Smartphone App Usage Analysis : Datasets, Methods, and Applications

Li, Tong

2022

Li , T , Xia , T , Wang , H , Tu , Z , Tarkoma , S , Han , Z & Hui , P 2022 , ' Smartphone App Usage Analysis : Datasets, Methods, and Applications ' , IEEE Communications Surveys and Tutorials , vol. 24 , no. 2 , pp. 937-966 . https://doi.org/10.1109/COMST.2022.3163176

http://hdl.handle.net/10138/354371
https://doi.org/10.1109/COMST.2022.3163176

unspecified
publishedVersion

Smartphone App Usage Analysis: Datasets, Methods, and Applications

Tong Li, *Member, IEEE*, Tong Xia, Huandong Wang, *Member, IEEE*, Zhen Tu, Sasu Tarkoma, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*, and Pan Hui, *Fellow, IEEE*

***Abstract*—As smartphones have become indispensable personal devices, the number of smartphone users has increased dramatically over the last decade. These personal devices, which are supported by a variety of smartphone apps, allow people to access Internet services in a convenient and ubiquitous manner. App developers and service providers can collect fine-grained app usage traces, revealing connections between users, apps, and smartphones. We present a comprehensive review of the most recent research on smartphone app usage analysis in this survey. Our survey summarizes advanced technologies and key patterns in smartphone app usage behaviors, all of which have significant implications for all relevant stakeholders, including academia and industry. We begin by describing four data collection methods: surveys, monitoring apps, network operators, and app stores, as well as nine publicly available app usage datasets. We then systematically summarize the related studies of app usage analysis in three domains: app domain, user domain, and smartphone domain. We make a detailed taxonomy of the problem studied, the datasets used, the methods used, and the significant results obtained in each domain. Finally, we discuss future directions in this exciting field by highlighting research challenges.**

***Index Terms*—Smartphone device, mobile app, app usage, behavior analysis, data mining.**

Manuscript received March 10, 2021; revised September 15, 2021 and January 27, 2022; accepted March 18, 2022. Date of publication March 31, 2022; date of current version May 24, 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFA0711403; in part by the National Natural Science Foundation of China under Grant 61972223, Grant 61971267, Grant 62171260, Grant U1936217, Grant U20B2060, and Grant U21B2036; in part by the International Postdoctoral Exchange Fellowship Program (Talent-Introduction Program) under Grant YJ20210274; in part by the Academy of Finland under Project 319669, Project 319670, Project 325570, Project 326305, Project 325774, and Project 335934; and in part by NSF under Grant CNS-2107216 and Grant CNS-2128368. *(Corresponding author: Pan Hui.)*

Tong Li is with the Department of Computer Science, University of Helsinki, 00100 Helsinki, Finland, also with the System and Media Laboratory, Hong Kong University of Science and Technology, Hong Kong, and also with the Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100190, China (e-mail: tong.li@helsinki.fi).

Tong Xia, Huandong Wang, and Zhen Tu are with the Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: wanghuandong@tsinghua.edu.cn).

Sasu Tarkoma is with the Department of Computer Science, University of Helsinki, 00100 Helsinki, Finland (e-mail: sasu.tarkoma@helsinki.fi).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: hanzhu22@gmail.com).

Pan Hui is with the Computational Media and Arts Trust, Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511458, China, also with the Division of Emerging Interdisciplinary Area, Hong Kong University of Science and Technology, Hong Kong, and also with the Department of Computer Science, University of Helsinki, 00100 Helsinki, Finland (e-mail: panhui@ust.hk).

Digital Object Identifier 10.1109/COMST.2022.3163176

I. Introduction

PEOPLE can now use their smartphone apps to access a variety of Internet services, including instant messaging (e.g., WhatsApp, WeChat), online socializing (e.g., Twitter, Weibo), electronic commerce (e.g., Amazon, Taobao), and online payment (e.g., PayPal, Alipay). These services have become an important part of the infrastructure of the modern information society, making smartphone apps a necessity in daily life [1]–[3]. According to a report from Statista [4], the number of apps available in Google Play, the official app store of Android, has increased exponentially from 16,000 in December 2009 to 2,893,806 in July 2021. The app market is expected to generate 935.2 billion US dollars in business value by 2023 [5]. Such a vast and vital app market has attracted developers and service providers to investigate app usage behavior to better develop and deliver mobile apps.

Understanding app usage behaviors has significant implications for all relevant stakeholders, including smartphone manufacturers, network operators, market intermediaries, app developers, and end consumers [6]. To improve device performance and extend usage time, smartphone manufacturers can optimize the scheduling of various smartphone resources, such as CPU, memory, and battery power, based on the usage patterns of specific apps [7], [8]. Based on app traffic patterns, network operators can dynamically optimize traffic offloading schemes and improve network services [9], [10]. Furthermore, network operators and market intermediaries can provide personalized services, such as accurate recommendations and targeted advertisements, by profiling mobile users' preferences and interests from their app usage behaviors. By doing so, operators and intermediaries can improve the quality of experience (QoE) while increasing profits [11], [12]. App developers can better understand customer satisfaction and market trends by analyzing app usage and profiling app popularity, which may provide excellent guidance for upgrading existing apps and designing new apps [13].

In recent years, extensive research efforts have been invested in understanding user behaviors using data-driven methods based on mobile app usage data. As shown in Fig. 1, we counted the number of papers published in the field of app usage data analysis. We can see that this research area began

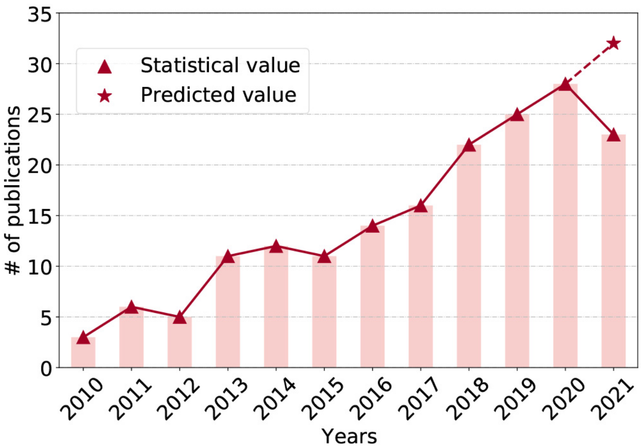


Fig. 1. Volume of publications in the field of app usage data analysis. The statistical values are up to Aug. 2021. The predicted value for 2021 is based on previous years' values.

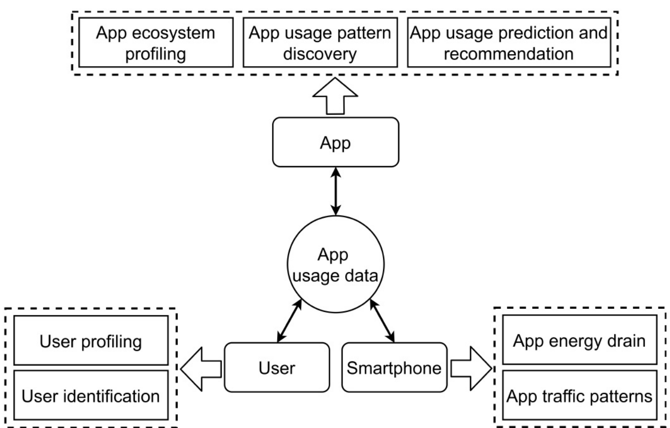


Fig. 2. Structure of existing studies for app usage analysis.

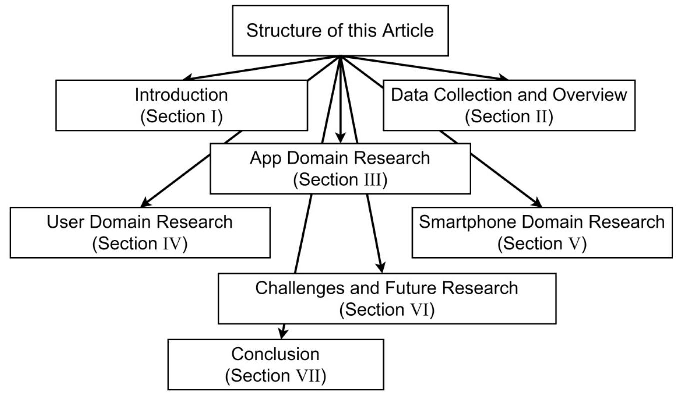


Fig. 3. Structure of the survey article.

in 2010, then grew slowly until 2012. Most notably, after 2012, there was a rapid increase in the number of publications. Overall, the volume of publications in this field is expanding, indicating a burgeoning research field. Thus, we aim to conduct a comprehensive literature survey in response to significant recent progress in mobile app usage analysis.

The app usage data are collected through surveys [14], monitoring apps [8], network operators [15], and app stores [16]. These data form a cross-domain and multi-view data ecosystem that includes various app usage behaviors, e.g., downloading, installing, launching, uninstalling apps, contextual information about app usage, e.g., time, location, traffic, energy consumption, and information about apps, e.g., app description, app rating, and user reviews. As a result, the app usage data reflects the characteristics of apps, users, and smartphones.

Existing studies for app usage data analysis fall into three domains: app domain, user domain, and smartphone domain, as shown in Fig. 2. From the app perspective, app domain research aims to reveal app features from app usage data. The research topics in this area include app ecosystem profiling, app usage pattern discovery, and app usage prediction and recommendation. App ecosystem profiling looks into the inherent characteristics of mobile apps and app markets, such as app categorization and popularity modeling. The goal of app usage pattern discovery is to find regularities in mobile app usage behaviors. Predicting which apps will be released and inferring users' preferences for undiscovered new apps are the goals of app usage prediction and recommendation.

The goal of user domain studies is to discover the link between user characteristics and app usage behaviors. User profiling and user identification are two research topics in this domain. User profiling focuses on group-level user attributes like gender, occupation, income, and personality. User identification, focusing on the individual level, aims to identify a user based on his or her app usage behaviors.

The smartphone domain research focuses on smartphone characteristics, with two main areas of investigation: app energy drain and app traffic patterns. These two areas aim to improve smartphone performance by analyzing app energy and traffic consumption patterns.

In this survey, we aim to answer the following questions. How should one collect mobile app data? What are the benefits and limitations of the various data collection methods? What are the critical features of app usage datasets that have been used in existing studies? What are the main research topics in mobile app usage analysis? What are the most common data-driven methods used, the significant findings, and the main results in existing studies? What are the challenges and promising future directions in mobile app usage analysis?

To answer these questions and draw conclusions, we systematically review the recent research. Fig. 3 shows our survey paper's structure. In Section II, we describe and compare data collection methods. We also present a set of public datasets and discuss privacy and ethical issues. Section III summarizes existing studies in the app domain. Section IV reviews research efforts in the user domain. Section V reviews existing studies in the smartphone domain. We compare app usage datasets, methods, and findings across all research topics. In Section VI, we discuss the challenges and future research in mobile app usage analysis. Section VII concludes the survey.

The main contributions of this paper can be summarized as follows.

TABLE I
SUMMARY OF EXISTING SURVEY ARTICLES ON APP USAGE ANALYSIS

Survey paper	Contributions	Limitations
[17]	A summary of studies on app usage prediction and recommendations.	This survey is limited to a single topic and does not provide a complete picture for studies analyzing app usage data. As a result, it fails to connect app usage prediction and recommendations to other important topics such as app usage pattern discovery. Recent advances in mobile app usage prediction and recommendations, such as deep-learning and graph-learning based methods, are not covered in this survey.
[11]	An overview of studies working on user profiling based on their app usage behavior.	This survey is limited to a single topic and does not provide a complete picture for studies analyzing app usage data. The relationship between descriptive and predictive user profiling is not discussed in this survey, which focuses primarily on predictive profiling. Furthermore, the key findings of the relationships between user attributes and app usage behaviors are not summarized in this survey.
[13]	A survey of mobile app user understanding and marketing, based on a crowdsourced data-driven perspective.	This survey only looks at studies that use crowdsourced datasets and does not compare various data collection methods. This survey does not reveal connections between different topics of mobile app usage analysis. The key findings gleaned from app usage behaviors have yet to be summarized in this survey.

- A comprehensive survey of data collection methods used to collect app usage data is presented, including surveys, monitoring apps, network operators, and app stores. We discuss data characteristics and compare the benefits and limitations of various data collection methods. We also gather and introduce prominent app usage datasets to the research community for further study.
- A comprehensive review of current research in the field of smartphone app usage data analysis is presented, including primary research topics in the app, user, and smartphone domains. As for each research topic, we make a thorough comparison in terms of the problem they studied, the datasets they used, the methods they applied, and the significant results they achieved.
- Advanced technologies and key findings gleaned from app usage behaviors are thoroughly summarized, with significant implications for all relevant stakeholders in the research community and industries.
- The challenges and future research opportunities concerning mobile app usage analysis are identified. We give a thorough discussion of the challenges in data and methods. We also go over future research topics, such as app evolution, context-aware app usage modeling, deep reasoning of app usage behaviors, linking physical activities and app usage, and location-based services and urban computing.

There are only a few surveys on related topics. Cao and Lin [17] surveyed studies that looked at app usage prediction and recommendations. Zhao *et al.* [11] summarized studies on user profiling based on their app usage behaviors. Guo *et al.* [13] conducted a review of the literature using app usage data gathered through crowdsourced methods. Existing surveys are either limited to a single data source [13] or a single small topic [11], [17]. Table I summarizes the contributions and limitations of existing survey articles on app usage analysis. In comparison, our survey examines the state-of-the-art studies in all app, user, and smartphone domains, as well as the connections between them. We provide a systematic review of existing datasets, commonly used analytical methods, and key findings and results.

II. DATA COLLECTION AND OVERVIEW

The foundation and core parts of smartphone app usage analysis are the real-world app usage data. We will introduce and compare different data collection methods in this section, and we will present a set of public datasets. We will also discuss privacy and ethical concerns.

A. Data Collection Methods

1) Surveys: Surveys are a simple and important method to collect information about app usage. Researchers generally create a questionnaire based on their goals and then collect data from the participants' responses. In practice, researchers should pay close attention to questionnaire design, particularly question wording, to reduce respondent bias. The American Association for Public Opinion Research offers some tips on how to conduct a high-quality survey.[1]

There are several methods for conducting a survey, including telephone, mail, and the Internet. In-app surveys, which post questionnaires and collect user feedback directly inside mobile apps, have been popular in recent years and are often used to obtain app usage data. WhatsApp and Skype, for example, ask users to rate the quality of their audio calls at the end of the call. YouTube uses questionnaires to gather information about its users' preferences. In-app surveys require less maintenance than other methods such as telephone and face-to-face surveys [18]. In-app surveys can capture a lot of data because they can reach a large number of people lightly. However, in-app surveys, respondents only represent one group, namely app users, which may limit the generalizability and representativeness of the data obtained. In other words, in-app surveys offer a mode impact, whereas different survey modes result in different data being collected.

Surveys can collect only the coarse-grained app usage behaviors of users, such as which app store they used and

[1]https://www.aapor.org/Standards-Ethics/Best-Practices.aspx

the number of apps they downloaded per month [19], [20]. However, fine-grained app usage traces, such as when and what apps are launched, are difficult to collect with surveys. Participants may be hesitant to share sensitive information and may also give socially acceptable responses to the questions posed. As a result, there are biases in the app usage data gathered through surveys. Researchers should be aware of this and take proactive steps to mitigate biases, such as data sampling.

2) Monitoring Apps: One common data collection method is to use monitoring apps installed on participants' smartphones to record fine-grained app usage behaviors automatically. Researchers can use this data collection method on a small scale by recruiting volunteers [21] and on a large scale by publishing monitoring apps in app stores [8]. Recruiting volunteers can focus on a particular group of users, e.g., students [22] and older adults [23]. In addition, recruiting volunteers, such as selecting involved users based on their backgrounds and properties, can pre-control the quality of data. Publishing in app stores can improve data quality by filtering out noisy data and reducing bias by taking advantage of a large number of active users [24]. Notably, as a result of globalization, it is now easier to collect app usage data from multiple countries by publishing monitoring apps in international app stores such as Google Play and Apple Store. The results of such an analysis will be more general and representative.

Monitoring apps generally use an event-triggered collection scheme, which means they collect app usage data when an event occurs. The event can be user actions [25] (e.g., turning on the screen, launching apps, or typing), messages received [26] (e.g., notifications, e-mails), network requests [27], and hardware status [28] (e.g., CPU usage, battery levels). Researchers can collect a variety of user behaviors and control the granularity of data collection by properly selecting trigger events. Smartphones also have a variety of sensors [29], such as an accelerometer, a gyroscope, and a GPS. Monitoring apps can gather enough sensor data, such as CPU usage, movement status, GPS location, and battery status. This sensor data can be used to analyze app usage by providing sensor contextual information. It is worth noting that monitoring apps are available to obtain almost all kinds of smartphone usage behaviors as long as the user's consent is obtained.

3) Network Operators: Most apps nowadays rely on the Internet to provide their services. As a result, network operators can gather network data and deduce app usage traces from it. Network data can be collected and extracted from multiple network interfaces, such as from the serving gateway (SGW), the mobility management entity (MME), the access & mobility management function (AMF), or the session management function (SMF), as shown in Fig. 4. Deep packet inspection and deep flow inspection are typically used to infer app usage information from traffic flow records collected from the SGi and N6 network interfaces [30].

This data collection method is typically used by network operators and in large-scale measurements. The data collected generally cover most mobile users in an entire city [31] or a country [32]. Due to the large volume of network traffic

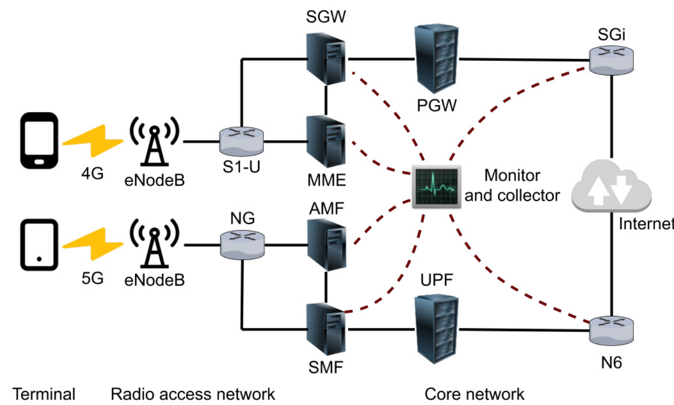


Fig. 4. Data-collecting points in networks. MME: Mobility Management Entity; SGW: Serving Gateway; PGW: Packet Data Network Gateway; SGi: Serving Gateway interface; AMF: Access & Mobility Management Function; SMF: Session Management Function; UPF: User Plane Function.

data, network operators have to use sample strategies, collecting network traffic records at regular intervals like every hour [33] or every several minutes [34]. The datasets collected by network operators typically include location data for app usage records. The location is approximated by the GPS location function of the associated base station.

An important step in this data collection method is to infer app usage activity from network traffic flows. One popular method is to use the HTTP header of network packets [35], [36]. Yao *et al.* [35], for example, created SAMPLE, a systematic tool that inspects the HTTP head and uses the destination domain and user-agent as the app identifier. SAMPLE can automatically generate the conjunctive rules that identify over 90% of apps with an average accuracy of 99%. However, due to security concerns, most apps now use the HTTPS protocol to transmit packets. Some research has focused on identifying app usage traces from encrypted data traffic [36]–[38]. Taylor *et al.* [37], for example, developed AppScanner, a system that used packets' side information, such as packet size, direction, and delay, to identify the app label. AppScanner was able to correctly identify the top 110 most popular apps with a 96% accuracy rate. Most importantly, when collecting data from network operators, researchers should be aware of the mode effect. This data collection method cannot be used to track app usage that does not generate network traffic.

4) App Stores: App stores have access to a variety of app usage data. The app store, which is an app management tool, records users' app management behaviors, such as downloading, updating, installing, and uninstalling apps. User preferences and app popularity are implicitly reflected in these management activities. Due to the user account mechanism, app stores can track the same user's behaviors across multiple devices, which is difficult to do with surveys and network measurements. Additionally, some app stores, such as Wandoujia,[2] a free Android store in China, support the monitoring of smartphone sensors. As a result, app stores can still provide sensor contextual information about app usage, such

[2]https://www.wandoujia.com/

TABLE II
COMPARISON OF DIFFERENT DATA COLLECTION METHODS

Collection methods	**Scale**	**App usage behaviors**	**Side information**	**Advantages**	**Limitations**
Surveys	Small-scale, large-scale	Coarse-grained usage behaviors	User profiles	• Available to collect user profiles. • Easier to conduct worldwide collections.	• Hard to conduct collections of millions of users. • Inevitable biases in the collected data. • Hard to collect fine-grained usage traces, like when and what apps are launched.
Monitoring apps	Small-scale, large-scale	All kinds of usage behaviors	Sensor data	• Easier to conduct worldwide collections. • Available to collect all kinds of app usage behaviors. • Available to collect smartphone sensor data.	• Hard to conduct collections of millions of users. • Hard to collect the data principally covering a certain area, like a city.
Network operators	Large-scale	Network access logs of mobile apps	Associated base stations	• Easier to conduct large scale collections. • Easier to obtain user mobility information. • Available to collect data for a specific area.	• Cannot cover usage records that do not generate network traffic. • Cannot collect smartphone sensor data. • Hard to conduct worldwide collections. • Hard to distinguish the same user with multiple devices.
App stores	Small-scale, large-scale	App management behaviors	App metadata	• Available to distinguish the same user with multiple devices. • Easier to conduct worldwide and large scale collections. • Available to access app metadata.	• Hard to cover all smartphone operating systems. • Hard to collect fine-grained usage behaviors, like launching and typing. • Hard to collect the data principally covering a certain area.

as network traffic, CPU usage, memory usage, battery status. However, because of the security concerns of iOS, iPhones cannot support third-party app stores. Thus, this data collection method is most concentrated on Android users.

App stores can provide various types of app metadata, such as app name, app category, app description, app rating, and user reviews. App metadata produces valuable information about smartphone apps and allows the collection of user feedback about them. App descriptions, for example, are frequently used to extract app functions for app categorization. For user profiling, app categories provide semantic meanings to explain usage activities. App ratings and user reviews are used to model app popularity and recommend apps based on their popularity and usage experience. App metadata can be crawled from app stores, such as Google Play and Apple Store, or provided directly by app stores.

Table II compares the four data collection methods mentioned above. In terms of data collection scale, all four approaches are capable of large-scale measurements involving tens of thousands of users, thanks to advancements in communication and network technologies. However, surveys and monitoring apps are limited by response rates or app popularity, which makes it difficult to collect millions of users. By recruiting volunteers, both surveys and monitoring apps can collect small-scale datasets. It is also possible to conduct control studies by carefully selecting participants, which is impracticable for network operators and app stores. Monitoring apps outperform the competition in terms of collectible app usage behavior. It is difficult for surveys to collect fine-grained usage traces because they are limited by questionnaires. Because users are directly responsible for their responses, they may be hesitant to share sensitive information but willing to provide socially desirable responses, resulting in biases in the data collected. Although other methods collect data automatically and avoid biases, network operators and app stores are still limited to specific types of behaviors, such as network access and app management. Monitoring apps, which are installed on smartphones and based on event-triggered collection, can be used to collect all types of app usage behaviors by selecting different trigger events. Different data collection methods can provide additional side information about app usage behaviors, such as user profiles, sensor data, location, and app metadata.

Every coin has two sides. Different collection methods have their own set of benefits and drawbacks. In practice, we must select data collection methods with care and precision to answer the research question. For example, surveys, monitoring apps, and app stores are better choices than network operators for investigating country and culture differences in mobile app usage behaviors because it is difficult to conduct worldwide data collection from network operators. App stores and monitoring apps are also better if we want to track a user's usage behaviors over time because they can link the same user based on user accounts even if the user changes smartphones. Different collection methods can be combined at times. For example, operator data can be used to discover app usage patterns across millions of users and then conduct control studies on small-scale datasets collected from monitoring apps or surveys to verify the discovered patterns.

B. Public Datasets

In this section, we present nine datasets that are publicly available for further research to the research community.

We introduce the collection method, collection period and collection items of these datasets in detail.

1) Worldwide Survey Dataset: This dataset was conducted through a survey methodology in 2012, collecting mobile app usage information of 10,208 people from more than 15 countries [19]. The countries include the United States, Canada, Mexico, the United Kingdom, Australia, France, Germany, Italy, Spain, Brazil, Russia, India, Japan, China, and South Korea. The collected items for mobile app usage behaviors include the most frequently used app stores and app categories, the reasons that lead users to find apps, and the reasons that lead users to download and abandon apps. It is worth noting that this dataset only contains usage data for app categories instead of individual apps. The dataset includes participants' demographics, including age, gender, nationality, marital status, country of residence, first language, ethnicity, education level, occupation, and income. This dataset also collects participants' personality traits by using the Big-Five personality measurement. This dataset and corresponding questionnaire are public, available at 'http://www0.cs.ucl.ac.uk/staff/S.Lim/app_user_survey/'.

2) Mobile Data Challenge (MDC) Dataset: The Lausanne Data Collection Campaign (LDCC) developed a specific monitoring app based on Nokia platforms to implement smartphone data collection. From October 2009 to March 2011, they collected a longitudinal smartphone dataset from nearly 200 volunteers. The volunteers are primarily dispersed throughout the Lake Geneva region of Switzerland. Large amounts of smartphone data, such as app usage behaviors (launch, foreground, close, and view), location (GPS, WLAN), motion (accelerometer), and proximity (Bluetooth), are all recorded in the dataset. The dataset was released to the research community in the Mobile Data Challenge (MDC) [39] held by Nokia, which is available at 'https://www.idiap.ch/dataset/mdc'.

3) LiveLab Dataset: This dataset was collected through a monitoring app called LiveLab [40]. LiveLab is an iOS app used to gather smartphone usage data from 24 iPhone users over the course of a year, from February 2010 to February 2011. The dataset specifically tracks two types of app usage behaviors: app launches and changes to the foreground app. The dataset also includes contextual information about user behaviors gathered from iPhone sensors such as the accelerometer, GPS, battery, Bluetooth, and WiFi status. It is important to note that this data was collected from skewed samples; all 24 volunteers are Rice University students. This dataset was released to the public and can be accessed via 'http://yecl.org/livelab/traces.html'.

4) Carat Dataset: Carat [8], a monitoring app, was used to collect this dataset. Carat applies an event-triggered collection scheme, gathering a data sample every time the battery level changes by 1%. Each data sample contains a list of apps used, a user-specific identifier, and a timestamp. Carat also gathers sensor contextual data, such as battery level, battery status, mobile country code. Carat is available on both Google Play and Apple Store, which the developers hope will increase the number of people who participate in data collection. In this way, they collect data from all over the world. Carat has collected data from over 500,000 mobile users in over 100 countries so far.[3] The dataset owners made a long-term app usage dataset available to the research community. The top 1,000 users ranked by total time spent using Carat from 2014 to 2018 are included in the public long-terms app usage dataset. The user with the longest duration in the public dataset has 18,146,042 time-series records spanning 4.65 years, and even the user with the shortest duration has more than two years of records. The public dataset is available at 'https://www.cs.helsinki.fi/group/carat/data-sharing/'.

5) TalkingData: The dataset is collected using the TalkingData Software Development Kit (SDK), which is integrated into monitoring apps. This dataset contains the complete app usage traces for over 70 thousand users from May 01, 2016, to May 07, 2016. Each app usage log includes an anonymous user ID, timestamp, app ID, and location information (longitude and latitude). This dataset contains demographic information about the users involved, such as age and gender. Researchers should be aware of the mode effect and biases in the analysis when using this public dataset because the users involved are mostly from mainland China. This dataset was released in a Kaggle challenge, TalkingData Mobile User Demographics, and can be accessed via 'https://www.kaggle.com/c/talkingdata-mobile-user-demographics/'.

6) PhoneStudy Dataset: The PhoneStudy smartphone research app for Android was used to collect this data [41]. Between September 2014 and January 2018, 743 volunteers were recruited through forums, social media, blackboards, flyers, and direct recruitment. The activities of users on their smartphones were logged in the form of time-stamped logs of events. Those events included calls, app starts/installations, screen de/activations, contact entries, texting, global positioning system (GPS) locations, etc. The dataset is publicly available at 'https://osf.io/kqjhr/'.

7) Context-Aware App Usage Dataset: This dataset was collected by a primary Internet Service Provider (ISP) in China [42]. The data was collected over the course of one week in April 2016, and it covered the entire metropolitan area of Shanghai, one of the world's largest cities. The SAMPLES [35] tool was used to identify app usage records from network metadata, and the traffic generated by the 2,000 most popular apps was successfully recognized. Each app usage record contains an anonymized user ID, timestamp, base station ID, app ID, and traffic volume. The top 1,000 active users in the app usage dataset are released to the research community. The public dataset is available at 'http://fi.ee.tsinghua.edu.cn/appusage/''.

8) Wandoujia Dataset: Wandoujia is a leading app store in China [43]. The dataset was collected from over 17 million users and recorded their app management activities, such as downloading, updating, and uninstalling apps, from May 1 to September 30, 2014. Wandoujia also gathered daily network activities for each app, such as traffic volume and Wi-Fi and cellular network connection time. A portion of this dataset has been made available to the public. The public dataset is available at 'http://www.liuxuanzhe.com/appdata/'.

[3] http://carat.cs.helsinki.fi/.

9) iOS Apps Dataset: This dataset [44] was extracted from the Apple Store and covers the metadata of removed iOS apps for a period of 1.5 years, from January 1, 2019, to April 30, 2020. There are 1,129,615 app records in total, corresponding to 1,033,488 unique mobile apps. The dataset contains app objective data such as app name and release date, app subjective data such as app ratings and reviews, and app popularity data such as app ranking. The information is open to the public and can be found on 'https://github.com/LuckyFQ/iOS-Removed-Apps-Dataset'.

In Table III, we summarize the above seven public datasets in terms of collection method, collection area, collection date, number of apps and users, collected items, and availability status.

C. Privacy and Ethical Considerations

In 2018, the General Data Protection Regulation (GDPR) became enforceable in the European Union. GDPR gives European citizens greater control over their data and raises awareness of privacy issues and the value of their personal data [45]. GDPR, as a strong regulation for collecting and processing personal data, imposes a number of requirements. The following are the most relevant and important data collection principles. First, personal data can only be collected if the user has given consent for a specific purpose. Namely, user consent is only used for a specific purpose. Second, the data minimization principle limits data collection to the minimum required to achieve the app's purpose. Third, individuals have the right to withdraw their consent and erase their personal data. The GDPR has had a significant impact on how mobile app usage data is collected. According to a recent study [46], the number of permission items used by apps decreased significantly after the GDPR was implemented. This finding suggests that monitoring apps are cautious about the data they collect to follow to the GDPR's data minimization principle. Furthermore, the data collection section is also responsible for consent management. Many app vendors and app stores have mechanisms in place to respond to user requests, such as erasing data and withdrawing consent [47].

Smartphone app usage data, as a type of sensitive personal data, must be handled with caution when it comes to ethics. When collecting, processing, and analyzing app usage data, researchers must adhere to the principle of Data for Good [48]. First, app usage data should be used for the greater good of humanity and society. We should use data to improve people's lives, such as by improving interpersonal relationships, providing convenient and ubiquitous services, improving user experience. However, some dangerous and terrible things have occurred in recent years. Cambridge Analytica, for example, abused the data of millions of Facebook users [49]. To prevent such things from happening again, apart from government regulations, we still need data scientists to hold themselves to a high standard. Second, we should make good use of app usage data. We should adhere to the principle of end-user transparency by outlining the items collected, the purpose of data collection, the potential privacy risk. Collection and analysis should be done in a way that protects people's privacy. We must also take into account the ethical decisions that developed systems will make.

III. App Domain Research

App features such as functionality and popularity have a significant impact on mobile app usage behavior. Many efforts have been made in recent years to examine mobile app usage data from the standpoint of apps by analyzing app features. App ecosystem profiling, app usage pattern discovery, and app usage prediction and recommendation are the three main topics in app domain research, according to our careful review of existing studies. In this section, we will summarize existing app domain studies on the three topics and compare their datasets, methods, and key findings.

A. App Ecosystem Profiling

Millions of mobile apps form a symbiotic mobile app ecosystem [50]–[52]. App ecosystem profiling aims to investigate the inherent characteristics of mobile apps, focusing on two main sub-problems: app categorization and popularity modeling.

1) App Categorization: An app category groups apps that perform similar functions. WhatsApp and Skype, for example, are both classified as communication apps on Google Play. App categories are currently determined by app developers when they release apps in app stores [54], [55]. However, deciding on a category for an app can be difficult at times [56]. WeChat, for example, can be used for social networking as well as messaging. As a result, both the communication and social app categories appear to be appropriate. Moreover, manually selecting app categories opens the door to deliberate gaming [57], [58]. App developers may desire to avoid fierce competition by selecting less appropriate categories to improve their app's ranking. For the reasons stated above, many mobile apps are miscategorized in app stores. Hence, it is expected to have an effective and automated method for categorizing mobile apps.

Several studies looked into the possibility of automatically categorizing apps based on their descriptions. They generally modeled app categorization as a problem of short text topic modeling, which they solved with Natural Language Processing (NLP) tools. Al-Subaihin *et al.* [59] used Term Frequency-Inverse Document Frequency (TF-IDF), while Ochiai *et al.* [60] used a word embedding model to extract app features from app descriptions. Fuad and Al-Yahya [61] further used Latent Dirichlet Allocation (LDA) to analyze the app descriptions of 13,282 apps in Google Play. They discovered some new categories, including occasions, pandemics, and languages, that were not present in the original app store classification. Surian *et al.* [62], on the other hand, concentrated on the app market's miscategorization issue. To refine the category of apps, they developed a von Mises-Fisher distribution-based probabilistic topic model. They discovered that $0.35 \sim 1.10$ percent of game apps are miscategorized after running experiments on 48,663 game apps crawled from Google Play.

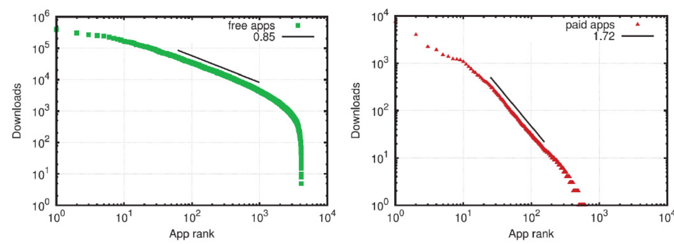
TABLE III
PUBLIC APP USAGE DATASETS

Name	Collection method	Collection area	Collection date	# Apps	# Users	Collected items	Public
Worldwide survey dataset	Survey	Worldwide	26 Sept. 2012-26 Nov. 2012	-	10,208	Used app stores, used app categories, demographics of users.	Whole dataset
MDC dataset	Monitoring app	Lake Geneva region, Switzerland	2009-2012	-	200	App usage, location, motion, and proximity.	Whole dataset
LiveLab dataset	Monitoring app	Houston, the USA	Feb. 2010-Feb. 2011	2,302	24	App usage, accelerometer, GPS, battery, Bluetooth, and WiFi state.	Whole dataset
Carat dataset	Monitoring app	Worldwide	2014-Present	Over 200,000	Over 500,000	App usage, accelerometer, GPS, battery, Bluetooth, and WiFi state.	Whole dataset
TalkingData	Monitoring app	China	01 May 2016-07 May 2016	Over 400,000	74,646	App usage, timestamp, location, user demographics.	Whole dataset
PhoneStudy Dataset	App store	Europe	Sept. 2014-Jan. 2018	-	743	App usage, timestamp, calls, screen de/activations, contact entries, texting, locations.	Samples
Context-aware app usage dataset	Network operator	Shanghai, China	19 April 2016 - 26 April 2016	2,000	1.6 million	App usage, app category, timestamp, traffic, cellular location, POI.	Samples
Wandoujia dataset	App store	China	1 May 2014 - 30 Sept. 2014	Over 1 million	17 million	App download, update, uninstall, device model, app rating, network activity.	Samples
iOS apps dataset	App store	Worldwide	1 Jan. 2019 - 30 Apr. 2020	1,033,488	-	App name, release date, app ratings, app reviews, and app ranking.	Samples

Untrusted developers may manipulate app descriptions, affecting categorization accuracy. As a result, some researchers used customer-generated app usage data to improve app categorization, assuming that different types of apps have different usage patterns. For example, Zhu *et al.* [63] extracted contextual features of mobile apps from usage records, such as time, location, and battery level. They combined all of the features into a Maximum Entropy model [64] for app categorization. Radosavljevic *et al.* [65] took advantage of users' app installation behavior to categorize apps. He *et al.* [66] explored sequential characteristics of app usage. They used time series to formalize users' app-launching behavior and extracted app features using dynamic time warping [67] across time series. After conducting experiments on 3,086 apps, they concluded that app usage traces are a good source for app categorization.

2) Popularity Modeling: App popularity, as measured by chart rankings, user ratings, and downloads, reflects the user experience of mobile apps. App popularity modeling is critical for developers and app market intermediaries to understand the factors that influence app adoptions and then make appropriate decisions about which apps to develop, release, and remove [44], [68].

Modeling app popularity by depicting popularity distributions is one of the most common approaches used in existing research. Many studies have demonstrated that app popularity has a power-law distribution [69]–[72]. Petsas *et al.* [53] further discovered that paid and free apps have different popularity distributions. Paid app popularity follows a power-law distribution with truncated edges, as shown in Fig. 5, whereas free app popularity follows a clear power-law distribution. Such a discovery provides valuable prior knowledge for downstream applications. For example, the recommendation system for free apps does not need to consider the long tail effect. Li *et al.* [73], [74] studied the long-term evolution of app popularity from 2012 to 2017, spanning six years. They discovered that the popularity of apps exhibits a typical Pareto effect. Shen *et al.* [75] discovered that, in addition to the Pareto effect, app popularity exhibits a Matthew effect when considering time-varying data. In other words, high-ranked apps tend to receive higher ratings and have more consistency in top app lists.



(a) Free apps. (b) Paid apps.

Fig. 5. The downloads of free apps follow a power-law distribution with truncated edges, while paid apps follow a clear power-law distribution [53].

Some researchers investigated the factors that influence app popularity. Fu *et al.* [76] used LDA to extract key topics from over 13 million user reviews of 171,493 apps on Google Play. They identified the top five user concerns as attractiveness, stability, accuracy, compatibility, and connectivity. Potharaju *et al.* [77] investigated the relationship between the price and popularity of paid apps and discovered that the price of paid apps does not affect app popularity. Lu *et al.* [78] noticed that different mobile device models have an impact

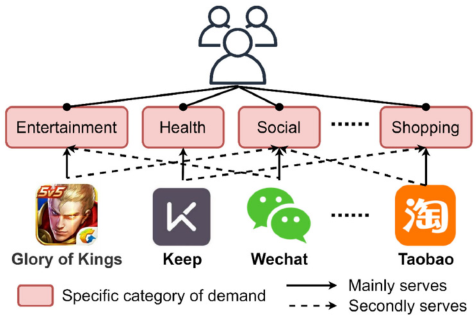


Fig. 6. A three-level hierarchy structure to model app adoption [79].

on app adoption. Lower-end users, for example, prefer Opera Mini, whereas higher-end users prefer Chrome. Shen *et al.* [75] discovered that app release strategy has an impact on app popularity. A timely subsequent release has a higher chance of turning a descending rating trend than a late release.

Some studies have made efforts towards app popularity prediction. Zhu *et al.* [80] used chart rankings and user ratings to define app popularity states and proposed a popularity-based hidden Markov model (PHMM) to predict the popularity states of mobile apps. Wang *et al.* [79] proposed an evolutionary hierarchical competition model (EHCM) to predict app downloads by using a three-level hierarchy structure to model app adoption. As shown in Fig. 6, the top-level in the app adoption model is users; the middle-level is categories of user demands; and the bottom-level is apps. Each app serves several categories of user demands and competes with other apps by providing engaging functions. In terms of model design, EHCM has higher interpretability than PHMM. Ouyang *et al.* [81] further proposed a Multivariate Hawkes Process-based prediction model (MHP). The MHP jointly considered exogenous stimuli, e.g., updates, reviews, ratings, and endogenous excitations, e.g., historical popularity, app ages. MHP outperforms PHMM and EHCM, according to extensive experiments. Zhang *et al.* [82] leveraged deep learning to predict app popularity and proposed DeePoP, an RNN-based prediction model. DeePoP considered time-varying app interactions and modeled their impact on app popularity. DeePOP outperforms MHP and EHCM, effectively lowering the Root Mean Square Error (RMSE) to 0.088. Li *et al.* [83], on the other hand, argued that app ratings do not accurately reflect app quality or user preferences. As a result, they proposed modeling user preferences and predicting app popularity using uninstallation-downloading sequential activities.

3) Discussion: Existing research primarily examined the app ecosystem in terms of app categories and popularity. Table IV summarizes prominent literature by dataset size, data type, methods, results, and findings. As can be seen in Table IV, the study scale ranges from 600 to over 1 million apps. The majority of the datasets used were obtained from app stores. App descriptions are commonly used for app categorization. To extract app features from app descriptions, NLP tools such as TF-IDF, LDA, and probabilistic topic models are used. However, because untrustworthy developers may operate app descriptions, there is a trend to use app usage data, such as installation and usage behaviors, to categorize apps more reliably.

In popularity modeling, app downloads, rankings, and ratings are commonly used to indicate app popularity. Descriptive statistics were used to depict the distribution of app popularity and analyze factors influencing app popularity. Time-series methods, like the Markov model, Wavelet transform, Hawkes Process, and Recurrent Neural Network, are used to predict app popularity. In addition to positive activities such as downloading and installing apps, some studies tried to use negative behaviors such as uninstalling apps to model app popularity.

B. App Usage Pattern Discovery

App usage pattern discovery aims to find regularities in usage behaviors to learn typical usage habits and improve the quality of user experience. Existing studies focused primarily on two sub-fields: contextual pattern discovery and temporal pattern discovery.

1) Contextual Pattern Discovery: The goal of contextual pattern discovery is to investigate the connection between app usage and contextual factors [84]. As shown in Table V, there are three different types of context: sensor context, usage context, and social context. Sensor context, collected from smartphone sensors, includes location, time, battery levels, movement status of users, WiFi or cell connectivity, etc. Usage context refers to the prior and posterior used apps. Social context refers to the social environments in which an app is used, such as with friends, family members, strangers, coworkers, and so on.

Locations, as a sensor context, have a significant impact on app usage. Mehrotra *et al.* [85] discovered that student users are more attentive to app notifications at college, in libraries, on the streets, and in residential areas. However, users at religious institutions are less receptive to app notifications. Do *et al.* [86] and Böhmer *et al.* [87] found that while waiting for and during trips, users prefer to use Web and multimedia apps. Graells-Garrido *et al.* [88] further discovered that street types have an impact on app usage. Message apps, for example, generate more traffic on main streets, whereas dating apps are more popular on pedestrian streets. Xia and Li [89] and Ren *et al.* [90] demonstrated how the function of a location influences app usage in that location.

Time is also an important sensor contextual factor that influences app usage. Böhmer *et al.* [87] discovered that news apps are most popular in the morning, while game apps are at night. Van Canneyt *et al.* [91] and Li *et al.* [92] noted that some specific dates and events, such as New year's day, UEFA European Championship, and Covid-19, will disrupt users' regular app usage patterns. Several studies investigated the relationship between app usage and other sensor context factors [86], [93], like Bluetooth data, battery levels, and movement status. For example, Do *et al.* [86] studies how Bluetooth density affects app usage and discovered that communication apps are more likely to be used with high Bluetooth density.

App usage is still strongly linked to the usage context, i.e., prior and posterior used apps. The fact that several apps need

TABLE IV
LITERATURE REVIEW IN APP ECOSYSTEM PROFILING

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	Methods	Results and findings
App categorization	[59]	17,877	-	-	App store	App description	TF-IDF, hierarchical clustering	300 apps were manually verified, revealing that current categorizations could be improved.
	[61]	13,282	-	-	App store	App description	LDA	New categories such as occasions, pandemics, and languages were discovered.
	[62]	48,663	-	-	App store	App description	Probabilistic topic model	$0.35 \sim 1.10$ percent of game apps were detected to be miscategorized.
	[63]	680	443	1 year	Monitoring app	App description, contextual features (e.g., time, location, battery level.)	Maximum entropy model	Contextual features can help improve the accuracy of app categorization.
	[65]	200,000	-	-	Monitoring app, app store	App description, app install sequence	Skip-gram, kNN classification	App categorization benefits from app installation behavior.
	[66]	3,086	3,000	-	Network operator	App usage sequence	Dynamic time warping, SVM	App usage traces are a good data source for app categorization.
Popularity modeling	[69]	260,172	8,112,145	1 month	App store	App download	Descriptive statistics	App popularity follows a power-law distribution.
	[53]	20,000	-	6 months	App store	App download, app price	Descriptive statistics	Paid and free apps exhibit different popularity distribution.
	[75]	17,820	-	105 days	App store	App ranking, rating, version information	Descriptive statistics, multinomial naive bayes model	App popularity exhibits a Matthew effect. App release strategy affects app popularity.
	[73]	18,000	1,465	6 years	Monitoring app, app store	App usage	Descriptive statistics	The popularity of apps exhibits a typical Pareto effect.
	[76]	171,493	-	-	App store	App rating, user review	LDA	Attractiveness, stability, accuracy, compatibility, and connectivity are the top five user concerns in app adoption.
	[77]	247,223	-	6 months	App store	App download, app price	Descriptive statistics	Changing app price does not affect app downloads.
	[78]	238,231	4,775,293	3 months	App store	App adoption, device model	Collaborative filtering	Different mobile device models have an impact on app adoption.
	[80]	600	-	21 months	App store	App download, ranking, rating, review topics	PHMM	Predict popularity states of apps.
	[79]	20,000	-	550 days	App store	App download, app categories	EHCM	Predict app downloads from both user demand category level and product level.
	[81]	1,023	-	13 months	App store	App update, user reviews, rating, download	MHP	MHP outperforms PHMM and EHCM.
	[82]	1,023	-	1 year	App store	App download, user reviews	DeePoP	DeePOP outperforms MHP and EHCM, effectively lowering the RMSE to 0.088.
	[83]	1,054,969	1 million	5 months	App store	App rating, install, uninstall, update	Ridge regression, Lasso regression	Users' uninstallation downloading sequential activities can be used to predict app popularity.

TABLE V
TYPES OF CONTEXTUAL FACTORS

Sensor context	Usage context	Social context
Time, Location, battery levels, movement status, etc.	Prior and posterior used apps	With friends, family members, coworkers, etc.

to work together to complete a single task causes such correlations [94], [95]. Rahmati *et al.* [96] first demonstrated the app usage dependency of one-nearest prior used apps and found that such a dependency remains relatively constant for one to three months. By analyzing a dataset spanning seven years, Fan *et al.* [97] looked into how usage context changes over time. Huang *et al.* [34] and Liu *et al.* [98] identified frequent cooccurrence app sets. They discovered that e-commerce and online payment apps, such as Taobao and Alipay, are frequently used together to complete the task of online shopping. Tseng and Hsu [99] found that usage context dependency has sequential characteristics. For example, using the Camera app first and then the Album app is twice as likely as using the Album app first and then the Camera app.

App usage is also influenced by the social context. Ferreira *et al.* [100] classified the social context into four types alone, with friends, with strangers, and others. They observed a significant correlation between app usage and social context with a p-value of 0.002. Shema and Acuna [101] used app usage sequences to predict social context, achieving an accuracy of 70% with a random forest model. Kloumann *et al.* [102] and Taylor *et al.* [103] discovered that social contexts have varying degrees of influence on mobile app usage. The context of friendship has a greater impact when compared to family members.

2) Temporal Pattern Discovery: Unlike contextual patterns, which focus on static analysis, temporal patterns look into the dynamics of app usage. The diurnal pattern is a basic temporal pattern of app usage behavior that has been discovered by numerous studies [104]–[106]. A typical diurnal

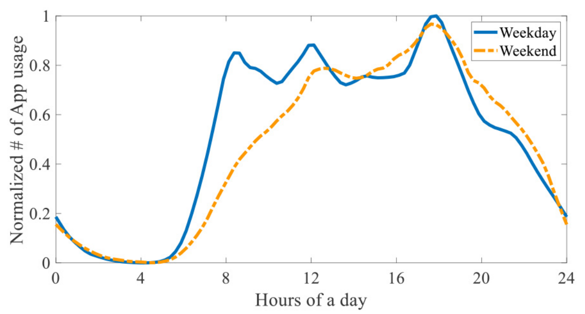


Fig. 7. A typical diurnal pattern of app usage [104].

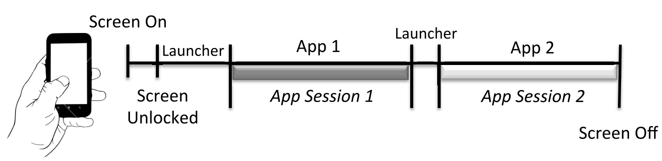


Fig. 8. Diagram of app usage sessions [113]. For each app usage, only one app is in the foreground and the usage will last for a period.

pattern of app usage is shown in Fig. 7. The intensity of app usage increases during the day and decreases during the night. However, in different scenarios, such a diurnal pattern may change. The diurnal pattern presents differently at various granularities [105], such as bytes, packets, flows, and users. For example, in terms of the number of bytes and packets, app usage is still highly active at night from 20.00 to 24.00, but not observed in flows and users. Also, different app categories may have distinct diurnal patterns. In contrast to other app categories, the diurnal usage pattern of transportation apps has more than two peaks on weekends [106]. Meanwhile, some researchers attempted to model the temporal patterns of app usage over the course of a single day. Kostakos *et al.* [107] used a time-sensitive Markov model, while Do and Gatica-Perez [108] represented app usage traces in one day as a bag-of-apps to infer the underlying structure of daily app usage.

The temporal traces of mobile app usage, on the other hand, can be represented as a series of app sessions. An app session, as shown in Fig. 8, is a period during which a user is actively engaged with the app. Some studies looked into the temporal patterns of app sessions. Silva *et al.* [109] conducted description statistics on app session time and discovered that map and media apps have longer app sessions. As users' activities for app advertisement are dependent on app session time [110], Rula *et al.* [113] proposed using decision trees to predict app session time for better advertising. Jones *et al.* [111] and Cao *et al.* [114] analyzed users' app re-visitation patterns across sessions and discovered three distinct patterns: checkers, waiters, and responsives. Users who exhibit brief revisit patterns of fast re-visitation are referred to as checkers (less than one hour). Users with longer revisit patterns, uniformly distributed between short-medium re-visitations (between 1min and 4hrs) and long re-visitations, are referred to as waiters (from 2hrs to 3days). Users who exhibit both brief and long revisit patterns are referred to as responsives. Leiva *et al.* [112] looked into the disrupted patterns of app sessions and discovered that app sessions disrupted by phone calls resulted in task completion delays of up to four times.

3) Discussion: Contextual pattern discovery and temporal pattern discovery are two sub-topics of existing research on app usage pattern discovery. Table VI summarizes prominent literature in terms of dataset information, methods, and patterns discovered. We can see that datasets vary greatly according to size and duration. The majority of datasets are collected from monitoring apps and network operators because pattern discovery requires fine-grained usage behaviors, such as launching apps and switching apps. The studies on contextual pattern discovery focus on static patterns to discover relationships between app usage and context factors. Statistical methods such as Pearson correlation, analysis of variance, hypothesis testing, and posterior probability are commonly used to investigate correlations. These studies pave the way for context-aware app prediction and recommendation. Temporal pattern discovery, on the other hand, focuses on dynamic patterns to investigate how app usage changes over time. In this subtopic, descriptive statistics and time series analysis are frequently used methods. It is worth noting that some studies looked into temporal patterns and used them to predict and recommend app usage. In the following section, we will go over these studies in detail.

C. App Usage Prediction and Recommendation

1) App Usage Prediction: App usage prediction aims to forecast the apps that will be launched so that smartphones can load app-related resources in advance to cut down the searching and loading time [115]–[117].

App usage prediction is solved by examining usage rules based on historical user behavior. Predicting the most frequently used (MFU) and most recently used (MRU) apps [118], as the basic benchmark method, takes advantage of the most basic regularity of user behavior. Furthermore, probabilistic models are used to investigate complex regularity [119]–[121]. Natarajan *et al.* [119] modeled the app usage sequence as a Markov chain and utilized the one-order transition probability for prediction. Zou *et al.* [120] considered high-order app transitions to improve the prediction accuracy by leveraging a bayesian network model. Baeza-Yates *et al.* [121] used all apps used in the recent time window for prediction. They achieved an accuracy of 85.7% by using a parallel tree augmented Naive Bayesian network. Thanks to recent advances in deep learning, Xu *et al.* [122] and Lee *et al.* [123] proposed using long short-term memory (LSTM) to capture the temporal regularity for app usage prediction.

App usage is also influenced by a variety of sensor contexts, as shown in Section III-B. As a result, many existing studies used various kinds of sensor contextual information, such as time, location, and movement status, to improve prediction performance instead of solely relying on users' historical app usage traces.

Time is widely used for app usage prediction as a fundamental sensor contextual feature, indicating the hour of one day or

TABLE VI
LITERATURE REVIEW IN APP USAGE PATTERN DISCOVERY

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	Methods	Results and findings
Contextual pattern discovery	[85]	-	26	2 weeks	Monitoring app	App launching, app notification, time, location	ANOVA	Users' attention towards app notifications is related to locations.
	[86]	11	77	9 months	Monitoring app	App usage, time, location, Bluetooth data	Pearson correlation	App usage depends on locations and bluetooth density.
	[87]	22,626	4,125	5 months	Monitoring app	App usage, time, location	Descriptive statistics	App usage is related to time of a day and locations.
	[88]	32	-	15 days	Network operator	App usage, time, location	Negative Binomial regression	Different types of streets have different effects on app usage.
	[89]	2,000	1,700,000	1 week	Network operator	App usage, time, location	Hidden Markov Model	The function of a location influences app usage in that location.
	[91]	230,000	600 million	-	Monitoring app	App usage, time, event	Statistic significance	Special events disrupt users' normal app usage patterns.
	[96]	100	24	1 year	Monitoring app	App usage, sensor context, time	Posterior probability	The dependency of usage context remains relatively constant for one to three months.
	[34]	2,000	1.7 million	1 week	Network operator	App usage, time	FP-Growth	Identified frequent app usage set, like Taobao and Alipay, WeChat and Taobao.
	[99]	1,132	25	7 months	Monitoring app	App usage, time	Apriori algorithm, PrefixSpan	The dependency of usage context has sequential features.
	[100]	-	21	3 weeks	Monitoring app	App session, time, location, screen state, social context	Hypothesis testing	A significant correlation between app usage and social context.
	[103]	-	108	2 years	Survey	App used, social relation	Correlation, logistic regression	Compared with family members, friends have a stronger effect on app usage.
Temporal pattern discovery	[105]	100	-	32 hours	Network operator	App usage, time, traffic data	Descriptive statistics	There is a slight difference between the diurnal patterns at different granularities.
	[108]	5	111	230,000 hours	Monitoring app	App usage, time	Author-topic model	Represent usage traces in one day as a bag-of-apps to infer the underlying structure of daily app usage.
	[107]	-	218	90 days	Monitoring app	App usage, screen state, time	Time-sensitive Markov model	Model temporal patterns of app usage.
	[109]	7	5,342	1 year	Monitoring app	App session, time, traffic	Descriptive statistics	Map and media apps have a larger duration of app sessions.
	[110]	206	15,423	8 months	Monitoring app	App session, time, purchase activity	Hypothesis testing	Users with a longer app session time in media apps are less active in app advertisement.
	[111]	1,527	165	3 months	Monitoring app	App session, screen state	k-means	Identify three distinct re-visitation patterns: checkers, waiters, and responsives.
	[112]	22,626	4,125	532 days	Monitoring app	App session, phone call	Descriptive statistics	App sessions interrupted by calls will lead to more overhead, delaying completion of a task by up to 4 times.

the day of one week of usage behavior. Jiang *et al.* [124] added time as a new feature and used the most similar app usage case from history to generate prediction results. Wang *et al.* [125] used time as a condition in their probabilistic graphical models. Zhao *et al.* [126] developed a deep learning-based model for predicting app usage based on time context. As illustrated in Fig. 9, they concatenated the time feature with historical used apps and fed the time feature into the last layer of the neural network to introduce time context. It is a state-of-the-art deep learning-based app usage prediction model, with a Recall@5 of 84.47%.

Location is also an important contextual factor to consider when predicting app usage. To achieve location-aware app prediction, Parate *et al.* [127] split app usage sequences into a variable-length Markov chain according to location features, while Wang *et al.* [128] used the location factor as a condition in Bayesian networks. Furthermore, deep learning techniques are used in location-aware prediction. Xia *et al.* [129], for example, developed a recurrent neural network-based model to predict the next used app and visited location at the same time. Chen *et al.* [130] proposed CAP, a graph embedding-based model for learning node embeddings from app-location,

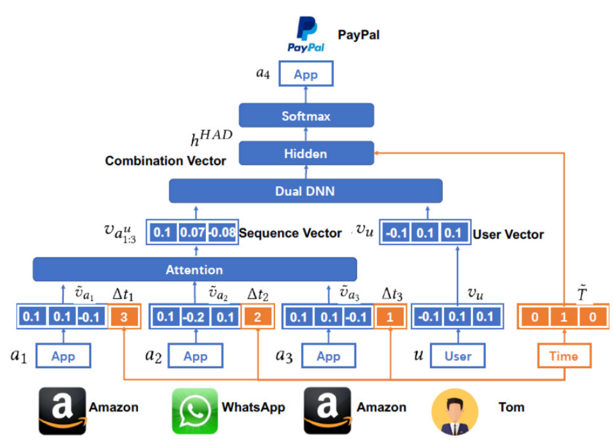


Fig. 9. A deep learning model for time context-aware app usage prediction [126].

app-time, and app-category subgraphs. Yu *et al.* [131] used a graph neural network to learn the node embeddings on an app-location-time graph. They then used the learned node embeddings to predict app usage. Zhou *et al.* [132] further proposed a heterogeneous graph-based model that learns embeddings for apps, locations, and time, to achieve an end-to-end prediction. The above frameworks based on graph embedding open the door to modeling correlations between app, location, and time through spatiotemporal app usage graphs.

Movement status, network mode, and battery level are also used for app usage prediction. In [133] and [134], movement status, including mobility entropy and travel patterns, is extracted from users' trajectories. The above information and app usage sequences are jointly utilized for training a classifier for app usage prediction. Do and Gatica-Perez [135] considered network mode and battery level when building the probabilistic model for app usage prediction. Xu *et al.* [136] enriched query vector using screen state and network mode to recall the most likely used apps as prediction results.

2) App Recommendation: App recommendation aims to infer users' preferences towards unobserved new apps and then recommend favorable ones [137], [138]. Regarding whether contextual information is used, app recommendations can be classified into non-contextualized recommendations and context-aware recommendations.

Non-contextualized app recommendations are made solely based on the user's previous app usage behavior. Yan and Chen [139] presented AppJoy, a framework for tracking how users install and use apps. The AppJoy system uses an item-based collaborative filtering (CF) model to predict usage scores and make recommendations. To deal with sparse datasets, Shi and Ali [140] proposed a PCA-based model for app recommendation, which uses PCA to extract features from the user-item matrix before performing item-based CF. Instead of collaborative filtering, Liang *et al.* [141] and Ouyang *et al.* [142] used graph neural network models to depict the similarity of apps and make recommendations. Specifically, Liang *et al.* [141] considered only app installation behavior, while Ouyang *et al.* [142] took both installation and search behaviors into account.

Many studies extended to a context-aware case in which apps are recommended based on contextual information such as time, location, movement status, and app metadata. One of the most widely used methods for utilizing contextual information is context-aware collaborative filtering (CAF), which extends conventional collaborative filtering by multiplexing user identities through context features. Böhmer *et al.* [144] compared different recommendation systems and demonstrated that context-aware collaborative filtering outperformed non-contextualized collaborative filtering.

Tu *et al.* [145] took time as a contextual factor and considered changes in user preferences over time. They proposed a personal interest evolution network to model user dynamic preference and make recommendations. Xu *et al.* [143] leveraged usage context, prior and posterior used apps, to infer users' preference to recommend apps. The proposed model, as shown in Fig. 10, contains two paths to combine user features and app usage context. They started by merging the user and app embeddings. The merged embedding was then used as inputs to predict context apps (on the right path) and to predict user preference (on the left path). Pan *et al.* [14] used social contextual features for app recommendations by creating four different social networks for users, including a call log network, a Bluetooth proximity network, a friendship network, and an affiliation network. They then proposed a graph model to predict the probability of installing a new app.

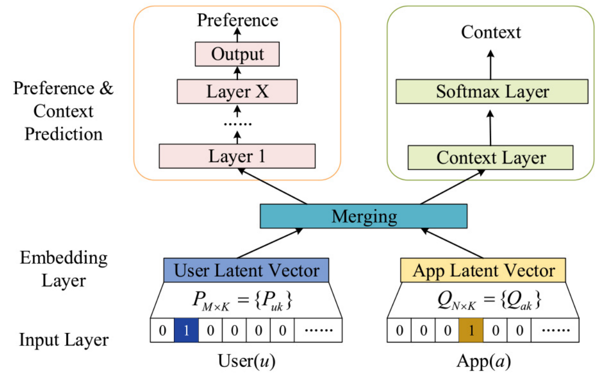


Fig. 10. A deep learning model for usage context-aware app recommendations [143].

App metadata like category, description, and permission are frequently used in context-aware app recommendations. Liu *et al.* [146] leveraged app category information for app recommendation based on a kernelized non-negative matrix factorization (KNMF) model. Liang *et al.* [147] proposed characterizing app features based on three factors: categories, permissions, and descriptions. They then used a tensor-based framework to integrate the multi-view features of apps to achieve app recommendations. Lin *et al.* [148] and Cao *et al.* [149] looked into version-sensitive app recommendation by considering the version update of apps. Rocha *et al.* [151] and Gao *et al.* [150] proposed using app permissions, to improve security degree of recommendations.

3) Discussion: We reviewed existing studies on app usage prediction and recommendation in this subsection. Table VII summarizes prominent literature in terms of dataset information, methods, and system performance.

In the field of app usage prediction, sensor context-based cases are the mainstream because sensor context features, such as time and location, can provide a significant performance boost. Most of the datasets are collected from monitoring apps and network operators because these two collection methods can support sensor context data. There is a clear progression from basic statistical learning models (e.g., nearest neighbor, random forest, Markov model) to probabilistic models (e.g., Bayesian networks, probabilistic graphical models), and finally to deep learning models (e.g., deep neural networks, graph embedding, graph neural networks). This is because deep learning models can explore contextual information in-depth and extract coupling relationships between users, apps, time, and locations.

In the field of app recommendations, contextual information, such as location, time, and app metadata, is frequently used. To support supplementary data, the datasets used are primarily collected from monitoring apps and app stores. Regarding the methods applied, collaborative filtering (CF) is the most basic algorithm framework. Tensor factorization is a popular and effective way to address

TABLE VII
LITERATURE REVIEW IN APP USAGE PREDICTION AND RECOMMENDATION

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	Context	Methods	Results and findings
App usage prediction	[118]	177	111	3 months	Monitoring app	App usage	Prior used app	MFU, MRU	Accuracy: 82.5%∼86.5%
	[119]	9,583	17,062	1 year	Monitoring app	App usage sequence	Prior used app	Markov model	Recall@5: 67.01%
	[120]	59	80	-	Monitoring app	App usage sequence	Prior used app	Bayesian network	Hitrate@5: 85%
	[121]	7,000	480	7 months	Monitoring app	App session	Prior used app	Bayesian network	Accuracy: 85.7%
	[124]	1,003	25,376	3 months	Network operator	App usage	Prior used app, time	Nearest neighbor	F1-score: 0.605
	[126]	9,373	2,000	3 months	Network operator	App usage sequence	Prior used app, time	Deep neural network	Recall@5: 84.47%
	[127]	-	7,630	-	Monitoring app	App usage sequence	Prior used app, time, location	Variable-length Markov	Accuracy: 81.125%
	[128]	3,500	1.7 million	1 week	Network operator	App usage	Prior used app, time, location	Bayesian mixture model	Residual: 0.39
	[129]	2,000	10,000	1 week	Network operator	App usage sequence	Prior used app, time, location	Deep neural network	Precision: 54.62%
	[130]	2,000	1,788	1 week	Network operator	App usage	Prior used app, time, location, app category	Graph embedding	Accuracy@5: 82%
	[131]	2,000	11,170	1 week	Network operator	App usage	Prior used app, time, location, app category	Graph neural network	Accuracy@10: 61%
	[132]	-	24	1 year	Monitoring app	App usage	Prior used app, time, location, app category	Graph neural network	Accuracy@4: 70%
	[133]	-	1,000	14 days	Network operator	App usage	Prior used app, time, location, mobility and travel pattern	Random forest	Accuracy: 90.3%
	[135]	9	71	17 months	Monitoring app	App usage	Prior used app, time, location, battery level, network mode	Probabilistic model	Accuracy: 54.3%
	[136]	16,000	4,606	millions of hours	Monitoring app	App usage	Prior used app, time, location, network mode, app category, screen state	Nearest neighbor	Accuracy: 62%
App recommendation	[139]	16,950	4,606	10 million hours	Monitoring app	App usage, installation	-	CF	Viewed Ratio: 24.2%, RMSE: 0.9749.
	[140]	55,020	101,106	1 month	Monitoring app	App usage	-	CF	Recall@50: 0.27.
	[141]	6,756	8,875	-	App store	App installation	-	Graph neural network	MAE: 0.8567, RMSE: 1.2507.
	[142]	18,229	1,011,567	31 days	App store	App installation, search	-	Graph neural network	Recall@10: 0.1024.
	[144]	-	45	3 months	Monitoring app	App usage frequency	Time, location	CF, CAF	Conversion rate: 7%-24%
	[145]	25,372	40	9 months	Monitoring app	App usage, installation	Time	Deep neural network	AUC: 0.6248.
	[143]	7,783	56,672	2 months	-	App session	Prior and posterior used app	Deep neural network	Recall@6: 0.4.
	[14]	821	55	5 months	Monitoring app, survey	App installation	Bluetooth, friendship, call log, affiliation	Graph model	RMSE: 0.25, MP@5: 0.31, F1 Score: 0.43
	[146]	6,527	11,166	2 months	Monitoring app, app store	App usage	App category	KNMF	MAP@20: 0.63, NDGG@20: 0.734
	[147]	6,200	16,300	-	App store	App installation	App rating, category, permission, description	Tensor analysis	MAE: 0.7765, RMSE: 1.1419
	[148]	6,524	9,797	-	App store	App installation	App rating, category, version, description	Topic model	Recall@100: 0.73
	[149]	600	-	20 months	App store	App installation	App rating, version, description, category	Matrix factorization, evolution progress modeling	MAE: 0.8963, RMSE: 1.1953
	[150]	47,264	10	6 months	Monitoring app, app store	App installation	App reviews, permissions, descriptions	Clustering, rule-based model	Recall@5: 0.73

context-aware app recommendations by taking location and time as context information and adding additional dimensions to represent them. Deep learning models, such as deep neural networks and graph neural networks, have become increasingly popular in recent years for mobile app recommendations.

TABLE VIII
USER ATTRIBUTES THAT CAN BE PROFILED FROM MOBILE APP USAGE DATA

Demographic characteristics	Personality traits	Psychological status	Personal interests	Life status
Gender, age, income level, occupation, nationality, culture, etc.	Extraversion, agreeableness, conscientiousness, neuroticism, openness to experience, etc.	Stress level, well-being, emotion, etc.	Sports, music, cooking, reading, etc.	Life events, workout activities, life stages, social status, etc.

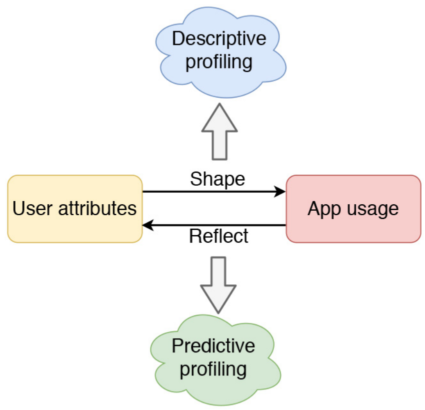


Fig. 11. The relationship between user attributes and app usage.

IV. USER DOMAIN RESEARCH

Personal characteristics, such as interests, age, gender, and occupation, have a significant impact on mobile app usage. Many studies analyzed mobile app usage data from the user's perspective to reveal a link between user characteristics and app usage behavior. This section summarizes user domain research on two primary topics: user profiling and user identification.

A. User Profiling

User profiling aims to infer personal attributes from user-generated data, which is essential for personalized services such as personalized search, recommendations, and advertisements [152]–[154]. Demographic characteristics, personality traits, psychological status, personal interests, and life status are five categories of user features profiled from mobile app usage data, as shown in Table VIII. Demographic characteristics refer to a person's inherent properties, such as gender, age, income level, nationality, and occupation. Personality traits reflect people's characteristic patterns of thoughts, feelings, and social adjustments. The Big-Five model is the most widely used personality trait measurement system [41], [155], which includes five broad traits: extraversion, agreeableness, conscientiousness, neuroticism, and openness. Psychological status is a mental state that influences people's behavior and decisions, such as stress, well-being, and emotion. Personal interests are the objects or events that users prefer to concentrate on. Individuals' life status denotes their preferred way of life, which includes things like life events and life stages.

The relationship between user attributes and app usage is depicted in Fig. 11. On the one hand, user attributes shape how people use mobile apps. On the other hand, app usage behavior also reflects individual attributes. Based on different directions of the relationship, existing studies can be separated into two groups: descriptive profiling and predictive profiling. In descriptive profiling, researchers focus on how user attributes affect their app usage behavior. Alternatively, researchers aim to use app usage data to predict users' profile labels in predictive profiling.

1) Descriptive Profiling: References [19], [23], [156]–[160] investigated how demographic characteristics affect app usage behavior. Andone *et al.* [156] looked into the effects of age and gender. They analyzed the app usage patterns of 30 thousand users for 30 days. According to the study, females spent more time on communication and social apps, while males spent more time playing games. Teenagers aged 12 to 17 spend the most time on communication, social media, and gaming apps, averaging over 40 minutes per day. Nonetheless, participants over the age of 30 spend less than ten minutes using these apps. Gordon *et al.* [23] pointed out that older adults using fewer apps is caused by cognitive decline. As a result, they recommended that developers consider cognitive function to better support older adults. Zhao *et al.* [157] looked at how app usage differed by income level. They discovered that users with a higher income use apps in the categories of shopping, finance, travel, and business more frequently. Users with a low-income level use game and video apps more frequently. Tu *et al.* [158] expanded the analysis to include education level and discovered that Ph.D. users use more communication apps but fewer sports apps than other education levels. References [19], [159], [160] investigated the differences in mobile app usage by country. Lim *et al.* [19] gathered data on app adoption in an online survey of 10,208 participants from over 15 countries. They discovered significant differences in different countries. Users in the United States, for example, are more likely to download medical apps. Peltonen *et al.* [159] suggested that cross-cultural differences could explain differences in mobile app usage behavior across countries. They discovered three main clusters of countries with distinct usage patterns by using app category usage as features and applying the hierarchical clustering algorithm. The countries within the same cluster, as shown in Fig. 12, have similar cultural backgrounds, with three main clusters: European, English-speaking, and mixed. Guzman *et al.* [160] also demonstrated that cross-cultural differences in mobile app reviews still exist.

The impact of personality traits on mobile app usage was investigated in [161], [162]. They used the Big-Five framework, which consists of five bipolar factors: extraversion, agreeableness, neuroticism, conscientiousness, and openness, to represent users' personality traits. The five factors in Table IX are explained using adjective examples to describe each trait. The impact of personality traits on app category usage was highlighted by Huseynov [161]. Extraverts use

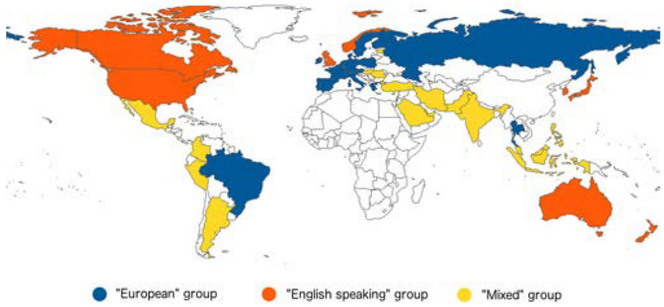


Fig. 12. Countries colored by cluster labels in terms of app category usage [159].

TABLE IX
THE BIG-FIVE PERSONALITY FRAMEWORK AND EXAMPLES OF ADJECTIVES DESCRIBING EACH TRAIT

Personality traits	**Examples of adjectives**
Extraversion	Active, Energetic, Enthusiastic, Outgoing.
Agreeableness	Appreciative, Forgiving, Generous, Kind.
Conscientiousness	Efficient, Organized, Reliable, Responsible.
Neuroticism	Anxious, Tense, Unstable, Worrying.
Openness	Artistic, Imaginative, Insightful, Wide Interests.

more photography and video editing apps. Agreeableness individuals are negatively associated with the use of health and lifestyle mobile apps, while e-commerce-related apps are usually avoided by conscientious individuals. Beierle *et al.* [162] looked at the link between personality traits and app session characteristics. They discovered that extraversion and neuroticism are linked to more app sessions, whereas conscientiousness is linked to a shorter average session time.

Some researchers looked into the psychological status of users [163]–[165]. Gao *et al.* [163] analyzed the links between app use and social anxiety and loneliness. Using the Wilcoxon-Mann-Whitney test to examine differences in user behaviors, they discovered that people with social anxiety or loneliness receive fewer incoming calls and use health apps more frequently. People with higher levels of social anxiety receive fewer messages and utilize camera apps less often. Katevas *et al.* [164] focused on users' well-being. Based on data from 340 participants over the course of four weeks, they offered evidence that intense app usage alone does not indicate negative well-being. However, nightly usage behavior shows a strong link to a reduced sense of well-being. According to [165], user emotions and app usage behaviors have bidirectional causality. App usage, on the one hand, influences emotions. Entertainment apps, like YouTube and music apps, evoke positive emotions. On the other hand, emotions drive app usage. When users are experiencing happy emotions, they are more active in using social apps.

Personal interests are the subject of [33], [166], [167]. Zhao *et al.* [33], [166] conducted a large-scale study in which they assessed 105,762 users' app usage traces for one month. They identified 382 unique groups of users based on their temporal patterns of app usage using K-means clustering. Each user group is assigned a meaningful label, such as night communicators, evening learners, or financial users. Lee *et al.* [167], on the other hand, undertook a small-scale investigation. They used Seq2Seq to extract app usage sequence embeddings from 180 volunteers' app usage sequences. They found eleven user groups using the K-means clustering method on embeddings: conversationalists, utilitarians, social stars, photographers, music fans, news and magazine readers, video streamers, gaming buffs, power users, and beginners. Both large-scale and small-scale analyses in the previous studies highlighted the diversity of personal interests in app usage.

Some researchers have indicated links between app usage and life status [168], [169]. Frey *et al.* [168] are the first to recognize that a user's current life stage has a significant impact on their app adoptions. For example, when a person's life stage changes from without children to with children, the use of travel and entertainment apps drops dramatically. Chen *et al.* [169] used fitness app data to examine people's workout activities and discover their exercise styles. They found that people in the central business district (CBD) walk or jog more than those in rural areas and that people with longer mobility spans exercise less.

2) Predictive Profiling: Given users' app usage data, predictive profiling aims to predict their attribute labels. References [170]–[174], [176], [177], [184] estimated users' demographic characteristics, such as gender, age, and income, based on their app usage behaviors. Seneviratne *et al.* [170] gathered app adoption data and gender labels from 200 users and found that using only app installation lists could accurately estimate users' gender by 75%. They then expanded their research to include other characteristics such as religion, spoken languages, and nationality [171], where precision reached over 90%. Malmi and Weber [172] further used a larger dataset of 3,760 users to confirm Seneviratne's findings and considered new demographics, such as wealth and race. According to their study, gender is the most predictable attribute, with an accuracy of 82.3%, while wealth is the most difficult to forecast, with an accuracy of 60.3%. Instead of using the list of users' apps directly as features, Zhao *et al.* [173] constructed a Boolean user-app matrix to represent user app adoptions and used Boolean matrix factorization (BMF) to extract hidden representations of users. Hidden representations had a higher gender prediction accuracy of 77.2% than app lists (75.2%). Zhao *et al.* [174] also looked into the sequential characteristics of app usage and learned user embeddings from app usage sequences. With the help of users' embeddings, they improved gender prediction accuracy to 82.49%. Bian *et al.* [175] considered user-app interactions as well as user-item interactions on apps to learn user embeddings and infer users' occupations with a 70% precision. To improve prediction accuracy, some studies use additional information such as app descriptions and sensory data. Zhao *et al.* [176] used the LDA model to extract topic features from app descriptions as side information. They then used both topic features and app lists to predict users' gender labels. Yu *et al.* [177] used sensory data such as screen and battery status to predict gender with a precision of 91.70 percent and an RSME of 4.3696 in age estimation.

TABLE X
LITERATURE REVIEW IN DESCRIPTIVE PROFILING

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	User attributes	Methods	Results and findings
Descriptive profiling	[156]	-	30,677	28 days	Monitoring app	App usage, app category	Gender, age	Descriptive statistics	Females spend more time on communication and social apps. Males spend more time on games. Teenagers have the highest usage time on communication, social media, and games.
	[23]	-	84	3 months	Monitoring app	App launch, call, unlock, message	Age	Descriptive statistics	Older adults use fewer apps and take longer to complete tasks on apps.
	[157]	77,685	106,762	30 days	Network operator	App usage, time	Gender, age, income level	Descriptive statistics	Higher-income users use more shopping, finance, travel, and business apps. Low-income users use game and video apps more.
	[158]	100,000	1,608	7 years	Monitoring app, survey	App usage, time	Gender, age, education level, salary	Descriptive statistics	Ph.D. users use more communication apps but fewer sports apps than other education levels.
	[19]	-	4,824	2 months	Survey	App rating, adoption, price	Nationality	Descriptive statistics	Users from the USA prefer to download medical apps. Users from Japan and Australia are less likely to rate apps.
	[159]	54,776	25,323	1 year	Monitoring app, survey	App category usage	Nationality, culture	Hierarchical clustering	Cross-cultural differences cause different app usage behavior across countries. Discovered three distinct groups, European group, English speaking group, and mixed group.
	[160]	7	2,560	2 months	App store	App review	Nationality, culture	Tukey-Kramer test	There are cross-cultural differences in app reviews.
	[161]	52	101	1 week	Survey, monitoring app	App usage	Personality trait	Regression analysis	Extraverts use more photography and video editing apps, while agreeableness individuals are negatively associated with the use of health and lifestyle apps.
	[162]	-	526	48 days	Monitoring app, survey	App usage sessions	Personality trait, age, gender	Random-intercept, random-slope multilevel regression analysis	Extraversion and neuroticism are linked to more app sessions. Conscientiousness is linked to a shorter average session time.
	[163]	-	127	30 days	Monitoring app, survey	App usage, SMS, call, screen state	Social anxiety, loneliness	Wilcoxon-Mann-Whitney test	Individuals with social anxiety or loneliness receive less incoming calls and use healthy apps more frequently.
	[164]	-	340	4 weeks	Monitoring app, survey	App usage, call, screen state, data activity	Well-being	Spectral clustering	Intense usage does not represent negative well-being. Nightly usage behavior has a high correlation with lower well-being.
	[165]	-	30	2 weeks	Monitoring app, survey	App usage	Emotions	Convergent cross mapping	User emotions and app usage behaviors have bidirectional causality.
	[166]	27,779	106,762	30 days	Network operator	App usage, time	Personal interest	K-means	Determined 382 distinct groups of users based on their temporal patterns of app usage.
	[167]	-	180	-	Monitoring app	App usage sequence	Personal interest	Seq2Seq, K-means	Identified ten user groups, i.e., conversationalists, utilitarians, social stars, photographers, music lovers, news readers, video streamers, gaming buffs, power users, and beginners.
	[168]	-	1,435	4 months	Monitoring app, survey	App adoption	Life stage	Descriptive statistics	Users' current life stage strongly influences their app adoption patterns.
	[169]	-	4,000	1 week	Network operator	App usage, location	Work out activity	Descriptive statistics	People in the CBD walk or jog more than people in rural areas. People with larger spans tend to work out less.

In terms of personality trait prediction, Xu *et al.* [178] developed a prototype, Personality Test, which automatically evaluates users' personality traits based on apps they installed. Personality Test tracks the use of seven app categories: social apps, games, music and video apps, shopping apps, photography apps, finance apps, and personalization apps. With a random forest classifier, the Personality Test can obtain an average prediction precision of 42.72%. Peltonen *et al.* [179] went on to collect data on app usage across all categories. They showed that category-level aggregated app usage can predict Big Five personality traits with a prediction fit of 86%-96%. In addition to app usage data, Chittaranjan *et al.* [180] considered call logs, SMS logs, and Bluetooth scan logs. They achieved an average F1-score of 0.57 using an SVM classifier with a radial basis function (RBF) kernel. Kambham *et al.* [181] and Gao *et al.* [182] characterized personality trait prediction as a regression problem by placing participants on a Big Five Personality Inventory continuum. They used app usage and smartphone sensory data and achieved RMSEs ranging from 12.7% to 22.2% for different traits.

Ochiai *et al.* [183] looked at app usage patterns to see if they might forecast users' psychological states, such as stress levels. They considered the order in which the apps were used and created an app usage graph for each user to depict the contextual relationship between apps. Based on the Graph Isomorphism Network (GIN), they classified app usage graphs to predict users' stress levels and achieved an accuracy of 54.5%.

3) Discussion: We summarized recent studies on user profiling based on app usage behaviors in this section. In Tables X and XI, we review prominent literature on descriptive profiling and predictive profiling. The tables include research topics,

TABLE XI
LITERATURE REVIEW IN PREDICTIVE PROFILING

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	User attributes	Methods	Results and findings
Predictive profiling	[170]	4,167	200	-	Monitoring app	App adoption	Gender	Naive Bayes, SVM	Accuracy: 0.75.
	[171]	4,167	200	-	Monitoring app	App adoption	Religion, nationality, spoken language	SVM	Precision: 0.9.
	[172]	8,840	3,760	1 month	Monitoring app	App adoption	Gender, age, race, income	Naive bayes, SVM, LR, random forests	Accuracy: 0.603∼0.823.
	[173]	9,326	15,000	-	Network operator, survey	App adoption	Gender	BMF, SVM, LR, GBDT, DNN	Accuracy: 0.772.
	[174]	2,000	10,000	-	Network operator, survey	App usage sequence	Gender	Embedding, SVM, LR, GBDT, DNN	Accuracy: 0.8249
	[175]	560	19,863	-	Monitoring app	App interactions	Occupation	Embedding, Transformer	Precision: 0.695
	[176]	1,000	15,000	-	Network operator, survey	App adoption, app description	Gender	LDA, SVM, LR, GBDT, DNN	Accuracy: 0.7662.
	[177]	-	84	2 weeks	Monitoring app	App adoption, screen status, battery status	Gender, age	Random forest	Gender prediction: precision: 0.9170, Age estimation: RSME: 4.3696.
	[178]	155,187	2,043	6 days	Monitoring app, survey	App adoption	Personality trait	Random forest	Precision: 0.4272.
	[179]	7,852	843	6 months	Monitoring app, survey	App adoption	Personality trait	Random forest, SVR, DNN	Prediction fit: 86%∼96%.
	[180]	6	117	17 months	Monitoring app, survey	App usage, Bluetooth scan, call logs, SMS	Personality trait	SVM with a RBF kernel	F1-score: 0.57.
	[181]	-	375	1 month	Monitoring app, survey	App usage, WiFi status, screen status, battery status	Personality trait	MLP	RMSE: 12.7%∼22.2%.
	[182]	40	183	-	Monitoring app, survey	App usage, call and SMS logs	Personality trait	DNN, attention	RMSE: 30.5%∼64.7%.
	[183]	319	28	457 days	Monitoring app	App usage sequence	Stress level	GIN	Accuracy: 0.545.

dataset information, methods, and findings and results. We can see that the dataset scale ranges from 28 users to over 100,000 users. App usage data is generally collected through monitoring apps and network operators, while user attributes are collected through online surveys. Descriptive statistics are frequently used in descriptive profiling to reveal the relationship between app usage and user characteristics. The statistical significance of discovered relationship is demonstrated using correlation test metrics such as Pearson correlation, Wilcoxon-Mann-Whitney test, and Tukey-Kramer test. In predictive profiling, SVM, LR, MLP, and random forest are commonly used classification methods. A few studies recently looked into advanced machine learning techniques, such as transformer, attention, and embedding, indicating a hot direction for predictive user profiling.

B. User Identification

Although data collection agencies have anonymized user IDs to protect users' privacy, mining or sharing app usage datasets still poses a significant privacy risk. Many studies have demonstrated that users can be identified or re-identified from anonymized datasets based on their app usage behaviors. In 2010, Falaki *et al.* [185] were the first ones to demonstrate the diversity of smartphone and app usage among individuals, paving the way for user identification research. Welke *et al.* [186] studied app adoption of 46,726 participants and discovered that using 500 of the world's most popular apps could distinguish unique app-signatures for 99.67% of users. Tu *et al.* [187] analyzed a larger dataset of 1.37 million Chinese users and found that using only four apps can uniquely identify 88% of users. Sekara *et al.* [188] further compared the size of app-signatures of users in different countries. They discovered that the identification rate is heavily influenced by the country of users. When the size of app-signatures is set to 5, the average identification rate varies dramatically between countries, ranging from 41.2% in Finland to 66.5% in the United States. These observations present severe privacy concerns as most individuals are identifiable and trackable, even in large-scale anonymized datasets.

Some researchers use such identifiable characteristics for active authentication to secure users' smartphones. The basic idea is to use a user's app usage patterns to verify his or her identity [189]. For each user, Ashibani and Mahmoud [190] chose the eight most-used apps and used the timestamp of app usage as a feature to continuously identify users. Their authentication model verifies users with an average F1-score of 96.5% after conducting experiments on ten users. They then improved the F1-score to 98% by including traffic patterns while accessing apps as features in their model [191]. For more reliable authentication, Bassu *et al.* [192] proposed combining contextual information, such as location and time,

TABLE XII
LITERATURE REVIEW IN USER IDENTIFICATION

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	Methods	Results and findings
User identification	[185]	-	255	7∼28 weeks	Monitoring app	App adoption, screen state,network traffic, battery state	Descriptive statistics	The diversity of smartphone and app usage among individuals.
	[186]	100,000	46,726	216 days	Monitoring app	App adoption	Descriptive statistics	Using 500 globally most popular apps can distinguish unique signatures for 99.67% of users.
	[187]	2,000	1.37 million	7 days	Network operator	App adoption	Descriptive statistics	88% of users can be uniquely identified by four apps.
	[188]	1,129,110	3,556,083	12 months	Monitoring app	App adoption	Descriptive statistics	The identification rate is heavily influenced by the country of users.
	[189]	1,992	34	365 days	Monitoring app	App usage, time	RF, KNN, SVM	Equal error rate: 3%.
	[190]	8	10	152 days	Monitoring app	App usage, time	RF, DT, LR, KNN, SVM, MLP	Average F1-score: 96.5%.
	[191]	530	10	20 weeks	Monitoring app	App usage, time, traffic	RF	F1-score: 98%.
	[193]	20	200	30 days	Monitoring app	App usage, text, website, location	SVM	Equal error rate: 5%.
	[194]	-	263	-	Monitoring app	App usage, time	HMM	The appearance of unknown apps would cause an increase in the number of false negatives.

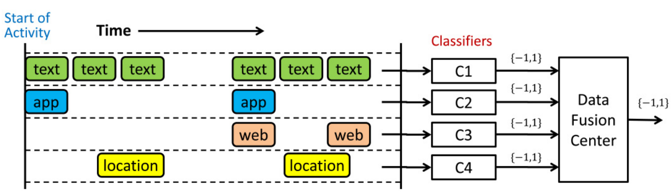


Fig. 13. User authentication by using text entered via soft keyboards, apps used, websites visited, locations visited [193].

with users' uploading, downloading, and updating behaviors. Fridman *et al.* [193] also looked at various types of biometric behaviors, including text entered via soft keyboards, apps used, websites visited, and locations visited. Each distinct biometric behavior is fed into a classifier, as shown in Fig. 13. A data fusion center then fuses the local binary decisions from each classifier to produce a single global binary decision. The authentication system can achieve an equal error rate (ERR) of 5% after conducting experiments on 200 users. Notable, unexpected events, such as newly installed apps, can have a significant impact on the performance of authentication systems. Mahbub *et al.* [194] investigated the impact of unknown apps on the verification task. In their case, the unknown apps refer to the test set's apps but not in the train set. They discovered that the appearance of unknown apps increased the number of false negatives.

User identification can be regarded as individual-level user profiling, i.e., predicting a user's specific identity label instead of an attribute label. Table XII summarizes prominent literature related to user identification. We can observe that users' identifiable characteristics based on their app usage behaviors are demonstrated in both small-scale [185] and large-scale [186]–[188] datasets. Several studies have also looked into using such identifiable characteristics for user authentication. By utilizing simple classifiers, like RF, LR, SVM, they achieved good performance with an F1-score of 98% or an equal error rate of 3%. However, existing frameworks continue to rely on a centralized learning structure, in which all users' data is uploaded to a server, putting user privacy at risk during data uploading and computing processes.

V. SMARTPHONE DOMAIN RESEARCH

Many researchers analyzed app usage data to improve system performance and solve problems in the smartphone domain. This section summarizes smartphone domain research in terms of two perspectives: app energy drain and app traffic patterns.

A. App Energy Drain

The smartphone is a limited-energy device whose battery life is a critical performance and user experience metric [195]. Both the research community and the industry are working on ways to extend the life of batteries. Increasing battery capacity is one simple solution. However, due to hardware limitations, we need to improve the energy efficiency of smartphone apps parallelly. Some researchers have studied the energy drain of app usage in order to improve energy efficiency.

Song *et al.* [196] empirically measured energy consumption for ten mobile apps. They discovered that average energy consumption in the studied apps varies significantly, ranging from 0.25 Joule/second to 1.25 Joule/second. Their research emphasizes the importance of optimizing app design to reduce energy consumption. To gain a thorough understanding of app energy consumption patterns, Oliner *et al.* [8] developed Carat, an app for both iOS and Android platforms, which detects and diagnoses energy anomalies in apps. On stock devices, Carat runs as a user-level app that collects information such as battery level, running app names, memory status, and device model. Oliner *et al.* discovered 10,110 hogs and 233,258 buggy apps by analyzing energy drain data from over 500,000 devices. An app is a hog if it consumes a lot of energy to run. Pandora Radio, Skype, Live Wallpapers are typical hogs. Apps that consume significantly more energy on some clients than others are referred to as bug apps. That may be caused by the different usage habits of users. For example, some users prefer to use Kindle with networks to synchronize notes and bookmarks, while others prefer not to do so to save money on mobile data. The energy consumption will differ significantly depending on whether the network connection is turned on or off. In this way, Kindle, Facebook, Youtube are typical bug apps. Based on these findings, they then created actionable diagnosis trees to recommend practical actions such as turning on/off WiFi/GPS, avoiding using certain apps under certain conditions, and upgrading/keeping app or operating system versions.

Chen *et al.* [197] created eStar, a free Android app that collects energy drain data, and used it to gather an energy trace dataset from 1,520 Galaxy S3 and S4 devices covering 800 different apps in the wild. They discovered that background apps and services consume 16.1% of total energy during the screen-off period. This finding implies that energy consumption can be reduced by improving the app scheduling algorithm and killing background apps when the screen is off. However, if we disable all apps without considering the differences in background activities, the user experience may be adversely affected. In [198], Chen *et al.* looked into this issue in depth. By analyzing the same dataset, they discovered 76 no-sleep apps for which the CPU was never suspended. They then designed a metric, Background-Foreground Correlation (BFC), to assess the utility of background app activities. A simple yet effective screen-off energy optimizer based on BFC was developed to learn and suppress useless background activities of apps automatically. Li *et al.* [199] pointed out that the unnecessary workload of apps is a major root cause of energy issues. Many apps, they discovered, perform computations that do not provide users with discernible benefits, resulting in unnecessary workload and energy consumption.

The above studies look at the energy drain pattern of a single app. However, because most smartphones have multi-core CPUs, parallelism execution results in different app combinations having different power consumption patterns. The energy consumption of one app's threads will affect the energy consumption of other apps' threads. By studying different apps as a group, Rex *et al.* [200] attempted to reschedule app threads and correlate their influence on energy consumption. Because the number of total possible combinations of apps is enormous, they explored the frequent app usage sets based on users' usage habits and then optimize scheduling sequences for frequent app usage sets.

To help developers create 'greener' apps that use less energy, some researchers looked at energy drain patterns of specific app events or functions. Li *et al.* [210] developed a prototype tool called vLens to collect energy data at the nanosecond and millisecond levels. As a result, vLens can calculate the energy consumption of smartphone apps at the source line level and track high-energy events like thread switching and garbage collection. In [201], they then used vLens to collect the energy drain patterns of 405 apps. They discovered that, on average, apps waste 61% of their energy in idle states, and the network is the most energy-consuming component, confirming the findings in [197]. Most importantly, they discovered that called system APIs account for the majority of an app's energy consumption, implying that app developers should exercise caution when using them.

B. App Traffic Patterns

Smartphones, which are supported by a variety of network-based apps, allow users to stay connected to the ubiquitous Internet [211]. According to a Cisco report [212], smartphone traffic accounted for 73% of all traffic in 2018 and will rise to 89% by 2023. Because of the high volume and growing importance of mobile traffic, researchers and network operators are interested in understanding how mobile apps use network resources.

Xu *et al.* [32] conducted anonymous network measurements from a tier-1 cellular carrier in the United States, covering over 600 thousand users, and used HTTP signatures to identify traffic from various apps. In terms of traffic sources, they discovered that mobile apps could be divided into two groups: local apps and national apps. The majority of the traffic of local apps comes from a single region, whereas the traffic of national apps comes from diverse regions. This finding suggests that content optimization in access networks could be significant, with local app content being placed on servers closer to end-users. Jiang *et al.* [202] looked into the relationship between app popularity and network performance. They discovered that over 60% of popular apps optimize network delay under 350ms, highlighting the importance of app service network performance. Falaki *et al.* [213] and Li *et al.* [203] looked at a small and large scale network dataset, respectively, and found that smartphone operating systems have an impact on app traffic consumption. When compared to Android and Windows Phone, iOS consumes more traffic. Jin *et al.* [204] also discovered that different data collection purposes resulted in different app network traffic usage. As a result, they created MobiPupose, a system that could track app network requests and classify data collection purposes based on app traffic patterns. Walelgne *et al.* [205] and Okic *et al.* [206] showed that different app categories have different traffic patterns. In terms of traffic volume, entertainment and social media apps are the most traffic-intensive, while education and weather apps are the least traffic-intensive [205]. Moreover, in terms of temporal patterns, music and shopping apps see the most traffic in the morning, while e-mail and game apps see the most traffic during lunchtime [206]. Some studies leveraged app traffic patterns to improve network performance. Ouyang and Yan [207] developed AppWiR, a crowdsourcing-based system that gathers app usage data and derives relationships between app usage, network traffic, and network resources. AppWiR can predict network resources occupied by apps with a mean absolute percentage error of 12.54% sim13.39$. Zeng *et al.* [10] analyzed the temporal traffic consumption patterns of various app categories. They then used a linear regression model to forecast the amount of traffic consumed by the most popular app categories over a given period of base stations and devised an edge caching strategy to cache the content for the most popular app services. Aceto *et al.* [208] defined network

TABLE XIII
LITERATURE REVIEW IN SMARTPHONE DOMAIN RESEARCH

Topic	Ref.	# Apps	# Users	Duration	Collection method	Data	Methods	Results and findings
App energy drain	[196]	10	-	-	Monitoring app	App usage, battery state	Descriptive statistics	Average energy consumption in apps varies significantly, ranging from 0.25 Joule/second to 1.25 Joule/second.
	[8]	243,368	500,000	-	Monitoring app	App usage, battery state, device model, network connection, time	Collaborative inference	10,110 hogs and 233,258 buggy apps.
	[197]	800	1,520	40 days	Monitoring app	App usage, screen status, time, battery consumption of CPU, GPU and WiFi	Descriptive statistics	Background apps and services during screen-off period cost 16.1% of the total energy.
	[198]	800	2,000	40 days	Monitoring app	App usage, screen status, time, battery consumption of CPU, GPU and WiFi	Linear regression, descriptive statistics	76 no-sleep apps. A simple yet effective screen-off energy optimizer, HUSH.
	[199]	89	-	11 years	User report	App name, energy issues	Descriptive statistics	Unnecessary workload of apps is a major root cause of energy issues.
	[200]	10	-	-	Simulator	App name, thread power	AppSets	Optimize scheduling sequences for frequent app usage sets.
	[201]	405	-	-	Monitoring app	Nanosecond and millisecond level energy information	Descriptive statistics	On average apps spend 61% of their energy in idle states. The network is the most energy consuming component. A few system APIs are of significant energy consuming.
App traffic patterns	[32]	22,000	600,000	1 week	Network operator	App usage, traffic volume, time, location	Descriptive statistics	In terms of traffic sources, mobile apps could be divided into local apps and national apps.
	[202]	23	-	-	Monitoring app	App usage, network delay	DBSCAN clustering	Over 60% of popular apps optimize network delay under 350ms.
	[203]	29	2 million	2 days	Network operator	App usage, traffic, operating system	Descriptive statistics	IOS apps cause more traffic consumption compared with Android and Windows Phone.
	[204]	185,173	-	50 days	Monitoring app	App usage, traffic, IP, time, device model	SVM, decision tree, maximum entropy	Predicting data collection purpose of apps with an average precision of 84%.
	[205]	-	65,117	1 month	Monitoring app	App usage, traffic	Descriptive statistics, clustering	Entertainment and social media apps are the most traffic-intensive, while education and weather apps are the least traffic-intensive.
	[206]	7,215	125,609	1 month	Network operator	App usage, traffic, time, location	Descriptive statistics	Music and shopping apps see the most traffic in the morning, while e-mail and game apps see the most traffic during lunchtime.
	[207]	-	50	2 months	Monitoring app	App usage, data throughput, connection type, duration, location	Random forest	Predicting network resources occupied by apps with a mean absolute percentage error of 12.54%$\sim$13.39%.
	[10]	2,000	1,1188,000	6 days	Network operator	App usage, app category, base station, traffic, time	TF-IDF, K-means, linear regression	Predicting popular apps in a certain base station with an accuracy of 60%.
	[208]	40	3	2 years	Monitoring app	App usage, packet data, flow metadata	Hidden markov models	Predicting app network traffic of packet and message levels with a G-mean of 0.8.
	[209]	2,000	1 million	1 week	Network operator	App usage, base station, traffic, time	K-means, convex optimization	Optimizing data pricing strategies in terms of mobile app usage patterns.

traffic at various levels, such as packet and message levels, and proposed using the Markov model to predict app network traffic at different levels. Yin *et al.* [209] incorporated mobile app usage patterns into data pricing strategies, creating a pricing scheme that is updated based on user satisfaction and operational conditions.

C. Discussion

This section summarized relevant smartphone domain studies from two principal fields: app energy drain and app traffic patterns. Table XIII lists the most important literature in terms of research topics, dataset information, methods, and key findings. We can see that most datasets for energy drain analysis are gathered from monitoring apps to support battery state collection. Excessive energy consumption is primarily caused by hogs and background apps and unnecessary workload. These findings suggest that app developers should carefully select system APIs to reduce unnecessary power consumption and that operating systems should improve app scheduling algorithms. Some studies have attempted to optimize app scheduling; however, their optimizer is still unable to support personalized requirements, such as adjusting optimization strategies for different user habits. In the field of app traffic patterns, the datasets are collected from network operators and monitoring apps. Network operator datasets typically cover millions of users, which is significantly more than those of monitoring apps. Clustering algorithms, such as K-means and DBSCAN, are used to discover app traffic patterns. In the meantime, classification and regression methods, such as SVM, decision tree, markov model, and random forest, are used to forecast app traffic consumption to optimize network resource allocation and data pricing strategies.

VI. Challenges and Future Research

Smartphone app usage analysis is a promising direction. A significant number of studies have been done in recent years and obtained a lot of achievements. However, there are still several active challenges that have not been well addressed. This section will discuss the challenges first and then look forward to the future research of smartphone app usage analysis.

A. Challenge in Data

1) Data Collection: The collection of app usage data has become much more difficult after the implementation of GDPR. A major reason is that GDPR principles and data collection for big-data analysis are not always mutually exclusive. For example, a large amount of data is preferred in the data collection to reduce deviation and bias in analysis results. However, it goes against the data minimization principle. New hypotheses are frequently introduced after data collection in data analyses. However, it is constrained by the purpose of the users' initial consent. One possible way to overcome such contradictions is to collect anonymized data because anonymized data is excluded from GDPR [214]. While the GDPR has a strict definition of anonymized data, which means that the data subject is not or is no longer identifiable. In other words, it must be impossible to obtain personal information from anonymized data. As a result, anonymized app usage data can only be used in the app and smartphone domain research rather than user domain research.

We have discovered that GDPR has some unintended consequences. Collecting data from network operators, for example, has become extremely difficult due to the difficulty in obtaining user consent. Researchers are hesitant to publish datasets for fear of penalties, especially when European citizens are involved. Furthermore, because data collection in the European Union is difficult, there is a risk that researchers will prefer to use data from places with fewer or no regulations. A few trends have emerged. Recent studies from the last three years, i.e., 2018, 2019, and 2020, are primarily based on data from China and the United States. Only a few studies are based on data from Europe.

2) Mode Effect and Data Bias: Existing app usage datasets vary greatly in terms of collection methods, scales, collected items, and duration. These differences could lead to biases in their studies and make research more difficult to replicate. Different collection methods will result in a mode effect, which will introduce biases into the collected data. App stores and monitoring apps, for example, can only collect data from their users. Furthermore, monitoring apps can impact users' normal app usage behavior due to privacy concerns, resulting in data collection biases. Approximately 1% of participants changed their normal usage behavior during data collection [215]. As a result, it is critical to address participants' privacy concerns and develop more unobtrusive monitoring apps. The population difference across studies also leads to data bias. It is hard to organize a smartphone app usage dataset wholly and accurately reflecting real-world user behaviors. Most existing studies are based on biased populations. The users involved in [34], [43], [69], [78], [169], [187] are all Chinese. The participants in [39], [101], [216] all live in the Lake Geneva region in Switzerland. University students are only considered in [40], [217], [218]. To alleviate the biases, improving the diversity of participants and sampling technology are useful methods to increase the robustness of results.

The replicability of research will be hampered by data bias. Church *et al.* [2], for example, attempted to duplicate previous work. However, their findings differed significantly from previous research, revealing replicability issues and biases across datasets. To solve this problem, we need to put a lot more effort into constructing open datasets. Researchers can use unified public datasets to compare their new algorithms and discoveries to state-of-the-art ones, especially in method-oriented topics like app usage prediction, app recommendation, and predictive user profiling. The standard platform can help to reduce the number of unrepresentative repeated experiments and boost academic loyalty. Thankfully, we have taken the first step toward creating benchmark datasets. As we mentioned in Section II, some researchers have made their app usage datasets public. We believe their public datasets will be strong candidates for benchmark datasets in terms of data quality and impact.

B. Challenge in Methods

1) Heterogeneous Data Fusion: Other types of data, such as smartphone sensor data and app metadata, are also important and helpful for analyzing smartphone app usage. Although some approaches for fusing heterogeneous data in app usage analysis have been proposed, current methods can be improved by increasing their effectiveness and generalizability. For example, there are two widely used methods in existing studies for fusing context data into app usage records. The first one is to use the tensor structure by adding additional dimensions to represent context information. The second one is to take context features as conditions that are parallelly input into a probabilistic model or Bayesian networks. However, when the scale of the context grows large, these two methods will suffer from the curse of dimensionality and high computational cost.

Most existing data fusion models are task-specific, making it difficult to generalize results across different studies. App features, for example, are commonly used in the tasks of app categorization and app recommendations, which are determined based on app descriptions and app usage behaviors. However, the model is difficult to transfer directly due to different dataset settings [2]. The question of how to better fuse heterogeneous data is still an active challenge. Embedding technology could be a promising method. Embedding technology aims to map high-dimensional data into low-dimensional vectors in latent embedding space while maintaining similarity across data points. A few studies [167], [174] have tried to employ embedding to model app usage sequences. Heterogeneous data fusion can also benefit from embedding technology. For example, to describe the relationship between different types of data, we can create a heterogeneous graph. We can learn a low-dimensional data representation that embeds node features and relations together using graph embedding technologies like Metapath2Vec [219] and HAN [220], which improves the effectiveness of data fusion. We can learn universal data representations in this way, such as app embeddings that capture global patterns of app usage, context information, and app metadata. The learned embeddings can be reused in multiple models.

2) Privacy Preserving Analysis: Privacy is always an important consideration when accessing, using, and sharing mobile app usage data. As we discussed in Section IV-A, users' attributes can be inferred from their app usage data. Thus, privacy-preserving technologies must be used when dealing with such sensitive data. Anonymization is a common technique. Böhmer *et al.* [87], for example, used hash functions to replace all participants' personal identifiers with unrecognizable symbols. This method can delink multiple sets of users, reducing the risk of individual identities being leaked along with other public auxiliary data. However, simply anonymizing user IDs is insufficient to ensure privacy. Based on anonymized app usage records, previous studies have shown that most people can be re-identified and tracked [187], [188]. In the privacy-preserving analysis, the problem of how to create a well-established anonymization mechanism remains unresolved.

Most existing studies rely on centralized data processing, which is bound to raise privacy concerns. Collecting app usage data and centralized processing has become more difficult as a result of recently enacted strict privacy protection regulations. The European GDPR, for example, requires data controllers to disclose any data collection and to declare the lawful basis and purpose for data processing. In addition, after iOS 9, the iOS system removed most data collection APIs to address user privacy concerns. As a result, converting to a decentralized analysis framework is critical to protect user privacy and facilitate the research community's long-term development. Federated Learning could be a promising analysis framework [221]–[223]. Federated Learning allows mobiles to learn a shared model collaboratively while keeping all data on the local device. Users are not required to upload their usage data to cloud servers, and their sensitive information will be well protected [224]. Researchers can also take proactive steps to build privacy-preserving mechanisms by making the analysis framework transparent and giving users complete control over their data. These measures will help to protect user privacy to some extent and alleviate privacy concerns.

C. Future Research

1) App Evolution Globalization VS. Localization: Since the release of the first iPhone in 2007, the app ecosystem has evolved significantly over the last two decades. Starting with a few system-embedded apps, the app ecosystem has grown to include over 3 billion apps that cater to a wide range of user needs. Many apps have appeared in recent decades, attracting millions of users before eventually disappearing. Exploring the app ecosystem's evolutionary processes and extracting general rules and impact factors behind app usage is critical for all relevant stakeholders, including smartphone manufacturers, service providers, and app developers, to provide valuable guidance.

Globalization has a significant impact on the app ecosystem's development. Globalization has accelerated interaction and integration among people from various countries and cultural backgrounds in recent decades. The app's usage also reflects such interaction and integration. Apps are easier to distribute around the globe and attract a large number of users. For instance, Pokémon Go, a popular smartphone game, was first released in the United States on July 6, 2016, and quickly spread around the world. The study of how popular apps spread is important for developing business strategies, predicting app popularity, and profiling the app ecosystem.

Localization, on the other hand, is also a trend in the app ecosystem. Many apps provide access to local information and are useful to both tourists and residents. For example, residents of Shenzhen, one of China's largest cities, can use a parking app called Yitingche. KYOTO Trip+[4] is an official app that provides information to visitors and residents of Kyoto, Japan's largest city. According to our careful investigation, a few previous studies have worked on local apps. Ochiai *et al.* [60], for example, created a framework for identifying local apps in app stores and making app recommendations to residents. However, we pay less attention to local apps in existing studies than we do to global apps. There could be two reasons for this. Local apps, tied to a specific location, usually have fewer users and have a less international impact than global apps. Second, conducting studies on local apps is more difficult due to a lack of data. Nonetheless, we would like to emphasize that, in contrast to global apps that support a broad range of interests, local apps focus on local life services and deserve more research attention.

2) Context-Aware App Usage Modeling: Context-aware app usage modeling is a critical but difficult problem to solve. Breakthroughs in mobile app usage pattern discovery, app usage prediction, and app recommendations will be made if this problem is solved. App usage modeling can also help with the creation of synthetic datasets for benchmarking systems. However, because app usage varies depending on the context, modeling such complex and dynamic behaviors is difficult. Thankfully, a few existing studies have taken the first steps. Yu *et al.* [131] depicted the co-occurrence of apps, locations, and time to model app usage behaviors. Zhao *et al.* [126] considered user characteristics. However, existing research either ignores the sequential features of app usage or the location context. Thus, modeling the sequential features of app usage in various contexts, such as location, time, and motion, is still a work in progress. The spatiotemporal graph neural network is one possible solution, in which we can use graph structure to model the co-occurrence of app usage and context factors, and a recurrent neural network (RNN) model to capture sequential patterns.

[4]https://play.google.com/store/apps/details?id=jp.kyoto.pref.visitkyoto&hl

3) Deep Reasoning App Usage Behaviors: Existing research is limited to correlation analysis. Even though correlation analysis can reveal co-occurrence, it lacks interpretability in some cases. When we find a user who uses travel apps frequently, for example, we do not know if he/she is traveling or just planning to travel. Thus, it is difficult to provide direct guidance to service providers. Furthermore, the correlation analysis may become stuck in Simpson's Paradox, resulting in incorrect analysis results. Therefore, adding causal analysis to deep reasoning app usage behaviors is required. Fortunately, causal inference has progressed significantly in recent years. A number of promising methods, such as individual treatment effects estimation [225] and shared causal model [226], have emerged. For reasoning analysis of app usage behaviors, these advanced methods provide robust technology and algorithm support.

Sophisticated deep learning methods are now used to investigate app usage behaviors to improve system performance. These black-box models, however, are unable to interpret the discovered features and how they relate to the significant results obtained. Several deep learning explainers [227] have recently been developed by advanced machine learning studies in an attempt to open the black-box models. Incorporating these deep learning explainer models into app usage analysis will be a promising step toward improving the interoperability of profiling results and providing reliable guidance to relevant stakeholders.

4) Linking User Activities and App Usage: Smartphones, which are supported by a diverse set of mobile apps, allow people to do things like order food, shop, manage their finances, and socialize more conveniently. These spatiotemporal app usage traces have the potential to infer users' physical world activities and can be used for user profiling and authentication. A strong link between physical activities and mobile app usage has been demonstrated in a few studies. Li *et al.* [104] used app usage data to identify seven user activities. However, they focused solely on app usage, ignoring the impact of time and locations. Different physical world activities may be reflected by the same app usage behavior at different times and in different places. As a result, it is necessary to introduce spatial and temporal factors to infer users' physical activities better. This is difficult due to the high complexity of spatiotemporal features. The probabilistic graphical model, which is scalable to model multiple factors and can apply different distribution functions to capture various activity patterns, is one possible solution to this problem.

5) Location-Based Service and Urban Computing: There is a strong link between app usage behaviors and locations, according to numerous existing studies. However, the majority of them were only interested in using locations as contextual features to improve app usage prediction and recommendations. By leveraging app-location relationships, app usage data can be used in location-based services, such as urban computing, in addition to app-oriented tasks. As we discussed in Section II, the app usage dataset gathered from network operators can cover the majority of mobile users in a given area, such as a city or state. The datasets usually include location data derived from associated base stations. Such high-coverage and fine-grained datasets provide rich user behavioral data, such as app usage and user mobility [228], for conducting urban computing studies. There have been a few studies that have focused on this promising area. Xia and Li [89], for example, used users' online behavior, such as app usage, as well as offline behavior, such as human mobility, to uncover urban dynamics and identify urban zone functions.

VII. Conclusion

The research efforts on smartphone app usage analysis are surveyed and summarized in this paper. We first introduced and compared various data sources, such as surveys, monitoring apps, network operators, and app stores. For the research community, we presented a set of public datasets and discussed privacy and ethical issues. The related studies in the app, user, and smartphone domains were surveyed, respectively. We made a detailed taxonomy for each research domain based on the problem investigated, the characteristics of the datasets used, the methods used, and the key results obtained. Finally, we discussed two current research challenges and identified five future research directions for this hot topic.

Acknowledgment

The authors would like to thank Prof. Yong Li for all of the support and valuable discussions.

References

[1] N. Oliver *et al.*, "Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle," *Sci. Adv.*, vol. 6, no. 23, 2020, Art. no. eabc0764.

[2] K. Church, D. Ferreira, N. Banovic, and K. Lyons, "Understanding the challenges of mobile phone usage data," in *Proc. 17th Int. Conf. Human–Comput. Interact. Mobile Devices Services*, 2015, pp. 504–514.

[3] T. Li, Y. Li, T. Xia, and P. Hui, "Finding spatiotemporal patterns of mobile application usage," *IEEE Trans. Netw. Sci. Eng.*, early access, Nov. 30, 2021, doi: 10.1109/TNSE.2021.3131194.

[4] Statista. "Number of available applications in the google play store from December 2009 to July 2021." 2021. [Online]. Available: https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/

[5] Statista. "Worldwide mobile app revenues in 2014 to 2023." 2021. [Online]. Available: https://www.statista.com/statistics/269025/worldwide-mobile-app-revenue-forecasts

[6] J. Liu, Y. Fu, J. Ming, Y. Ren, L. Sun, and H. Xiong, "Effective and real-time in-app activity analysis in encrypted Internet traffic streams," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2017, pp. 335–344.

[7] Y. Chen, X. Jin, J. Sun, R. Zhang, and Y. Zhang, "POWERFUL: Mobile app fingerprinting via power analysis," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, 2017, pp. 1–9.

[8] A. J. Oliner, A. P. Iyer, I. Stoica, E. Lagerspetz, and S. Tarkoma, "Carat: Collaborative energy diagnosis for mobile devices," in *Proc. 11th ACM Conf. Embedded Netw. Sensor Syst.*, 2013, p. 10.

[9] F. Xu, Y. Li, H. Wang, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 1147–1161, Apr. 2017.

[10] M. Zeng *et al.*, "Temporal–spatial mobile application usage understanding and popularity prediction for edge caching," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 36–42, Jun. 2018.

[11] S. Zhao *et al.*, "User profiling from their use of smartphone applications: A survey," *Pervasive Mobile Comput.*, vol. 59, Oct. 2019, Art. no. 101052.

[12] A. De Masi and K. Wac, "You're using this app for what? A MQOL living lab study," in *Proc. ACM Int. Joint Conf. Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput.*, 2018, pp. 612–617.

[13] B. Guo, Y. Ouyang, T. Guo, L. Cao, and Z. Yu, "Enhancing mobile app user understanding and marketing with heterogeneous crowdsourced data: A review," *IEEE Access*, vol. 7, pp. 68557–68571, 2019.

[14] W. Pan, N. Aharony, and A. Pentland, "Composite social network for predicting mobile apps installation," in *Proc. 25th AAAI Conf. Artif. Intell.*, 2011, pp. 821–827.

[15] D. Li, W. Li, X. Wang, C.-T. Nguyen, and S. Lu, "App trajectory recognition over encrypted Internet traffic based on deep neural network," *Comput. Netw.*, vol. 179, Oct. 2020, Art. no. 107372.

[16] X. Lu *et al.*, "Mining usage data from large-scale Android users: Challenges and opportunities," in *Proc. IEEE/ACM Int. Conf. Mobile Softw. Eng. Syst. (MOBILESoft)*, 2016, pp. 301–302.

[17] H. Cao and M. Lin, "Mining smartphone data for app usage prediction and recommendations: A survey," *Pervasive Mobile Comput.*, vol. 37, pp. 1–22, Jun. 2017.

[18] F. Kreuter, G.-C. Haas, F. Keusch, S. Bähr, and M. Trappmann, "Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent," *Soc. Sci. Comput. Rev.*, vol. 38, no. 5, pp. 533–549, 2020.

[19] S. L. Lim, P. J. Bentley, N. Kanakam, F. Ishikawa, and S. Honiden, "Investigating country differences in mobile app user behavior and challenges for software engineering," *IEEE Trans. Softw. Eng.*, vol. 41, no. 1, pp. 40–64, Jan. 2015.

[20] G. R. Jesse, "Smartphone and app usage among college students: Using smartphones effectively for social and educational needs," in *Proc. EDSIG Conf.*, 2015, p. 3424.

[21] A. Hiniker, S. N. Patel, T. Kohno, and J. A. Kientz, "Why would you do that? Predicting the uses and gratifications behind smartphone-usage behaviors," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 634–645.

[22] R. Wang *et al.*, "StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 3–14.

[23] M. L. Gordon, L. Gatys, C. Guestrin, J. P. Bigham, A. Trister, and K. Patel, "App usage predicts cognitive ability in older adults," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2019, pp. 1–12.

[24] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 171–209, 2014.

[25] I. Andone, K. Błaszkiewicz, M. Eibes, B. Trendafilov, C. Montag, and A. Markowetz, "Menthal: A framework for mobile data collection and analysis," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Adjunct*, 2016, pp. 624–629.

[26] B. Finley and T. Soikkeli, "Mobile device type substitution," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–20, 2018.

[27] X. Su, D. Zhang, W. Li, and X. Wang, "AndroGenerator: An automated and configurable Android app network traffic generation system," *Security Commun. Netw.*, vol. 8, no. 18, pp. 4273–4288, 2015.

[28] A. Zuniga *et al.*, "Tortoise or hare? Quantifying the effects of performance on mobile app retention," in *Proc. World Wide Web Conf.*, 2019, pp. 2517–2528.

[29] M. Krichen, "Anomalies detection through smartphone sensors: A review," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7207–7217, Mar. 2021.

[30] Q. Xu *et al.*, "Automatic generation of mobile app signatures from traffic observations," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2015, pp. 1481–1489.

[31] J. Yang, Y. Qiao, X. Zhang, H. He, F. Liu, and G. Cheng, "Characterizing user behavior in mobile Internet," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 1, pp. 95–106, Mar. 2015.

[32] Q. Xu, J. Erman, A. Gerber, Z. M. Mao, J. Pang, and S. Venkataraman, "Identifying diverse usage behaviors of smartphone apps," in *Proc. ACM SIGCOMM Conf. Internet Meas. Conf.*, 2011, pp. 329–344.

[33] S. Zhao *et al.*, "Discovering different kinds of smartphone users through their application usage behaviors," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 498–509.

[34] J. Huang, F. Xu, Y. Lin, and Y. Li, "On the understanding of interdependency of mobile app usage," in *Proc. IEEE 14th Int. Conf. Mobile Ad Hoc Sensor Syst. (MASS)*, 2017, pp. 471–475.

[35] H. Yao, G. Ranjan, A. Tongaonkar, Y. Liao, and Z. M. Mao, "SAMPLES: Self adaptive mining of persistent lexical snippets for classifying mobile application traffic," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 439–451.

[36] X. Wang, S. Chen, and J. Su, "Real network traffic collection and deep learning for mobile app identification," *Wireless Commun. Mobile Comput.*, vol. 2020, pp. 1–14, Feb. 2020.

[37] V. F. Taylor, R. Spolaor, M. Conti, and I. Martinovic, "Robust smartphone app identification via encrypted network traffic analysis," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 63–78, Jan. 2018.

[38] C. Xiang, Q. Chen, M. Xue, and H. Zhu, "APPCLASSIFIER: Automated app inference on encrypted traffic via meta data analysis," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–7.

[39] J. K. Laurila *et al.*, "The mobile data challenge: Big data for mobile computing research," in *Proc. Pervasive Computing*, 2012, pp. 1–8.

[40] C. Shepard, A. Rahmati, C. Tossell, L. Zhong, and P. Kortum, "LiveLab: Measuring wireless networks and smartphone users in the field," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 38, no. 3, pp. 15–20, 2011.

[41] C. Stachl *et al.*, "Predicting personality from patterns of behavior collected with smartphones," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 30, pp. 17680–17687, 2020.

[42] H. Wang, F. Xu, Y. Li, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," in *Proc. Internet Meas. Conf.*, 2015, pp. 225–238.

[43] X. Liu *et al.*, "Understanding diverse usage patterns from large-scale appstore-service profiles," *IEEE Trans. Softw. Eng.*, vol. 44, no. 4, pp. 384–411, Apr. 2018.

[44] F. Lin, H. Wang, L. Wang, and X. Liu, "A longitudinal study of removed apps in IoS app store," in *Proc. Web Conf.*, 2021, pp. 1435–1446.

[45] S. Greengard, "Weighing the impact of GDPR," *Commun. ACM*, vol. 61, no. 11, pp. 16–18, Oct. 2018.

[46] N. Momen, M. Hatamian, and L. Fritsch, "Did app privacy improve after the GDPR?" *IEEE Security Privacy*, vol. 17, no. 6, pp. 10–20, Nov./Dec. 2019.

[47] J. L. Kröger, J. Lindemann, and D. Herrmann, "How do app vendors respond to subject access requests? A longitudinal privacy study on IoS and Android apps," in *Proc. 15th Int. Conf. Availability Rel. Security*, 2020, pp. 1–10.

[48] J. M. Wing, "Data for good," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2018, p. 4.

[49] J. Isaak and M. J. Hanna, "User data privacy: Facebook, cambridge analytica, and privacy protection," *Computer*, vol. 51, no. 8, pp. 56–59, 2018.

[50] T. Petsas, A. Papadogiannakis, M. Polychronakis, E. P. Markatos, and T. Karagiannis, "Rise of the planet of the apps: A systematic study of the mobile app ecosystem," in *Proc. Conf. Internet Meas. Conf.*, 2013, pp. 277–290.

[51] H. Mendelson and K. Moon, "Modeling success and engagement for the app economy," in *Proc. World Wide Web Conf.*, 2018, pp. 569–578.

[52] H. Wang, H. Li, and Y. Guo, "Understanding the evolution of mobile app ecosystems: A longitudinal measurement study of google play," in *Proc. World Wide Web Conf.*, 2019, pp. 1988–1999.

[53] T. Petsas, A. Papadogiannakis, M. Polychronakis, E. P. Markatos, and T. Karagiannis, "Measurement, modeling, and analysis of the mobile app ecosystem," *ACM Trans. Model. Perform. Eval. Comput. Syst.*, vol. 2, no. 2, pp. 1–33, 2017.

[54] C. Jesdabodi and W. Maalej, "Understanding usage states on mobile devices," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2015, pp. 1221–1225.

[55] W. Sun, S. Wu, and Z. Xue, "Clustering mobile apps based on design and manufacturing genre," in *Proc. IEEE 6th Int. Conf. Comput. Commun. (ICCC)*, 2020, pp. 1956–1960.

[56] H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian, "Exploiting enriched contextual information for mobile app classification," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manag.*, 2012, pp. 1617–1621.

[57] S. Xue, L. Zhang, A. Li, X.-Y. Li, C. Ruan, and W. Huang, "AppDNA: App behavior profiling via graph-based deep learning," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, 2018, pp. 1475–1483.

[58] M. Shamsujjoha, J. Grundy, L. Li, H. Khalajzadeh, and Q. Lu, "Checking app behavior against app descriptions: What if there are no app descriptions?" in *Proc. IEEE/ACM 29th Int. Conf. Program Comprehension (ICPC) (ICPC)*, May 2021, pp. 422–432.

[59] A. A. Al-Subaihin *et al.*, "Clustering mobile apps based on mined textual features," in *Proc. 10th ACM/IEEE Int. Symp. Empirical Softw. Eng. Meas.*, 2016, pp. 1–10.

[60] K. Ochiai, F. Putri, and Y. Fukazawa, "Local app classification using deep neural network based on mobile app market data," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. PerCom*, 2019, pp. 186–191.

[61] A. Fuad and M. Al-Yahya, "Analysis and classification of mobile apps using topic modeling: A case study on google play arabic apps," *Complexity*, vol. 2021, Feb. 2021, Art. no. 6677413.

[62] D. Surian, S. Seneviratne, A. Seneviratne, and S. Chawla, "App miscategorization detection: A case study on google play," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 8, pp. 1591–1604, Aug. 2017.

[63] H. Zhu, E. Chen, H. Xiong, H. Cao, and J. Tian, "Mobile app classification with enriched contextual information," *IEEE Trans. Mobile Comput.*, vol. 13, no. 7, pp. 1550–1563, Jul. 2014.

[64] A. Berger, S. A. D. Pietra, and V. J. D. Pietra, "A maximum entropy approach to natural language processing," *Comput. Linguist.*, vol. 22, no. 1, pp. 39–71, 1996.

[65] V. Radosavljevic *et al.*, "Smartphone app categorization for interest targeting in advertising marketplace," in *Proc. 25th Int. Conf. Companion World Wide Web*, 2016, pp. 93–94.

[66] Y. He, C. Wang, G. Xu, W. Lian, H. Xian, and W. Wang, "Privacy-preserving categorization of mobile applications based on large-scale usage data," *Inf. Sci.*, vol. 514, pp. 557–570, Apr. 2020.

[67] E. J. Keogh and C. A. Ratanamahatana, "Exact indexing of dynamic time warping," *Knowl. Inf. Syst.*, vol. 7, no. 3, pp. 358–386, 2005.

[68] A. Morrison, X. Xiong, M. Higgs, M. Bell, and M. Chalmers, "A large-scale study of iphone app launch behaviour," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2018, pp. 1–13.

[69] H. Li *et al.*, "Characterizing smartphone usage patterns from millions of Android users," in *Proc. Internet Meas. Conf.*, 2015, pp. 459–472.

[70] H. Wang *et al.*, "Beyond google play: A large-scale comparative study of Chinese Android app markets," in *Proc. Internet Meas. Conf.*, 2018, pp. 293–307.

[71] S. Sigg, E. Lagerspetz, E. Peltonen, P. Nurmi, and S. Tarkoma, "Exploiting usage to predict instantaneous app popularity: Trend filters and retention rates," *ACM Trans. Web*, vol. 13, no. 2, pp. 1–25, 2019.

[72] Y. Tian, K. Zhou, and D. Pelleg, "What and how long: Prediction of mobile app engagement," *ACM Trans. Inf. Syst.*, vol. 40, no. 1, pp. 1–38, 2022.

[73] T. Li, M. Zhang, H. Cao, Y. Li, S. Tarkoma, and P. Hui, "What apps did you use? Understanding the long-term evolution of mobile app usage," in *Proc. Web Conf.*, 2020, pp. 66–76.

[74] T. Li, Y. Fan, Y. Li, S. Tarkoma, and P. Hui, "Understanding the long-term evolution of mobile app usage," *IEEE Trans. Mobile Comput.*, early access, Jul. 21, 2021, doi: 10.1109/TMC.2021.3098664.

[75] S. Shen, X. Lu, Z. Hu, and X. Liu, "Towards release strategy optimization for apps in google play," in *Proc. 9th Asia–Pac. Symp. Internetware*, 2017, pp. 1–10.

[76] B. Fu, J. Lin, L. Li, C. Faloutsos, J. Hong, and N. Sadeh, "Why people hate your app: Making sense of user feedback in a mobile app store," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2013, pp. 1276–1284.

[77] R. Potharaju, M. Rahman, and B. Carbunar, "A longitudinal study of google play," *IEEE Trans. Comput. Soc. Syst.*, vol. 4, no. 3, pp. 135–149, Sep. 2017.

[78] X. Lu *et al.*, "PRADA: Prioritizing Android devices for apps by mining large-scale usage data," in *Proc. IEEE/ACM 38th Int. Conf. Softw. Eng. (ICSE)*, 2016, pp. 3–13.

[79] Y. Wang, N. J. Yuan, Y. Sun, C. Qin, and X. Xie, "App download forecasting: An evolutionary hierarchical competition approach," in *Proc. IJCAI*, 2017, pp. 2978–2984.

[80] H. Zhu, C. Liu, Y. Ge, H. Xiong, and E. Chen, "Popularity modeling for mobile apps: A sequential approach," *IEEE Trans. Cybern.*, vol. 45, no. 7, pp. 1303–1314, Jul. 2015.

[81] Y. Ouyang, B. Guo, T. Guo, L. Cao, and Z. Yu, "Modeling and forecasting the popularity evolution of mobile apps: A multivariate hawkes process approach," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–23, 2018.

[82] Y. Zhang, J. Liu, B. Guo, Z. Wang, Y. Liang, and Z. Yu, "App popularity prediction by incorporating time-varying hierarchical interactions," *IEEE Trans. Mobile Comput.*, vol. 21, no. 5, pp. 1566–1579, May 2022.

[83] H. Li *et al.*, "Voting with their feet: Inferring user preferences from app management activities," in *Proc. 25th Int. Conf. World Wide Web*, 2016, pp. 1351–1362.

[84] P. Papapetrou and G. Roussos, "Social context discovery from temporal app use patterns," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Adjunct Publication*, 2014, pp. 397–402.

[85] A. Mehrotra *et al.*, "Understanding the role of places and activities on mobile phone interaction and usage patterns," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 1–22, 2017.

[86] T. M. T. Do, J. Blom, and D. Gatica-Perez, "Smartphone usage in the wild: A large-scale analysis of applications and context," in *Proc. 13th Int. Conf. Multimodal Interfaces*, 2011, pp. 353–360.

[87] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer, "Falling asleep with angry birds, facebook and kindle: A large scale study on mobile application usage," in *Proc. 13th Int. Conf. Human Comput. Interact. Mobile Devices Services*, 2011, pp. 47–56.

[88] E. Graells-Garrido, D. Caro, O. Miranda, R. Schifanella, and O. F. Peredo, "The WWW (and an H) of mobile application usage in the city: The what, where, when, and how," in *Proc. Web Conf.*, 2018, pp. 1221–1229.

[89] T. Xia and Y. Li, "Revealing urban dynamics by learning online and offline behaviours together," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 1, pp. 1–25, 2019.

[90] Y. Ren, T. Xia, Y. Li, and X. Chen, "Predicting socio-economic levels of urban regions via offline and online indicators," *PLoS ONE*, vol. 14, no. 7, 2019, Art. no. e0219058.

[91] S. Van Canneyt, M. Bron, A. Haines, and M. Lalmas, "Describing patterns and disruptions in large scale mobile app usage data," in *Proc. 26th Int. Conf. World Wide Web Companion*, 2017, pp. 1579–1584.

[92] T. Li, M. Zhang, Y. Li, E. Lagerspetz, S. Tarkoma, and P. Hui, "The impact of COVID-19 on smartphone usage," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16723–16733, Dec. 2021.

[93] V. Srinivasan, S. Moghaddam, A. Mukherji, K. K. Rachuri, C. Xu, and E. M. Tapia, "MobileMiner: Mining your frequent patterns on your phone," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 389–400.

[94] Y. Tian, K. Zhou, M. Lalmas, and D. Pelleg, "Identifying tasks from mobile app usage patterns," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2020, pp. 2357–2366.

[95] B. Friedrichs, L. D. Turner, and S. M. Allen, "Discovering types of smartphone usage sessions from user-app interactions," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, 2021, pp. 459–464.

[96] A. Rahmati, C. Shepard, C. Tossell, L. Zhong, and P. Kortum, "Practical context awareness: Measuring and utilizing the context dependency of mobile usage," *IEEE Trans. Mobile Comput.*, vol. 14, no. 9, pp. 1932–1946, Sep. 2015.

[97] Y. Fan *et al.*, "Understanding the long-term dynamics of mobile app usage context via graph embedding," *IEEE Trans. Knowl. Data Eng.*, early access, Sep. 3, 2021, doi: 10.1109/TKDE.2021.3110141.

[98] Y. Liu, K. Yu, X. Wu, Y. Shi, and Y. Tan, "Association rules mining analysis of app usage based on mobile traffic flow data," in *Proc. IEEE 3rd Int. Conf. Big Data Anal. (ICBDA)*, 2018, pp. 55–60.

[99] W.-R. Tseng and K.-W. Hsu, "Smartphone app usage log mining," *Int. J. Comput. Elect. Eng.*, vol. 6, no. 2, pp. 151–156, 2014.

[100] D. Ferreira, J. Goncalves, V. Kostakos, L. Barkhuus, and A. K. Dey, "Contextual experience sampling of mobile application micro-usage," in *Proc. 16th Int. Conf. Human–Comput. Interact. Mobile Devices Services*, 2014, pp. 91–100.

[101] A. Shema and D. E. Acuna, "Show me your app usage and I will tell who your close friends are: Predicting user's context from simple cellphone activity," in *Proc. CHI Conf. Extended Abstracts Human Factors Comput. Syst.*, 2017, pp. 2929–2935.

[102] I. M. Kloumann, L. A. Adamic, J. M. Kleinberg, and S. Wu, "The lifecycles of apps in a social ecosystem," in *Proc. 24th Int. Conf. World Wide Web*, 2015, pp. 581–591.

[103] D. G. Taylor, T. A. Voelker, and I. Pentina, *Mobile Application Adoption by Young Adults: A Social Network Perspective*, WCOB Faculty, Jack Welch College Bus., Fairfield CT, USA, 2011.

[104] T. Li, Y. Li, M. A. Hoque, T. Xia, S. Tarkoma, and P. Hui, "To what extent we repeat ourselves? Discovering daily activity patterns across mobile app usage," *IEEE Trans. Mobile Comput.*, vol. 21, no. 4, pp. 1492–1507, Apr. 2022.

[105] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang, "Characterizing geospatial dynamics of application usage in a 3G cellular data network," in *Proc. IEEE INFOCOM*, 2012, pp. 1341–1349.

[106] N. Ghahramani and C. Brakewood, "Trends in mobile transit information utilization: An exploratory analysis of transit app in New York City," *J. Public Transp.*, vol. 19, no. 3, p. 9, 2016.

[107] V. Kostakos, D. Ferreira, J. Goncalves, and S. Hosio, "Modelling smartphone usage: A Markov state transition model," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 486–497.

[108] T.-M.-T. Do and D. Gatica-Perez, "By their apps you shall understand them: Mining large-scale patterns of mobile phone usage," in *Proc. 9th Int. Conf. Mobile Ubiquitous Multimedia*, 2010, pp. 1–10.

[109] F. A. Silva, A. C. S. A. Domingues, and T. R. B. Silva, "Discovering mobile application usage patterns from a large-scale dataset," *ACM Trans. Knowl. Disc. Data*, vol. 12, no. 5, pp. 1–36, 2018.

[110] J. Lee and D.-H. Shin, "Targeting potential active users for mobile app install advertising: An exploratory study," *Int. J. Human–Comput. Interact.*, vol. 32, no. 11, pp. 827–834, 2016.
[111] S. L. Jones, D. Ferreira, S. Hosio, J. Goncalves, and V. Kostakos, "Revisitation analysis of smartphone app use," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2015, pp. 1197–1208.
[112] L. A. Leiva, M. Böhmer, S. Gehring, and A. Krüger, "Back to the app: The costs of mobile application interruptions," in *Proc. 14th Int. Conf. Human–Comput. Interact. Mobile Devices Services*, 2012, pp. 291–294.
[113] J. P. Rula, B. Jun, and F. E. Bustamante, "Mobile AD (D) estimating mobile app session times for better ads," in *Proc. 16th Int. Workshop Mobile Comput. Syst. Appl.*, 2015, pp. 123–128.
[114] H. Cao, Z. Chen, F. Xu, Y. Li, and V. Kostakos, "Revisitation in urban space vs. online: A comparison across pois, websites, and smartphone apps," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–24, 2018.
[115] Y. Tian, K. Zhou, M. Lalmas, Y. Liu, and D. Pelleg, "Cohort modeling based app category usage prediction," in *Proc. 28th ACM Conf. User Model. Adapt. Pers.*, 2020, pp. 248–256.
[116] Z. Tu, Y. Li, P. Hui, L. Su, and D. Jin, "Personalized mobile app prediction by learning user's interest from social media," *IEEE Trans. Mobile Comput.*, vol. 19, no. 11, pp. 2670–2683, Nov. 2020.
[117] D. Yu, Y. Li, F. Xu, P. Zhang, and V. Kostakos, "Smartphone app usage prediction using points of interest," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–21, 2018.
[118] C. Shin, J.-H. Hong, and A. K. Dey, "Understanding and prediction of mobile application usage for smart phones," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 173–182.
[119] N. Natarajan, D. Shin, and I. S. Dhillon, "Which app will you use next? Collaborative filtering with interactional context," in *Proc. 7th ACM Conf. Recommender Syst.*, 2013, pp. 201–208.
[120] X. Zou, W. Zhang, S. Li, and G. Pan, "Prophet: What app you wish to use next," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publication*, 2013, pp. 167–170.
[121] R. Baeza-Yates, D. Jiang, F. Silvestri, and B. Harrison, "Predicting the next app that you are going to use," in *Proc. 8th ACM Int. Conf. Web Search Data Min.*, 2015, pp. 285–294.
[122] S. Xu, W. Li, X. Zhang, S. Gao, T. Zhan, and S. Lu, "Predicting and recommending the next smartphone apps based on recurrent neural network," *CCF Trans. Pervasive Comput. Interact.*, vol. 2, no. 4, pp. 314–328, 2020.
[123] Y. Lee, S. Cho, and J. Choi, "App usage prediction for dual display device via two-phase sequence modeling," *Pervasive Mobile Comput.*, vol. 58, Aug. 2019, Art. no. 101025.
[124] Y. Jiang, X. Du, and T. Jin, "Using combined network information to predict mobile application usage," *Physica A Stat. Mech. Appl.*, vol. 515, pp. 430–439, Feb. 2019.
[125] H. Wang, Y. Li, M. Du, Z. Li, and D. Jin, "App2Vec: Context-aware application usage prediction," *ACM Trans. Knowl. Disc. Data*, vol. 15, no. 6, pp. 1–21, 2021.
[126] S. Zhao *et al.*, "AppUsage2Vec: Modeling smartphone app usage for prediction," in *Proc. IEEE 35th Int. Conf. Data Eng. (ICDE)*, 2019, pp. 1322–1333.
[127] A. Parate, M. Böhmer, D. Chu, D. Ganesan, and B. M. Marlin, "Practical prediction and prefetch for faster access to applications on mobile phones," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2013, pp. 275–284.
[128] H. Wang *et al.*, "Modeling spatio-temporal app usage for a large user population," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 1, pp. 1–23, 2019.
[129] T. Xia *et al.*, "DeepApp: Predicting personalized smartphone app usage via context-aware multi-task learning," *ACM Trans. Intell. Syst. Technol.*, vol. 11, no. 6, pp. 1–12, 2020.
[130] X. Chen, Y. Wang, J. He, S. Pan, Y. Li, and P. Zhang, "CAP: Context-aware app usage prediction with heterogeneous graph embedding," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 1, pp. 1–25, 2019.
[131] Y. Yu, T. Xia, H. Wang, J. Feng, and Y. Li, "Semantic-aware spatio-temporal app usage representation via graph convolutional network," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–24, 2020.
[132] Y. Zhou, S. Li, and Y. Liu, "Graph-based method for app usage prediction with attributed heterogeneous network embedding," *Future Internet*, vol. 12, no. 3, p. 58, 2020.
[133] X.-X. Zhao, Y. Qiao, Z. Si, J. Yang, and A. Lindgren, "Prediction of user app usage behavior from geo-spatial data," in *Proc. 3rd Int. ACM SIGMOD Workshop Manag. Min. Enriched Geo-Spatial Data*, 2016, pp. 1–6.
[134] M. De Nadai, A. Cardoso, A. Lima, B. Lepri, and N. Oliver, "Strategies and limitations in app usage and human mobility," *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, 2019.
[135] T. M. T. Do and D. Gatica-Perez, "Where and what: Using smartphones to predict next locations and applications in daily life," *Pervasive Mobile Comput.*, vol. 12, pp. 79–91, Jun. 2014.
[136] Y. Xu *et al.*, "Preference, context and communities: A multi-faceted approach to predicting smartphone app usage patterns," in *Proc. Int. Symp. Wearable Comput.*, 2013, pp. 69–76.
[137] D. Han, J. Li, L. Yang, and Z. Zeng, "A recommender system to address the cold start problem for app usage prediction," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 9, pp. 2257–2268, 2019.
[138] R. Jisha, J. Amrita, A. R. Vijay, and G. Indhu, "Mobile app recommendation system using machine learning classification," in *Proc. IEEE 4th Int. Conf. Comput. Methodol. Commun. (ICCMC)*, 2020, pp. 940–943.
[139] B. Yan and G. Chen, "Appjoy: Personalized mobile application discovery," in *Proc. 9th Int. Conf. Mobile Syst. Appl. Services*, 2011, pp. 113–126.
[140] K. Shi and K. Ali, "Getjar mobile application recommendations with very sparse datasets," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2012, pp. 204–212.
[141] T. Liang *et al.*, "Mobile app recommendation via heterogeneous graph neural network in edge computing," *Appl. Soft Comput.*, vol. 103, May 2021, Art. no. 107162.
[142] Y. Ouyang, B. Guo, X. Tang, X. He, J. Xiong, and Z. Yu, "Mobile app cross-domain recommendation with multi-graph neural network," *ACM Trans. Knowl. Disc. Data*, vol. 15, no. 4, pp. 1–21, 2021.
[143] Y. Xu, Y. Zhu, Y. Shen, and J. Yu, "Leveraging app usage contexts for app recommendation: A neural approach," *World Wide Web*, vol. 22, no. 6, pp. 2721–2745, 2019.
[144] M. Böhmer, L. Ganev, and A. Krüger, "AppFunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications," in *Proc. Int. Conf. Intell. User Interfaces*, 2013, pp. 267–276.
[145] Z. Tu, B. Duan, Z. Wang, and X. Xu, "Bidirectional sensing of user preferences and application changes for dynamic mobile app recommendations," *Neural Comput. Appl.*, vol. 33, no. 16, pp. 9791–9803, 2021.
[146] C. Liu, J. Cao, and S. Feng, "Leveraging kernel-incorporated matrix factorization for app recommendation," *ACM Trans. Knowl. Disc. Data*, vol. 13, no. 3, pp. 1–27, 2019.
[147] T. Liang *et al.*, "A broad learning approach for context-aware mobile application recommendation," in *Proc. IEEE Int. Conf. Data Min. (ICDM)*, 2017, pp. 955–960.
[148] J. Lin, K. Sugiyama, M.-Y. Kan, and T.-S. Chua, "New and improved: Modeling versions to improve app recommendation," in *Proc. 37th Int. ACM SIGIR Conf. Res Develop. Inf. Retrieval*, 2014, pp. 647–656.
[149] D. Cao *et al.*, "Version-sensitive mobile app recommendation," *Inf. Sci.*, vol. 381, pp. 161–175, Mar. 2017.
[150] S. Gao, L. Liu, Y. Liu, H. Liu, Y. Wang, and P. Liu, "App recommendation based on both quality and security," *J. Softw. Evol. Process*, vol. 33, no. 3, 2021, Art. no. e2325.
[151] T. Rocha, E. Souto, and K. El-Khatib, "Functionality-based mobile application recommendation system with security and privacy awareness," *Comput. Security*, vol. 97, Oct. 2020, Art. no. 101972.
[152] H. Wang, Z. Tu, Y. Fu, Z. Wang, and X. Xu, "Time-aware user profiling from personal service ecosystem," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3597–3619, 2021.
[153] H. Kwon, S. Lee, and D. Jeong, "User profiling via application usage pattern on digital devices for digital forensics," *Exp. Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114488.
[154] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, "Deep learning based inference of private information using embedded sensors in smart devices," *IEEE Netw.*, vol. 32, no. 4, pp. 8–14, Jul./Aug. 2018.
[155] R. Razavi, "Personality segmentation of users through mining their mobile usage patterns," *Int. J. Human–Comput. Stud.*, vol. 143, Nov. 2020, Art. no. 102470.
[156] I. Andone, K. Błaszkiewicz, M. Eibes, B. Trendafilov, C. Montag, and A. Markowetz, "How age and gender affect smartphone usage," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Adjunct*, 2016, pp. 9–12.

[157] S. Zhao *et al.*, "Investigating smartphone user differences in their application usage behaviors: An empirical study," *CCF Trans. Pervasive Comput. Interact.*, vol. 1, no. 2, pp. 140–161, 2019.

[158] Z. Tu *et al.*, "Demographics of mobile app usage: Long-term analysis of mobile app usage," *CCF Trans. Pervasive Comput. Interact.*, vol. 3, pp. 235–252, Apr. 2021.

[159] E. Peltonen *et al.*, "The hidden image of mobile apps: Geographic, demographic, and cultural factors in mobile usage," in *Proc. 20th Int. Conf. Human–Comput. Interact. Mobile Devices Services*, 2018, pp. 1–12.

[160] E. Guzman, L. Oliveira, Y. Steiner, L. C. Wagner, and M. Glinz, "User feedback in the app store: A cross-cultural study," in *Proc. IEEE/ACM 40th Int. Conf. Softw. Eng. Softw. Eng. Soc. (ICSE-SEIS)*, 2018, pp. 13–22.

[161] F. Huseynov, "Understanding usage behavior of different mobile application categories based on personality traits," *Interact. Comput.*, vol. 32, no. 1, pp. 66–80, 2020.

[162] F. Beierle *et al.*, "Frequency and duration of daily smartphone usage in relation to personality traits," *Digit. Psychol.*, vol. 1, no. 1, pp. 20–28, 2020.

[163] Y. Gao, A. Li, T. Zhu, X. Liu, and X. Liu, "How smartphone usage correlates with social anxiety and loneliness," *PeerJ*, vol. 4, Jul. 2016, Art. no. e2197.

[164] K. Katevas, I. Arapakis, and M. Pielot, "Typical phone use habits: Intense use does not predict negative well-being," in *Proc. 20th Int. Conf. Human–Comput. Interact. Mobile Devices Services*, 2018, pp. 1–13.

[165] Z. Sarsenbayeva *et al.*, "Does smartphone use drive our emotions or vice versa? A causal analysis," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2020, pp. 1–15.

[166] S. Zhao *et al.*, "Who are the smartphone users? Identifying user groups with apps usage behaviors," *Mobile Comput. Commun.*, vol. 21, no. 2, pp. 31–34, 2017.

[167] Y. Lee, I. Park, S. Cho, and J. Choi, "Smartphone user segmentation based on app usage sequence with neural networks," *Telemat. Informat.*, vol. 35, no. 2, pp. 329–339, 2018.

[168] R. M. Frey, R. Xu, and A. Ilic, "Mobile app adoption in different life stages: An empirical analysis," *Pervasive Mobile Comput.*, vol. 40, pp. 512–527, Sep. 2017.

[169] X. Chen, Z. Zhu, M. Chen, and Y. Li, "Large-scale mobile fitness app usage analysis for smart health," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 46–52, Apr. 2018.

[170] S. Seneviratne, A. Seneviratne, P. Mohapatra, and A. Mahanti, "Your installed apps reveal your gender and more!" *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 18, no. 3, pp. 55–61, 2015.

[171] S. Seneviratne, A. Seneviratne, P. Mohapatra, and A. Mahanti, "Predicting user traits from a snapshot of apps installed on a smartphone," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 18, no. 2, pp. 1–8, 2014.

[172] E. Malmi and I. Weber, "You are what apps you use: Demographic prediction based on user's apps," in *Proc. 10th Int. AAAI Conf. Web Soc. Media*, 2016, pp. 635–638.

[173] S. Zhao, G. Pan, J. Tao, Z. Luo, S. Li, and Z. Wu, "Understanding smartphone users from installed app lists using Boolean matrix factorization," *IEEE Trans. Cybern.*, vol. 52, no. 1, pp. 384–397, Jan. 2022.

[174] S. Zhao, F. Xu, Z. Luo, S. Li, and G. Pan, "Demographic attributes prediction through app usage behaviors on smartphones," in *Proc. ACM Int. Joint Conf. Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput.*, 2018, pp. 870–877.

[175] S. Bian *et al.*, "A novel macro-micro fusion network for user representation learning on mobile apps," in *Proc. Web Conf.*, 2021, pp. 3199–3209.

[176] S. Zhao *et al.*, "Gender profiling from a single snapshot of apps installed on a smartphone: An empirical study," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1330–1342, Feb. 2020.

[177] Z. Yu, E. Xu, H. Du, B. Guo, and L. Yao, "Inferring user profile attributes from multidimensional mobile phone sensory data," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5152–5162, Jun. 2019.

[178] R. Xu, R. M. Frey, E. Fleisch, and A. Ilic, "Understanding the impact of personality traits on mobile app adoption–insights from a large-scale field study," *Comput. Human Behav.*, vol. 62, pp. 244–256, Sep. 2016.

[179] E. Peltonen, P. Sharmila, K. O. Asare, A. Visuri, E. Lagerspetz, and D. Ferreira, "When phones get personal: Predicting big five personality traits from application usage," *Pervasive Mobile Comput.*, vol. 69, Nov. 2020, Art. no. 101269.

[180] G. Chittaranjan, J. Blom, and D. Gatica-Perez, "Mining large-scale smartphone data for personality studies," *Pers. Ubiquitous Comput.*, vol. 17, no. 3, pp. 433–450, 2013.

[181] N. K. Kambham, K. G. Stanley, and S. Bell, "Predicting personality traits using smartphone sensor data and app usage data," in *Proc. IEEE 9th Annu. Inf. Technol. Electron. Mobile Commun. Conf. (IEMCON)*, 2018, pp. 125–132.

[182] S. Gao, W. Li, L. J. Song, X. Zhang, M. Lin, and S. Lu, "PersonalitySensing: A multi-view multi-task learning approach for personality detection based on smartphone usage," in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 2862–2870.

[183] K. Ochiai, N. Yamamoto, T. Hamatani, Y. Fukazawa, and T. Yamaguchi, "Exploiting graph convolutional networks for representation learning of mobile app usage," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, 2019, pp. 5379–5383.

[184] S. Zhao *et al.*, "Mining user attributes using large-scale app lists of smartphones," *IEEE Syst. J.*, vol. 11, no. 1, pp. 315–323, Mar. 2017.

[185] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *Proc. 8th Int. Conf. Mobile Syst. Appl. Services*, 2010, pp. 179–194.

[186] P. Welke, I. Andone, K. Blaszkiewicz, and A. Markowetz, "Differentiating smartphone users by app usage," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 519–523.

[187] Z. Tu *et al.*, "Your apps give you away: Distinguishing mobile users by their app usage fingerprints," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–23, 2018.

[188] V. Sekara, L. Alessandretti, E. Mones, and H. Jonsson, "Temporal and cultural limits of privacy in smartphone app usage," *Sci. Rep.*, vol. 11, no. 1, pp. 1–9, 2021.

[189] Y. Ashibani and Q. H. Mahmoud, "A multi-feature user authentication model based on mobile app interactions," *IEEE Access*, vol. 8, pp. 96322–96339, 2020.

[190] Y. Ashibani and Q. H. Mahmoud, "A behavior profiling model for user authentication in iot networks based on app usage patterns," in *Proc. IEEE 44th Annu. Conf. Ind. Electron. Soc. (IECON)*, 2018, pp. 2841–2846.

[191] Y. Ashibani and Q. H. Mahmoud, "Design and evaluation of a user authentication model for iot networks based on app event patterns," *Clust. Comput.*, vol. 24, no. 2, pp. 837–850, 2021.

[192] D. Bassu, M. Cochinwala, and A. Jain, "A new mobile biometric based upon usage context," in *Proc. IEEE Int. Conf. Technol. Homeland Security (HST)*, 2013, pp. 441–446.

[193] L. Fridman, S. Weber, R. Greenstadt, and M. Kam, "Active authentication on mobile devices via stylometry, application usage, Web browsing, and GPS location," *IEEE Syst. J.*, vol. 11, no. 2, pp. 513–521, Jun. 2017.

[194] U. Mahbub, J. Komulainen, D. Ferreira, and R. Chellappa, "Continuous authentication of smartphones based on application usage," *IEEE Trans. Biometr. Behav. Identity Sci.*, vol. 1, no. 3, pp. 165–180, Jul. 2019.

[195] R. Pereira *et al.*, "GreenHub: A large-scale collaborative dataset to battery consumption analysis of Android devices," *Empirical Softw. Eng.*, vol. 26, no. 3, pp. 1–55, 2021.

[196] S. Song, F. Wedyan, and Y. Jararweh, "Empirical evaluation of energy consumption for mobile applications," in *Proc. IEEE 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, 2021, pp. 352–357.

[197] X. Chen, N. Ding, A. Jindal, Y. C. Hu, M. Gupta, and R. Vannithamby, "Smartphone energy drain in the wild: Analysis and implications," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 43, no. 1, pp. 151–164, 2015.

[198] X. Chen, A. Jindal, N. Ding, Y. C. Hu, M. Gupta, and R. Vannithamby, "Smartphone background activities in the wild: Origin, energy drain, and optimization," in *Proc. ACM 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 40–52.

[199] X. Li, Y. Yang, Y. Liu, J. P. Gallagher, and K. Wu, "Detecting and diagnosing energy issues for mobile applications," in *Proc. 29th ACM SIGSOFT Int. Symp. Softw. Test. Anal.*, 2020, pp. 115–127.

[200] H. Z. Q. Rex, J. C. Chuen, and A. Herkersdorf, "Apps-usage driven energy management for multicore mobile computing systems," in *Proc. IEEE Int. Symp. Integr. Circuits (ISIC)*, 2014, pp. 472–475.

[201] D. Li, S. Hao, J. Gui, and W. G. Halfond, "An empirical study of the energy consumption of Android applications," in *Proc. IEEE Int. Conf. Softw. Maintenance Evol.*, 2014, pp. 121–130.

[202] Z. Jiang, R. Kuang, J. Gong, H. Yin, Y. Lyu, and X. Zhang, "What makes a great mobile app? A quantitative study using a new mobile crawler," in *Proc. IEEE Symp. Service Orient. Syst. Eng. (SOSE)*, 2018, pp. 222–227.

[203] Y. Li, J. Yang, and N. Ansari, "Cellular smartphone traffic and user behavior analysis," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2014, pp. 1326–1331.

[204] H. Jin *et al.*, "Why are they collecting my data? Inferring the purposes of network traffic in mobile apps," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–27, 2018.

[205] E. A. Walelgne, A. S. Asrese, J. Manner, V. Bajpai, and J. Ott, "Clustering and predicting the data usage patterns of geographically diverse mobile users," *Comput. Netw.*, vol. 187, Mar. 2021, Art. no. 107737.

[206] A. Okic, A. E. Redondi, I. Galimberti, F. Foglia, and L. Venturini, "Analyzing different mobile applications in time and space: A city-wide scenario," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2019, pp. 1–6.

[207] Y. Ouyang and T. Yan, "Profiling wireless resource usage for mobile apps via crowdsourcing-based network analytics," *IEEE Internet Things J.*, vol. 2, no. 5, pp. 391–398, Oct. 2015.

[208] G. Aceto, G. Bovenzi, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapé, "Characterization and prediction of mobile-app traffic using Markov modeling," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 1, pp. 907–925, Mar. 2021.

[209] J. Yin *et al.*, "Mobile app usage patterns aware smart data pricing," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 4, pp. 645–654, Apr. 2020.

[210] D. Li, S. Hao, W. G. Halfond, and R. Govindan, "Calculating source line level energy information for Android applications," in *Proc. ACM Int. Symp. Softw. Test Anal.*, 2013, pp. 78–89.

[211] S. Rezaei, B. Kroencke, and X. Liu, "Large-scale mobile app identification using deep learning," *IEEE Access*, vol. 8, pp. 348–362, 2019.

[212] Cisco. "Cisco Annual Internet Report (2018–2023) White Paper." 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html

[213] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin, "A first look at traffic on smartphones," in *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, 2010, pp. 281–287.

[214] D. Peloquin, M. DiMaio, B. Bierer, and M. Barnes, "Disruptive and avoidable: GDPR challenges to secondary research uses of data," *Eur. J. Human Genet.*, vol. 28, pp. 697–705, Mar. 2020.

[215] M. de Reuver, H. Bouwman, N. Heerschap, and H. Verkasalo, "Smartphone measurement: Do people use mobile applications as they say they do?" in *Proc. ICMB*, 2012, p. 2.

[216] N. Kiukkonen, J. Blom, O. Dousse, D. Gatica-Perez, and J. Laurila, "Towards rich mobile phone datasets: Lausanne data collection campaign," in *Proc. ICPS*, vol. 68, 2010, pp. 1–7.

[217] A. Rahmati, C. Tossell, C. Shepard, P. Kortum, and L. Zhong, "Exploring iPhone usage: The influence of socioeconomic differences on smartphone adoption, usage and usability," in *Proc. 14th Int. Conf. Human–Comput. Interact. Mobile Devices Services (MobileHCI)*, 2012, pp. 11–20.

[218] T. Yan, D. Chu, D. Ganesan, A. Kansal, and J. Liu, "Fast app launching for mobile devices using predictive user context," in *Proc. 10th Int. Conf. Mobile Syst. Appl. Services*, 2012, pp. 113–126.

[219] Y. Dong, N. V. Chawla, and A. Swami, "MetaPath2Vec: Scalable representation learning for heterogeneous networks," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2017, pp. 135–144.

[220] X. Wang *et al.*, "Heterogeneous graph attention network," in *Proc. World Wide Web Conf.*, 2019, pp. 2022–2032.

[221] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Gener. Comput. Syst.*, vol. 115, pp. 619–640, Feb. 2021.

[222] W. Y. B. Lim *et al.*, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.

[223] D. Xu *et al.*, "Edge intelligence: Empowering intelligence to the edge of network," *Proc. IEEE*, vol. 109, no. 11, pp. 1778–1837, Nov. 2021.

[224] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. Vincent Poor, "Federated learning for Internet of Things: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1622–1658, 3rd Quart., 2021.

[225] L. Yao, S. Li, Y. Li, M. Huai, J. Gao, and A. Zhang, "Representation learning for treatment effect estimation from observational data," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 2633–2643.

[226] B. Huang, K. Zhang, P. Xie, M. Gong, E. P. Xing, and C. Glymour, "Specific and shared causal relation modeling and mechanism-based clustering," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 13510–13521.

[227] G. Jaume *et al.*, "Quantifying explainers of graph neural networks in computational pathology," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 8106–8116.

[228] M. Zhang, T. Li, Y. Li, and P. Hui, "Multi-view joint graph representation learning for urban region embedding," in *Proc. 29th Int. Conf. Int. Joint Conf. Artif. Intell.*, 2021, pp. 4431–4437.

Tong Li (Member, IEEE) received the B.S. and M.S. degrees in communication engineering from Hunan University, China, in 2014 and 2017. He is currently pursuing the Ph.D. degree with the Hong Kong University of Science and Technology and University of Helsinki. His research interests include data mining and machine learning, especially with applications to mobile big data and urban computing.

Tong Xia received the B.S. degree in electrical engineering from the School of Electrical Information, Wuhan University, Wuhan, China, in 2017, and the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2020. She is currently pursuing the Ph.D. degree in machine learning with the Department of Computer Science, University of Cambridge, Cambridge, U.K. Her research interests include user behavior modeling, ubiquitous computing, and artificial intelligence for healthcare.

Huandong Wang (Member, IEEE) received the first B.S. degree in electronic engineering, the second B.S. degree in mathematical sciences, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2014, 2015, and 2019, respectively, where he is currently a Postdoctoral Research Fellow with the Department of Electronic Engineering. His research interests include urban computing, mobile big data mining, and machine learning.

Zhen Tu received the first B.S. degree in electronics and information engineering and the second B.S. degree in economics from Wuhan University, Wuhan, China, in 2016. She is currently pursuing the master's degree with the Electronic Engineering Department, Tsinghua University, Beijing, China. Her research interests include mobile big data mining, user behavior modeling, data privacy, and security.

Sasu Tarkoma (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in computer science from the Department of Computer Science, University of Helsinki, where he is currently the Dean of the Faculty of Science and a Professor of Computer Science. He is also affiliated with the Helsinki Institute for Information Technology HIIT. He has authored four textbooks and has published over 240 scientific articles. He has ten granted U.S. patents. His research has received a number of best paper awards and mentions, such as at IEEE PerCom, ACM CCR, and ACM OSR.

Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Idaho. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department as well as with the Computer Science Department, University of Houston, Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received the NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, IEEE Leonard G. Abraham Prize in the Field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018, an AAAS Fellow since 2019, and an ACM distinguished Member since 2019. He is a 1% highly cited researcher since 2017 according to Web of Science. He is also the Winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."

Pan Hui (Fellow, IEEE) received the bachelor's and M.Phil. degrees from the University of Hong Kong and the Ph.D. degree from the Computer Laboratory, University of Cambridge. He is a Professor of Computational Media and Arts with the Hong Kong University of Science and Technology and the Nokia Chair Professor of Data Science with the University of Helsinki. He was a Distinguished Scientist for Telekom Innovation Laboratories Germany and an Adjunct Professor of Social Computing and Networking with Aalto University. His industrial profile also includes his research with Intel Research Cambridge and Thomson Research Paris. He has published more than 400 research papers and with over 22,500 citations. He has 30 granted and filed European and U.S. patents in the areas of augmented reality, data science, and mobile computing. He has been serving on the organizing and technical program committee of numerous top international conferences, including ACM SIGCOMM, MobiCom, MobiSys, IEEE Infocom, ICNP, IJCAI, AAAI, ICWSM, and WWW. He served as an Associate Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING and IEEE TRANSACTIONS ON CLOUD COMPUTING. He is an International Fellow of the Royal Academy of Engineering, an ACM Distinguished Scientist, and a member of the Academia Europaea.