

IMPORTANT ASPECTS OF DATA MINING & DATA PRIVACY ISSUES

**K.P Jayant, **Ajit kumar singh yadav, #Devendra Kumar*

**Research Scholar JJT University Rajasthan*

***Research CMJ University*

#Research CMJ University

ABSTRACT

It has made the world a smaller place and has opened up previously inaccessible markets of companies. The internet has brought about a huge change in the way that business is conducted the world over. On the down side though, indicate that many things about us, such as our taste in magazines, are finding their way into databases and are no longer so personal and private as many of us would prefer them to be. Data mining is the process of analyzing data from different criteria and summarizing it into useful information. This information can be used to increase revenue, cut costs, or both.

Data mining is part of a technological, social, and economic revolution that is making the world smaller, more connected, more service driven, and providing unprecedented levels of prosperity. At the same time, more information is known, stored and transmitted about us as individuals than ever. This section will provide a brief overview of some of the privacy issues and controversies that surround the use of data mining in business today.

Keywords: *Data Mining, Privacy, Issues, Criteria, Technology, Information, Change*

INTRODUCTION

Data are any facts, numbers, or text that can be processed by a computer. The patterns, associations, or relationships among all this *data* can provide *information*. Information can be converted into *knowledge* about historical patterns and future trends. Thus data mining process extract information from huge amount of data.

It has made the world a smaller place and has opened up previously inaccessible markets of companies. The internet has brought about a huge change in the way that business is conducted the world over. On the down side though, indicate that many things about us, such as our taste in magazines, are finding their way into databases and are no longer so personal and private as many of us would prefer them to be. Data mining is part of a technological, social, and economic revolution that is making the world smaller, more connected, more service driven, and providing unprecedented levels of prosperity. At the same time, more information is known, stored and transmitted about us as individuals than ever. This section will provide a brief overview of some of the privacy issues and controversies that surround the use of data mining in business today.

DATA MINING & PRIVACY

Privacy of data is a complex issue that, because of technology, is increasingly becoming a social issue [1] [2]. The Cambridge Advance Learner's

Dictionary (2004), defines the word privacy as someone's right to keep their personal matters and relationships secret, Today, every form of commerce leaves an electronic trail, and acts that were once considered private or at least quickly forgotten, are stored for future reference.

It is an important issue to consider both as individuals and in the work we do that may intrude on the privacy of other.

Limits are already placed on privacy by the social contact, and the issue is really how much information should be collected and who is in control of the information.

Every person has a different perspective on privacy.

Different levels of tolerance with regard to information about them being available to others. Technology plays a role in defining privacy, protecting it, and intruding on privacy.

THE ROLE OF DATA MINING

The data mining is a competence that addresses the strategic need of businesses to manage their customer relationships and run more efficiently. Most of the uses of data mining are in the area of marketing. Although not all of the aspects of data mining pose potential threats to individuals' privacy. The following two very important considerations:

Privacy violations may incur legal liability that could result in expensive law suits.

Privacy violations may result in bad press that can do considerable damage to corporate or brand image.

THREAT OF DATA MINING TO PRIVACY

In order to better understand the impact of data mining on privacy, consider the following example of its potential application in the telecommunication industry. A cellular phone service provider has the technological ability to determine the location of any switched on cell phone in its coverage area. The cell phone service provider collects information about all its subscribers during the sale of a contract. Typical subscriber information that would be collected may include the following:

Age Occupation.

Income

Banking Details

The ability of the cell phone service provider to track the location of the cell phone, and therefore its owner, might yield the following information:

The route typically traveled to and from work by the subscriber.

Whether the subscriber travels during business hours, or spends most of the day in the office. Which shopping centers the subscriber visits over weekends or after hours.

The cell phone service provider could make use of the collected information to position its own advertising billboards more strategically, or to situate its different branches at the correct shopping

centers. At the same time however, the organization might decide to benefit from this knowledge by selling it off to other organizations. The information could for instance be sold to other organizations who also want to be able to position their billboard more effectively, or a fast food chain could send out advertising messages to subscribers as soon as they come into close proximity of one of their outlets. One potential application of the above information could be the use of the information by marketers selling billboard advertising on the side of the road. Knowing the age, income and occupation of the people who travel a specific route could improve the effectiveness of such marketing campaigns even further. This example highlights an interesting application of data mining, but it also shows the potential threats that data mining pose to privacy.

The main areas of concern with regard to data mining and privacy are therefore found in the followings:

What kind of information do you collect about your customer? Who is ultimately in control of that information?

It is up to the organization employing data mining to ensure that their actions result in neither of the negative effects, namely, incurring legal liability or obtaining bad press as a result of privacy violations associated with their data mining effort (New 2004) [2]. Awareness project aimed at applying data mining to commercial databases for information on potential terrorists, due to a lack of consideration that was shown for privacy issues. The consumers might for instance be aware of the fact that collected information about them is used for billing purposes, but that they did not necessarily implicitly agree to allow the organization to use the data in a data mining scenario, thereby exceeding the original intent of the data collection. To this end it is important to pay particular attention to how the data used in data mining was obtained in the first place, and whether it's used could result in a violation of privacy.

PRIVACY PRESERVING DATA MINING

Privacy preserving data mining is a novel research direction in data mining and statistical databases, where data mining algorithms are analyzed for the side-effects they incur in data privacy [3][4]. The main objective of PPDM is to develop algorithms for modifying the original data in some way, so that the private data and private knowledge remain private even after the mining process, (Verykios et al) [6]. In essence this implies that the data to be mined would be stripped of all information that could be used to identify a specific individual, and that the same would be done to the resulting knowledge gained from the data mining effort. Privacy preserving data mining is still in its infancy, and whether it will be able to address all the privacy concerns in data mining is debatable. Another question that one needs to ask is how valuable data mining results would be to marketing, if the individual customers that the marketing efforts should be directed at, cannot be identified. In the meantime, organizations should consider the following in order to protect themselves from legal liabilities or bad press resulting from irresponsible data mining efforts:

Provide customers with an opt-out option whereby they have the ability to exclude themselves from being used in data mining or from being the target of directed marketing.

Ensure that you only buy data from reputable organizations, and that the necessary permission has been obtained for making use of that data.

Inform you customers of potential use of their information for data mining purposes, and obtain their consent prior to releasing this information to other organizations.

One thing that information technology experts and business professional much realize that following ethical practices and respecting the privacy of individuals makes good business sense. The bad publicity associate with a single incident can taint an organization's reputation for years, even when the organization has followed the law and done everything that it perceives possible to ensure the privacy of those from who the data was collected (Wang, 2003;397) [7].

CONCLUSION

In this paper we have discussed different issues on the privacy of data and services. In data mining data acts an important role to determine any fact of information. We have also review points on privacy preserving data mining.

REFERENCES

1. *Hao Li, Leping Liu. Application of data mining in physical statistics. Journal of east institute of technology. Vol.23, No.2, 2004.06*
2. *Chen Li and Claus Pahl. Security in the Web Services Framework. Dublin City University.*
3. *R. Agrawal and R. Srikant. Privacy-preserving data mining. In Proceedings of the 2000 ACM SIGMOD Conference on Management of Data, pages 439–450, Dallas, TX, May 14-19 2000. ACM.*
4. *Special issue on constraints in data mining. SIGKDD Explorations, 4(1), June 2002.*
5. *C. Clifton. Using sample size to limit exposure to datamining. Journal of Computer Security, 8(4):281–307, Nov.2000.*
6. *M. Atallah, E. Bertino, A. Elmagarmid, M. Ibrahim, and V. Verykios. Disclosure limitation of sensitive rules. In Knowledge and Data Engineering Exchange Workshop (KDEX'99), pages 25–32, Chicago, Illinois, Nov. 8 1999.*
7. *Wang J. 2003. Data Mining: Challenges and Opportunities. London, IRM Press.*

Author would like to extend their heartfelt thanks to the academic and infrastructural support received from the University.