Exploiting Mobile Big Data: Sources, Features, and Applications

Xiang Cheng, Luoyang Fang, Xuemin Hong, and Liuqing Yang

Abstract

The worldwide rollout of 4G LTE mobile communication networks has accelerated the proliferation of the mobile Internet and spurred a new wave of mobile applications on smartphones. This new wave has provided mobile operators an enormous opportunity to collect a huge amount of data to monitor the technical and transactional aspects of their networks. Recent research on mobile big data mining have shown its great potential for diverse purposes ranging from improving traffic management, enabling personal and contextual services, to monitoring city dynamics and so on. The mobile big data research has a multi-disciplinary nature that demands distinct knowledge from mobile communications, signal processing, and data mining. The research field of mobile big data has emerged quickly in recent years, but is somewhat fragmented. This article aims to provide an integrated picture of this emerging field to bridge multiple disciplines and hopefully to inspire future research.

Introduction

In the past, non-structured data fragments have usually been considered as useless by-products merely to facilitate the proper flow of structured data. The purpose of big data is to piece together such data fragments to gain insight on user behaviors, and to reveal underlying routines that may potentially lead to much more informed decisions. Drastically different from traditional practice, where a service determines and defines the data, in the big data era, data is becoming a standalone entity that may drive and even create new services.

The worldwide rollout of 4G LTE mobile communication networks has accelerated the proliferation of mobile Internet and spurred a new wave of mobile applications on smartphones. This new wave has provided mobile operators an enormous opportunity to collect a huge amount of data to monitor the technical and transactional aspects of their networks. Compared to the so-called 4V characteristic (volume, variety, velocity, and value), mobile big data is distinct in its unique multi-dimensional, personalized, multi-sensory, and real-time (MPMR) features. The practical usage of such data, however, has rarely gone beyond network management and billing. It has been recently recognized that mobile big data (its overview is shown in Fig 1) could well be an under-exploited gold mine from various perspectives.

Recent research on mobile big data mining has shown its great potential for diverse purposes ranging from improving traffic management and enabling personal and contextual services to monitoring city dynamics and more. The unique value of mobile big data comes from its ubiquity and context richness. On one hand, it has been evident that mobile Internet not only offers traditional services running on the fixed Internet, but enables a wide range of new applications that allows the Internet to be involved in almost every aspect of modern life. On the other hand, compared to fixed Internet, mobile Internet traffic carries a much richer context, which pinpoints the time, location, activity, social relationship, and surrounding environment of a mobile user.

Today, GPS is becoming part of the default configuration of any mobile device, rendering location information readily available. Even in the lack of exact location information when GPS is not enabled, the coarse location can still be inferred from the *network-level* data. The location information can enable a great variety of applications to provide personal service (context-aware recommendation, next location prediction-based traffic time estimation, etc.), and to assist public service study and planning (e.g., traffic flow analysis, transportation plan, city zone recognition). As smartphones are equipped with multiple sensors, personal behavior can be learned and monitored. Activity recognition is essential for healthcare applications [1]. Also, the usage pattern of smartphones could be utilized to learn the mental status of users [2]. Furthermore, mobile data can in turn provide good information to optimize resource allocation for communications networks (paging efficiency, future data rate prediction, resource deployment, etc.).

Mobile big data research has a multi-disciplinary nature that demands distinct knowledge from mobile communications, signal processing, and data mining. The research field of mobile big data has emerged quickly in recent years, but is somewhat fragmented. This article aims to provide an integrated picture of this emerging field to bridge multiple disciplines and hopefully to inspire future research.

The article is organized as follows. We analyze the data available to mobile operators. We introduce the unique features of mobile big data. We overview various applications that exploit these data. Summarizing remarks are presented.

Data Sources and Acquisition

Mobile data is usually divided into two categories [3]: one is the *app-level* data directly from sensors of mobile phones collected by mobile app vendors. As sensor technologies are ubiquitously equipped in a smartphone (e.g., GPS,

Xiang Cheng is with State Key Laboratory of Advanced Optical Communication Systems and Networks, Peking University, and National Mobile Communications Research Laboratory, Southeast University.

Luoyang Fang and Liuqing Yang are with Colorado State University.

Xuemin Hong is with Xiamen University.

Digital Object Identifier: 10.1109/MNET.2017.1500295NM

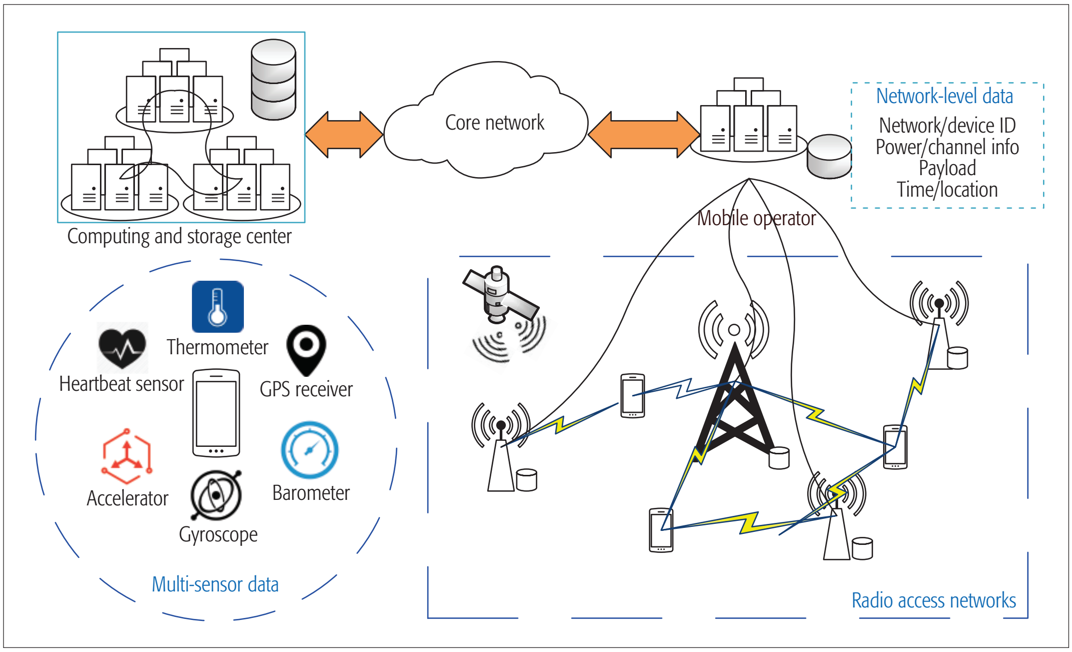

FIGURE 1. Overview of mobile big data.

accelerometer, magnetic field, gyroscope), the latter usually acts like a sensor hub with enriched connectivity for data collection and transmission. The other is the *network-level* one traditionally collected by mobile telecoms, which involves vast amounts of spatio-temporal mobile broadband data about their customers. This type of data records the service requests and information of users (user ID, location, device type, timestamps, type of service, etc.). This section introduces the diverse types of data available to mobile operators.

Data from the Radio Access Network

In the radio access network, data mainly comes from the interaction or measurement between the mobile terminal and the base station. Their interactions involve cell search, synchronization, link establishment, uplink and downlink data transfer, handover, and system information broadcast. These lead to the exchange of a variety of information involving multiple network layers. Correspondingly, the data could be classified into the following categories:

- Network and device identity
- Power/channel information
- Payload and transmission
- Time information
- Location information

Data from the Server

At the server side, the actual data depends on the protocol involved. For example, the DNS server collects users' domain access queries and responds with the corresponding IP addresses. Under FTP, the incoming request corresponds to users' file requests, and the outgoing is the corresponding responses. For HTTP, a user is allowed to access any type of data. Under Simple Message Transmission Protocol (SMTP), service requests, such as authorization, sender, and receiver, are exchanged between the sending and receiver servers.

In the radio access network, data mainly comes from the interaction or measurement between the mobile terminal and the base station. Their interactions involve cell search, synchronization, link establishment, uplink and downlink data transfer, handover, and system information broadcast.

Data from the Mobile Terminal

There have been quite a few projects (summarized in Table 1) focusing on the collection of data from mobile terminals. Reality Mining, carried out by the MIT Human Dynamics Lab over nine months in 2004 was among the earliest efforts. Seventy-five faculty members and students with the MIT Media Laboratory and 25 students at the MIT Sloan Business School participated using 100 Nokia 6600 smartphones. In this experiment, call logs, bluetooth devices in proximity, cell tower IDs, phone status (charging or idle), and popular application usage data were collected. In the more recent Mobile Data Challenge (MDC) by Nokia, 200 volunteers participated using Nokia N95 in the Lake Geneva region from October 2009 to March 2011. Data collected include calls, SMS, photos, videos, application events, calendar entries, location points, unique cell towers, accelerometer samples, Bluetooth observations, unique Bluetooth devices, WLAN observations, unique WLAN access points, and audio samples. Since March 2011, the Device Analyzer experiment, on a much larger scale, involving 12,500 Android devices was carried out the the Computer Laboratory at the University of Cambridge. The covered countries, phone types, OS versions, device settings, installed applications, system properties, Bluetooth devices, WiFi networks, disk storage, energy and charging, telephony, data usage, CPU and memory, alarms, media, and contacts, as well as sensors have been collected and analyzed.

Unique Features of Mobile Big Data

As mobile devices (e.g., smartphones, wearable devices) are becoming central to almost everyone's daily life, mining the sheer volume of data from

Project	Time	Organization	Data Collected
Reality Mining	2004	MIT Human Dynamic Lab	Call logs, Bluetooth devices in proximity, cell tow IDs, phone status, popular application usage data
Mobile Data Challenge (MDC)	2009–2011	Nokia	Calls, SMS, photos, videos, application events, calendar entries, location points, unique cell towers, accelerometer samples, etc.
Device Analyzer Experiment	2011–~	Computer Laboratory at the University of Cambridge	Covered countries, phone types, OS versions, device settings, installed applications, system properties, Bluetooth devices, WiFi networks, disk storage, energy and charging, telephony, data usage, CPU and memory, alarms, media and contacts

TABLE 1. Summary of mobile data collection projects.

The potential of mobile data driven applications could be roughly categorized into two types: one is mining on the individuals data to provide personal service. The other is mining on the aggregation of mobile data to learn and analyze the pattern of human activities.

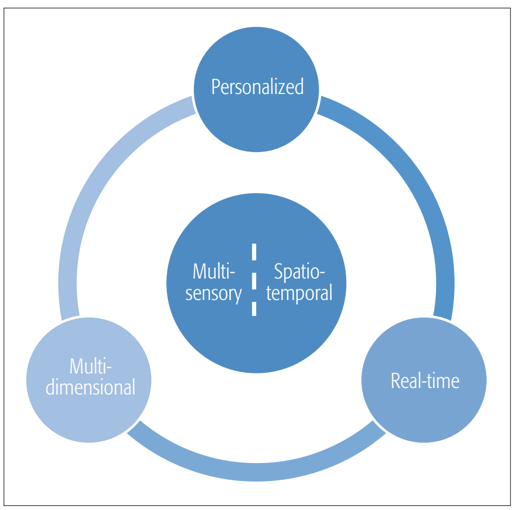

FIGURE 2. Features of mobile big data.

mobile devices has attracted great interest from different research areas, including data mining, signal processing, communications, machine learning, sociology, geography and so on. Besides the 4V characteristics of traditional big data [4], there are distinct features (shown in Fig. 2) of mobile big data.

Multi-Dimensional

As almost every smartphone is equipped with a GPS signal receiver, location and time are distinct characteristics inherent in mobile big data. Even when the location service of a smartphone is not enabled, the rough location of users can still be inferred by the base station location contained in the call data record (CDR) [5]. The multi-dimensional spatio-temporal data provides the information of users' trajectories, which will have a broad range of potential for various applications in location-based social networks, intelligent transportation systems, and urban computing, just to name a few.

Personalized

Mobile data directly collected from user devices or from a mobile network (e.g., gateway, base station) always contains the user identity. Other than the explicit identity information, the mobile data itself is usually highly personalized and relevant to users' location and context. In addition, the trend in mobile big data analytics is that mobile big data are increasingly used not just for analyzing the past or understanding the present, but also for predicting the future [6], which will provide precise personal services (e.g., smart context-aware personal service). However, the data itself and its potential applications will largely reveal and intrude on the privacy of users if it is not protected well, which may in turn jeopardize the availability of data for people who may end up unwilling to share their data.

Multi-Sensory

Almost all smartphones are equipped with a rich set of embedded sensors [4], such as accelerometers, compasses, gyroscopes, GPS signal receivers, and ambient light sensors. In addition, with powerful connectivity for data collection and transmission, smartphones often serve as sensor hubs for wearable sensors (heartbeat, pedometer, etc.). The high-dimensional data from multiple sensors provides vast possibilities and great potential for health monitoring, activity recognition, and context-aware services, but it also inherits some drawbacks from sensor data (e.g., incomplete datasets, missing data, and outliers) due to failure of sensors and communications.

Real-Time

A typical requirement of analytics based on massive mobile data is that the location-based and highly personalized information and services need to be delivered to mobile users in near real time [3]. However, the limitation of computing capability and storage of mobile devices could hardly fulfill the intensive computing demands on massive mobile data. Therefore, mobile cloud computing may be a solution that distributes the computation between the mobile devices and cloud computing [4].

Applications of Mobile Big Data

Analytics and mining on mobile big data with time and location information will provide great opportunities for new services. The potential of mobile data driven applications could be roughly categorized into two types. One is mining on individuals' data to provide personal service (e.g., context-aware recommendation, point of interests, activity recognition). The other is mining on aggregated mobile data to learn and analyze the pattern of human activities, which aims to understand human behaviors in order to help

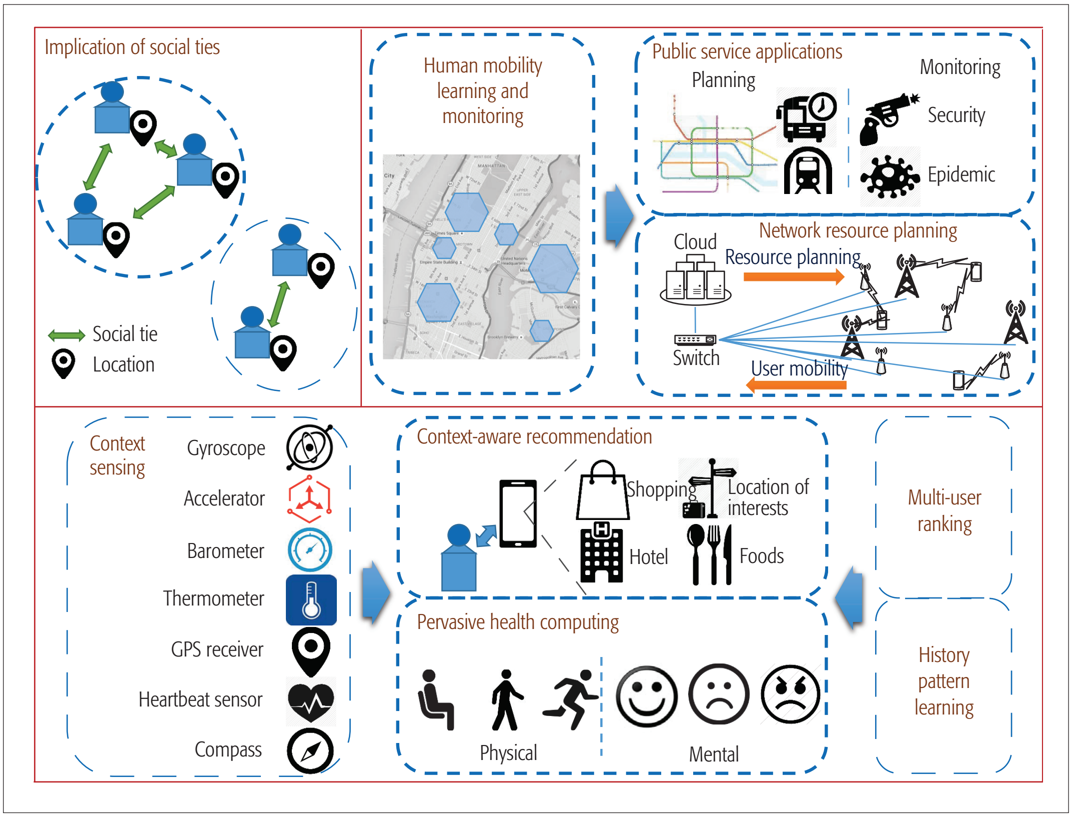

FIGURE 3. Applications of mobile big data.

public service planning and city monitoring (social response after social events or disasters, anomaly detection, traffic flow pattern learning, city zone characterization, etc.). Details of mobile big data applications (shown in Fig. 3) are discussed in the following.

Location Learning

In mobile applications, location information is often used to facilitate various recommendation services. For such, not only is GPS is not available in indoor environments, but also the raw location data alone containing longitude and latitude information is often not quite meaningful for, say, service recommendation applications. As a result, location inference and mapping between raw location data to points of interest, assisted by the mining of mobile big data from various sources, are of great interest. Such efforts include simultaneous localization and mapping based on indoor WiFi , fine-grained location inference based on hashed mobile IP addresses, and the establishment of inter-call user position and mobility models in order to achieve comprehensive spatio-temporal refinement of CDR [5].

Location Prediction and Human Mobility

With semantically meaningful location information, user location prediction is an interesting application of mobile big data. Such interest is further enhanced by the high predictability of human mobility. As suggested in [7], based on the study of 100,000 anonymized mobile phone users whose position is tracked for a six-month period, it was observed that human mobility is highly regularized rather than randomized in terms of both temporal and spatial aspects. In other words, each individual follows reproducible patterns in terms of characteristic travel distance (exploration) and significant probability of return to highly frequented locations (preferential return). In [8], it was further indicated that human mobility is predictable with up to 93 percent accuracy. Based on these, location prediction has been reported for short, medium, and long terms into the future. We expect that the full exploitation of location prediction based on human mobility will potentially sprout a great variety of applications.

> A model integrating mobility and social interactions has the potential to reveal characteristics of the network topology in terms of the total number of connected components, the distance distribution among connected users, the properties of social overlaps, and the distance-based user clusters.

Implications of Social Ties

Intuitively, human mobility is usually influenced by social ties. As a result, physical location and social relationships are intrinsically entangled. In [9], a model was introduced by explicitly considering the feedback of mobility on the formation of social ties using data from Twitter, Gowalla, and Brightkite. A model integrating mobility and social interactions has the potential of revealing characteristics of the network topology in terms of the total number of connected components, the distance distribution among connected users, the properties of social overlaps, and the distance-based user clusters. Exploitation of the implication of social ties are expected to bring two-fold benefits: more regularized location prediction for individuals or social groups with improved accuracy and precision; and the enhanced identification

The context-aware recommendation is a straightforward application of mobile big data with spatio-temporal information. It aims to provide precise personal recommendation to users, given the context of a user and the user's personal preferences.

of social ties and interest groups that in turn facilitate better targeted services.

Context-Aware Recommendation

The context-aware recommendation is a straightforward application of mobile big data with spatio-temporal information. It aims to provide precise personal recommendation to users, given the context of a user (e.g., location, timestamp, velocity, acceleration) and the user's personal preference. Context-aware recommendation is essentially an interaction between the preference of a specific user, and the common interests and ranking of multiple users in a certain context, which results from analytics and mining on mobile big data of a huge number of users. For example, the social role of users is studied on the shared preferences or behavior patterns among multiple users [10], which could be utilized for recommendation or advertisement services. Not surprisingly, nearly all context-aware services require certain location information. The integration of the actual location, the inferred points of interest, the predicted location based on personal location and/or social relationships, and the inferred/refined social roles based on location information, as discussed before, will undoubtedly enhance the consumer experience of context-aware recommendation.

Pervasive Health Computing

With the multi-sensor data from smartphones, physical activity monitoring [1] can be achieved as well as mental health monitoring [11]. For instance, either fall detection coupled with an alert system could be implemented at common commercially available electronic devices to both detect falls and alert authorities, or the fog computing paradigm could be utilized by splitting the detection task between the edge devices and the servers in the cloud.

Although physical activity monitoring can be considered as a modernized integration of biomedical sensors into personal mobile communication devices, mental health monitoring hinges much more heavily on context sensing. Based on where we have been, with whom we communicate, which applications we use, and how we use our mobile devices, various learning algorithms can be developed to exploit these tens of mobile phone sensor values. These learning models can adapt to and predict a mobile user's mood, emotions, cognitive/motivational states, activities, environmental context, and social context.

Besides psychological monitoring and care of patients, such information can be employed for a vast array of health/mood related applications. For example, video and music recommendations, context-aware advertisements, and social networks would all significantly benefit from these.

Public Service Applications

Besides the personal applications and services discussed above, the aggregation of mobile big data will provide a great tool to exhibit the big picture of human being social behaviors, that is, human mobility, or social response and propagation of events, diseases, or disasters. For instance, mobile big data can help reveal regions of different functionalities in urban areas [12] in terms of the pattern of the aggregation of human mobility, based on which the urban planning (e.g., public transportation planning) can be effectively designed. Traffic patterns can be inferred and different traffic zones can be divided based on the mobile big data, such as the CDR. The distribution of travel time from sensors of mobile devices (e.g., accelerator) and GPS trajectories can help monitor traffic conditions and provide the fastest route based on the current traffic status from mobile big data. In addition, the social and pathological response and propagation of events, diseases, and disasters can be studied based on the social stream, physical and emotional sensing, and CDR to help improve public security, as well as emergency response and recovery.

Network Resource Planning

Naturally, based on the aforementioned features, mobile big data will in turn provide excellent information to feed the resource allocation and optimization for communications networks. Mobile big data can be used to extract, model, and predict mobile traffic patterns. The capability of mobile traffic prediction has great potential to guide network planning and resource management. Mobile big data can also be used to gain individual insight on demographic attributes, mobility patterns, personal preferences, and instant context. Such insight can be used to optimize personal content delivery, contextual services, and mobile advertisement. Such knowledge not only can improve the network planning from the traditional perspective in terms of access point planning and flexible radio resource management, but also will be essential for emerging trends, such as adaptive content and cloud distribution, by providing content consumption cartography [13].

In summary, mobile big data with unique features will have unprecedented potential in a great variety of applications for both personal and public services. As described previously, mobile big data is an inter-disciplininary area including statistics, machine learning, high performance computing, signal processing, communications, and so on.

Research Opportunities and Challenges

Mobile big data provides tremendous amounts of raw samples, facilitating new personal and public applications as well as new opportunities in almost every research field. Big data's 4V characteristics and unique mobile features (spacial-temporal, multi-sensory, etc.) will bring great challenges to the current infrastructure of computing, communications, and networking. Therefore, new paradigms supporting data collection, data processing, and data sharing are open research opportunities. In addition, data security and privacy, especially in the mobile scenario, are always important issues, highly related to data collection and storage. Furthermore, the challenge from mobile big data mining is that the applications of mobile big data should be context-aware and need to adapt to dynamic mobile environments.

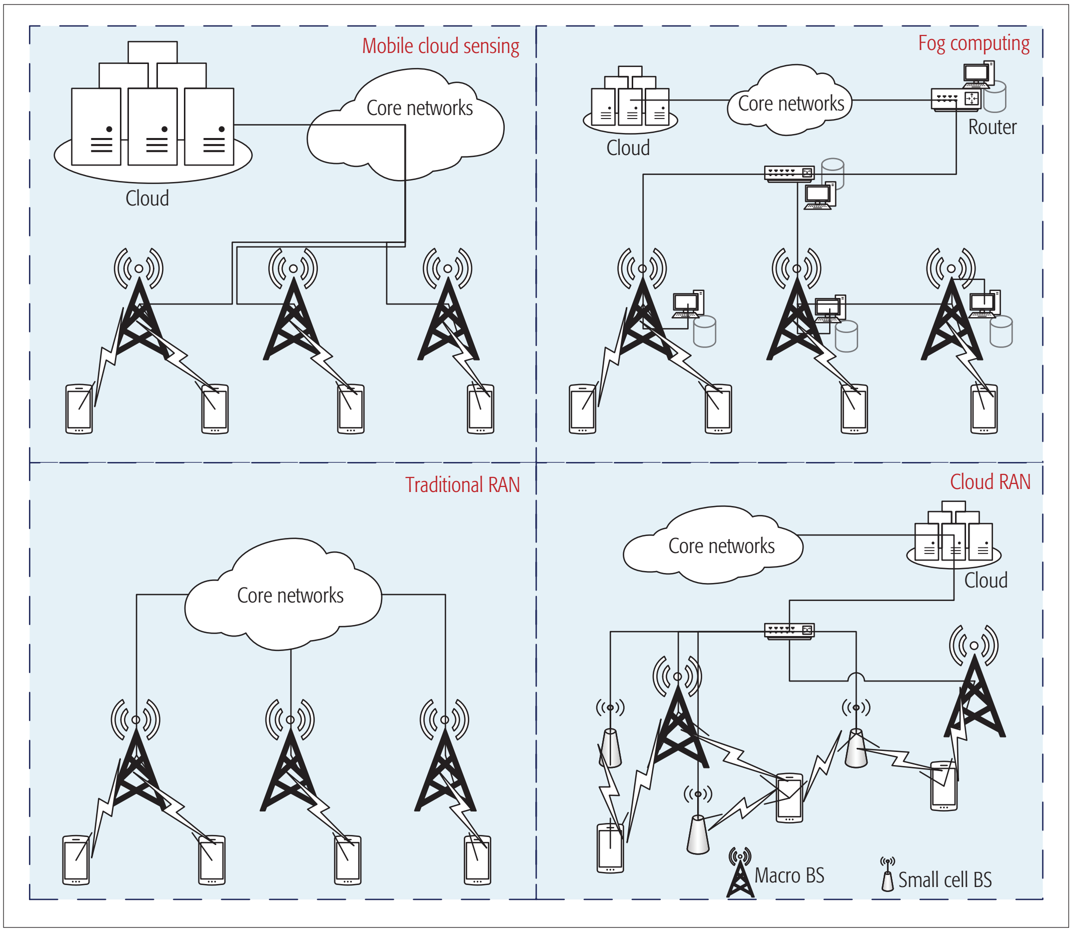

FIGURE 4. Infrastructures for computing, communications, and networks.

Infrastructures for Computing, Communications, and Networks

One major source of mobile big data is from smartphones with multiple sensors, and its applications are mostly using smartphones as the platform to interact with users. However, due to the limited computing resource and memory, mobile data cannot be processed locally, that is, by the smartphone itself. In fact, battery capacity in smartphones is a bottleneck in current smartphone development, which also hinders complex data processing in smartphones. Therefore, infrastructures on computing, communications, and networks (shown in Fig. 4) to support mobile big data are of great importance.

Mobile Cloud Sensing: The concept of mobile cloud sensing [4] is proposed to solve the problem of mobile big data processing, combining mobile sensing, cloud computing, and mobile big data. The intensive computing workload and high-volume data storage demands of mobile big data processing are offloaded to the cloud via wireless communications and network interfaces, including Wi-Fi, cellular, and so on.

With the idea of mobile cloud sensing, the problem of intensive computing and storage demand of mobile big data processing is converted to the communication problem between the mobile devices and the cloud. Actually, the network should be able to handle massive data transmission due to the tremendous volume of mobile big data, as well as massive simultaneous device connection requests.

> One major source of mobile big data is from smartphones with multiple sensors, and its applications are mostly using smartphones as the platform to interact with users. However, due to limited computing resources and memory, mobile data cannot be processed locally, that is, by the smartphone itself.

The current cellular networks cannot satisfy the intensive needs of mobile big data transmission. Furthermore, network delay is a vital factor for mobile big data applications, especially when interaction between the mobile terminal and the cloud is required in real time. However, reducing high latency due to the network delay in the context of mobile big data is a great challenge.

Fog/Edge Computing: In order to reduce the network delay coming from the core network, the concept of fog computing [1] is proposed to bring the computing and storage capability closer to the mobile devices at the edge of the network. In other words, devices located at the edge of the Internet, such as routers, switches, base stations, and access points, will be equipped with computing and storage infrastructures. In fact, fog computing extends the scheme of cloud computing from the core of the network to the edge of the network.

In the context of mobile big data, the fog computing paradigm can deal with data acquisition, data aggregation, and preprocessing, and even data mining, without suffering the high latency as in the paradigm of mobile cloud sensing. However, the heterogeneous computing resource

Mobile big data containing the information of time, location, and even activities of users is always privacy-sensitive. Furthermore, privacy-sensitive mobile big data is located not just in user terminals or remote servers, but everywhere with computing and storage capability in the network.

management in the hierarchical network will bring challenges and provide great research opportunities. As fog computing is generally not powerful enough compared to cloud computing, the interaction and coordination between them is another open research problem.

Although the paradigm of fog computing can reduce the latency from the core of the Internet, the bandwidth and connectivity limitations in the current structure of wireless access networks (especially the widely used cellular networks) still exist.

Cloud Radio Access Networks — The unprecedented mobile big data traffic will bring great challenges to radio access networks (i.e., cellular networks), which are generally used in mobile data collection and transmission. The current network bandwidth and capacity of a radio access network is not able to fulfill the demand of mobile big data applications. Therefore, the radio access network paradigm needs to be revolutionized.

In the traditional radio access network, base stations with limited numbers of antennas can only serve a fixed coverage, which leads to low utilization of network resources in terms of both space and time. In the evolution of the radio access network, small cells are preferred to increase the utilization of spatial spectrum reuse. However, the interference management and coordination in the hierarchical cell structure pose great challenges to researchers. In addition, the computing resource in traditional base stations may not be able to fulfill the demands of dynamic resource management.

The concept of the cloud radio access network [14] is proposed to move the computing-intensive functions (baseband processing and resource allocation) to the backend cloud connected to base stations via high-capacity backhauls. Meanwhile, the function that remains in base stations is RF-level wireless access. In addition, the network device parameters could be reconfigured online, which will provide great flexibility and fully utilization of the network resources for high-capacity and reliable radio access services.

With the decoupling of the wireless access and computing functions, multi-cell joint dynamic resource allocation could be facilitated, according to users' mobility patterns. Furthermore, collaborative radio processing will also be enabled in the centralized computing paradigm of the radio access network, as mobile users could usually be served by multiple small cells. Nevertheless, the design of dynamic resource allocation and collaboration of radio processing to support the real-time high-data-rate applications of mobile big data are open problems.

On the other hand, the computing cloud will be able to learn the behavior of users and even to predict them with the availability of spatial and temporal mobile data characterizing its users. The knowledge learned will in turn provide the information of users to adjust network structure and reconfigure device parameters so that the network performance and quality of service can be optimized in the scenario of a reconfigurable cloud radio access network. However, it is challenging to identify and extract useful features on mobile big data, as well as to discover the relationship connecting mobile user behavior and network performance.

Data Security and Privacy

Mobile big data containing the information of time, location, and even activities of users is always privacy-sensitive. Furthermore, privacy-sensitive mobile big data is located not just in user terminals or remote servers, but everywhere with computing and storage capability in the network (e.g., routers, switches, access points) due to the distributed nature of fog/edge computing. Therefore, how to protect the data and preserve the privacy of users while providing services by collecting mobile big data will post great challenges.

Authentication — User authentication and data encryption are two key factors to ensure the privacy of users. Biometric authentication, such as fingerprint and facial authentication, may be applied to protect privacy-sensitive data. {If the biometric authentication is processed locally, the biometric key is well protected. However, when the biometric authentication process is outsourced to clouds [15], how to protect the biometric key and design an efficient encrypting algorithm are open problems.}

Access Control — The distributed computing and storage resources may serve as infrastructures for commercial leasing in the future. That is, one node with computing and storage resources could be accessed by multiple agents. Then the data life cycle in the distributed storage should be carefully designed to prevent data leakage. In addition, valuable user data would attract great interest in attacking and hacking. How to identify and detect the compromised computing nodes is an interesting issue.

Data Mining and Knowledge Discovery

In fact, almost every application or service based on mobile big data will be developed from raw data to insightful information or knowledge discovery, and from discovered information to decision making. Actually, one may be caught in the trap of simple and misleading correlation between statistics of data, which may lead to false knowledge discovery. The problem of false knowledge discovery requires deep understanding in the domain of application and careful verification and validation.

Moreover, the contradiction between the real-time response requirement of mobile big data applications and extremely large-volume and high-velocity mobile big data not only places great challenges on infrastructures to support the applications of mobile big data, but also brings challenges to the algorithm design on mobile big data mining, which should be scalable and adaptable to dynamic mobile environments of mobile users. In addition, how to extract and select features from multi-sensor mobile big data according to specific requirements of applications or services is of great importance in order to reduce the computing complexity and workload in data collection and data processing.

Conclusions

In this article, we discuss the unique features of mobile big data, and its sources and various applications. Based on all these, we analyze the main research opportunities and challenges in fulfilling the great potential of mobile big data. These include computing, communications and networking infrastructure, data security and privacy, as well as data mining and knowledge discovery. The main purpose is to give a comprehensive overview of this emerging research field with promising potential.

The distributed computing and storage resources may serve as infrastructures for commercial leasing in the future. That is, one node with computing and storage resources could be accessed by multiple agents.

Acknowledgment

This work is in part supported by the National Natural Science Foundation of China (NSFC) under Grant 61622101, 61571020, and 61571378, the National 973 Project under Grant 2013CB336700, the National 863 Project under Grant 2015AA01A708, the open research fund of National Mobile Communications Research Laboratory under Grant 2016D03, Southeast University, and the NSF under Grant CNS1343189.

References

[1] Y. Cao *et al.*, "Distributed Analytics and Edge Intelligence: Pervasive Health Monitoring at the Era of Fog Computing," *Proc. 2015 Wksp. Mobile Big Data*, Hangzhou, China, June 22–25, 2015, pp. 43–48.

[2] R. LiKamWa *et al.*, "Moodscope: Building A Mood Sensor from Smartphone Usage Patterns," *Proc. 11th Annual Int'l. Conf. Mobile Systems, Applications, and Services*, Taipei, Taiwan, June 25–28, 2013, pp. 389–402.

[3] D. Z. Yazti and S. Krishnaswamy, "Mobile Big Data Analytics: Research, Practice, and Opportunities," *Proc. 15th IEEE Int'l. Conf. Mobile Data Management*}, Brisbane, Australia, July 14–18, 2014, pp. 1–2.

[4] Q. Han, S. Liang, and H. Zhang, "Mobile Cloud Sensing, Big Data, and 5G Networks Make an Intelligent and Smart World," *IEEE Network*, vol. 29, no. 2, Mar. 2015, pp. 40–45.

[5] M. Ficek and L. Kencl, "Inter-Call Mobility Model: A Spatio-Temporal Refinement of Call Data Records Using a Gaussian Mixture Model," *Proc. IEEE INFOCOM*, Orlando, FL, Mar. 25–30, 2012, pp. 469–77.

[6] M. Musolesi, "Big Mobile Data Mining: Good or Evil?," *IEEE Internet Computing*, vol. 18, no. 1, Jan. 2014, pp. 78–81.

[7] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding Individual Human Mobility Patterns," *Nature*, vol. 453, no. 7196, Mar. 2008, pp. 779–82.

[8] C. Song *et al.*, "Limits of Predictability in Human Mobility," *Science*, vol. 327, no. 5968, Feb. 2010, pp. 1018–21.

[9] P. A. Grabowicz *et al.*, "Entangling Mobility and Interactions in Social Media," *PLoS ONE*, vol. 9, no. 3, Mar. 2014, p. e92196.

[10] R. K. Wong, V. W. Chu, and T. Hao, "Online Role Mining for Context-Aware Mobile Service Recommendation," *Personal Ubiquitous Computing*, vol. 18, no. 5, June 2014, pp. 1029–46.

[11] M. N. Burns *et al.*, "Harnessing Context Sensing to Develop a Mobile Intervention for Depression," *J. Medical Internet Research*, vol. 13, no. 3, 2011.

[12] S. Grauwin *et al.*, "Towards a Comparative Science of Cities: Using Mobile Traffic Records in New York, London, and Hong Kong," *Computational Approaches for Urban Environments*, ser. Geotechnologies and the Environment, M. Helbich, J. Jokar Arsanjani, and M. Leitner, Eds., Springer, Nov. 2015, vol. 13, pp. 363–87.

[13] S. Hoteit *et al.*, "Content Consumption Cartography of the Paris Urban Region Using Cellular Probe Data," *Proc. 1st Wksp. Urban Networking*, Nice, France, Dec. 10–13, 2012, pp. 43–48.

[14] S. Bi *et al.*, "Wireless Communications in the Era of Big Data," *IEEE Commun. Mag.*, vol. 53, no. 10, Aug. 2015, pp. 190–99.

[15] H. Chun *et al.*, "Outsourceable Two-Party Privacy-Preserving Biometric Authentication," *Proc. 9th ACM Symp. Info., Computer and Commun. Security*, Kyoto, Japan, June 4–6, 2014, pp. 401–12.

Biographies

Xiang Cheng [S'05, M'10, SM'13] (xiangcheng@pku.edu.cn) received his Ph.D. degree from Heriot-Watt University and the University of Edinburgh, United Kingdom, in 2009, where he received the Postgraduate Research Thesis Prize. He is currently an associate professor at Peking University. His general research interests are in the areas of channel modeling and communications. He was the recipient of the IEEE Asia Pacific Outstanding Young Researcher Award in 2015, and Best Paper Awards at IEEE ITST '12, ICCC '13, ITSC '14, and ICC '16. He has served as Symposium Leading-Chair, Co-Chair, and a member of the Technical Program Committee for several international conferences. He is now an Associate Editor of *IEEE Transactions on Intelligent Transportation Systems*.

Luoyang Fang [S'12] (Luoyang.Fang@colostate.edu) received his B.S. degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2011, and is now pursuing his Ph.D. degree in the Department of Electrical and Computer Engineering, Colorado State University. His research interests include big data, distributed storage systems, network coding, rateless erasure coding, and information-centric networking.

Xuemin Hong [S'05, M'12] (xuemin.hong@xmu.edu.cn) received his Ph.D. degree from Heriot-Watt University in 2008. He is currently a professor with the Department of Communications Engineering at Xiamen University, China. His research interests include cognitive radio and green radio networks. He has published more than 40 technical papers in major international journals and conferences.

Liuqing Yang [S'02, M'04, SM'06, F'15] (lqyang@engr.colostate.edu) received her Ph.D. degree from the University of Minnesota, Minneapolis, in 2004. She is currently a professor with the Department of Electrical and Computer Engineering at Colorado State University. Her main research interests include communications and signal processing. She has been actively serving the technical community, including the organization of many IEEE international conferences, and on the Editorial Boards of a number of journals, including *IEEE Transactions on Communications*, *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Intelligent Transportation Systems*, and *IEEE Transactions on Signal Processing*. She received the Office of Naval Research Young Investigator Program Award in 2007, the National Science Foundation Career Award in 2009, the IEEE GLOBECOM Outstanding Service Award in 2010, the George T. Abell Outstanding Mid-Career Faculty Award at Colorado State University in 2012, and Best Paper Awards at IEEE ICUWB '06, ICCC '13, ITSC '14, GLOBECOM '14, ICC '16, and WCSP '16.