# NextMe: Localization Using Cellular Traces in Internet of Things

Daqiang Zhang, *Senior Member, IEEE*, Shengjie Zhao, *Senior Member, IEEE*, Laurence T. Yang, Min Chen, Yunsheng Wang, and Huazhong Liu

***Abstract*—The Internet of Things (IoT) opens up tremendous opportunities to location-based industrial applications that leverage both Internet-resident resources and phones' processing power and sensors to provide location information. Location-based service is one of the vital applications in commercial, economic, and public domains. In this paper, we propose a novel localization scheme called *NextMe*, which is based on cellular phone traces. We find that the mobile call patterns are strongly correlated with the co-locate patterns. We extract such correlation as social interplay from cellular calls, and use it for location prediction from temporal and spatial perspectives. *NextMe* consists of data preprocessing, call pattern recognition, and a hybrid predictor. To design the call pattern recognition module, we introduce the notions of critical calls and corresponding patterns. In addition, *NextMe* does not require that the cell tower addresses should be bounded with concrete coordinates, e.g., *global positioning system (GPS) coordinates*. We validate *NextMe* across MIT Reality Mining Dataset, involving 500 000 h of continuous behavior information and 112 508 cellular calls. Experimental results show that *NextMe* achieves fine-grained prediction accuracy at cell tower level in the forthcoming 1–6 h with 12% accuracy enhancement averagely from cellular calls.**

***Index Terms*—Cell towers, Internet of Things (IoT), localization, location prediction, mobile calls.**

## I. Introduction

INTERNET of Things (IoT) connects uniquely identifiable objects into an Internet-like structure and seamlessly integrates the physician world with the digital world based on the participation of billions of networking sensors [1]–[3]. Such sensors are deployed into devices and machines in real world. They may collect various kinds of data, such as environmental data, geographical data, astronomical data, and logistic data. Mobile equipment, transportation facilities, public facilities and home appliances all can be data acquisition in IoT.

IoT has fostered increasing attention to location-based telecommunication applications and services with lots of human digital traces. These traces are captured by such as mobile phones, embedded sensors, and radio-frequency identification (RFID) [4], [5]. Digital traces ranging from cell IDs, mobile calls, short messages to global system for mobile communications (GSM) traces reflect many facets of user mobility and interaction. With the prior knowledge of user location, telecommunication operators can serve users with customized services. Thus, they can attract more users and boost their loyalty.

With the development of IoT and cloud computing technology, predicting mobile phone users' location is essential to many mobile applications, including location-based services [6]–[10], mobile multimedia quality of service (QoS) provision [11], [12], as well as the resource management [13] for mobile computation [14], and storage [15].

Recently, a variety of location prediction schemes have been proposed, e.g., evaluating mobility models for temporal prediction (ETP) [16], NextPlace [17], home-cell community-based mobility model (HCMM) [18], time-variant community model (TVCM) [19], and Markov-based schemes [20]. Most existing schemes predict user location based on user behavioral regularity. This is because users exhibit spatial and temporal regularity, when visiting certain locations such as *homes and offices*. However, users follow the regular mobility patterns in a very loose manner. Therefore, the prediction schemes based on spatial and temporal regularity are limited. As uncovered in [21], the predictability of these schemes is bounded. Some efforts [22], [23] have divulged that social relationship has a significant impact on user mobility patterns, and thus it can be used for location prediction. Nevertheless, they all require external data sources rather than mobile phone traces to characterize user relationships, e.g., *getting user relationship information from Facebook* in [24]. Moreover, they identify neither the moment social relationships nor the encounter duration.

We investigate the location prediction problem by using MIT Reality Mining Dataset (*abbr. Reality dataset*), which involves 2 667 895 GSM traces and 112 508 cellular calls [25]. We find that human mobility exhibits randomness to some degree, and cellular calls are one of factors to result in such randomness. To be specific, the call pattern between a pair of users is highly correlated with their "encounters" (i.e., *simultaneously co-locating at the same region*), and the influence of the cellular

Manuscript received October 17, 2013; revised September 09, 2014; accepted January 03, 2015. Date of publication January 09, 2015; date of current version March 27, 2015. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2014CB340404; in part by the National Science Foundation of China under Grant 61472283, Grant 61471267, and Grant 61103185; in part by the Fok Ying-Tong Education Foundation, China, under Grant 142006; and in part by the Fundamental Research Funds for the Central Universities under Grant 2013KJ034. This project is sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. Paper no. TII-14-0417. (*Corresponding author: Shengjie Zhao.*)

D. Zhang is with the School of Software Engineering, Tongji University, Shanghai 201804, China.

S. Zhao is with the Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 201804, China (e-mail: shengjiezhao@tongji.edu.cn).

L. T. Yang, M. Chen, and H. Liu are with the School of Computer Science, Huazhong University of Science and Technology, Wuhan 430074, China.

Y. Wang is with the Computer Science Department, Kettering University, MI 48504 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2015.2389656

calls on user mobility is short term. Each pair of users who have successive cellular calls and following certain call patterns would co-locate soon with a high probability in a short period of time after the most recent call. Consequently, we introduce the social interplay to reveal the social relationships embedded in the cellular calls. We further investigate the impact of the social interplay on the prediction of user mobility. To invoke the social interplay for location prediction, we introduce the concepts of critical call (CC) and CC pattern.

Note that MIT Reality Mining Dataset brings forth two issues.

1) Locations of cell towers are represented by code names (i.e., *symbolic locations*). This makes some existing prediction schemes fail, e.g., [17]. For a cell tower address "5188.41097" from MIT Reality Mining Dataset, "5188" and "41097" are code names of a cell area and a cell tower, respectively. Apparently, they do not support any arithmetic or logic operation, as global positioning system (GPS) coordinates do.
2) There are a large number of cell tower handover data in user mobility traces as cell towers are highly overlapped.

To this end, we propose NextMe—a location prediction scheme that leverages cellular calls to enhance location prediction accuracy. NextMe comprises several components: 1) data preprocessing; 2) call pattern recognition; 3) periodicity-based and social interplay-based prediction modules; and 4) aggregation unit. To remove the noise and handover data of cell towers, we design a data preprocessing component. Then, NextMe identifies CCs by the call pattern recognition component. Based on the recognition results, NextMe will invoke the prediction module based on either user behavior periodicity or social interplay. Finally, it aggregates the recommendation for prediction results.

Compared with those traditional standard location tracking techniques (e.g., GPS, GSM, and WiFi localization), the proposed method adopts some core technologies in new industry, such as location-based service, digital footprint, and mobile computing. Due to widely usage of mobile devices, mobile-based applications and services are highly required as they reach most of users without additional efforts. Unlike smart phones, which use extra resources (e.g., WAP, WiFi, and Internet applications), basic cellular phone-based mobile location applications depend on existing cellular network infrastructure, which do not need additional resources [26].

Meanwhile, the development of IoT, particularly the explosion of sensor-equipped mobile phones, has led to an unprecedented accumulation of digital footprints—the digital traces that people have left while interacting with cyber-physical spaces. These digital traces have spurred numerous innovative applications in mobile and ubiquitous computing field. By analyzing pervasive data streams collected from personal mobile phones, we can mine the individuals and groups active mode, large-scale human activities, and urban dynamic regular patterns. Researchers adopt the innovative services in human health, public safety, city resource management, environmental monitoring, and transportation management [27]. In addition, the proposed method takes the characteristics of the periodicity and social interplay in cellular traces into consideration simultaneously in the prediction process.

To summarize, the main contributions of this paper are fourfold.

1) NextMe extracts social interplay from telecommunication call records, which reveals the underlying correlation between human mobility patterns and cellular call patterns. It further comes up with a novel prediction module using cellular calls.
2) NextMe introduces the call pattern recognition component, which is based on the proposed concepts of CCs and relevant patterns. The recognition component determines the moment the social interplay-based prediction module works.
3) NextMe is capable of handling symbolic cell tower locations by converting them to regions. This abates the complexity of detecting cell tower topology and transforming the code name of cell towers to computable coordinates. This also removes much cell tower handover noise in the raw dataset.
4) Experimental results demonstrate that NextMe achieves much higher prediction accuracy than the periodicity-based location schemes at region level. The social relationships contribute to 12% on average prediction accuracy.

The rest of this paper is organized as follows. Section II briefly overviews the related work. Section III conducts empirical study with MIT Reality Mining Dataset to reveal the existence of social interplay and its impact on user mobility patterns. The design of the proposed scheme is introduced in detail in Section IV. Section V reports experimental results. Section VI concludes this paper.

## II. Related Work

The IoT evolves from wireless sensor networks, mobile computing to ubiquitous computing. In some extent, IoT can refer to previous research paradigms for localization. In recent years, there are several schemes close to our work, which address the location prediction problem using cellular records. Hence, we mainly review these schemes. In general, they fall into two broad categories: 1) *regularity-based schemes;* and 2) *nonregularity-based schemes*.

### A. Regularity-Based Schemes

These schemes take advantage of temporal and spatial regularities that are exhibited in user daily lives. They foresee user location by detecting periodic patterns in user traces. The periodic mobility model (PMM) [23] was based on an intuition that the majority of human mobility was periodic among a small set of locations. Relying on the time of the day, PMM predicted a user location that was in the location set. Various techniques in artificial intelligence and machine learning have been exploited to discover the mobility regularity, such as *hierarchical clustering technique* [17], *Markov models* [20], and *nonlinear time series technique* [28].

Nevertheless, regularity-based schemes implicitly assume that user mobility periodicity is static. This assumption might not be always held in reality because users usually follow

TABLE I
ABBREVIATIONS USED IN THIS PAPER

| Abbreviation | Full name |
|---|---|
| CC | Critical call |
| CCCP | Critical cellular call pattern |
| CDF | Cumulative distribution function |
| PDF | Probability density function |

floating periodicity. Moreover, they miss a prediction engine, which leverages social relationships.

### *B. Nonregularity-Based Schemes*

The interaction among emerging areas such as social networks, mobile computing, wireless communication, and ubiquitous computing has fostered a new research direction for location prediction from the social perspective. Social network-based schemes uncover that user mobility is partially driven by social relationships. HCMM [18] incorporated social community and user location preferences for prediction. On top of HCMM, Hossmann *et al*. [22] proposed a new way to extract contact lists in mobile devices as social relationships. Backstrom *et al*. [24] predicted user location using user-side address information and connections between Facebook members.

In the latest couple of years, several user behavior studies based on mobile phone traces have been reported, e.g., Barabási's work [29]. In [25], Eagle *et al*. showed the existence of social relationships and the behavioral similarity among the frequent call pairs. Calabrese *et al*. [30] analyzed cellular data and found that 70% of users having reciprocal cellular calls would co-locate at the same cell tower. Both of them, *however*, neither modeled the social relationships based on call records nor proposed any specific location prediction models. In contrast, Cho *et al*. [23] modeled user relationships as a function of distance that users would travel. Then, with the user current location available in checking-in websites as input, they computed the probability using the function that the user would move.

To sum up, the aforementioned schemes have exploited various data sources, rather than cellular call records, to extract social relationships. For instance, the check-in information at Foursquare was exploited to get social relationships in [23], and the friend information in Facebook was used in [24]. In this paper, we find a new kind of social relationship—social interplay—that exists in cellular call records. According to the interplay, we design an engine of location prediction.

## III. EMPIRICAL STUDY

In this section, we conduct an empirical research on the social interplay, as well as its influence on user mobility patterns. Table I lists the abbreviations used in this paper. In order to find the relationships between user social interplay and mobility patterns, we conduct empirical study aiming to answer the following two questions.

1) Does the social interplay exist in cellular calls?
2) When will the social interplay affect the user mobility patterns?

TABLE II
STATISTICS OF MIT REALITY MINING DATASET

| Item | Description |
|---|---|
| Starting time | January 2004 |
| Ending time | July 2005 |
| No. of users | 106 |
| No. of faculties | 11 |
| No. of cell towers | 32 579 |
| No. of areas | 1 027 |
| No. of GSM traces | 2 667 895 |
| Avg. no. of GSM traces perpers on perday | 46.7 |
| No. of mobile calls | 112 508 |
| No. of mobile contacts | 26 965 |
| Logical location | AreaID.CellID |
| Physical coordinates | Not available |

We examine these two questions using MIT Reality Mining Dataset. We choose the dataset due to its popularity and the unavailability of other large-scale cellular trace datasets. This dataset consists of 112 508 cellular calls and 500 000 h of human behavior information, e.g., user location, co-location, proximity, and communication information. Table II summarizes the statistics of this dataset. We extract three kinds of events from MIT Reality Mining Dataset.

These events are as follows.

1) *The cellular call event* that refers to a directed call happened from one user to another.
2) *The face-face encounter event* that denotes two users are connected with Bluetooth.
3) *The coregion event, i.e., the co-locating event* that stands for two users visiting the same region during the same period of time.

We ignore the *face-face encounter event*, because it is out of the scope of this paper. As a matter of fact, we consider the *cellular call events*, *coregion events*, and user cell tower location traces for location prediction.

### *A. Does the Social Interplay Exist in Cellular Calls?*

In order to answer this question, we present two concepts: 1) the inter-coregion time; and 2) inter-call-coregion time. Given two users, they are given as Definition 1.

*Definition 1:* The inter-coregion time is the interval between two consecutive co-locating events, whereas the inter-call-coregion time refers to the interval between the current cellular call time and the forthcoming co-location time.

In fact, the inter-coregion time reflects the periodic co-locating patterns for user pairs and the inter-call-coregion time reveals the co-locating patterns with the influence of cellular calls.

In order to validate whether the social interplay exists in cellular calls, we conduct an experiment across the MIT Reality Mining Dataset. Fig. 1 shows the CDFs of the inter-coregion time and inter-call-coregion time. Overall, the CDF curves of the inter-coregion time and inter-call-coregion time exhibit the similar upward trend in the next few hours. When the prediction time falls into the interval from 0 to 12, the CDF of the inter-call-coregion time is higher than that of the inter-coregion time. When the prediction hour is beyond that interval, the CDF of the inter-call-coregion time becomes less than that of the inter-coregion time. This figure has two important implications.

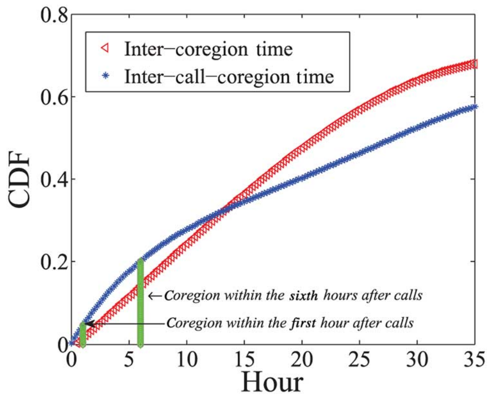



Fig. 1. CDFs of the inter-coregion time and inter-call-coregion time, exhibiting the similar upward trend in the next few hours. When the prediction time falls into the interval from 0 to 12, the CDF of the inter-call-coregion time is higher than that of the inter-coregion time, and when the prediction hour is beyond that interval, the CDF of the inter-call-coregion time becomes less than that of the inter-coregion time.

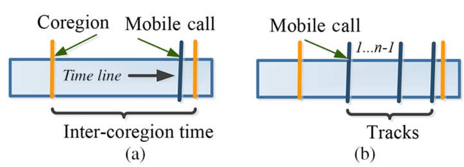



Fig. 2. Cellular call patterns are classified into two categories: (a) there is only one cellular call between two successive co-locating events; (b) there are more than $n$ ($n >= 2$) cellular calls occurred between two successive co-locating events.

1) Social relationships are closely related with user coregion events. We name such social relationships as social interplay.
2) Social interplay has a remarkable impact on user short-term mobility.

Its influence on user mobility will drive user to move within next few hours after cellular calls, e.g., *8 h on average in MIT Reality Mining Dataset*. Hereby, we limit the prediction period from the forthcoming 1–6 h.

### *B. When Will the Social Interplay Affect User Mobility in the Forthcoming 1–6 h?*

As a result of further study, a certain cellular call pattern would coincide with a user mobility pattern i.e., *leading to the coregion of call pairs*. Note that some call pairs share certain call patterns. They have reciprocal yet strong social relationship, but might never co-locate after cellular calls. In this section, we investigate the interval that the social interplay affects user mobility in short. In other words, we would like to answer what kinds of cellular call patterns would lead to the co-locating events.

We classify all cellular calls between two successive co-locating events into two categories, as illustrated in Fig. 2. One category is that there is only one cellular call between two co-locating events. The other category is that there are more than one cellular call between two co-locating events. Through further analysis, we figure out that the first category with one call leads to less than 5% of the total co-locating events, whereas the second category causes the rest of the co-locating events.

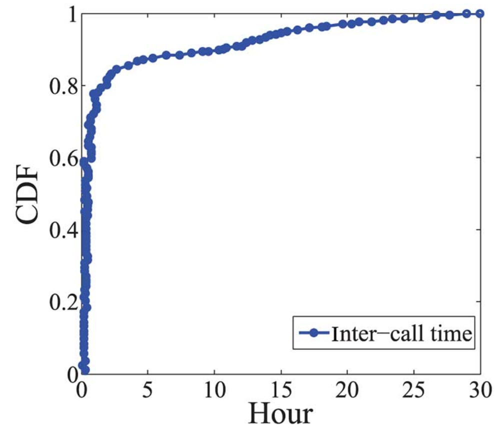



Fig. 3. CDF of inter-call time of cellular calls in CC patterns, indicating that there are about 80% cellular calls occurred successively within 2 h and most of the intervals between two successive calls are short.

Moreover, we find a specific cellular call pattern named as critical cellular call pattern (CCCP), which reveals a user pair to have an immediate co-locating event with a very high likelihood. Specifically, the CCCP is characterized by the number of cellular calls and interval between the last two cellular calls. There must be more than two calls between two successive co-locating events. Meanwhile, the interval between the last two cellular calls should be short.

Given two successive co-locating events $e_i$ and $e_{i+1}$ for a call pair, let:

1) $n$ be the number of cellular calls occurred within the interval of two co-locating events $e_i$ and $e_{i+1}$;
2) $\sigma$ be the threshold of the number of cellular calls;
3) $\rho$ be the interval between the last two cellular calls (e.g., *the* $n-1$ *call and the* $n$ *call*) occurred between $e_i$ and $e_{i+1}$; and
4) $\varrho$ be the threshold of the interval between the last two cellular calls.

Then, the CCCP and CCs are given as Definitions 2 and 3, respectively.

*Definition 2:* A CCCP between two successive co-locating events is a call sequence that satisfies: 1) $n >= \sigma$; and 2) $\rho <= \varrho$.

*Definition 3:* Given a CCCP, the CC refers to the last cellular call.

We further introduce the *inter-call time* concept, which denotes that the interval between two successive cellular calls. Fig. 3 shows the CDF of the inter-call time of cellular calls in CC patterns. We observe that there are about 80% cellular calls occurred successively within 2 h, implying that most of the intervals between two successive calls are short.

In summary, we have shown that social interplay does affect user short-term mobility with a great probability when the CCCP appears. What follows is the design of the proposed scheme.

## IV. NextMe: A Novel Location Prediction Scheme for IoT

According to our findings, we propose a newly featured localization scheme, called NextMe, using mobile traces

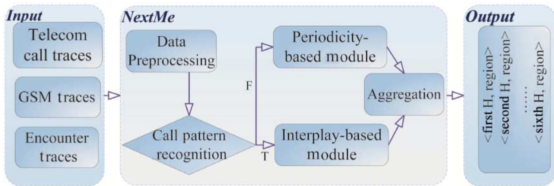



Fig. 4. Architecture of the NextMe scheme, consisting of five separate components: 1) data preprocessing; 2) call pattern recognition; 3) periodicity-based module; 4) social interplay-based module, and 5) aggregation. The data preprocessing module removes the noise and handover data of cell towers. The call pattern recognition module decides to activate corresponding prediction module, if there is no CCCP happening, the periodicity-based prediction module is activated, otherwise, the social interplay-based module is invoked to estimate user location. The aggregation module aggregates the prediction results of two separate modules.

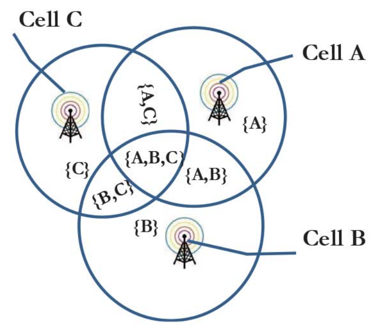



Fig. 5. Region extraction from cell towers. When there are three cell towers A, B, and C and every two cell towers is overlapped, seven regions are extracted as {A}, {B}, {C}, {A, B}, {A, C}, {B, C}, and {A, B, C}.

for IoT, which can achieve fine-grained localization accuracy. NextMe incorporates user behavior regularity and social relationship from cellular calls into location prediction. The regularity refers to the spatial and temporal features of user behaviors. We will detail the scheme in the following sections. Note that we tend to provide localization accuracy at region level. A region is a part of the covering area of a cell tower. We use the "users" and "customers" interchangeably in the contexts without ambiguity.

### A. Overview

NextMe aims to predict user location at region level in the forthcoming 1–6 h. It consists of two subgoals—detecting the CCCP patterns and CCs and then foreseeing user location.

Fig. 4 shows the system architecture of the proposed scheme, which consists of five separate components: 1) data preprocessing; 2) call pattern recognition; 3) periodicity-based module; 4) social interplay-based module; and 5) prediction aggregation. The data preprocessing module filters the noise in the dataset and converts symbolic-represented cell tower locations to our regions. The call pattern recognition module detects whether the CCCPs and CCs take place at the current instant. Suppose there is no happening CCCP, the periodicity-based prediction module is activated. Otherwise, the social interplay-based module is invoked to estimate user location. Finally, NextMe aggregates the prediction results of two separate modules and delivers the estimated location sequence to applications.

### B. Data Preprocessing

The dataset imposes several problems on the proposed scheme. In order to identify the interference from users accurately, we need to handle them.

1) The cell tower IDs are represented as code names rather than physical coordinates. Hence, we do not allow any arithmetic and logic operations on cell tower IDs.
2) The dataset contains lots of redundant data. As each cell might overlap with several other cells, cell handover frequently takes place.

For an overlapping area, only one cell tower is logged in at a moment. In the next moment, another cell tower might be logged in. Therefore, we introduce the concept of "region" that is a part of the covering area of a cell tower. Fig. 5 gives an example, where there are three cell towers $A$, $B$, and $C$ and each pair of two cell towers is overlapped. NextMe models both the independent area (e.g., $\{C\}$ *shown in the diagram*) and the overlapped area (e.g., $\{A, C\}$) as a region, correspondingly. Thus, there are totally seven regions coming from three overlapped cells. NextMe generates 89 764 regions from 32 579 cell towers in MIT Reality Mining Dataset. In the evaluation stage, NextMe regards the prediction as correct when the predicted region sequence is the same as the ground truths.

In addition, the dataset contains many outlier cell towers because it does not provide us with the topology of cell towers. We first collect the times all cell towers appeared in the selected traces for a time period $t$. For each cell tower, we count its occurrence number over the period. Let the set $\{n_1, n_2, \ldots, n_m\}$ ($m$ is the number of the cell towers) denote the occurrence numbers collected. We define the neutral value of the occurrence number $u_n$ as the expected occurrence number when there is no interference and failure, i.e.,

$$u_n = \frac{\sum_{i=1}^{m} n_i}{m}.$$

The sensitivity of the occurrence number of cell towers is measured by the standard deviation of the time series, i.e.,

$$\sigma_n = \sqrt{\left(\sum (n_i - u_n)^2\right)/m}.$$

With the neutral number $k$ ($k > 1$) as the threshold, we can determine whether a cell tower is an outlier or not. Technically, we have the following theorem.

*Theorem 1:* Let $u$ and $\sigma$ be the neutral value and the sensitivity of the occurrence number of a cell tower, respectively. In the given period, if a cell tower appears $n$ times, and $|n - u| \geq k\sigma$, the probability that a cell tower is not an outlier is at least $(1 - \frac{1}{k^2})$.

*Proof:* Derived from Chebyshev's inequality directly. ■

For the same reason, we can identify the overlapped covering areas of cell towers and formulate these areas as regions. After the data preprocessing step, we get the dataset where a user is associated with a region for an hour. Note that a user may visit more than one region in an hour. To be simple, we select the region where the user stays for the longest time as the location for this hour.

**Algorithm 1.** The Process of Checking Whether the Cellular Calls After the Latest Coregion Event $e_i$ Follow a CCCP Pattern or Not For a User Pair $x$ and $y$

```
Input:
    e_i is a coregion event between user x and y
    n is the number of cellular calls occurred after the e_i
      coregion
    σ is the threshold of the number of cellular calls
    ρ is the interval between the last two cellular calls after
      e_i
    ϱ is the interval threshold between the last two cellular
      calls after e_i
Output:
    Yes, return the critical call information
    Otherwise, return NULL
1  begin
2    if isExistCCCP(x, y) then
3      /* Parameter estimation */
4      σ ⟵ estimatePoissonMean(n of all samples);
5      ϱ ⟵ estimatePoissonStandardVar(ρ of all
       samples);
6      /* Get coregion event time */
7      t_i ⟵ getEventTime(e_i);
8      /* Get the number of cellular calls
          happened during the period of time
          t_i */
9      n ⟵ getCallNumber(t_i, x, y);
10     /* the call number of the CCCP
          pattern must satisfy the
          Definition 2. */
11     if n >= σ then
12       ρ ⟵ getLastInterval(t_i, x, y);
13       /* Check the CCCP call pattern */
14       if ρ <= ϱ then
15         return getCriticalCall(t_i, x, y);
16       else
17         return NULL;
18     else
19       return NULL;
20   else
21     return NULL;
22 end
```

### C. Call Pattern Recognition

This component attempts to identify when the social interplay-based module works. To be specific, it has two objectives: 1) detecting the CCCP patterns; and 2) identifying the CCs. Note that a CCCP is defined as a cellular call pattern that would lead to the next coregion event for two users. In this paper, we only consider cellular calls.

According to Definition 2, NextMe recognizes a CCCP from three perspectives: 1) the number of cellular calls; 2) the interval between the last two cellular calls; and 3) the co-locating history. Without co-locating history, NextMe also reproduces the CCCP for a user pair. This slightly degrades the prediction accuracy. With the dataset, the CCCP discovery for all call pairs can be achieved in an offline manner. By Definition 3, NextMe can easily identify the CCs that trigger the social interplay-based prediction module. Algorithm 1 gives the pseudocode of detecting the CCCP pattern for a user pair $x$ and $y$. Lines 3–5 compute the value of the thresholds $\sigma$ and $\varrho$ before the pattern recognition. Evidently, each of them follows the Poisson distribution. We resort to the Poisson parameter estimation technique for model establishment. Lines 11–19 show the CC detection based on the Definitions 2 and 3.

Recap that Algorithm 1 checks the CCCP in an online manner as we do not know the next co-locating time and region. We may get false positive errors. To avoid such errors as possible, we introduce the confidence level as 0.95 to filter many cellular calls in the process of parameter estimation. These calls affect user mobility with a small probability. Thus, the probability $P^{1-6}_{x,y}$ of the call pair $(x, y)$ coregioning in the forthcoming 1–6 h approaches to 1.

### D. Periodicity-Based Module

The periodicity-based module takes the historic location traces of a user as input and foretells user future location as output. It exploits the temporal and spatial periodicity embedded in user mobility traces. As mentioned before, the cell towers in MIT Reality Mining Dataset are symbol-represented so that the arithmetic and logic operations cannot be directly applied. We have to take this into consideration.

We design a periodicity-based module using Kullback–Leibler divergence (KLD). This module consists of three stages.

1) It exploits a combined method of Fourier transform and auto-correlation to retrieve the periodic movements with scattered periods from several reference locations.
2) It uses a probabilistic model to characterize the periodic behaviors based on the detected periods. Then by measuring the KLD-based distance between the movements, the module clusters the scattered periodic movements into integral periodic behavioral patterns.
3) It predicts the user location by using periodic behavioral patterns.

### E. Social Interplay-Based Module

In prediction stage, once a call sequence is detected as a CCCP, NextMe will immediately invoke the social interplay-based module. To reduce the workloads, we can turn OFF the periodicity-based module. In order to design the prediction module according to the social interplay of a user pair, we need to answer the following three questions.

1) When will the call pair co-locate?
2) Where will the call pair co-locate?
3) How long will the pair co-locate?

By answering the above questions, we are able to forecast both user location and the coregion duration. In the rest of this section, we will answer these questions in detail.

*1) When Will the Call Pair Co-Locate?:* In general, given a call pair, their inter-call-coregion time complies with the Poisson distribution. Thus, NextMe employs the Poisson parameter estimation to get the mean and standard variance of

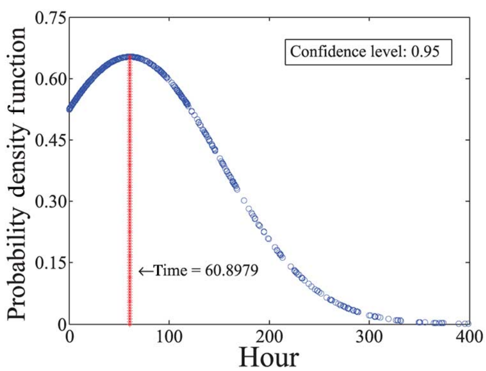



Fig. 6. Pdf of the call pair user 4 and user 8, indicating that the probability of user 4 and user 8 co-locate is more than 60% within 60.8979 h, where the confidence level is set to 0.95.

the inter-call-coregion time. With the mean and standard variance, NextMe obtains the probability density function (pdf). Fig. 6 illustrates the pdf diagram for the selected call pair of users 4 and 8 in the dataset. According to the calling time, the prediction duration and pdf, NextMe estimates the moment the call pair might co-locate at the same region.

To be specific, NextMe collects the inter-call-coregion time of a given call pair from MIT Reality Mining Dataset. Then, it obtains the Poisson distribution of the inter-call-coregion time using parameter estimation. In light of the Poisson formula, NextMe computes the pdf of inter-call-coregion time. According to the time that CCs occurred and pdf, NextMe infers the moment that the call pair might co-locate. The proposed scheme also takes the temporal constraint into consideration that user mobility is subject to the current time, i.e., the next visit to a place relies on the current time. Given a call pair $(x, y)$, let:

1) $\tau$ be the prediction duration (*i.e., its value as 1, 2, 3, 4, 5 or 6*);
2) $C_i$ be the $i$th call in the $n$ successive calls in a CCCP;
3) $K_{x,y}(\tau)$ be a KLD-based likelihood function for the $\tau$th prediction hour; and
4) $G_{x,y}$ be the Poisson function.

NextMe obtains the coregion probability $P_{x,y}(\tau)$ for the user pair $x$ and $y$ at the $\tau$th hour as

$$\arg\max_{\tau} P_{x,y}(\tau) \quad \text{and} \quad \tau \in \{1, 2, \ldots, 6\}$$
$$P_{x,y}(\tau) = P_{x,y}^{1-6} \cdot \frac{K_{x,y}(\tau)}{\sum_{\tau \in \{1,2,\ldots,6\}} K_{x,y}(\tau)} \tag{1}$$

where $K_{x,y}(\tau)$ is computed as

$$K_{x,y}(\tau) = \prod_{1 \le i \le n} G_{x,y}(C_n - C_i + \tau), \tau \in (1, 6) \tag{2}$$

where $G_{x,y}(k)$ is given as

$$\mathrm{G}_{x,y}(k) = \frac{e^{-\sigma}\sigma^k}{k!} \tag{3}$$

where the parameter $\sigma$ is the mean and standard variance for the Poisson function. The value of the parameter is obtained by Poisson parameter estimation on the raw data of the inter-call-coregion time.

To alleviate the computation workload in (1), NextMe generates a social interplay matrix beforehand, where each item is either 1 or 0, representing the related users have contacted with each other or not. By this matrix, NextMe can quickly identify the candidate set of users that have calls with the current user, thus considerably narrowing the searching space of the candidate users. When there are more than one critical call from several candidate users, for the sake of simplicity, NextMe exploits the latest CC for prediction.

*2) Where Will the Call Pair Co-Locate?:* Usually, the above step gets a set of time that the call pair might co-locate. However, there is a spatial constraint that user mobility is affected by his/her current location or preferred locations. That means there is a close relationship between the next place user would visit and the current location. Consequently, the proposed scheme designs the following process for every call pair.

NextMe first predicts the regions that the call pair might co-locate. As users usually prefer to co-locate at nearby regions, it incorporates the region preference into the prediction results. To be specific, NextMe estimates the most possible region for co-locating in two steps.

1) We build two indices for all 683 383 co-locating records. One is the hour index $h$ which records the co-locating hour in a day. The other is the region index $Rg$ by occurring regions.
2) From these two indices, we can easily get the co-locating probability of every region. Equation (4) addresses the probability of the user pair $x$ and $y$ co-locate at the region $r$ at the moment $t'$, where $h_{x,y}(t \bmod 24)$ is the set of co-locating records for the user pair $(x, y)$ in the corresponding section of time $t$ and $\mathrm{Rg}_{x,y}(r)$ refers to the set of co-locating at the region $r$ for the pair $(x, y)$. Because of the fuzziness and freedom of user mobility, NextMe takes $\pm 1$ h slots around the $t'$ into account

$$P_{x,y}(r|t') = \frac{\sum\limits_{t \in \{t'-1\ldots t'+1\}} |h_{x,y}(t \bmod 24) \cap \mathrm{Rg}_{x,y}(r)|}{\sum\limits_{t \in \{t'-1\ldots t'+1\}} |h_{x,y}(t \bmod 24)|}. \tag{4}$$

Therefore, the probability of the call pair $x$ and $y$ co-locate at the region $r$ at time $t$ is calculated as the product of (1) and (4). NextMe deduces the region that a given user would like to move. Note that when there is no co-locating records for a user pair, we randomly select a current location of one user as the encounter place. To keep concise, we set the probability of co-locating at the picked location as the value of one to the total number of regions that both users in the call pair have visited before.

*3) How Long Will the Call Pair Co-Locate?:* NextMe is targeted to forecast user location in the forthcoming 1–6 h. Hence, after prediction for the first hour, it needs to estimate the co-locating duration.

NextMe takes advantage of two sources to estimate the duration: 1) the average time of the user pair co-locate; and 2) location inference. Given a user pair, the former source is computed as the ratio of the total co-locating time to their co-locating times. In contrast, the latter source is from GSM traces that reveals the user location in most time. User location is associated with region addresses and sample time in GSM traces. By continuously sampling user location in GSM traces, we infer the duration of a user staying at a cell tower. For example, a user stays at a region $r_1$ for co-location at time $t_1$ and appears at another nearby region $r_2$ at time $t_2$. Thus, we estimate the co-locating duration. We adopt the average co-locating time as the basic estimation of the co-locating duration, and improve it when the second source is available. We count the co-locating duration by hour, as around 91% user pairs will co-locate for more than one hour.

### F. Prediction Aggregation

NextMe makes prediction by fusing the location prediction results from its two modules. The interplay-based module predicts the short-term locations that the user pairs would visit, whereas the periodicity-based module generates both short-term and long-term locations.

When initiating the prediction process, NextMe will detect the CCs. Once a CC is detected, it will turn on the social interplay-based module. Suppose NextMe detects a CC at 00:45:00 for the user $Tomas$. It foretells his location using the interplay-based module. Its prediction is a region $R$ at the fourth hour after the CC. The co-locating duration will be 1 h. Thus, NextMe will predict all his location before the co-locating time (e.g., *04:45:00*) by the periodicity-based module. His location in the next hour after the co-location (e.g., *05:45:00*) will still be the region $R$.

So far we have finished the introduction to the proposed scheme. Among all the modules, the call pattern recognition is fundamental. To reduce the computation complexity, it is necessary to generate social interplay matrix beforehand.

## V. Experiments

To examine the performance of the proposed scheme, we have conducted a series of experiments with a real-world user telecom trace that enables scientists to investigate human mobility and interaction using user footprints recorded by smart phones. Empirical study is run in MATLAB, and the prediction model is implemented in Java SE Development Kit 7. Another implementation using public cloud is given in [31], [32].

In particular, we would like to answer the following questions.

1) How is the overall performance of the proposed scheme? Does it work better than the state-of-the-art scheme?
2) How is the performance of the periodicity-based module?
3) How much does the social interplay affect the prediction accuracy of the proposed scheme?

### A. Metrics

We select the prediction accuracy to measure the performance of the proposed scheme. For every user, NextMe will deliver six locations as the predicted results for the forthcoming 1–6 h. Let $E_j^i$ be the prediction result at the $i$th hour for the user $j$. The $E_j^i$ is valued 1 when the predicted region is the same as the one that is really visited, otherwise it goes to 0. Equation (5) defines the prediction accuracy at the $i$th hour in a prediction process

$$\text{Accuracy}^i = \frac{\sum_{j \in s} E_j^i}{||s||} \tag{5}$$

where $s$ is the times we repeat the experiments with different activity monitoring time, and the $||s||$ is the size of $s$.

### B. Experiment Design

We select MIT Reality Mining Dataset [25] as the raw dataset, which is generated from 106 users in 1.5 years. Its statistical information is given in Table II. We extract the 2 666 897 GSM traces, call logs and parts of activity information records for this study.

We select two periodicity-based schemes: 1) periodicity-based (Perio) predictor; and 2) collective behavioral patterns (CBP) predictor [33] as the baseline. Perio scheme is described in detail in Section IV-D, which is built on top of the state-of-the-art periodicity detection technique—KLD [28]. CBP scheme aims to forecast ones locations in next 6 h based on the locations of other users. First, the scheme observes the existence of collective behavioral patterns through association pattern mining, and uncovers the association rules of CBP. The association rules identify the correlation of user's locations at the same moment. CBP can acquire such correlations between the current locations of crowds and the locations of target user in the next few hours. Then the scheme extends the collective behavioral patterns to associate users' locations from crowds to individuals with time shifting. Finally, the scheme builds a Bayesian model to learn the correlations with time-shifting from the mobility data of crowds, and applies it to localization.

### C. Overall Performance

We do several experiments for all participants to verify the overall performance of the proposed scheme. We use the aggregated prediction accuracy as the metric. The aggregated prediction accuracy is computed as the average accuracy of all predicted users for every prediction period.

Fig. 7 illustrates the overall performance of the proposed scheme, where the *x*-axis and *y*-axis represent the forthcoming hour and prediction accuracy respectively. With the value of *x* varies from 1 to 6, the values of *y* for Perio, CBP, and NextMe schemes decrease. Compared with the Perio scheme, NextMe exhibits a slower downward trend and CBP maintains relatively stable trend. In all prediction periods, NextMe achieves higher prediction accuracy than the Perio and CBP schemes. This shows the superiority of the proposed scheme regarding prediction accuracy. Moreover, Fig. 7 demonstrates that the social interplay significantly contributes to the prediction accuracy. This is attributed to the fact that the cellular calls affect user

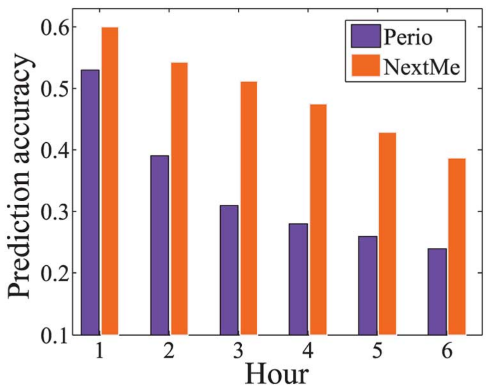



Fig. 7. Overall performance of Perio, CBP, and NextMe schemes, indicating that the prediction accuracy for Perio, CBP, and NextMe schemes decrease in the forthcoming 1–6 h, compared with the Perio scheme, NextMe exhibits a slower downward trend and CBP maintains relatively stable trend. In all prediction periods, NextMe achieves higher prediction accuracy than the Perio and CBP schemes.

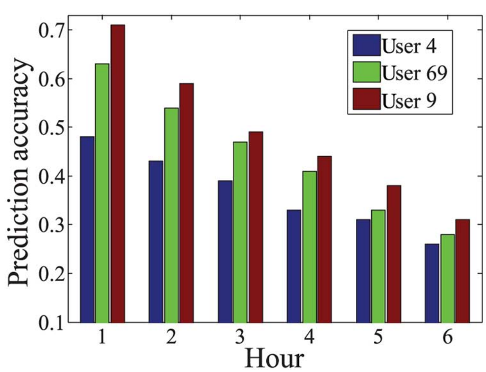



Fig. 8. Case study: the performance of CBP and Perio for active user 4, semi-active user 69 and passive user 9, demonstrating that Perio scheme can obtain more than 58% accuracy in the forthcoming 1–2 h for user 9. Compared with CBP scheme, Perio gains higher accuracy for the active user 4 and the semi-active user 69, but the accuracy becomes lower for the passive user 9. The accuracy of the user 4 is less than that of the user 9 for both CBP and Perio schemes, which is because the active users usually exhibit high-level randomness and freedom of their mobility.

short-term mobility. In addition, NextMe copes well with the traces with symbolic locations by the data preprocessing module.

### D. How Is the Performance of the Periodicity-Based Module

This section aims to validate the periodicity-based module of NextMe scheme. The users in MIT Reality Mining Dataset could be classified into three categories: 1) active; 2) semi-active; and 3) passive users. Active users are the users who frequently communicate with other persons, whereas the passive users are the users who have fewer social connections. To represent these user categories, we extract users 4, 69, and 9 as exemplars.

Fig. 8 illustrates the prediction accuracy for user 4, 69, and 9 adopted Perio and CBP schemes respectively. Overall, for the passive user 9, Perio scheme can obtain more than 58% accuracy in the forthcoming 1–2 h. This explains that NextMe is

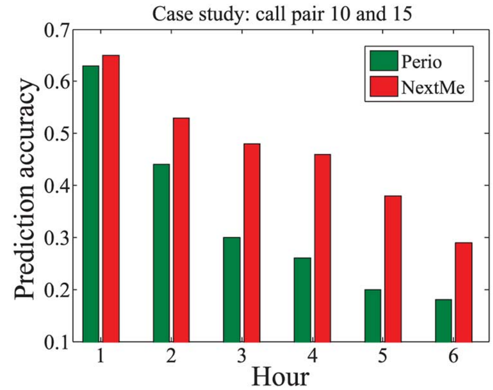



Fig. 9. Case study: the performance of Perio and NextMe schemes for call pair—active users 10 and 15, indicating that the prediction accuracy is improved from 3% to 14% in the forthcoming 1–6 h and its average is 9.1%.

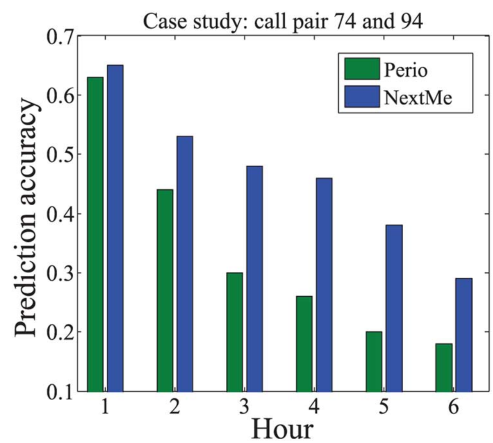



Fig. 10. Case study: the performance of Perio and NextMe schemes for call pair—passive users 74 and 94, demonstrating that the prediction accuracy is enhanced from 2% to 20% in the forthcoming 1–6 h and its average is 13.6%.

appropriate to make prediction based on the mobility regularity. Compared with CBP scheme, Perio gains higher accuracy for the active user 4 and the semi-active user 69, but the accuracy becomes lower for the passive user 9. Furthermore, the accuracy of the active user 4 is less than that of the passive user 9 for both Perio and CBP schemes. This is because the user 9 shows a higher-level regularity than the user 4. In fact, the active users usually exhibit high-level randomness and freedom of their mobility. Thus, it appears challenging to forecast their location only by behavior regularity.

### E. How Much Does the Social Interplay Affect the User Mobility

In this section, we examine the affection of social interplay to user mobility. We extract two classes of call pairs: 1) active users 10 and 15; and 2) passive users 74 and 94. We carry out experiments by treating every call pair as a new "user" (i.e., *the prediction accuracy is the average accuracy for the two users in a call pair*).

Figs. 9 and 10 present the prediction accuracy of Perio and NextMe schemes for the call pairs {10, 15} and {74, 94}, correspondingly. As expected, NextMe scheme obtains higher prediction accuracy than the Perio scheme in these two figures. In Fig. 9, the average improvement of the prediction accuracy is from 3% to 14%, and its average is 9.1%. In Fig. 10, the enhancement of the prediction accuracy falls into the interval of 2% to 20%, and its average is 13.6%. Consequently, we infer that the social interplay is a fundamental driver to user mobility, and its contribution is 12% on average for call pairs in MIT Reality Mining Dataset.

## VI. Conclusion

As the penetration of IoT goes up rapidly, location-based telecommunication services are vital to telecommunication operators. In this paper, we have investigated the large-scale mobile traces from Telecom logs and introduced the social interplay that affects user short-term mobility. We further propose a prediction scheme named NextMe that can predict user location at region level in the forthcoming 1–6 h.

In the future, we will incorporate more telecommunication records, e.g., short messages, into the system design. Also, we would like to develop a practical system and mobile applications, enabling a series of IoT services in telecommunication related fields.

## References


[1] S. Li, L. Xu, and S. Zhao, "The internet of things: A survey," *Inf. Syst. Frontiers*, pp. 1–17, 2014, doi: 10.1007/s10796-014-9492-7.

[2] K. M. Lim, W. Bahr, and C. H. Leung, "RFID in the warehouse: A literature analysis (1995–2010) of its applications, benefits, challenges and future trends," *Int. J. Prod. Econ.*, vol. 145, no. 1, pp. 409–430, 2013.

[3] J. Wan, D. Zhang, S. Zhao, L. T. Yang, and J. Lloret, "Context-aware vehicular cyber-physical systems with cloud support: Architecture, challenges, and solutions," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 106–113, Aug. 2014.

[4] S. Ivanov and E. Nett, "Localization-based radio model calibration for fault-tolerant wireless mesh networks," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 246–253, Jun. 2013.

[5] A. Whitmore, A. Agarwal, and L. Da Xu, "The internet of things: A survey of topics and trends," *Inf. Syst. Frontiers*, pp. 1–14, 2014, doi: 10.1007/s10796-014-9489-2.

[6] L. Zhang *et al.*, "Cloud manufacturing: A new manufacturing paradigm," *Enterp. Inf. Syst.*, vol. 8, no. 2, pp. 167–187, May 2012.

[7] C. Jin, F. Li, M. Wilamowska-Korsak, L. Li, and L. Fu, "BAP-GA: A new genetic algorithm for system optimization and excellent schema selection," *Syst. Res. Behav. Sci.*, vol. 31, no. 3, pp. 337–352, 2014.

[8] B. Huang, C. Li, and F. Tao, "A chaos control optimal algorithm for Qos-based service composition selection in cloud manufacturing system," *Enterp. Inf. Syst.*, vol. 8, no. 4, pp. 445–463, 2014.

[9] D. Zhang, M. Chen, M. Guizani, H. Xiong, and D. Zhang, "Mobility prediction in telecom cloud using mobile calls," *IEEE Wireless Commun.*, vol. 21, no. 1, pp. 26–32, Feb. 2014.

[10] D. Zhang, A. V. Vasilakos, and H. Xiong, "Predicting location using mobile phone calls," in *Proc. ACM SIGCOMM Conf. Appl., Technol., Archit., Protoc. Comput. Commun.*, Helsinki, Finland, Aug. 2012, pp. 295–296.

[11] W.-S. Soh and H. S. Kim, "Qos provisioning in cellular networks based on mobility prediction techniques," *IEEE Commun. Mag.*, vol. 41, no. 1, pp. 86–92, Jan. 2003.

[12] W. Tan *et al.*, "A trust evaluation model for e-learning systems," *Syst. Res. Behav. Sci.*, vol. 31, no. 3, pp. 353–365, 2014.

[13] A. Roy, S. K. Das, and A. Misra, "Exploiting information theory for adaptive mobility and resource management in future cellular networks," *IEEE Wireless Commun.*, vol. 11, no. 4, pp. 59–65, Aug. 2004.

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

[15] P. Stuedi, I. Mohomed, and D. Terry, "Wherestore: Location-based data storage for mobile devices interacting with the cloud," in *Proc. 1st ACM Workshop Mobile Cloud Comput. Services: Social Netw. Beyond*, 2010, p. 1.

[16] Y. Chon, H. Shin, E. Talipov, and H. Cha, "Evaluating mobility models for temporal prediction with high-granularity mobility data," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, Lugano, Switzerland, 2012, pp. 61–70.

[17] S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell, "Nextplace: A spatio-temporal prediction framework for pervasive systems," in *Proc. 9th Pervasive*, Helsinki, Finland, 2011, pp. 152–169.

[18] C. Boldrini and A. Pass, "Modelling spatial and temporal properties of human mobility driven by users' social relationships," *Comput. Comm.*, vol. 33, pp. 1056–1074, 2010.

[19] W. J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling time-variant user mobility in wireless mobile networks," in *Proc. 27st IEEE Infocom*, Anchorage, AK, USA, 2007, pp. 758–766.

[20] H. Si, Y. Wang, J. Yuan, and X. Shan, "Mobility prediction in cellular network using hidden markov model," in *Proc. 7th IEEE Consumer Commun. Netw. Conf.*, Mannheim, Germany, 2010, pp. 1–5.

[21] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[22] T. Hossmann, T. Spyr, and F. Legendre, "Putting contacts into context: Mobility modeling beyond inter-contact time," in *Proc. 12th MobiHoc*, Paris, France, 2011, pp. 199–208.

[23] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, San Diego, CA, USA, 2011, pp. 1082–1090.

[24] L. Backstrom, E. Sun, and C. Marlow, "Find me if you can: Improving geographical prediction with social and spatial proximity," in *Proc. 19th Int. Conf. World Wide Web*, Raleigh, NC, USA, 2010, pp. 61–70.

[25] N. Eagle, A. S. Pentlandb, and D. Lazerc, "Inferring social network structure using mobile phone data," in *Proc. Natl. Acad. Sci.*, 2009, pp. 15 274–15 278.

[26] M. Abo-Zahhad, S. M. Ahmed, and M. Mourad, "Services and applications based on mobile user's location detection and prediction." *Int. J. Commun. Netw. Syst. Sci.*, vol. 6, no. 4, pp. 167–175, 2013.

[27] A. Abu-Mahfouz and G. Hancke, "Distance bounding: A practical security solution for real-time location systems," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 16–27, Sep. 2013.

[28] Z. Li, B. Ding, J. Han, R. Kays, and P. Nye, "Mining periodic behaviors for moving objects," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, 2010, pp. 1099–1108.

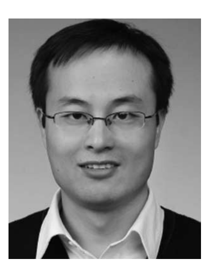
[29] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabási, "Human mobility, social ties, and link prediction," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, 2011, pp. 1100–1108.

[30] F. Calabresel, Z. Smoreda, V. D. Blondel, and C. Ratti, "Interplay between telecommunications and face-to-face interactions: A study using mobile phone data," *PLoS ONE*, vol. 6, no. 7, p. e20 814, 2011.

[31] D. Zhang, D. Zhang, H. Xiong, L. T. Yang, and V. Gauither, "Nextcell: Predicting location using social interplay from cell phone traces," *IEEE Trans. Comput.*, Nov. 2013, doi: 10.1109/TC.2013.223.

[32] D. Zhang, D. Zhang, H. Xiong, C.-H. Hsu, and A. Vasilakos, "Basa: Building mobile ad-hoc social networks on top of android," *IEEE Netw.*, vol. 28, no. 1, pp. 4–9, Jan. 2014.

[33] H. Xiong, D. Zhang, D. Zhang, and V. Gauthier, "Predicting mobile phone user locations by exploiting collective behavioral patterns," in *Proc. IEEE UIC/ATC*, 2012, pp. 164–171.

**Daqiang Zhang** (SM'14) received the Bachelor of Management degree in management science in 2003; the M.S. degree in computer science in 2006; and the Ph.D. degree in computer science in 2010.

He is an Associate Professor with the School of Software Engineering, Tongji University, Shangai, China. He has authored more than 60 papers in major journals and international conferences in his research areas. His research interests include mobile computing, distributed computing, and wireless sensor networks.

Prof. Zhang is on the Editorial Board of *Telecommunication Systems*, *European Transactions, on Telecommunications*, *KSII Transactions on Internet and Information Systems*, and *New Review of Hypermedia and Multimedia*. He was the recipient of Best Paper Award from ACCV'2009 and UIC'2012.

**Shengjie Zhao** (SM'09) received the B.S. degree in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 1988; the M.S. degree in electrical and computer engineering from China Aerospace Institute, Beijing, China, in 1991; and the Ph.D. degree in electrical and computer engineering from Texas A&M University, College Station, TX, USA, in 2004.

He is a Professor with the College of Electronics and Information Engineering, Tongji University, Shangai, China. In previous postings, he conducted research at Lucent Technologies, Whippany, NJ, USA, and China Aerospace Science and Industry Corporation, Beijing, China. His research interests include big data, wireless communications, image processing, and signal processing.

Dr. Zhao is a PI of 4th Topic of Project of Big Data and the National Basic Research Program of China (973 Program), and was a PI of TD-LTE-Advanced System Research and Development, National Science and Technology Specific Projects of China. He is a Fellow of the Thousand Talents Program of China and is a Member of Sigma Xi.

**Laurence T. Yang** received the B.S. degree in computer science in 1994, and the Ph.D. degree in computer science in 1999.

He is a Professor with the School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China. He has authored many papers in various refereed journals, conference proceedings, and book chapters in these areas.

Prof. Yang has been involved actively in conferences and workshops as a Program/General/Steering Conference Chair (mainly as the Steering Co-Chair of the IEEE UIC/ATC, IEEE CSE, IEEE HPCC, IEEE/IFIP EUC, IEEE ISPA, IEEE PiCom, IEEE EmbeddedCom, IEEE iThings, and IEEE GreenCom), and numerous conference and workshops as a Program Committee Member. He served as the Vice-Chair of the IEEE Technical Committee of Supercomputing Applications (TCSA) until 2004, was the Chair (elected in 2008 and 2010) of the IEEE Technical Committee of Scalable Computing (TCSC), and the Chair of the IEEE Task Force on Ubiquitous Computing and Intelligence (2009–present). He was also on the steering committee of IEEE/ACM Supercomputing Conference Series (2008–2011), and the National Resource Allocation Committee (NRAC) of Compute Canada (2009–2013).

**Min Chen** received the B.S. degree in electrical engineering in 1999, and the Ph.D. degree in electrical engineering in 2004.

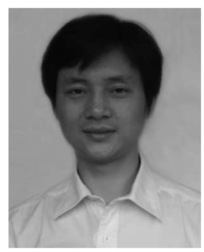
He was an Assistant Professor with the School of Computer Science and Engineering, Seoul National University (SNU), Seoul, South Korea, from September 2009 to February 2012. He worked as a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, BC, Canada. Before joining UBC, he was a Post-Doctoral Fellow with SNU. Currently, he is a Professor with the School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China. He has authored more than 170 papers.

Prof. Chen is a Guest Editor for *The Network* and IEEE WIRELESS COMMUNICATIONS. He is the Symposium Co-Chair for the IEEE ICC 2012 and the IEEE ICC 2013. He is General Co-Chair for the IEEE CIT 2012. He is a TPC Member for the IEEE INFOCOM 2014. He is a Keynote Speaker for CyberC 2012 and Mobiquitous 2012. He was the recipient of the Best Paper Award from the IEEE ICC 2012, and the Best Paper Runner-up Award from QShine 2008.

**Yunsheng Wang** received the B.Eng. degree in electronic and information engineering from the Dalian University of Technology, Dalian, China, in 2007; the M.Res. degree in telecommunication from University College London, London, U.K., in 2008; and the Ph.D. degree in computer and information sciences from Temple University, Philadelphia, PA, USA, in 2013.

He is an Assistant Professor with the Department of Computer Science, Kettering University, Flint, MI, USA. His research interests include various topics in the application and protocols of wireless networks, efficient communication in delay tolerant networks, mobile opportunistic social networks, and vehicular networks. He has authored more than 20 papers in major journals and international conferences in the above areas.

Dr. Wang is the Editor of the *International Journal of Ad Hoc and Ubiquitous Computing*. His research is supported by the U.S. National Science Foundation and the U.S. Department of Justice.



**Huazhong Liu** received the B.S. and M.S. degrees in computer science from the School of Computer Science and Technology, Jiangxi Normal University, Hunan Normal University, Changsha, China, in 2004 and 2009, respectively. Currently, he is pursuing the Ph.D. degree at the School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China.

His research interests include big data, cloud computing, and Internet of Things.