Fog Computing and Its Applications in 5G

Longxiang Gao, Tom H. Luan, Bo Liu, Wanlei Zhou, and Shui Yu

Abstract With smartphones becoming our everyday companions, high-quality mobile applications have become an important integral of people's lives. The intensive and ubiquitous use of mobile applications have led to explosive growth of mobile data traffics. To accommodate the surge mobile traffic yet providing the guaranteed service quality to mobile users represent a key issue of 5G mobile networks. This motivates the emergence of Fog computing as a promising, practical and efficient solution tailored to serving mobile traffics. Fog computing deploys highly virtualized computing and communication facilities at the proximity of mobile users. Dedicated to serving the mobile users, Fog computing explores the predictable service demand patterns of mobile users and typically provides desirable localized services accordingly. Stitching above features, Fog computing can provide mobile users with the demanded services through low-latency and short-distance local connections. In this chapter, we introduce the main features of Fog computing and describe its concept, architecture and design goals. Lastly, we discuss on the potential research issues from the perspective of 5G networking.

1 Introduction

Smartphones have already become our everyday companions. In 2011, the smartphone shipment worldwide overtook that of PCs for the first time in history, and now the smartphone penetration has reached 75 % in US. It is envisioned by Cisco that the average number of connected mobile devices per person will hit 6.56 in 2020, due to the proliferating use of "Internet of Things" applications, e.g., smart home, smart community, and emerging mobile electronics, e.g., wearable devices.

Smart devices have brought rich computing and communication capability to the palm of our hand. As a result, rich mobile applications are enabled to enhance our day-to-day experiences by enabling productivity, connectivity and achievement

L. Gao • T.H. Luan (✉) • B. Liu • W. Zhou • S. Yu
School of Information Technology, Deakin University, Burwood, Victoria, Australia
e-mail: longx@deakin.edu.au; tom.luan@deakin.edu.au; bo.liu@deakin.edu.au; wanlei.zhou@deakin.edu.au; shui.yu@deakin.edu.au

W. Xiang et al. (eds.), *5G Mobile Communications*,
DOI 10.1007/978-3-319-34208-5_21

of our goals. For example, we may be already addicted to mobile applications everyday for social connectivity and to fulfill our multimedia messaging needs. On each passing day, mobile applications connect with wearable devices to read our heartbeat and track our health conditions, adjust the temperature and light of our room by communicating with smart home facility, coerce us to take that morning walk, offer a different route that will help us avoid the rush hour traffic, and become increasingly intelligent by understanding our mobilities, gestures and social activities. Apparently, from changing the way we communicate to revolutionizing the way we work and live, mobile electronics and applications pervade our daily lives everywhere.

The proliferation and pervasive use of mobile applications inevitably leads to the explosive growth of the mobile data traffic. To accommodate the surge mobile traffic and in the meantime provide guaranteed service quality to mobile users represent the key issue of next generation mobile networks. This motivates the emergence of Fog computing as a promising, practical and efficient solution that extends cloud computing to better serving mobile traffics. The term "Fog computing" was first proposed by Cisco in 2012 [1]. Similar systems typically known as edge computing, such as Cyber Foraging [2], Cloudlets [3] can date back to early 2000.

In this chapter, we introduce why the Fog computing is promising, the main features of Fog computing and describe its concept, architecture and design goals. Then we demonstrate a case study on how the Fog computing can improve the network performance in 5G environment, followed by a discussion on the potential research issues from the perspective of 5G networking.

2 Fog Computing Architecture

Fog computing extends the cloud-based Internet by introducing an intermediate layer between mobile devices and cloud, aiming at the smooth, low-latency service delivery from the cloud to mobile. This accordingly leads to a three hierarchy Mobile-Fog-Cloud architecture as depicted in Fig. 1.

The intermediate Fog layer is composed of geo-distributed Fog servers which are deployed at the edge of networks, e.g., parks, bus terminals, shopping centers, etc. Each Fog server is a highly virtualized computing system, similar to a light-weight cloud server, and is equipped with the on-board large-volume data storage, compute and wireless communication facility. The role of Fog servers is to bridge the mobile users and cloud. On one hand, Fog servers directly communicate with the mobile users through single-hop wireless connections using the off-the-shelf wireless interfaces, such as WiFi and Bluetooth. With the on-board compute facility and pre-cached contents, they can independently provide pre-defined service applications to mobile users without assistances from cloud or Internet. On the other hand, the Fog servers can be connected to the cloud so as to leverage the rich functions and application tools of the cloud. The next section describes some typical examples of Fog computing in details.

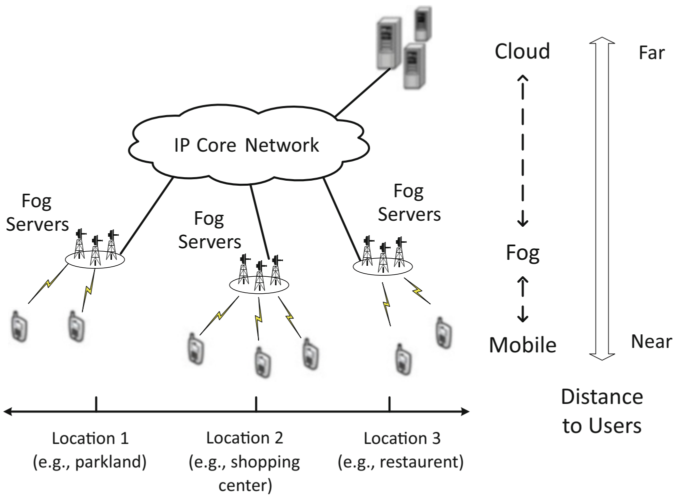

Fig. 1 Fog computing architecture

To summarize, the purpose of Fog computing is to place a handful of compute, storage and communication resources in the proximity of mobile users, and therefore to serve mobile users with the local short-distance high-rate connections. This overcomes the drawback of cloud which is far to mobile users with elongated service delays. Therefore, the fog is interpreted as "the cloud close to the ground" [1].

3 Why Fog Computing?

Nowadays, the evolving of Internet has shown two obvious trends. First, the could-based architecture is adopted to host major applications and storage. As predicted by Cisco Global Cloud Index, the global cloud traffic will account for more than three-fourths of total data center traffic by 2018. Second, the Internet users have shifted predominantly from using desktop computers to smartphones and tablets. With cloud becoming the overarching approach for service delivery and information retrieval, and mobile users becoming the major service consumers, the seamless interconnection of cloud computing and mobile applications therefore represents a key issue in the 5G mobile networks and motivates the emergence of Fog computing.

To show the rationale of Fog computing, in what follows we take a retrospect study by revisiting the design of cloud-based Internet and service requirements of mobile users.

3.1 Global and Local Information

The cloud computing represents an efficient and scalable centralized solution for information management and distribution. It is efficient to serve the information requests from the traditional desktop users. To be specific, the desktop users, typically accessing Internet at homes and offices are often interested in the information which is irrelevant to their locations, such as the world news, stock market at different cities or countries, to name a few. We refer to such information as the *global information*. As a contrast, we refer to the location-based service information related to the location of users as the *local information*. Cloud computing favors desktop users with an optimized approach for serving the global information services. With a scalable and efficient approach to store and manage the information originated at different locations of the world and using a static public IP, cloud computing conveniently distribute the cached global information from a remote central server to desktops worldwide.

The mobile users, however, have distinguished service requirements from the desktop users. This requires the current cloud-based Internet to be modified accordingly to cater to the specific service requirements of mobile users. In specific, different from desktop users, mobile users, particularly smartphones, are typically in the outdoor environments. This makes their service requirements closely related to their current locations. In other words, mobile users are more interested in the *local information* around them. For example, a mobile user in a shopping center tends to be interested in the sales, open hour, restaurants and events inside the attended shopping center; such information become useless once he/she leaves the shopping center. In another example, a traveller to a city would seek for information on the places of interest, local news and weather conditions of the specific city, while such information of other places is useless. The massive demand of location-based mobile services is also reported in [4].

3.2 Physical and Communication Distance

The cloud-based Internet can be inefficient to serve the local information desired by mobile users. As a motivating example shown in Fig. 2a, assuming that a mobile user inside a shopping center intends to retrieve flyers of stores within the shopping center. To do this using the cloud-based Internet, the stores may need to first upload their flyers to a remote cloud server over Internet, and then direct mobile users to retrieve the desired information from the remote cloud server. In other words, although the *physical distance* between the mobile user (destination) and stores (original source) is short, using the remote cloud as the information depot, the actual *communication distance* can be far, e.g., from the cloud server to mobile user in this example.

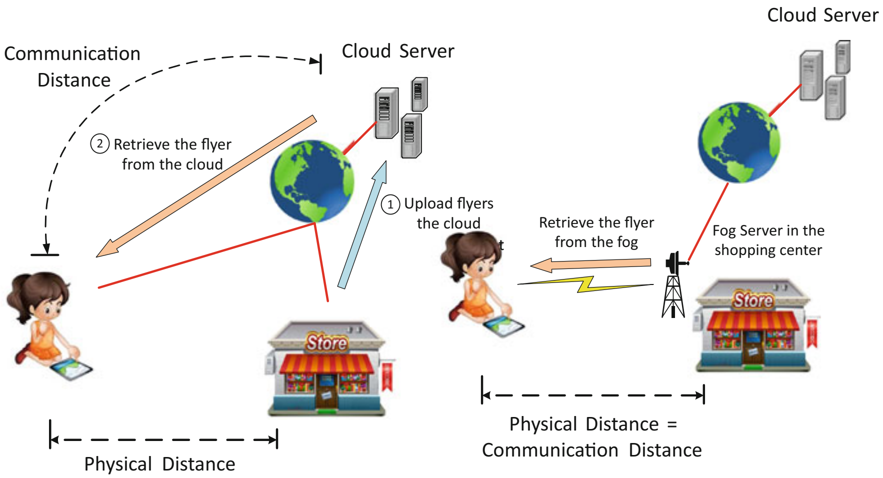

Fig. 2 Example: download the flyer of a nearby store. (**a**) Retrieving the flyer from the cloud. (**b**) Retrieving the flyer from the fog

The Fog computing paradigm represents a practical and efficient solution to resolve the mismatch between physical and communication distances. As a remedy shown in Fig. 2b, a Fog server can be deployed inside the shopping center and to distribute the local store flyers to mobile users. As such, the physical distance is equal to the communication distance and users can acquire low-latency desirable services.

By minimizing the communication distance, the Fog computing therefore brings the following two advantages:

- **To mobile users**: compared to cloud, the Fog computing can provide enhanced service quality with much increased data rate and reduced latency and response time. Moreover, by reducing the bandwidth cost of data transmission in the backbone, the users can also be benefited from the reduced service cost.
- **To network**: by avoiding the duplicated back and forth traffic between cloud and mobile user, not only the backbone bandwidth can be significantly saved, the energy consumption of core networks can also be greatly reduced, which contributes to the sustainable development of networking.

4 Components of Fog Computing

The Fog thus behaves as a surrogate of Cloud or a private Cloud at the user's premises. This enables Fog servers to be more efficient to handle the localized computation requests. Therefore, Fog computing targets to deliver the localized and location-based service applications to mobile users. In what follows, we showcase some examples of Fog computing implementation from this perspective, and discuss on the features of a Fog server as a comparison to Cloud server.

4.1 Exemplary Implementations

4.1.1 Shopping Center

Assuming that a number of Fog servers are deployed inside a multi-floor shopping center, which collectively form an integrated localized information system. The Fog servers at different floors can pre-cache floor-related contents, such as the layout and ads of stores on a particular floor. The Fog servers can deliver engaged services including indoor navigation, ads distribution and feedback collections to mobile users through WiFi.

4.1.2 Parkland

The Fog computing system can be deployed in the parkland to provide localized travel services. For instance, Fog servers can be deployed at the entrance and other important locations of the park. The Fog server at the entrance can pre-cache information including park map, travel guide and local accommodations; other Fog servers at different locations inside the park can be incorporated with sensor networks for environment monitoring and provide navigation to travellers. By connecting the Fog servers to the park administration office and cloud, the Fog servers can be used as an information gateway to send timely alerts and notifications to travellers.

4.1.3 Inter-State Bus

Greyhound has launched "BLUE" [5], an on-board Fog computing system over inter-state buses for entertainment services. As an example illustrated in Fig. 3, a Fog server can be deployed inside the bus and provides on-board video streaming, gaming and social networking services to travellers using WiFi. The on-board Fog server connects to the Cloud through cellular networks to refresh the pre-catched contents and update application services. Using its computing facility, the Fog server can also collect and process user's data, such as number of travellers and their feedbacks, and report to cloud.

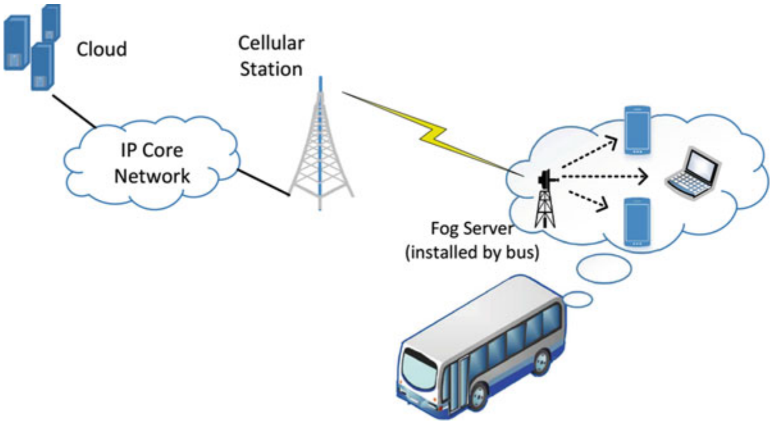

Fig. 3 On-board fog computing system

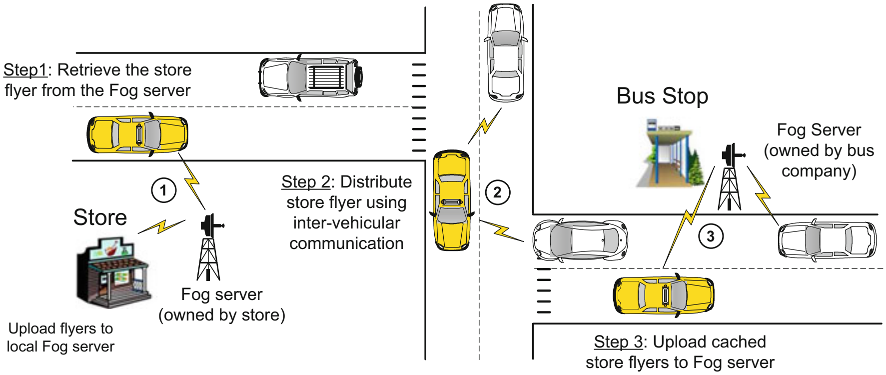

Fig. 4 Fog computing for content distribution in vehicular ad hoc networks

4.1.4 Vehicular Fog Computing Networks

Luan et al. [6] present the application of Fog computing as an integrated large-scale network for localized content disseminations. Figure 4 shows a motivating scenario. Assuming that a store installs a Fog server at its parking lot with the purpose to distribute the store flyer. In step 1, the store uploads flyers to the Fog server via wireless connections, and the Fog server distributes the flyers wirelessly to vehicles driving through its coverage using wireless communications. With the vehicle moving to different locations, it can further disseminate the cached flyers to other vehicles using wireless communications, as depicted in Step 2. In Step 3, the flyers can also be retrieved and cached at other Fog servers deployed at different locations, e.g., bus stop, and further propagated in the network.

4.2 Comparison to Cloud Computing

Fog computing is featured by the dedication to serving the localized and location-based applications. To this end, a Fog server manages its on-board resources to fully explore the location information and predictable user demand with the following functions.

- **Wireless**: Fog computing is dedicated to serving the mobile users. Each Fog server typically has limited wireless coverage, e.g., 200 m using WiFi, and directly interacts with mobile users using the single-hop wireless connections.
- **Local Services**: Fog computing is dedicated to serving the localized information and providing location-based service applications. For example, the Fog computing system deployed in a specific park only provide the navigation services within the park.
- **Distributed Management**: A Fog computing system may typically be deployed and managed by the local business, with the purpose to deliver designated contents and services to specific user groups.

Using the example in Fig. 4 for illustration, the Fog servers provide localized content distribution using wireless communications, which matches the first two features. The Fog server deployed nearby the store may be installed and managed by the store owner for the distribution of store flyers; the Fog server at the bus stop may be managed by the bus company for the distribution of bus information, e.g., bus schedules, safety manual, etc., to mobile users waiting for the bus. The Fog computing system in [6] is therefore distributedly constructed with Fog servers distributed installed and managed by different entities to serve their own purposes, which matches the third feature.

Table 1 summarizes the differences between Fog computing and Cloud computing.

By targeting to different user groups at different locations, Fog computing extends Cloud computing to better serve local mobile traffics. As the system architecture shown in Fig. 1, the Fog servers deployed at different locations would be used to deliver engaged services specified by their owners. The Fog servers at different locations can connect to the same cloud and form an integrated Fog computing system in a wide region.

4.3 Components of Fog Computing

4.3.1 Storage

In a predefined service area, a Fog server predicts the mobile user's demand on information and pre-cache the desirable information accordingly using a proactive way in its storage. Such information can be either retrieved from the Cloud or uploaded by its owner. For example, the Fog server installed at a restaurant can

Table 1 Comparison of fog computing and cloud computing

	Fog computing	Cloud computing
Target user	Mobile users	General Internet users
Service type	Limited localized information services related to specific deployment locations	Global information collected from worldwide
Hardware	Limited storage, compute power and wireless interface	Ample and scalable storage space and compute power
Distance to users	In the physical proximity and communicate through single-hop wireless connection	Faraway from users and communicate through IP networks
Working environment	Outdoor (streets, parklands, etc.) or indoor (restaurants, shopping malls, etc.)	Warehouse-size building with air conditioning systems
Deployment	Centralized or distributed in reginal areas by local business (local telecommunication vendor, shopping mall retailer, etc.)	Centralized and maintained by Amazon, Google, etc.

pre-cache the menu of the restaurant and dish recipes to serve the mobile users inside the restaurant. In another example, the Fog servers deployed in the airport can pre-cache the flight and local transportation information which is desirable to travellers in the airport. Therefore, the key design issue of Fog computing is to predict the user's demand and proactively select the contents to cache in the geo-distributed Fog servers based on the specific locations.

The Content Delivery Network (CDN) [7] represents the most mature catch networks and extensively pursued in both academic and industry. CDN is the Internet-based cache network by deploying cache servers at the edge of Internet to reduce the download delay of contents from remote sites. CDN mainly targets to serve traditional desktop Internet users, which have much broader interests and blur service demands that are more difficult to predict than those of mobile users. With precise service region, Fog computing has more clear target users of specific service demand. It is thus key for Fog servers to explore this feature to fully utilize its storage and computing resources to provide the most desirable services to mobile users.

Similar to Fog computing, the Information Centric Network (ICN) [8] is also a wireless cache infrastructure which provides content distribution services to mobile users with distributed cache servers. Different from the cache servers in ICN, the Fog servers are intelligent computing unit. Therefore, the Fog servers are not only used for caching, but also as a computing infrastructure to interact with mobile users and devices for real-time data processing. The Fog servers can be connected to the cloud and accordingly utilize the extensive computing power and big data tools for rich applications other than content distribution, such as internet of things, vehicular communications and smart grid applications [1].

Baştuğ et al. [9] also show that the information demand patterns of mobile users are predictable to an extent and propose to proactively pre-cache the desirable information before users request it. The social relations and device to device communications are leveraged. Unlike Fog computing, the proactive caching scheme in [9] is not explicitly used to serve local information services. As a more broad and generic paradigm, Fog computing can incorporate the proactive caching framework as described in [9].

4.3.2 Compute

A salient feature that differentiates Fog computing from the traditional cache networks is that Fog servers are intelligent compute system. This allow a Fog server to autonomously and independently serve local computation and data processing requests from mobile users. Satyanarayanan et al. [10] shows the applications of Fog computing in the cognitive applications. In another example, a Fog server inside the shopping mall or parkland can maintain an on-board geographic information system, and provide the real-time navigation and video streaming to connect mobile users.

Bridging the mobile and Cloud, a Fog server can also be conveniently used to collect the environmental data or demographic data from mobile users at the deployed spot, and transport the collected big data to Cloud for in-depth data analysis; the results can be provided to third party for strategic and valuable insights on business and government event planning, execution and measurement.

Despite of the high computing power, the Cloud is faraway from mobile users and can hardly support real-time computing intensive applications due to the bandwidth-constrained IP networks. The demand of real-time resource-intensive mobile applications, e.g., cognitive and internet-of-things applications, motivates the design of ubiquitous edge computing system [10, 11]. Cloudlets [3, 10] adopt the same framework of Fog computing, in which a Cloudlet server, similar to the Fog server, is deployed in the proximity of mobile users and processes the computing requests of mobile devices at real-time for video streaming and data processing. A comparison of processing delays using Cloudlets and Amazon Clouds is shown in http://elijah.cs.cmu.edu/demo.html. Transparent computing [11] is a highly virtualized system, which targets to develop the computing system transparent to users with cross-platform and cross-application support.

The Fog computing is a generic platform for edge computing and focuses on the localized service applications and computation requests. The prototype and techniques in [10, 11] can be incorporated in Fog computing framework.

4.3.3 Communication

Fog server can equip with different wireless interfaces, e.g., WiFi, Bluetooth and visible light communications [12] according to the specific application scenarios. The Fog computing differs from traditional radio access networks, e.g., WiFi and Femtocell networks, in two important ways.

Cross-Layer Design Unlike traditional WiFi access points, the Fog server manages an autonomous, all-inclusive network by providing both service applications and wireless communications to mobile users in the coverage. Therefore, a Fog server can work without Internet connections as that in [6]. Note that the Fog computing tailors its applications based on the specific deployment location and environment, and therefore is highly service-oriented. To this end, a Fog server can manage all the communication layers and effectively enables the cross-layer design [13] to provide the best service quality to users. For example, as in "BLUE" [5], a Fog server can cache a number of videos and deliver Youtube-like video streaming services to mobile users in the proximity. In this case, based on the context, wireless channel and video popularity information, the video services can be conveniently adapted towards the optimal performance via cross-layer adjustifications.

Predictable Location-Based Service The key of Fog computing is to provide the localized network and information applications to mobile users, whereas the traditional radio access networks focus on the provision of Internet applications and global information. With this distinguished feature, the design of Fog computing communications needs to consider the specific deployment environment and the features of mobile users in the considered scenario. For example, a Fog computing system deployed in the shopping mall needs to address the diverse mobilities of users, whereas the similar system deployed in the inter-state bus [5] only needs to consider static on-board passengers.

5 Case Study: Hybrid Data Dissemination in Fog Computing

In this section, we demonstrate a case study based on Fog computing and show how Fog computing can be incorporated into the 5G network towards improved performance to mobile users.

5G technique will make streaming applications becoming more and more popular, however the long latency may severely affect the user experience and is not tolerable. To address this issue, Fog computing can move Cloud services from remote Internet to the edge of networks and makes streaming content much closer to mobile users, which significantly decreases the streaming latency. On the contrary, the data dissemination from Cloud to every Fog servers can be expensive, which may takes the hudge 5G bandwidth resource. In addition, note that since the majority of streaming applications are video based, such as movies, teleplay and product advertisement, such contents are not always necessary to be strictly up-to-date and 1 or 2 days latency is affordable. Therefore, the Fog computing and delay tolerant network (DTN) techniques can be combined together to improve the performance of 5G network.

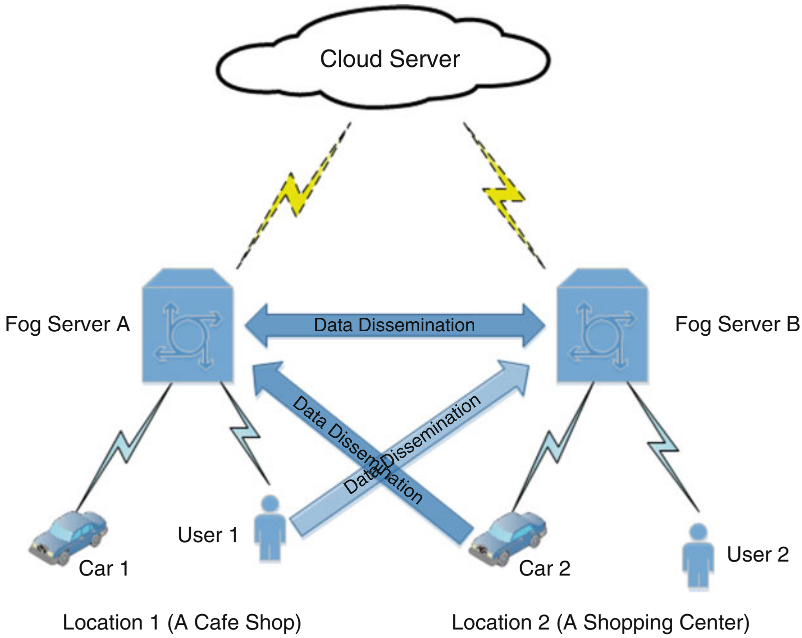

Fig. 5 Data dissemination in fog computing based on delay tolerant network technique

5.1 System Model

To have an efficient data dissemination in Fog computing, DTN is used to offload data among Fog servers. For example, in Fig. 5, if Fog server A has some small content, such as a new video ads, but Fog server B does not. In previous, Fog server B needs to get the update from a Cloud server directly. With DTN technique, user 1 could download this content when he has a coffee in this cafe shop. After a couple of hours, he goes to shopping center for shopping. When he moves into the transmission range of Fog server B, the content storied in user 1's mobile device is automatically upload to Fog server B and this store-carry-forward process is completed. For a large content, e.g. a high definition movie, transmitted from Fog server B to Fog server A, vehicle based DTN is used. In this example, if car 2 is parked in shopping center and within the Fog server B's transmission range, it downloads this movie into its local storage. Once this vehicle moves to the transmission range of Fog server A, this movie is uploaded to it and this data dissemination is completed.

In addition of the above DTN based data dissemination between mobile user and Fog server, direct data dissemination from Fog server to Fog server based on DTN technique is also available. As shown in Fig. 6, there is an on-board Fog server on a tourism bus, where all passengers on this bus can access its Fog server to watch movies or play games. If a passenger wants to find some new stuff, such as "just in"

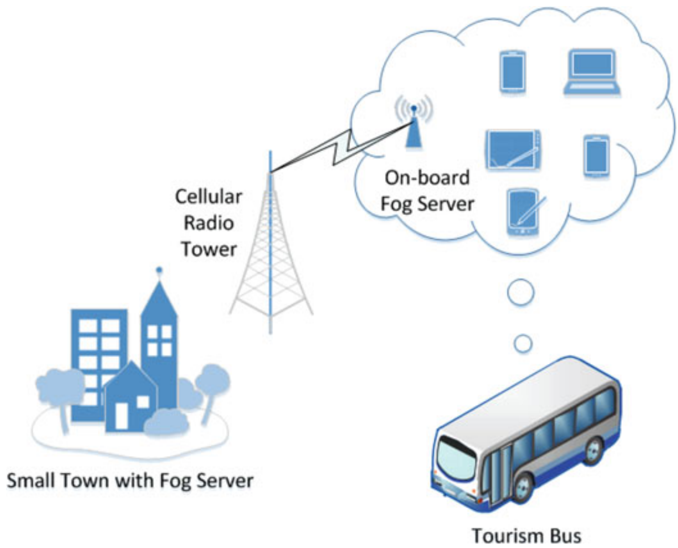

Fig. 6 Data dissemination between fog servers

news, Fog server can use the cellular network to retrieve this content immediately. When this bus travel to a small town along its route, it can synchronize its content with the Fog server located in this small town to update both servers' content list. If the Fog server in this small town has some content while other small towns along the route do not have, this tourism bus could download these contents into its local storage and carry to those small towns where they need these contents.

With the above DTN based data dissemination techniques, we propose a hybrid data dissemination model, as shown in Fig. 7. This hybrid model not only includes the normal data dissemination between Cloud servers and Fog servers, but also involves large amount low-cost DTN based data disseminations, which can be used between Fog servers and mobile users and among Fog servers. To organize these data dissemination, we re-identify the function of Cloud servers. In this model, the main function of Cloud server is to act as the "control plane" to determine the Fog server needed to be updated with the required content and control data dissemination process, as shown in Fig. 8. Fog servers and part of Cloud servers are treated as "data plane" to provide data dissemination service.

This data dissemination model has three components, namely as *Data Structures*, *Protocol Messages* and *Algorithms*. Data structures use tables to store key information which is used to determine the path of data dissemination. Protocol messages use various tapes of messages to discover content and mobile devices associated with a Fog server, exchange content list, and other tasks to learn and maintain accurate information about the network. Algorithms are used to calculate the best data dissemination path.

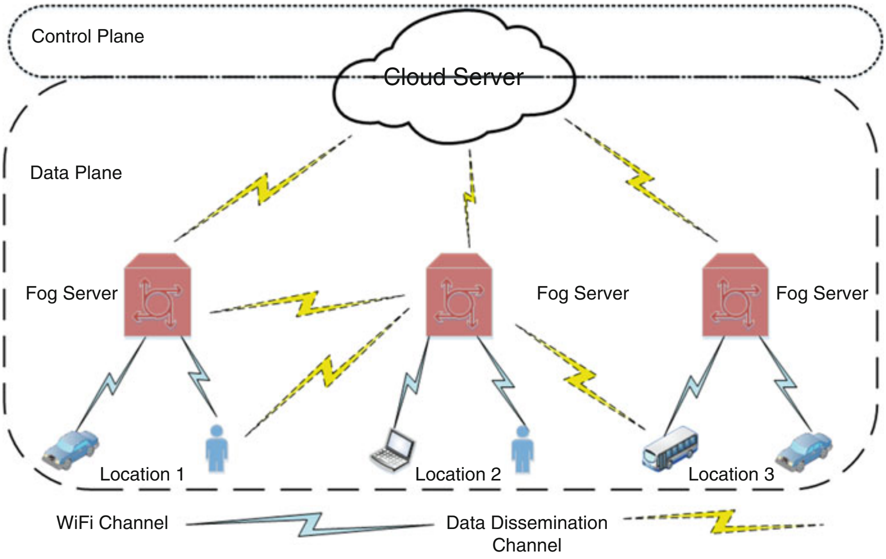

Fig. 7 Hybrid data dissemination model in fog computing

Cloud server in this model needs to have an overall information and its data structures include the following tables:

- Fog Server List Table: a table to record all Fog servers, which are managed by Cloud server. This table includes Fog server's ID, content ID in each of Fog servers, mobile device ID associated with each of Fog servers.
- Global Content List Table: a table to record all public contents (not include these private content created by Fog server owner) in Fog servers or supposed to be in Fog servers. This table includes content ID, the size of each contents, Fog server's ID (for these Fog servers who have this content), date of update, validation time.
- Mobile Devices's Movement Pattern Table: a table to record mobile device's ID, Fog server ID (whom mobile device linked before), linked time, social attribute, geographic movement pattern.

A Fog server needs to have a table to record content ID, the size of this content, mobile device's ID which linked with this Fog server, the linked duration of this mobile device. For mobile devices, they need to record the content ID which they carry on, their movement path, time duration with a Fog server and its ID.

In order to collect and exchange the above information, several data messages are used in this model.

- Hello Message between Fog Servers and Cloud Server: an update message from a Fog server to its Cloud provider, which includes its content ID and associated mobile devices ID. This is a triggered message.

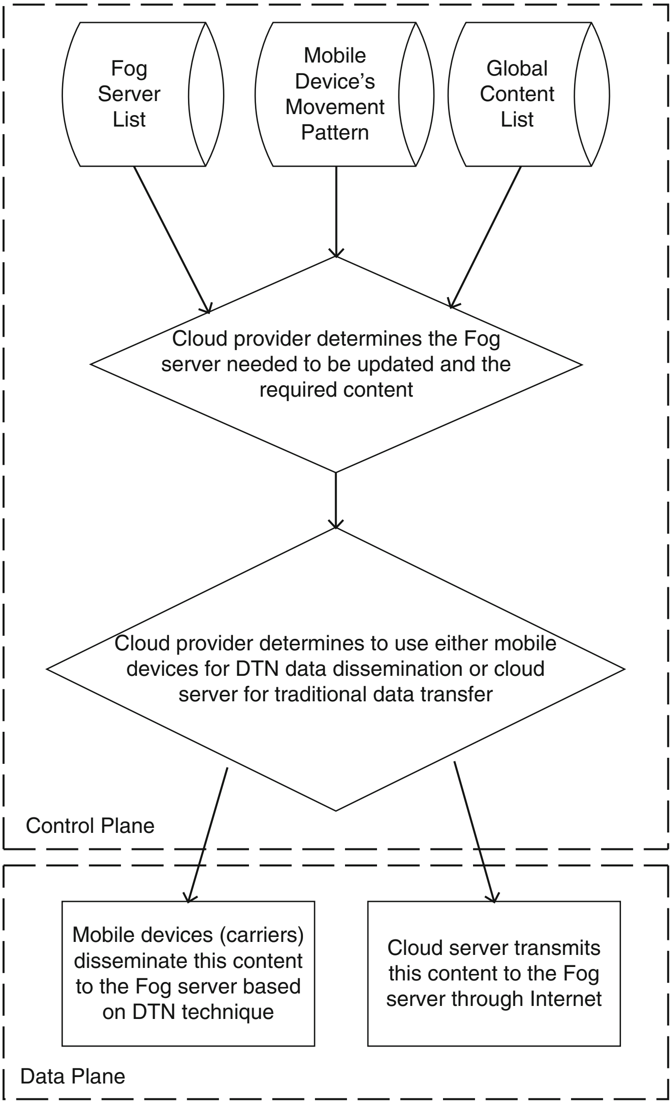

Fig. 8 Control plane and data plane in this model

- DTN Data Dissemination Request Message: a message sent from Cloud provider to a Fog server. When a Cloud provider determines there is a content need to be updated for Fog server A, if Fog server B has this content and its associated mobile device has potential move to Fog server A, a DTN data dissemination request message is sent to Fog server B to ask it to disseminate the content through that mobile device.

- DTN Data Dissemination Accept Message: a message sent from Fog server to its Cloud provider. Once a Fog server received a DTN data dissemination request message and it sends the content to the corresponding mobile device, it sends this DTN data dissemination accept message to its Cloud provider to confirm that this content has been sent out.
- DTN Data Dissemination Decline Message: a message sent from Fog server to its Cloud provider. When a Fog server received a DTN data dissemination request message, it's associated carrier (mobile device), for some reason, does not receive the complete content before this carrier leave the current Fog server. In this case, the DTN data dissemination decline message is sent back to the Cloud provider.
- DTN Data Dissemination Acknowledgement Message: a message sent from Fog server to its Cloud provider. When a Fog server receives the assigned content from carrier, it sends this acknowledgement message to Cloud provider to confirm that this DTN data dissemination is completed.

Algorithms together with other operations in this hybrid data dissemination model are described in the following sub-section.

5.2 Data Dissemination

Hybrid data dissemination is determined by control plane, as shown in Fig. 8, where Cloud provider in the control plane has the global information to control the data plane. The main data flow control algorithm conducted by Cloud provider is illustrated in Algorithm 1, where Cloud provider checks its global Fog server and content lists to determine the Fog server that needed to be updated, and the required content. It also checks which Fog server has this content. If none of Fog server has this content, Cloud provider sends updated content to that Fog server directly by using traditional Cloud based techniques, such as broadband and cellular networks [14]. Otherwise, the DTN based data dissemination is applied by using Algorithm 2 to choose mobile devices, which are connecting with these selected Fog servers, as carriers to provider DTN based data dissemination services.

The carriers selection is based on their delivery time and delivery probability (Algorithm 3). A pre-determined content delay threshold, T_{delay}, which is an attribute of this content and can also be treated as content delivery priority, is provided by the Cloud provider. Only those mobile devices (carriers) with a shorter delivery time compared with the pre-determined delay time, and a higher delivery probability are selected as potential DTN based data dissemination carriers.

Once these carriers are chosen, Cloud provider sends DTN data dissemination request message to each of the selected Fog servers to ask them send the required content to the Fog server which is needed to be updated. If the content is transmitted to the carried successfully, a DTN data dissemination accept message is sent back

Algorithm 1 Data Flow Control Algorithm

Step 1: Cloud provider compares its "Global Content List" table with "Fog Server List" table to determine which Fog server is needed to be updated. In this case, Fog sever F_d is determined and content C is needed to be updated.
if There is no other Fog server has this content **then**
  Cloud provider sends this content to the Fog server directly, which is the same as traditional Cloud service and this dissemination process is finished.
else
  Move to the next step
end if
Step 2: Cloud provider determines a list of Fog servers, $< F_{c1}, F_{c2}, F_{c3}, \cdots >$, which have this content.
Step 3: Algorithm 2 is used to select n most suitable carrier, $Carrier_n$, to provide this DTN dissemination service.
if n greater than 0 **then**
  Move to the next step
else
  Cloud provider sends this content to the Fog server directly, which is the same as traditional Cloud service and this dissemination process is finished.
end if
Step 4: Cloud provider sends "DTN Data Dissemination Request" message to each of selected Fog servers to ask them send the content C to F_d by using the carrier (mobile device) determined in **Step 3**.
Step 5: Once a Fog server receives "DTN Data Dissemination Request" message, it sends the content C along with the destination, F_d, to the selected carrier.
if Content C is transferred to the selected carrier completely **then**
  This Fog server sends the "DTN Data Dissemination Accept" message to its Cloud provider and move to the next step.
else
  This Fog server send the "DTN Data Dissemination Decline" message to its Cloud provider. When Cloud provider receives this message, it repeats *Step 3* to get the "n+1" Fog server, if it has, and continue from *Step 4*.
end if
Step 6: Once F_d receives the content C, it sends "DTN Data Dissemination Acknowledgement" message to Cloud provider.
Step 7:
if Cloud provider receives the "DTN Data Dissemination Acknowledgement" message within a pre-defined period, T_{delay}, in Algorithm 2 **then**
  It updates "Fog Server List" and "Global Content List" tables, and this data dissemination is finished
else
  It repeats from the *Step 1*
end if

to Cloud provider confirming this content has been sent out. Otherwise, a DTN data dissemination decline message is sent out. For example, a mobile device (carrier) left the Fog server's coverage area.

When the Fog server, who needs this content, receives the content, it sends a DTN data dissemination acknowledgement message to Cloud provider to confirm it has received the content and this DTN based data dissemination process is finished.

Algorithm 2 DTN Data Dissemination Carrier Selection Algorithm

Step 1: Cloud provider determines the affordable delay time, T_{delay}, of this content. T_{delay} Cloud be treated as the priority of this content.

Step 2: Cloud provider checks its "Fog Server List" and "Mobile Devices's Movement Pattern" tables to find the list of mobile users accessed F_d before, $MobList_d(i)$, and their average connection time with F_d, $Time_d(i)$, where i is the ID of connected mobile device.

Step 3: These mobile devices with a short connection time are filtered out, as they are not able to upload the content to the Fog server:

for each of mobile device in $MobList_d(i)$ **do**

  if

$$\frac{Size_C}{Speed_i} > Time_d(i)$$

  then

    This mobile device is filtered out from $MobList_d(i)$ and a new list $MobList'_d(r)$ is formed, where r is the number of mobile device meeting the above condition

  end if

end for

Step 4: $MobList'_d(r)$ is further classified into two categories, *scheduled* and *non-scheduled* visit lists. Scheduled visit list stores these mobile devices which are pre-determined to visit a particular Fog server, such as airport shuttle bus. The rest of filtered mobile devices are classified into the non-scheduled visit list.

Step 5: For scheduled mobile devices, Si, as long as its delivery time, which is the time from now to its next scheduled visit time, is within the T_{delay}, it is added into the DTN data dissemination carrier list, $< Carrier_{S1}, Carrier_{S2}, \cdots, Carrier_{Sx} >$, where x the total number of carriers selected to add into the carrier list.

Step 6: Non-scheduled mobile device list, NSi, is re-ordered by mobile devices' delivery probability to F_d based on Algorithm 3. Cloud provider select the top y mobile devices according their delivery probabilities to add them into the DTN dissemination carrier list, where the number of y is the largest number to satisfy the following condition:

$$\frac{\sum_{i=1}^{x} DeliTime_{Si} + \sum_{i=1}^{y} DeliTime_{NSi}}{x+y} < T_{delay}$$

End: Now $x + y$ mobile devices are selected as carriers to provide DTN data dissemination service.

Otherwise, Cloud provider needs to re-select mobile nodes as carriers or directly sends the content using traditional method. Detailed hybrid data dissemination processed are illustrated in Algorithms 1–3, and notations used in these three algorithm are explained in Table 2.

6 Future Research Topics of Fog Computing in 5G

Based on the Mobile-Fog-Cloud hierarchy shown in Fig. 9, we envision potential research directions from the communication efficiency's viewpoint as follows.

Algorithm 3 Mobile Device Delivery Probability

Step 1: For each of mobile devices, m, Cloud provider collects its contact frequency with F_d, $ConFre_m$, geographic locations and visit times of the three most recently visited Fog servers, $Loc_m < Lan, Lon, T >$.
Step 2: Based on the three most recent visited history, $Loc_m < Lan, Lon, T >$, and real distance from the map, average movement speed and direction of mobile device m could be generated as $Speed_m$ and $Direction_m$.
Step 3: The expected delivery time from mobile device m to Fog server F_d, $DeliTime_m$, is calculated by using $Speed_m$, $Direction_m$ and the geographic distance between both them.
if $DeliTime_m > T_{delay}$ **then**
  The delivery probability of this mobile device, $DeliProb_m$, is marked as 0 and this algorithm is finished
else
  Move to next step
end if
Step 4: Assume there are n suitable mobile devices left in this step. The overall delivery probability of mobile device m is calculated as:

$$DeliProb_m = \frac{ConFre_m}{\sum_{i=1}^{n} ConFre_i} \times (1 - \frac{DeliTime_m}{\sum_{i=1}^{n} DeliTime_i})$$

and each of them is added into the delivery probability list, $DeliProbList[n]$.
Step 5: Sort $DeliProbList[i]$ in ascending order
Set $u = 1, j = n$
while $u \leq n$ **do**
  while $j > u$ **do**
    if $DeliProbList[j-1] > DeliProbList[j]$ **then**
      swap($DeliProbList[j-1]$, $DeliProbList[j]$)
    end if
    $j--$
  end while
  $u++$
end while
This delivery probability list is ready to be used for Algorithm 2.

6.1 Communications Between Mobile and Fog

Note that a Fog server manages 3-D resources including storage, computing and communication. The service quality acquired by users relies on the collective performance of resource utilization from all the three dimensions. Moreover, as Fog computing typically provides pre-defined application services and targets to specific user groups, the service-oriented resource allocation customized to the specific deployment environments is thus necessary. For example, considering the on-board Fog computing system inside the inter-state bus as in Fig. 3, three types of traffics may coexist including video streaming, gaming and web surfing delivered through the same Fog server. As such, a cross-layer MAC design at the Fog server can be devised based on the application's information. Considering that Fog servers have limited storage and deliver limited localized services only, another key design issue is how to optimally select the desirable information contents to cache at each Fog server and determine the appropriate service applications which cause the least service failure rates to mobile users. The solution needs to consider the predictable pattern of mobile service requests, available storage and compute power of a Fog server.

The Fog computing can also be incorporated with the 5G cellular networks. In this case, by making the cellular base station a Fog server with on-board storage and compute facility, the entire Fog system can provide greater coverage and dedicated services to cellular users.

6.2 Communications Between Fog and Cloud

The cloud performs two roles in the integrated Fog computing system. First, the cloud is the central controller of Fog servers deployed at different locations. With each Fog server focusing on the service delivery to mobile users at specific locations, the cloud manages and coordinates the geo-distributed Fog server clusters at different regions. Second, the cloud is the central information depot. The Fog servers at different locations select the information contents from the cloud and then deliver the copied contents from its cache to the mobile users. With above two roles, the design goal of the communications between fog and cloud can be twofold: (1) how to enable the reliable and scalable control of Fog servers at the cloud; and (2) how to develop the scalable data routing scheme from cloud to Fog server for content updates.

Note that the dual functions of cloud as stated above well match the architecture of a software-defined networking (SDN) [15, 16], which decouples the traffic routing to the control plane and data plane. It is thus promising to apply the SDN scheme for the control of Fog computing.

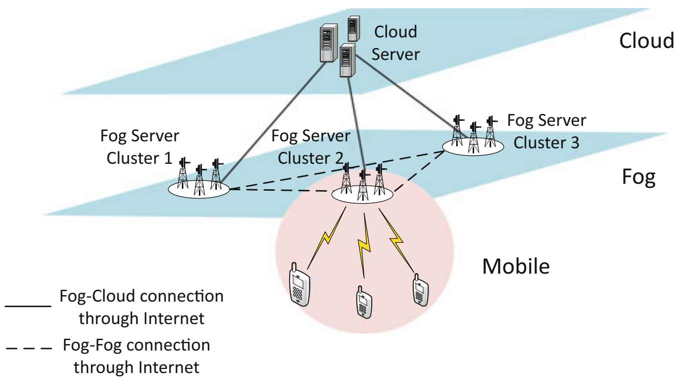

Fig. 9 Mobile-fog-cloud architecture of fog computing

Table 2 Notation used in Sect. 5.2

Notation	Definition
F_d	A fog server needed to be updated (the destination of data dissemination)
C	A content needed to be updated and disseminated
F_i	The fog server i
$Carrier_i$	A content carrier (mobile device) i to provide DTN based data dissemination service
T_{delay}	The maximum affordable delay time of a content. It is a time period from now to its must updated time, such as a shopping center promotion video must be released on every Wednesday
$MobList_d(i)$	The *ith* mobile device (carrier) in the mobile list attached with fog server d
$Time_d(i)$	An average connection period between mobile device i and fog server (d). It is a maximum time window used to upload/download a content to a fog server
$Size_C$	The size of content C
$Speed_i$	The wireless transmission speed of mobile device i
$Loc_i < Lan, Lon, T >$	The geographic location vector of mobile device i visited at latitude Lan and longitude Lon on time T
$Direction_i$	The expected movement direction of mobile device i
$DeliTime_i$	An average content delivery time of Carrier i to the destined fog server
$ConFre_m$	A contact frequency between the mobile device m and its destined fog server
$DeliProbList[n]$	A list to store all carriers (mobile devices) delivery probability to the destined fog server

6.3 Internet-of-Things Applications

As Fog servers are deployed at the physical spot close to mobile users and can be equipped with sensors, it is convenient to incorporate the Fog computing with the Internet-of-things applications. Bonomi [1] and Stojmenovic and Wen [17] present the examples of adopting Fog computing in the applications of smartgrid, vehicular networks and sensor networks.

7 Conclusion

This book chapter introduced the Fog computing under 5G environment.

This article presents Fog computing, a new networking frontier dedicated to serving mobile users. By deploying reserved compute and communication resources at the edge, Fog computing absorbs the intensive mobile traffic using local fast-rate connections and relieves the long back and forth data transmissions among cloud and mobile devices. This significantly improves the service quality perceived by mobile users and, more importantly, greatly save both the bandwidth cost and energy consumptions inside the Internet backbone. Therefore, Fog computing represents a scalable, sustainable and efficient solution to enable the convergence of cloud-based Internet and the mobile computing. The purpose of this article is to investigate on the major motivation and design goals of Fog computing from the networking's perspective. We emphasis that the emergence of Fog computing is motivated by the predictable service demands of mobile users, and Fog computing is thus mainly used to fulfill the service requests on localized information. As a Fog server possesses hardware resources in three-dimensions (storage, compute and communications), the 3-D service-oriented resource allocations are therefore the key of Fog computing. Moreover, with the three-tier Mobile-Fog-Cloud architecture and rich potential applications in both mobile networking and Internet-of-things, the Fog computing also opens broad research issues on network management, traffic engineering, big data and novel service delivery. Therefore, we envision a bright future of Fog computing.

References

1. F. Bonomi, R. Milito, J. Zhu, S. Addepalli, Fog computing and its role in the internet of things, in *Proceedings of ACM MCC* (2012), pp. 13–16
2. R. Balan, J. Flinn, M. Satyanarayanan, S. Sinnamohideen, H.I. Yang, The case for cyber foraging, in *Proceedings of ACM SIGOPS* (2002), pp. 87–92
3. M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The case for vm-based cloudlets in mobile computing, No. 99 (2011)
4. Location-Based Mobile Services Have Huge Untapped Potential Worldwide (2012)
5. G.M. Relations, Greyhound launches BLUE, an exclusive Wi-Fi enabled onboard entertainment system. https://www.greyhound.com/en/newsroom/viewrelease.aspx?id=528&year=2013. Accessed on Dec 2014
6. T.H. Luan, L.X. Cai, J. Chen, X. Shen, F. Bai, Vtube: towards the media rich city life with autonomous vehicular content distribution, in *Proceedings of IEEE SECON* (2011), pp. 359–367
7. G. Peng, CDN: content distribution network (2004). arXiv preprint cs/0411069
8. B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman, A survey of information-centric networking. IEEE Commun. Mag. **50**(7), 26–36 (2012)
9. E. Baştuğ, M. Bennis, M. Debbah, Living on the edge: the role of proactive caching in 5G wireless networks. IEEE Commun. Mag. **52**, 82–89 (2014)
10. M. Satyanarayanan, Z. Chen, K. Ha, W. Hu, W. Richter, P. Pillai, Cloudlets: at the leading edge of mobile-cloud convergence, in *Proceedings of MobiCASE* (2014)
11. Y. Zhang, Y. Zhou, Transparent computing: a new paradigm for pervasive computing, in *Ubiquitous Intelligence and Computing: Third International Conference* (Springer, Berlin/Heidelberg, 2006), pp. 1–11
12. T. Komine, M. Nakagawa, Fundamental analysis for visible-light communication system using LED lights. IEEE Trans. Consum. Electron. **50**(1), 100–107 (2004)

13. C.X. Lin, X. Shen, J.W. Mark, L. Cai, Y. Xiao, Voice capacity analysis of WLAN with unbalanced traffic. IEEE Trans. Veh. Technol. **55**(3), 752–761 (2006)
14. K. Zheng, F. Hu, W. Wang, W. Xiang, M. Dohler, Radio resource allocation in lte-advanced cellular networks with m2m communications. IEEE Commun. Mag. **50**, 184–192 (2012)
15. H. Kim, N. Feamster, Improving network management with software defined networking. IEEE Commun. Mag. **51**, 114–119 (2013)
16. K. Zheng, L. Hou, H. Meng, Q. Zheng, N. Lu, L. Lei, Soft-defined heterogeneous vehicular network: architecture and challenges. CoRR (2015). abs/1510.06579
17. I. Stojmenovic, S. Wen, The fog computing paradigm: scenarios and security issues, in *Proceedings of FedCSIS* (2014)