A Method of Sanitizing Privacy-Sensitive Sequence Pattern Networks Mined From Trajectories Released

Haitao Zhang, Nanjing University of Posts and Telecommunications, Nanjing, China
Yunhong Zhu, Nanjing University of Posts and Telecommunications, Nanjing, China

ABSTRACT

Mobility patterns mined from released trajectories can help to allocate resources and provide personalized services, although these also pose a threat to personal location privacy. As the existing sanitization methods cannot deal with the problems of location privacy inference attacks based on privacy-sensitive sequence pattern networks, the authors proposed a method of sanitizing the privacy-sensitive sequence pattern networks mined from trajectories released by identifying and removing influential nodes from the networks. The authors conducted extensive experiments and the results were shown that by adjusting the parameter of the proportional factors, the proposed method can thoroughly sanitize privacy-sensitive sequence pattern networks and achieve the optimal values for security degree and connectivity degree measurements. In addition, the performance of the proposed method was shown to be stable for multiple networks with basically the same privacy-sensitive node ratio and be scalable for batches of networks with different sensitive nodes ratios.

KEYWORDS

Influential Nodes, Location Privacy Inference Attacks, Privacy-Sensitive Sequence Pattern Network, Sanitized, Trajectory

INTRODUCTION

With a widespread use of GPS (Global Positioning System) positioning devices in automotive and terminal equipment in addition to the fast development of social networks and location-based services, industry sectors can collect and store large amounts of trajectories in a variety of ways (Zhu, Zheng., & Wong, 2019), meaning that this type of data grows rapidly in daily life (Giannotti, 2011; Williams, Thomas, Dunbar, Eagle, & Dobra, 2015; Dobra, Williams, & Eagle, 2015). Analyzing trajectories using data mining tools can discover interesting patterns and regularities, which will help to provide auxiliary decisions for relevant industry applications (Gabrielli, Fadda, Rossetti, Nanni, Piccinini, Pedreschi et al., 2018; Blondel, Decuyper, & Krings, 2015), promote personalized medical care and precision marketing. In addition, trajectory data as a new type of data can also assist scientific workers to carry out intelligent transportation (Kujala, Aledavood, & Saramäki, 2016), urban planning (Louail, Lenormand, Ros, Picornell, Herranz, Friasmartinez et al., 2014; Li, Sun, Cao, He, & Zhu, 2016) and other research works (Ortale, Ritacco, Pelekis, Trasarti, Costa, Giannotti et al., 2008).

As technologies are intended to be neutral, they harbor neither benevolent nor malevolent intent with respect to the individuals using them. In particular, a curious or malicious user can also use the trajectory data mining tools to find non-interesting patterns. Specifically, this can include privacy-
sensitive mobility patterns (i.e., mobility patterns involve privacy-sensitive spatial regions, such as military restricted areas, religious sites, private houses, private clubs, red-light-district, etc.), which will pose a threat to the location privacy of specific users (Giannotti & Pedreschi, 2008; de Montjoye, Hidalgo, Verleysen, & Blondel, 2013). Privacy-preserving data mining outsourcing (Liu, Wang, Shang, Li, & Zhang, 2017; Monreale, Rinzivillo, Pratesi, Giannotti, & Pedreschi, 2014) and privacy-preserving distributed data analytics (Monreale, Rinzivillo, Pratesi, Giannotti, & Pedreschi, 2014) are two methods to ensure that privacy-sensitive patterns are not detected by attackers in systems with trusted central servers and untrusted central servers, respectively. While, when a collector (i.e., location service provider) of trajectories wants to release (i.e., publish and share) the trajectories with a third party, the sanitization methods based on the strategy of knowledge hiding will be adopted, that is, (s)he must sanitize the trajectories to eliminate privacy-sensitive mobility patterns to prevent a threat to the privacy of the users whose trajectories were collected.

The existing sanitization methods for privacy-sensitive mobility patterns mainly aim to hide privacy-sensitive mobility patterns (Rajesh, Sujatha, & Lawrence, 2017; Bonchi & Ferrari, 2010; Aggarwal & Yu, 2008), while changing the original trajectories as little as possible. In addition, these methods are specified to certain types of mining techniques, which include association rule hiding (Tsai, Wang, Song, & Ting, 2016), sequence pattern hiding (Quang, Tai, Huynh, & Le, 2016), sequence rule hiding (Zhang, Wu, Chen, Liu, & Zhu, 2017) and so on.

However, these sanitization methods (Tsai, Wang, Song, & Ting, 2016; Quang, Tai, Huynh, & Le, 2016; Zhang, Wu, Chen, Liu, & Zhu, 2017) cannot effectively prevent location privacy inference attacks based on analyzing the relationship between privacy-sensitive mobility patterns. Specifically, an attacker can connect some single privacy-sensitive mobility patterns to construct a privacy-sensitive mobility pattern network and perform location privacy inference attacks based on the network connectivity analysis. In fact, there is a high probability of the occurrence of attacks on the privacy of locations based on privacy-sensitive mobility patterns, as researchers are more likely to study human mobility patterns from a network view. For example, previous studies (Cho, Myers, & Leskovec, 2011; Nguyen, & Szymanski, 2012) suggested highlighting to what extent human movements affect social dynamics and how social interactions influence the way people move, which was achieved by studying the interplay between human mobility networks and social networks. Furthermore, by exploring the interplay between human mobility networks and social networks, macroscopic characteristics of many complex geographic systems, such as traffic (Bajardi, Poletto, Ramasco, Tizzoni, Colizza, & Vespignani, 2011), energy (Louail, Lenormand, Picornell, Cantú, Herranz, Friasmartinez et al., 2015) and population (Balcan, Colizza, Gonçalves, Hu, Ramasco, Vespignani et al, 2009; Brockmann, & Helbing, 2013), can be also discovered as a mobility pattern network is an abstraction of spatial topological relations of a complex geographical system.

Therefore, it is necessary from a network view to analyze inference attacks based on the privacy-sensitive mobility patterns and design the corresponding countermeasures. In this paper, we aimed to study the location privacy inference attacks based on the privacy-sensitive sequence pattern network mined from trajectories and to design a method of sanitizing the privacy-sensitive sequence pattern network.

The remainder of this article is organized as follows. Section Preliminaries provides necessary preliminary information and the basic concepts utilized in our research. In Section Location Privacy Inference Attacks Based on Privacy-Sensitive Sequence Pattern Network, we define privacy attacks formally. In Section Proposed Sanitization Method, we present a method of sanitizing privacy-sensitive sequence pattern networks for defending against these attacks. In Section Experiments and discussion, we describe our comprehensive experiments and provide an analysis of the results. Section Conclusions and future work concludes the paper and discusses further work.
Adaptation of Error Adjusted Bagging Method for Prediction
Selen Yılmaz Isikhan, Erdem Karabulut, Afshin Samadi and Saadettin Klçkap (2019).
*International Journal of Data Warehousing and Mining* (pp. 28-45).
[www.igi-global.com/article/adaptation-of-error-adjusted-bagging-method-for-prediction/228936?camid=4v1a](www.igi-global.com/article/adaptation-of-error-adjusted-bagging-method-for-prediction/228936?camid=4v1a)

Power System Load Frequency Control Using Combined Intelligent Techniques
[www.igi-global.com/chapter/power-system-load-frequency-control/42365?camid=4v1a](www.igi-global.com/chapter/power-system-load-frequency-control/42365?camid=4v1a)