



International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Big Mobile Data Mining Challenges and Future A Review

Kausar Jahan

M.C.A – Millennium Institute of Management

Dr. Babasaheb Ambedkar Marathwada University Aurangabad, Maharashtra, India

In this technological era data sets are increasing every minute and so the Big Data has become integral part of diverse sectors from social media to business and corporate industries, medicine, astronomy etc. Data mining techniques are used to process this data to extract useful information. Big Mobile Data provide opportunities to Data miners to discover unseen information and patterns of knowledge which will be useful to understand human behavior and nature with historical as well as real time data.

Keywords— Big data, Data mining, Mobile data, Challenges, Internet of things

I. INTRODUCTION

We are surrounded by massive data sets which could be structured or unstructured. Increased use of computers and cost effective data storage devices are may be the reason for this voluminous data sets. These large structured or unstructured data sets for which traditional data processing tools are inadequate are known as Big Data. Data can be in the form of image, video, e-mail or records maintained in ledger etc. big data doesn't refer to any specific quantity or specific type of data. Big data is from diverse heterogeneous sources and in different format, this data is autonomous and there is no centralized control over it. To convert this data in to useful data sets numerous organizations are using Data mining techniques to comprehend relationship between the data through which they can improve productivity and efficiency of business. This article represents the big mobile data sources, its challenges and what can be the future of big mobile data mining.

II. BIG DATA SOURCES AND CHARACTERISTICS

Collection of large amount of data is Big data, it can be generated either from one source or from different sources like social media sites facebook, linkedin, twitter, feedback form, satellite generated data, organizations relational database which include information about employee, traditional data which is stored in some journal or ledger, Mobile data which is generated via sensors etc. Big data could be any type of data like image, video, e-mail, text, and numeric and could be in any format. Though there is no fixed size and type specification for big data it is usually in terabytes and petabytes.

Below are some major characteristics of big data;

Volume: Big data is massive in size and this volume challenge is one of the reasons that it can't be processed with traditional data processing and management tools.

Velocity: The speed of big data generation is very high and it need to be processed in very fast manner so that results will be useful to make real time decisions like traffic and earthquake signals, weather forecast etc.

Variety: Big data contains data of different format and sizes. Data can be structured (relational or tabular), semi structured (XML or SQL) or unstructured (web text or log files). Merging this different types of data in to one single data set and finding relationship between this data is difficult task.

Heterogeneous and autonomous: Big data is from diverse sources this is the reason that big data is in different format and heterogeneous. As data is stored at different places there is no centralized control over it and this makes it autonomous. Complex relationship between data: with the size of big data relationships between the data goes increasing which makes the data more complex and challenging for knowledge discovery.

III. BIG MOBILE DATA MINING

World has become technology freak now. Mobile devices, tablets and ipads are like small computers, they are integral part of our life. By 2017 the number of mobile phone users is forecast to reach 4.77 billion (Statista Inc.). So the vast amount of digital data in streams is generated every minute. Telecommunication companies have user's personal data like call history and SMS. Smartphones apps like truecaller, whatsapp, wechat, hike etc. fetches highly personalized data of its users to the app vendors. No doubt that this Massive amount of big mobile data will keep exploding day by day as researches are going on smartphones to make them more intelligent and more interesting and already mobile devices has become necessity to its users, increased use of technology, Internet and smartphone apps have given access to information at fingertip.

Type of data present in big mobile data

Big mobile data is generated from diverse heterogeneous sources, this data can be used for various purposes like Telecommunication Company can use users call history details to offer calling plans to user. From there call history they

get information about to who they call frequently, how long they make call, calls are international or national and all other details like incoming, outgoing and missed calls. Geo-locations can be used for traffic rule apps. This information will be of great help to Telecommunication Company to offer suitable plan to customer.

Few main types of Big Mobile Data[4]

Calls (in/out/missed),SMS (in/out/failed/pending),Photos, Videos, Application events, Calendar entries, Phone book entries ,Location points, Unique cell towers, Accelerometer samples, Bluetooth observations, Unique Bluetooth devices, WLAN observations, Unique WLAN access points, Audio samples etc.

Mobile data can be of structured, semi structured, unstructured. Data mining is done on this Big Mobile data to find out interesting patterns and hidden knowledge. Multitude of sensors on mobile devices generate vast amount of data.

IV. ASPECTS OR CHALLENGES THAT SHOULD BE CONSIDERED WHILE MOBILE DATA MINING

When it comes to the Privacy data mining becomes a debatable topic. Big Mobile data contains users highly private information about their work, home, family, places they visit etc. it can provide information about location of user, Internet access by user, app usage by user and many more. There should be some strict measures to protect the privacy of data.

1. Legal duties: Ethical and legal requirements should be considered while data mining. Mobile data contains highly personal and sometimes sensitive data. Every country has different rules and regulations about data privacy which should be followed to avoid future legal affairs.
2. Mobile users concerns about data privacy : Permission from mobile user to use their personal information; In Data mining researchers and data miners get access to the private information some users may be don't want to give access to it due to security reasons.
3. Mobile Users Data Rights: mobile users should have access to their own collected data when it is used for data mining or research purposes so that they will be aware of what information is used about them.
4. Data Security: Repositories of Mobile Data Storage should be safe enough to avoid any data theft or illegal attack of hackers. Best available industry practices should be followed for data storage.
5. Data anonymization : data anonymization should be done to protect sensitive information from cyber and social crime. anonymization techniques like pseudonyms and hashing provide security to data to some extent and it makes difficult to trace back the data.
6. Researchers and Data miners should be committed to privacy policy of Big Mobile data. To find out new facts and to make business predictions organizations need Big Mobile data which provides historical as well as real time information and sometimes considered as best source of Big Data. Though we protect data technically at some point we need to give its full access to Researchers and Data miners and for that reson, getting written commitments from the responsible authorities is crucial.

V. BIG MOBILE DATA MINING WILL FADE OR WILL IT BE IN HIKE

More technological development and application development will lead to more big data generation. Before unfolding what's in the treasure of Big mobile data mining, First let us consider todays broadband and mobile use with respect to social networking sites like facebook and twitter ,1.3 billion mobile users use Facebook on mobile and approximately 250 million people monthly are the active users of twitter. Broadband plays a crucial role in Big Mobile Data development its presence in every corner of society holds lot more in its pocket for future, in 2009 Huawei has commenced its 5G mobile internet platform and IOT means internet of things where 'Everything on mobile, Everything connected, Every function virtualized' and its expected that by 2020 it will be in commercial use. Typical usage scenarios of 5G include ultra-HD video, virtual reality, self-driving (automated parking), and industrial control etc. 5G is not just an existing technology upgrade or a mere innovation. It combines existing technologies with revolutionary innovations. This will lead to immense set of data generation. More data with more unknown patterns and relationships of data attributes will give further opportunities to research and data miners.



Source: Mario Morales, IDC

Fig. 1 Expected future of Mobile Applications use and big data generation with respect to IOT

Here I have mentioned Huawei's 5G innovation as an example to describe how technological development could take big mobile data mining to new heights. Innovations will always be there and smartphones and other computer gadgets use will fade is less likely because features they provide like easy to carry and easy connectivity to world and every information at fingertip have made them necessity to match the speed of this fast technological era. In simple words more technological innovations and application development will lead to more big mobile data generation and data mining.

VI. CONCLUSIONS

Big data and data mining has emerged as crucial tool for making future prediction in various applications and Mobile data provides somewhat accurate, real time and highly personalized information to researchers. If Mobile Data mining is done in the bound of data protection and security then it is very useful to generate new information from historic data. Without data mining I guess Big Data will be of no use. Big data can't be used as it is for knowledge discovery due to its large size and complexity. If information from big mobile data aggregated and anonymized carefully then this secure mined data can be used in various applications like traffic management, business location selection, marketing optimization etc. This data is useful to understand human behavior, interests and social awareness. Big mobile data is leveraging for data miners as information flow will surely keep increasing. And it will also be challenging for data miners to process vast amount of data and protect privacy because mobile data can contain highly sensitive data. But no doubt challenges bring bigger opportunity.

REFERENCES

- [1] Xindong Wu, Xingquan Zhu, Gong Qing Wu, Wei Ding, 'Data mining with Big data', IEEE, Volume 26, Issue 1, January 2014
- [2] Bharti Thakur, Manish Mann, 'Data Mining for Big Data- A Review', IJARCSSE, Volume 4, Issue 5, May 2014.
- [3] Demetrios Zeinalipour-Yazti, Shonali Krishnaswamy, 'Mobile Big Data Analytics: Research, Practice and Opportunities', Mobile Data Management (MDM), IEEE 15th International Conference , Volume 1, 2014.
- [4] Juha K. Laurila , Daniel Gatica-Perez , Imad Aad , Jan Blom , Olivier Bornet, Trinh-Minh-Tri Do, Olivier Dousse, Julien Eberle, Markus Miettinen, 'The Mobile Data Challenge: Big Data for Mobile Computing Research', idiap publications – Online [Available] www.idiap.ch/~gatica/publications/LaurilaGaticaEtal-mdc12.pdf
- [5] Linoff G S, *Data mining techniques*, Published by Wiley India Pvt Ltd, 3rd edition
- [6] The Wordstream Blog, 9 Jan 2015. Online [Available] <http://www.wordstream.com/blog/ws/2015/01/09/the-internet-of-things>
- [7] Jiawei Han and Micheline Kamber (2006), *Data Mining Concepts and Techniques*, published by Morgan Kauffman, 2nd edition