

Using the Sensitive Mobility Technology of Spark Big Data Platform to Purify the Complex Network Environment of Campus Network

Zhang Meijun

Department of Economics and Management Sichuan TOP IT Vocational Institute, Chengdu, Sichuan, 611743, China

email: zhangmeijun666666 @163.com

Keywords: Sensitive Mobile Knowledge Network, Complex Network, Data Purification

Abstract: When the mobility rules in orbit data cross the confidential information of external data, it will bring privacy threat to users who meet the precondition of mobility rules. Traditional research on privacy protection of high-sensitivity mobility rules adopts a method based on the analysis of sensitive mobile networks to eliminate the high-sensitivity knowledge of a single mobility model that cannot effectively resist inference attacks. At the same time, the current privacy protection methods for social networks are limited to identification attacks, and they can not deal with the mobility attacks of sensitive mobile knowledge networks. Therefore, this paper proposes a purification method to deal with inference attacks on sensitive mobile knowledge networks. In this paper, users of communication big data focus on the mobile track data. First of all, this paper studies the method of acquiring mobility knowledge from the data mining of communication track and building a mobility weighted knowledge network with high confidentiality. Secondly, the perceptual component with reasoning attack mode and script based on knowledge network are analyzed. Finally, based on the super big data platform, the algorithm of purifying the mobile mode with high confidentiality in the network, aggravating the network composition, determining the network type, identifying the key nodes, and evaluating the house supply and security is implemented.

1. Introduction

With the rapid development of the Internet, cloud computing, Internet and intelligent terminals and data are in the explosive growth stage. China's total data volume is growing at an average annual rate of 40%, and is expected to reach 21% of the global total by 2020 [1]. China is a big data resource country. For today's countries, data is a strategic resource, and its position is equivalent to the oil resources in the industrial era. It is one of the standards to measure a country's comprehensive national strength. For enterprises, big data is their core competitiveness, which determines their long-term development. The coverage of communication market is generally very high, about 99%. From a monopoly point of view, countries generally have three or four operators. In addition, each of the three operators in China has more than 2 million participants. When users make calls, they send text messages, chat online, play games or navigate, and generate data every 1 second. Even if the user connects the mobile phone terminal to the network of the operator, data such as call time, billing, location and mobile speed can also be generated [2]. Moreover, only telecom operators have the ability to collect such a large amount of data related to user behavior.

2. Sensitive Mobile Knowledge Network

2.1 Simple Mobility Knowledge

At present, most of the research on mobility knowledge is from the point of view of spatiotemporal database in the field of computer science [3]. Association rules and sequence pattern mining are used to discover the regularity between different item sets of sequence database (the regularity between nodes in the path). The goal of mining related rules is to find the parallel

relationship between data in the transaction database. It does not include the order of nodes, that is, the mining is based on all tracks that users pass. Therefore, mining association rules does not consider the order of positions. In the research of track data based on communication data, the order of users' moving position is very important for analysis. Therefore, in this draft, although the mining method of association rules is not used, the mining results of sequential patterns are processed as simple mobile knowledge. Because complex mobility knowledge can be formed by combining simple mobility knowledge, the complex mobility knowledge of track data is combined according to the mining results of sequential patterns.



Figure 1 Mobile knowledge network

2.2 Network Model

Academic circles at home and abroad have been exploring the formation mechanism and development direction of the actual network. Researchers continue to use the complex network model for deduction, through the study of the structural characteristics of the network model, found the unique statistical law and development mechanism of the network [4]. In the classical network model, random network, small world network, scale-free network and BBV network. Erdo's others. First of all, we establish a stochastic network model, assuming that there are n nodes in the network. The edge probability between each node pair is p , that is to say, a random network is formed. Probabilistic networks are the theoretical basis of the development of complex networks, but most of them are inconsistent with the characteristics of actual networks. As a result, the network is not completely randomized, so it is difficult to solve the practical problems. Watts et al. Proposed a small world model. The concept of small world describes the basic phenomenon of complex network: large network (large number of nodes), and represents large cohesion coefficient, which is between any two nodes. Keep a relatively short path, that is, the average shortest path is smaller. The most popular form of expression in the small world is the concept of six degrees of separation [5]. Watt found that as long as there are six acquaintances in the United States, these six people can connect with each other. The small world is characterized by a complex network that can almost be expressed: for example, the average three stars of Hollywood actors are cooperative relationships, or the chemical substances of cells are usually divided into three different reactions. Small world network model is a network with high integration characteristics and small average shortest path. The free network model of Skell is proposed by Barabasi et al. The concept of scale-free only refers to some links, which means that most nodes in the network, some nodes have multiple links with other nodes, and are allocated multiple times. The form of the formula follows the power distribution. Several nodes that often maintain links are often referred to as "distribution nodes," which can contain thousands, tens of thousands, or millions of links. Because the network containing "distribution nodes" has no obvious characteristics in node length, it can be considered as scale-free network. The BBV model proposed by Barrat et al. Is a weighted network model. From the perspective of weight, it is the most influential research model [6]. It continues to use the functionality of the skelefree network model, but adds weight to each edge. When the connection between nodes is updated, the weight of the edge changes accordingly. For the complex network represented by BBV model, the order and edge weight distribution follow the power rule.

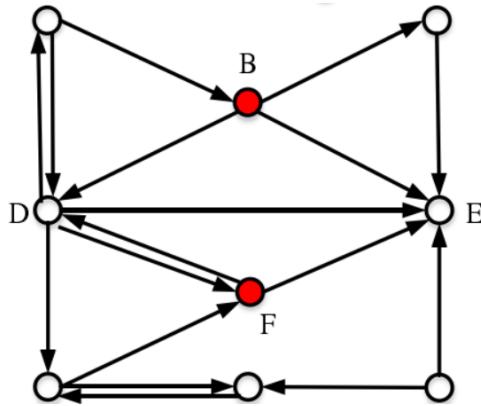


Figure 2 Topological abstraction of psspn

3. Spark Introduction

Spark system is a general open source high-speed parallel computing framework developed by researchers from Berkeley Institute, University of California [7]. The main features are as follows. Spark has the same advantages as Hadoop, and it is distributed according to map algorithm. However, unlike Hadoop, the intermediate output and calculation results of jobs can be stored in memory without having to read and write HDFS repeatedly. Therefore, the direction of spark reduction algorithm based on repeated mapping has a strong applicability. Iterative learning is usually carried out in the direction of machine learning, data mining or other algorithms. Therefore, at this point, Spark's performance is about 100 times higher than Hadoop's [8]. Develop spark Scala language and implement it with necessary functions. Scala is a new language based on Java virtual machine. This provides a reliable and fast environment for data processing. Spark and scala are unified, so spark can be generally used as the basic framework of applications. Developers can use Scala to simply set data, and local data objects run on the same distributed database can be written. At the same time, spark also supports Java, Python and other languages to provide convenience for users. Spark has a rich API interface, which makes it easy for programmers to implement related algorithms. In general, the code that implements the same function in a spark is Hadoop's code 1 / 10 or 1 / 100. In fact, the time spent developing code by programmers is directly proportional to the amount of code, so the development efficiency can be greatly improved. In a word, this paper chooses spark technology, which is based on the purification algorithm design of spark based location recognition privacy inference attack. First of all, it provides a large number of API interfaces, which can quickly develop related algorithms.

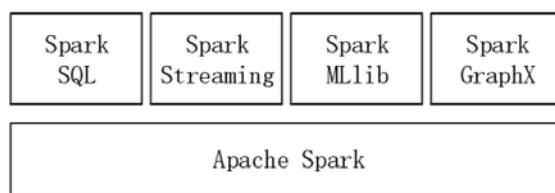


Figure 3 Spark ecosystem

4. Conclusion

With the rapid development of communication and satellite positioning technology, the data volume of electric power communication has increased dramatically, and it is also facing a huge data privacy threat era. Therefore, the purpose of this study is to use a big data platform with high reliable rules from the perspective of highly reliable mobile weighted knowledge network. In this paper, a method of constructing a high sensitivity mobility weighted knowledge network is proposed after the track data in the communication data is graphized in turn [9]. Based on the

identification of the confidential information of mobile geographic network, mobile knowledge network is transformed into sensitive network. Then, the edge weight data is used to transform it into a mobile weighted knowledge network of interest. In this paper, the establishment of mobile reinstallation knowledge chart, the ranking of golden key center nodes, and the new chart after node deletion are used to build. The high-sensitivity mobility component is accompanied by knowledge network in order to encapsulate the purification algorithm, and suggestions for the use of super big data computing platform. As well as the effectiveness and safety of purification assessment and other technologies. The existing knowledge law and relatively hidden, this paper, with a limited model and social network, but the analysis needs of the era of big data, suitable for mobile to aggravate the knowledge network point of view from the privacy attack. Aiming at the research content, through the experiment, it is verified that the joint method destroys the tight interaction between the node and the network, and keeps the complex "small world" characteristic in the continuous refining process [10]. At the same time, this method can completely resist the inference attack of sensitive mobility weighted knowledge network, because it can gradually eliminate the sensitivity of the network. It is reported that the proposed method is suitable for network structures with different sensitivity and complexity. As the premise of adjusting the influencing factors, the balance effect of utilization performance and safety after network refining can be achieved. The research results of this paper are very important for promoting the construction of information security platform in the field of information communication, and promoting the privacy protection through the geographic knowledge mining and analysis in the field of geographic information science.

References

- [1] Jo Rycroft-Malone. From Linear to Complicated to Complex Comment on "Using Complexity and Network Concepts to Inform Healthcare Knowledge Translation". International Journal of Health Policy & Management, vol. 7, no. 6, pp. 566-568, 2018.
- [2] Wenzhang Guo, Hao Wang, Zhengping Wu. Robustness analysis of complex networks with power decentralization strategy via flow-sensitive centrality against cascading failures. Physica A Statistical Mechanics & Its Applications, vol. 494C, no. 2018, pp. 186–199, 2017.
- [3] Hill Steven T, Kuintzle Rachael, Teegarden Amy,. A deep recurrent neural network discovers complex biological rules to decipher RNA protein-coding potential. Nucleic Acids Research, no. 16, pp. 16, 2018.
- [4] Yi Wang, Yang Long, Dun Zhang. Facile in Situ Growth of High Strong BiOI Network Films on Metal Wire Meshes with Photocatalytic Activity. Acs Sustainable Chemistry & Engineering, vol. 5, no. 3, 2017.
- [5] Juan Han,, Yun Wang,, Tong Chen,. Heat-induced coacervation for purification of Lycium barbarum, polysaccharide based on amphiphilic polymer–protein complex formation. Canadian Journal of Chemistry, vol. 95, no. 8, 2017.
- [6] Anoop Kumar Saini, Mohit Saraf, Pratibha Kumari,. A highly selective and sensitive chemosensor for L-tryptophan by employing a Schiff based Cu(II) complex. New Journal of Chemistry, vol. 42, 2018.
- [7] Elham Afzali, Hamid Forootanfar, Touba Eslaminejad,. Enhancing purification of α -amylase by superparamagnetic complex with alginate/chitosan/ β -cyclodextrin/TPP. Biocatalysis and Biotransformation, vol. 37, no. 3, pp. 1-9, 2018.
- [8] Virginia E Glazier, Thomas Murante, Daniel Murante,. Genetic analysis of the *Candida albicans* biofilm transcription factor network using simple and complex haploinsufficiency. Plos Genetics, vol. 13, no. 8, pp. e1006948, 2017.
- [9] Kapadia Carl, Chang Albert, Sotiropoulou Georgia,. Human Kallikrein 13: Production and

Purification of Recombinant Protein and Monoclonal and Polyclonal Antibodies, and Development of a Sensitive and Specific Immunofluorometric Assay. Clinical Chemistry, no. 1, pp. 1, 2020.

[10] Hutchinson Shirley, Luo Liu-Ying, Yousef George M., Purification of Human Kallikrein 6 from Biological Fluids and Identification of its Complex with α 1-Antichymotrypsin. Clinical Chemistry, vol. 5, pp. 5, 2020.