Xiaomin Xu

Prof. Busker

ENG 102, 052W

4 April 2015

Restrict Web-Service Company Freely Using Our Information

As information technology becomes more advanced in today's society, information retrieval and security have been a very popular topic. From last year statistic data, there are more than 71 million people in the United States that became cybercrime victims last year, which cost about $20.7 billion loss(Norton). With new big data technology Hadoop, dealing with TB(1TB = 1024GB) level web data is much easier than before. Therefore, web-service companies like Facebook, Google, and Twitter eager to use their services to collect users’ data, process data by Hadoop and share with third party companies or do experiments to improve their services freely. There are no restrictions for web-service companies to collect and use users' data as declared in register agreements. This definitely causes big issues for users. For examples, in 2015, a lot of people became victims in fraudulent electronic tax in Tax Turbo because of all their sensitive information like SSN and birthday is known by bad guys who get data from some web-service companies. Setting up information protecting laws and forgotten strategy (which means companies must delete users data in a specific time like one month) are really important for protecting users and warn these companies.

Introduction of Information Retrieval

There are mainly three stages in process of information retrieval.

The first stage is information collection, which means gathering user information.

Web-service companies basically use the web and mobile applications to gather information. For example, Facebook gather all information by users’ clicking and typing messages. Once users send their messages to friends, their messages also record in Facebook server and then their database. Especially, they are interested about the link that you clicked, they will record the website title and URL address of link. If users also use other Facebook services like Facebook games, they will record every bit data like what kind of game you played before, how much time you play each time, and other useful data. Basically, all information related to specific users Facebook want to collect.

The second stage is information process. This is very complicated process and all companies use their own methods to deal with their business services. The information companies have different algorithms to do it. Their goals are similar, which is trying to retrieve useful information. With new machine learning technology, software engineers can do more things than what you can imagine. For example, from the basically sample data they collect from hundreds of time of one user. They know the possibility about which locations the user will be at particular time slots(Since there is specific pattern that computer can analyze from data). In current industry area, the machine learning skills are used everywhere to improve the services of companies. The machine learning can give the feedback of users. It is very important technology for advertisements.

The third stage of the process is that distribute the information that web-service companies retrieve from users. As we all know that, the web-service companies such as Google, Yahoo and Facebook revenue are mainly from advertisements, they have also a lot of third partners. They will share part of information they retrieve to the third party companies. Their partners freely get or buy users' information from these companies.

The big data technology greatly improved the information retrieval process in the last few years that makes the information retrieve and sharing process much quicker than before. Currently, if company want to improve the speed of processing data, they just need to buy more computers or use Amazon Web Services, which is cheap online payable computing resources that I currently use for processing data.

Who benefits from the process.

1. The internet companies definitely benefits from the process directly since they can get advertisements revenue and make use of data to improve their services further increasing their competition ability in the market. (quote Facebook revenue income) that's why Facebook can buy the most expansive land in the earth in Palo Alto. Also Yahoo benefits the process by keep tracking of these kind of information. And even a car company benefits from the user information feedback. For example, "Automakers are obtaining data through real-time navigation functions and other on-board location services and storing it for undetermined lengths of time. They need to provide motorists with more information on how and why they're collecting and sharing data, according to a government report"(autoblog) There is a lot of advertisements revenue, "mobile advertisers are trying to find ways to entice drivers near their stores to shop with coupons and deals sent through such on-board systems."

2. As a user, we also benefits from the process. Companies will offer more precise and convenient services for us since get useful information from us.

For example, after feedback from users, leetcode company is free online service for coding practice, it gets a lot of solution about one coding questions, now it offers users solution about particular question and they also have explaination about particular coding questions. The users know saves time for searching answers about particular questions. Also, another example is my website, after collecting enough from users, I can make better recommendation for users. This is because each time users click the paper that I recommend to them and then know they click to specific papers and I will records which paper is more popular and next time I will score the paper higher and recommend to other users.

A lot of companies like Amazon would like to use user information to make recommendation.

3. Government, A lot of government organization are based on information, for example, FBI, NSA the spying organization in the U.S. they either hack company or directly get access as partners to retrieve all data companies. Now these organizations know much more things than before with so much data

that each companies retrieved.

Who lost in the process:

In this topic, it mainly talked about that Facebook did a study of how its users’ response to various kinds of stores in their feedbacks, namely by looking at emotional words in posts and giving different users different mixes of posts. The results that Facebook post is that there is slightly different but measurable effect in what people write stuff afterwards. This psychology was revelt by the National Academy of Sciences that 700,000 users are taking this “emotional contagion through social networks” tests.

Several people were immediately outraged that Facebook would do a thing like this and publish it. The policy vacuums is that Facebook let participants do the user psychology experiments that devised by their data scientists without declaring the experiments to users. Usually, researchers do a psychology experiment they will tell you that they will a experiments or they need to get particular approval if they need to keep users unknown about experiments.

A lot of company do their psychology experiments to understand their business better and make more money, just not publish them.

The moral principles or rules are involved is when a psychology experiments need to be done. The target users should always know they are doing a psychology experiments. However,

Facebook did his experiments did not tell the users.

Startup companies will continue lose the competitive in a very long term. The earlier years for startup to become a big company will be more struggler for them since they don't have users information as resources. That's why there is no another search engine can be competively with Google since google gather so much users information and feedback of users information will greatly help google to improve their services, no one can exceed Google even their technology is more advanced.

Other company like TurboTax, this is because once the sensitive information that internet companies collected is stolen by some bad hackers, there is a big trouble. The fraudulent electronic tax events were dramatically increased in the busy tax season.

Intuit, the company that made TurboTax complained being hit so hard by bad guys. Intuit had to close their online tax service for a whole day. And even worse, Minnesota directly stopped accepting filings from TurboTax. The fraudulent tax event increased dramatically exactly after the Anthem breach, in which millions of people’s sensitive data were stolen. And these sensitive data are very important for registering new accounts or login the account in TurboTax. As some experts suggest, after bad guys got a lot of sensitive data, they attack TurboTax and implement fraudulent electronic tax.

Before we dive into the ethical analysis, we first identify the CIA objectives. In these attacks, confidentially objective was involved because people’s sensitive information like social security number were stolen. And these bad guys made use of sensitive information to create millions of accounts online and they write code to modify their credit card information in order to change the receiving refund account. In this way, all refund money will directly enter into fraudulent accounts. For instance, Jeff Parish found his online filing had already calculated a federal refund but refund to another account. The result of bank account information changing cause Jeff lost the tune of more than $5,000 refund. Therefore, integrity objective was lost in these attacks since a lot of bank accounts and tax information were unauthenticated modified by bad guys. (Tax Scam 2015)

Sony Hackers

In conclusion, based on duty-based perspective, all these internet companies should be responsible for protected the sensitive information they collected, the should not be able to collected. Currently there is no precise law says that. From users perspective, our information rights are violated in a way that even ourselves did not realize. The Tax Scam events told the we should not let internet companies freely use our data and share our data.

Work Sited

Facebook, Inc. "Data Policy." Data Policy. *Facebook*, 30 Jan. 2015. Web. 04 Apr. 2015.

[Bigelow, Peter. "Car Companies Are Collecting Drivers' Location Data." Autoblog. AOL Inc, 30 Mar. 2015. Web. 04 Apr. 2015.](http://www.easybib.com/cite/clipboard/id/1428126702_551f7beef40c12.43669374/style/mla7" \t "_blank)

[Auerbach, David. "Google vs. the FTC." Google FTC Report. Slate.com, 25 Mar. 2015. Web. 04 Apr. 2015.](http://www.easybib.com/cite/clipboard/id/1428126968_551f7cf8013129.24536015/style/mla7" \t "_blank)

[Osborne, Charlie. "Norton: Cybercrime Cost $110 Billion Last Year" ZDNet. Zero Day, 5 Sept. 2012. Web. 04 Apr. 2015.](http://www.easybib.com/cite/clipboard/id/1428127075_551f7d633e3f36.10573092/style/mla7" \t "_blank)