DATA MINING AND PRIVACY ISSUES

Amritam Anand

NALSAR University of Law, Hyderabad

INTRODUCTION

William O. Douglas, the renowned American jurist were of an opinion that "***the right to let alone is indeed the beginning of all freedom***".[1] This right to be let alone is what we recognize as the right to privacy. Privacy is also recognized as "***protection from unreasonable use of state and corporate power"***.[2] Violation of personal space has become increasingly facile in the world of social media and E-commerce. Identity theft, voyeurism, Stalking, Hate speech are few of the many threats posed towards the rightful claim of an individual to determine the extent to which he wishes to share himself with others.[3] The surveillance capabilities of new technology is tremendous and is expanding expeditiously. The more we indulge with social media and E-commerce at a daily level, higher the vulnerability. There is no transparency, most of the time the user doesn't even know that he is being followed. Individual privacy is being commodified and sold by social media giants in collaboration with government, e-commerce wolves, etc... Every activity of a user on Internet, leaves a NET print which contains small portion of his individualistic information. These NET prints are then collected by the data brokers by a process called data mining. The mined data is thereafter analyzed and synchronized to create a Data profile of each and every person.[4] This is a continuous process, data is updated after every online activity of a user. Idea is similar to criminal profiling done by the investigating agencies.[5] Advancement in the technology has led to setting up of data servers which can amass huge amount of personal information.[6] This data can be instantaneously analyzed. Problem per se lies in the fact that, most of the time the person being profiled is unaware

1 Public Utilities Commission V. Pollak, 343 U.S. 451, 467 (1952)
2 Nick Harkaway, The blind Giant.
3 A.C. Breckenridge, The Right to privacy.
4 International Journal of Computer Trends and Technology- volume4Issue2- 2013 , ISSN: 2231-2803, Data Security and Privacy in Data Mining: Research Issues & Preparation
5 Criminal Profiling, A Viable Investigative Tool Against Violent crime, by- JOHN E. DOUGLAS, M.S. Special Agent/Program Manager Profiling and Consultation Program Behavioral Science Investigative Support unit National Center for the Analysis of Violent Crime FBI Academy Quantico, VA and ALAN E. BURGESS, M.Ed. Special Agent/Unit chief Behavioral Science Investigative Support unit and Deputy Administrator National Center for the Analysis
6 http://www.webopedia.com/TERM/D/data_server.html

of such a process being conducted on him. He is unaware that, there is a data center[7] somewhere far away from him, which contains servers amassing all the personal information about him. Infringement of privacy lies in the fact that the individual never gave his consent to mining or profiling. Even if he did, he was largely unaware of, what did he agreed to in form of those long user agreements.[8] He is using social media and e-commerce services being unaware of the sword of corporate surveillance. It is in this regard, the write-up looks to analyze data mining and privacy issues in India. For the purpose of this write-up, I will restrict, to the cases of data mining by the social media and e-commerce websites. How, this data mining is against the individual's right to privacy and is being exploited for commercial benefit.

DATA MINING

E-commerce

Suppose, two months back, I bought an iPad from Amazon. Thereafter last month, I bought a Television, again from Amazon. Now, Amazon will store this data, they will analyze it and will try to predict the next thing I need. So, from this moment onwards, I will see advertisement of different accessories connecting an iPad to a television. Such as HDMI cable connecting the two. Question arises, how is this problematic? Because it is in one way helping me. My information is being used to provide me with ease of business. This analogy seems genuine and true but problem doesn't persists in the purpose the data is being put to right now but lies in the numerous use it can be put into. Problem lies in the fact that, apart from Self-Regulation on part of these data mining agencies, there is no specific protection for an individual under the Indian law. To further explain, we can refer to the case of Toysmart.com.[9] Toy smart used to sell toys to children. They wanted to create a database of children so that they can analyze it, in order to design the toys as per their need. For this purpose, they started giving free toys to children for which kids have to fill a form. The personal data carved from the form was analyzed and then further used for the business purpose by the company. After few years, the company went into loss. In time of this financial crisis, the company decided to sell this data. This shows the ugly face of the corporate world, self-regulation will fade, the moment some financial crisis or a lucrative incentive is provided.

Social Media

Most of the social networking is based on the assumption that, whatever a person is doing on the social media is private unless the person decided to make it public. Privacy settings

7 http://www.cbronline.com/news/data-centre/top-10-biggest-data-centres-from-around-the-world-4545356/
8 https://www.theguardian.com/money/2011/may/11/terms-conditions-small-print-big-problems
9 FTC v. Toysmart.com, LLC, No. 00-11341-RGS

are provided by the social media sites, specifically for this purpose. So, we presume that we are not tracked or watched. User is asked to provide his personal information whenever he signs up and is constantly asked to update it. Fear of being tracked doesn't arise because of the security guarantee a social media website provides. The security guarantees build up a trust in the user. Trust also arises from the fact that any violation of privacy by the social media sites will lead to public back lash and loss of user base which will lead to financial loss for them. However, social media sites amasses a lot of information about you on a day to day basis. [10] The pages one likes, tell about his fashion choice, religious leanings, ideology, ethnicity, sexual orientation etc. your comments tell a lot about your views and opinion on a particular issue. In a world where a large chunk of world population is on social media, having control over personal data of people is in itself very dangerous. Personal data can be used by the owners of the data against their business rivals. Data brokers can use it to promote the political ideology they support. Most recent example of this was U.S. presidential elections of 2016-17. All, those who voted for Hillary were shocked when the results came out. They state that, all the social media feed on their network showed Hillary in positive light. Even opinion post and exit polls shown on their news feed, suggested Hillary to win. The reason behind this lies in the data mining. Social media sites use user's data to analyze their choices and ideological leanings. Then they decide which candidate will be more suitable to a particular user's ideology. Thereafter user's news-feed is designed keeping the analyzed information in mind. So, someone with a liberal ideology will have news feed supporting Hillary and someone with a conservative ideology will have their news feed supporting trump. This works against the very notion of deliberative democracy where a person has right to make choices and again, even if he is being influenced then he should be influenced logically and not by intentionally keeping him away from other side's argument. Social media, created a bubble for supporters of Hillary, which busted when trump won. Though it would had been true for trump supporters too, if Hillary had won but we cannot negate the possibility of misuse of such information. [11]A political party affiliated with a social media site can blackmail its rival or his fund-raiser by using his/her personal information. This raises the arguments for right to be forgotten on Internet. Proponents of such a right claim that an individual should have a right to get his data profile deleted from the servers of data brokers.[12] Deleting the data profile omits the person's existence online, which is necessary in the case if a person thinks that the data stored can be used against him/her in any manner. The user has little or no control over his data profile, or the information flow. Right to be forgotten, provides the

10 http://www.bbc.com/news/uk-36701297

11 http://www.businessinsider.com/escape-your-bubble-facebook-news-feed-2016-12?IR=T

12 Court of Justice of the European Union PRESS RELEASE No 70/14 Luxembourg, 13 May 2014, Judgment in Case C-131/12 Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González

user with an option to opt out of such a process. It is evident from above discussion that contemporary construct of privacy is political and commercial.

Privacy Concerns

In modern times, data is being unconsciously served by us on the plate to the data brokers. These data brokers then consciously use these data to create hidden databases of personal information about a user. These data brokers collaborate with e-commerce sites and social media sites to do an on-line profiling by the means of data mining. Data mining can be outsourced to data brokers or the e-commerce companies and social media companies themselves can play the role. User login option through social networking website is used to identify a user's activity with the basic public profile of the user created on the social media. This provides, a single node of identification for the data brokers. This, helps them relate user's each and every online activity to our social media profile. A computer which is used to surf the Internet, is a private property but its use may not be private and data brokers do not factor it in. there is a large possibility that, more than one user shares a single computer and single login setting for surfing. In, that case the click-bait process of mining may lead to creating a data profile which might not entirely belong to the user whose login settings were used. Further it is not necessary that a user might be interested in all the activity he is doing online. The object behind a user's activity on internet may not be personal at all point of time. So, when an online-profile is created using a click-bait method in collaboration with social media companies then it is highly possible that such a profile isn't accurate and is dubious in nature. This data can be easily used to incriminate someone or can be used for recruitment purposes. We have already seen, employer monitoring employee's activity via his social media account.[13] It is not far away that data brokers will be hired to monitor employee's activity. Technology has made, online activities an extension of person's identity. Right to privacy recognizes the notion that an individual can decide what he can share and what he cannot.[14] Divergence of information about someone is that person's own prerogative. Hence, when data mining is used to diverge such an information by collecting the small information trails a user leaves through his online activity, it leads to invasion of an individual's private sphere which per se is violation of right to privacy.

RIGHT TO PRIVACY

Right to privacy is not specifically mentioned under the constitution. It has been recognized through judicial pronouncements. It is also a right guaranteed under common law[15] and so as far as non-state entities are concerned, they will have a tortious liability for violation of

13 http://www.businessinsider.com/empoyers-ask-for-facebook-password-2012-3?IR=T
14 Supra Note 3
15 Jones v. Tsige, 2012 ONCA 32, 108 OR (3d) 214.

the Right to Privacy. In ***People's union for civil liberty V. union of India***[16], Supreme Court held right to privacy can be derived from right to life and personal liberty enshrined under article 21 of the constitution of India. Thereafter in 1995, Supreme Court for the first time provides propositions defining right to privacy in case of ***R.Rajagopal V. State of Tamilnadu***[17]. The court states it as a right to be let alone and provides that no one can publish anything related to an individual's private domain without his/her consent. Such a publication will be in violation of his right to privacy. In this case, court also created an exception, under which publication won't be violative, if such a publication is based on public records. Premise of such an exception being that as soon as a matter enters the public domain it ceases to be private therefore a claim of right to privacy doesn't arise. Though court only mentions it to be a considered as broad principle in a case of infringement of privacy which is not exhaustive and were of an opinion that right to privacy should be be developed case by case.

An argument can be made that each and every user activity online is in public domain, hence there cannot be a claim of right to privacy here. But something can be said to be in public domain when everyone have an equal access to the same and person has willingly made that information available in public domain.[18] And if not willingly then there was some legal enforcement making the information available in public. As far as data brokers are concerned, the data is not available to everyone equally, further person is not willingly making data available to entire world.

Kharak Singh V. State of U.P.,[19] the case dealt with police surveillance. Court was of an opinion that, invasion of privacy must be both tangible and and direct and a mere right to protect personal sensitiveness is not secured under constitution.

In ***Gobind V. State of M.P.***[20]***,*** Court allowed domiciliary visits in the case of police surveillance but restricted such surveillance to those cases where there is evident danger to the public security due to the person.

In ***People's Union for Civil Liberties V. Union of India***[21]***,*** court recognized the use of technology for invasion of privacy. Telephone tapping was termed as technological eavesdropping and hence was held to be violative of right to privacy.

All these, cases acknowledges that right to privacy is not an absolute right and it can be restricted in favour of public security as per as a procedure established by law. They are also

16 People's union for civil liberty V. union of India, AIR 1991 SC 207
17 R.Rajagopal V. State of Tamilnadu, AIR 1995 SC 264
18 Information in the public domain, Freedom of Information Act Environmental Information Regulations
19 Kharak Singh V. State of U.P., AIR 1963 SC 1295
20Gobind V. State of M.P., AIR 1975 SC 1378
21People's Union for Civil Liberties V. Union of India, AIR 1997 SC 568

the cases, initiated against public authorities, which do not deal with the issue, where a private entity is engaged in privacy infringement. As far as invasion of privacy by the private authority is concerned, the information and technology act, 2000 lays down provisions regarding the same. It extends the right to privacy to protection of data stored in form of electronic records. Any disclosure of data contained in electronic record without consent of person concerned has been made a punishable offence. Section 72 of the Information & Technology Act, 2000, recognizes the constitutional principle of privacy and provides penalty for same. The act, is premised on the idea that an individual privacy should be governed as per as his consent and not otherwise. In-spite of having both constitutional and statutory recognition of right to privacy, India for now lacks a specific data protection statute. Due, to which current Information technology act is insufficient to deal with issue of data mining which requires laying down of specific guidelines for both public and private entities engaged in the field of data mining. Information & Technology Act fails to safeguard a user in case of data mining because a user while using Internet provides his consent in form of those long and never read user-agreement. Therefore, a new act laying down guidelines for data mining is must in modern world. An individual should be specifically told that their activity on Internet will be tracked and for what purpose it will be tracked for. Tracking of activity should not be a pre-condition for using the service. A user should be given the right to choose whether he want to be tracked by the data brokers or not. Further if he wants to be tracked, then whether he want his data to be put into a database or not. If he assents to it being put into a data base then he should have control over any kind of amendment made to the data base. He should also have right to choose what kind of use his data will be put to.

Right to be Forgotten

This right is enforced in those cases where a person wants any specific personal information deleted from the Internet or from an online database.[22] The right was recently recognized by Karnataka high court, but the application of this right in India doesn't apply to any information which is stored in an online database but only those information which might be sensitive. [23]It could be something where a prisoner would like the Internet to forget his criminal record, so that he could integrate better in the society. Or sensitive information would include, online data base to not contain history of a woman being victim of rape or voyeurism. Sensitive information should also extend to inculcating that information which leads to creation of a stereotype against the person. Such a stereotype can cause a lot of stigma for a person in the society he belongs to. Right to be forgotten, demands the deleting of such sensitive information about a person which has been published online as per as the public records. They do not talk about that information which has been collected from a

22 Supra Note 12

23http://www.deccanchronicle.com/nation/current-affairs/300117/karnataka-high-court-upholds-womans-right-to-be-forgotten.html

user by tracking his activity online. In such case this is a private information which a user has only shared with the data broker as per as the user agreement. Secondary use of such data depends on the term of user agreement. Right to be forgotten cannot be extended to private agreements a user has entered into.

Right to be forgotten arises from the basic premise of right to privacy which provides an individual to be the sole proprietor of his personal information. European Union data protection directive of 1995, recognizes the right to forgotten. European Union Court of justice, while applying this right allowed users to ask search engines to remove links to web-pages that contains information about them.[24] Right to be forgotten is also based on the premise that it should be for a user to decide the relevancy of the information he wishes to have online. Right to be forgotten is still a western concept which is not at all developed in India. On one hand, Karnataka High-court upheld it in cases where some sensitive information is being placed online and on other hand Gujarat high court refused to recognize the right in India.[25] There is no explicit backing for right to be forgotten under any statutory or constitutional provisions in India. This calls for new privacy laws to deal with privacy issues arising out of data mining and online profiling in India.

CONCLUSION

Small information trails left by an individual ire collected by the data brokers and then analysed to form an online profile of the individual. This entire process came to be known as data mining. As creepy as it sounds, it also gives rise to a lot of privacy issues. Self-Regulation, is the only safeguard put on the use of the data mined by the brokers. Indian privacy laws are not updated to deal with the advancement in technology. India, requires a statutory data protection act which regulates the private authorities from collecting the personal information. Regulation should create a balance between right to privacy and right to know. It should also, provide for use an information could be put to. Write-up has illustrated that data mining has a tendency to store wrong information about an individual or a sensitive information about an individual, in such cases individual right to be forgotten should be recognized. This could only be provided by a statutory provision recognizing such rights of an individual. India is far behind in terms of laws for data protection in India and needs to expedite the process to compete with ever growing technological advancements.

24 Supra Note 12

25 http://www.legallyindia.com/views/entry/two-takes-on-the-right-to-be-forgotten