited nodes. This model aims to minimize latency over wireless links by utilizing an intercept technique that allows applications to optimize communications. In order to address this issue, WebExpress uses two key components: the Client Side Intercept (CSI) deployed on mobile nodes and the Server Side Intercept (SSI) deployed on the wired network. The CSI intercepts the mobile node request and co-operate with SSI so as to reduce data transmissions and improve content availability over the wireless link. Several optimization methods are used by WebExpress to meet these challenges such as caching of content both at the client (CSI) and server (SSI) side, computation of difference between cached based objects and new responses, multiplex of all requests over a single session, and HTTP header reduction.

In our view, WebExpress and more specifically the client/intercept model can be adopted in mobile CDNs by moving CSI from origin servers to surrogate servers. This relocation has the added benefits of further *minimizing the communication latency* during mobile users Web access, and *reducing the network traffic* over wireless links due to the optimization methods mentioned above.

Dikaiakos et al. [13] propose an independent information transfer protocol for the dissemination of information among vehicles and fixed road-side nodes participating in a VANET. The *Vehicular Information Transfer Protocol (VITP)* is an application layer, stateless protocol, that specifies the syntax and semantics of messages exchanged between vehicular peers, while at the same time it pertains its independency from the underlying VANET protocols. The authors proposed the deployment of VITP on vehicles for the support of location-based, traffic-oriented, services to drivers. Queries in VITP are location-oriented. This means that queries are transmitted to a specific geographic area of interest in order to retrieve some desired content. Inside this destination (target) location, a dynamic collection of VITP peers, called Virtual Ad-Hoc Server (VAHS), is established on the fly in order to compute the reply to the location query. The VITP peers that participate in the VAHS co-operate to generate the final reply to the VITP request until a return condition is met. VITP queries follow a request-response methodology. In addition, the protocol includes support for data caching through cache-control headers which can be included in VITP queries. If there is a hit in a peer's cache for a VITP query, this peer can use the aforementioned headers to decide whether to respond based on its local cache or whether to retransmit the query towards to the target location.

In the context of mobile CDNs, fixed, road-side VITP peers can be used as surrogate servers. In contrast to vehicles, such peers have increased processing and storage capacity, higher wireless-radio coverage and continuous high-bandwidth Internet connection. Each road-side peer can be appointed as the responsible surrogate server for the area it covers. Therefore, these surrogate servers could cache the responses to frequently asked VITP messages and location information.

Furthermore, mobile CDNs may offer their services to several existing vehicular applications, such as CarNet [24] and TrafficView [26]. Specifically, CarNet was one of the first proposed applications that utilized efficient information dissemination techniques to provide vehicular services. On the other hand, TrafficView is a proposed framework, for inter-vehicle information exchange with regard to traffic and road conditions, route scheduling and emergency messages dissemination. TrafficView is able to display a real-time, dynamic view of the road traffic further down and overall act as a compliment of traditional GPS systems installed on many cars today.

Finally, Daly and Haahr [11] support that information dissemination in MANETs is improved by using intermediaries running the SimBet Routing algorithm. Specifically, SimBet uses a metric based on social networks analysis to locally determine the centrality of a mobile node within the network. For finding the central node, the notion of *betweenness centrality* is used. Therefore, if the destination node is not known to the sending node, then information is routed to the most central node. As far as mobile CDN is concerned, this scheme may be applied on such an infrastructure so as to forward the requested content to the most central mobile node of the wireless network; this is useful in case the user requesting the content is temporarily out of coverage (or offline).

6 Implementation and Experimentation Perspectives

The first step prior the actual design and implementation of new technologies, policies and applications to real-time, production, CDN infrastructures, are considered to be simulations. Simulations are usually employed during the initial development phase of any of the above since it is quite difficult to experiment in real-time environments. From an implementation perspective, the following simulation methodologies may be used in order to evaluate the performance of a mobile CDN infrastructure:

- **Simulation testbeds:** Large scale simulation testbeds provide the necessary infrastructure for the implementation and evaluation of new CDN technologies, policies and applications. One such testbed is PlanetLab[5]. PlanetLab is a global research overlay network in which participating nodes are strategically located at sites around the world. PlanetLab forms a CDN testbed for creating and deploying planetary-scale services and massive applications that span a significant part of the globe. This CDN testbed consists of a network of high-performance proxy servers. Such proxy (*surrogate*) servers have been deployed on many PlanetLab nodes. Two academic CDNs have already been built on top of PlanetLab (CoDeeN [6] and Coral [7]). A detailed discussion of academic CDNs is given in Chapter 1.

- **Simulation software:** As a rule of thumb, in all scientific research areas, software simulators are a must-have tool during the design and testing phase of any product. CDN simulators pose no exception to the above rule and are highly valued among the CDN-oriented research and industry communities. Specifically, these tools provide the means for simulating CDN applications and infrastructures without the financial burden of acquiring, installing and configuring the actual underlying hardware. Simulation results are reproducible and easier to analyze since simulated environments are free from any other unpredictable and uncontrollable interfering factors (i.e. unwanted external traffic), which researchers may encounter while experimenting on real infrastructures. CDN simulators, simulate the behavior of a dedicate set of machines to reliable and efficiently distribute content to users on behalf of an origin server. Chapter 5 of this book presents an analytic simulation tool for CDNs called *CDNsim*.

However, the above CDN simulation software and testbeds are proposed having in mind fixed-wired infrastructures. None of the above frameworks inherently supports the simulation and evaluation of mobile CDNs due to the fact that user mobility, the wireless medium and other distinguished characteristics of mobile CDNs described in Section 3, are not taken under consideration. As per the view of the authors, in terms of CDN simulation software, practitioners should extend existing tools (such as CDNsim) through the development of new extensible add-on modules that will allow the support of mobile CDNs. Such modules could provide:

- **Realistic Mobility Traces:** Mobile nodes movement behavior can be described using a large set of mobility traces generated from well-known and accepted mobility models. For instance, a module that supports generation of mobility traces could incorporate the functionality provided by SUMO [20]. SUMO is a well-known microscopic, highly portable traffic simulator which produces realistic mobility traces for simulating vehicular behavior in urban environments. Moreover, developers of such module could also consider CosMos, a communication and mobility scenario

[5] Planetlab: http://www.planet-lab.org
[6] A CDN for PlanetLab: http://codeen.cs.princeton.edu
[7] Coral CDN: http://www.coralcdn.org

generator proposed by Günes and Siekermann [16]. Unlike other mobility generators that provide traces based on a single mobility model, CosMos integrates several mobility models thus providing more realistic mobility patterns for wireless network simulations.

- **Support for wireless environments:** As discussed previously in this chapter, the wireless environment exhibits different characteristics than the wired environment. For this reason, simulating the characteristics of the wireless environment is a crucial milestone towards the correct simulation and evaluation of mobile CDNs. New modules that provide support for wireless environments should be able to correctly and accurately simulate characteristics such as oscillating signal strength, bandwidth, intermittent, connections in both infrastructure and ad-hoc environments. Even more, such models can leap a step forward by possibly simulating phenomena associated with wave propagation such as multi-path, fading, diffraction etc.
- **Support for mobile resource-limited nodes:** Mobile CDN simulators should also take into account a key characteristic of mobile nodes: the limitation of available resources. More specifically, the majority of such nodes that make up a mobile CDN are limited in terms of energy availability. Hence, models that describe the power consumption behaviour of such nodes (changes between sleep, transmit, receive and idle modes) under different scenarios should be designed and implemented in new add-on modules. Ultimately, such modules will aid in the correct simulation of mobile nodes behaviour under the above limitations and will allow the experimentation on techniques that aim in energy conservation.

In terms of simulation testbeds, ORBIT [8] is a wireless network testbed that supports research in wireless technologies. It consists of large-scale wireless networks made up from 400 high-speed cellular (3G) and IEEE 802.11 enabled nodes interconnected in a grid layout. The obvious advantage of ORBIT over other large-scale testbeds is the support of mobility through well known mobility models such as the Brownian motion or the random waypoint model. This provides the ability to examine various wireless applications and technologies such as MANETs and location-oriented services.

From an experimentation perspective, we expect that mobile CDNs would improve the mean response time of mobile users as well as increase the byte hit ratio in the cache of surrogate servers. On one hand, the mean response time represents the users waiting time in order to serve their requests. On the other hand, the byte hit ratio provides an indication for the performance of the network. Moreover, we expect that surrogate servers in mobile CDNs would have low replica redundancy due to their high degree of cooperation. This is a critical issue if we take under consideration that high data redundancy leads to waste of money for CDN providers [27]. Finally, it is expected that

[8] Open-Access Research Testbed for Next-Generation Wireless Networks(ORBIT): http://www.orbit-lab.org

mobile CDNs can improve the QoS of mobile users by serving the majority of their requests. Without such an infrastructure, many mobile users might face denial of services.

7 Future Research Directions

Mobile wireless networks are characterized by the high variability in user demand as well as the high demand for dynamic content and media applications. For example, a driver wants to be alerted for any emergency situations during his/her journey using multimedia data. Another characteristic of mobile networks is the scarcity of resources. For instance, consider a person who wishes to upload a video to a server in order to free up his/her camera so as to accept more pictures/videos. Therefore, crucial issues should be addressed. In the following subsections, we discuss the future research directions for mobile CDNs in terms of content placement techniques, disseminating dynamic content and mobile streaming media.

7.1 Content Placement Techniques

The content placement problem is to decide where content is to be replicated so that some objective function is optimized based on requests and resource constraints. This problem has been extensively studied for static user demands [19, 28, 44, 51]. However, due to mobility and resource constraints, the existing schemes are not applicable to a mobile wireless environment. Specifically, new approaches should be investigated, which deal with the high demand for dynamic content, media applications and user mobility.

In this context, the content placement approaches for dynamic user demands are of interest [8, 25, 31, 33]. An algorithm that dynamically places replicas and organizes them into an application-level multicast tree with limited knowledge of the network topology was presented by Chen et al. [8]. This algorithm aims at satisfying both user-perceived latency and server capacity. In [25] the authors presented another framework for dynamic content placement. In this approach, the problem of optimal dynamic content placement has been described as a semi-Markov decision process, where the user requests are assumed to follow a Markovian model. Presti et. al. [31] address the dynamic content replication by using a non-linear integer programming formulation. Specifically, the decision on how the system should evolve is the outcome of a nonlinear integer programming formulation of the problem. Rabinovich et.al [33], presented an application CDN (called ACDN) which is dedicated to deliver dynamic content. They proposed a heuristic algorithm which dynamically places replicas based on past observed demand.

However, none of the above content-placement techniques has been evaluated under a wireless network infrastructure. Motivated by this fact, the authors in [1] presented an online heuristic algorithm for dynamic placement of content replicas in a mobile CDN. The proposed algorithm (called online MDCDN) is based on a statistical forecasting method, called Double Exponential Smoothing (DES). Taking user demand variations into account, this method predicts the future demand at each surrogate server. These predictions are used to decide whether to add a new content replica or remove an existing one in order to minimize the total traffic over the backbone. For the experiments, the authors used a mobility simulator [22] and modeled a 20-km radial city, divided into area zones based on population density and natural limits (e.g. rivers, highways, etc.).

A distributed algorithm, called DbC, was proposed by Miranda et al. [23] in order to place the content as evenly as possible among all the servers that belong to a wireless ad-hoc network. Thus, the replicas of the data items are sufficiently distant from each other to prevent excessive redundancy. On the other hand, the replicas remain close enough to each end-user to improve information dissemination. Simulation results showed that DbC improves the dissemination of items throughout the network. Three efficient content placement techniques for a wireless ad hoc network infrastructure have also been proposed by [17]. These methods place the replicas on mobile hosts taking into account either the PT values[9] of objects (E-SAF - Extended Static Access Frequency) or the PT values of objects and the neighboring mobile hosts (E-DAFN - Extended Dynamic Access Frequency and Neighborhood), or the PT values of objects and the whole network topology (E-DCG - Extended Dynamic Connectivity based Grouping). Experiments performed in [17] have shown that E-DCG gives the highest accessibility and the E-SAF method gives the lowest traffic.

7.2 Disseminating Dynamic Content

The applications that disseminate dynamic content have high sensitivity to delays. For instance, a driver wants to know which road is better to follow in order to avoid any delays due to a traffic jam. In such a case, a delay - even a few seconds - may be intolerable. Thus, the efficient dissemination of dynamic content in mobile environments is a critical issue. To this end, a wide range of proposed approaches and techniques have been proposed under the CDN infrastructure in order to accelerate the generation and dissemination of dynamic content to mobile users [1, 5, 34]. Many of the proposed approaches

[9] PT value is defined as the product of the popularity of the object and the time remaining until the object is updated next.

are implemented in commercial systems (Websphere Edge services of IBM[10], EdgeSuite network of Akamai[11]) proving in this way the importance and applicability of dynamic content technology.

Fragment-based policies have received considerable attention from the research community in recent years [5, 34], since experiments have shown that the typical overhead of composing a Web page from fragments is minor to the overhead of constructing the whole page from scratch [5]. Akamai has also enhanced fragment-based policies using the Edge Side Includes technology [14]. A novel scheme to automatically detect and flag "interesting" fragments in dynamically generated Web pages that are cost-effective cache units, is proposed by Ramaswamy et al. [34]. A fragment is considered to be interesting if it is shared among multiple pages or if it has distinct lifetime or personalization characteristics.

Instead of caching fragments of Web pages, another common approach for disseminating dynamic content is to cache the means which are used in order to generate the pages over the surrogate servers. This approach is based on the fact that generation of content on demand needs more time than simply fetching any other dynamic content, since it requires to issue one or more queries to a database. Thus, a simple idea is to cache the application code at the CDN surrogate servers, and keep the data centralized. This technique is the basic module of the *Edge Computing* product from Akamai and ACDN [33]. A disadvantage of this approach is that all users requests should be served by a central database. This centralized architecture leads to performance bottlenecks. In order to address this issue, another approach is to create a partial copy of the database in each surrogate server. The systems that have been implemented this approach are known as Content-Aware Caching systems (CAC). Finally, another approach is to cache the result of database queries as they are issued by the application code. This approach is known as Content-Blind query Caching (CBC) [35]. When a query is issued, it is checked whether the query result is cached or not. Experiments results presented by Sivasubramanian et al. [41], have shown that the CBC presents the best performance when the query workload exhibits high locality. On the other hand, the Edge Computing and the CAC present better performance than the CBC when the query workload exhibits low locality.

7.3 Disseminating Mobile Streaming Media

The increased bandwidth of next-generation mobile systems makes streaming media a critical component of future content delivery services. Thus, the efficient dissemination of streaming media in mobile environments is a

[10] IBM WebSphere Application Server:
http://www-306.ibm.com/software/webservers/appserv/was/
[11] Akamai: http://www.akamai.com/

challenging issue. Typically, streaming media content can be categorized as follows:

- *Live content:* the content is disseminated "instantly" from the encoder to the media server and then onto the end-user.
- *On-demand content:* the content is encoded and then stored as streaming media files on media servers. The content is then available for request by the end-users.

The media server is a specialized one which consists of a software that runs on a general-purpose server. Its task is to serve the digitized and encoded content to end-users. When a user requests a certain content, the media server responds to the query with the specific video or audio stream. Design requirements of such servers under the context of CDNs are discussed by Roy et.al [36]. However, the dissemination of media content to a large number of mobile users creates serious problems due to the strict requirements of streaming media, the inherent limitations of wireless networks and the mobility of users.

In this context, a mobile CDN may address these limitations by distributing the high demands of mobile users for streaming media to its surrogate servers. Furthermore, CDN surrogate servers improve the dissemination of streaming content by employing state-of-the-art compression techniques (caching, encoding). Specifically, a mobile streaming media CDN should consider the following issues:

- Surrogate server selection: When a user requests an object, his/her request should be directed to a surrogate server for serving the requested content. To achieve this, Yoshimura et al. [49], proposed a mobile streaming media CDN architecture in which the surrogate server selection is determined by a SMIL[12] file modification. The SMIL file is stored in the streaming media server and contains information about the status of surrogate servers. In this architecture, the mobile users read the SMIL file in order to select the best surrogate server in CDN.
- Media Caching: Earlier research efforts have shown that CDN performance is improved when caching techniques are integrated in a CDN. The problem is to determine what media streams should be cached in surrogate servers disks. A solution would be to store all the media streams. However, such a solution is not feasible since media streams require huge amounts of storage. Therefore, efficient data management schemes should be considered in the context of mobile streaming CDNs. In such a scheme, the cache manager manages efficiently the content that has been replicated to each surrogate server. A simple idea is to partition the objects into segments.

[12] The SMIL (Synchronized Multimedia Integration Language) is a W3C Recommended XML markup language for describing multimedia presentations. It defines markup for timing, layout, animations, visual transitions, and media embedding among other things.

In particular, the media objects are divided into smaller units so that only partial units are cached. This leads to efficient storage and network utilization. In the literature, several variants have been proposed for caching segmentations [7]. Prefix caching [39], and variable sized segmentation [47] are some indicative techniques for caching segmentation.

- Managing Session Handoffs: In streaming media, the user sessions are usually long-lived. However, the long-lived nature of streaming sessions in a mobility environment has raised the issue of managing in an efficient way the session handoffs among surrogate servers. During a handoff, no frames should be lost, and the data stream to the video player should be kept as smooth as possible. In [49], a SMIL file is used to control the session handoffs.

A mobile streaming media CDN (MSM-CDN) architecture, which enables media delivery over next-generation mobile networks has been described by Apostolopoulos et al. [3]. The MSM-CDN architecture has been designed to be modular and interoperable with other systems and underlying networks. In this scheme, the overlay servers are the basic components of the MSM-CDN; their task is to cache media streams. Here, an overlay server can be considered as surrogate server. The delivery of media is done through streaming and data-transfer interfaces. Specifically, the interfaces in MSM-CDN facilitate the delivery of media streams to mobile users. Another mobile streaming media CDN architecture has been proposed in [49]. The originality of this architecture is that all the technologies related to CDN are enabled by SMIL modification.

8 Conclusion

The recent advances in mobile content networking (e.g. GSM/3G, WiFi, etc.) enable the wireless network infrastructures to support bandwidth-intensive services such as streaming media, mobile TV etc. The information which is usually requested by mobile users is either dynamic content or media applications. Taking into account that mobile user appetites for information is expected to keep growing, we need innovative techniques that can improve information dissemination. However, the emergence of typical CDNs cannot meet these challenges in a mobile wireless network due to the inherent limitations of such an infrastructure as well as the distinguished characteristics of mobile devices. Traditional CDNs do not take the mobility of users into account. In this context, *mobile CDNs* may address these issues by accelerating the information dissemination in mobile wireless network infrastructures. A mobile CDN differentiates from typical CDNs in terms of the topology of surrogate servers, content outsourcing policy, and application services.

In this chapter, we presented a pathway for mobile CDNs, in order to understand their role in the recent evolution of the mobile networking infrastructure, as well as to investigate how the information dissemination can be improved by the emerging mobile CDN practices. In this context, the main characteristics of mobile CDNs were given. Next, we presented the most popular methodologies and implementations of mobile CDNs in terms of disseminating dynamic content, content placement techniques and mobile streaming media. Finally, we presented the network infrastructures of mobile CDNs and explored existing information dissemination techniques.

In summary, information dissemination in a mobile wireless environment is an interesting, useful, and challenging problem. The emergence of mobile CDNs has opened new perspectives in the research community. Although several research works exist, there is a lot of room for improvement in both theoretical and practical applications, since technology in the areas of mobile computing and mobile networking is still evolving.

References

1. Aioffi, W.M., Mateus, G.R., Almeida, J.M., Loureiro, A.A.F.: Dynamic content distribution for mobile enterprise networks. IEEE Journal on Selected Areas on Communication **23**(10) (2005)
2. Anda, J., LeBrum, J., Ghosal, D., Chuah, C.N., Zhang, M.: Vgrid: Vehicular adhoc networking and computing grid for intelligent traffic control. In: Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st, vol. 5, pp. 2905– 2909 (2005)
3. Apostolopoulos, J.G., Wee, S., tian Tan, W.: Performance of a multiple description streaming media content delivery network. In: Proceedings of the 2002 International Conference on Image Processing (ICIP 2002), pp. 189–192. Rochester, New York, USA (2002)
4. Chae, M., Kim, J.: What's so different about the mobile internet? Commun. ACM **46**(12), 240–247 (2003)
5. Challenger, J., Dantzig, P., Iyengar, A., Witting, K.: A fragment-based approach for efficiently creating dynamic web content. ACM Trans. Inter. Tech. **5**(2), 359–389 (2005)
6. Chand, N., Joshi, R.C., Misra, M.: Cooperative caching in mobile ad hoc networks based on data utility. Mobile Information Systems **3**(1), 19–37 (2007)
7. Chen, S., Wang, H., Zhang, X., Shen, B., Wee, S.: Segment-based proxy caching for internet streaming media delivery. IEEE MultiMedia **12**(3), 59–67 (2005)
8. Chen, Y., Katz, R.H., Kubiatowicz, J.: Dynamic replica placement for scalable content delivery. In: Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS), pp. 306–318. Cambridge, USA (2002)
9. Chen, Y., Qiu, L., Chen, W., Nguyen, L., Katz, R.: Efficient and adaptive web replication using content clustering. IEEE Journal on Selected Areas in Communications **21**(6), 979–994 (2003)
10. Chisalita, I., Shahmehri, N.: A peer-to-peer approach to vehicular communication for the support of traffic safety applications. In: Proceedings of IEEE 5th International Conference on Intelligent Transportation Systems, pp. 336– 341 (2002)
11. Daly, E.M., Haahr, M.: Social network analysis for routing in disconnected delay-tolerant manets. In: Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing, pp. 32–40 (2007)

12. Dikaiakos, M.D.: Intermediary infrastructures for the world wide web. Computer Networks **45**(4), 421–447 (2004)
13. Dikaiakos, M.D., Florides, A., Nadeem, T., Iftode, L.: Location-aware services over vehicular ad-hoc networks using car-to-car communication. IEEE Journal on Selected Areas in Communications **25**(8) (2007)
14. Dilley, J., Maggs, B.M., Parikh, J., Prokop, H., Sitaraman, R.K., Weihl, W.E.: Globally distributed content delivery. IEEE Internet Computing **6**(5), 50–58 (2002)
15. Dow, C.R., Lin, P.J., Chen, S.C., Lin, J.H., Hwang, S.F.: A study of recent research trends and experimental guidelines in mobile ad hoc networks. In: AINA '05: Proceedings of the 19th International Conference on Advanced Information Networking and Applications, pp. 72–77 (2005)
16. Günes, M., Siekermann, J.: Cosmos - communication scenario and mobility scenario generator for mobile ad-hoc networks. In: Proceedings of the 2nd Int. Workshop on MANETs and Interoperability Issues (MANETII'05) (2005)
17. Hara, T.: Replica allocation methods in ad-hoc networks with data update. Mobile Networks and Applications (MONET) **8**(4), 343–354 (2003)
18. Housel, B.C., Samaras, G., Lindquist, D.B.: Webexpress: a client/intercept based system for optimizing web browsing in a wireless environment. Mobile Networking and Applications **3**(4), 419–431 (1998)
19. Kangasharju, J., Roberts, J.W., Ross, K.W.: Object replication strategies in content distribution networks. Computer Communications **25**(4), 376–383 (2002)
20. Krajzewicz, D., Hertkorn, G., Rossel, C., Wagner, P.: Sumo (simulation of urban mobility); an open-source traffic simulation. In: 4th Middle East Symposium on Simulation and Modelling (MESM2002), pp. 183–187. SCS European Publishing House (2002). Http://sumo.sourceforge.net/ (last accessed January 2007)
21. Mahdy, A.M., Deogun, J.S., Wang, J.: Mobile ad hoc networks: a hybrid approach for the selection of super peers. In: Proceedings of the 2nd IFIP International Conference on Wireless and Optical Communications Networks (WOCN 2005), pp. 280–284 (2005)
22. Mateus, G.R., Goussevskaia, O., Loureiro, A.A.F.: Simulating demand-driven server and service location in third generation mobile networks. In: Proceedings of the 9th International Euro-Par Conference, pp. 1118–1128. Klagenfurt, Austria (2003)
23. Miranda, H., Leggio, S., Rodrigues, L., Raatikainen, K.E.E.: An algorithm for dissemination and retrieval of information in wireless ad-hoc networks. In: Proceedings of the 13th International Euro-Par Conference, pp. 891–900. Rennes, France (2007)
24. Morris, R., Jannotti, J., Kaashoek, F., Li, J., Couto, D.D.: Carnet: A scalable ad hoc wireless network system. In: Proceedings of the 9th ACM SIGOPS European workshop: Beyond the PC: New Challenges for the Operating System, pp. 61–65. Kolding, Denmark (2000)
25. N. Bartolini, F.L.P., Petrioli, C.: Optimal dynamic replica placement in content delivery networks. In: 11th IEEE International Conference on Networks (ICON 2003), pp. 125–130. Sydney, Australia (2003)
26. Nadeem, T., Dashtinezhad, S., Liao, C., Iftode, L.: Trafficview: traffic data dissemination using car-to-car communication. SIGMOBILE Mob. Comput. Commun. Rev. **8**(3), 6–19 (2004)
27. Pallis, G., Vakali, A.: Insight and perspectives for content delivery networks. Commun. ACM **49**(1), 101–106 (2006)
28. Pallis, G., Vakali, A., Stamos, K., Sidiropoulos, A., Katsaros, D., Manolopoulos, Y.: A latency-based object placement approach in content distribution networks. In: Third Latin American Web Congress (LA-Web 2005), pp. 140–147. Buenos Aires, Argentina (2005)

29. Pathan, A.M.K., Buyya, R.: A taxonomy and survey of content delivery networks. Technical Report, GRIDS-TR-2007-4, Grid Computing and Distributed Systems Laboratory, The University of Melbourne (2007)
30. Pitoura, E., Chrysanthis, P.K.: Caching and replication in mobile data management. IEEE Data Eng. Bull. **30**(3), 13–20 (2007)
31. Presti, F.L., Petrioli, C., Vicari, C.: Dynamic replica placement in content delivery networks. In: Proceedings of the 13th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2005), pp. 357–360. Atlanta, GA, USA (2005)
32. Rabinovich, M., Spatscheck, O.: Web Caching and Replication. Addison Wesley (2002)
33. Rabinovich, M., Xiao, Z., Douglis, F., Kalmanek, C.R.: Moving edge-side includes to the real edge - the clients. In: Proceedings of the USENIX Symposium on Internet Technologies and Systems, pp. 12–26. Seattle, Washington, USA (2003)
34. Ramaswamy, L., Iyengar, A., Liu, L., Douglis, F.: Automatic fragment detection in dynamic web pages and its impact on caching. IEEE Trans. Knowl. Data Eng. **17**(6), 859–874 (2005)
35. Rilling, L., Sivasubramanian, S., Pierre, G.: High availability and scalability support for web applications. In: SAINT '07: Proceedings of the 2007 International Symposium on Applications and the Internet. IEEE Computer Society, Washington, DC, USA (2007)
36. Roy, S., Ankcorn, J., Wee, S.: Architecture of a modular streaming media server for content delivery networks. In: Proceedings of the 2003 International Conference on Multimedia and Expo (ICME '03), vol. 3, pp. 569–572. IEEE Computer Society, Washington, DC, USA (2003)
37. Royer, E., Toh, C.: A review of current routing protocols for ad-hoc mobile wireless networks. Personal Communications, IEEE **6**(2), 46–55 (1999)
38. Sailhan, F., Issarny, V.: Energy-aware web caching for mobile terminals. In: Proceedings of the 22nd International Conference on Distributed Computing Systems, pp. 820–825. IEEE Computer Society (2002)
39. Sen, S., Rexford, J., Towsley, D.F.: Proxy prefix caching for multimedia streams. In: Proceedings of the IEEE INFOCOM, pp. 1310–1319. New York, USA (1999)
40. Sidiropoulos, A., Pallis, G., Katsaros, D., Stamos, K., Vakali, A., Manolopoulos, Y.: Prefetching in content distribution networks via web communities identification and outsourcing. World Wide Web (2008)
41. Sivasubramanian, S., Pierre, G., van Steen, M., Alonso, G.: Analysis of caching and replication strategies for web applications. IEEE Internet Computing **11**(1), 60–66 (2007)
42. T Nadeem, P.S., Iftode., L.: A comparative study of data dissemination models for vanets. Proceedings of the 3rd Annual International Conference on Mobile and Ubiquitous Systems (MOBIQUITOUS) (2006)
43. Teng, W.G., Chang, C.Y., Chen, M.S.: Integrating web caching and web prefetching in client-side proxies. IEEE Trans. Parallel Distrib. Syst. **16**(5), 444–455 (2005)
44. Tse, S.S.H.: Approximate algorithms for document placement in distributed web servers. IEEE Trans. Parallel Distrib. Syst. **16**(6), 489–496 (2005)
45. Vakali, A., Pallis, G.: Content delivery networks: status and trends. IEEE Internet Computing **7**(6), 68–74 (2003)
46. Wu, T., Dixit, S.: The content driven mobile internet. Wireless Personal Communications: An International Journal **26**(2-3), 135–147 (2003)
47. Xu, Z., Guo, X., Wang, Z., Pang, Y.: The dynamic cache algorithm of proxy for streaming media. In: Proceedings of the International Conference on Intelligent Computing (ICIC 2005), pp. 1065–1074. Hefei, China (2005)
48. Yin, L., Cao, G.: Supporting cooperative caching in ad-hoc networks. IEEE Trans. Mob. Comput. **5**(1), 77–89 (2006)

49. Yoshimura, T., Yonemoto, Y., Ohya, T., Etoh, M., Wee, S.: Mobile streaming media cdn enabled by dynamic smil. In: Proceedings of the 11th International World Wide Web Conference, WWW2002, pp. 651–661. Honolulu, Hawaii, USA (2002)
50. Zemlianov, A., de Veciana, G.: Capacity of ad hoc wireless networks with infrastructure support. IEEE Journal on Selected Areas in Communications **23**(3), 657–667 (2005)
51. Zhuo, L., Wang, C.L., Lau, F.C.M.: Load balancing in distributed web server systems with partial document replication. In: Proceedings of the 31st International Conference on Parallel Processing (ICPP), p. 305. Vancouver, Canada (2002)