Big data security problems threaten consumers' privacy

March 23 2016, by Jungwoo Ryoo

Credit: AI-generated image (disclaimer)

As more personal information is collected up by ever-more-powerful computers, giant sets of data – big data – have become available for not only legitimate uses but also abuses.

Big data has an enormous potential to revolutionize our lives with its

predictive power. Imagine a future in which you know what your weather will be like with 95 percent accuracy 48 hours ahead of time. But due to the possibility of malicious use, there are both security and privacy threats of big data you should be concerned about, especially as you spend more time on the Internet.

What threats are emerging? How should we address these growing concerns without denying society the benefits big data can bring?

The size of the potential problem

First of all, due to the sheer scale of people involved in big data security incidents, the stakes are higher than ever. When the professional development system at Arkansas University was breached in 2014, just 50,000 people were affected. That's a large number, but compare it with 145 million people whose birth dates, home and email addresses, and other information were stolen in a data breach at eBay that same year.

From the perspective of a security professional, protecting big data sets is also more daunting. This is partly due to the nature of the underlying technologies used to store and process the information.

Big data companies like Amazon heavily rely on distributed computing, which typically involves data centers geographically dispersed across the whole world. Amazon divides its global operations into 12 regions each containing multiple data centers and being potentially subject to both physical attacks and persistent cyberattacks against the tens of thousands of individual servers housed inside.

Difficulties with access control

One of the best strategies for controlling access to information or

physical space is having a single access point, which is much easier to secure than hundreds of them. The fact that big data is stored in such widely spread places runs against this principle. Its vulnerability is far higher because of its size, distribution and broad range of access.

In addition, many sophisticated software components do not take security seriously enough, including parts of companies' big data infrastructure. This opens a further avenue of potential attack.

For instance, Hadoop is a collection of software components that allows programmers to process a large amount of data in a distributed computing infrastructure. When first introduced, Hadoop had very basic security features suitable for a system used by only a few users. Many big companies have adopted Hadoop as their corporate data platform, despite the fact that its access control mechanism wasn't designed for large-scale adoption.

Consumer demand drives security and privacy

For consumers, then, it is critical to demand a heightened level of security through vehicles such as terms and conditions, service level agreements, and security trust seals from organizations collecting and using big data.

What can companies do to protect personal information? Countermeasures such as encryption, access control, intrusion detection, backups, auditing and corporate procedures can prevent data from being breached and falling into the wrong hands. As such, security can promote your privacy.

At the same time, heightened security can also hurt your privacy: it can provide legitimate excuses to collect more private information such as employees' web surfing history on work computers.

When law enforcement agencies collect information in the name of improved security, everyone is treated as a potential criminal or terrorist, whose information may eventually be used against them. The authorities already know a lot about us but could ask companies such as Apple, Google and Amazon to provide more intelligence such as a decrypted version of our data, what search terms we are using and what we are buying online.

The fundamental security principle used to justify this type of blanket surveillance (which is now more affordable and feasible due to the use of big data technologies) is "nobody can be trusted." Once collected, those data join the rest of the information in being susceptible to abuse and breaches, as demonstrated in snooping incidents involving National Security Agency employees.

And yet when used properly, big data can help enhance your privacy by allowing more information to be leveraged and eventually improve the quality (especially, the accuracy) of intelligence on potential attacks and attackers in cyberspace.

For example, in an ideal world we don't have to worry about fraudulent emails (also called phishing) because a big data analytics engine would be able to pick out malicious emails with pinpoint accuracy.

How big data is used – for you or against you

There are also other privacy concerns about big data. Companies are eager to deliver targeted advertising to you and tracking your every online move. Big data makes this tracking easier to do, less expensive and more easily analyzed.

A service like IBM's Personality Insights can build a detailed profile of you, moving well beyond basic demographics or location information.

Your online habits can reveal aspects of your personality, such as whether you are outgoing, environmentally conscious, politically conservative or enjoy travel in Africa.

Industry representatives make benign claims about this capability, saying it improves users' online experiences. But it is not hard to imagine that the same information could be very easily used against us.

For example, insurance companies could start questioning coverage to consumers based on these sorts of big-data profiles, which has already begun to happen.

Banning large-scale data collection is unlikely to be a realistic option to solve the problem. Whether we like it or not, the age of big data has already arrived. We should find the best way of protecting our privacy while allowing legitimate uses of big data, which can make our lives much safer, richer and more productive.

For example, when used legitimately and securely, big data technology can drastically improve the effectiveness of fraud detection, which, in turn, frees us from worrying about stolen identities and potential monetary loss.

Transparency is the key to letting us harness the power of big data while addressing its security and privacy challenges. Handlers of big data should disclose information on what they gather and for what purposes.

In addition, consumers must know how the data is stored, who has access to it and how that access is granted. Finally, big data companies can earn public trust by giving specific explanations about the security controls they use to protect the data they manage.

This article was originally published on The Conversation. *Read the*

original article.

Source: The Conversation