Towards Energy-Efficient Wireless Networking in the Big Data Era: A Survey

Xianghui Cao, *Senior Member, IEEE*, Lu Liu, *Student Member, IEEE*, Yu Cheng, *Senior Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*—With the proliferation of wireless devices, wireless networks in various forms have become global information infrastructure and an important part of our daily life, which, at the same time, incur fast escalations of both data volumes and energy demand. In other words, energy-efficient wireless networking is a critical and challenging issue in the big data era. In this paper, we provide a comprehensive survey of recent developments on energy-efficient wireless networking technologies that are effective or promisingly effective in addressing the challenges raised by big data. We categorize existing research into two main parts depending on the roles of big data. The first part focuses on energy-efficient wireless networking techniques in dealing with big data and covers studies in big data acquisition, communication, storage, and computation; while the second part investigates recent approaches based on big data analytics that are promising to enhance energy efficiency of wireless networks. In addition, we identify a number of open issues and discuss future research directions for enhancing energy efficiency of wireless networks in the big data era.**

***Index Terms*—Wireless networks, big data, energy efficiency, data acquisition, data communication, data storage, data computation, machine learning, open issues.**

I. Introduction

WIRELESS communication networks in various forms (e.g., cellular networks, wireless local area networks (WLANs), wireless personal area networks (WPANs), wireless sensor networks (WSNs) and vehicular ad hoc networks) have been developing rapidly. In contrast to wired networks, wireless networks offer conspicuous convenience for ubiquitous and efficient data communication, leading to their growing market share and making them important elements of our daily life. For example, many people spend a long time every day for Web-browsing, on-line instant messaging and information sharing with their mobile phones. Advances in wireless networks also bring the possibility of new applications such as mobile social networking and crowdsensing that are almost impossible with wired networks. In both industry and academia, wireless network technologies keep being developed and will remain a hot research area for a long time.

With the proliferation of wireless devices (e.g., mobile phones, wireless sensors, wireless smart meters, unmanned vehicles and drones), wireless networks become a global information infrastructure incurring a fast escalation of data volume known as big data. For example, today's smartphones are equipped with various sensors such as camera, audio, accelerometer, GPS, gyroscope, compass, and ambient light sensors, making each smartphone a big data source. Nowadays, tremendous amounts of big data traffic are brought by broadband downloading, social connections and content sharing, online business and entertainment, behavior monitoring, health sensing, distributed storage, computing services, and cloud radio access infrastructure, and so on [1]. As predicted by Cisco, there will be 11.6 billion mobile devices by the year 2020 (i.e., 1.5 per capita) and an average smartphone will generate 4.4 GB data per month [2]. The total traffic of mobile big data per month will grow to about 30.6 Exabyte (30.6×10^{18} bytes). Such high-volume and high-generating-speed big data impose significant burdens on wireless networks in terms of both networking paradigms and every big data handling stage including big data acquisition, communication, storage, and computation, which have attracted a large amount of research and development efforts recently [3]–[5].

At the same time, the energy expenditure of wireless networks is huge. For example, Verizon consumes 8.9TWh energy which amounts to 0.24% of the total energy consumption of the U.S. [6]. The energy consumption of communication infrastructures grows exponentially [7], which means fast increasing capital costs for both network operators and end users and vast environmental expense. It is estimated that around 3% of global energy expenditure and 2% of global CO_2 emissions are from information and communication technology, among which mobile and wireless networks take 57% of the energy consumption [8]. It was predicted that, in the year 2020, networks and related infrastructures will contribute around 320 Million tons of CO_2 emissions, where around half of them pertain to mobile communications [9]. In the

Manuscript received January 14, 2017; revised August 9, 2017; accepted October 29, 2017. Date of publication November 8, 2017; date of current version February 26, 2018. This work was supported in part by the U.S. National Science Foundation under Grant CNS-1320736 and Grant ECCS-1610874, in part by the National Natural Science Foundation of China under Grant 61573103 and Grant 61628107, in part by the State Key Laboratory of Synthetical Automation for Process Industries, and in part by the Fundamental Research Funds for the Central Universities under Grant 2242016K41068.
(Corresponding author: Xianghui Cao.)

X. Cao is with the School of Automation, Southeast University, Nanjing 210096, China (e-mail: xhcao@seu.edu.cn).

L. Liu and Y. Cheng are with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: lliu41@hawk.iit.edu; cheng@iit.edu).

X. Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/COMST.2017.2771534

big data era, the rapidly developing and wide spreading wireless networks are going to consume more energy in the future, urging worldwide researchers and engineers to consider how to lower the energy consumption in every networking scenario seriously.

With the energy problem becoming a globally urgent issue, a significant trend of wireless networks is to evolve to "green" ones with high energy efficiency (EE) in a wide variety of aspects including CMOS level re-engineering and green technologies for big data acquisition, communication and computation [3], [7]. The revolution for energy-efficient wireless networking (EEWN) is also spurred by the practice of energy-constrained big data generating devices [10], [11], e.g., battery-powered smartphones and wireless sensors. In addition, protocols and mechanisms that efficiently utilize renewable energy (e.g., solar and wind energy) can reduce the cost of energy supplied by power lines [12], [13].

While saving energy is a critical demand, it should not sacrifice too much performance in order to support big data applications. In other words, the key problem for wireless network energy efficiency is the tradeoff between energy consumption and achieved performance (e.g., throughput or quality of service (QoS)). Technologies towards EEWN have established a long research history where most studies focused on energy-efficient data communications. Since the wireless medium is open, interference is a big obstacle lies in the way towards better performance and a significant source of energy waste. The essence of many energy-efficient technologies is to exploit the network spatio-temporal diversities in order to mitigate interference [14], [15]. For instance, by properly scheduling the network devices to transmit at different time, data delivery reliability can be improved while the energy waste due to collisions can be reduced. Meanwhile, the power supply can be gradually decaying for battery-powered devices or time varying for energy-harvesting devices. The computation, storage and communication capabilities of a wireless device are also limited. These constraints raise critical challenges for wireless network resource allocation and call for sophisticated techniques in order to balance energy consumption and performance. Recently, owing to the widespread of wireless networks and the emergence of new wireless and networking technologies (e.g., device-to-device (D2D) communication, software-defined network (SDN) and cloud), energy efficiency remains hot and has attracted a significant amount of research efforts.

In the big data era, aside from data communication services, today's wireless networks also play an active and important role in data acquisition (e.g., WSNs), storage (e.g., wireless data-center networks (DCNs) and mobile networks with caching) and computation (e.g., mobile cloud networks). In these scenarios, enhancing the energy efficiency of the wireless networks encounters many obstacles such as the large scale, rapid generation and high diversity features of big data. For example, how to gather spatio-temporally correlated data with low energy cost, how to cache data within wireless networks to support timely data querying and retrieving, and how to exert collective power of resource-constrained wireless devices in performing large-scale big data computing. Moreover, in such wireless networks, low-complexity and distributed decisioning algorithms that offer fast data delivery and processing are preferred. In the case of social big data generated from thousands of smartphones where each user's quality of experience (QoE) matters, more flexible wireless network architectures and user-centric networking paradigms are demanded.

This paper presents a comprehensive survey of recent advances in EEWN in the big data era. We investigate existing achievements in EEWN for big data applications as well as promising technologies and opportunities that can be applied in future EEWN, with an attempt to manifest the fundamental techniques in enhancing energy efficiency of wireless networks. Specifically, the survey is divided into two main parts depending on the roles of big data, i.e., EEWN for big data (N4B) that focuses on energy-efficient techniques in handling big data sets, and big data for EEWN (B4N) that accounts for big data based learning methods offering the opportunities of improving energy efficiency of wireless networks. We focus our major attention on studies in recent 5-6 years and cover many emerging technologies such as cognitive radio networks (CRNs), future cellular networks integrated with D2D communications, mobile social networks, crowdsensing networks, cloud networks, and SDNs. In addition, we present issues and challenging problems that remain open to encourage future research studies.

In the literature, energy efficiency issues of wireless networks have been studied for a long time, and a number of surveys have been published [14]–[24] (as compared with this work in Table I). Many of them focus on either specific wireless networks such as cellular networks and WLANs or specific network protocols such as routing, medium access control (MAC) or physical layer protocols. On the other hand, some recent surveys have discussed the wireless networking aspects of big data [3], [5]. However, to the authors' best knowledge, there is short of a systematic survey of recent developments for EEWN in the big data era.

The remainder of this paper is organized as follows. Section II overviews the energy efficiency problem and challenges due to big data. We cover technologies in the N4B category in Sections III-VII and B4N technologies in Section VIII. Section IX discusses open issues and Section X concludes the paper.

II. OVERVIEW

Big data come from a large variety of longitudinal and distributed sources, in which wireless big data that are generated and handled in wireless networks contribute to an important portion. Typical sources of wireless big data are mobile data downloading, mobile social networking, mobile business, distributed storage and computing, smart grid communications, vehicular networking, and Internet of Things (IoT) applications [1], [3]. In this section, we present the architectures of wireless networks in the big data era and the induced challenges in designing EEWN schemes. In addition, we present a classification of energy-efficient technologies towards N4B

TABLE I
COMPARISON OF THIS WORK AND OTHER RECENT RELATED SURVEY PAPERS

Ref.	Target networks	Main approach or contributions	Taxonomy of energy-efficient techniques	Coverage		
				EEWN	Big Data	B4N
[15]	cellular networks	selectively turns base stations into sleep mode to save energy	types of enabling techniques (user association, self-organizing network, cell breathing, and heterogeneous deployment)	√	—	—
[17]	cellular networks	energy-efficiency on selected topics (e.g., MIMO, OFDM, and CoMP)	radio resource management, network deployment, and cross-layer optimization methods	√	—	—
[18]	cellular networks	investigates sources of network energy inefficiency and corresponding energy-efficient solutions	types of solutions, including constant envelope OFDM, power amplifier, transmission modes and link rate adaptation, traffic-adaptive cells, relays and cooperation, multi-antenna, sleep scheduling, etc.	√	—	—
[19]	multi-cell cellular networks	allocates resources to handle inter-cell interferences and maximize EE	types of networks (homogeneous, heterogeneous, and cooperative networks)	√	—	—
[20]	mobile cellular networks	energy-efficient solutions and analytical models for both network operators and mobile users	on/off scheduling in low traffic load scenarios, and scheduling (e.g., small cells and power supplies) in high load scenarios	√	—	—
[21]	CRNs	energy-efficient techniques for CR-based spectrum management and CR-based wireless networking	energy-efficient spectrum sharing methods, CR-based wireless access networks, energy-harvesting and green energy based CRNs	√	—	—
[22]	5G radio access networks	radio interference and resource management to save energy in 5G networks	types of energy-saving techniques, i.e., energy harvesting, dynamic power saving, cooperative communications, cross-layer design, and joint design solutions	√	—	—
[14]	cellular networks and WLANs	mainly to schedule the operation time of network components to save energy	five criteria: cellular or WLANs, metrics, online or offline, centralized or distributed, and evaluation method	√	—	—
[23]	cellular and Wi-Fi networks	energy-aware multimedia reception and delivery techniques for mobile devices	Internet protocol layers (physical, link, application and cross-layer)	√	—	—
[24]	WLANs, WSNs, cellular networks, CRNs	trade-off between EE and aspects such as SE, throughput, delay, stability, routing, and transport efficiency	layers (physical, MAC, network, transport, application) and infrastructure domain	√	—	—
[5]	both wired and wireless networks	investigates networking aspects for big data representation, and techniques for supporting big data such as scheduling, modeling and queueing	—	—	√	—
[25]	WSNs	machine learning techniques for solving various functional and non-functional issues of WSNs	machine-learning based energy-aware approaches for data acquisition, event detection, MAC and routing protocols	√	—	√*
This work	WLANs, WSNs, cellular, mobile cloud, crowdsensing networks, etc.	energy-efficient techniques in the big data era	N4B techniques in big data acquisition, communication, storing and computing processes and B4N techniques with big data resources	√	√	√

* The machine learning methods can be applied for B4N, although the paper does not focus on big data.

and B4N. We summarize the main abbreviations used globally throughout this paper in Table II.

A. Architectures of Wireless Networks in the Big Data Era

A direct extension of traditional wireless networking structures to a paradigm that embraces big-data generating devices, components for data computing and storage and considerations on energy efficiency is illustrated in Fig. 1.

The bottom tier contains various types of hardware that generate and collect big data at any time, any location and in different manners. In big data applications powered by wireless networks, sensors (in various forms of wireless sensors in WSNs, body sensors, and smart meters) and smartphones are two major types of devices that generate big data. Smartphones are undoubtedly big-data devices, as they have become the main tools of mobile communications. As for sensors, although the amount of data generated by each sensor may be insignificant, the overall traffic of large-scale WSNs will form important source of big data [26]. Above these and other data generating devices are wireless access networks such as cellular networks, WLANs and WPANs. Through them, big data can be uploaded to or downloaded from the Internet; and lying behind the Internet are DCNs. Private networks, e.g., enterprise private cloud, may allow WLANs or WPANs to directly connect to data-centers. Across the tiers, big data should be effectively collected, stored, delivered and processed where data computation can be adaptively carried out at local devices or remote centralized units in order to support big data applications [1]. The overall system should be able to balance between achieved performance in terms of QoS, users' QoE and system's energy efficiency.

However, in traditional networks, routers and switches with embedded control are self-closed systems offering little

TABLE II
MAIN ABBREVIATIONS

Abbreviation	Meaning
EE/SE	energy efficiency / spectral efficiency
EEWN	energy-efficient wireless networking
QoS/QoE/QoP	quality of service/experience/protection
N4B	EEWN technologies for handling big data
B4N	big data based technologies for improving EEWN
WLAN	wireless local area network
WPAN	wireless personal area network
WSN	wireless sensor network
VANET	vehicular ad hoc network
CRN	cognitive radio network
IoT	Internet of Things
SDN	software-defined network
DCN	data-center network
MAC	medium access control
SINR	signal-to-interference-plus-noise ratio
D2D	device-to-device
MIMO	multi-input multi-output
OFDM	orthogonal frequency-division multiplexing
BS / AP	base station / access point
MR-MC	multi-radio multi-channel
ML	machine learning

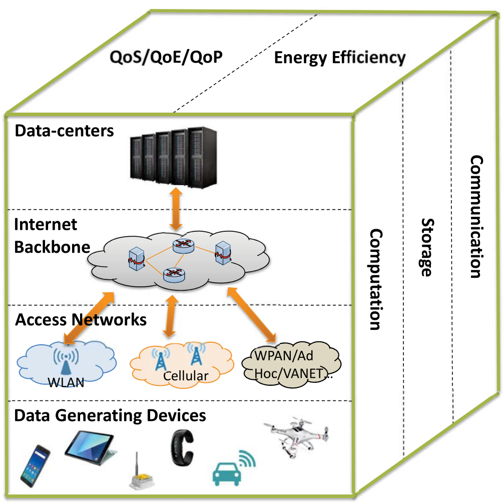


Fig. 1. Architecture of wireless networks in the big data era.

opportunity to deploy different communication protocols and control services. To address the infrastructure closure and ossification, a new network paradigm coined as SDN has been proposed in which the network data plane and control plane are decoupled to allow flexible and programmable control such that new protocols and applications can be implemented more easily [27]. Although SDN was firstly defined for Internet systems, many efforts have been devoted toward software-defined wireless networks (SDWN) including OpenRadio and OpenRoad systems and SDWN based on various wireless technologies such as LTE cellular networks, WLANs, WPANs and WSNs [28], [29]. An overview of SDWN is illustrated in Fig. 2.

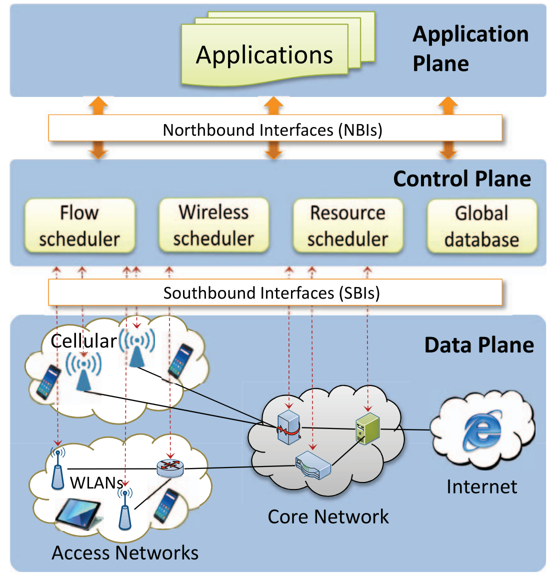


Fig. 2. A conceptual architecture of SDWN [31].

Due to its flexibility, programmability and controllability, the SDN technology can improve the energy efficiency of wireless networks in the following aspects. First, the separation of control and data makes it convenient to deploy low-cost and energy-efficient protocols in wireless networks. Second, SDN promises a convenient way to reduce protocol operation overhead as well as network management and maintenance overhead. For instance, routing decisions made by controllers can be carried out as flow rules in the flow tables of nodes, such that the application-specific overhead of normal nodes is reduced [30]. Third, controllers can have global (or semi-global) view of the network information, which will benefit network energy optimization through joint resource allocation (to be discussed later in Section IV-E).

B. Conventional Energy Efficiency (EE) Concepts

A wireless node may be powered in different ways that generally fall in following three categories: (1) Constant supply ensures a constant power input, e.g., BSs in cellular systems and APs in Wi-Fi networks that are directly connected to power lines. (2) Power-constrained supply such as capacity-constrained batteries will continuously decrease as the node is on. (3) Renewable energy supply varies according to external renewable energy input. Advanced wireless devices can harvest environmental energy such as solar and wind to support data processing and communication activities. In these scenarios, the environmental supply may vary from time to time [13], [32]. For devices with constant supply, although we do not need to worry about energy shortage, their energy efficiency is critical since they usually contribute to the most significant portions of energy consumption in corresponding wireless networks. For example, in cellular networks, it has been shown that the BSs consume 80% of the total energy; hence reducing the energy cost at BSs is in the kernel of improving energy efficiency of such networks [33]. On the other hand, energy efficiency of terminal devices is also important in the viewpoint of end-users. For distributed networks with energy-constrained nodes, the energy efficiency of an average node characterizes the whole network's energy efficiency.

Since energy efficiency is a compound of network performance and energy expenditure, whether a network can be considered energy-efficient depends on how we treat the two aspects together. Most existing studies on EEWN can be categorized into the following three classes, which respectively correspond to three optimization paradigms.

- *Type I (e.g., [34] and [35]):* bits per Joule maximization. The most popular definition of energy efficiency is the network capacity/throughput per unit of energy, which is also known as "bits per Joule" [34], [35]. With such a definition, both capacity and energy expenditure are combined together as a single function with clean physical meanings. However, since the function is nonlinear basically, its concavity and consequently the existence and/or uniqueness of an optimal solution are generally not guaranteed.
- *Type II (e.g., [36]):* performance guaranteed energy minimization. From the viewpoint of network QoS or users' QoE, the most energy-efficient way is to provision a required level of QoS/QoE utilizing energy as low as possible [36], resulting in minimization problems with QoS/QoE constraints. Moreover, as the security and privacy concerns of big data applications are becoming more and more critical, quality of protection (QoP) [37] will be another metric of network performance.
- *Type III (e.g., [38] and [39]):* performance optimization under energy constraint. In some energy-constrained cases, such as battery-powered smartphone networks and energy harvesting networks, the problem of energy efficiency concentrates on delivering best performance (e.g., in terms of throughput or users' QoE) under a certain energy budget or limited energy recharging rate.

C. Challenges of EEWN in the Big Data Era

The significance of big data is manifested in many aspects far beyond the large-scale feature. Below is a popular 4Vs model for characterizing big data:

- **Volume** captures the large amount of data.
- **Velocity** means that big data not only are generated at a high speed but also should be processed, transported and analyzed at a high rate that matches the speed of data generation in order to deliver certain QoE to users.
- **Variety**: data generating devices are highly diverse which results in high data variety in terms of various modalities such as audio, video and SMS and also in terms of various types such as structured, unstructured, semi-structured and mixed data [40].
- **Value** that is conveyed in big data sets usually outweighs the data amounts. Therefore, it is important to explore and exploit the huge hidden value in big data sets to improve users' QoE, help business and to improve network security [3].

From a big data analytics point of view, the above characteristics incur critical challenges such as how to extract meaningful information from unstructured big data sets, how to design an efficient file system for large-scale mixed data so that they can be timely retrieved, how to manage mixed big data, and how to efficiently execute big data analytics [40]. However, these problems, though important, have little to do with wireless networking and hence lie out of this paper's scope. Instead, we focus on EEWN technologies to handle and utilize big data where the major challenges can be summarized as follows.

- *Energy-efficient information acquisition:* In most cases, it is not only cost-expensive but also unnecessary to collect all the redundant and highly correlated raw data from their generating devices. Instead, we are more caring about the value hidden behind. That is, how to efficiently gather information of value with low energy expenditure and low data amounts or using less devices in both static and mobile wireless networks.
- *Energy-efficient data communication:* As the primary task of wireless networks, data communications are obviously under great pressure due to high volumes and velocities of big data. How to deliver big data in a timely and energy-efficient manner is the key challenge, which calls for energy-optimized allocations of network resources such as power, time and spectrum. In addition, in view of the variety feature, data may be assigned different priorities, which introduces service differentiation concerns in the design of communication protocols.
- *Energy-efficient data storage:* Storing large-scale data demands data-centers for which efficient resource management can save a great amount of electrical energy. Meanwhile, although normal devices have limited storage spaces, it is possible to utilize their own resources to cache large-scale data collectively inside the wireless networks instead of transporting the data to remote hardware. Such a way is particularly helpful but also challenging for applications (e.g., mobile social networks [4]) when data are frequently accessed.
- *Energy-efficient data computation:* Within the scope of wireless networks, we focus on big data computation by taking advantage of the collectiveness of the networked devices. Therefore, the challenge is how to accomplish big data computation goals with less energy under that each participating device has only limited ability.
- *Data-driven energy efficiency optimization:* Depending on theoretical models of target wireless networks, existing energy efficiency optimizations depend on model accuracy and are inflexible and non-adaptive to environment changes, yielding a considerable theory-practice gap. In contrast, optimizations driven by big data of network measurements are able to learn the practical network operation status by means of machine-learning algorithms and utilize energy more efficiently and adaptively.

D. Organization

In the following, we investigate recent research and techniques for EEWN that are effective or promising to address the above listed challenges. Corresponding to the above challenges, the survey is divided into two main parts: EEWN

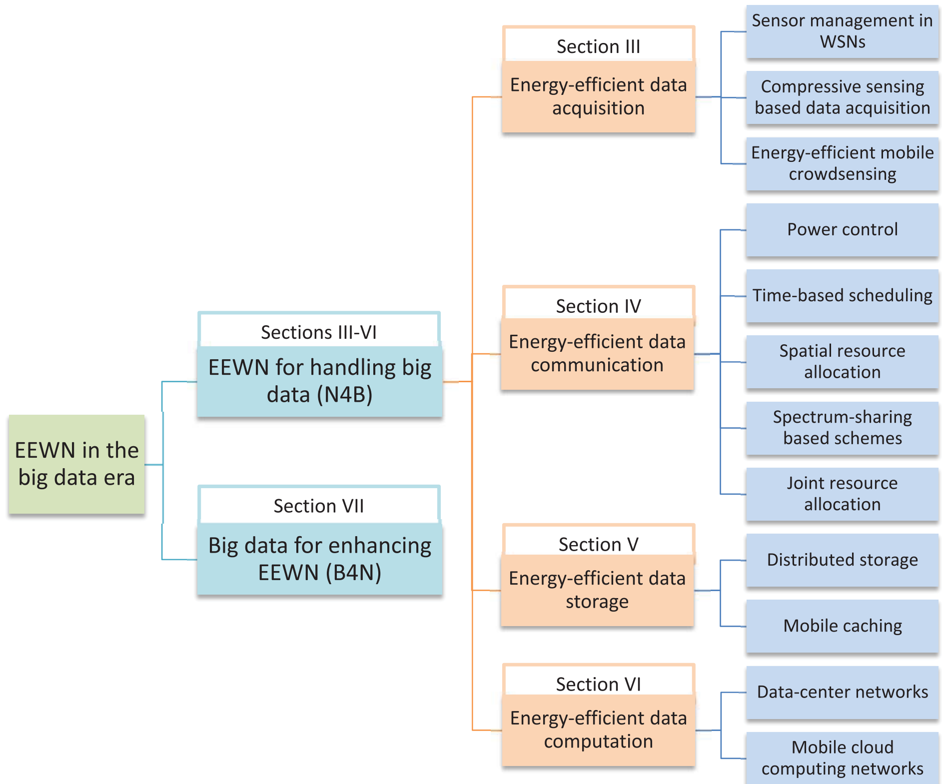


Fig. 3. Overview of EEWN techniques in the big data era.

technologies for handling big data (N4B) and big data based technologies for EEWN (B4N). The N4B part reviews recent studies on EEWN for collecting, storing and processing wireless big data. It is further divided into four sections (i.e., Sections III–VI) roughly following a typical procedure of how big data are handled in wireless network environments, i.e., from big data acquisition, communication, storage to computation, each of which corresponds to each of the first four challenges discussed above. The B4N part in Section VII reviews recent studies on learning based approaches that effectively utilize big data sets for improving (or promisingly applicable in improving) the energy efficiency of wireless networks. An outlook of our categorization of the surveyed techniques is shown in Fig. 3.

III. Energy-Efficient Data Acquisition

This section investigates recent technologies on energy-efficiently acquiring big data at source devices, which is the first step of N4B. Data generated from different devices may be redundant and mutually correlated. Therefore, energy efficiency of wireless networks in the acquisition stage can be achieved by scheduling wireless devices and reducing redundant data in view of their spatio-temporal correlations. As aforementioned, sensors and smartphones are two major types of devices that generate big data. In the following, we shall focus on WSNs and mobile crowdsensing networks.

A. Sensor Management in WSNs

In WSNs, sensors are deployed mainly to monitor (typically in periodic manners) its ambient environment and detect events such as environment changes, target movements and human body activities. During this process, coverage is an important concern, which ensures that regular or abnormal events can be detected with high probabilities. The definition of coverage depends on specific applications and usually can be categorized into coverage of points of interest (PoIs), regions of interest (RoIs) and barriers [41]. The goal of PoIs (or RoIs) coverage is to ensure all the PoIs (or RoIs) are within the sensing ranges of activated sensors. Barrier coverage aims to monitor a boundary region such that the undetected probability of an intruder penetrating through the region is small enough.

To support big data acquisition in WSNs, coverage and connectivity should be ensured at the network planning stage. Intuitively, it may be expected to deploy a large number of sensors in WSNs for ubiquitous data acquisition, but at the cost of high interference and energy consumption. In [42], the placement of static sensors to minimize total network energy consumption or maximize network lifetime was investigated. Unlike static sensor network deployment which requires human participation, mobile sensor networks usually can achieve the above deployment requirements automatically based on dynamic sensor replacement [43], [44]. Given an initial deployment, mobile sensors can apply a distributed virtual force to achieve an even deployment in the way that close sensors repel while faraway sensors attract each other. Song *et al.* [44] proposed several distributed deployment algorithms based on virtual forces among mobile sensors to optimize overall network energy efficiency in terms of $\alpha \frac{\varphi_C}{\varphi_U \varphi_E}$, where φ_C, φ_U and φ_E represent the metrics for coverage, network lifetime and sensors moving energy, respectively.

To further conserve energy at the network running stage, the mainstream approach for energy-efficient coverage is to schedule the sensor activities such that only a subset of sensors are activated at each time while others keep in sleep mode to save energy. The idea behind sensor selection is that sensors are often deployed with high density so that commanding redundantly covering nodes to enter their low power modes can conserve energy without significantly affecting the performance of coverage. The problem of sensor scheduling for energy-efficient coverage can be formulated as follows [45]: Given a set of sensors $\mathcal{S} = \{s_1, \ldots, s_n\}$ each of which has a limited initial energy, find a set of connected subsets of the sensors $\{\mathcal{C}_t | \mathcal{C}_t \in \mathcal{S}, t = 1, 2, \ldots\}$ that maximizes the network lifetime (or minimizes energy consumption) while guaranteeing a desired level of coverage. Since this problem is NP-hard basically, existing approaches usually resort to iterative methods for finding the optimal solutions. In [45], an ant colony optimization algorithm was designed which takes three types of local and global pheromones to search for the optimal active sensor sets where probabilistic sensor detection models were considered.

In applications such as dynamic event capturing, the PoIs are not necessary to be always covered by active sensors; in fact, sensors can be activated only when some event happens within their sensing ranges. Given the knowledge of stochastic properties of the event dynamics, sensors can coordinate their sleep and wakeup activities periodically in either a synchronous or an asynchronous manner to save energy while guaranteeing a certain event detection probability [46]. Full coverage of a large RoI requires a large number of sensors. However, in cases that small coverage holes are acceptable, we can turn off many sensors with the remaindering sensor network still satisfying coverage requirement. In [47], sensor activation for trap coverage was studied where the goal was to maximize network lifetime while guaranteeing that the diameter of each uncovered hole is no greater than a given threshold (thus, any possible intruders will be either detected or trapped within some small areas).

In dense networks, the spatial diversity is a rich resource to be utilized jointly with other resources. For example, in WSNs for environment monitoring, Nikolov and Haas [48] explored spatial correlation among sensor nodes and divided them into subgroups, each of which was responsible for a predetermined interval of measurement. A sensor only transmits a true indicator (1bit) to the BS if its measurement falls within its dedicated interval, thus saving a significant amount of energy than simply sending a full measurement packet.

B. Compressive Sensing Based Data Acquisition

In response to the high volumes of big data, data compression and redundancy reduction are effective technologies to alleviate the burdens on big data acquisition [3]. Compressive sensing is an important one of such technologies applicable in WSNs for data acquisition. With compressive sensing, for some signals that are sparse in a certain basis, they can be reconstructed from a smaller number of measurements, which offers the opportunity of using fewer sensor measurements to acquire high-dimension signals at good accuracy, and hence saving energy. Specifically, if a signal x of dimension N can be represented by a linear combination of some basis vectors $\{\Psi_i\}$ that are mutually orthogonal, i.e.,

$$x = \sum_{i=1}^{N} c_i \Psi_i = \Psi c, \quad (1)$$

where Ψ is the matrix constructed by $\{\Psi_i\}$ and $\{c_i\}$ are coefficients. If the coefficient vector c has K nonzero elements, x is said to be of K-sparsity. Then, based on the compressive sensing theory [49], x can be reconstructed based on y consisting of $M = O(K \log N)$ measurements:

$$y = \Phi x + \epsilon, \quad (2)$$

where Φ is the sensing matrix and ϵ is the noise during acquisition. Accordingly, for acquiring a high-dimension signal using the compressive sensing strategy, when the signal sparsity level is low, the total energy consumption of a sensor node is much lower than that of conventional acquisition schemes [50]. Roughly speaking, compared with the method that directly acquires x, $1 - \frac{M}{N}$ percent of energy can be saved by using compressive sensing. The smaller the sparsity level K is, the fewer measurements are required. Compressive sensing has been proved effective in image/video processing and has been successfully applied in medical imaging, holography and mobile phone camera sensors [51].

Leveraging the compressive sensing theory in the sense of data acquisition using fewer measurements and accordingly less computing and transmission costs, energy saving in sensor networks can be achieved in many ways. One example is to save energy by taking advantage of temporal correlations of sensor samples. For monitoring a scalar environment information, the work in [52] proposed a three-phase algorithm to tune sensors' sampling rate, where they relaxed the conventional assumption that the sparsity of the monitored signal was known in advance. Each sensor firstly applies a random sampling scheme and the generated data are then processed to recover the monitored signal at a fusion center based on compressive sensing. Then, the fusion center evaluates the reconstruction quality in terms of a metric called RQI and based on this information it commands the sensors to adjust their sampling rates dynamically, in order to ensure the reconstruction error within an appropriate range, as shown in Fig. 4. Based on real world temperature sensor data, Chen and Wassell [52] showed that, the proposed method based on compressive sensing can reduce both the number of required sensor samples and hence the energy consumption of the

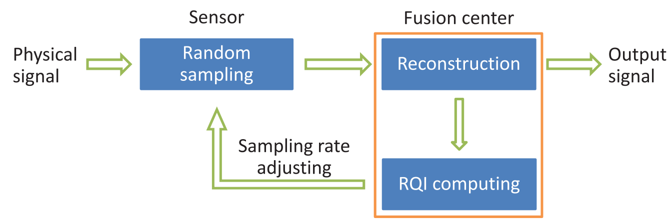


Fig. 4. Compressive sensing based data acquisition [52].

network for environment temperature acquisition by 75%, as compared to the direct sampling scheme without compressive sensing.

Compressive sensing has been successfully applied in collecting wireless big data and has been shown energy-efficient. An energy-efficient wireless data collection framework was proposed in [53], in which the sensors report data to the cloud. The collected data can be viewed as x in Eq. (1) with each element be the sensory data obtained from a particular sensor at a particular time. To save energy, the cloud runs an online learning algorithm that determines the average data collection probability for each sensor, while each sensor applies a local adaptive control law to decide its own collection probability. Specifically, the learning algorithm at the cloud side first predicts the amount of principle data which corresponds to the sparsity K defined below Eq. (1), and then computes the average collection probability P by $P = \frac{K}{N}$ where N is the dimension of the big data x. After receiving P, each sensor adjusts its collection probability based on the data dynamics, unexpected issues, neighbor status, residual energy and link quality. For instance, if the sensed data varies quickly along time, the collection probability is increased.

The work in [54] investigated the quantization effect in compressed sensing and proposed a configurable quantization method to determine sampling rate and quantization rate to improve battery efficiency in body sensor networks. A similar approach was taken in [55] to reduce sensor's average sampling frequency based on a random down sampling matrix, thus reducing energy consumption and extending system's lifetime.

Another example is sparse event detection in a certain area, where the spatial correlation of sensor data is considered. In [56], it was considered that events might randomly happen at a large number of locations termed as sources. If the event generating probabilities at the sources are low, i.e., the events happen sparsely, instead of deploying a large number of sensors with each to monitor a source, such sparsity offers a way of using or activating only a much smaller number of sensors. By applying the compressive sensing strategy, the number of active sensors can be at the same lever of the number of events. Similar as above, simulations demonstrated that the achieved event detection probability increased as the sparsity decreased. Furthermore, a recent work in [57] exploited both spatial and temporal correlations of sensors based on compressive sensing and selectively activated working sensors. The authors developed an active sensors selection approach through minimizing both the reconstruction error and the energy consumption of active sensors. It was shown in this paper that, with the proposed approach, the network lifetime can be significantly prolonged while with a relative low reconstruction error.

C. Energy-Efficient Mobile Crowdsensing

In the emerging mobile crowdsensing systems, a large number of smartphones sense data from their vicinity environment and report to the data analysis center(s) or cloud. The proliferation of sensor-enabled smartphones provides this novel crowdsensing paradigm for real-time and ubiquitous big data acquisition in applications (e.g., gathering real-time population

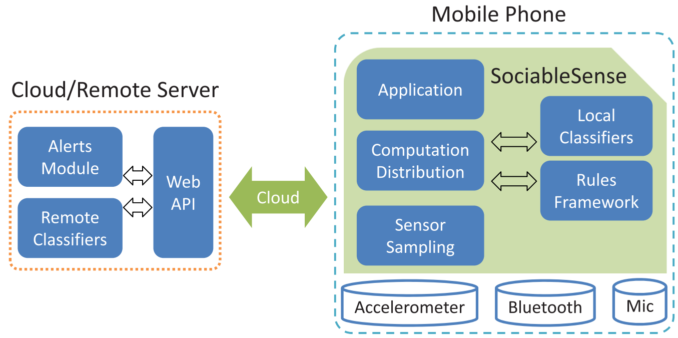


Fig. 5. The SociableSense system [61].

density at bus/subway stations, road traffic congestion/accident information, and urban air quality information) that are too expensive or even impossible by conventional infrastructure based methods. However, both sensing and data communications are significant energy consumers for energy constrained mobile smartphones, which may make users unwilling to engage in too many activities in a crowdsensing system.

Human behavior is perhaps the most important factor that influence the performance of their smartphones participating in a crowdsensing system. Automatic learning, predicting and wisely utilizing the information of human activities and environment conditions can better schedule the smartphone sensors and hence save energy. For example, in location based crowdsensing applications, the simplest approach is always letting the GPS sensors on, in which state the battery of a smartphone may be quickly depleted in hours. However, this always-on operation is unnecessary in practice. A smart GPS scheduling method, called SensTrack, can adaptively adjust the GPS sampling rate and remove unnecessary sampling activities based on the information from acceleration and orientation sensors [58]. When a mobile user is moving without changing his/her orientation, for example, future locations can be estimated based on the motion's inertial property without performing GPS sampling. In a similar idea, the ENRAPT algorithm adaptively decides the sensing rate according to accelerometer readings that indicate whether the user is driving, walking or running [59]. Compared to the always-on operation, ENRAPT can prolong the battery lifetime up to 8 times. Environmental conditions can be also exploited to determine sensor activities. For example, when the light intensity is excessively low, video recording modules may become unnecessary and can be turned off to save energy [60].

As shown in Fig. 5, an adaptive sampling algorithm and a computation distribution scheme were integrated in a smartphone-based social sensing platform called SociableSense [61]. A smartphone dynamically tunes the sampling rates of its sensors (e.g., accelerometer, Bluetooth and microphone) based on the theory of learning automata. Specifically, the sensing probability of each sensor, denoted as p_i, is dynamically tuned as follows.

$$p_i = \begin{cases} p_i + \alpha(1 - p_i), & \text{if senses an unmissable event} \\ (1 - \alpha)p_i, & \text{if senses a missable event} \end{cases} \tag{3}$$

where $\alpha \in (0, 1)$. In the above, an unmissable event means that the sensory data from this sensor indicate some interesting phenomenon in the environment that should not be missed. In this case, the sensor's sampling rate is increased as shown above. Otherwise, a missable event happens (e.g., no motion nor voice), the sensor can tune down its rate until sleep to save energy. The two event types are determined by a classification algorithm and are application dependent. The computation distribution component decides whether locally on the phone or remotely in the cloud to execute energy-consuming sensory data processing. A multi-criteria decision approach was proposed which maximized a combinatory utility function of energy saving, latency and data amount uploaded [61]. Taking smartphone resource changes such as battery charge/discharge cycles and data plan availability into consideration, the solution resulted in a dynamic computation distribution and can well balance smartphone's energy and performance.

Instead of using the high-energy cellular communications for sensed data uploading, smartphones can explore available opportunities of low-cost communication venues such as Wi-Fi. On the other hand, continuously sampling and data uploading—which are demanded in real-time applications—incur the expensive task of handling big data which may be unaffordable to smartphone users with both limited battery supply and limited data plan [62]. However, in some cases that the sensed data are delay-tolerable (e.g., the MIT Reality Mining project that collects user data to analyze their interests and activities does not require real-time operation), users with and without data plans can cooperatively upload data. Such an idea was realized in the effSense framework in which users without data plans can either upload data to the cloud or transfer data to other users using low-cost Bluetooth and Wi-Fi and request them to help forwarding the data [62]. The results showed that, with effSense, users without data plans could upload around 50% data without extra cost.

In situations that a certain sensor or data processing/communication unit is already turned on and the data are already available (e.g., when a user is making a phone call or browsing the GoogleMap), simultaneously carrying out sampling or data uploading for crowdsensing purpose can save a large amount of energy. To exploit such sporadic opportunities, a piggyback crowdsensing scheme based on predictive models of smartphone usage patterns was proposed in [63]. The prediction models enable selection of those opportunities to perform either one of sensing, data uploading and computation. Based on a large-scale data set of over 1000 smartphone users, the experiment results showed that the piggyback scheme could save more than 10% of energy compared with benchmark strategies such as periodic sampling and context-driven sampling.

IV. Energy-Efficient Data Communication

After the data acquisition stage, the communication stage is required to transport big data from their source devices to management and processing units efficiently. In addition, energy-efficient data communication is necessary for wireless network based big data storage and computation, to be discussed later.

For the communication stage in applications such as mobile big data, the high volume and velocity features of big data require wireless networks to have high throughput and low delay, which coincides with the design goals of most resource allocation schemes in wireless networks. In other big data applications such as IoT networks for long-term monitoring, although the volume and velocity of each node's traffic are not high, the aggregated volume of the network through long-term operation will be high [26]. This requires the network to have long lifetime or devices in low energy states can be timely replenished. Therefore, in this section, we shall focus on recent technological advances in energy-efficient resource allocation able to provide high-capacity, low-delay and long-term wireless networking. Although many surveyed works did not explicitly discuss big data, the proposed methods are effective or promisingly effective to support big data communications in wireless networks.

The energy of a wireless device is mainly consumed by the computation and wireless communication components, and a close examination of the communication component reveals that its major energy consumption is determined by its operating power, spectrum and duration in active state. Viewed at the network level, the total energy consumption also depends on the spatial network deployment and the different roles of the devices. In this sense, the key for saving communication energy and improving network energy efficiency lies in wise utilization of several fundamental resources including power, spectrum, time and spatial resources. Corresponding techniques can be categorized into following four schemes.

- *Power control schemes* basically aim to utilize available transmission power more efficiently.
- *Time-based scheduling schemes* aim to exploit temporal diversity to mitigate interference and save energy in low-power modes of network nodes.
- *Spatial resource allocation schemes* seek to activate or assign tasks to the most appropriate devices to improve energy efficiency of the whole network.
- *Spectrum sharing based schemes* allocates spectrum resources including techniques of cognitive radio, D2D, OFDM and general MR-MC networks.
- *Joint resource allocation schemes* simultaneously consider multiple resources for energy-efficient data communications.

Note that the above categorization may yield some overlapping. For example, in networks with spatio-temporal correlations in data traffic, spatial resource allocations are often combined with time-based scheduling considerations. For convenience, some of such cases shall be covered in either the second or the third category in the above depending on whether time or spatial resource plays the most important part in improving network energy efficiency. Others will be considered in the last category.

A. Power Control for EEWN

Power is a basic type of radio resource that greatly affects network connectivity, interferences and SINR. Controlling the power usage of all the involved nodes such as sensors, user equipments and BSs is one of the most direct ways to save energy of wireless networks. Big data communications may involve a large number of distributed devices. Hence, unlike centralized schemes that often require high signaling overhead and have relatively low scalability, distributed and cooperative power control is promising. In this subsection, we discuss recent developments in distributed and cooperative power-based resource allocation schemes. In addition, energy management in energy-harvesting networks is also accounted.

1) Transmit Power Control: Wisely allocating transmit power of each node plays a vital role in improving both network capacity and energy efficiency. Power control has been long studied in the past decades in various wireless networks [64], [65]. Based on different techniques, recent advances in distributed power control are achieved in the following aspects.

- *Convex optimization based distributed power control:* Conventional power control problem is to maximize network utility under constraints of nodes' available power, where the utility (e.g., throughput) is usually a function of effective transmissions, which can be connected to the transmit power of the nodes via SINR. By using the Lagrangian dual technique, the overall problem can be decomposable if the utility is convex, and the optimal (or suboptimal) power control solutions can be obtained by distributively finding the optimal Lagrangian multipliers based on method such as gradient iterative searching [66]. The work in [66] studied hyper-dense small cell networks and proposed a two-phase algorithm based on the Lagrangian dual technique to address the problem of distributed inter-cell power control. This study is particularly useful in the big data era since the deployment of hyper-dense small cell networks is envisioned to significantly enhance user data rate and hence the proposed method is able to handle the high volume and velocity issues of big data and improve network energy efficiency. However, such distributed power control approaches can only (optimally) handle a certain type of problems such as convex optimization, while for others the optimality is difficult to guarantee due to duality gap.
- *Game-based distributed power control:* In a fully distributed network context without centralized coordination, nodes may content for transmissions based on power usage. Modeling the contention as a game, distributed power control based on game theory has been shown to be promising in optimizing network performance and thereby improving energy efficiency. The Nash equilibrium of such a game can provide a distributed power allocation for each node. According to the characteristics of node behavior in wireless networks, distributed power control is often modeled by a regular non-cooperative game where each node decides its transmit power based on optimizing its own utility as a function of SINR, referring to a survey in [67]. In [68], distributed power control in dense femtocell networks was studied, which relates to big data for the reason mentioned above. A non-cooperative game based on potential game theory was established, which was shown to improve the network throughput by 7% and at the same time reduce the average energy consumption by 50% compared to existing methods. Whereas, game-based approaches can achieve fully distributed operations of power control of individual nodes, but often at the cost of sacrificing some energy efficiency due to price of anarchy—the achieved equilibrium points often deviate from the optimal solutions of the energy efficiency optimization problems.

2) Energy Management in Energy Harvesting Networks: Advances in techniques to harvest energy from ambient environment (e.g., solar, wind, vibration, thermal energy and RF radiation energy) and human movements (e.g., finger motion, footfalls and exhalation) have brought to us the energy harvesting wireless devices and networks [12], [13], [69]–[71]. The ability of harvesting energy is also important for the so-called wireless big data [72], where battery limited sensors become able to work continuously to generate a large amount of data.

In these networks, the problem of energy efficiency is often formulated as to maximize performance under energy availability constraints. The transmit power of an energy-harvesting node at any time must be feasible, i.e., all currently available energy is able to support the attempting transmission. Such energy neutrality to keep balance between harvested and utilized energy becomes a critical new constraint for power control in energy harvesting networks [73]. In [39], for throughput maximization under energy neutralization, it was shown that the optimal power allocation policy should try best to keep the power as constant as possible and that the optimal power decreased (increased) only at energy arrivals when the battery was full (depleted). Renewable energy (e.g., solar and wind energy) availability often varies along time and locations, and may be highly stochastic, making it difficult to design deterministic power utilization strategies. Huang and Neely [74] formulated a joint power allocation and transmission rate control problem for network utility maximization, where the harvestable energy followed either independent and identically distribution or Markov process. With the Lyapunov optimization approach, the problem was solved through an on-line transmission scheduling algorithm. Based on the energy queue technique, the design of online algorithms can avoid assuming specific distribution of the stochastic energy arrivals [75]. For both offline and online algorithms, interested readers are referred to the survey in [13].

Recently, wireless RF charging technique is shown to be promising to address the spatial diversity of energy levels by allowing some nodes to wirelessly transfer energy to other energy harvesting nodes [76]. In this field, one of the main research challenges is the charging efficiency optimization, and an effective way to improve the efficiency is to use multi-antenna systems with energy beamforming strategy. Multiple chargers can work collaboratively for energy transferring, where the charging cost can be reduced by minimizing the number of active chargers, referring to the survey in [77].

3) Remarks: One lesson learned from above is that the key strategy for power-based resource allocation is to concentrate power utilization on effective transmissions and coordinate to reduce mutual interferences, during which the temporal and spatial diversities in terms of energy availability and channel conditions should be taken into account. In this sense, the power-based allocation can improve network throughput, thus partially solving the large volume issue of big data. However, due to capacity constraint, excessively large volumes of big data cannot be supported by the network even with power control and other allocation schemes.

In the context of big data, some important issues remain challenging, e.g., distributed power control in dynamic networks, low-complexity algorithms with a certain level of performance guarantee, and fairness preservation among users. These challenges are also common for the other resource allocation schemes, and will be discussed in Section VIII. Another challenge lies in the dynamic power management in networks with hybrid energy supplies, e.g., in heterogeneous cellular networks, a small cell BS can be powered by on-grid energy, renewable energy or both [78].

B. Time-Based Scheduling for EEWN

To support big data communications over wireless networks, it naturally calls for utilizing resources (including time) as much as possible. However, there is still space for time-based scheduling to improve network energy efficiency by reducing energy consumption through alleviating mutual interference and switching idle nodes to sleep modes if not significantly disturbing the network performance. A typical energy consumption breakdown for a wireless card shows that the modes of transmitting and receiving data consume much higher energy than the sleep mode [79]. Therefore, the key challenge of such energy-saving schemes is when a node should transmit or sleep.

1) Contention Control: In IoT and mobile big data scenarios with densely distributed devices, intensified channel contention among sensors or mobile phones may result in low throughput, which is unwanted. For WLANs, IEEE 802.11 defines the Distributed Coordination Function (DCF) as the MAC protocol for contention control, in which the contention window size plays a vital role. From both throughput and energy points of view, the optimal contention window size should be able to reflect the true contention intensity and balance the idle sensing time and the collisions. Since the power for idle sensing is close to the power for transmission, the optimal contention window sizes with respect to throughput and energy efficiency are similar [80]. This shows a promising fact that, for WLANs with DCF, energy efficiency can be improved without sacrificing too much throughput. As long as the channel contention degree can be measured by each contending node, the optimal contention window sizes can be determined and dynamically tuned for higher energy efficiency [81]. The channel access scheme in DCF is essentially random access where each node may have little information of others' states or actions. This motivates people to model the network as a non-cooperative game, where the access probability of a node can be viewed as the strategy taken by itself and a combined objective of both throughput and energy efficiency can be viewed as the utility function. Then with such a game model and by exploring the corresponding Nash equilibrium, the optimal parameters of the protocols can be derived, referring to strategies in [82] and [83].

IEEE 802.15.4 networks are prevalent in WSNs and IoT systems [84]–[86]. For the contention control in such networks, the work in [84] proposed an energy-efficient adaptive algorithm that allows each node to dynamically tune its MAC parameters including backoff window size, backoff times and retransmission times, based on average delivery ratio and loss ratio estimates, in order to keep the network reliability around some application-required level. In this way, channel access contention and retransmissions are appropriately controlled so that the reliability and energy consumption can be better balanced [84].

2) Exploiting Power-Saving Modes: The power-saving modes in networks such as WLANs and WiMAX provide an effective way for saving energy in applications that can tolerate a certain level of delay. For example, in voice-over-IP (VoIP) applications, the VoIP packets may be received before its play out deadline. This offers the opportunity of saving energy by turning off nodes according to their packets spare time before the deadline in multi-media big data applications. With the power-saving mode defined in IEEE 802.11 standard, packets destined to a sleeping node will be temporally buffered at the AP. The node periodically wakes up to contend the channel in order to retrieve buffered downlink packets from the AP or sends uplink packets. In [87], an energy-efficient sleep scheduling mechanism was proposed which aimed at maximizing the sleep time while ensuring packet delay below some tolerable value. The sleep requests from the nodes are coordinated by the AP in order to avoid conflicting downloading periods. A more flexible and dynamic sleeping scheme is defined as Automatic Power Saving Delivery (APSD) in IEEE 802.11e. In the scheduled version of APSD, AP schedules the Service Period (SP) for each mobile node and therefore each node only needs to stay awake during its own SP. By minimizing the possible overlap among SPs, higher energy efficiency can be achieved [88]. A device can be put to sleep mode when the channel is busy transmitting to other devices, as described in transmission opportunity power save mode (TXOP PSM) in IEEE 802.11ac. Considering the time and energy consumption incurred when mode transition, a very short sleep period may not be desirable. In [89], analysis on the achievable energy efficiency was provided and a burst transmission scheme was proposed to overcome this issue.

Similar power saving classes are defined in IEEE 802.16e mainly for reducing energy consumption in applications with real-time traffic (e.g., VoIP). One of the power saving classes adopts exponentially increasing sleep windows that can allow both small and large inactive periods during traffic sessions, suitable for treating traffic bursts. By exploiting the characteristics of traffic, the sleep window parameters can be further optimized to enhance energy efficiency [90]. In view of that the above power saving class can result in unpredictably large latency when the sleep window grows large, another power saving class featuring in periodically alternating between listening and sleep states of relatively short fixed durations. With these two power saving classes, a hybrid scheme can be formed that different classes are applied for silent periods and talk-spurt periods, respectively [91].

3) BS Scheduling in Cellular Networks: In cellular networks, the traffic load may exhibit fluctuations and thus incur energy waste due to under-utilization of BSs. Accordingly, BSs carrying light traffic can be powered off to save energy. In a heterogeneous network scenario, how many small cell traffic BSs can be turned off according to the traffic fluctuations was investigated in [92], and the results indicated that denser deployment of macro-BSs would allow turning off more small cell BSs and hence might improve network energy efficiency. The works in [93] and [94] further investigated the BS operation in heterogeneous cellular networks and the energy efficiency of switching-off macrocell BSs. In [95], a software-defined hyber-cellular networking scheme was designed in which the BSs were classified into control BSs (responsible for control functions such as network access) and traffic BSs (responsible for data communications). In this way, control BSs can make decisions and dynamically switch traffic BSs on/off for significant energy savings.

Cell-breathing is another technique for cellular networks which can improve the energy efficiency by dynamically adjusting the cell size according to traffic demands. In [96], cell breathing was exploited by putting cells with low loads into sleep mode, where the antenna tilting was re-optimized correspondingly such that the energy saving of sleeping would not degrade the network performance. Based on radio-over fiber technology, Gomes *et al.* [97] proposed a dynamic multi-tier architecture for mobile wireless networks where cells at different tiers have different sizes of coverage. With such a framework, the authors proposed to split (or merge) cells by turning on (or off) the BSs according to the traffic demand of mobile users, with objectives such as minimization of BS number, maximization of served user number and energy consumption minimization. Simulation results showed that the energy consumption was reduced by two thirds while the number of served users increased by 17%, compared to a single-tier scheme.

4) Energy-Efficient Throughput-Optimal Flow Control: The problem of transmission scheduling can be generalized to flow control in which link scheduling is jointly considered with flow assignment. Jiang *et al.* [98] formulated the problem of throughput optimization under energy constraint as a non-linear program, and proposed a linear near optimal solution based on piece-wise linear approximation to address the flow control. They further considered the throughput-optimal and energy-minimal problem as a multi-criteria optimization, with which they sought Pareto-optimal solutions and finally achieved a throughput-energy curve, where each point on the curve indicated a weakly Pareto-optimal solution. A generic framework for energy-efficient flow control was formulated in [99], in which the authors developed a multi-objective optimization problem based on multi-commodity flow formulation augmented with scheduling constraints, in order to jointly solve the optimal flow allocation and independent set scheduling that can maximize network capacity with minimal energy consumption. Delay column generation was leveraged to effectively solve the optimization problem, based on which the proposed algorithm can improve energy efficiency with low computation overhead.

5) Remarks: Time-based resource allocation can achieve significant energy savings especially in low-load networks, proving its effectiveness. However, applying these techniques to support big data communications is challenging.

As discussed above, the time-based allocation mainly coordinates the activities of the nodes over time. The above schemes such as scheduling according to traffic demand and the utilization of power-saving classes are able to deal with big data with bursting traffic, which is relating to the velocity feature of big data.

Most time-based allocation schemes are sensitive to network traffic fluctuations that would be common in the big data era. Therefore, efficient schemes should be able to either predict traffic information and take proactive actions or adopt adaptive MAC protocols closely reacting to traffic fluctuations.

Big data are generated in various devices and may be assigned different priorities during communications, corresponding to the variety feature of big data. Although some MAC protocols, e.g., IEEE 802.11e, already can handle priority diversity [100], it remains challenging in the design of priority-aware sleep scheduling and flow control for energy efficiency optimization.

C. Spatial Resource Allocation for EEWN

A wireless network can leverage spatial diversity by designating or electing an appropriate subset of nodes out of all others to perform required tasks with low energy consumption. This involves issues such as node deployment, selection and routing.

1) BS Deployment in Cellular Networks: To increase the cellular network coverage and capacity, more and more BSs are deployed which amount to around 80% of total network energy consumption [101]. To save energy in cellular networks, a pioneering work in [102] investigated the energy efficiency of BS deployment and proposed area power consumption (APC) as a metric of the network performance. The APC is defined as the ratio of total energy consumption of the macro and micro BSs over the cell coverage area and has a unit of Watt per square kilometer. The results showed that the network APC strongly depended on the BS density: both sparse and dense deployment of BSs would result in high APC, while the optimal APC was achieved at a moderate density. Moreover, by adding micro BSs, the network APC can be improved to some extent. It was further convinced in [103] that deploying micro BSs was able to significantly reduce the network APC without sacrificing the network throughput performance. In [104], the optimal density of micro BSs was studied for improving energy efficiency. The authors showed that, if the micro BSs energy cost was lower than a threshold, more micro BSs can be deployed. In these studies, BS density is optimized only on peak traffic load. Wu and Niu [105] proposed an analytical approach based on linear topologies to analyze the energy-optimal BS deployment problem considering the traffic load variations. They showed that their deployment scheme could save more than

20% of energy consumption compared to the schemes focusing on peak traffic load. Macrocell offloading by deploying femtocells is able to improve network energy efficiency [106]. However, dense deployment of femtocells may increase the total network energy consumption instead.

2) User Association in Cellular Networks: In cellular networks, especially for downlink transmissions that contribute to the major portion of mobile big data, a user should be associated with a BS. The cellular network energy efficiency can be improved by selecting appropriate serving BSs for mobile users. In [107], the user association problem was formulated as to optimize a linear combination of flow-level performance (which accounts for delay and traffic loads) and energy consumption. The optimal solution showed that, if the objective function degenerated to pure energy consumption, then the optimal BS for each mobile user was the one that had highest energy efficiency in terms of bits per joule.

3) Energy-Efficient Clustering: Clustering to form backbone links in large-scale spatially distributed wireless networks is an efficient way to save energy, since the transmission distances of ordinary nodes are significantly reduced. Selecting the optimal cluster heads is one of the key issues for energy-efficient clustering. Prior work has proposed many algorithms for static networks. In [108], a particle swarm based multi-objective optimization algorithm was proposed to search for the Pareto-optimal solutions, where the objectives included the number of clusters, energy consumption and a network load-balancing factor. The proposed algorithm can be applied in mobile ad hoc networks—it finds cluster heads and their associated members and update the particle positions iteratively by taking node mobility into account.

Clustering can also save energy in large-volume big data collection with a mobile sink in dense WSNs. In [26], a clustering method based on expected-maximization was proposed and the optimal number of clusters that minimizes the energy consumption was derived analytically. The results demonstrated that the proposed method achieved higher energy efficiency than two existing ones.

4) Energy-Efficient Data Routing: Energy-efficient routing in static wireless networks has been extensively studied in the literature (see the survey in [109]). In the following, we shall focus on energy-efficient routing in mobile networks, where static routes may not always exist.

- *Dynamic routing:* Dynamic routing in mobile ad hoc networks aims at dynamically searching and maintaining energy-efficient paths. For example, the routing protocol proposed in [110] first selects from available paths a minimum-energy shortest path between the source and destination pair during the discovery phase. Then, based on the selected path, a dynamic route maintenance mechanism was applied to adjust the actual transmission path with low overhead.
- *Opportunistic routing:* Traditional routing protocols often require the knowledge of path quality before making route decisions, which could be impractical since many factors such as fading, interference and multipath effects can lead to rapid path quality fluctuation. Therefore, opportunistic routing, which allows nodes that overhear the transmission to participate in packet forwarding, becomes an efficient solution. Selecting appropriate forwarders is the key to improving energy efficiency in this case. In [111], the forwarders were prioritized and the optimal list was selected such that the induced total energy consumption for forwarding as well as that for reaching an agreement amount of the potential forwarders was minimized. In [112], considering opportunistic routing, the authors proposed an analytical model for the average total energy consumption for transmitting a packet and for the end-to-end throughput. In their design, forwarders were selected and prioritized recursively based on their resultant normalized energy consumption.
- *Routing in social networks:* With the explosively increasing use of smartphones and tablets, social data sharing is an important source of mobile big data [4]. In order to increase the data delivery probability for end-to-end communications, many studies apply the "epidemic" approach by employing multiple relays for data forwarding [113]; however, the energy consumption could be very high. Taking advantage of the small-world phenomenon, a more energy-efficient routing approach by limiting the number of forwarding hops and using a light-weight relay selection strategy was proposed in [114]. A non-destination node was selected as the relay if the delivery probability of either its own or one of its neighbors was the largest. Such a way of including neighbors into account is based on the "high cluster coefficient" small-world phenomenon that people are likely to make friends with friends' friends. In [115], in order to disseminate the information of mobile users' common interests in multicast aided Pico-cells, five relay selection methods were presented and compared, among which the SSD method (the relays that have shorter distances away from one of the destination mobile users will be selected) was found to outperform others in terms of both delay and energy consumption.

5) Remarks: Existing studies on spatial-based resource allocation have proven their capability of improving network energy efficiency, though their primary goals are usually on spectral efficiency (SE). To apply these methods to support big data applications over wireless networks, there are several challenges as below.

- The methods of node deployment in the network planning stage and selection in the running stage may be combined to provide better solutions for network energy efficiency. On the other hand, most deployment/selection methods are centralized ones that do not suit many big data scenarios (e.g., crowdsoursing) where nodes behavior is decided individually.
- Many existing path/relay selection schemes rely on the channel state information (CSI) and assume it remains unchanged in a frame time. However, such memoryless channel assumption may not hold, especially in a mobile environment. Alternatively, [116] considered a Markov channel model and proposed a distributed relay selection policy. After the handshaking between source and destination, candidate relays will compete by CSMA to

broadcast their candidate index (indicating their residual relay energy and the channel condition). The source node will select the one with the smallest index to cooperate, among all the relays that successfully send their indices to the source.
- Dense wireless networks will likely generate big data. The network density is an important factor for the network energy efficiency. Dense networks may have low energy efficiency due to high interference and significantly increased energy cost. Improving the energy efficiency of dense networks requires jointly allocation power, time, spectrum and spatial resources.

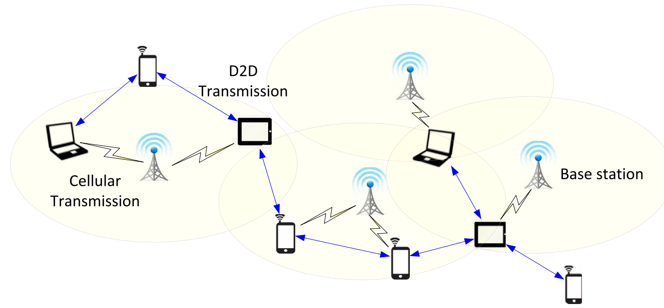


Fig. 6. Device-to-Device (D2D) communications.

D. Spectrum Sharing Based Schemes for EEWN

The capacity limitation of wireless channels makes it challenging to transport high volumes of wireless big data. The scarcity of spectrum resources and the objective of higher energy efficiency call for more flexible spectrum utilization schemes. Basically, technologies such as CRNs, D2D communications and OFDMA networks are primarily designed to improve the SE for higher capacity, and hence providing effective ways for delivering large-volume big data. In this subsection, we briefly discuss recent developments in improving the energy efficiency of such spectrum-based schemes.

1) Cognitive Radio Networks (CRNs): In order to exploit the underexploited spectrum resource (e.g., TV white bands), the cognitive radio (CR) technique allows unlicensed users (i.e., secondary users or SUs) to dynamically access the licensed spectrum when licensed users (i.e., primary users or PUs) are temporally absent. With CR, the optimization of spectrum utilization will improve the network throughput and, as a side effect, its energy efficiency. However, it is also noteworthy that an extra amount of energy is consumed in performing spectrum sensing and channel switching by CR nodes, which raises the tradeoff problem between spectrum sensing and transmission [117]–[119]. The solution is usually derived based on the outcome of spectrum sensing and an SU can adaptively select the operation mode among sensing, switching to another channel, and transmission [117]. Moreover, aside from scheduling between sensing and transmitting states, the problem is often coupled with deciding channel sensing order and sensing duration in order to achieve optimal energy utilization [118].

2) D2D in Cellular Networks: D2D communications underlaying cellular networks (see Fig. 6) can significantly increase network performance by better utilizing radio and spectrum resources. D2D communications can be applied to offload cellular traffic in mobile big data sensing [120].

Since D2D communications share the same spectrum with cellular communications, proper resource allocation for such spectrum sharing is a critical issue for D2D communications [121]–[124]. D2D communications have the potential to enhance the network energy efficiency by switching between cellular and D2D modes [125]. In [126], a joint resource allocation and mode selection optimization scheme in D2D-integrated OFDMA system was proposed. The optimization problem aimed to minimize total downlink transmit power under the user QoS constraints, by subcarrier allocation, adaptive modulation and mode selection solved from a heuristic scheme. The work in [121] also focused on the energy-efficient mode selection and power allocation for D2D system underplayed cellular networks, but with an exhaustive search of all possible mode combinations. The energy efficiency of all the modes of devices were first obtained and then by searching among all the combinations, the optimal solution was obtained along with the corresponding power allocation. Similarly, in [127], an optimization problem that jointly solves mode selection, scheduling and power control was proposed, where a distributed scheme was utilized that can achieve near optimal performance in terms of energy efficiency and fairness. Mode selection and power control were also considered in [122] with a comprehensive and tractable analytical framework. In the proposed scheme, the authors took both D2D link distance and cellular link distance into consideration, and the effect from D2D communication on cellular network was investigated. In [128], based on the findings that the cellular interface consumes lower amount of energy in connectivity maintenance but higher energy in data transferring than a Wi-Fi interface, the authors employed Wi-Fi in cellular networks to improve the energy efficiency.

3) OFDMA Networks: Orthogonal frequency-division multiple access (OFDMA) divides the spectrum into orthogonal sub-channels and assigns them to individual users so that multiple users can simultaneously transmit data with low interference. By doing so, the SE is improved. To further improve OFDMA-network energy-efficient, both uplink and downlink have been studied. For example, the work in [129] developed low complexity schemes for uplink OFDMA systems, which allocated bandwidth to users to optimize the energy efficiency of the network. Both energy-efficient link adaption and subchannel assignment schemes of low complexity were developed and shown to achieve near optimal energy efficiency where the optimal solutions were obtained by exhaustive searching.

The energy efficiency of downlink OFDMA networks was considered in [130]–[132]. In [130], OFDMA downlink network with a large number of transmit antennas was studied. A non-convex optimization problem was formulated to derive optimal rate adaptation policies as well as antenna and subcarrier allocation to maximize energy efficiency. The work in [131] studied the tradeoff between energy efficiency and spectral efficiency for downlink OFDMA networks with cooperative BSs using a similar optimization approach as above. It used dual decomposition to derive closed-form power allocation solutions that maximize energy efficiency. The tradeoff between energy and spectral efficiency was also analyzed in [132], which proved that EE is quasi-concave in SE. A low-complexity and near-optimal algorithm was developed to achieve the desired tradeoff.

Resource allocation for both uplink and downlink OFDMA cellular networks was considered in [133], which aimed at optimizing energy under QoS constraints. Both the optimal solution and a low-complexity suboptimal solution were derived.

4) General MR-MC Networks: One mathematical generalization of CRNs, OFDMA networks and WiMAX systems is multi-radio multi-channel (MR-MC) wireless networks, where each node has multiple radios that can operate on multiple channels [134]–[136]. Based on the observation that idle radios incur energy waste, the work in [137] proposed an approach to turn off unneeded radios if the network performance is not impaired. Both a heuristic algorithm and a mixed-integer linear program were proposed in the paper to find optimal channel assignment and routing solutions. The optimality of energy efficiency in generic MR-MC networks was exploited in [99], in which the conditions to achieve optimal energy efficiency were analyzed. With this work, the optimal multi-dimensional resource allocation to achieve maximized energy efficiency at full network capacity can be obtained by selecting proper number of active radios and channels in the network.

5) Remarks: With advances in self-interference cancellation, full duplex radios can transmit and receive data simultaneously, thus reusing the scarce frequency resource. Deploying full duplex relaying systems is expected to improve SE by a factor at most 2 as compared with traditional half-duplex ones [138]. Whereas, the EE of full duplex wireless networks has not been well explored.

Nowadays, wireless networks present a tendency of exploiting multiple radios or channels. In general, resource allocation problems in MR-MC networks are of higher complexity due to the multiple dimension nature [139]. Energy efficiency optimization problems in generic MR-MC networks calls for low-complexity and efficient algorithms to solve the coupled problems of radio and channel allocation and transmission scheduling, which remains an open issue.

E. Joint Resource Allocation for EEWN

Taking advantage of the full resource space information and jointly allocating different resources, significant improvement to the overall network EE can be anticipated. For example, as a combination of two techniques, MIMO-OFDMA systems [140] can bring more benefit but demand for joint spatial and frequency allocation. In the above, some joint resource allocation algorithms have been already discussed such as those exploiting spatio-temporal diversities of wireless networks. Another representative application scenario of jointly allocating all the above resources is heterogeneous networks (HetNets) that are deployed with both macro cells and small cells (e.g., femtocells) for better coverage, higher capacity and QoE. In HetNets, femtocell BSs operate with much lower power than macro cell ones, thus improving the network EE demands wise deployment and power management of femtocell BSs to mitigate inter-cell interference under coverage constraints. Scheduling these BSs to adapt to network traffic conditions (e.g., turning off some femtocell BSs in areas with no traffic demand) calls for novel load-balancing techniques. In addition, enhancing network EE calls for joint power and spectrum allocation schemes [141].

Besides, there is a large volume of research on cross-layer optimization approaches for designing energy-efficient wireless network protocols. For instance, crossing the bottom MAC and PHY layers will allow jointly allocating power, spectrum and time resources in order to achieve higher energy efficiency [142]. Since the fundamental energy-preserving techniques are almost mentioned above, we refer interested readers to excellent surveys on joint allocation schemes in [17], [22], [142], and [143].

1) Remarks: Most joint resource allocation approaches result in solving complex optimization problems over the whole network. Nevertheless, they may encounter significant challenges in practical implementations due to lack of sufficient management of network heterogeneity, complexity and consistency [144], [145]. For example, in mobile big data networks, the BSs are often unaware of user status, making it difficult to obtain the optimal solutions. In addition, since network service providers focus on their own profit, to implement the optimization solutions on devices belonging to different providers is challenging.

In addition, in the big data era, new features such as high volume data traffic loads and great pressure on spectrum utilization call for more flexible resource allocation schemes across the network. To this send, a prominent solution is the SDN framework. With data and control separation, controllers are able to jointly allocate network resources. For example, in the SDN based VANET concept as proposed in [29], the system can make more informed decisions on path selection, channel/frequency coordination, and power control.

V. Energy-Efficient Data Storage in Wireless Networks

The acquired big data must be stored for querying, accessing and computing. Storing big data is a vital and challenging problem which involves issues such as file systems (e.g., Google's GFS) and database technologies (e.g., NoSQL) [3]. Although important for big data, such issues are not closely related to wireless networking. In this section, we investigate the wireless networking aspect of big data storage and hence, based on the above energy-efficient wireless communications, we shall focus on the in-network data caching technologies.

Conventionally, data from their source devices are transmitted to some remote servers for storing and processing. However, such a centralized paradigm is significantly energy-inefficient over wireless networks in the big data era due to high communication cost. Therefore, communication cost and data storage cost should be balanced in order to improve energy efficiency. On the other hand, big data generated from IoT sensors/devices and mobile devices may be frequently accessed by on-site controllers, end-users and edge nodes in mobile networks, demanding that the data are cached closer to them for fast querying. To this end, distributed storage in WSNs and mobile caching in mobile big data scenarios are two important ways.

A. Distributed Data Storage in WSNs

In big data applications in IoT and WSNs, using fixed sinks for collecting all sensory data is almost impractical since the fast energy depletion of nodes around the sinks will create energy holes and break network connectivity as a result. In addition, in hostile, harsh and catastrophic application environments, the WSNs may frequently loss connectivity due to failure of either wireless communication links or sensor nodes. Some studies have suggested the use of mobile sinks to collect sensor data, but it is hard to guarantee that all the data are collected reliably and with low latency. To tackle these challenges, distributed data storage is to store sensory data (with redundancy) at the sensor nodes themselves to ensure reliable data reconstruction for the mobile sink nodes by visiting only a subset of the sensors. Energy-efficient distributed storage in WSNs focuses on saving energy through reducing number of transmissions and receptions for data collection and dissemination, while maintaining the reliability of sensor data recovery.

1) Compressed Network Coding: Data dissemination contributes the major portion of power consumption in distributed storage of WSNs, which includes data transmission and reception. In order to reduce energy consumption, an effective way is to reduce the number of data transmissions during dissemination. In densely distributed WSNs, sensing results of neighboring sensors usually have high correlation. Taking advantage of this feature, compressed sensing can be exploited to reduce data transmission. In addition, by allowing node to forward linearly combined data, network coding technique can be applied to further reduce the number of transmissions. Based on these approaches, Yang *et al.* [146] proposed a compressed network coding based distributed storage scheme for WSNs. They exploited the correlations of sensor readings based on compressed sensing theory and network coding, and designed a data format to control data reception. This can achieve reduction in both transmissions and receptions while guaranteeing the reliable recovery of data. Further, in order to improve the efficiency, the same authors proposed an adaptive method in [147], where, based on the random geometric graph theory, the expressions of the numbers of transmissions and receptions were derived, which indicated that the number of transmissions can be reduced by using a smaller forwarding probability. On the other hand, the forwarding probability should be large enough to maintain acceptable level of recovery error. It was observed in [146] that using the same forwarding probability for all the nodes was not optimal since each node's ability of disseminating messages is affected by neighbors. A node with a larger number of neighbors can use a smaller forwarding probability, while it should more frequently forward messages when with less neighbors. Based on these, an adaptive scheme was proposed with which the forwarding probability of nodes was adjusted according to the number of neighbors of each node.

In addition to the spatial correlation of neighboring sensors, Gong *et al.* [148] took temporal correlations over time slots into consideration for sensor reading collection, which can further reduce the transmissions required for data recovery. Specifically, sensor readings from consecutive time slots were linearly combined with a network coding scheme. Then to recover the readings, only a subset of nodes need to be visited, which greatly reduces the number of transmissions. It was shown that the proposed scheme could achieve the same recovery performance with much less transmissions and receptions.

2) Information-Based Querying: In view of the "value" dimension of the big data, the information stored in WSNs may be more important than the big data themselves. For example, instead of storing all generated measurement data, a sensor node may only store the average or maximum values of the data [149]. Targeting at highly connected and dynamic WSNs with massive data, Bergelt *et al.* [149] proposed a database-orientation system for information querying in WSNs, where the whole network was viewed as a virtual database with each sensor corresponding to a row. Together with a wake-up mechanism and an aggregation strategy, the proposed system enables efficient use of energy.

In complex environments where the sensing data are with multiple attributes, i.e., with multi-dimensional data, Tissera *et al.* [150] proposed an energy-efficient querying mechanism that resolved two types of queries: ANY-type query was resolved when the query packet reached a sensor node that contained data relating to an attribute, while ALL-type query was resolved when this sensor node had all data of the attribute. Then, a load balancing problem was considered and a distributed algorithm was applied to construct multiple trees from information source node. Simulation results showed that the proposed mechanism reduces query response time and energy cost.

3) Practical Considerations: Considering the practical scenarios where sensor readings may not be compressible in the discrete cosine transformation domain, the work in [151] improved the spatio-temporal compressive coding scheme to guarantee data recovery by training dictionaries adaptively to achieve sparse representation and optimize measurement accuracy. In [151], a two-dimensional dictionary training method was proposed and the measurement matrices were redesigned for both the spatial and temporal dictionaries. Considering both sparseness and incoherence, an adaptive column combination method was developed that can achieve better energy efficiency with more accurate data recovery.

Another practical consideration is to estimate the real distribution and addresses of sensor nodes. In [152], sensing data is mapped to clusters such that the storage among sensors can be balanced, according to the estimated distribution. Then a cluster-based routing is performed, where the cluster head is rotated among cluster members to balance the energy consumption within each cluster. With this method, the sensor

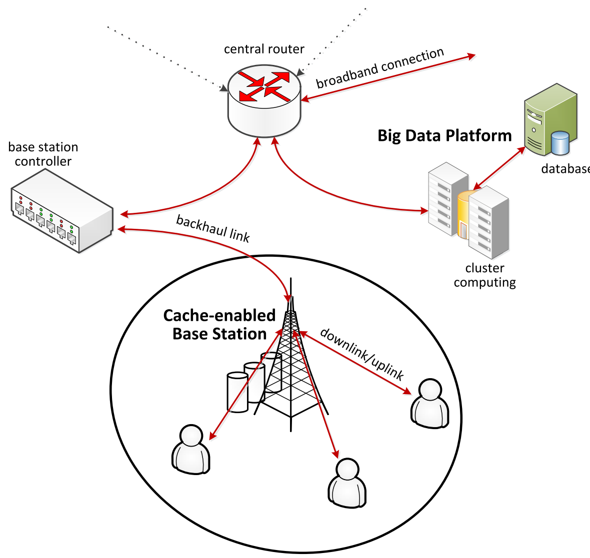


Fig. 7. Architecture of proactive caching based on a big data platform [153].

nodes are not required to equip with localization system, thus further energy savings can be achieved.

B. Mobile Caching of Big Data

In mobile networks, it is observed that duplicate downloads of some popular contents (e.g., pop music and videos) contribute to an important portion of big data traffic. Motivated by this, mobile caching aims at proactively caching popular contents at edge devices in order to reduce both the traffic and energy cost of backhaul links, especially during peak traffic hours. Meanwhile, big data in mobile networks face not only the challenge of huge data amount, but also the limited storage space and power supply, which should be considered in the design of energy-efficient mobile caching.

1) Proactive Caching: Since human behavior is highly predictable, proactive caching can be utilized in mobile networks for caching at the edges, which are usually at BSs. With human behavior prediction and predictive resource management, proactive caching moves content closer to users, and hence is promising in improving user experiences and backhaul offloading gains. At the same time, a large amount of data enables big data analytics and machine learning (ML) techniques, based on which content popularity estimation can be carried out to enhance caching efficiency and performance.

As shown in Fig. 7, a proactive caching scheme to handle huge amounts of big data was proposed in [153] and [154], where the authors leveraged big data analysis for content popularity computation such that strategic contents can be cached at BSs to improve user satisfaction and backhaul offloading. Popularity estimation under such a scenario is highly challenging due to high sparsity of spatio-temporal user behavior, large number of users and content catalog. As a solution, the authors exploit ML tools to enable parallel computations of content popularity on big data platform (details to be discussed in Section VII). The results showed that 100% user request satisfaction and 98% reduction on backhaul usage were achieved.

Further, the work in [155] reduced the energy consumption of mobile user devices in proactive content caching by increasing transmission time of requests and dynamically scheduling when to download the content according to channel conditions. At the same time, since users' cache memory is usually limited, the energy efficiency of proactive caching is also limited by the cache capacity constraint. In this case, an offline optimization was performed to derive the upper bound of the proactive caching gain, together with a backward water-filling algorithm to obtain the optimal caching strategy.

2) Social Networking for Caching Mobile Big Data: Since mobile networks are highly related to human activities, social networking plays an important role in mobile caching. For example, the work in [156] leveraged social networking and mobile caching network to improve the energy efficiency of edge nodes in fog computing systems. Considering the energy conservation and QoS requirements of fog computing, the paper exploited social and spatial structure of the network, and designed an edge node selection scheme based on the social centrality of users, which was obtained from encountering history information and location information of end users. Then the content placement among edge nodes was optimized by minimizing the energy consumption in the network.

In order to store a very large volume of mobile big data in mobile social networks, a content-centric framework with caching in content store was proposed in [4]. Each node has a content stores and the mobile big data are distributed and stored at the content stores cooperatively. In this way, a data request from a mobile user can be fulfilled more rapidly with less traffic whenever a replica of the data is cached in a nearby content store. Further, it was proposed that the mobile big data can be served with different priorities based on information from user plans and potential profits of data. Then the data value can be explored for decision making and optimization of the network in order to support fast management of the big data, which resolves the issues regarding variety of value (aside from volume) features of big data.

3) Context Awareness: In practice, mobile caching faces many challenges. Due to user mobility, the caching scheme has to be carefully designed to balance the communication and storage costs for big data. Besides, the storage decision should be made to avoid bottleneck problems in which a user is accessed by too many neighbors. In addition, users may have different data access preferences that should be accounted in the design. These issues can be effectively tackled with context-aware storage schemes, where context information such as user mobility patterns, network condition and access preferences can be jointly considered. For example, the work in [157] developed a context-aware scheme that accounted these factors in a mobile cloud computing environment and studied the problem of minimizing the expected global transmission time in data access of all users.

In addition, content popularity, preference, user characteristics and operator objective were also considered as context information in [158]. The content popularity was first learned online by observing users' requests on cache content without assuming priori knowledge. Then, the content popularity was modeled with users' personal characteristics, equipments, and external factors, based on which the context-specific content popularity can be learned. In addition, the proposed scheme also enables service differentiation with customer prioritization. With the learned content popularity, cache content can be updated accordingly and thus content placement strategy can be optimized.

VI. Energy-Efficient Data Computation

Based on the data acquisition, communication and storage technologies, the collected massive data will be computed for extracting meaningful information, which is the core part of big data research. In this section, since we are focusing on the wireless networking aspect, we narrow our scope to energy-efficient big data computation platforms based on wireless networks, such as wireless cloud networks and wireless data-center networks, and shall focus on recent studies for better service provisioning, smarter resource utilization, higher flexibility and plausibly lower costs in these platforms.

A. Data-Center Networks (DCNs)

Data-centers carry out the most data storage and computation services and probably consume the most significant part of energy in handling big data. As reported in [159], a normal data-center consumes roughly the same energy as 25,000 households. Therefore, the energy efficiency of data center networks is of great importance for big data.

In the literature, the most commonly mentioned data-centers are based on wired communications. Three primary obstacles that limit the application of most wireless technologies in DCNs are low data rate, unreliability due to wireless interference and fading, and security issues such as information leakage [160]. However, driven by the demand for flexible and low cost data-centers in contrast to wired ones, recently, there have been a few studies attempting to develop wireless DCNs. An important method is to establish wireless links based on the emerging 60 GHz RF technology which can achieve throughput as high as 15Gbps [161]. The 60 GHz RF can protect information security since the emitted wireless signals will not penetrate walls if the data-centers are properly placed in concrete rooms [160]. Technologies such as directional antenna, beamforming and beamsteering are integrated to further enhance reliability and mitigate interference [160], [162]. Based on the 60 GHz technology, it is demonstrated that fully wireless data-centers (except for the power supply wires) are practically feasible and, compared with traditional wired data-centers, are promising in improving latency and reducing energy consumption [161], [163].

In the literature, the energy efficiency aspect of wireless DCNs has not been well studied. However, to improve the energy efficiency, many existing mechanisms originally designed for wired data-centers might be applied. While energy saving can be achieved by reducing the active component in DCNs, it potentially increases the appearance probability of hotspot and congestion. Considering such a trade-off, approaches have been taken from the following aspects.

1) Sleep Scheduling Based on Virtualization: Data-centers are often over provisioned in order to accommodate traffic upsurges, which leads to that the data-centers are underutilized in most time. Virtualization in DCNs allows for dynamic migration of virtual machines (VMs) among physical nodes according to workload demands and performance requirements. Within a data-center, VMs can be consolidated to reduce the number of active physical servers so that idle servers can be turned off to save energy without violating service level agreements (SLA). However, aggressive VM placement strategies may concentrate workloads on a subset of the machines and hence cause hot spots and congestion, which may not only degrade the system performance but also incur extra energy cost for cooling. In [164], a two-level control approach was considered, where the local controller of each VM allocated resources (e.g., CPU, memory and storage) to guarantee application performance, while the data-center employed a global controller to determine VM placement at the initial stage. The VM placement was formulated as a multi-objective optimization problem aiming at jointly minimizing conflicting objectives including energy consumption, thermal dissipating cost and total resource wastage. A grouping genetic algorithm was designed to search for optimal solutions where a fuzzy logic based approach was embedded to evaluate each solution. To dynamically determine VM placement to save energy, a double-threshold idea was proposed in [165], in which: 1) a server would be commanded to sleep and all its VMs would be migrated to others if its CPU utilization was below the lower threshold: 2) some of the server's VMs would be migrated to others if its CPU utilization exceeded the upper threshold. In order to save communication energy of VMs, one useful way is to consolidate groups of communicating VMs in a small area of a data center network in order to reduce the path length of the flows. To solve such a VM consolidation problem that is NP-hard, a topology-aware recursive algorithm was proposed which prevented the formation of network bottlenecks and had the advantage of handling a large number of VMs without sacrificing energy efficiency [166].

Network virtualization was exploited in [167] and a heuristic scheduling algorithm was proposed for mapping in virtual networks to turn off idle servers and network resources. Similarly, the work in [168] proposed a greedy approach to find a sub-network according to traffic condition which can fulfill flow demands, and the rest of the network can be turned off to reduce energy consumption. Considering the energy consumption on routers and switches, Nam *et al.* [169] combined server consolidation and idle logic, which can rapidly turn on/off sub-components, based on dynamic traffic activities. The fluctuation pattern of traffic can also be explored to power on/off idle devices, such as powering off devices during off-peak time. As discussed in [170], both the hosts and the network were considered, and a joint optimization of virtual machine placement and flow routing was studied.

2) Traffic Routing: The VM placement problem is often coupled with traffic routing in order to handle dynamic workloads in which the key idea for energy saving is to shutdown unneeded network devices such as idle switches. Under given routing requirement, the energy-aware routing problem aiming at minimizing the number of switches while guaranteeing a certain level of throughput was studied in [171]. Given the network traffic flow matrix, it was shown that the problem is NP-hard. A heuristic algorithm was proposed that, based on an initial basic routing, iteratively removed switches until the achieved throughput decreased to its threshold value. With the objective of minimizing total energy consumption, a joint VM placement and routing problem was formulated and proved to be NP-complete in [172]. The problem was then decomposed into three sub-problems including traffic-aware VM grouping, distance-aware VM-group to server-rack mapping and power-aware inter-VM traffic flow routing. The third sub-problem was transformed into a multi-commodity flow problem that was solved based on Greedy Bin-Packing algorithm.

Wang *et al.* [173] analyzed the network-as-a-service model, based on which the energy-efficient routing was formulated as an optimization problem under multiple resources. Flows were selected progressively with a greedy algorithm to exhaust node capacity, while paths assignment was done according to residual capacities and flow demand. The work also leveraged regularity of DCN structure for a topology-aware method. Considering deadline constrained flows, the work in [174] proposed an approximation algorithm to jointly perform flow scheduling and routing based on relaxation and randomized rounding methods. Further, dependency based virtual deadline approach was proposed in [175], which set virtual deadlines according to the dependencies of tasks for scheduling, thus tasks can be dynamically assigned to servers based on link and server load.

B. Mobile Cloud Computing Networks

Cloud computing is a new service provisioning paradigm that offers dynamic and scalable computing services in different manners referred to as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) and so on. Users can purchase and enjoy these services anywhere and anytime on demand. The cloud structure can be integrated with wireless networks yielding, for example, cloud base station networks and mobile clouds, which provide flexible and rich computing platforms for storing and computing mobile big data. On the other hand, the increasing amount of mobile big data due to proliferation of mobile devices and increase of mobile data demand are accelerating the development of mobile cloud computing systems. It is hence important to develop energy-efficient mobile cloud computing networks to be applied in the big data era. In this subsection, we present a survey of recent such efforts.

1) Cloud Base Stations: In traditional cellular networks, BSs are managed roughly in a distributed manner and the computation resources are provisioned according to traffic load conditions at each BS to ensure certain QoS. However, without close coordination among the BSs, there is an inevitable amount of computation resource waste when handling big data due to highly dynamic traffic loads. The cloud radio access networks (C-RAN) and Wireless Network Cloud (WNC) projects promoted the transition from distributed to centralized cloud framework by consolidating the baseband units of multiple BSs into a virtual one. In this manner, the baseband cellular signals can be processed in a central location, which potentially saves energy. As an example, the eBase system for baseband unit clustering can save around 20% of the total energy [176]. Exploring the load variations of the BSs, a cloud based framework, called CloudIQ, was proposed which partitioned the BSs into groups and scheduled them to meet real-time requirements [177].

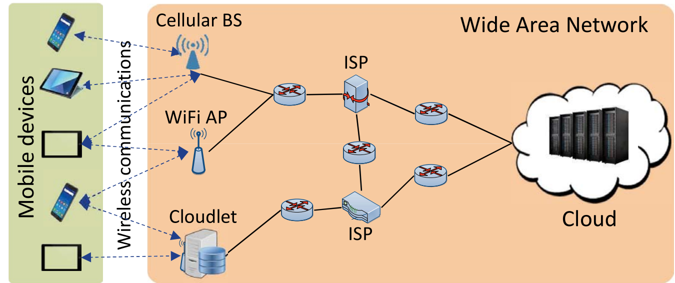


Fig. 8. Architecture of mobile cloud computing [181].

With the virtualization technology based on software-defined hardware, the computation resources in cloud-based cellular systems can be pooled and dynamically allocated. In such virtual BSs, investigating the energy efficiency should take both computation and communication costs into account. In [178], a power-delay tradeoff problem was studied in which the total power consisted of baseband signal processing power and the radio (including radio circuits) power. The optimization solution can jointly optimize the data rate and the number of CPU cores. Such a computation-resource-aware approach over virtualized BSs can save much more energy (more than 60% as demonstrated in [178]) than conventional BS systems.

To overcome the performance limitation of C-RAN as constrained by the limited-capacity fronthaul, heterogeneous C-RAN (H-CRAN) has been proposed that takes advantages of both C-RAN and heterogeneous networks. H-CRAN has emerged as a cost-efficient network architecture for supporting mobile big data, and achieves higher energy efficiency than C-RAN [179], [180].

2) Mobile Edge Computing: Mobile edge computing (MEC), sometimes also termed as mobile cloud computing (MCC), extends cloud computing services to mobile end users. As illustrated in Fig. 8, through wireless accessing networks and cloudlets (can be viewed as small-scale wireless datacenters), MEC systems can offer ubiquitous computing service access with high scalability [181] and has emerged as a useful platform for mobile big data computing [182].

In an MEC system, computation intensive applications can be migrated from resource-constrained mobile devices to resourceful cloud such that the servers or other mobile devices can help carry out the computation. By means of such offloading, we may expect that the energy expenditure of the mobile node can be saved and its battery lifetime can be prolonged. However, in handling mobile big data, the problem of energy efficiency remains challenging for offloading in such a mobile cloud due to high communication overhead, despite many existing energy-efficient communication protocols and mechanisms for cellular networks and Wi-Fi networks (referring to relevant sections above). To illustrate this, consider the

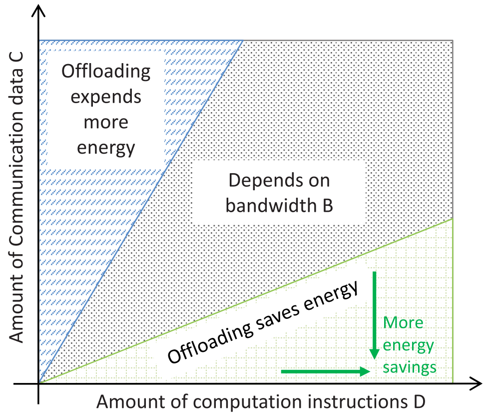


Fig. 9. When does offloading save energy? [183].

example in [183] where a mobile node sends D bytes of data to a server to offload an application with C instructions for computation. It then comes that the amount of energy that the node can save by performing the offloading can be written as

$$\Delta E = \alpha C - \frac{\beta}{B} D, \tag{4}$$

where α and β are constants depending on the power consumption of the node and the cloud server, respectively. B is the available bandwidth for communication. As suggested in (4), whether the offloading can save energy depends on the amounts of both computation instructions D and communication data C. As illustrated in Fig. 9, when D is much larger than C, it is intuitively more energy-efficient to offload such a computation intensive application to the server. In contrast, when communication is expensive, this application is better carried out at the mobile node itself. It is worth noticing that the condition also depends on the bandwidth B in the sense that a large B can save communication time (and hence energy) and thus can improve the energy efficiency.

The above simple example demonstrates the basic tradeoff problem between communication and computation costs when investigating the benefit of offloading in MEC networks. As suggested by Fig. 9, more energy may be saved by either reducing the amounts of communications or increasing the amount of offloaded computations. In this sense, energy-efficient offloading methods can be categorized into three classes as follows.

- *Computation-based approaches:* By means of some fine-grained code management schemes (e.g., MAUI [184] and ThankAir [185]) which partition the whole offloading application program and (adaptively) offload those computation-intensive partitioned parts, the average amount of computation D to be offloaded can be increased. To reduce the burden on programmers, the entire program can migrate from the mobile device to cloud servers, which, however, saves less energy.
- *Communication-based approaches:* Another way is to reduce the amount of communication by means such as data aggregation (to merge sporadic data packets) and data compression (to reduce the number of communication bits) [186], [187]. Also, the cloud can store certain data and perform computation on it, making the mobile node more convenient and energy-efficient to send short data pointers instead of transmitting the whole program or data to cloud servers [183]. In practice, wireless communications suffer from time-varying connectivity and available bandwidth constraints, which further affect the energy efficiency of offloading in MEC. Taking the time-varying communication quality into account, Shu *et al.* [188] proposed an energy-efficient data transmission strategy *eTime* for offloading in MEC. *eTime* dynamically seizes opportunities of good wireless connectivity of 3G or Wi-Fi to prefetch frequently used data while deferring delay-tolerant data to save energy.
- *Hybrid and Joint Optimizations:* In some cases, rather than completely offloading or executing a certain program locally on the mobile node, a hybrid method can better utilize the cloud resource and may yield higher energy efficiency. Zhang *et al.* [189] considered that the tasks of an application can be collaboratively executed on the mobile node and the cloud, and investigated the problem of minimizing the energy consumption of the mobile node while guaranteeing the total execution of the tasks. Based on a tree-topology task model and a stochastic channel model, they showed that the most favorable task scheduling policy was to conduct task offloading at most once. Their results also advocated the collaborative tasks execution over completely offloading or no offloading. Partial computation offloading can be realized by applying the dynamic voltage scaling technique which can adaptively adjust computational speed based on computation load to reduce energy consumption. Considering the partial computation offloading problem of a mobile device in MEC, Wang *et al.* [190] studied two joint optimization problems: one was energy consumption minimization and the other was execution latency minimization, where the decision variables for the mobile device were its computational speed f_l, transmit power P_t and the fraction of data bits to be executed locally λ. Specifically, in the first minimization problem, the objective was to minimize total energy consumption which accounts for the local computing cost (depending on both f_l and λ) and the transmitting and receiving energy (depending on both λ and P_t) during the offloading process between the device and the cloud server. The minimization program is subject to a constraint on the application execution latency which is modeled as the larger value of local computing time and the total time with offloading to the server (including execution and round-trip communication time), since the two processes run in parallel. The second minimization problem aims at minimizing the total application execution latency while guaranteeing that the total energy is bounded. The results showed that partial offloading was beneficial in terms of energy saving. In multi-user offloading cases, a joint optimization of communication resources (in terms of users' transmitting precoding matrices relating to data rate) and computation resources (in terms of CPU cycles) of each mobile users was studied in [191]. The problem was formulated as a nonlinear

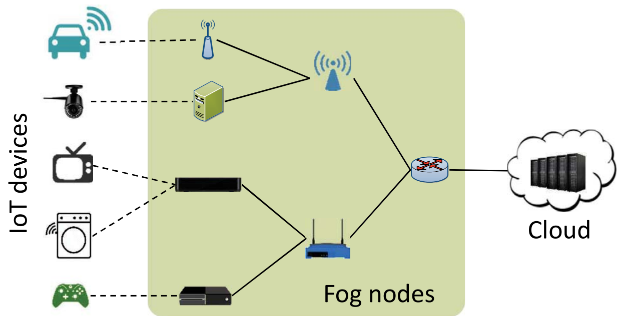


Fig. 10. Fog computing for IoT [192].

program to minimize the total energy consumption of each mobile user while ensuring bounded overall latency experienced by each user and under power budget constraints. The results confirmed that the joint optimization outperforms disjoint optimization approaches.

3) Fog Computing: In IoT networks, devices (e.g., sensors) are often densely distributed that will generate and process big data sets. For service real-timeliness and energy efficiency, it is advisable to process big data within the vicinity of the data source devices. This can be achieved with fog computing, which is usually referred to as a distributed paradigm that offers real-time computing, distribution and storage services to IoT devices. As shown in Fig. 10, a fog computing system consists of both edge and core networking devices (e.g., routers, APs, BSs and switches devices) and, unlike MEC, can also extend cloud services such as IaaS, PaaS and SaaS to edge devices [192].

Compared to centralized cloud computing platforms, the distributed fog computing can provide more convenient services to proximal IoT devices. However, whether such a paradigm shift from centralized to distributed schemes will save energy has not been well studied. In [193], the energy consumption of a service provided by centralized data-centers (DCs) in traditional cloud and nano data-centers (nDCs, smaller servers that can host and distribute data and applications in a peer-to-peer manner) in fog computing was analyzed and compared, where the energy consumption accounted for all that in IoT devices, access network, edge and core networks and DCs or nDCs. The results showed that nDCs can save a small amount of energy in some cases, but generally the energy saving depends on the type of access networks, nano server scheduling and type of applications. It was also shown that the best energy efficiency was achieved in applications generating and distributing IoT big data that were infrequently accessed.

4) Ad Hoc Cloud Computing: The current prominent cloud computing systems necessitate a cluster of expensive and dedicated cloud servers, incurring significant capital and energy costs. Moreover, accessing the cloud infrastructure may not be always available especially when the mobile devices move out of the coverage of cellular BSs, Wi-Fi APs or cloudlets. On the other hand, the proliferation of mobile devices such as smartphones and tablets provides us the opportunity to utilize voluntary untapped proximal computing and storage resources, forming a new cloud computing paradigm termed as mobile ad hoc cloud (MAHC) [194], [195].

With MAHC, battery-constrained mobile devices can offload computations to other nearby mobile devices that are resource rich or underloaded. In [196], a virtual cloud computing framework was proposed for mobile phones, where a task can be partitioned into portions such that some can be executed locally while others can be offloaded to nearby mobile devices. It was shown that the task processing time was reduced, implying the reduction of energy consumption. As designed in [197], mobile devices having spare resources can form a dynamic MAHC to provide computing services to others. A mobile device that is unaffordable for executing certain tasks can launch a resource discovery process to find an appropriate node in the MAHC for executing the task. In order to save energy, the target node is selected as the one having minimum estimated task execution cost.

In MAHC, human activities (e.g., mobility, social ties, smartphone recharging, smartphone usage habits, and data demand preferences) play a vital role. For example, weather and news data may be more demanded in the morning. Considering such information would help make fine-grained resource allocation decisions and may achieve more energy saving. Specifically, sociality-aware neighbor discovery and transmission scheduling for D2D communications can be further explored in the future [198]. Social ties and reputation information may be utilized in computation outsourcing applications for better efficiency.

VII. Big Data Analytics for Enhancing Energy Efficiency of Wireless Networks

Although big data will continue to impose great pressure on wireless networks in terms of storage, communication and computation, the benefits that big data can bring are also significant. Most existing approaches for network energy efficiency optimization are model-based which strongly rely on model accuracy; whereas in practice, various kinds of noises, obstacles, interferences and other factors hardly accounted in the models will limit the optimality (or even feasibility) of these approaches. On the other hand, network optimization problems based on accurate models are often too complicated to be solved analytically or the solutions are difficult to be implemented due to high complexity. However, the value conveyed in the big data can be learned by means of some data mining and ML algorithms and further used to develop relatively low-complexity solutions, and hence will improve the network performance including energy efficiency. More importantly, learning-based approaches are able to learn and adapt to time-varying wireless environment, which is an uneasy task for model-based approaches.

In the literature, data analytics have been successfully applied in WSNs [25] and CRNs [199] dealing with issues such as data aggregation, routing, localization, clustering, security, dynamic channel allocation, transmission control and cross-layer resource allocation. However, there are not many works explicitly discussing the application of data analytics for enhancing EE of wireless networks in the big data context. In this section, we present a review of recent studies closely relating to the B4N category.

A. Mining Valuable Information

It was reported that 4.2 Exabytes of mobile data traffic were generated per month in 2015 and this number was predicted to increase to 24.3 Exabytes per month in 2019 [4]. The big data generated by sensors, smartphones and vehicles may convey useful information of users, e.g., user habit, behavior, preferences, and ambient environment, making big data themselves as precious resources [200]. Under regulation and privacy guarantee, such information can be extracted and utilized to improve the network configurations for better performance. For example, Xu *et al.* [201] proposed to extract spatial and temporal information from social media big data to identify urban emergency events. Such a method can be useful in mobile crowdsensing networks, e.g., we can predict network traffic bursts if we learn the emergence of certain social events.

Further, with ML based methods, valuable information (e.g., primary user's state in cognitive radio systems) from big data sets to achieve various goals such as signal classification, MAC protocol identification, attack detection and network throughput improvement [202]. Big data were exploited in [203]–[205] for anomaly detection in the network. In [203], a big data analytics platform was presented for anomaly detection and root cause analysis in mobile wireless networks. The proposed method learned the symptoms of network anomalies and built a knowledge database with historic data. Then an ML approach was employed to identify new anomalies and mapped the detected anomalies to the database for root cause analysis. Similarly, big data analytics were exploited in [204], where an effective QoS management was proposed to perform root cause analysis and predict traffic congestion. Moreover, Parwez *et al.* [205] employed big data analytics with ML tools for user anomaly detection, which helped in identifying regions of interest for resource allocation and fault avoidance solution. The work in [206] provided an overview of the challenges in mobile big data, and proposed exploiting deep learning in extracting meaningful information and hidden patterns from big data. Specifically, a Spark based framework was developed to execute distributed deep learning that was time-efficient in large-scale mobile systems, which can speed up the learning rate of deep network with many hidden layers and millions of parameters.

The vast amount of geolocated data generated from pervasive mobile devices can be exploited in analyzing human behavior, but with some technical challenges such as the collection and storage of data, noise removing and analyzer. To this end, a big mobile data analytical framework was developed in [207], which defined data processing rules, constructed user trajectories by extracting user location data from different sources and reduced oscillations to remove data noise.

The learned valuable information can be further utilized to enhance the energy efficiency of wireless networks. A proactive caching scheme that utilized cache-enabled unmanned aerial vehicles (UAVs) was proposed in [208]. In order to provide required QoE to mobile users in C-RAN with minimum transmit power of the UAVs, an ML algorithm was proposed, which leveraged human-centric information to predict content request distribution and mobility pattern of each user and used the prediction results to determine the UAVs placement and the content to cache. The human information is the data of gender, occupation, age, and type of mobile device in use. That paper showed a good example of how to improve wireless network energy efficiency based on learning useful information from big data. In [209], a reinforcement learning based MAC protocol was proposed for WSNs, in which the learning algorithm helped a node to infer the states of other nodes. By doing so, a near-optimal MAC policy can be learned that achieves high throughput and low power consumption.

B. ML for Performance Evaluation

High volumes of user data also provide a credible way to evaluate users' QoE that is important for user-centric networks. For example, in cellular handover management, existing methods that usually select cells merely based on signal strength may result in bad QoE after handover. To tackle this problem, the work in [210] proposed an ML scheme consisting of two levels of feed-forward artificial neural networks to learn the impact of handover on user QoE from historic data. The first-level neural network outputs the successful file downloading rate of users; while the second-level neural network outputs the file downloading time which is further used as the key QoE to decide the handover cell. The two neural networks were trained with measurements obtained from user equipments. Their simulation study showed that the proposed algorithm was able to select the handover cell with better expected QoE and the achieved performance was close to the optimal one. Similarly, the performance of LTE network was investigated in [211] with big data techniques. In [211], a large amount of network measurements and diagnosis data were utilized to evaluate and predict the network capacity. A forecasting algorithm was proposed, which can predict the network resource consumptions based on network traffic and service growth.

We envision that such methods can provide credible models of network performance in terms of QoE, capacity and EE, and hence facilitate the formulation of EE optimization under QoE constraints when accurate models of the EE and QoE are unavailable. Such an EE optimization strategy is worthy of future investigations.

C. ML for Resource Allocation

Big data assisted resource allocation was investigated in [212], for improving the system capacity in LTE networks. By learning from a large amount of network measurements and diagnosis data, an interference management algorithm using big data analytics was proposed that can cluster users based on specified metric and perform resource allocation accordingly. Further, in mobile networks especially in the context of self-organizing networking, allocating resources in order to provision required QoS and QoE for users is a challenging issue. A learning-based approach that dynamically assigns frequency and bandwidth resources in LTE small cell networks was developed in [213]. The key idea is shown in Fig. 11. The prediction engine estimates the network key performance

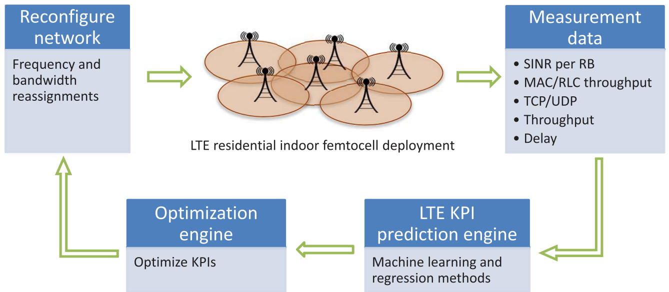


Fig. 11. A learning-based resource allocation scheme [213].

indicators (KPIs) that relate to changes in QoS and QoE, based on measurements (e.g., throughput, delay and SINR per resource block) obtained from users and the small cells. It also predicts the network KPI by means of some ML and regression methods. Then, the Optimization Engine decides the resource allocation based on both the PE output and the measurements; specially, it updates the network parameters to reconfigure frequency and bandwidth assignments if by doing so, some performance benefits can be achieved. The main novelty lies in the prediction engine for which the authors tested and compared various ML algorithms including linear regression, bagging tree, boosted tree, K-nearest neighbor (KNN), support vector machine, Kohonen networks and projection pursuit regression. Their results showed that the learning-based approach can achieve 95% of the optimal network performance and that the bagging tree based method outperformed others. Although energy efficiency is not an issue in that work, the proposed learning-based approach can be applied to optimize network energy efficiency if it is accounted as a KPI.

As network operators possess a large amount of data relating to user behavior and network performance, they can operate the network in a proactive manner if equipped with big data learning abilities. In other words, the network paradigm transforms from base-station-centric to user-centric. In this regard, in order to improve cellular backhaul offloading in cache-enabled 5G wireless networks, a proactive scheme to cache contents at BSs was developed in [153]. The problem was formulated as a minimization of backhaul load, which requires joint optimization of cache decisions (i.e., the caching time and BS of each content) and content popularity (which characterizes user demand of the contents). To this end, the authors proposed a Hadoop platform to process the big data of BSs and estimate the contention popularity based on ML tools. The platform analyzes over 20 billion downlink packets and over 15 billion uplink packets, which amount to over 80 TByte of total data daily. Numerical results demonstrated improvements in user satisfaction of backhaul load. The paper did not explicitly mention network energy cost, however the idea manifested can be further extended to study energy efficiency optimization of cache placement at BSs in cellular networks.

Learning based frameworks are promising in addressing data processing and resource management in IoT. However, the unique features in IoT such as resource constraints, heterogeneity and strict QoS requirements may limit the application of learning frameworks. This issue was discussed in [214], where several emerging learning frameworks were presented with their advantages, limitations and applications. Particularly, by introducing the cognitive hierarchy theory, a novel framework was proposed to overcome the heterogeneity issue in IoT, which mapped different devices to multiple levels of rationality such that different learning frameworks can be used according to resource availability.

D. Discussion

As we can see from the above literature, ML based big data analytics are powerful tools in enhancing network performance. However, applying such methods for improving wireless network energy efficiency, though promising, have not been well studied. There are a hand of tools available to develop such learning-based energy-efficient algorithms; however, which tool is most suitable depends on specific networks and their application scenarios. On the other hand, since the learning approaches often require a training stage to gradually adjust important algorithm parameters, an extra amount of computation energy is incurred which calls for future investigations to account it in the network energy efficiency optimization problems. In [215], in view of the great complexity of the flow allocation problem for EE optimization in MR-MC wireless networks, a deep learning based algorithm was developed, which first evaluates the links at the training stage and then allocates flows by solving a reduced optimization problem that only considers the links with high scores.

As big data may be continuously generating over time, the performance of the big data analytics tools used should be able to scale with the increase of data volume. To this end, deep learning has been shown more scalable than traditional techniques such as back propagation neural networks and support vector machines.

As discussed in the big data acquisition and storage sections, the collection of both training and testing big data is itself a challenge in large-scale and dense wireless networks. For applying big data analytics in such networks, learning algorithms that can be distributedly implemented (e.g., sequential learning and reinforcement learning [214]) are preferred.

VIII. Open Issues

With the proliferation of wireless devices and rapidly generating big data, the demand for more EEWN technologies is vast. There has been a much larger volume of literature on techniques that are promising for EEWN in the coming big data era than what have been surveyed in the above. Still, there are many research issues remaining open and calling for further investigations.

A. Energy Efficiency Optimization

Rethink the EE optimization paradigms mentioned in Section II-B. It is known that minimizing energy consumption and maximizing network performance are often conflicting objectives. As a result, each paradigm has its own issues.

Since both link capacity and communication energy consumption can be modeled as functions of physical layer parameters such as transmit power and antenna configuration of the wireless nodes, this offers great mathematical convenience for formulating the energy-efficient power control problem under the Type-I optimization paradigm with objective function in form of bits per Joule. Despite the convenience, such a formulation encounters two challenges. Firstly, the objective function is generally non-convex, which limits the application of many mature convex optimization methods. Distributed solutions based on system decomposition are either mathematically difficult to derive or hard to guarantee global optimality. Secondly, unconstrained optimal solution may be found at physically less meaningful points (e.g., to achieve high bits-per-Joule performance, nodes may tend to use excessively low transmit power, which incurs unsatisfactorily low capacity). Therefore, the optimal solutions should be searched within a subspace of the whole space spanned by all possible transmit power allocation, where the subspace can be constrained by energy budget or desired level of capacity. This is perhaps an important reason that optimizing other network parameters such as link scheduling, routing, and flow control usually do not take the Type-I EE concept.

Usually, utilizing more resource that implies spending more energy will result in better performance, and vice versa. In this case, the Type-II and Type-III EE optimization paradigms will degenerate to searching for the energy-minimal resource allocation within the subspace where performance exactly meets its requirement and searching for the performance-optimal resource allocation within the subspace where energy expenditure exactly meets its budget, respectively. However, those optimal allocation points may not be optimal from a bits-per-Joule perspective. Take the Type-II case for example. It is possible that, at those optimal points, slightly increasing the energy by exploiting more resources will introduce significant performance improvement.

In future wireless networks, especially when dealing with big data of large volumes and high variety, we may more often talk about multi-objective optimization or apply different optimization frameworks at different performance/energy intervals. Such a multi-interval hybrid approach may circumvent the shortcomings of the above three optimization paradigms.

B. Mobility Management

A significant portion of wireless big data is generated by mobile devices such as mobile phones, wearable sensors and vehicles. In such mobile networks, connectivity, channel quality and traffic flows become highly dynamic, making energy-efficient protocols and algorithms difficult to design. For instance, in many transmission time based scheduling techniques [216], [217], the scheduling decisions are made based on prediction or estimation of channel quality, user mobility or traffic conditions. Consequently, the accuracy of prediction plays an important role in the final performance. Therefore, the design of effective on-line learning algorithms with high accuracy and low power consumption is a promising but challenging issue for EE optimization.

C. Scalability and Fairness

In the big data era, scalability of protocols and algorithms is an important practical issue in potentially large-scale networks such as crowdsensing networks, vehicular networks and cloud computing networks. To improve EE, the energy overhead for running the optimization algorithms should not exceed the energy savings offered by them. Sophisticated algorithms may be discouraged if they incur too much signaling and computation overhead. Centralized schemes, which require powerful central units to retrieve and process the whole network information, may result in high overhead and consequently low EE and low scalability. Alternatively, distributed and cooperative methods, which require only local communications and information processing, could be more energy-efficient and scalable. However, a key challenge for distributed and scalable methods lies in the coordination among nodes within their local subnetwork (e.g., neighbors). As another fundamental challenge, distributed methods often can only achieve sub-optimality of EE optimizations. A plausible solution can be a hybrid scheme in which distributed cooperation among neighboring nodes can be coordinated by some central units with (partial) global view of network operations.

Besides, for big data applications over wireless networks, it is often required to provide fair QoE, in a long-term run, for all participating users. However, with such a fairness constraint, the network may deviate from its optimal operation points, which incurs the tradeoff problem between fairness and network energy efficiency.

D. Energy-Efficient Data Security Provisioning

Data security and privacy are becoming more and more important concerns of network users. To enhance the network robustness against various security attacks and preserve the privacy of sensitive information against possible eavesdroppers, we have to spend significant amounts of energy for computation- and communication- expensive tasks such as data encryption/description, authentication, attack detection and defending [85]. For example, transmitting and receiving security-related packet headers incur a considerable extra amount of communication overhead. A challenging research direction for future study is to reduce energy cost while preserving required quality of protection (QoP) [37]. For example, in the big data era, it becomes quite inefficient and expensive to strictly protect every piece of data. Instead, in view of the "variety" and "value" dimensions of big data, it calls for system-level solutions with security service differentiations by taking the information conveyed in big data into account.

E. Energy Efficiency in Wireless Cyber-Physical Systems

In most existing EEWN protocols, data are treated as of equal values and importance. For example, network throughput accounts for all transmitted data. However, in the big data era, data may be of different meanings and of different values. This is particularly true in wireless cyber-physical systems (CPS) in which physical dynamic systems (e.g., habitant environments, smart grid, drones) are controlled in real-time and the information exchanges among sensors, actuators, routers and controllers are carried out by wireless networks [218]–[221]. In such systems, the performance may be no longer judged by the network throughput where data are deemed of the same value, but by the monitoring and control quality where the reliability and real-timeliness of critical data transmissions are more emphasized [218]. From the networking aspect, new allocation methods involving all the above resources are demanded to improve the system EE. For example, event-triggering strategy that transmits data only when necessary can save energy while preserving a certain quality of real-time control [220]. An approach that jointly allocates temporal and spectral resources was shown to be promising to improve the EE of remote state estimation systems [118]. In particular, systematic approaches that co-design communication, networking and control are demanded in wireless cyber-physical systems.

F. Big Data As Resources

Through the above survey, we may learn that the essence of EEWN is to optimally utilize available resources to provide satisfactory performance with low energy consumption. Traditional EEWN techniques usually take power, time, spectrum and spatial resources into account, but only a few are aware of that storage and computing resources should be seriously considered in the design of energy-efficient protocols and algorithms. In the big data era, all the resources may be jointly optimized for higher network energy efficiency. Whereas, the resulting optimization problems may be too complex to be solved optimally. Instead of searching for sub-optimal solutions, the idea of B4N as discussed in Section VII provides a promising approach. In fact, big data themselves can be viewed as resources that can provide network designers with valuable information about the network operation status, service quality and quality of users' experience. Assisted with advanced ML techniques, big data can be utilized to solve those complex optimization problems of joint resource allocations, calling for interdisciplinary research.

IX. Conclusion

We have presented a survey of recent studies that are effective or promising for developing energy-efficient wireless networks for big data. It is concluded that the essence of most energy-efficient protocols and algorithms is to utilize network resources (e.g., nodes' power, time, spatial, spectrum, computing and storage resources and also the network generated data themselves) in an optimized manner, i.e., to provision best performance at lowest cost. We also have identified several open issues critical for enhancing wireless network energy efficiency in the big data era, including issues such as energy-efficient data security and energy efficiency in wireless cyber-physical systems. This survey should spur more novel perspectives and design approaches for energy efficiency enhancement in future wireless networks with big data.

References

[1] S. Bi, R. Zhang, Z. Ding, and S. Cui, "Wireless communications in the era of big data," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 190–199, Oct. 2015.

[2] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020*, White Paper, CISCO, San Jose, CA, USA, 2016.

[3] M. Chen, S. Mao, Y. Zhang, and V. C. M. Leung, *Big Data: Related Technologies, Challenges and Future Prospects*. Cham, Switzerland: Springer, 2014.

[4] Z. Su, Q. Xu, and Q. Qi, "Big data in mobile social networks: A QoE-oriented framework," *IEEE Netw.*, vol. 30, no. 1, pp. 52–57, Jan./Feb. 2016.

[5] S. Yu, M. Liu, W. Dou, X. Liu, and S. Zhou, "Networking for big data: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 531–549, 1st Quart., 2017,

[6] S. N. Roy, "Energy logic: A road map to reducing energy consumption in telecom munications networks," in *Proc. IEEE Int. Telecommun. Energy Conf.*, San Diego, CA, USA, 2008, pp. 1–9.

[7] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 2, pp. 223–244, 2nd Quart., 2011.

[8] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, and T. T. Kwon, "A survey of green mobile networks: Opportunities and challenges," *Mobile Netw. Appl.*, vol. 17, no. 1, pp. 4–20, 2012.

[9] Global eSustainability Initiative (GeSI). (2008). *SMART 2020: Enabling the Low Carbon Economy in the Information Age*. [Online]. Available: http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf

[10] A. P. Miettinen and J. K. Nurminen, "Energy efficiency of mobile clients in cloud computing," in *Proc. USENIX Conf. Hot Topics Cloud Comput.*, 2010, p. 4.

[11] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Netw.*, vol. 7, no. 3, pp. 537–568, 2009.

[12] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, 3rd Quart., 2011.

[13] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[14] Ł. Budzisz *et al.*, "Dynamic resource provisioning for energy efficiency in wireless access networks: A survey and an outlook," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2259–2285, 4th Quart., 2014.

[15] J. Wu, Y. Zhang, M. Zukerman, and E. K.-N. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 803–826, 2nd Quart., 2015.

[16] M. Kennedy, A. Ksentini, Y. Hadjadj-Aoul, and G.-M. Muntean, "Adaptive energy optimization in multimedia-centric wireless devices: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 768–786, 2nd Quart., 2013.

[17] D. Feng *et al.*, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, 1st Quart., 2013.

[18] K. Davaslioglu and E. Ayanoglu, "Quantifying potential energy efficiency gain in green cellular wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2065–2091, 4th Quart., 2014.

[19] J. B. Rao and A. O. Fapojuwo, "A survey of energy efficient resource management techniques for multicell cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 154–180, 1st Quart., 2014.

[20] M. Ismail, W. Zhuang, E. Serpedin, and K. Qaraqe, "A survey on green mobile networking: From the perspectives of network operators and mobile users," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1535–1556, 3rd Quart., 2015.

[21] X. Huang, T. Han, and N. Ansari, "On green-energy-powered cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 827–842, 2nd Quart., 2015.

[22] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1656–1686, 3rd Quart., 2016.

[23] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Energy efficient multimedia streaming to mobile devices—A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 579–597, 1st Quart., 2014.

[24] R. Mahapatra, Y. Nijsure, G. Kaddoum, N. U. Hassan, and C. Yuen, "Energy efficiency tradeoff mechanism towards wireless green communication: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 686–705, 1st Quart., 2016.

[25] M. A. Alsheikh, S. Lin, D. Niyato, and H.-P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1996–2018, 4th Quart., 2014.

[26] D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, "Toward energy efficient big data gathering in densely distributed sensor networks," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 388–397, Sep. 2014.

[27] F. Hu, Q. Hao, and K. Bao, "A survey on software defined networking (SDN) and OpenFlow: From concept to implementation," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2181–2206, 4th Quart., 2014.

[28] S. Tomovic, M. Pejanovic-Djurisic, and I. Radusinovic, "SDN based mobile networks: Concepts and benefits," *Wireless Pers. Commun.*, vol. 78, no. 3, pp. 1629–1644, 2014.

[29] I. Ku *et al.*, "Towards software-defined VANET: Architecture and services," in *Proc. Annu. Mediterranean Ad Hoc Netw. Workshop (MED HOC NET)*, 2014, pp. 103–110.

[30] A. De Gante, M. Aslan, and A. Matrawy, "Smart wireless sensor network management based on software-defined networking," in *Proc. Biennial Symp. Commun.*, 2014, pp. 71–75.

[31] M. Yang *et al.*, "Software-defined and virtualized future mobile and wireless networks: A survey," *Mobile Netw. Appl.*, vol. 20, no. 1, pp. 4–18, 2015.

[32] M.-L. Ku, W. Li, Y. Chen, and K. J. R. Liu, "Advances in energy harvesting communications: Past, present, and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1384–1412, 2nd Quart., 2016.

[33] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[34] F. Meshkati, H. V. Poor, S. C. Schwartz, and R. V. Balan, "Energy-efficient resource allocation in wireless networks with quality-of-service constraints," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3406–3414, Nov. 2009.

[35] G. Y. Li *et al.*, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.

[36] A. A. Khalek and Z. Dawy, "Energy-efficient cooperative video distribution with statistical QoS provisions over wireless networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 7, pp. 1223–1236, Jul. 2012.

[37] H. Zhu, K.-H. Liu, W. He, and K. Ota, "Quality of experience and quality of protection provisions in emerging mobile networks," *IEEE Wireless Commun.*, vol. 22, no. 4, pp. 8–9, Aug. 2015.

[38] X. Lu, P. Wang, D. Niyato, and Z. Han, "Resource allocation in wireless networks with RF energy harvesting and transfer," *IEEE Netw.*, vol. 29, no. 6, pp. 68–75, Nov./Dec. 2015.

[39] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1180–1189, Mar. 2012.

[40] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. Netto, and R. Buyya, "Big data computing and clouds: Trends and future directions," *J. Parallel Distrib. Comput.*, vols. 79–80, pp. 3–15, May 2015.

[41] S. He, J. Chen, J. Li, and Y. Sun, *Energy-Efficient Area Coverage for Intruder Detection in Sensor Networks* (Briefs in Computer Science). Cham, Switzerland: Springer, 2014.

[42] M. Patel, R. Chandrasekaran, and S. Venkatesan, "Energy efficient sensor, relay and base station placements for coverage, connectivity and routing," in *Proc. IEEE Int. Perform. Comput. Commun. Conf.*, Phoenix, AZ, USA, 2005, pp. 581–586.

[43] P. Cheng, X. Cao, J. Bai, and Y. Sun, "On optimizing sensing quality with guaranteed coverage in autonomous mobile sensor networks," *Comput. Commun.*, vol. 35, no. 9, pp. 1107–1114, 2012.

[44] Y. Song, B. Wang, Z. Shi, K. R. Pattipati, and S. Gupta, "Distributed algorithms for energy-efficient even self-deployment in mobile sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1035–1047, May 2014.

[45] J.-W. Lee, B.-S. Choi, and J.-J. Lee, "Energy-efficient coverage of wireless sensor networks using ant colony optimization with three types of pheromones," *IEEE Trans. Ind. Informat.*, vol. 7, no. 3, pp. 419–427, Aug. 2011.

[46] S. He, J. Chen, D. K. Y. Yau, H. Shao, and Y. Sun, "Energy-efficient capture of stochastic events by global-and local-periodic network coverage," in *Proc. ACM MobiHoc*, New Orleans, LA, USA, 2009, pp. 155–164.

[47] J. Chen *et al.*, "On energy-efficient trap coverage in wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 10, no. 1, p. 2, 2013.

[48] M. Nikolov and Z. J. Haas, "Collaborating with correlation for energy efficient WSN," in *Proc. ACM Int. Workshop Mission Orient. Wireless Sensor Netw.*, 2012, pp. 27–32.

[49] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[50] C. Karakus, A. C. Gurbuz, and B. Tavli, "Analysis of energy efficiency of compressive sensing in wireless sensor networks," *IEEE Sensors J.*, vol. 13, no. 5, pp. 1999–2008, May 2013.

[51] D. Schneider, "New camera chip captures only what it needs," *IEEE Spectr.*, vol. 3, no. 50, pp. 13–14, Mar. 2013.

[52] W. Chen and I. J. Wassell, "Energy-efficient signal acquisition in wireless sensor networks: A compressive sensing framework," *IET Wireless Sensor Syst.*, vol. 2, no. 1, pp. 1–8, Mar. 2012.

[53] L. Kong *et al.*, "Embracing big data with compressive sensing: A green approach in industrial wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 10, pp. 53–59, Oct. 2016.

[54] A. Wang, F. Lin, Z. Jin, and W. Xu, "A configurable energy-efficient compressed sensing architecture with its application on body sensor networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 15–27, Feb. 2016.

[55] J. Li, X. Xu, H. Zhao, Y. Hu, and T. Z. Qiu, "An energy-efficient sub-Nyquist sampling method based on compressed sensing in wireless sensor network for vehicle detection," in *Proc. Int. Conf. Connected Veh. Expo (ICCVE)*, Las Vegas, NV, USA, 2013, pp. 314–319.

[56] M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Sparse event detection in wireless sensor networks using analog fountain codes," in *Proc. IEEE Globecom*, Austin, TX, USA, 2014, pp. 3520–3525.

[57] W. Chen and I. J. Wassell, "Optimized node selection for compressive sleeping wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 827–836, Feb. 2016.

[58] L. Zhang, J. Liu, H. Jiang, and Y. Guan, "SensTrack: Energy-efficient location tracking with smartphone sensors," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3775–3784, Oct. 2013.

[59] R. Frank, F. Melakessou, G. Castignani, and T. Engel, "Energy-efficient rate-adaptive passive traffic sensing using smartphones," in *Proc. Annu. Mediterranean Ad Hoc Netw. Workshop (MED HOC NET)*, Ajaccio, France, 2013, pp. 67–74.

[60] W. Sherchan *et al.*, "Using on-the-move mining for mobile crowdsensing," in *Proc. IEEE Int. Conf. Mobile Data Manag. (MDM)*, 2012, pp. 115–124.

[61] K. K. Rachuri, C. Mascolo, M. Musolesi, and P. J. Rentfrow, "SociableSense: Exploring the trade-offs of adaptive sampling and computation offloading for social sensing," in *Proc. ACM MobiCom*, Las Vegas, NV, USA, 2011, pp. 73–84.

[62] L. Wang, D. Zhang, and H. Xiong, "effSense: Energy-efficient and cost-effective data uploading in mobile crowdsensing," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publ.*, Zürich, Switzerland, 2013, pp. 1075–1086.

[63] N. D. Lane *et al.*, "Piggyback crowdsensing (PCS): Energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities," in *Proc. ACM SenSys*, Rome, Italy, 2013, Art. no. 7.

[64] N. A. Pantazis and D. D. Vergados, "A survey on power control issues in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 9, no. 4, pp. 86–107, 4th Quart., 2007.

[65] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Found. Trends® Netw.*, vol. 2, no. 4, pp. 381–533, 2008.

[66] J. Xu *et al.*, "Cooperative distributed optimization for the hyper-dense small cell deployment," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 61–67, May 2014.

[67] D. E. Charilas and A. D. Panagopoulos, "A survey on game theory applications in wireless networks," *Comput. Netw.*, vol. 54, no. 18, pp. 3421–3430, 2010.

[68] X. Wang, W. Zheng, Z. Lu, X. Wen, and W. Li, "Dense femtocell networks power self-optimization: An exact potential game approach," *Int. J. Commun. Syst.*, vol. 29, no. 1, pp. 16–32, 2016.

[69] L. Lei *et al.*, "Optimal reliability in energy harvesting industrial wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5399–5413, Aug. 2016.

[70] R. Deng, Y. Zhang, S. He, J. Chen, and X. Shen, "Maximizing network utility of rechargeable sensor networks with spatiotemporally coupled constraints," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1307–1319, May 2016.

[71] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.

[72] K. Yang, Q. Yu, S. Leng, B. Fan, and F. Wu, "Data and energy integrated communication networks for wireless big data," *IEEE Access*, vol. 4, pp. 713–723, 2016.

[73] S. Peng and C. P. Low, "Prediction free energy neutral power management for energy harvesting wireless sensor nodes," *Ad Hoc Netw.*, vol. 13, pp. 351–367, Feb. 2014.

[74] L. Huang and M. J. Neely, "Utility optimal scheduling in energy-harvesting networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 4, pp. 1117–1130, Aug. 2013.

[75] D. Zhang *et al.*, "Utility-optimal resource management and allocation algorithm for energy harvesting cognitive radio sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3552–3565, Dec. 2016.

[76] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[77] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless charging technologies: Fundamentals, standards, and network applications," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1413–1452, 2nd Quart., 2016.

[78] S. Zhang *et al.*, "Energy-aware traffic offloading for green heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1116–1129, May 2016.

[79] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," in *Proc. IEEE INFOCOM*, vol. 3. Anchorage, AK, USA, 2001, pp. 1548–1557.

[80] R. Bruno, M. Conti, and E. Gregori, "Optimization of efficiency and energy consumption in p-persistent CSMA-based wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 99, no. 1, pp. 10–31, Jan./Mar. 2002.

[81] T.-H. Lee, A. Marshall, and B. Zhou, "An energy efficient cross-layer design by using the contention window optimization scheme for ad-hoc wireless networks operating in error-prone channel conditions," in *Proc. Int. Conf. Genet. Evol. Comput.*, Shenzhen, China, 2010, pp. 505–508.

[82] L. Chen, S. H. Low, and J. C. Doyle, "Contention control: A game-theoretic approach," in *Proc. IEEE CDC*, New Orleans, LA, USA, 2007, pp. 3428–3434.

[83] T. Cui, L. Chen, and S. H. Low, "A game-theoretic framework for medium access control," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1116–1127, Sep. 2008.

[84] M. Di Francesco, G. Anastasi, M. Conti, S. K. Das, and V. Neri, "Reliability and energy-efficiency in IEEE 802.15.4/ZigBee sensor networks: An adaptive and cross-layer approach," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1508–1524, Sep. 2011.

[85] X. Cao *et al.*, "Ghost-in-ZigBee: Energy depletion attack on ZigBee-based wireless networks," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 816–829, Oct. 2016.

[86] Y. Liu, C. Yuen, X. Cao, N. Ul Hassan, and J. Chen, "Design of a scalable hybrid MAC protocol for heterogeneous M2M networks," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 99–111, Feb. 2014.

[87] L. Liu, X. Cao, Y. Cheng, and Z. Niu, "Energy-efficient sleep scheduling for delay-constrained applications over WLANs," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2048–2058, Jun. 2014.

[88] T.-H. Lee and J.-R. Hsieh, "Low-complexity class-based scheduling algorithm for scheduled automatic power-save delivery for wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 12, no. 3, pp. 571–580, Mar. 2013.

[89] R. Palacios-Trujillo, J. Alonso-Zarate, N. L. S. da Fonseca, and F. Granelli, "Maximum achievable energy efficiency of TXOP power save mode in IEEE 802.11ac WLANs," in *Proc. IEEE Globecom*, Washington, DC, USA, 2016, pp. 1–7.

[90] X.-H. Lin, L. Liu, H. Wang, and Y.-K. Kwok, "On exploiting the on-off characteristics of human speech to conserve energy for the downlink VoIP in WiMAX systems," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Istanbul, Turkey, 2011, pp. 337–342.

[91] J.-R. Lee and D.-H. Cho, "Dual power-saving modes for voice over IP traffic supporting voice activity detection," *IET Commun.*, vol. 3, no. 7, pp. 1239–1249, 2009.

[92] S. Zhang, J. Gong, S. Zhou, and Z. Niu, "How many small cells can be turned off via vertical offloading under a separation architecture?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5440–5453, Oct. 2015.

[93] Y. S. Soh, T. Q. S. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 840–850, May 2013.

[94] D. Cao, S. Zhou, and Z. Niu, "Optimal combination of base station densities for energy-efficient two-tier heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4350–4362, Sep. 2013.

[95] S. Zhou, T. Zhao, Z. Niu, and S. Zhou, "Software-defined hyper-cellular architecture for green and elastic wireless access," *IEEE Commun. Mag.*, vol. 54, no. 1, pp. 12–19, Jan. 2016.

[96] G. Micallef, P. Mogensen, and H.-O. Scheck, "Cell size breathing and possibilities to introduce cell sleep mode," in *Proc. Eur. Wireless Conf.*, Lucca, Italy, 2010, pp. 111–115.

[97] P. H. Gomes, N. L. S. da Fonseca, and O. C. Branquinho, "Radio resource allocation and green operation for mobile access networks based on radio-over-fiber," *IEEE Trans. Mobile Comput.*, vol. 13, no. 4, pp. 894–906, Apr. 2014.

[98] C. Jiang, Y. Shi, Y. T. Hou, and W. Lou, "Cherish every joule: Maximizing throughput with an eye on network-wide energy consumption," in *Proc. IEEE INFOCOM*, Orlando, FL, USA, 2012, pp. 1934–1941.

[99] L. Liu *et al.*, "Energy-efficient capacity optimization in wireless networks," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 2014, pp. 1384–1392.

[100] X. Cao *et al.*, "Real-time misbehavior detection and mitigation in cyber-physical systems over WLANs," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 186–197, Feb. 2017.

[101] J. T. Louhi, "Energy efficiency of modern cellular base stations," in *Proc. Int. Telecommun. Energy Conf.*, Rome, Italy, 2007, pp. 475–476.

[102] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. IEEE VTC Fall*, Anchorage, AK, USA, 2009, pp. 1–5.

[103] A. J. Fehske, F. Richter, and G. P. Fettweis, "Energy efficiency improvements through micro sites in cellular mobile radio networks," in *Proc. IEEE Globecom Workshops*, Honolulu, HI, USA, 2009, pp. 1–5.

[104] D. Cao, S. Zhou, and Z. Niu, "Improving the energy efficiency of two-tier heterogeneous cellular networks through partial spectrum reuse," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4129–4141, Aug. 2013.

[105] Y. Wu and Z. Niu, "Energy efficient base station deployment in green cellular networks with traffic variations," in *Proc. IEEE ICCC*, Beijing, China, 2012, pp. 399–404.

[106] A. De Domenico, E. C. Strinati, and A. Duda, "An energy efficient cell selection scheme for open access femtocell networks," in *Proc. IEEE PIMRC*, Sydney, NSW, Australia, 2012, pp. 436–441.

[107] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1525–1536, Sep. 2011.

[108] H. Ali, W. Shahzad, and F. A. Khan, "Energy-efficient clustering in mobile ad-hoc networks using multi-objective particle swarm optimization," *Appl. Soft Comput.*, vol. 12, no. 7, pp. 1913–1928, 2012.

[109] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados, "Energy-efficient routing protocols in wireless sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 551–591, 2nd Quart., 2013.

[110] J. Zhu and X. Wang, "Model and protocol for energy-efficient routing over mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1546–1557, Nov. 2011.

[111] X. Mao, X.-Y. Li, W.-Z. Song, P. Xu, and K. Moaveni-Nejad, "Energy efficient opportunistic routing in wireless networks," in *Proc. ACM MSWiM*, 2009, pp. 253–260.

[112] J. Zuo *et al.*, "Cross-layer aided energy-efficient opportunistic routing in ad hoc networks," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 522–535, Feb. 2014.

[113] K. Wei, X. Liang, and K. Xu, "A survey of social-aware routing protocols in delay tolerant networks: Applications, taxonomy and design-related issues," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 556–578, 1st Quart., 2014.

[114] K. Wei, D. Zeng, S. Guo, and K. Xu, "Social-aware relay node selection in delay tolerant networks," in *Proc. Int. Conf. Comput. Commun. Netw. (ICCCN)*, Nassau, The Bahamas, 2013, pp. 1–7.

[115] J. Hu, L.-L. Yang, and L. Hanzo, "Cooperative multicast aided picocellular hybrid information dissemination in mobile social networks: Delay/energy evaluation and relay selection," in *Proc. IEEE WCNC*, Istanbul, Turkey, 2014, pp. 3207–3212.

[116] Y. Wei, F. R. Yu, and M. Song, "Distributed optimal relay selection in wireless cooperative networks with finite-state Markov channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2149–2158, Jun. 2010.

[117] R. Deng *et al.*, "Sensing-performance tradeoff in cognitive radio enabled smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 302–310, Mar. 2013.

[118] X. Cao, X. Zhou, L. Liu, and Y. Cheng, "Energy-efficient spectrum sensing for cognitive radio enabled remote state estimation over wireless channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2058–2071, Apr. 2015.

[119] J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. Shen, "Dynamic channel access to improve energy efficiency in cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3143–3156, May 2016.

[120] H. Li, T. Li, X. Shi, and Y. Wang, "Data collection through device-to-device communications for mobile big data sensing," in *Proc. 1st Workshop Mission Crit. Big Data Anal.*, 2016. [Online]. Available: https://www.semanticscholar.org/paper/Data-Collection-through-Device-to-Device-Communica-Li-Li/7b4639dbad094618a49ac066e57329d63b336365?tab=abstract

[121] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (D2D) communication," in *Proc. IEEE VTC Spring*, Yokohama, Japan, 2012, pp. 1–5.

[122] H. ElSawy, E. Hossain, and M.-S. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4147–4161, Nov. 2014.

[123] X. Cao, L. Liu, Y. Cheng, L. X. Cai, and C. Sun, "On optimal device-to-device resource allocation for minimizing end-to-end delay in VANETs," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7905–7916, Oct. 2016.

[124] Y. Wu, J. Chen, L. P. Qian, J. Huang, and X. S. Shen, "Energy-aware cooperative traffic offloading via device-to-device cooperations: An analytical approach," *IEEE Trans. Mobile Comput.*, vol. 16, no. 1, pp. 97–114, Jan. 2017.

[125] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, 4th Quart., 2014.

[126] X. Xiao, X. Tao, and J. Lu, "A QoS-aware power optimization scheme in OFDMA systems with integrated device-to-device (D2D) communications," in *Proc. IEEE VTC Fall*, San Francisco, CA, USA, 2011, pp. 1–5.

[127] M. Belleschi, G. Fodor, and A. Abrardo, "Performance analysis of a distributed resource allocation scheme for D2D communications," in *Proc. IEEE Globecom Workshops*, Houston, TX, USA, 2011, pp. 358–362.

[128] A. Rahmati and L. Zhong, "Context-for-wireless: Context-sensitive energy-efficient wireless data transfer," in *Proc. ACM MobiSys*, 2007, pp. 165–178.

[129] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Low-complexity energy-efficient scheduling for uplink OFDMA," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 112–120, Jan. 2012.

[130] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.

[131] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.

[132] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.

[133] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec. 2012.

[134] H. Li, Y. Cheng, C. Zhou, and W. Zhuang, "Minimizing end-to-end delay: A novel routing metric for multi-radio wireless mesh networks," in *Proc. IEEE INFOCOM*, 2009, pp. 46–54.

[135] Y. Cheng, H. Li, D. M. Shila, and X. Cao, "A systematic study of maximal scheduling algorithms in multiradio multichannel wireless networks," *IEEE/ACM Trans. Netw.*, vol. 23, no. 4, pp. 1342–1355, Aug. 2015.

[136] X. Cao, L. Liu, W. Shen, and Y. Cheng, "Distributed scheduling and delay-aware routing in multihop MR-MC wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6330–6342, Aug. 2016.

[137] S. Avallone, "Turning off radios to save power in multi-radio wireless mesh networks," in *Proc. ACM MobiHoc*, Paris, France, 2011, p. 22.

[138] A. Sabharwal *et al.*, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[139] H. Li, Y. Cheng, C. Zhou, and P. Wan, "Multi-dimensional conflict graph based computing for optimal capacity in MR-MC wireless networks," in *Proc. IEEE ICDCS*, Genova, Italy, 2010, pp. 774–783.

[140] Z. Xu *et al.*, "Energy-efficient configuration of spatial and frequency resources in MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 564–575, Feb. 2013.

[141] M. Ismail, A. T. Gamage, W. Zhuang, and X. S. Shen, "Energy efficient uplink resource allocation in a heterogeneous wireless medium," in *Proc. IEEE ICC*, Sydney, NSW, Australia, 2014, pp. 5275–5280.

[142] G. Miao, N. Himayat, Y. G. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: A survey," *Wireless Commun. Mobile Comput.*, vol. 9, no. 4, pp. 529–542, 2009.

[143] L. Liu, X. Cao, W. Shen, Y. Cheng, and L. Cai, "DAFEE: A decomposed approach for energy efficient networking in multi-radio multi-channel wireless networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, 2016, pp. 1–6.

[144] T. Chen, M. Matinmikko, X. Chen, X. Zhou, and P. Ahokangas, "Software defined mobile networks: Concept, survey, and research directions," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 126–133, Nov. 2015.

[145] D. Zhang, Z. Chang, M. Zolotukhin, and T. Hämäläinen, "Energy efficient resource allocation in heterogeneous software defined network: A reverse combinatorial auction approach," in *Proc. IEEE/CIC ICCC*, Shenzhen, China, 2015, pp. 1–6.

[146] X. Yang *et al.*, "Compressed network coding for distributed storage in wireless sensor networks," in *Proc. Int. Symp. Commun. Inf. Technol.*, Gold Coast, QLD, Australia, 2012, pp. 816–821.

[147] X. Yang *et al.*, "Energy-efficient distributed data storage for wireless sensor networks based on compressed sensing and network coding," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5087–5099, Oct. 2013.

[148] B. Gong *et al.*, "Spatiotemporal compressive network coding for energy-efficient distributed data storage in wireless sensor networks," *IEEE Commun. Lett.*, vol. 19, no. 5, pp. 803–806, May 2015.

[149] R. Bergelt, M. Vodel, and W. Hardt, "Energy efficient handling of big data in embedded, wireless sensor networks," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Queenstown, New Zealand, 2014, pp. 53–58.

[150] M. Tissera, R. Doss, G. Li, and L. Batten, "Fast and energy efficient data storage for information discovery in multi-dimensional WSNs," in *Proc. Int. Telecommun. Netw. Appl. Conf.*, Sydney, NSW, Australia, 2015, pp. 88–93.

[151] C. Wang, P. Cheng, Z. Chen, N. Liu, and L. Gui, "Practical spatiotemporal compressive network coding for energy-efficient distributed data storage in wireless sensor networks," in *Proc. IEEE VTC Spring*, Glasgow, U.K., 2015, pp. 1–6.

[152] J. Xu, S. Guo, B. Xiao, and J. He, "Energy-efficient big data storage and retrieval for wireless sensor networks with nonuniform node distribution," *Concurrency Comput. Pract. Exp.*, vol. 27, no. 18, pp. 5765–5779, 2015.

[153] M. A. Kader *et al.*, "Leveraging big data analytics for cache-enabled wireless networks," in *Proc. IEEE Globecom Workshops*, San Diego, CA, USA, 2015, pp. 1–6.

[154] E. Zeydan *et al.*, "Big data caching for networking: Moving from cloud to edge," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 36–42, Sep. 2016.

[155] A. Güngör and D. Gündüz, "Proactive wireless caching at mobile user devices for energy efficiency," in *Proc. IEEE Int. Symp. Wireless Commun. Syst.*, Brussels, Belgium, 2015, pp. 186–190.

[156] S. Wang, X. Huang, Y. Liu, and R. Yu, "CachinMobile: An energy-efficient users caching scheme for fog computing," in *Proc. IEEE/CIC ICCC*, Chengdu, China, 2016, pp. 1–6.

[157] D. Han, Y. Yan, and T. Shu, "Context-aware distributed storage in mobile cloud computing," in *Proc. IEEE Globecom Workshops*, San Diego, CA, USA, 2015, pp. 1–6.

[158] S. Müller, O. Atan, M. van der Schaar, and A. Klein, "Context-aware proactive content caching with service differentiation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1024–1036, Feb. 2016.

[159] J. M. Kaplan, W. Forrest, and N. Kindler, "Revolutionizing data center energy efficiency," McKinsey & Company, New York, NY, USA, Tech. Rep., Jul. 2008. [Online]. Available: https://www.sallan.org/pdf-docs/McKinsey_Data_Center_Efficiency.pdf

[160] Y. Katayama, K. Takano, Y. Kohda, N. Ohba, and D. Nakano, "Wireless data center networking with steered-beam mmWave links," in *Proc. IEEE WCNC*, 2011, pp. 2179–2184.

[161] J.-Y. Shin, E. G. Sirer, H. Weatherspoon, and D. Kirovski, "On the feasibility of completely wireless datacenters," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1666–1679, Oct. 2013.

[162] X. Zhou *et al.*, "Mirror mirror on the ceiling: Flexible wireless links for data centers," in *Proc. ACM SIGCOMM*, Helsinki, Finland, 2012, pp. 443–454.

[163] Y. Zhu *et al.*, "Cutting the cord: A robust wireless facilities network for data centers," in *Proc. ACM MobiCom*, 2014, pp. 581–592.

[164] J. Xu and J. A. B. Fortes, "Multi-objective virtual machine placement in virtualized data center environments," in *Proc. IEEE/ACM Conf. Green Comput. Commun. Conf. Cyber Phys. Soc. Comput.*, Hangzhou, China, 2010, pp. 179–188.

[165] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Gener. Comput. Syst.*, vol. 28, no. 5, pp. 755–768, 2012.

[166] R. A. C. da Silva and N. L. S. da Fonseca, "Topology-aware virtual machine placement in data centers," *J. Grid Comput.*, vol. 14, no. 1, pp. 75–90, 2016.

[167] X. Guan, B.-Y. Choi, and S. Song, "Topology and migration-aware energy efficient virtual network embedding for green data centers," in *Proc. Int. Conf. Comput. Commun. Netw. (ICCCN)*, Shanghai, China, 2014, pp. 1–8.

[168] A. R. A. Kumar, S. V. Rao, and D. Goswami, "Greening 4-4, 1-4 data center network: A Greedy approach for finding an energy efficient subnetwork," in *Proc. IEEE Int. Conf. Cloud Comput. Technol. Sci.*, 2014, pp. 869–874.

[169] T. M. Nam, N. H. Thanh, and D. A. Tuan, "Green data center using centralized power-management of network and servers," in *Proc. Int. Conf. Electron. Inf. Commun.*, Da Nang, Vietnam, 2016, pp. 1–4.

[170] H. Jin *et al.*, "Joint host-network optimization for energy-efficient data center networking," in *Proc. Int. Symp. Parallel Distrib. Process. (IPDPS)*, Boston, MA, USA, 2013, pp. 623–634.

[171] Y. Shang, D. Li, and M. Xu, "Energy-aware routing in data center network," in *Proc. ACM SIGCOMM Workshop Green Netw.*, New Delhi, India, 2010, pp. 1–8.

[172] W. Fang, X. Liang, S. Li, L. Chiaraviglio, and N. Xiong, "VMPlanner: Optimizing virtual machine placement and traffic flow routing to reduce network power costs in cloud data centers," *Comput. Netw.*, vol. 57, no. 1, pp. 179–196, 2013.

[173] L. Wang, A. F. Anta, F. Zhang, J. Wu, Z. Liu, "Multi-resource energy-efficient routing in cloud data centers with network-as-a-service," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Larnaca, Cyprus, 2015, pp. 694–699.

[174] L. Wang *et al.*, "Energy-efficient flow scheduling and routing with hard deadlines in data center networks," in *Proc. IEEE ICDCS*, Madrid, Spain, 2014, pp. 248–257.

[175] M. Żotkiewicz, M. Guzek, D. Kliazovich, and P. Bouvry, "Minimum dependencies energy-efficient scheduling in data centers," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 12, pp. 3561–3574, Dec. 2016.

[176] Z. Kong, J. Gong, C.-Z. Xu, K. Wang, and J. Rao, "eBase: A baseband unit cluster testbed to improve energy-efficiency for cloud radio access network," in *Proc. IEEE ICC*, Budapest, Hungary, 2013, pp. 4222–4227.

[177] S. Bhaumik *et al.*, "CloudIQ: A framework for processing base stations in a data center," in *Proc. ACM MobiCom*, Istanbul, Turkey, 2012, pp. 125–136.

[178] T. Zhao, J. Wu, S. Zhou, and Z. Niu, "Energy-delay tradeoffs of virtual base stations with a computational-resource-aware energy consumption model," in *Proc. IEEE ICCS*, 2014, pp. 26–30.

[179] Q. Liu *et al.*, "SH-CRAN: Hierarchical framework to support mobile big data computing in a secure manner," in *Proc. Workshop Mobile Big Data*, Hangzhou, China, 2015, pp. 19–24.

[180] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.

[181] F. Liu *et al.*, "Gearing resource-poor mobile devices with powerful clouds: Architectures, challenges, and applications," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 14–22, Jun. 2013.

[182] A. Ahmed and E. Ahmed, "A survey on mobile edge computing," in *Proc. Int. Conf. Intell. Syst. Control (ISCO)*, Coimbatore, India, 2016, pp. 1–8.

[183] K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?" *Computer*, vol. 43, no. 4, pp. 51–56, Apr. 2010.

[184] E. Cuervo *et al.*, "MAUI: Making smartphones last longer with code offload," in *Proc. ACM MobiSys*, San Francisco, CA, USA, 2010, pp. 49–62.

[185] S. Kosta, A. Aucinas, P. Hui, R. Mortier, and X. Zhang, "ThinkAir: Dynamic resource allocation and parallel execution in the cloud for mobile code offloading," in *Proc. IEEE INFOCOM*, Orlando, FL, USA, 2012, pp. 945–953.

[186] B. Aggarwal *et al.*, "Stratus: Energy-efficient mobile communication using cloud support," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 4, pp. 477–478, 2010.

[187] Y. Xu and S. Mao, "A survey of mobile cloud computing for rich media applications," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 46–53, Jun. 2013.

[188] P. Shu *et al.*, "eTime: Energy-efficient transmission between cloud and mobile devices," in *Proc. IEEE INFOCOM*, Turin, Italy, 2013, pp. 195–199.

[189] W. Zhang, Y. Wen, and D. O. Wu, "Collaborative task execution in mobile cloud computing under a stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 81–93, Jan. 2015.

[190] Y. Wang, M. Sheng, X. Wang, L. Wang, and J. Li, "Mobile-edge computing: Partial computation offloading using dynamic voltage scaling," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4268–4282, Oct. 2016.

[191] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 1, no. 2, pp. 89–103, Jun. 2015.

[192] R. Mahmud, R. Kotagiri, and R. Buyya, "Fog computing: A taxonomy, survey and future directions," in *Internet of Everything: Algorithms, Methodologies, Technologies and Perspectives*, B. Di Martino, K.-C. Li, L. T. Yang, and A. Esposito, Eds. Singapore: Springer, 2018.

[193] F. Jalali, K. Hinton, R. Ayre, T. Alpcan, and R. S. Tucker, "Fog computing may help to save energy in cloud computing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1728–1739, May 2016.

[194] D. M. Shila, W. Shen, Y. Cheng, X. Tian, and X. S. Shen, "AMCloud: Toward a secure autonomic mobile ad hoc cloud computing system," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 74–81, Apr. 2016.

[195] W. Shen, B. Yin, X. Cao, Y. Cheng, and X. S. Shen, "A distributed secure outsourcing scheme for solving linear algebraic equations in ad hoc clouds," *IEEE Trans. Cloud Comput.*, to be published, doi: TCC.2016.2647718.

[196] G. Huerta-Canepa and D. Lee, "A virtual cloud computing provider for mobile devices," in *Proc. ACM Workshop Mobile Cloud Comput. Services Soc. Netw. Beyond*, 2010, p. 6.

[197] A. Malhotra, S. K. Dhurandher, and B. Kumar, "Resource allocation in multi-hop mobile ad hoc cloud," in *Proc. Recent Adv. Eng. Comput. Sci.*, Chandigarh, India, 2014, pp. 1–6.

[198] L. Wang, L. Lu, X. Cao, X. Tian, and Y. Cheng, "Socially-aware resource allocation for device-to-device communications in cellular networks," *IET Commun.*, vol. 9, no. 3, pp. 342–349, Apr. 2015.

[199] W. Wang, A. Kwasinski, D. Niyato, and Z. Han, "A survey on applications of model-free strategy learning in cognitive wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1717–1757, 3rd Quart., 2016.

[200] C. Hu, H. Li, Y. Jiang, Y. Cheng, and P. Heegaard, "Deep semantics inspection over big network data at wire speed," *IEEE Netw.*, vol. 30, no. 1, pp. 18–23, Jan./Feb. 2016.

[201] Z. Xu *et al.*, "Crowdsourcing based description of urban emergency events using social media big data," *IEEE Trans. Cloud Comput.*, to be published, doi: 10.1109/TCC.2016.2517638.

[202] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, 3rd Quart., 2013.

[203] K. Yang, R. Liu, Y. Sun, J. Yang, X. Chen, "Deep network analyzer (DNA): A big data analytics platform for cellular networks," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2016.2624761.

[204] S. Jain, M. Khandelwal, A. Katkar, and J. Nygate, "Applying big data technologies to manage QoS in an SDN," in *Proc. Int. Conf. Netw. Service Manag. (CNSM)*, Montreal, QC, Canada, 2016, pp. 302–306.

[205] M. S. Parwez, D. B. Rawat, and M. Garuba, "Big data analytics for user-activity analysis and user-anomaly detection in mobile wireless network," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2058–2065, Aug. 2017.

[206] M. A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and apache spark," *IEEE Netw.*, vol. 30, no. 3, pp. 22–29, May/Jun. 2016.

[207] Y. Qiao, Y. Cheng, J. Yang, J. Liu, and N. Kato, "A mobility analytical framework for big mobile data in densely populated area," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1443–1455, Feb. 2017.

[208] M. Chen *et al.*, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017.

[209] Z. Liu and I. Elhanany, "RL-MAC: A reinforcement learning based MAC protocol for wireless sensor networks," *Int. J. Sensor Netw.*, vol. 1, nos. 3–4, pp. 117–124, 2006.

[210] Z. Ali, N. Baldo, J. Mangues-Bafalluy, and L. Giupponi, "Machine learning based handover management for improved QoE in LTE," in *Proc. IEEE/IFIP Netw. Oper. Manag. Symp.*, Istanbul, Turkey, 2016, pp. 794–798.

[211] S. Hu, Y. Ouyang, Y.-D. Yao, M. H. Fallah, and W. Lu, "A study of LTE network performance based on data analytics and statistical modeling," in *Proc. Wireless Opt. Commun. Conf. (WOCC)*, Newark, NJ, USA, 2014, pp. 1–6.

[212] J. Gao, X. Cheng, L. Xu, and H. Ye, "An interference management algorithm using big data analytics in LTE cellular networks," in *Proc. Int. Symp. Commun. Inf. Technol. (ISCIT)*, Qingdao, China, 2016, pp. 246–251.

[213] B. Bojović, E. Meshkova, N. Baldo, J. Riihijärvi, and M. Petrova, "Machine learning-based dynamic frequency and bandwidth allocation in self-organized LTE dense small cell deployments," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, p. 183, 2016.

[214] T. Park, N. Abuzainab, and W. Saad, "Learning how to communicate in the Internet of Things: Finite resources and heterogeneity," *IEEE Access*, vol. 4, pp. 7063–7073, 2016.

[215] L. Liu, Y. Cheng, L. Cai, S. Zhou, and Z. Niu, "Deep learning based optimization in wireless network," in *Proc. IEEE ICC*, Paris, France, 2017, pp. 1–6.

[216] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "PW-MAC: An energy-efficient predictive-wakeup MAC protocol for wireless sensor networks," in *Proc. IEEE INFOCOM*, Shanghai, China, 2011, pp. 1305–1313.

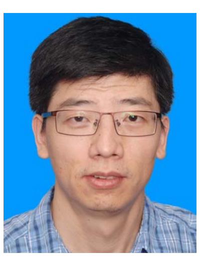
[217] J. Paek, J. Kim, and R. Govindan, "Energy-efficient rate-adaptive GPS-based positioning for smartphones," in *Proc. ACM MobiSys*, San Francisco, CA, USA, 2010, pp. 299–314.

[218] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *Proc. IEEE WCSP*, Nanjing, China, 2011, pp. 1–6.

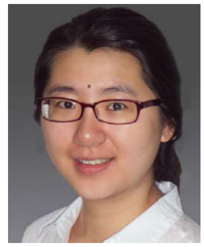
[219] X. Cao, P. Cheng, J. Chen, and Y. Sun, "An online optimization approach for control and communication codesign in networked cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 439–450, Feb. 2013.

[220] J. Araújo, M. Mazo, A. Anta, P. Tabuada, and K. H. Johansson, "System architectures, protocols and algorithms for aperiodic wireless control systems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 175–184, Feb. 2014.

[221] C. Lu *et al.*, "Real-time wireless sensor-actuator networks for industrial cyber-physical systems," *Proc. IEEE*, vol. 104, no. 5, pp. 1013–1024, May 2016.

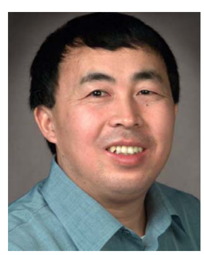


Xianghui Cao (S'08–M'11–SM'16) received the B.S. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2006 and 2011, respectively. From 2007 to 2009, he was a Visiting Scholar with the Department of Computer Science, University of Alabama, Tuscaloosa, AL, USA. From 2012 to 2015, he was a Senior Research Associate with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA. He is currently an Associate Professor with the School of Automation, Southeast University, Nanjing, China. His research interests include cyber-physical systems, wireless network performance analysis, wireless networked control, and network security. He serves as the Publicity Co-Chair for ACM MobiHoc 2015, Symposium Co-Chair for ICNC 2017 and IEEE/CIC ICCC 2015, and a TPC Member for a number of conferences. He also serves as an Associate Editor for several journals, including *ACTA Automatica Sinica*, the IEEE/CAA *Journal of Automatica Sinica*, *KSII Transactions on Internet and Information Systems*, *Security and Communication Networks*, and the *International Journal of Ad Hoc and Ubiquitous Computing*. He was a recipient of the Best Paper Runner-Up Award from ACM MobiHoc 2014.



Lu Liu (S'13) received the B.S. degree in automation from Tsinghua University, Beijing, China, in 2010 and the M.S. degree in electrical engineering from the Illinois Institute of Technology, Chicago, IL, USA, in 2012, where she is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. Her current research interests include energy efficient networking and communication, resource allocation and protocol design of wireless networks, and machine learning-based network optimization.

Yu Cheng (S'01–M'04–SM'09) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University in 1995 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2003. He is currently a Full Professor with the Department of Electrical and Computer Engineering, Illinois Institute of Technology. His research interests include wireless network performance analysis, network security, big data, and cloud computing. He was a recipient of the Best Paper Award at QShine 2007, the IEEE ICC 2011, the Runner-Up Best Paper Award at ACM MobiHoc 2014, the National Science Foundation CAREER Award in 2011, and the IIT Sigma Xi Research Award in the junior faculty division in 2013. He has served as a Symposium Co-Chairs for IEEE ICC and IEEE GLOBECOM, and a Technical Program Committee Co-Chair for WASA 2011 and ICNC 2015. He is a founding Vice Chair of the IEEE ComSoc Technical Subcommittee on Green Communications and Computing. He is an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is an IEEE ComSoc Distinguished Lecturer.



Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree in electrical engineering from Dalian Maritime University, China, in 1982 and the M.Sc. and Ph.D. degrees in electrical engineering from Rutgers University, NJ, USA, in 1987 and 1990, respectively. He is a University Professor and the Associate Chair for Graduate Studies, Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on resource management, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He was an Elected Member of IEEE ComSoc Board of Governor, and the Chair of Distinguished Lecturers Selection Committee. He served as the Technical Program Committee Chair/Co-Chair for IEEE Globecom'16, Infocom'14, IEEE VTC'10 Fall, and Globecom'07, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC'08, the General Co-Chair for ACM Mobihoc'15, Chinacom'07 and QShine'06, and the Chair for IEEE Communications Society Technical Committee on Wireless Communications. He also serves/served as the Editor-in-Chief for the IEEE INTERNET OF THINGS JOURNAL, the IEEE NETWORK, *Peer-to-Peer Networking and Application*, and *IET Communications*; a Founding Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Computer Networks*, and *ACM/Wireless Networks*. He was a recipient of the Excellent Graduate Supervision Award in 2006 and the Premier's Research Excellence Award in 2003 from the province of Ontario, Canada. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.