Search, Mining, and Their Applications on Mobile Devices: Introduction to the Special Issue

HONGNING WANG, Department of Computer Science, University of Virginia
RUI LI, Google Inc.
MILAD SHOKOUHI, Microsoft
HANG LI, Noah's Ark Lab of Huawei Technologies
YI CHANG, Huawei Research America

In recent years, mobile devices have become the most popular interface for users to retrieve and access information: recent reports show that users spend significantly more time and issue more search queries on mobile devices than on desktops in the United States.[1] The accelerated growth of mobile usage brings unique opportunities to the information retrieval and data mining research communities.

Mobile devices capture rich contextual and personal signals that can be leveraged to accurately predict users' intent for serving more relevant content and can even proactively provide novel zero-query recommendations. Apple Siri, Google Now, and Microsoft Cortana are recent examples of such emerging systems. Furthermore, mobile devices constantly generate a huge amount of sensor footprints (e.g., GPS, motion sensors) and user activity data (e.g., used apps) that are often missing from their desktop counterparts. These new sources of implicit and explicit user feedback are valuable for discovering actionable knowledge, and designing better systems that serve each individual the right content at the right time and location. In addition, by aggregating mobile interactions across individuals, one can infer interesting conclusions beyond search and recommendation. Generating real-time traffic estimates is one example of such applications.

This special issue focuses on research problems of search, mining, and their applications in mobile devices. Topics of interest in this special issue include but are not limited to mobile data mining and management, mobile search, personalization and recommendation, mobile user interfaces and human-computer interaction, and new applications in the mobile environment. The aim of this special issue is to bring together top experts across multiple disciplines, including information retrieval, data mining, mobile computing, and cyberphysical systems, such that academic and industrial researchers can exchange ideas and share the latest developments on the state of the art and practice of mobile search and mobile data mining.

CCS Concepts: • **Information systems → Mobile information processing systems;** *Data mining; Information retrieval;*

Additional Key Words and Phrases: Personalization, recommendation, mobile user interfaces, new applications in mobile environment

[1]https://techcrunch.com/2017/04/03/statcounter-android-windows/.

This work was partially supported by the National Science Foundation under grant IIS-1553568.
Authors' addresses: H. Wang, Department of Computer Science, University of Virginia, 85 Engineer's Way, Charlottesville, VA, 22904, USA; email: hw5x@virginia.edu ; R. Li, Google Inc., 1600 Amphitheatre Pkwy, Mountview, CA, 94043, USA; email: ruili@google.com; M. Shokouhi, Microsoft, Cambridge, UK; email: milads@microsoft.com; H. Li, Noah's Ark Lab of Huawei Technologies, China; email: hangli.hl@huawei.com; Y. Chang, Huawei Research America, Search Technology Labs, Sunnyvale, CA, 22903, USA; email: yichang@acm.org.

DOI: http://dx.doi.org/10.1145/3086665

ACM Reference Format:
Hongning Wang, Rui Li, Milad Shokouhi, Hang Li, and Yi Chang. 2017. Search, mining, and their applications on mobile devices: Introduction to the special issue. ACM Trans. Inf. Syst. 35, 4, Article 29 (August 2017), 17 pages.
DOI: http://dx.doi.org/10.1145/3086665

1. MOBILE SEARCH

With the prevalence of mobile devices, users' search activities have shifted from desktop to mobile devices at an incredibly fast pace. In 2015, Google announced that more searches took place on mobile devices than on desktops in the United States [12]. As an introduction to the special issue, we provide a review of mobile search research over the past 10 years to show how mobile search technologies have evolved.

1.1. Understanding Mobile Search

Initial research studies on mobile search focus on understanding the differences between mobile search and desktop search [2, 25, 26]. Those studies mainly come from researchers from Google, Yahoo!, and Microsoft, who have witnessed the growth of mobile search traffic from the very beginning.

The first mobile search analysis was published by Google in 2006. Kamvar and Baluja [25] analyzed about 1 million randomly sampled Google mobile searches and report some important statistics about mobile search, such as query length, query categories, session length, and session duration. Table I gives a summary of key stats reported in that study. In particular, they found the following key differences between mobile and desktop searches:

- The diversity of queries in mobile search was far less than in desktop search.
- Top categories of mobile queries were quite different from those of desktop queries. Adult queries were particularly popular in mobile search.
- Although words/characters per query were fairly similar, the number of queries per session in mobile search was significantly lower than in desktop search,
- The click rate in mobile search was low, and users were less likely to explore mobile search results.

The authors attributed these differences to (1) the difficulty of typing query on small screens and (2) the limited access of search results in mobile devices at that time. They also believed that those differences might evolve as mobile devices become more accessible in terms of costs, input methods, and search results. In 2007, the same group of researchers conducted a follow-up study based on newly acquired search logs [26]. The study not only confirmed their previous findings but also validated their assumption that mobile search is evolving. For instance, mobile queries had become relatively more diverse, and there were more high-end devices reported. Due to the rise of high-end mobile devices (e.g., iPhone), they further compared search patterns across three platforms: computers, conventional mobile phones, and high-end phones in another work [27]. They found that search behavior on high-end phones resembles computer-based search behavior more than conventional mobile phones.

At almost the same time, researchers from Yahoo! started to focus on mobile search as well. Particularly, Baeza-Yates et al. [2] conducted a study based on Yahoo! Japan search logs. Although most of their findings were consistent with statistics reported in the work of Kamvar and Baluja [25], they observed some different characteristics of mobile search in Japan. For instance, the number of characters per query in mobile search was on average less than that in desktop search, which can be attributed to

Table I. Mobile Search Top-Level Statistics Reported in Kamvar and Baluja [25]

	Query Length	Top Query Categories	Session Length
Mobile	2.3 words	“Adult”, “Entertainment”	1.6 queries
Desktop	2.35 words	“Commerce”, “People, Places”	2.02 queries

difference between languages. Yi et al. [68] further conducted a study on a larger dataset consisting of 20 million queries from the United States, Europe, and Asia through Yahoo! mobile search. Their main findings can be summarized as follows: (1) personal entertainment is the most popular category in mobile search, (2) there exist meaningful variations of mobile search in different regions, and (3) mobile search patterns are evolving. Yi and Maghoul [67] further conducted a follow-up study to discover trends of mobile search in 2011, where they found that mobile search has drastically changed. In particular, (1) mobile queries had become much more frequent and diverse, and (2) user information needs had substantially shifted from personal entertainment to local intent queries.

The most recent analysis about mobile search is from Microsoft in 2013. Song et al. [57] performed a series of analysis about mobile search with recent search logs. They observed that some characteristics have been changed since 2007. For instance, mobile users in general issue longer queries than tablet and desktop users. The authors attributed such changes to the evolution of mobile input methods. In addition, they provided an analysis of mobile search patterns from new perspectives, such as time distribution, search locality, click patterns, and browse patterns.

1.2. Improving Mobile Search

As mobile search becomes more and more important, many efforts have been made to improve different aspects of users’ mobile search experiences.

1.2.1. Improving Querying Experience. As discussed previously, the query input method is the major limitation of mobile search [25, 68]. Text input in mobile devices is generally slow and clumsy, which results in many key differences observed between mobile and desktop search (e.g., the number of queries per session in mobile search is significantly lower than in desktop search). Thus, technologies that can assist users to formulate queries, such as query auto completion (QAC) and query suggestion (QS), become very important for mobile search. Although QAC and QS have been widely studied in desktop search [3, 64], only a few relatively novel approaches have been proposed specifically for mobile search.

One direction is to utilize mobile-related signals with traditional algorithms. For instance, Zhang et al. [70] explore exclusive signals from mobile devices, such as installed apps and recently opened apps, to improve QAC. Specifically, they demonstrate the correlation between search queries and app-related signals by joining search logs and app usage logs, and propose a model to learn those correlations. The learned model is used to rerank QAC candidates and achieves significant improvement as measured by offline metrics. Zhao et al. [71] explore location signals available on mobile devices to improve QS. They formulate the problem as a tensor function learning problem with side information by exploring correlations among user, query, and location.

Vargas et al. [62] argue that classical QAC mechanisms, which aim for whole query completion, will give suboptimal results in mobile devices due to the smaller screen size and the clumsier input method of mobile devices. To overcome these limitations, they propose the idea of term-by-term QAC for mobile search, which suggests one term at a time. Technically, they design a structure called the *query-term graph*, which allows efficient storage and retrieval of query completions based on query logs, and a user model, which retrieves and ranks terms based on different measurements.

In addition to traditional text input methods, some researchers have started to focus on voice searches [20, 24]. Guy [20] compares the spoken queries to the typed-in queries in Yahoo search logs and presents semantic and syntactic characteristics of the two types of queries. For example, voice queries focus more on question answering and less on adult domains; voice queries are closer to natural language than typed queries. Jiang et al. [24] analyze typical voice input errors and users' corresponding reformulation strategies, and utilize the reformulations to further improve the performance of voice searches.

1.2.2. Improving Search Results. Improving relevance is always the key problem for search systems. To further improve mobile search relevance, researchers explore different new signals available on mobile devices. For example, Song et al. [57] design a set of mobile specific features, such as CTR of a query on mobile devices, assess how likely the URL gets clicked at location for mobile ranking, and propose a novel knowledge transfer framework to train a mobile-specific relevance model by incorporating the training labels from desktop data. Guo et al. [19] propose to leverage touch interactions such as pinching and swiping on mobile devices. Specifically, they first evaluate the quality of various touch interactions as implicit relevance feedback based on a user study and then utilize touch interaction data (e.g., dwell time, zooming, swiping) as features to train ranking models.

The search industry also puts a huge amount of effort into improving accessibility of search results, as it is another major limitation for mobile search discussed in the previous studies [25, 68]. Commercial search engines begin to use mobile-friendliness (e.g., how fast a result can be load and whether it can be correctly rendered in mobile browsers) as a factor to rank search results [66]. Further, they have developed many features to directly serve users' information needs in search results without requiring any additional clicks. For example, users can now directly get weather, stock quotes, local information, images, and flight information in search results.

Recently, there has been emerging interests in cross-device search [21, 40], where searches are conducted on multiple platforms such as desktop computers and mobile devices. Researchers aim to understand how these devices are used and how people transit between them during information seeking. For example, Montanez et al. [40] study characteristics of multidevice search, including aspects of search behavior on each device and characteristics of device transitions, and learn models to predict the next device used for search. Han et al. [21] study the problem of how to use interactions on mobile devices to improve subsequent search queries on desktop computers.

1.2.3. Vertical Mobile Search. Local search is the most important vertical of mobile search [25, 57, 68]. For example, Song et al. [57] find that 38% of mobile information needs are locally related. To fully understand mobile local search, Teevan et al. [60] conducted a comprehensive user study. They found that in addition to geographic features, the temporal aspects and the social context of users are also important to mobile local search. West et al. [63] explore how mobile queries, users' locations, and the context of those locations are related based on real-world search logs. They further propose a statistical model to predict whether a user is soon observed at the searched location based on the discovered correlations. Lv et al. [38] propose new ranking signals extracted from query logs to improve ranking models of local mobile search.

Mobile app search, which retrieves mobile apps for a given query, has become an important vertical in mobile search. However, as app names or their description are often too short, traditional search algorithms can hardly use this type of information. Park et al. [43] propose a solution to this novel problem based on leveraging users' reviews data. Specifically, they design a probabilistic topic model to jointly model user reviews and unstructured product information, and use the learned topics to enrich the representation of apps. More importantly, they provide the first quantitative evaluation of this problem with a real-world dataset crawled from Google Play. The dataset is open for the community to use. Park et al. [42] further extend their approach by leverage users status and other information in social media. We note that mobile app recommendation also becomes an important problem in the community [37, 69]. For instance, Liu et al. [37] introduce a new and important perspective—privacy—into app recommendations on mobile devices.

1.3. Trends in Mobile Search

Next we discuss some new research problems in the context of mobile search systems to reveal the "next generation" of mobile search systems. In general, we observe that mobile search systems will become more and more intelligent, which can provide direct answers instead of links for users information needs and can predict users' information needs with context.

1.3.1. Good Abandonment. Li et al. [34] present the seminal study about "good abandonment" in desktop and mobile search. They define a *good abandonment* as an abandoned query for which the information need is successfully addressed by the result page and present three key findings for good abandonment based on manually labeled search logs:

- First, a large percentage (around 19% to 50%) of abandoned queries are good abandonment.
- The good abandonment rate from mobile search is significantly higher than that from desktop search.
- Several types of information needs are the major sources of good abandonment. For example, for mobile search in the U.S. market, they are Local (42.9%), Answer (22.3%), and Stock (11.9%); for mobile search in the Japan market, they are Local (24.8%), Answer (20.1%), and Celebrities (17.8%).

This research highlights the existence of good abandonment in mobile search and points out that simply taking abandoned queries as negative signals for modeling and evaluation is incorrect.

To differentiate good and bad abandonment, Lagun et al. [30] propose a new concept "viewport," which is based on tracking the visible portion of result pages on mobile phones. In particular, they identify that increased scrolling past an answer and increased time below an answer as clear, measurable signals of user dissatisfaction with answers. Williams et al. [65] further study this problem in a comprehensive setting. They analyze the correlation between good abandonment and several types of features, including viewport features (e.g., visible region), gestures features (e.g., number of swipes, distance between swipes), search result features (e.g., the existence of images), and query features. They build a machine learning model to automatically identify good abandonment in mobile search with an accuracy of 75%. The identified that good abandonment can be further used in many search systems, such as learning relevance models and evaluating the quality of search results [1].

1.3.2. Proactive Search. In addition to providing users with direct answers in search result pages, search engines have evolved to *proactive* systems (e.g., Google Now, Apple Siri, and Microsoft Cortana), which directly show relevant content to users based on the context without queries. Some initial studies have been conducted for proactive search systems.

The first study of proactive search systems appears in the work of Shokouhi and Guo [54]. The authors analyze how users interact with information cards in proactive systems and propose a new model for ranking proactive cards. Specifically, they present two key observations: (1) usage patterns of proactive information cards strongly depend on time and location, and (2) the topics that a user interacts with are consistent across platforms. Based on those observations, they designed a supervised learning framework for reranking proactive cards based on the user's context and past history. As proactive cards do not need users to click, the authors also leverage the features mentioned in the previous section, such as viewport duration and swipes, to infer pseudorelevance labels for cards.

To enrich use cases of proactive search systems, Song and Guo [56] study the problem of how to identify repetitive tasks from search logs and predict when users will perform those tasks in the future. In particular, they leverage a structured learning framework and an information-theoretic model to identify repetitive tasks from search logs, and propose a deep neural network to classify how likely s user will perform such a task at a certain time based on a set of novel features, which are designed for characterizing task repetition. Sun et al. [59] model the same prediction problem as monitoring users' real-time intent with spatiotemporal context and propose a collaborative nowcasting model for the problem. As we will discuss in Section 3, the various contributions to this special issue add to our knowledge and understanding of mobile search.

2. MOBILE DATA MINING

As in mobile search, mobile data mining has also witnessed significant growth in research activity in recent years. With the unprecedented fast development of mobile devices and applications designed for such devices (also known as apps), immense amounts of data about mobile device users (e.g., call record data, GPS location data, app usage data, and social network data) have become accessible. The availability of such data introduces various types of interesting new data mining problems (e.g., modeling crowd mobility patterns in urban areas [14, 46], mobile crowdsourcing [7], and predicting mobile app usage [28]), and has drawn great attention among researchers within data mining communities.

In this section, to situate the contributions on mobile data mining included in this special issue, we review research on mobile data mining and highlight the latest trends.

2.1. Mining Human Mobility Patterns

In 2013, the number of mobile phone subscriptions reached 6.8 billion, corresponding to a global penetration of 96% [50]. The pervasiveness of mobile phones is spreading fast, with the number of subscriptions reaching 7.3 billion by 2014, from a report by the International Telecommunications Union (ITU) at the 2013 Mobile World Congress [14]. The deep penetration of mobile phones provides a unique opportunity for city planning, crowd management, and emergency response. With real-time and historical data analytics, anonymized mobile geolocation data has the potential for providing a new powerful social microscope, which may help us better understand human mobility and discover the hidden principles that characterize the trajectories defining human movement patterns.

Dong et al. [14] analyze call detail record (CDR) data to detect and monitor unusual events, which involve a large amount of people who expose an unusual mobility behavior. Mining CDR data has its own unique challenges, as CDR data only records the user location when a call or text message is made or received, and it is temporally and spatially sparse since call or message frequency of users is usually low and unpredictable. The authors define a new concept called the *cylindrical cluster* to capture sparse spatiotemporal CDR data across individual users and extract crowd events based on the volume of such clusters. A visual analytics system to support the exploration of the unusual crowd events is developed based on the proposed solution. The system allows end users, such as analysts and city managers, to analyze the formation and evolution of crowds, and study the impact of different parameters.

Candia et al. [6] study the individual and collective human dynamics from mobile phone records. They partition the space using a regular grid and measure the aggregated call activity inside each square bin forming the grid. By considering a series of consecutive time slices, they investigate the rise, clustering, and decay of spatially extended anomalous events, which could be relevant, for example, in real-time detection of emergency situations.

Poonawala et al. [46] develop a joint telco-and-farecard–based learning approach to understand urban mobility patterns. The proposed approach addresses the sparsity issue in telecommunications data by leveraging it jointly with real-time farecard data, which provides highly accurate quantitative knowledge on station-to-station travel times and volumes. Anonymized geolocation data from `StarHub.com` mobile and the Singapore Land Transport Authority (LTA) farecard are analyzed in IBM's *City in Motion* system. They identified that travel patterns improve the prediction of the initial (first mile) and final (last mile) segment of user trajectories before and after completing a probable train journey, which facilities the design of on-demand public transit feeders and main public transit lines.

A collaborative filtering technique [5, 51] has been leveraged to perform citywide human mobility completion from sparse call record data [17]. Specifically, a topic model–based spatiotemporal collaborative filtering solution is developed to simultaneously infer the topic distribution over users, time of days, days, and locations. The underlying assumption in this model is that users not only tend to repeat their own historical spatiotemporal visitation patterns but also share such patterns with users of similar historical movement trajectories. To predict the missing spatiotemporal data, the model uses the topic distributions to estimate a posterior over locations and infer the optimal location sequence in a hidden Markov model (HMM) considering the spatiotemporal continuity. However, map information is not used in this work. The road network, point-of-interest (POI) distribution, and land use can enforce a strong prior on the user visitation patterns. In Song et al. [55], such information (i.e., the transportation network data) is explicitly exploited to simulate and predict human mobility and the transportation mode at a citywide level. A multitask deep learning architecture based on a long short-term memory (LSTM) neural network structure [22, 32] is developed to jointly capture the long temporal dependency in human mobility and transportation patterns. Improved accuracy was reported in predicting an individual person's possible movements and transportation mode.

Mobile data has also been analyzed for anomaly detection and crime prevention. Qin et al. [47] take a generative approach to combine group anomaly detection and a topic segmentation method for segmenting a mobile subscriber's sequence of behaviors into explainable routines, and they report those unexplainable segments as anomalies simultaneously. Lin et al. [36] compute the statistical ambiance features on WiFi and cell tower data from location anonymized datasets and customize an HMM to capture the spatiotemporal patterns of each user's mobility behaviors. Their model is then applied to predict if a particular user is the owner of a given mobile device. Bogomolov et al. [4] use human behavioral data derived from anonymized and aggregated mobile network activity, combined with demographics, to predict crime hotspots. Traunmueller et al. [61] analyze footfall counts as recorded by mobile data and extract metrics that act as proxies of urban crime theories. Using correlation analysis between such proxies and crime activity derived from open crime data records, they reveal to what extent different theories of urban crime hold and where.

In addition to mining mobile data for human mobility and transportation mode modeling in the scenario of urban planning, crime prevention, and emergency response, mobile data has also been used in detecting spatial regions where the population is exposed to a significantly higher risk of disease than expected. Lan et al. [31] cluster diseases in georegions by extending Kulldorff's spatial scan statistic to mobility data. Their underlying assumption is that the probability that an individual becomes sick is a logistic function of a weighted sum of the disease risks at the visited locations. A log-likelihood ratio test is performed to measure if a given subregion has a significantly higher disease risk than expected.

2.2. Mining Mobile Phone Usage Patterns

Mobile devices record huge amounts of user behavioral data, particularly users' daily communications with others. This provides us with an unprecedented opportunity to study how people behave differently and form different groups. Dong et al. [15] analyze a real-world large mobile network of more than 7 million users and more than 1 billion communication records (CALL and SMS). Several interesting social strategies that mobile users frequently use to maintain their social connections were identified. First, young people are very active in broadening their social circles, whereas seniors tend to keep close but more stable connections. Second, female users put more attention on cross-generation interactions than male users, although interactions between male and female users are frequent. Third, a persistent same-gender triadic pattern over one's lifetime has been discovered for the first time, whereas more complex opposite-gender triadic patterns have been only exhibited among young people. We further study the extent of users' demographics. Based on such findings, it is also reported that users' demographics such as gender and age can be inferred from their mobile communications.

Success in the use of mobile phone data as a proxy for social interaction has been proved in several recent investigations. Eagle et al. [16] compares observational data from mobile phones with standard self-report survey data and finds that the information from these two data sources is overlapping but distinct. It demonstrates that it is possible to accurately infer friendships based on the observational data alone, where friend *dyads* demonstrate distinctive temporal and spatial patterns in their physical proximity and calling patterns. Phithakkitnukoon et al. [45], based on nearly 1 million mobile phone records of users in the central Metro-Boston area, report a strong correlation in daily activity patterns within the group of people who share a common work area profile. In addition, within the group itself, the similarity in activity patterns decreases as the users' work places become apart.

LiKamWa et al. [35] demonstrate the feasibility of inferring mood from smartphone usage, paving the way for energy-efficient, privacy-preserving systems that automatically infer user mood. Signals from mobile phone users' social interaction records (e.g., phone calls, text messages, and emails), and routine activity records (e.g., browser history, application usage, and location history) are extracted to fit a multilinear regression model. Important features for mood prediction are identified via regression analysis.

2.3. Mining Mobile App Usage Patterns

The prevalence of mobile apps has generated a huge volume of app usage data. Understanding user behaviors of using mobile apps is meaningful to all stakeholders in the ecosystem, including app developers, mobile phone users, marketplace operators, and device manufacturers.

Studies show evidence about different users having different patterns of using mobile apps [48, 53]. Do and Gatica-Perez [13] apply the author topic model [49] on the mobile app usage data by treating the app phone log data as a bag of apps. Apps of similar usage patterns are clustered into the same topic, and users who use those apps in similar patterns will have similar topic distribution. As a result, both the apps and users are clustered. Srinivasan et al. [58] consider mobile app usage pattern mining into its actual usage context. Association rule mining is performed to identify a user's longitudinal, multimodal context behavior patterns. For example, users tend to use news app to read news in the morning when they are at home. As the pattern mining algorithm is executed on the mobile device, special attention is given to energy efficiency and computational complexity. The identified patterns are shown to be useful in predicting the next contact called or the next app launched by the user. Chittaranjan et al. [11] conducted a similar analysis to understand the dependency between the use of mobile applications and the location and social context.

Li et al. [33] present a comprehensive analysis of app usage behaviors collected from millions of Android users. They collected mobile app usage data covering over 0.2 million Android apps from Wandoujia,[2] a leading Android app marketplace in China. Their analysis mainly focuses on two types of user behaviors: (1) app management activities (i.e., install, uninstall, update) and (2) app network usage. Several interesting findings are reported in this work. For example, by analyzing the app download history, user subscription, and network traffic, the authors find that the popularity distribution of apps typically follows the "Pareto" effect. A large number of apps have a small number of downloads and contribute marginally to network usage. In addition, they also found that some apps are frequently installed together, such as those in the "communication" and "social" categories. Apps that come from the same vendor or have similar functionality are more likely to be installed together. Based on the qualitative findings, the authors give suggestions to different stakeholders in the mobile app ecosystem. For example, the marketplace operators can improve recommendation systems, such as by clustering the frequently coinstalled apps at a close location to allow end users to download them quickly.

Mobile app usage mining also contributes to personality analysis. Chittaranjan et al. [11] extract behavior features from five different modalities, such as SMS logs, call logs, and app logs, to perform correlation analysis between those features and the self-perceived personality of the participants. The study finds that that the office app was more likely to be used by conscientious participants. The Internet was found to be more likely used by introverts and those who are disagreeable. Disagreeable and conscientious users were more likely to use the Mail application, whereas extroverts and nonconscientious participants were less likely to use the YouTube app. The Video/Audio/Music apps were more likely to be used by users who score higher on openness and low on conscientiousness. Interestingly, it was found that the frequency of usage of SMS was positively correlated to emotional stability among females and negatively among males.

2.4. Mining Mobile App Marketplace Data

The proliferation of smartphones is driving the rapid growth of mobile app marketplaces. In 2013, Google Play Store, the official and the largest Android app repository, had more than 700,000 mobile apps, mostly developed by third-party companies, organizations, and individual developers.[3] Mining text, image, and user rating data available in the app marketplace provide a new venue for knowledge discovery and decision optimization.

[2]https://www.wandoujia.com/.
[3]http://mashable.com/2013/07/24/google-play-1-million#QmCUYIekBiqw.

Fu et al. [18] look into user-generated app reviews to discover inconsistencies in reviews, identify reasons why users like or dislike a given app, and provide an interactive and zoomable view of how users' reviews evolve over time. They provide valuable insights into the entire app market, identifying users' major concerns and preferences for different types of apps. To facilitate mobile app developers to discover the most "informative" user reviews from a large and rapidly increasing pool of user reviews, Chen et al. [10] automatically group informative user reviews by first filtering noisy and irrelevant ones. The selected reviews are ranked and visualized before presenting to the app developers. Palomba et al. [41] trace informative crowd reviews onto source code changes and monitor the extent to which developers accommodate crowd requests and follow-up user reactions as reflected in their ratings.

Chen et al. [8] study the problem of inferring the semantic similarity between apps based on multiple modalities of their relatedness, such as their names, developer, category, description, and permission request. Later on, based on the learned similarity between different apps, the authors study the problem of app tagging, such as assigning short topical labels to a given mobile app [9]. In particular, the text tags are selected from similar apps via phase-based text summarization techniques.

By looking into a mobile app's functionality and textual description information, Kong and Jin [29] propose to predict the permission requests of a new mobile app, with a goal to help users be aware of the privacy risks of mobile apps. A multilabel classifier is built on structure feature learning, which incorporates both textual description and functionality information, and learns the feature projection to the permission request patterns.

Seneviratne et al. [52] track the removed apps from online marketplaces to identify spam apps. Features are extracted from a set of manually crafted rules to describe the spam apps, on which an adaptive boost classifier is built for early detection of spam apps. Hu et al. [23] build a multilabel classifier to predict the mature contents in a given app and then label the maturity level according to a rating policy. A set of keywords indicating the maturity content are first manually specified and then expanded via word embedding vectors [39, 44] learned on a large set of app descriptions. A multilabel text classifier is built based on the selected word features to predict the maturity level with regard to the app description.

3. ABOUT THIS SPECIAL ISSUE

For this special issue, we solicited articles reflecting the state of the art and emerging research trends in search, mining, and their applications on mobile devices. Manuscripts focusing on all areas of mobile search, mobile data mining, and new applications and their studies in mobile environment were encouraged. In particular, topics such as large-scale modeling of mobile social networks, multimodal mobile app recommendation, and search intent modeling in the mobile search environment emerge in this special issue. From 31 submissions, we selected 14 high-quality articles that represent current themes of research on mobile search, mobile data mining, and applications on mobile devices. In the following, we briefly summarize the accepted articles by classifying their topics into categories of mobile search and mobile data mining.

Mobile Search

The need for providing proactive service assistance in the mobile environment urges in-depth understanding of user intent. In the article "Collaborative Intent Nowcasting with Real-Time Dataflow," the authors appeal to the idea of "nowcasting" originated in meteorology and macroeconomics to predict users' real-time intent from context signal streams. A tensor decomposition based on the collaborative nowcasting model is developed to monitor the relatedness between users' latent intent and continuously arriving contextual signals. Due to the collaborative nature of the proposed solution, the computation can be distributed across users on their mobile devices. This reduces the communication overhead between the server and clients. A similar idea is explored in the article "A Time-Aware Personalized Point-of-Interest Recommendation via High-Order Tensor Factorization," where tensor decomposition was applied on users' historical movement trajectories to provide personalized POI recommendations. A fourth-order tensor factorization is performed to integrate the user's short-term influence, long-term influence, and time variant preference to effectively predict the category of the next location. As only the location category is predicted, the data sparsity issue is alleviated. Based on the predicted location category, location ranking is performed to provide the recommendations. In a related line of research, the article "Mining Exploratory Behavior to Improve Mobile App Recommendations" explores users' exploratory behavior to improve the quality of app recommendation. A Gaussian mixture model is imposed over the users' exploratory behavior sequences, such as downloading apps with similar functionality, and individual users' preference is modeled with a probabilistic topic model over the items.

From a different perspective of satisfying users' information needs based on contextual information monitored on a mobile device, the article "Curious Cat: Mobile, Context-Aware Conversational Crowdsourcing Knowledge Acquisition" studies the problem of crowdsourcing knowledge acquisition. The goal is to simultaneously satisfy users' immediate information needs while extending an existing knowledge base by user feedback. The developed system automatically constructs precisely targeted natural language crowdsourcing tasks ("question") for the right audience at the right moment based on the monitored context. To evaluate the utility of the proposed solution, the authors also built a proof-of-concept conversational assistant named *Curious Cat*, which provides users with interesting and useful information about the locations they visit while also supporting incidental conversation ranging over commonsense topics.

To assist users to easily clip articles from diverse mobile applications, the authors in the article "Search by Screenshots for Saving Articles in One Place from Different Mobile Apps" developed a unified solution named *UniClip*. UniClip allows a user to perform article search by screenshots. More specifically, it segments a screenshot into structural units called *blocks* and formulates keyword-based search queries by considering the role of each block. The learning-to-rank technique is used to aggregate the search results from multiple queries.

Mobile Data Mining

The availability of mobile phone data, such as mobile communication records and mobile trajectory data, enhances the study in social network analysis. The article "User Modeling on Demographic Attributes in Large-Scale Mobile Social Networks" analyzes a large collection of mobile communication records to infer the demographic profile of each user in the network. Based on some interesting findings in users' interaction patterns, a probabilistic factor graph model is developed to exploit the dependency between network features and users' gender and age attributes. In the article "A Neural Network Approach to Joint Modeling Social Networks and Mobile Trajectories," social network analysis is integrated with mobile trajectory modeling via a deep neural network model. Network embedding is first performed with regard to the observed network structure, and then it is used to model the users' sequential location visitation sequence. As a result, the learned embedding structure can be better calibrated to characterize each user in the network. In particular, four factors are explicitly considered in the modeled embedding, including user visit preference, influence of friends, short-term sequential contexts, and long-term sequential contexts.

Mining users' mobile phone usage data and mobile app review data provides researchers a unique opportunity to improve app recommendations. In the article "Cross-Platform App Recommendation by Jointly Modeling Ratings and Texts," the app review text content and ratings from multiple mobile device platforms (e.g., mobile phone and tablet) are jointly modeled to provide personalized app recommendations. A matrix factorization-based solution is developed to model users' cross-platform behaviors by sharing a common set of latent user factors. The article "Version-Aware Rating Prediction for Mobile App Recommendation" addresses the same app recommendation problem from a different perspective. To address the data sparsity issue, the authors exploit the user provided ratings in different versions of the same app in addition to the review content. Rating correlations between multiple versions within the same app and those between similar apps are leveraged to improve the recommendation quality. In the article "Derive User Preferences of Mobile Apps from Their Management Activities," the authors look into a very large app management log collected through a leading Android app marketplace in China. What is interesting in their qualitative study is that they find the metrics commonly used to rank apps in app stores, such as ratings and number of downloads, do not truly reflect the users' real attitudes. Based on such behavioral analysis, they identify several behavioral patterns from the app management activities that more accurately indicate user preferences of an app even when no explicit rating is available.

The emergence of large-scale human mobile data collected from mobile devices empowers research in urban planning, emergency management, and healthcare. In the article "Computing Urban Traffic Congestions by Incorporating Sparse GPS Probe Data and Social Media Data," the authors integrate the data from GPS probe and real-time events extracted from social media for traffic condition estimation. A traffic co-congestion pattern is extracted from the GPS probe data and diverse auxiliary information, including social events, road features, POI, and weather. The identified patterns are utilized to help estimate traffic congestions and detect anomalies in a transportation network. To fully exploit the multisourced data, a coupled matrix and tensor factorization model is developed to factorize the sparse traffic congestion matrix in a collaborative way. Promising congestion prediction accuracy is achieved via this joint factorization method. In the article "DeepMob: Learning Deep Knowledge of Human Emergency Behavior and Mobility from Big and Heterogeneous Data," a big and heterogeneous collection of data, including GPS records of 1.6 million users over 3 years, data on earthquakes occurring in Japan over 4 years, news report data, and transportation network data, is jointly modeled in a deep neural network–based system. The system aims at analyzing and predicting human evacuation behavior and mobility following different types of natural disasters. This potentially provides a better solution for disaster management. The article "Tweet to Be Fit: Integrating Wearable Sensors and Multiple Social Media for Wellness Profile Learning" utilizes mobile sensing data (e.g., altitude, geolocation, time, and heart rate) and social media data (e.g., text and images from Twitter and Instagram) to monitor personal wellness attributes. In particular, the body mass index (BMI) category and BMI trend are monitored through mobile sensing data and social media data. A multitask learning solution is developed to utilize the social network structure to improve prediction accuracy.

Protecting user privacy in mobile devices is receiving increasing attention in both data mining and security communities. The article "Understanding the Purpose of Permission Use in Mobile Apps" looks into app implementation code to extract useful features for inferring the purpose of required permissions. This enables early detection of a malicious permission request. Both static and dynamic analysis is performed for feature extraction. For example, the authors looked into compile Java class files for informative text descriptions by decompiling the apps. A manually created taxonomy is used to define two target purposes for analysis, namely the location tagging purpose and the contact list checking purpose. Complementary to previous solutions that use text content in app description for the inference purpose, this study provides a new perspective of understanding app permission requests.

4. CONCLUSION AND REMARKS FOR FUTURE DIRECTIONS

The bloom of mobile device and mobile applications stimulates research in mobile search and mobile data mining. We believe that we have collected a strong set of interesting works on these increasingly important topics. Each of the accepted papers is likely to spawn off follow-up research in the related area. We look forward to seeing how the field continues its rapid development as the articles are picked up by the community. Based on the work presented in this special issue, we believe that computational user intent and behavior modeling in the mobile environment, collaborative sensing and learning, and privacy and secure computation on mobile devices are particularly important directions that deserve considerable attention from both academic and industrial researchers.

ACKNOWLEDGMENTS

We would like to sincerely thank our reviewers for their insightful comments and contribution in this special issue. And we also want to thank the Editorial Board of ACM TOIS for their invaluable support and help.

REFERENCES

[1] Olga Arkhipova and Lidia Grauer. 2014. Evaluating mobile Web search performance by taking good abandonment into account. In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'14)*. ACM, New York, NY, 1043–1046. DOI:http://dx.doi.org/10.1145/2600428.2609505

[2] Ricardo Baeza-Yates, Georges Dupret, and Javier Velasco. 2007. A study of mobile search queries in Japan. In *Proceedings of the 16th International Conference on World Wide Web (WWW'07)*. ACM, New York, NY.

[3] Ziv Bar-Yossef and Naama Kraus. 2011. Context-sensitive query auto-completion. In *Proceedings of the 20th International Conference on World Wide Web (WWW'11)*. ACM, New York, NY, 107–116. DOI:http://dx.doi.org/10.1145/1963405.1963424

[4] Andrey Bogomolov, Bruno Lepri, Jacopo Staiano, Nuria Oliver, Fabio Pianesi, and Alex Pentland. 2014. Once upon a crime: Towards crime prediction from demographics and mobile data. In *Proceedings of the 16th International Conference on Multimodal Interaction*. ACM, New York, NY, 427–434.

[5] John S. Breese, David Heckerman, and Carl Kadie. 1998. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*. 43–52.

[6] Julián Candia, Marta C. González, Pu Wang, Timothy Schoenharl, Greg Madey, and Albert-László Barabási. 2008. Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical* 41, 22, 224015.

[7] Chen Cen, Shih-Fen Cheng, Hoong Chuin Lau, and Archan Misra. 2015. Towards city-scale mobile crowdsourcing: Task recommendations under trajectory uncertainties. In *Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI'15)*. 1113–1119.

[8] Ning Chen, Steven C. H. Hoi, Shaohua Li, and Xiaokui Xiao. 2015. SimApp: A framework for detecting similar mobile applications by online kernel learning. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, 305–314.

[9] Ning Chen, Steven C. H. Hoi, Shaohua Li, and Xiaokui Xiao. 2016. Mobile app tagging. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, 63–72.

[10] Ning Chen, Jialiu Lin, Steven C. H. Hoi, Xiaokui Xiao, and Boshen Zhang. 2014. AR-Miner: Mining informative reviews for developers from mobile app marketplace. In *Proceedings of the 36th International Conference on Software Engineering*. ACM, New York, NY, 767–778.

[11] Gokul Chittaranjan, Jan Blom, and Daniel Gatica-Perez. 2013. Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing* 17, 3, 433–450.

[12] Jerry Dischler. 2015. Building for the Next Moment. Retrieved May 15, 2017, from https://adwords.googleblog.com/2015/05/building-for-next-moment.html.

[13] Trinh-Minh-Tri Do and Daniel Gatica-Perez. 2010. By their apps you shall understand them: Mining large-scale patterns of mobile phone usage. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. ACM, New York, NY, 27.

[14] Yuxiao Dong, Fabio Pinelli, Yiannis Gkoufas, Zubair Nabi, Francesco Calabrese, and Nitesh V. Chawla. 2015. Inferring unusual crowd events from mobile phone call detail records. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. 474–492.

[15] Yuxiao Dong, Yang Yang, Jie Tang, Yang Yang, and Nitesh V. Chawla. 2014. Inferring user demographics and social strategies in mobile social networks. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 15–24.

[16] Nathan Eagle, Alex Sandy Pentland, and David Lazer. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences* 106, 36, 15274–15278.

[17] Zipei Fan, Ayumi Arai, Xuan Song, Apichon Witayangkurn, Hiroshi Kanasugi, and Ryosuke Shibasaki. A collaborative filtering approach to citywide human mobility completion from sparse call records. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*. 2500–2506.

[18] Bin Fu, Jialiu Lin, Lei Li, Christos Faloutsos, Jason Hong, and Norman Sadeh. 2013. Why people hate your app: Making sense of user feedback in a mobile app store. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 1276–1284.

[19] Qi Guo, Haojian Jin, Dmitry Lagun, Shuai Yuan, and Eugene Agichtein. 2013. Mining touch interaction data on mobile devices to predict Web search result relevance. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*. ACM, New York, NY, 153–162. DOI: http://dx.doi.org/10.1145/2484028.2484100

[20] Ido Guy. 2016. Searching by talking: Analysis of voice queries on mobile Web search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'16)*. ACM, New York, NY, 35–44. DOI: http://dx.doi.org/10.1145/2911451.2911525

[21] Shuguang Han, Zhen Yue, and Daqing He. 2015. Understanding and supporting cross-device Web search for exploratory tasks with mobile touch interactions. *ACM Transactions on Information Systems* 33, 4, Article No. 16. DOI: http://dx.doi.org/10.1145/2738036

[22] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9, 8, 1735–1780.

[23] Bing Hu, Bin Liu, Neil Zhenqiang Gong, Deguang Kong, and Hongxia Jin. 2015. Protecting your children from inappropriate content in mobile apps: An automatic maturity rating framework. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. ACM, New York, NY, 1111–1120.

[24] Jiepu Jiang, Wei Jeng, and Daqing He. 2013. How do users respond to voice input errors? Lexical and phonetic query reformulation in voice search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*. ACM, New York, NY, 143–152. DOI: http://dx.doi.org/10.1145/2484028.2484092

[25] Maryam Kamvar and Shumeet Baluja. 2006. A large scale study of wireless search behavior: Google mobile search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 701–709.

[26] Maryam Kamvar and Shumeet Baluja. 2007. Deciphering trends in mobile search. *Computer* 40, 8, 58–62. DOI: http://dx.doi.org/10.1109/MC.2007.270

[27] Maryam Kamvar, Melanie Kellar, Rajan Patel, and Ya Xu. 2009. Computers and iPhones and mobile phones, oh my! A logs-based comparison of search users on different devices. In *Proceedings of the 18th International Conference on World Wide Web (WWW'09)*. ACM, New York, NY, 801–810. DOI: http://dx.doi.org/10.1145/1526709.1526817

[28] Jingu Kim and Taneli Mielikäinen. 2014. Conditional log-linear models for mobile application usage prediction. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. 672–687.

[29] Deguang Kong and Hongxia Jin. 2015. Towards permission request prediction on mobile apps via structure feature learning. In *Proceedings of the 2015 SIAM International Conference on Data Mining*. 604–612.

[30] Dmitry Lagun, Chih-Hung Hsieh, Dale Webster, and Vidhya Navalpakkam. 2014. Towards better measurement of attention and satisfaction in mobile search. In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'14)*. ACM, New York, NY, 113–122. DOI: http://dx.doi.org/10.1145/2600428.2609631

[31] Liang Lan, Vuk Malbasa, and Slobodan Vucetic. 2014. Spatial scan for disease mapping on a mobile population. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14)*. 431–437.

[32] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature* 521, 7553, 436–444.

[33] Huoran Li, Xuan Lu, Xuanzhe Liu, Tao Xie, Kaigui Bian, Felix Xiaozhu Lin, Qiaozhu Mei, and Feng Feng. 2015. Characterizing smartphone usage patterns from millions of Android users. In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference*. ACM, New York, NY, 459–472.

[34] Jane Li, Scott Huffman, and Akihito Tokuda. 2009. Good abandonment in mobile and PC internet search. In *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'09)*. ACM, New York, NY, 43–50. DOI:http://dx.doi.org/10.1145/1571941.1571951

[35] Robert LiKamWa, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. 2013. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, New York, NY, 389–402.

[36] Miao Lin, Hong Cao, Vincent W. Zheng, Kevin Chen-Chuan Chang, and Shonali Krishnaswamy. 2015. Mobility profiling for user verification with anonymized location data. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*. 960–966.

[37] Bin Liu, Deguang Kong, Lei Cen, Neil Zhenqiang Gong, Hongxia Jin, and Hui Xiong. 2015. Personalized mobile app recommendation: Reconciling app functionality and user privacy preference. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining (WSDM'15)*. ACM, New York, NY, 315–324. DOI:http://dx.doi.org/10.1145/2684822.2685322

[38] Yuanhua Lv, Dimitrios Lymberopoulos, and Qiang Wu. 2012. An exploration of ranking heuristics in mobile local search. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'12)*. ACM, New York, NY, 295–304. DOI:http://dx.doi.org/10.1145/2348283.2348325

[39] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*. 3111–3119.

[40] George D. Montanez, Ryen W. White, and Xiao Huang. 2014. Cross-device search. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM'14)*. ACM, New York, NY, 1669–1678. DOI:http://dx.doi.org/10.1145/2661829.2661910

[41] Fabio Palomba, Mario Linares-Vásquez, Gabriele Bavota, Rocco Oliveto, Massimiliano Di Penta, Denys Poshyvanyk, and Andrea De Lucia. 2015. User reviews matter! Tracking crowdsourced reviews to support evolution of successful apps. In *Proceedings of the 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME'15)*. IEEE, Los Alamitos, CA, 291–300.

[42] Dae Hoon Park, Yi Fang, Mengwen Liu, and Cheng-Xiang Zhai. 2016. Mobile app retrieval for social media users via inference of implicit intent in social media text. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (CIKM'16)*. ACM, New York, NY, 959–968. DOI:http://dx.doi.org/10.1145/2983323.2983843

[43] Dae Hoon Park, Mengwen Liu, ChengXiang Zhai, and Haohong Wang. 2015. Leveraging user reviews to improve accuracy for mobile app retrieval. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, New York, NY, 533–542.

[44] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods on Natural Language Processing (EMNLP'14)*, Vol. 14. 1532–1543.

[45] Santi Phithakkitnukoon, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, and Carlo Ratti. 2010. Activity-aware map: Identifying human daily activity pattern using mobile phone data. In *Proceedings of the International Workshop on Human Behavior Understanding*. 14–25.

[46] Hasan Poonawala, Vinay Kolar, Sebastien Blandin, Laura Wynter, and Sambit Sahu. 2016. Singapore in motion: Insights on public transport service level through farecard and mobile data analytics. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 589–598.

[47] Tian Qin, Guojie Song, and Sizhen Du. 2016. Routine mining based anomaly detection in mobile phone data. In *Proceedings of the 2016 SIAM International Conference on Data Mining*. 180–188.

[48] Ahmad Rahmati and Lin Zhong. 2013. Studying smartphone usage: Lessons from a four-month field study. *IEEE Transactions on Mobile Computing* 12, 7, 1417–1427.

[49] Michal Rosen-Zvi, Thomas Griffiths, Mark Steyvers, and Padhraic Smyth. 2004. The author-topic model for authors and documents. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*. 487–494.

[50] Brahima Sanou. 2013. The world in 2013: ICT facts and figures. *International Telecommunications Union*. Retrieved May 15, 2017, from https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf.

[51] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web (WWW'01)*. ACM, New York, NY, 285–295.

[52] Suranga Seneviratne, Aruna Seneviratne, Mohamed Ali Kaafar, Anirban Mahanti, and Prasant Mohapatra. 2015. Early detection of spam mobile apps. In *Proceedings of the 24th International Conference on World Wide Web (WWW'15)*. ACM, New York, NY, 949–959.

[53] Choonsung Shin, Jin-Hyuk Hong, and Anind K. Dey. 2012. Understanding and prediction of mobile application usage for smart phones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, New York, NY, 173–182.

[54] Milad Shokouhi and Qi Guo. 2015. From queries to cards: Re-ranking proactive card recommendations based on reactive search history. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'15)*. ACM, New York, NY, 695–704. DOI:http://dx.doi.org/10.1145/2766462.2767705

[55] Xuan Song, Hiroshi Kanasugi, and Ryosuke Shibasaki. 2016. Deeptransport: Prediction and simulation of human mobility and transportation mode at a citywide level. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*. 2618–2624.

[56] Yang Song and Qi Guo. 2016. Query-Less: Predicting task repetition for nextgen proactive search and recommendation engines. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)*. 543–553. DOI:http://dx.doi.org/10.1145/2872427.2883020

[57] Yang Song, Hao Ma, Hongning Wang, and Kuansan Wang. 2013. Exploring and exploiting user search behavior on mobile and tablet devices to improve search relevance. In *Proceedings of the 22nd International Conference on World Wide Web (WWW'13)*. ACM, New York, NY, 1201–1212. DOI:http://dx.doi.org/10.1145/2488388.2488493

[58] Vijay Srinivasan, Saeed Moghaddam, Abhishek Mukherji, Kiran K. Rachuri, Chenren Xu, and Emmanuel Munguia Tapia. 2014. Mobileminer: Mining your frequent patterns on your phone. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, 389–400.

[59] Yu Sun, Nicholas Jing Yuan, Xing Xie, Kieran McDonald, and Rui Zhang. 2016. Collaborative nowcasting for contextual recommendation. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)*. 1407–1418. DOI:http://dx.doi.org/10.1145/2872427.2874812

[60] Jaime Teevan, Amy Karlson, Shahriyar Amini, A. J. Bernheim Brush, and John Krumm. 2011. Understanding the importance of location, time, and people in mobile local search behavior. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'11)*. ACM, New York, NY, 77–80. DOI:http://dx.doi.org/10.1145/2037373.2037386

[61] Martin Traunmueller, Giovanni Quattrone, and Licia Capra. 2014. Mining mobile phone data to investigate urban crime theories at scale. In *Proceedings of the International Conference on Social Informatics*. 396–411.

[62] Saúl Vargas, Roi Blanco, and Peter Mika. 2016. Term-by-term query auto-completion for mobile search. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining (WSDM'16)*. ACM, New York, NY, 143–152. DOI:http://dx.doi.org/10.1145/2835776.2835813

[63] Robert West, Ryen W. White, and Eric Horvitz. 2013. Here and there: Goals, activities, and predictions about location from geotagged queries. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*. ACM, New York, NY, 817–820. DOI:http://dx.doi.org/10.1145/2484028.2484125

[64] Stewart Whiting and Joemon M. Jose. 2014. Recent and robust query auto-completion. In *Proceedings of the 23rd International Conference on World Wide Web (WWW'14)*. ACM, New York, NY, 971–982. DOI:http://dx.doi.org/10.1145/2566486.2568009

[65] Kyle Williams, Julia Kiseleva, Aidan C. Crook, Imed Zitouni, Ahmed Hassan Awadallah, and Madian Khabsa. 2016. Detecting good abandonment in mobile search. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)*. 495–505. DOI:http://dx.doi.org/10.1145/2872427.2883074

[66] Rolfe Winkler. 2015. Google gives boost to mobile-friendly sites. *Wall Street Journal*. Retrieved May 15, 2017, from https://www.wsj.com/articles/google-gives-boost-to-mobile-friendly-sites-1429660022

[67] Jeonghe Yi and Farzin Maghoul. 2011. Mobile search pattern evolution: The trend and the impact of voice queries. In *Proceedings of the 20th International Conference Companion on World Wide Web (WWW'11)*. ACM, New York, NY, 165–166. DOI:http://dx.doi.org/10.1145/1963192.1963276

[68] Jeonghee Yi, Farzin Maghoul, and Jan Pedersen. 2008. Deciphering mobile search patterns: A study of Yahoo! mobile search queries. In *Proceedings of the 17th International Conference on World Wide Web (WWW'08)*. ACM, New York, NY, 257–266. DOI:http://dx.doi.org/10.1145/1367497.1367533

[69] Peifeng Yin, Ping Luo, Wang-Chien Lee, and Min Wang. 2013. App recommendation: A contest between satisfaction and temptation. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM'13)*. ACM, New York, NY, 395–404. DOI:http://dx.doi.org/10.1145/2433396.2433446

[70] Aston Zhang, Amit Goyal, Ricardo Baeza-Yates, Yi Chang, Jiawei Han, Carl A. Gunter, and Hongbo Deng. 2016. Towards mobile query auto-completion: An efficient mobile application-aware approach. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)*. 579–590. DOI:http://dx.doi.org/10.1145/2872427.2882977

[71] Zhou Zhao, Ruihua Song, Xing Xie, Xiaofei He, and Yueting Zhuang. 2015. Mobile query recommendation via tensor function learning. In *Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI'15)*. 4084–4090. http://dl.acm.org/citation.cfm?id=2832747.2832818

Received April 2007; accepted April 2017