Understanding Motivations behind Inaccurate Check-ins

FENGLI XU, GUOZHEN ZHANG, ZHILONG CHEN, JIAXIN HUANG, and
YONG LI*, Tsinghua University, China
DIYI YANG, Carnegie Mellon University, United States
BEN Y. ZHAO, University of Chicago, United States
FANCHAO MENG, Tencent Inc., China

Check-in data from social networks provide researchers a unique opportunity to model human dynamics at scale. However, it is unclear how indicative these check-in traces are of real human mobility. Prior work showed that a significant portion of Foursquare check-ins did not match with the physical mobility patterns of users, and suggested that misrepresented check-ins were incentivized by external rewards provided by the system.

In this paper, our goal is to understand the root cause of inaccurate check-in data, by studying its validity in social media platforms without external rewards for check-ins. We conduct a data-driven analysis using an empirical check-in trace of more than 276,000 users from WeChat Moments, with matching traces of their physical mobility. We develop a set of hypotheses on the underlying motivations behind people's inaccurate check-ins, and validate them using a detailed user study which includes both surveys and interviews. Our analysis reveals that there are a surprisingly large number of inaccurate check-ins even in the absence of rewards: 43% of total check-ins are inaccurate and 61% of survey participants report they have misrepresented their check-ins. We also find that inaccurate check-ins are often a result of user interface design as well as for convenience, commercial advertisement and self-presentation.

188

CCS Concepts: • **Information Systems → User/Machine Systems**; **User Interfaces**; • **Computer Applications → Social and Behavioral Sciences**;

Additional Key Words and Phrases: Inaccurate check-ins; social media; wechat; data driven; user study

ACM Reference Format:
Fengli Xu, Guozhen Zhang, Zhilong Chen, Jiaxin Huang, Yong Li, Diyi Yang, Ben Y. Zhao, and Fanchao Meng. 2018. Understanding Motivations behind Inaccurate Check-ins. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 188 (November 2018), 22 pages. https://doi.org/10.1145/3274457

*This is the corresponding author

This work was supported in part by The National Key Research and Development Program of China under grant 2017YFE0112300, the National Nature Science Foundation of China under 61861136003, 61621091 and 61673237, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology.
Authors' addresses: Fengli Xu, xfl15@mails.tsinghu.edu.cn; Guozhen Zhang; Zhilong Chen; Jiaxin Huang;
Yong Li, liyong07@tsinghua.edu.cn, Beijing National Research Center for Information Science and Technology(BNRist), Department of Electronic Engineering, Tsinghua University, Beijing, China; Diyi Yang, diyiy@cs.cmu.edu, Carnegie Mellon University, Pittsburgh, United States; Ben Y. Zhao, ravenben@cs.uchicago.edu, University of Chicago, Chicago, United States; Fanchao Meng, Tencent Inc. Beijing, China.

2573-0142/2018/11-ART188 $15.00
https://doi.org/10.1145/3274457

1 INTRODUCTION

In recent years, location sharing services, also known as check-in services, have become an increasingly popular feature integrated into top social media platforms such as Twitter, Facebook, Foursquare and WeChat. For example, more than 10 billion check-ins have been made on Foursquare platform by 2016. Check-in services not only allow their users to interact, but also provide a means for large user populations to properly document their daily activities with location tags and time stamps. The rise of check-in services provides a unique opportunity for research communities to study human dynamics. To date, researchers have tried to exploit available check-in data to facilitate a wide range of studies, including human mobility modeling [9, 37, 55], exploring the correlation between friendship and distance [38, 52], and optimizing viral marketing strategies [2, 29, 56].

Despite the benefits promised by the check-in data, recent empirical results have raised serious concerns about the accuracy of data collected from check-in service providers like Foursquare [27, 45, 53]. Prior measurements find evidence that nearly 75% of Foursquare check-ins do not match with the actual physical locations of users [53]. Later user studies examine why users misrepresent their locations in check-in data, and the results suggest that much of the misrepresentation is incentivized by rewards provided by the platform, such as badges, Mayorships, and financial incentives [45].

Our study is motivated by a deeper understanding of misrepresentation in location-based social networks (LBSNs). Our key questions are: are users of LBSNs truly driven to misrepresent their location by rewards such as Mayorships, badges, or financial incentives, *i.e.* discounts? Is user check-in behavior driven by more fundamental motivations, and how would users behave in the absence of external rewards? Given the potential value of accurate check-in data, answers to these questions could dramatically influence the future design of these systems, and shed light on whether accurate user mobility data can indeed be extracted from LBSN platforms.

To answer these questions, we conduct a study on check-in behavior on the largest social media platform in China, *i.e.*, WeChat. WeChat supports location-based check-in services through a feature called WeChat Moments. It is a representative of check-in services, since its check-in interface is similar to other mainstream social media platforms, *e.g.*, Twitter and Facebook. The primary difference is WeChat Moments does not offer explicit incentives for check-ins, which may reduce the amount of bias or misbehavior in its check-in data stream.

Specifically, we conduct a detailed study of check-in behavior using a combination of analysis on large user datasets and a comprehensive user study including both surveys and interviews. *First*, we collect a large-scale empirical check-in dataset from more than 250,000 users in WeChat Moments. This dataset is valuable because it contains both real users' check-in events and their actual physical mobility traces as reported by the localization module in WeChat Moments. By comparing the check-in data with physical geolocation data, we effectively gain empirical ground truth that allows us to evaluate the correctness of check-in events. We characterize the distinct features of inaccurate check-ins, and develop key hypotheses on the underlying motivations by combining empirical observations with established human behavior models. *Second*, to validate hypotheses derived from this analysis, we carefully design a user study including both user surveys and user interviews. The user survey is includes responses from 541 users of WeChat Moments, which allows us to quantitatively validate our hypotheses. To gain deeper insights on users' check-in behavior, we use open-ended questions to allow users to explain their motivations in the survey. We then follow up with a deeper analysis of specific users and motivations through an interview study with 8 participants.

Our findings can be summarized into three points. First, even on a social media platform without external rewards for check-ins, a significant amount of check-ins still exhibit large discrepancies

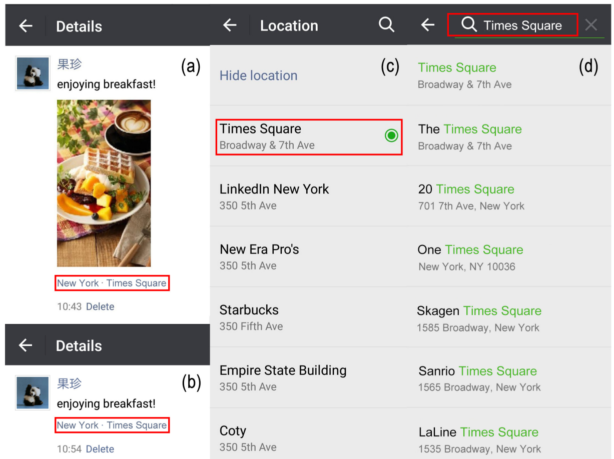


Fig. 1. Illustration of check-ins in WeChat Moments.

from user's real locations. More specifically, 43% of total check-ins do not match with physical mobility traces, and 61% of survey participants self-report that they have misrepresented their check-ins in social media before. Second, the interface that users interact with on check-in services can play a surprisingly important role in inaccurate check-ins. A key reason behind 67% of inaccurate check-ins is that the PoI suggestions by the check-in service can be imprecise, and more accurate tags matching users' locations can be ranked lower in the recommended list. Users often opt not to scroll deep into the list, and instead choose one of the "incorrect" options offered at the top of the list. Third, our user study identifies three key personal reasons that lead to inaccurate check-ins: delayed check-ins for convenience, advertising to achieve commercial gains, and misrepresentation to impress one's followers. In contrast, we are surprised to find that privacy concerns (revealing one's true physical location) do not play a significant role in the process. Finally, the interview study reveals concrete instances and detailed reasons of inaccurate check-ins, which are consistent with the data analysis and survey study. For example, users tend to delay their check-ins until they have spare time for photo modification, and they intend to check-in interesting location tags to appear cool in social media.

2 BACKGROUND: CHECK-IN SERVICES AND WECHAT MOMENTS

Check-in services in Social Media. With the rapid development of LBSNs, check-in services have become a popular feature widely adopted by social media platforms. Social platforms such as Facebook, Twitter and WeChat now allow their users to post and share content with location tags and time stamps. These check-in services can be classified into two categories: LBSNs that focus user interactions primarily around check-ins and physical co-location of their users, *e.g.* Foursquare, Yelp, Gowalla, and more general social media platforms such as Facebook and WeChat, which integrate check-ins as part of user-generated content. Dedicated LBSNs tend to incentivize user check-ins with external rewards (*e.g.* badges, Mayorships and monetary incentives), while larger social platforms offer no explicit rewards for check-ins.

WeChat Moments. Since its launch in 2011 originally as a messaging application, WeChat has become the most popular social media platform in China. Apart from messaging, WeChat supports a wide variety of functionality, including WeChat Moments, an immensely popular component that allows users to post pictures, comments and videos with their friends. WeChat Moments also integrates check-in services, allowing users to post content with location tags that refer to points-of-interest (PoIs). Details of check-ins in WeChat Moments are shown in Figure 1 (a) and (b).

There are two ways to generate check-ins in the platform. First, when prompted, the platform automatically generates a list of recommended PoIs, and users can select a PoI from the list to check-in to, as shown in Figure 1(c). Current recommendations first filters a candidate set of PoIs within some distance, and then ranks them based on popularity, personal preferences and other features. The top 5 PoIs are displayed in the first page, and users need to pull down the list for additional suggestions. Alternatively, users can also skip the recommended list, and check-in to a PoI by manually entering its name, as shown in Figure 1(d).

3 RELATED WORK

3.1 Motivations Behind Posting Check-ins

The dramatic success of mobile social media platforms has drawn significant research attentions to users' motivations behind the check-in behavior. Earlier research reveals that the intentions behind posting check-ins go beyond the simple need of keeping tracks of user's daily life [27]. Specifically, there are numerous factors that may encourage users to generate check-ins, including social interaction [36], self-disclosure [46], gaming [10], check-in for rewards [45] and location promotion [29]. On the other hand, factors of privacy, user preferences and social pressure may prevent users from posting check-ins [27, 45]. More recent works suggest that motivations of posting check-ins vary across different platforms, since users tend to present different information depending on the audience [6, 43]. Specifically, the check-in behaviors in general purpose social media differ significantly from those platforms that focus on check-in services, for example, the landscape of check-ins is changing from coordinate tags to location tags and most users post check-ins consciously [43].

In this paper, we contribute to the understanding of the motivation of posting check-ins on general purpose social media platforms, *i.e.*, WeChat Moments. Unlike most prior work, we focus on check-ins on platforms without rewards or incentives, so that we can gain a deeper understanding of other motivations for misrepresentations of their location.

3.2 Bias and Discrepancies in Social Media Data

The availability of large social media datasets has led to their rise as ideal platforms to study human dynamics at scale [9, 38, 56]. However, a number of recent studies have raised questions about the credibility of social media datasets. For example, it has been shown that a significant amount of social media posts consist of high quality rumors [28, 51]. In addition, social media datasets provide biased views of user opinions, due to behavioral differences of opinion groups and inappropriate system designs [23, 26].

In the more specific context of location-based networks, check-in datasets are somewhat unique in that errors and bias could be accurately quantified, since it is feasible for users' check-in data to be directly compared against ground truth of their physical locations. Early work finds that that check-ins in Foursquare are not good representatives of activities in reality, but rather communicative events between users [36]. As a result, the number of check-ins in a location has no clear correspondence with visitor numbers. A recent study quantitatively measures the discrepancy between Foursquare check-ins and physical mobility traces, and finds that about 75% of total check-ins do not match with physical mobility traces [53]. The authors further carry out a survey study and demonstrate that the primary cause for misrepresented check-ins are the external rewards provided by Foursquare [45]. Other work has shown that deceptive location disclosures have significant correlations with physical distance, tie strength and order of visibility [15].

Given these results, one might hypothesize that check-in events would accurately represent users' physical mobility in the absence of external rewards. Until now, however, this has not been validated or disproved. We note that given the continued growth of larger social platforms and the gradual decline of LBSNs like Foursquare, unrewarded check-in services on Facebook and WeChat are likely the dominant check-in services (and sources of data) for the foreseeable future. This makes understanding misrepresentation on these platforms even more critical.

3.3 Interplay between Social Media Design and User Behaviors

There is a long history of work to understand the interplay between social media system design and user behavior. Recent work demonstrates that users exhibit significant diverse needs when using different social media [54]. For instance, the need of self-disclosure closely correlates with the target social media platform [21] and the traits of users [32]. Users tend to employ different self-presentation strategies in different social media platforms [40]. These results indicate that understanding user behaviors can significantly improve the design of social media interfaces. In addition, the design of social media systems also plays an important role in user behavior. For example, the design of feed stream significantly impacts on users' choices of the content they make public [12, 13]; the recommendation algorithms on social network platforms can bias users' news consumption [35] and political opinions [23]. More positively, imitation theory and deterrence theory have been proven effective in encouraging positive behaviors and discouraging anti-social behaviors in online communities [39].

Our work adds to this area by identifying another example of how interface design in social media systems can introduce apparent bias in user behavior. While not an initial target question for this study, our user study shows that user interface plays a surprisingly strong role in encouraging inaccurate check-ins. We provide more details on these findings in Section 8.

4 RESEARCH QUESTION AND METHODS

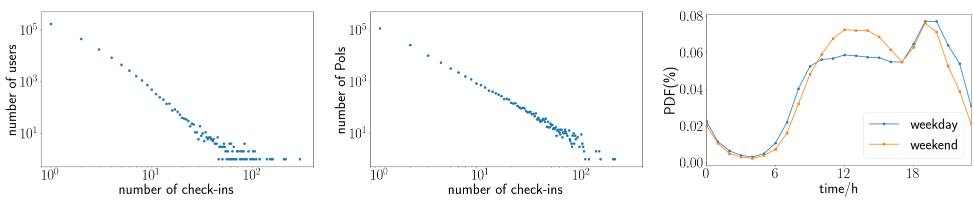
With the ever-increasing penetration rate of social media, the check-in services have become an important source to collect large scale mobility trace for wide range of research and applications. Previous works usually conduct researches based on the assumption that check-in data is an accurate and unbiased proxy to individual's actual mobility behavior. Although this assumption is fundamental to the scientific findings built on top of check-in data, the efforts made to validate it are inadequate. In fact, recent studies have revealed that there are significant bias and discrepancies between users' check-ins on Foursquare and their actual mobility traces, and the root cause is the external rewards provided by system [45, 53]. Therefore, an important question to ask is as follows:

Research Question 1: *Without explicit external rewards, will users misrepresent their check-ins in social media?*

Although the accuracy of check-in data plays an important role in downstream applications, a more interesting question to ask is why users will generate inaccurate check-ins in social media platforms, especially when external rewards are absent. Understanding the motivations behind inaccurate check-ins can not only help us improve users' experience in social media, but also provide insights on human's dishonest online behaviors. Such knowledge has potential to further our understanding in modelling users' behaviors in online communities, and encourage them to generate genuine interaction with each other in social media. Therefore, our further research question is:

Research Question 2: *What are the motivations behind users' inaccurate check-ins?*

To answer these research questions, we conduct our study in two steps. First, we collect a large scale empirical check-ins dataset, and conduct data-driven analyses to examine the presence of inaccurate check-ins and explore the potential motivations behind them. In addition, based on the



(a) Number of check-ins per user (b) Number of check-ins per PoI (c) Temporal distribution

Fig. 2. The statistical characteristics of the check-in trace.

empirical observations, we develop key hypotheses on the underlying motivations of inaccurate check-ins. Second, based on the proposed hypotheses, we design and carry out a user study, which allows us to properly evaluate the hypotheses and gain deeper insights on users' check-in behaviors.

5 DATASET

5.1 Data Collection

Through a close collaboration with Tencent incorporation, we were granted access to study the empirical check-in behavioral data on WeChat Moments platform. The dataset covers 276,346 users, which are randomly selected from the general user population spread across China. The data covers one month of usage, *i.e.*, from April 17, 2017 to May 17, 2017. It consists of two parallel traces for each user: a *check-in trace* and a *physical mobility trace*. The *check-in trace* records the detailed information of each check-in during the time period, including time stamp, the names and locations of check-in PoIs, and the lists of recommended PoIs that are displayed to the users. This dataset is large in that it contains more than 440,000 check-ins. The *physical mobility trace* contains more than 600 million physical location values collected during the same time period.

Whenever users invoke location based services, *e.g.* check-in to a location, WeChat's localization module records the user's physical location. Thus each check-in record has a corresponding ground truth record in *physical mobility trace*. So while the mobility trace portion of the dataset does not provide a continuous trace of the user's movements, it does provide an accurate matching record of the user's location during each check-in. Specifically, the localization module takes the signals of WiFi connection, GPS and base station connection as inputs and is able to determine the location with a error of a few meters in more than 90% of cases [19].

Points of Interest (PoIs) are classified into sixteen categories based on their functions, including *residence, shopping, business, education, food, tourist attractions, hotel, factory, transportation, life services, medical, landmark, institution, entertainment, culture* and *sports*. In addition, the Tencent online map service also provides the perimeter of the coverage area of each PoI [20]. When examining the accuracy of check-in events, we always consider the distance between the user's physical location (recorded by the localization module), and the edge of the PoI reported by the online map.

5.2 Ethical Considerations

We take careful steps to address privacy issues regarding the sharing and mining of user data. First, the Terms of Service for WeChat Moments include consent for research studies. Tencent, the parent company of WeChat, share user data with us after preprocessing the data to protect user privacy. All user identifiers have been replaced with secure hashcodes to improve anonymity. Second, our research protocol has been reviewed and approved by our local university institutional board. Third, all data is stored in a secure off-line server, with access limited to only authorized members of the research team bound by strict non disclosure agreements.

5.3 Basic Statistics of Dataset

To gain a basic understanding of the investigated dataset, we first delineate the distribution of number of check-ins posted by each user in Figure 2(a), and the number of check-ins per PoI in Figure 2(b). We can observe that both distributions follow a well-fitted power law, which indicates small fraction of users and PoIs cover a large portion of total check-ins. Figure 2(c) presents the temporal distribution of check-ins on weekdays and weekends. The number of check-ins on weekdays peak at around 8 PM, while more check-ins can be spotted during daytime on weekends than on weekdays, which both exhibit clear sleeping cycles. These observations are consistent with previous studies on check-ins of other social media platforms, such as Foursquare and Twitter [8], which indicates that our dataset is representative of typical check-in behaviors in social media.

6 ANALYZING EMPIRICAL CHECK-IN DATA

6.1 Classifying Inaccurate Check-ins

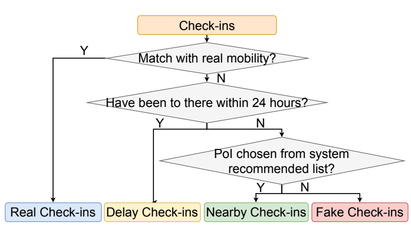


Fig. 3. The flow chart of check-in classification process.

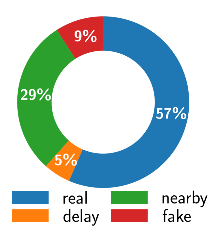


Fig. 4. Composition of overall check-ins.

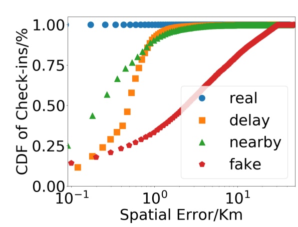


Fig. 5. Distribution of spatial error.

First, we check the accuracy of check-ins by comparing the perimeter of each posted PoIs' coverage area with the actual location coordinate of the corresponding users in *physical mobility trace*. Recall that the localization module introduces spatial error of less than 22.5 meters for more than 90% of cases [19], we set a threshold of 50 meters to accommodate for potential spatial errors introduced by WeChat's localization module. To be specific, if user's actual location is inside the posted PoI's coverage perimeter or within 50 meters to the edge of the perimeter, the check-in is classified as *real check-ins*, otherwise it is determined as an inaccurate check-in. Note that the threshold is designed to capture a lower bound of inaccurate check-ins, and we can be confident in our identification of inaccurate check-ins.

We further classify inaccurate check-ins into more fine-grained categories to better characterize the distinct features among them. The *physical mobility trace* captures high quality mobility records of users during the time period, where over 63% of users have more than 30 mobility records per day, and 95% of consecutive mobility records of each user have interval time of less than 1 hour. It facilitates us to analyze the delayed behavior in check-in service, which has been widely observed in human dynamics and social interactions [1, 42]. To be specific, we classify the inaccurate check-ins that match with users' trajectories within the previous 24 hours as the *delayed check-ins*, because users have physically visited the posted PoIs before. Furthermore, since there are two different approaches to post check-ins, *i.e.*, choosing from the app's recommended list or manually inputting by users, the inaccurate check-ins generated through different means are likely motivated by different reasons. Therefore, the rest of inaccurate check-ins that are generated through system recommended list are classified as *nearby check-ins*, since it always recommend nearby PoIs. Those generated through manual input are classified as *fake check-ins* since they are likely corresponding to deliberate misrepresentations. Finally, the flow chart of the classification process is shown in Figure 3. Formally, the definition of the four categories of check-ins are summarized as follows:

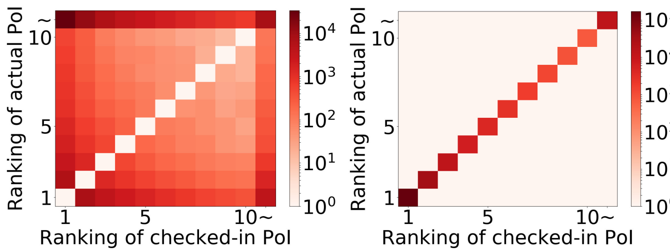
- *Real check-ins*: Users' actual location coordinates are inside the posted PoI's coverage perimeter, or within predefined distance from the edge of the perimeter.
- *Delayed check-ins*: At the check-in time, the user is beyond the predefined distance away from the perimeter of the posted PoIs's coverage area, but has visited the PoI within the previous 24 hours.
- *Nearby check-ins*: The user selects the PoI from the system recommended list to check-in, but is beyond the predefined distance away from the perimeter of the posted PoI's coverage area, and has not visited the PoI within the previous 24 hours.
- *Fake check-ins*: Similar to nearby check-ins, except that the user checks-in to the PoI through manual input.

The composition of the four categories of check-in is shown in Figure 4. Real check-ins account for 57% of total check-ins, while the rest 43% percent check-ins are inaccurate ones. The result indicates that there are surprisingly high percentage of inaccurate check-ins in social media platforms even without explicit rewards for generating check-ins. In addition, nearby check-ins cover 29% of total check-ins, being the largest component of inaccurate check-ins. Figure 5 demonstrates the cumulative distribution function of the spatial error between checked-in PoIs and *physical mobility trace*, which is measured by the distance between checked-in PoIs and the actual locations of the users. We can observe that real check-ins exhibit no spatial error. In addition, nearby check-ins have relatively smaller errors compared with other inaccurate check-ins, *i.e.*, more than 80% of them have spatial errors for less than 1 kilometer. On the other hand, fake check-ins exhibit the largest spatial errors, for over 25% of them have spatial errors larger than 10 kilometers.

6.2 Exploring Potential Motivations

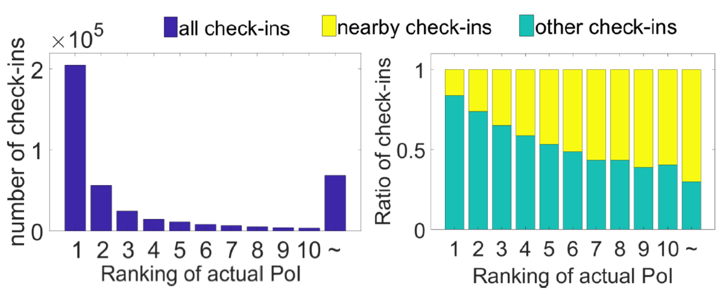
To reveal the underlying mechanisms behind inaccurate check-ins, we first explore the potential motivations through a data-driven approach. To be specific, we detect and characterize the distinct features of each category of inaccurate check-ins. Then, we analyze the features with established human behavior models, which help us derive the key hypotheses on the underlying motivations of inaccurate check-ins.

Nearby check-ins. First, we examine nearby check-ins, which correspond to the behavior that users select wrong PoIs from the recommended list, where the recommendation algorithm plays an important role. The app's recommendation algorithm first selects the PoIs within a predefined distance as candidates, and then ranks them based on the features of popularity, previous check-in history and other features. In addition, through associating the actual locations in physical mobility trace with the spatial coverage of PoIs, we are able to extract the actual PoIs users locate in when they post check-ins. To examine how users interact with the recommended list, we present the joint frequency distributions of the rankings of checked-in PoIs and actual PoIs in the recommended list as heatmaps in Figure 6, where deeper color indicates higher frequency and lighter color indicates lower frequency. For example, in Figure 6(a) the color in the slot of first column and second row means that there are about 3,000 nearby check-ins that checked-in PoIs rank first and actual PoIs rank second in system recommended list. From Figure 6(a), we can observe that most of the nearby check-ins distribute in the upper-left area, which suggests the actual PoIs have low rankings while the checked-in PoIs have high rankings in the recommended list. However, as for Figure 6(b), that



(a) Nearby check-in (b) Real check-in

Fig. 6. The heatmap of joint frequency distribution of the rankings of check-in PoIs and actual PoIs.



(a) Number of Check-ins (b) Percentage of Check-ins

Fig. 7. The number and percentage of nearby check-ins when actual PoIs are of different rankings.

most of the real check-ins distribute in the lower-left corner suggests that both the actual PoIs and check-in PoIs have high rankings in the recommended list.

To further investigate this feature, we present the number and the percentage of nearby check-ins when actual PoIs have different rankings in recommended list in Figure 7. From Figure 7(a), we can observe that in more than 200,000 check-ins the actual PoIs rank first in recommended list. However, the rankings of actual PoIs exhibit a fat-tailed distribution, where it is out of top-5 for more than 96,000 check-ins (*i.e.* 23.7% of total check-ins). Moreover, Figure 7(b) suggests that nearby check-ins are more likely to happen when the rankings of actual PoIs are low in recommended list. To be specific, the percentage of nearby check-ins gradually increases from 16% to 60% when rankings of actual PoIs increase from 1 to 10. These observations indicate strong correlation between the performance of recommend system and nearby check-in behaviors, where nearby check-ins are more likely to occur when the rankings of actual PoIs in the recommended list are low.

Similar behavior patterns have been discovered in multiple-choice question scenario [5]. The author proposed the term "position bias" to describe the bias caused by respondent's tendency in choosing the first few options regardless of their content. The term has later been extended to model user's click behavior in web searching, which predicts the click through rate of web pages decreases as their rankings in search engine decrease [57]. Moreover, the eye-tracking devices facilitate researchers to prove that users primarily go through the options top-to-bottom [7], and the allocated attention decreases significantly as the rankings decrease [22]. Researchers further find proof that users' attention can be manipulated by changing the presentation order of items [25]. Therefore, we propose the hypothesis on the reason of nearby check-ins as follows:

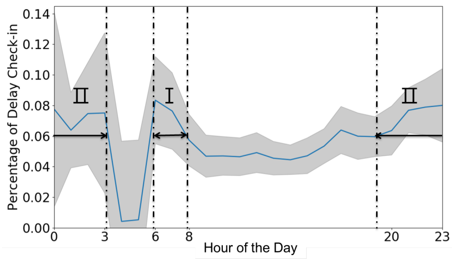


Fig. 8. Percentage of delayed check-ins in different hours of a day, where the blue line represents the average value and grey area represents standard deviation.

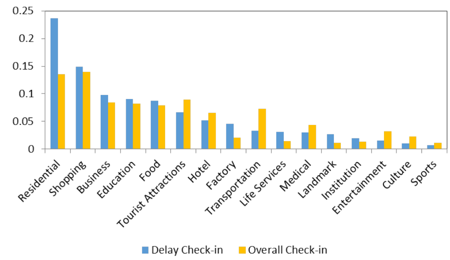


Fig. 9. The distribution of PoIs categories where delayed check-ins are initiated comparing with overall check-ins.

H1: *Nearby check-ins are generated primarily because the actual PoIs rank low in the recommended list and users only pay attention to the first few options.*

Delayed check-ins. Delayed check-ins refer to the behavior that users check-in to previously visited PoIs, where the temporal feature is an important aspect. To explore the potential motivations, we first examine the distribution of delayed check-ins on different temporal periods of a day. To be specific, we calculate the average percentage of delayed check-ins in overall check-ins of different periods of a day, and present them with the standard deviation in Figure 8. The figure reveals that there are two periods of high percentage of delayed check-ins, which are 6 AM to 8 AM and 8 PM to 3 AM. Compared with the daily schedule of a typical user, these time periods likely correspond to off-work hours.

Furthermore, we compare the distribution of the PoIs categories where the delayed check-ins are initiated with overall check-ins in Figure 9. Compared to overall check-ins, the delayed check-ins are more likely to take place when users are at PoIs of residential, shopping, business, education and food categories. PoIs of these categories usually have stable internet connection and users can conveniently access check-in services. These observations indicate that delayed check-ins have strong correlation with the locations and time periods, and users tend to delay their check-ins until the location and time are of their convenience. Previous works in human behavior modelling also indicated that human have the intent to prioritize their daily activities into queue and postpone the tasks of low importance [1]. Therefore, we propose a hypothesis as follows:

H2: *The primary motivation of delayed check-ins is that users tend to postpone the check-in behavior until they are in convenient locations and time periods, e.g., when they have stable internet connect or not busy.*

Fake check-ins. Our analysis reveals that fake check-ins can be best characterized by two different types of behaviors, which can be distinguished by features of the posted PoIs. The distinct features are the portion of fake check-ins and diversity of visitors of the posted PoIs. The diversity of visitors is measured by the number of users who have checked-in to the PoIs. We visualize the distribution of fake check-ins on these two features as a heatmap in Figure 10, where red color indicates more check-ins while blue color indicates fewer. The results demonstrate that most of the fake check-ins can be classified into two clusters, which are marked as *A* and *B*, respectively. Cluster *A* corresponds to the fake check-ins that associate with PoIs of low diversity, i.e., fewer than 3 users have checked-in to them. In addition, these posted PoIs have a surprisingly high portion of fake check-ins, where over 70% check-ins associate with them are fake check-ins. On the other

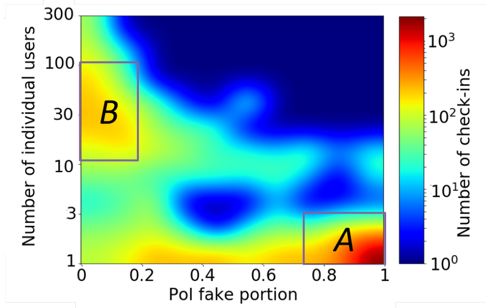


Fig. 10. The heatmap of fake check-ins' distribution on users and PoIs with different ratio of fake check-ins.

hand, the fake check-ins in cluster *B* are associated with PoIs of high diversity and low portion of fake check-ins.

To better understand the underlying motivations of these two behavior features, we present the PoI distribution of these two major clusters of fake check-ins in Figure 11. Comparing with overall check-ins, we observe that the fake check-ins in cluster *A* are significantly more likely to be associated with PoIs of business, life services, sports, shopping and food categories, and less likely to be the type of tourist attractions, residential, transportation, hotel, medical and culture. By manually going through a randomly sampled subset of the fake check-ins in cluster *A*, we find out that a significant amount of them are due to the advertisement of self-own businesses, such as grocery stores and life service shops. In fact, researchers have reported a growing trend of utilizing the check-in function in social media as a mean of advertising and location promotion [29, 33, 56]. Therefore, we derive a hypothesis as follows:

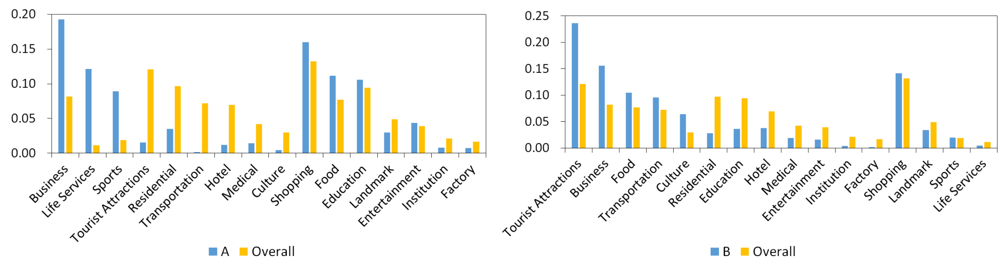
H3: *One primary motivation of fake check-ins is to advertise and promote the locations of users' businesses.*

On the other hand, Figure 11(b) demonstrates that the check-ins in cluster *B* have a significantly higher probability to be associated with PoIs in the categories of tourist attractions, business, food and culture, while have a lower probability to be associated with PoIs in the categories of residential, education, hotel, institution and factory. By manually examining the posted PoIs in these check-ins, we find out most of them are famous scenic sites, high-end working places and restaurants. One hypothesis is that users will claim to be in these locations to leave others an impression of interest or of high social status. This can be interpreted by self-presentation theory, which suggests people have the intention to manage others' impression of them and appear cool in community [14]. Prior work reports that self-presentation is becoming increasingly popular in users' online participation to social media [3, 17, 47]. Thus, we derive the final hypothesis as follows:

H4: *Self-presentation is another primary motivation for users to generate fake check-ins.*

7 USER STUDY

To properly evaluate the proposed hypotheses and further explore the mechanisms behind inaccurate check-ins, we design an online survey to collect users' self-reported reasons for misrepresenting check-ins. Finally, we also conduct a follow-up interview of a smaller user group to dive deeper into user motivations.



(a) PoI distribution of fake check-ins in *A* cluster compare to overall check-ins.

(b) PoI distribution of fake check-ins in *B* cluster compare to overall check-ins.

Fig. 11. The distribution of fake check-ins on different types of PoIs.

7.1 User Study Design

In our user survey, the designed questionnaire contains 17 questions in total, which can be summarized into two parts: validation questions and exploration questions.[1] To ensure the credibility of our survey, we add two validation questions, which are identical in meaning, but whose answer choices are arranged in different order. These two validation questions are spaced far apart in the survey to avoid detection. Furthermore, to make sure respondents are answering the questions carefully, we also check consistency across answers to identify users whose selections contain contradictions, *e.g.*, state they have never added fake check-ins while selecting in a later question choices for why they make fake check-ins. Finally, we set up a browser cookie examiner to prohibit multiple submissions from the same respondent.

We note that asking people's misbehavior is likely subject to social desirability bias: people may mask their dishonorable behaviors such as lying or cheating [44]. We mitigate this challenge with three steps. First, our questionnaire is self-administered, where anonymity can be more assured and the influence of social desirability can be lessened [30]. Second, referring to the method of proxy subjects [30], we design additional questions that ask participant's perception on the inaccurate check-ins of one of their typical friends, and examine their speculations on the underlying motivations. Participants' perceptions of inaccurate check-ins are thus measured through the procedure, which are regarded as a reflection of people's real thoughts. Third, when designing the questions, we avoid explicit expressions like "fake" or "casual" that could directly lead respondents to link specific actions with negative interpretations of user behavior. We deploy the survey on an online survey system similar to SurveyMonkey. Since the majority of Wechat's users are Chinese, the questionnaire is written in Chinese.

We also note that it is unreliable to ask people to recall their previous behaviors. Therefore, at the beginning of the questionnaire we ask participants to review their previous check-ins in WeChat Moments to build up a context. In addition, when it comes to specific questions concerning previous behaviors, we ask the participants to refer to the concrete instances in their previous check-in records. For example, when asking about their check-in frequency, we require participants to look back at their check-ins in WeChat Moments in the last year and then give an estimation.

We also conduct a semi-structured interview to supplement the survey study. We include a total of 6 seed questions concerning users' check-in habits, motivations behind their check-ins, reasons of inaccurate check-ins and observations on their friends' behaviors. The concrete instances and

[1]The full question list is available at https://www.dropbox.com/s/q5e0xa46gughq2x/Questionnaire.pdf?dl=0

specific reasons of their behaviors serve as our main focus. To diminish social desirability bias, we avoid using overtly negative expressions.

7.2 Recruiting Participants

For the survey study, we seek participants with an *improved snowball sampling method* [4]. Unlike original snowball sampling method, we set up several constraints to make the recruitment procedure controllable and minimize the bias shown as follows:

- First, we carefully select 40 initial participants, denoted by *root participants*, from researchers' acquaintances who have reasonable difference in identity background.
- Then, we give each of them a different link to our questionnaire, and ask them to recruit new participants with different backgrounds from their WeChat friends, which are denoted by *leaf participants*. They must follow the constraints that each root participant can only recruit 20 leaf participants at most and leaf participants can not recruit new participants.

For our user interviews, we send out interview invitations to the participants via emails after the survey. Then, we conduct on-site interviews on the participants that first accept our invitations within a time window.

7.3 Respondents and Verification

677 responses from all over China were received in our survey study, with about 17 leaf respondents per root respondent on average. We observe that the portion of legitimate responses drops dramatically when the finish time is less than 60 seconds (*i.e.* 54% finishing in 60 seconds, 89% otherwise), which suggests that responses completed in less than 60 seconds are most likely of poor quality. Thus, we screen the responses according to the following criteria: 1) passing the designed validation test, and 2) spending more than 60 seconds on the survey.

After screening, we obtain 541 valid responses from 28 provinces in China, with 222 (41%) male participants and 319 (59%) female participants. The participants cover a wide range of age groups, spanning from under 18 to above 60 years old. We compare the age distribution with overall WeChat users age distribution reported in WeChat global user report 2015, and find the Kullback-Leibler distance between them less than 0.04. Furthermore, the self-reported frequency of posting inaccurate check-ins in survey responses is consistent with the observations in our empirical check-ins dataset, where average frequency is 1.96 per month in survey and 1.78 per month in empirical check-ins data. In addition, we find that the distribution of self-reported check-in frequency is similar to our empirical observations. Specifically, the most active top 11.67% respondents contribute to 56.69% of the check-ins, and the 61.73% least active respondents post check-ins less than once a month. Our respondents include both power users and less active users. These observations indicate that the survey data is representative of typical WeChat users. Our interview study includes 8 participants, denoted by P1-P8, with 5 (62.5%) male participants and 3 (37.5%) female participants. Note that we focus on understanding inaccurate check-in behavior on per user basis in survey and interview study, which we believe is an important complementary to the empirical data analysis where most check-ins are contributed by small amount of power users.

8 FINDINGS

In this section, we analyze users' responses to examine the proposed hypotheses.

Impact of social desirability bias. We start with examining the impact of social desirability bias on the responses. To be specific, we compare respondents' self-reported inaccurate check-in behavior with the misbehavior they observe on one of their typical friends. Note that the observation on their friends will not be affected by the number of their friends, since we ask them to report the

Motivations	N	P
I have been to the place I checked in, but did not manage to post a check-in timely.	164	49.25%
For the sake of social interaction, for example, to impress my friends, I posted a check-in to a place I have never been to.	140	42.04%
To advertise my own business or to get some commercial benefits, such as food discount coupons or online game packages, I posted a check-in to a place I have never been to.	112	33.63%
I do not care if my check-in PoI matches my real location, so I selected a PoI randomly from the recommended list.	84	25.23%

Table 1. The most common motivations for posting inaccurate check-ins, and the number (N) and percentage (P) of participants who report the corresponding motivations.

behavior of one of their typical friends. We find that 134 (25%) respondents claim their friends will post inaccurate check-ins, but have not posted any themselves, while only 38 (7%) respondents report to have posted inaccurate check-ins themselves, but their friends will not conduct such behavior. The result is in line with [24] demonstrating people's tendency of self-concealment, *i.e.*, actively concealing those personal information will be perceived as negative. Respondents have the tendency to answer the questions in the manner that they will be viewed favorably by others. It suggests self-reported misbehavior could serve as a lower bound of their actual inaccurate check-ins, since we assume people will only try to hide their misbehavior but will not admit to misbehavior they have not committed.

Motivations for posting inaccurate check-ins. Respondents report a variety of reasons for posting inaccurate check-ins from both the options we provide and free-text responses. Table 1 presents the most popular motivations, where the number and the percentage of participants who report the corresponding motivations are sorted in descending order. Here each participant can report more than one motivation.

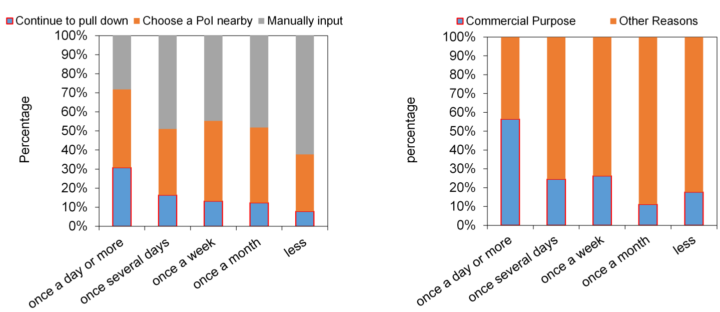
From Table 1, three key observations are obtained. First, we do observe a large portion of respondents who report to have posted at least one type of inaccurate check-ins themselves (N=333, 61%). This is consistent with our previous empirical findings, that even without external rewards, posting inaccurate check-ins is prevalent in social media. These answers shed light on *research question 1*. Second, the most common self-reported motivations support our previously proposed hypotheses. The top 4 self-reported motivations correspond to *H2*, *H4*, *H3* and *H1*, respectively. As is discussed, self-reported misbehavior should serve as a lower bound of the actual misbehavior. Therefore, the survey data indicates there are a significant number (at least 25%) of users exhibit each of the top 4 self-reported motivations, which quantitatively validates the proposed hypotheses. Third, interestingly, from respondents' perspectives the top 2 motivations for misrepresenting check-ins are delayed behavior and self-presentation, which are less likely to cause *nearby check-ins*. However, in the empirical check-in dataset, *nearby check-ins* account for the largest portion of inaccurate check-ins. It suggests that users are less aware of their *nearby check-ins* due to the nature of casual behavior. Sometimes they even do not know they have posted *nearby check-ins*.

8.1 Rationales for different types of inaccurate check-ins

To gain deeper understanding of users' motivations behind different types of inaccurate check-ins, we conduct correlation analysis on several important variables we observe through our survey. In addition, we dive deeper into the specific reasons that directly lead to different types of inaccurate check-ins, by asking the respondents more detailed and open-ended questions and through interviews. To examine the impact of social desirability bias, we also set parallel questions to ask respondents about similar misbehavior they observe on their typical friends.

Reasons	%(S)	%(F)
I want to convey an approximate message to my friends, like "I am back at XX city". So the accuracy of my check-in location is out of my consideration.	63.86%	68.79%
I just want to share my moments with a location, so I choose one at will.	60.24%	66.24%
I want to share my moments with a location, but in consideration of privacy, I will not choose my actual location to check-in.	46.99%	45.86%

Table 2. The specific reasons for posting nearby check-ins. For the tables below, "S" stands for self-report behavior and "F" stands for observation on friends.



(a) Reactions when actual PoI is not in recommended list VS check-in frequency.

(b) Advertising VS check-in frequency.

Fig. 12. The cross analysis of several important variables in survey data.

Nearby check-ins. According to hypothesis *H1*, we first look into the role of system recommended list in the inaccurate check-in behavior. To be specific, we ask users what they will do if the PoIs that they want to check-in are not found in the first page of the system recommended list. One of our key observations is that only a small portion of users (N=61, 11.28%) are willing to pull down the list to look for the desired PoIs, while a significant portion of respondents (N=181, 33.46%) prefer to choose a nearby PoI presented as the first few options of the list as an alternative. The result suggests that users are likely to generate *nearby check-ins* if the recommender system does not rank the actual PoIs in the top of recommended list, which supports the hypothesis *H1*.

Furthermore, we present the self-report specific reasons for generating nearby check-ins in Table 2. The two most popular reasons show the users do not care much about the precision of the location tags, which suggests they are more likely to check-in to the PoIs with high rankings. On the other hand, the third reason indicates that privacy is also a factor for nearby check-ins but of relatively smaller importance. In addition, Figure 12(a) presents the cross analysis between users' reactions when the PoI they want to check-in are not found and their frequency of using check-in service. We find out that the percentage of users who will pull down the list decrease as the frequency of posting check-ins decrease, which suggests inactive users are less likely to go through the recommended list to look for actual PoIs. These results indicate that the performance of recommender system plays a crucial role in users' nearby check-in behaviors, especially for less active users.

The interview study corroborates the aforementioned findings. For example, P1 answers: "Some of my WeChat Moments check-ins have nothing to do with my physical location. The reason I make some of my check-ins, is that the check-in list is shown to me and I just casually pick one from the first few rather than selecting nothing." This is in line with the finding that nearby check-ins are likely to be generated casually. In addition, nearby check-ins can also be attained intentionally. P1

Reasons	%(S)	%(F)
I have the habit of posting check-ins at my convenience after I organize the words posted with check-in.	57.06%	59.88%
For example, when I am at a meeting or without WiFi, it is not convenient to post a check-in without delay.	56.44%	61.08%
Because I do not come up with the thought of posting a check-in until I leave the place I finally post a check-in to.	53.99%	62.18%
In consideration of the protection of my privacy, I delay the check-in time consciously.	36.81%	42.51%
I want to use the check-ins to achieve certain purposes, such as pretending that I am at work, etc.	12.88%	28.14%

Table 3. The specific reasons for posting delayed check-ins.

Reasons	%(S)	%(F)
I want to advertise my own business or the business where I am working.	66.07%	77.13%
I want to get some commercial benefits, such as food discount coupons or online game packages.	58.92%	63.83%

Table 4. The specific reasons for posting fake check-ins for commercial purposes.

refers to checking-in to Hangzhou East Railway Station near the hotel he was at rather than the hotel itself, because he believes "checking-in at a better-known place nearby rather than my actual, less known, location means others will have a better idea of where I am." Similar motivations are also articulated by P3.

Delayed check-ins. Among users who self-report delayed check-in behavior, we ask them to describe the specific reasons of their behavior, which is presented in Table 3. We can observe that the top 2 reasons can be directly linked to hypothesis *H2*, since they express the intent on posting check-in at a more convenient time or location (*e.g.* when they get home). In addition, the other three reasons indicate users will also generate delayed check-ins simply because they forget about check-ins, want to preserve privacy, or intend to achieve some other means.

In our interview study, delayed check-ins are prevalent among the interviewees. The limitation of poor network signals (P2), the need to modify photos before sending out (P2, P7), the lateness of the ending of activity (P1) and the excitement of the activity (P3, P5, P8) all contribute to delays between users' activities and when they post to WeChat Moments. Despite these kinds of constraints, as long as "the activity is worth recording" (P5), people will still check-in later, thus producing a delayed check-in.

These results suggest that the portion of check-ins marked as delayed might be underestimated in the empirical data analysis. One possible reason is because the physical mobility trace does not constantly track the users. Some delayed check-ins will not have a matching location value in the physical mobility trace, leading them to be classified as inaccurate check-ins. In addition, the users may be more comfortable to admit to delayed check-in behavior, since they have been to the check-in locations.

Fake check-ins. Among the users who self-report to have posted fake check-ins for advertisement purposes, we find there exist two specific incentives: "advertising" and "getting commercial benefits like coupon" as shown in Table 4. In fact, "getting commercial benefits" could also be viewed as a service of location promotion that users trade to business owners. Therefore, the survey results are consistent with hypothesis *H3*. Moreover, we analyze the correlation between the advertising behavior and check-in frequency, and present them in Figure 12(b). We can observe that there is a sharp increase of the portion of users who generate fake check-ins for commercial purposes in the group of most active users. It indicates that active users are more likely to generate fake check-ins for commercial purposes.

Reasons	**%(S)**	**%(F)**
I want to make a good impression on my friends, for example, I have a very interesting social life.	48.91%	49.37%
I want to draw my friends' attention, or to keep or improve the relationships with my friends.	48.18%	43.04%
I want to give an impression to my friends on WeChat that matches with their impressions on me in reality.	44.53%	50.21%
I want to gain more "likes" from my WeChat friends.	**42.34%**	**62.03%**
I follow my friends as they do.	21.17%	34.60%
I want to set up a new social image among my friends.	18.25%	13.92%
I want to make up for the bad impression my friends had on me in the past.	14.60%	13.08%

Table 5. The specific reasons for posting fake check-ins for self-presentation.

On the other hand, among users who self-report to have posted fake check-ins for self-presentation, we summarize the specific reasons in Table 5. The most common self-reported reasons can be concluded as impression management and gaining their friends' attentions, which is consistent with hypothesis *H4*. The other reasons also include gaining virtual rewards (i.e. "likes") and simply following the actions of their friends.

Similar results are attained in the interview study as well. Motivations of commercial gain and self-presentation are observed in concrete instances of posting fake check-ins. P2 refers to one of her tour guide friends who checks-in at her business consistently to advertise her business, and P3 mentions a friend who checks-in only at the library no matter what he is doing, to leave an impression of diligence. What's more, some interviewees illustrate manually inputting "new" PoIs to check-in. P3 relates to his experience of manually inputting "Beijing West Railway Station South Square East" to check-in, the PoI of which combines the four directions altogether in Chinese. P1, P2, P4 and P6 mention their friends who prefer to check-in with self-invented locations like "XXX's bed" and "somewhere only 10 meters from the excavator". The main motivation shared by the interviewees of posting self-invented "new" PoIs is to show their sense and humour and appear cool in social media platforms.

Social desirability bias in different types of check-ins. Through comparing the self-reported misbehavior and participant's observations on their friends, we quantitatively evaluate the social desirability bias across different types of inaccurate check-ins. We observe that in most of the reasons for inaccurate check-ins the portion of self-reported participants is less than the portion of participants who claim to observe on their typical friends. This result is consistent with our previous assumption that users tend to conceal their misbehavior due to the social desirability bias. Furthermore, Table 6 presents Pearson correlation r and Kullback-Leibler distance (KL distance) D between the self-report behavior and participants' observations, where higher Pearson correlation and lower KL distance indicate the self-report behavior and observations are more consistent. Table 6 reveals that users' answers in self-presentation questions have the lowest Pearson correlation and highest KL distance, which indicates the social desirability bias is most significant in self-presentation questions.

	nearby check-in	delayed check-in	commercial purposes	self-presentation
r	0.986	0.995	1	**0.864**
D	0.024	0.002	0.001	**0.038**

Table 6. The consistency between self-reported behaviors and observations on friends.

Correlations between demographic factors and inaccurate check-ins. We aim to examine how the check-in behavior will deviate on users of different demographic groups, i.e. ages and genders. We make three key observations through the cross analyses. First, female users exhibit higher check-in frequency, where average check-in frequency is 2.17 for females and 1.66 for males. Second, the Welch t-test [11] result demonstrates that there is no significant difference between the inaccurate check-in behaviors of males and females. Third, the Welch t-test results show that users above 31 years old are significantly more likely to post fake check-ins out of commercial purposes, while users under 25 years old are significantly more likely to post fake check-ins for self-presentation.

9 DISCUSSION

Through our empirical data analysis and user study, we demonstrate that the discrepancy between check-ins and physical mobility trace is surprisingly high, even without check-in rewards or incentives. 43% of total check-ins do not match with physical location and 61% of survey participants self-report they have misrepresented check-ins before. In addition, our study reveals four primary causes for inaccurate check-ins: inappropriate design of suggestions in the check-in interface, delaying check-ins for convenience, advertising to achieve commercial benefits and self-presentation. These findings may provide important insights to the researchers who plan to exploit check-in data as well as developers for check-in services.

Implications for researchers. First, our work echos prior research efforts in quantifying the spatial discrepancies of check-in data [16, 45, 53]. Specifically, previous studies demonstrate the inaccurate check-ins are prevalent in LBSNs, where 75% of total check-ins do not match with physical mobility traces [53]. Researchers find evidence that the primary motivation of inaccurate check-ins are the external rewards provided by the Foursquare platform [45]. Our results augment prior results, and show that inaccurate check-ins are less prevalent, but still account for a significant portion of check-ins, even on platforms without rewards or incentives. 43% of overall check-ins exhibit more than 50 meters spatial discrepancies with physical mobility traces, while 21% of overall check-ins exhibit spatial discrepancies of over 500 meters. These findings confirm concerns on the increasingly common practice of modelling human mobility behavior with check-in data. For example, Cheng et al. utilize check-in data to explore the correlation between social relationship and human mobility [9], while Yuan et al. [50] and Huang et al. [18] leverage check-in in Twitter to discover spatiotemporal events and model regular human mobility patterns. We argue that utilizing check-in data without carefully considering the spatial discrepancies may result in bias in the scientific findings and hinder their reproducibility.

We also note that users often select nearby but inaccurate PoIs from system recommendations, which produce check-in events that are reasonably accurate in location but with incorrect PoI categories, *e.g.* 37% of the overall check-ins are posted with PoIs of wrong categories. This is not well studied by prior work [15, 45], and poses challenges to widely adopted practices that inferring semantics of check-in records based on associated PoI categories [31, 48, 49]. We strongly urge researchers to be more cautious about the limitation and potential bias of check-in data when leveraging them to model human dynamics.

Second, our work reveals more about underlying motivations of inaccurate check-in behaviors. Contrary to prior works, we find evidence that inaccurate check-ins are not solely incentivized by the external rewards, but also driven by more fundamental motivations: inappropriate design of suggestions in the check-in interface, delaying check-ins for convenience, advertising to achieve commercial benefits and self-presentation. Due to the position bias effect, users tend to choose the PoIs with high rankings in system recommended list regardless of their accuracy. Therefore, popular PoIs that rank high in recommended list are more likely to be associated with inaccurate check-ins. Due to the self-presentation tendency, users are more likely to post inaccurate check-ins with the PoIs that seem interesting, *i.e.* certain PoIs under tourist attractions, food and culture categories. Finally, users also use check-ins as a tool to promote their own businesses in social media.

Third, our study also contributes to the large research body of modelling online anti-social behaviors. Inaccurate check-in is a special form of deceptive behavior in online communities. Previous studies indicate that the deceptive behavior is a commonplace in individual's daily communications via text messaging [41]. Guha et al. further conduct a experience sampling study on the deceptive location disclosures in Foursquare [15]. Their work shows the deceptive location disclosures are closely correlated with physical distance, social tie strength and visibility of check-ins [15]. These findings are further rationalized by "butler lies" in unavailability management theory [34], where users intend to post deceptive check-ins that are less likely to be serendipitous encountered by their friends, *e.g.*, far away locations. Our findings indicate that inaccurate check-ins are not solely the results of deliberate misrepresentations, but also can be generated by unintentional behaviors, *i.e.*, casually selecting nearby PoIs from system recommendations. These findings also indicate that inaccurate check-in behaviors could be mitigated by improving the design of PoI tags recommend system, which echos previous research efforts that aim to discouraged anti-social behavior with better interface design [39].

Implications for check-in services developers. Our research proves that inaccurate check-ins are commonplace in general purpose social media platforms. Therefore, check-in services developers should not ignore these behaviors simply because their platform does not provide external incentives. Our findings suggest that these behaviors can be mitigated through modifications of current system designs.

First, our results suggest that recommendation mechanisms for points of interests have a significant impact on user check-ins: users who simply pick one of the highly ranked PoIs constitute 67.4% of total inaccurate check-ins. In our user study, only 11.28% of users are willing to pull down the system recommend list when their currently located PoIs are not in the first page. As an alternative to the current system that factors popularity and personal preferences, we suggest check-in developers should better tune their system to produce more accurate recommendations (possibly placing heavier emphasis on proximity to user), or otherwise modifying the user interface so that more PoI candidates can be offered at once. We note that prior results show that changing the mechanisms of check-in services could have negative influences on user engagement [45]. Therefore, developers must carefully strike a balance between user engagement and check-in quality.

Second, our study reveals that the inconvenience of posting check-ins is the main motivation of delayed check-ins. Our user study shows users typically want to delay check-ins until they have a more stable network, or until they have spare time or have finished modifying photos. One suggestion is to simplify the check-in process into a two stage process, a single one-button check-in to mark the time and physical location, with the rest of the check-in to be completed later. This is similar to Yelp's "review-later" functionality, and gives users a quick way to mark points of interest before they forget.

Limitations. We note several limitations of our work. First, the findings presented in our study are mainly derived from the population of Chinese users on one single platform, WeChat Moments. Obtaining large traces of check-ins with corresponding physical location data is challenging, and we are actively seeking secondary datasets to validate our findings. The results described here might be influenced by the context bias in terms of cultural background and social media platform. We leave more comprehensive cross-platform analysis to future work. Second, although we have taken steps to improve the breadth and diversity in recruiting survey participants, the *snowball* *sampling method* may produce sampling bias in our selection. However, the survey data produces conclusions consistent with those from large scale data analysis, adding further validation. Finally, since the physical mobility trace does not constantly track the users, the delayed check-ins might be underestimated in empirical data analysis, which is supported by survey results. However, the survey and interview provide a meaningful complementary study to make up the shortcoming of empirical data analysis.

10 CONCLUSION

In this paper, we seek to quantify and understand the discrepancies between the check-ins in social media and real physical locations. To achieve this goal, we first conduct a data-driven analysis on large scale empirical check-in data. Then, based on the empirical observations we develop key hypotheses on the motivations behind inaccurate check-ins, and carry out a detailed user study to further explore and validate them. Our study reveals that the discrepancy of check-in data is surprisingly severe even in the social media without external rewards for check-ins, where 43% of total check-ins do not match with real mobility and 61% of survey participants self-report they have misrepresented check-ins before. More importantly, the primary motivations behind inaccurate check-ins can be summarized into four factors: poorly designed interface in check-in apps, delaying the check-ins to convenient time and locations, advertising to achieve commercial benefits and self-presentations. We believe that this work provides insights on utilizing check-in data in downstream applications, and shed light on the design of social media platforms to encourage genuine online interactions.

ACKNOWLEDGMENTS

This work was supported in part by The National Key Research and Development Program of China under grant 2017YFE0112300, the National Nature Science Foundation of China under 61861136003, 61621091 and 61673237, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology.

REFERENCES

[1] Barabasi AL. 2005. The origin of bursts and heavy tails in human dynamics. *Nature* 435, 7039 (2005), 207.
[2] Jie Bao, Yu Zheng, David Wilkie, and Mohamed Mokbel. 2015. Recommendations in location-based social networks: a survey. *GeoInformatica* 19, 3 (2015), 525–565.
[3] Liad Bareket-Bojmel, Simone Moran, and Golan Shahar. 2016. *Strategic self-presentation on Facebook.* Elsevier Science Publishers B. V.
[4] Patrick Biernacki and Waldorf Dan. 1981. Snowball Sampling: Problems and Techniques of Chain Referral Sampling. *Sociological Methods Research* 10, 2 (1981), 141–163.
[5] Niels J Blunch. 1984. Position bias in multiple-choice questions. *Journal of Marketing Research* (1984), 216–220.
[6] Danah Michele Boyd. 2008. *Taken out of context: American teen sociality in networked publics.* University of California, Berkeley.
[7] Michael D Byrne, John R Anderson, Scott Douglass, and Michael Matessa. 1999. Eye tracking the visual search of click-down menus. In *SIGCHI.* ACM, 402–409.
[8] Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z Sui. 2011. Exploring millions of footprints in location sharing services. *ICWSM* 2011 (2011), 81–88.
[9] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *SIGKDD.* ACM, 1082–1090.
[10] Henriette Cramer, Mattias Rost, and Lars Erik Holmquist. 2011. Performing a check-in: emerging practices, norms and'conflicts' in location-sharing using foursquare. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services.* ACM, 57–66.
[11] Marie Delacre, Daniel Lakens, and Christophe Leys. 2017. Why psychologists should by default use WelchâĂŹs t-test instead of StudentâĂŹs t-test. *International Review of Social Psychology* 30, 1 (2017).

[12] Michael A DeVito, Jeremy Birnholtz, and Jeffery T Hancock. 2017. Platforms, people, and perception: Using affordances to understand self-presentation on social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 740–754.
[13] Michael A DeVito, Jeremy Birnholtz, Jeffery T Hancock, Megan French, and Sunny Liu. 2018. How People Form Folk Theories of Social Media Feeds and What It Means for How We Study Self-Presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 120.
[14] Goffrnan Erving. 1959. *The Presentation of Self in Everyday Life*. Doubleday,. 655 pages.
[15] Shion Guha and Stephen B Wicker. 2015. Spatial subterfuge: an experience sampling study to predict deceptive location disclosures. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1131–1135.
[16] B. Hecht and M. Stephens. 2014. A tale of cities: Urban biases in volunteered geographic information. In *ICWSM*. AAAI.
[17] Bernie Hogan. 2010. The presentation of self in the age of social media: Distinguishing performances and exhibitions online. *Bulletin of Science, Technology & Society* 30, 6 (2010), 377–386.
[18] Qunying Huang and David W. S. Wong. 2015. Modeling and Visualizing Regular Human Mobility Patterns with Uncertainty: An Example Using Twitter Data. *Annals of the Association of American Geographers* 105, 6 (2015), 1179–1197.
[19] Tencent Incorporation. 2017. Introduction to Tencent location based services. Retrieved April. 3, 2018 from http://lbs.qq.com/guides/loc.html
[20] Tencent Incorporation. 2017. Tencent map service. Retrieved April. 3, 2018 from http://map.qq.com/
[21] Kokil Jaidka, Sharath Chandra Guntuku, Anneke Buffone, H Andrew Schwartz, and Lyle Ungar. 2018. Facebook vs. twitter: Differences in self-disclosure and trait prediction. In *Proceedings of the International AAAI Conference on Web and Social Media*.
[22] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. 2005. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR*. Acm, 154–161.
[23] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 417–432.
[24] Dale G Larson and Robert L Chastain. 1990. Self-concealment: Conceptualization, measurement, and health implications. *Journal of Social Clinical Psychology* 9, 4 (1990), 439–455.
[25] Kristina Lerman and Tad Hogg. 2014. Leveraging position bias to improve peer recommendation. *PloS one* 9, 6 (2014), e98914.
[26] Q Vera Liao, Wai-Tat Fu, and Markus Strohmaier. 2016. # Snowden: Understanding Biases Introduced by Behavioral Differences of Opinion Groups on Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3352–3363.
[27] Janne Lindqvist, Justin Cranshaw, Jason Wiese, Jason Hong, and John Zimmerman. 2011. I'm the mayor of my house: examining why people use foursquare-a social-driven location sharing application. In *SIGCHI*. ACM, 2409–2418.
[28] Tanushree Mitra, Graham P Wright, and Eric Gilbert. 2017. A parsimonious language model of social media credibility across disparate events. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 126–145.
[29] Dmitry Namiot and Manfred Sneps-Sneppe. 2011. A new approach to advertising in social networks - business-centric check-ins. In *International Conference on Intelligence in Next Generation Networks*. 92–96.
[30] Anton J Nederhof. 1985. Methods of coping with social desirability bias: A review. *European Journal of Social Psychology* 15, 3 (1985), 263–280.
[31] Anastasios Noulas, Salvatore Scellato, Cecilia Mascolo, and Massimiliano Pontil. 2011. Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks. *The social mobile web* 11, 2 (2011).
[32] Kunwoo Park, Ingmar Weber, Meeyoung Cha, and Chul Lee. 2016. Persistent sharing of fitness app status on twitter. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 184–194.
[33] Dedi Rianto Rahadi and Leon Andretti Abdillah. 2013. The utilization of social networking as promotion media (Case study: Handicraft business in Palembang). *Seminar Nasional Sistem Informasi Indonesia* 37, 1 (2013), 43–42.
[34] Lindsay Reynolds, Madeline E Smith, Jeremy P Birnholtz, and Jeff T Hancock. 2013. *Butler lies from both sides: actions and perceptions of unavailability management in texting*. 769–778 pages.
[35] Filipe N Ribeiro, Lucas Henriqueo, Fabricio Benevenutoo, Abhijnan Chakraborty, Juhi Kulshrestha, Mahmoudreza Babaei, and Krishna P Gummadi. 2018. Media Bias Monitor: Quantifying Biases of Social Media News Outlets at Large-Scale. (2018).

[36] Mattias Rost, Louise Barkhuus, Henriette Cramer, and Barry Brown. 2013. Representation and communication: Challenges in interpreting large social media datasets. In *CSCW*. ACM, 357–362.
[37] Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, and Andrew T Campbell. 2011. Nextplace: a spatio-temporal prediction framework for pervasive systems. In *International Conference on Pervasive Computing*. Springer, 152–169.
[38] Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, and Cecilia Mascolo. 2011. Socio-spatial properties of online location-based social networks. *ICWSM* 11 (2011), 329–336.
[39] Joseph Seering, Robert Kraut, and Laura Dabbish. 2017. Shaping pro and anti-social behavior on twitch through moderation and example-setting. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 111–125.
[40] Bryan Semaan, Bryan Dosono, and Lauren M Britton. 2017. Impression Management in High Context Societies:'Saving Face'with ICT. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 712–725.
[41] Madeline E. Smith, Jeffrey T. Hancock, Lindsay Reynolds, and Jeremy Birnholtz. 2014. Everyday deception or a few prolific liars? The prevalence of lies in text messaging. *Computers in Human Behavior* 41, C (2014), 220–227.
[42] J Stehle, A Barrat, and G Bianconi. 2010. Dynamical and bursty interactions in social networks. *Physical Review E Statistical Nonlinear Soft Matter Physics* 81, 3 Pt 2 (2010), 035101.
[43] Dan Tasse, Zichen Liu, Alex Sciuto, and Jason I Hong. 2017. State of the Geotags: Motivations and Recent Changes.. In *ICWSM*. 250–259.
[44] Catalina L Toma and Jeffrey T Hancock. 2010. Reading between the lines: linguistic cues to deception in online dating profiles. In *CSCW*. ACM, 5–8.
[45] Gang Wang, Sarita Yardi Schoenebeck, Haitao Zheng, and Ben Y Zhao. 2016. " Will Check-in for Badges": Understanding Bias and Misbehavior on Location-Based Social Networks.. In *ICWSM*. 417–426.
[46] Yi-Chia Wang, Moira Burke, and Robert Kraut. 2016. Modeling self-disclosure in social networking sites. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 74–85.
[47] Yi Chia Wang, Hayley Hinsberger, and Robert E Kraut. 2016. Does Saying This Make Me Look Good?: How Posters and Outsiders Evaluate Facebook Updates. (2016), 125–129.
[48] Mao Ye, Dong Shou, Wang-Chien Lee, Peifeng Yin, and Krzysztof Janowicz. 2011. On the semantic annotation of places in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 520–528.
[49] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wen-Ning Kuo, and Vincent S Tseng. 2012. Urban point-of-interest recommendation by mining user check-in behaviors. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*. ACM, 63–70.
[50] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat-Thalmann. 2013. Who, where, when and what: Discover spatio-temporal topics for Twitter users. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 605–613.
[51] Muhammad Bilal Zafar, Parantapa Bhattacharya, Niloy Ganguly, Saptarshi Ghosh, and Krishna P Gummadi. 2016. On the wisdom of experts vs. crowds: discovering trustworthy topical news in microblogs. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 438–451.
[52] Chao Zhang, Lidan Shou, Ke Chen, Gang Chen, and Yijun Bei. 2012. Evaluating geo-social influence in location-based social networks. In *CIKM*. ACM, 1442–1451.
[53] Zengbin Zhang, Lin Zhou, Xiaohan Zhao, Gang Wang, Yu Su, Miriam Metzger, Haitao Zheng, and Ben Y Zhao. 2013. On the validity of geosocial mobility traces. In *Hotnets*. ACM, 11.
[54] Xuan Zhao, Cliff Lampe, and Nicole B Ellison. 2016. The social media ecology: User perceptions, strategies and challenges. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 89–100.
[55] Yu Zheng. 2011. Location-Based Social Networks: Users.
[56] Wen Yuan Zhu, Wen Chih Peng, Ling Jyh Chen, Kai Zheng, and Xiaofang Zhou. 2016. Exploiting Viral Marketing for Location Promotion in Location-Based Social Networks. *TKDD* 11, 2 (2016), 25.
[57] Zeyuan Allen Zhu, Weizhu Chen, Tom Minka, Chenguang Zhu, and Zheng Chen. 2010. A novel click model and its applications to online advertising. In *ICWSM*. ACM, 321–330.

Received April 2018; revised July 2018; accepted September 2018