Cloud risk communication on social media: The case of Premera Blue Cross

Jean Pierre Guy Gashami[1] **Christian Fernando Libaque-Saenz**[2]
Myeong-Cheol Park[1] **Jae Jeung Rho**[1]
[1] Korea Advanced Institute of Science and Technology
N22, 291 Daehak-ro, Yusong-Gu, Daejon 34141, Republic of Korea
[2] Universidad del Pacífico
Avenida Salaverry # 2020, Jesús María, Lima 11, Peru
jp.gashami@gmail.com cf.libaques@up.edu.pe
imcpark@kaist.ac.kr jjrho111@kaist.ac.kr

Abstract

Cloud computing has been growing at a fast pace. This growth has been fueled by this technology's inherent benefits such as cost reduction and convenience. However, the increasing amount and variety of data processed on the cloud have raised the number of security breaches. Although cloud providers were responsible for data security in the past, the new threats require that both cloud providers and users coordinate efforts to minimize losses and ensure data recovery. Our study aims to explore how cloud providers and users can leverage social media to mitigate data security breaches through effective risk communication. We analyzed public data collected from Twitter regarding the security breach faced by the Premera Blue Cross web application between January and April 2015. Preliminary results indicate that Premera Blue Cross (cloud provider) acted as an information source for Twitterers seeking relevant and accurate information during this security breach. Future steps for this study are discussed.

1 Introduction

Cloud computing is disrupting consumption models of information technology (IT) across industries. For example, around 65% of all major enterprises in USA are using some form of cloud computing (Verizon, 2014), while general spending on public cloud computing services is expected to grow by US$921 billion by 2017 (Gartner, 2011). The rapid increase in the use of cloud computing services by both enterprises and individual users is driven by benefits such as cost reduction, mobility, and convenience (Gashami et al., 2015). Indeed, users are increasingly relying on cloud providers to run hardware, software, and also to properly handle their data. However, having more data in the cloud, including sensitive data such as personal, financial, research, and health information means high potential risks for users (Zhou et al., 2010). Not surprisingly, data risk has been identified as a high threat to cloud computing (King and Raja, 2012). For instance, costs associated with data security breaches in the healthcare industry alone could reach US$5.6 billion annually (Experian, 2015). Undoubtedly, security breaches may occur in spite of cloud providers efforts to ensure data safety (Armbrust et al., 2010). Some research even argues that data security breaches are inevitable in the cloud (Staten et al., 2014). To face such security challenges, cloud providers have developed risk management frameworks which mainly focus on risk analysis, risk assessment, and risk mitigation (Zhang et al., 2010). These frameworks address technical and managerial issues; however, it is still unclear how cloud providers treat users throughout the analysis, assessment, and mitigation of security breaches. Existing research suggests that risk communication with all stakeholders is an important element of risk management in various contexts (Aguirre, 2004; Lagadec, 2002). On the cloud front, communication with users may play a crucial role in limiting potential damages by raising user awareness of data practices and protection. Indeed, communicating potential security breaches to users can lead to actions such as reinforcing weak passwords, using private keys, or enabling local backup of data (Rainie and Duggan, 2014). On the other hand, social media such as Facebook and Twitter have been signaled as the new avenues for channeling information during risk management due to their low-or no-cost policy and their worldwide usage (Wright and Hinson, 2009). Natural and health disasters are clear examples of populations and organizations relying on social media to alert, organize, or manage rescue efforts (Theocharis, 2013).

Despite the relevance of risk communication in the context of cloud computing and the potential of social media for information dissemination during a security breach, to the best of our knowledge there is no research on the usage of social media for risk communication during security breaches in the cloud. The objective of our study is to fill this gap in the literature. In this first step, we attempt to address the following research question:

RQ: Who are the key players disseminating information of cloud computing data security breaches on social media?

2 Literature Review

2.1 Cloud Computing

Cloud computing emerged as a computing model rooted in various technology innovations such as virtualization and web services (Foster et al., 2008). Cloud computing can be defined as a computing model that enables the provision of ubiquitous, network-based, and on-demand services to users (Armbrust et al., 2010). With cloud computing, services and infrastructure that were traditionally provided locally are remotely accessed, consumed and paid for through a web browser or an application interface (Marston et al., 2011). Cloud computing can be classified as: private model, where the cloud is solely operated by a single organization; public model, where it is open to the general public; community model, which allows organizations with common interests to set up and access the same cloud; and hybrid model, which is a combination of any of the three previous models (Mell and Grance, 2011). Additionally, cloud computing services can be categorized as: Software as a Service (SaaS), encompassing web applications; Platform as a Service (PaaS), which offers software development environments over the web; and Infrastructure as a Service (IaaS), which provides users with access to storage and computational power (Youseff et al., 2008). All these types of cloud computing come not only with benefits but also with potential risks for users (Ko et al., 2011).

Prior studies recognized data security risk as a serious threat to cloud computing (Jaeger et al., 2008). For example, research by Belian and Hess (2011) and Wu et al. (2011) found that security risks were negatively affecting SaaS use in enterprises. Zhang et al. (2010), on the other hand, developed an information security framework for cloud computing that emphasizes the role of risk analysis, assessment, and mitigation. Whereas Chan et al. (2012) proposed a risk framework made up of event identification, risk assessment, risk response, information and communication and monitoring. These studies, however, do not consider the involvement of users in data-protection initiatives.

2.2 Risk Communication on Social Media

Risk communication can be defined as a process of exchanging information among interested parties about the nature, magnitude, significance and control of a risk (Covello et al., 1998). Risk communication has become highly important in risk mitigation and damage control in areas such as homeland security (Jung and Park, 2014), and earthquake occurrence (Nigg, 2006).

With the rapid evolution of IT, risk communication is shifting towards social network sites (SNS). SNS can be defined as applications based on Web 2.0 that serve as platforms where users create and distribute content (Kaplan and Haenlein, 2010). These platforms facilitate sharing information in real time for a rapid diffusion. For example, Yates and Paquette (2011) studied the use of social media during the earthquake in Haiti in 2010. Likewise, Bird et al. (2012) addressed how citizens and rescue organizations relied on social media during the Queensland and Victorian floods. Goolsby (2010) also highlighted the heavy use of Twitter in communicating the areas to avoid during the Mumbai attack in 2009. In short, prior research focuses on the use of social media in high-risk environments with potential human or property loss. However, to the best of our knowledge, no study has been conducted to understand how this same channel can be used to prevent or mitigate data security risks.

2.3 The Premera Blue Cross Data Security Breach

Premera Blue Cross is a health insurance company based in Mountlake Terrace, Washington, USA. On March 17th, 2015, the company announced that it had suffered a security breach and that data from 11 million users might have been compromised (Matthews and Yadron, 2015). The Premera Blue Cross data security breach is an example of a typical cyber attack through a web application. Web applications and services are among cloud computing key core technologies (Marston et al., 2011). Hence, understanding data security breaches in this technology and the associated risk mitigation can accurately reflect cloud computing vulnerabilities (Grobauer et al., 2011).

3 Data Collection and Analysis

3.1 Data Collection

We collected public data from Twitter based on the keyword Premera from March 18th, 2015 to March 31st, 2015, spanning a 14-day period. Twitter is a microblogging SNS that allows users to send 140-characters messages known as tweets, respond to tweets using Retweets (RT), mentions (@user), and hashtags (#word) (Kwak et al., 2010). Twitter was chosen in our study because recent studies found that organizations and individuals rely heavily on Twitter for risk communication during protests, environmental disasters, homeland security risks, or political campaigns (Achrekar et al., 2011; Jung and Park, 2014). We used NodeXL for data collection. Node XL, developed by the Social Media Research Foundation, is a plugin for Microsoft Excel that allows the collection and analysis of multiple social media data (Smith et al., 2010). NodeXL was used in our study because it serves as a robust tool for data analysis and for deriving knowledge from complex social media interactions (Kim and Park, 2012).

3.2 Data Analysis

We relied on Social Network Analysis (SNA), a useful and reliable methodology for data analysis and visualization. Based on Perer and Shneidermans (2008) research, we measured various network indicators for the collected data. First, we examined social network graph metrics for the overall Premera Blue Cross data security breach, including number of vertices, edges, unique edges, and duplicate edges. Second, we analyzed vertices degrees, centrality measures, and page rank. Third, we plotted vertices metrics to identify information sources and brokers. Table 1 shows the definitions of the key terminologies related to SNA.

Metrics	Definitions
Vertex	A single element count of the primary entity of a network. In the case of Twitter, a vertex represents a Twitter user
Edge	An element that connects two vertices. In the Twitter context, an edge could be a tweet, a retweet (RT) or a mention (@)
Degree	This element measures the total number of edges connected to a particular vertex. In-degree measures the connections pointing inward to a vertex. Out-degree measures the connections originating from a vertex
Betweenness Centrality or Bridge Score	A metric that indicates how much disruption to other connections can cause the removal of a vertex in the network
Eigenvector Centrality	A metric that measures the quality of connections of a vertex. A vertex with higher connections yields a higher eigenvector value (PageRank is a variant of this metric)

Table 1: Definitions for metrics in SNA.

NodeXL calculates graph metrics related to SNS by using an algorithm developed by the Social Network Analysis Project (SNAP) at Stanford University (Leskovec et al., 2011).

4 Results and Discussion

Data collection yielded a total of 15 592 tweets from 8 689 unique Twitter accounts. Average geodesic distance is 6.16, with a maximum geodesic distance of 17, and a graph density of 0.00005451 (see Table 2). These results suggest low-affinity relationships between Twitter users in the Premera data security breach network. Table 3 and Table 4 show that top five words and hashtags were related to Premera Blue Cross data security breach, suggesting reliability of the collected data.

Graph Metric	Value
Graph Type	Directed
Vertices	8689
Unique Edges	7309
Edges with Duplicates	8283
Total Edges	15592
Maximum Geodesic Distance (Diameter)	17
Average Geodesic Distance	6.161185
Graph Density	5.5451E-05

Table 2: Overall graph metrics.

Top Words in Tweet in Entire Graph	Entire Graph Count
premera	10211
blue	4115
cross	3638
breach	3367
data	3287

Table 3: Top words counts in ”premera” network graph.

From an inspection of Figure 1, results indicate that Premera Blue Cross (@premera) took the lead on Twitter during communication of the crisis (Betweenness Centrality = 1753201.512, PageRank = 63.020546, In-degree = 205). This result suggests that the institution that received the attack (i.e., Premera) became the source of information for information seekers. Twitterers concerned about this data security breach turned to the Premera Blue Cross Twitter account to gather relevant and accurate information.

The Seattle Times (@Seattletimes) is a provider of news and information established in Seattle (Washington, USA), the same city in where Premera Blue Cross Headquarters is located. The geographical proximity between both institutions may explain the bridging role played by the former during the crisis. Considering that a great number of Premeras stakeholders are located in the Seattle area, results suggest that these stakeholders turned to this channel of local news for information about the Premera Blue Cross crisis.

Dark Reading (@darkreading), Brian Krebs (@briankrebs), TechCrunch (@techcrunch), Gary Davis (@garyjdavis), and SC Magazine (@scmagazine), hereinafter referred to as the tech community, represent IT security specialists or technology-specific news media. As Twitter participants recognize the tech community to be specialists in information security and hence a reliable source of information, they relayed information coming from their accounts, making them pivotal bridges during the Premera Blue Cross data security breach.

Top Hashtags in Tweet in Entire Graph	Entire Graph Count
infosec	701
premera	657
security	397
healthcare	357
databreach	308

Table 4: Top hashtags in ”premera” network graph.

In other words, Premera Blue Cross (@Premera) acted as a source of information on Twitter, while other key actors became intermediaries in relaying information about this data security breach. These findings are in line with previous research suggesting: (1) problem recognition and level of involvement predict information seeking and dissemination behavior (Yates and Paquette, 2011; Bird et al., 2012), and (2) a limited number of actors such as public figures, journalists, and mass media play an intermediary role during crisis communications (Perko, 2011).

5 Implications

Our study presents a new perspective that shows that data security is a cloud stakeholders issue rather than cloud providers responsibility. Also, our findings indicate that social media populations turned toward the application provider for accurate information during these events. Cloud providers should be prepared to take the lead, and this can be achieved by creating and reinforcing their social media presence. For instance, cloud providers could engage actively in risk communication by raising risk awareness on social media and educating their followers on security procedures.

Second, considering that the tech community and local news media play an intermediary role during risk communication on social media, cloud providers can engage in partnerships with IT secu-

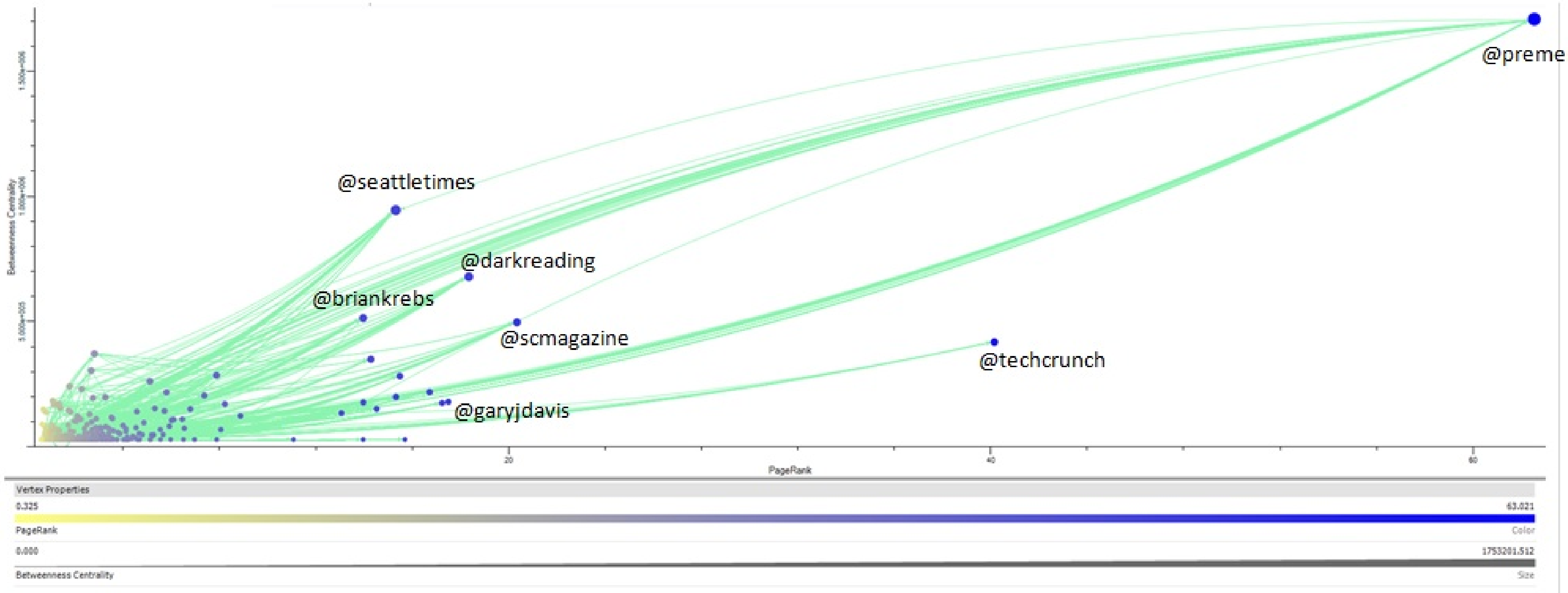

Figure 1: Vertex properties graph (PageRank on X-Axis and Betweenness Centrality on Y-Axis.

rity firms and specialists, local news media based on clients and partners location and recommend that all stakeholders follow those accounts for relevant and accurate information to safeguard data and mitigate damages.

6 Conclusions

This study highlights the need for a good risk communication during a security breach, which should involve all cloud-computing stakeholders. Our study makes recommendations on the steps to be taken by cloud providers to ensure that clients and partners remain reliably informed before and after any data security breach. Nevertheless, this study presents some limitations. First, the present study only analyzes risk communication for a data security breach coming from web applications, one of the core technologies of cloud computing. Second, the present study only considers risk communication on Twitter, a single popular social media.

Future steps in our study include the analysis of communication patterns and the inclusion of theories that may help to explain the phenomenon under study.

References

H. Achrekar, A. Gandhe, R. Lazarus, S. H. Yu, and B. Liu. 2011. Predicting flu trends using Twitter data. In *Proc. 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS 2011)*.

B. E. Aguirre. 2004. Homeland security warnings: Lessons learned and unlearned. *International Journal of Mass Emergencies and Disasters*, 22(2):103–115.

M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zahaira. 2010. Clearing the clouds away from the true potential and obstacles posed by this computing capability. *Communications of the ACM*, 53(4):50–58.

A. Benlian and T. Hess. 2011. Opportunities and risks of software-as-a-service: Findings from a survey of IT executives. *Decision Support Systems*, 52(1):232–246.

D. Bird, M. Ling, and K. Haynes. 2012. Flooding facebook: The use of social media during the queensland and victorian floods. *Australian Journal of Emergency Management*, 27(1):27–33.

W. Chan, E. Leung, and H. Pili. 2012. Enterprise risk management for cloud computing. *Committee of Sponsoring Organizations of the Treadway Commission*.

V. T. Covello, P. M. Sandman, and P. Slovic. 1998. *Risk communication, risk statistics and risk comparisons: A manual for plant managers*. Chemical Manufacturers Association, Washington D.C.

Experian. 2015. *Data breach industry forecast*. Experian Data Breach Resolution.

I. Foster, Y. Zhao, I. Raicu, and S. Lu. 2008. Cloud computing and grid computing 360-degree compared. In *Proc. Grid Computing Environments Workshop 2008 (GCE 2008)*.

Gartner. 2011. *Gartner identifies the top 10 strategic technologies for 2012*. Gartner.

J. P. G. Gashami, Y. Chang, J. J. Rho, and M.-C. Park. 2015. Privacy concerns and benefits in SaaS adoption by individual users: A trade-off approach. *Information Development*. doi:10.1177/0266666915571428.

R. Goolsby. 2010. Social media as crisis platform. *ACM Transactions on Intelligent Systems and Technology*, 1(1):1–11.

B. Grobauer, T. Walloschek, and E. Stcker. 2011. Understanding cloud computing vulnerabilities. *IEEE Security and Privacy*, 9(2):50–57.

P. T. Jaeger, J. Lin, and J. M. Grimes. 2008. Cloud computing and information policy: Computing in a policy cloud? *Journal of Information Technology and Politics*, 5(3):269–283.

K. Jung and H. W. Park. 2014. Citizens social media use and homeland security information policy: Some evidences from twitter users during the 2013 North Korea nuclear test. *Government Information Quarterly*, 31(4):563–573.

A. M. Kaplan and M. Haenlein. 2010. Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 53(1):59–68.

M. Kim and H. W. Park. 2012. Measuring twitter-based political participation and deliberation in the South Korean context by using social network and Triple Helix indicators. *Scientometrics*, 90(1):121–140.

N. J. King and V. T. Raja. 2012. Protecting the privacy and security of sensitive customer data in the cloud. *Computer Law and Security Review*, 28(3):308–319.

R. K. L. Ko, P. Jagadpramana, M. Mowbray, S. Pearson, M. Kirchberg, Q. Liang, and B. S. Lee. 2011. TrustCloud: A framework for accountability and trust in cloud computing. In *Proc. 2011 IEEE World Congress on Services (SERVICES 2011)*.

H. Kwak, C. Lee, H. Park, and S. Moon. 2010. What is Twitter: A Social Network or a News Media? In *Proc. International World Wide Web Conference Committee (IW3C2)*.

P. Lagadec. 2002. Crisis management in france: Trends, shifts and perspectives. *Journal of Contingencies and Crisis Management*, 10(4):159–172.

J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. 2011. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123.

S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi. 2011. Cloud computing the business perspective. *Decision Support Systems*, 51(1):176–189.

A. W. Matthews and D. Yadron. 2015. Premera Blue Cross says cyberattack could affect 11 million members. `http://www.wsj.com`. [Online; accessed 15-November-2015].

P. Mell and T. Grance. 2011. The NIST definition of cloud computing [Recommendations of the National Institute of Standards and Technology-Special Publication 800-145]. `http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf`. [Online; accessed 15-November-2015].

J. M. Nigg. 2006. Communication under conditions of uncertainty: Understanding earthquake forecasting. *Journal of Communication*, 32(1):27–36.

A. Perer and B. Shneiderman. 2008. Integrating statistics and visualization: Case Studies of gaining clarity during exploratory data analysis. In *Proc. Human Factors in Computing Systems (CHI'08)*.

T. Perko. 2011. Importance of risk communication during and after a nuclear accident. *Integrated Environmental Assessment and Management*, 7(3):388–392.

L. Rainie and M. Duggan. 2014. Heartbleeds Impact. `http://www.pewinternet.org/2014/04/30/heartbleeds-impact/`. [Online; accessed 15-March-2016].

M. Smith, B. Shneiderman, N. Milic-Frayling, E. M. Rodrigues, V. Barash, C. Dunne, T. Capone, A. Perer, and E. Gleave. 2010. NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010. `http://nodexl.codeplex.com/`. [Online; accessed 15-March-2016].

J. Staten, L. E. Nelson, D. Bartoletti, L. Herbert, W. Martorelli, and H. Baltazar. 2014. *Predictions 2015: The days of fighting the cloud are over*. Forrester.

Y. Theocharis. 2013. The wealth of (occupation) networks? communication patterns and information distribution in a twitter protest network. *Journal of Information Technology and Politics*, 10(1):35–56.

Verizon. 2014. 2014 data breach investigations report. *Verizon Business Journal*, 2014(1):1–60.

D. Wright and M. Hinson. 2009. An analysis of the increasing impact of social and other new media on public relations practice. In *Proc. 12th Annual International Public Relations Research Conference*.

W.-W. Wu, L. W. Lan, and Y.-T. Lee. 2011. Exploring decisive factors affecting an organizations saas adoption: A case study. *International Journal of Information Management*, 31(6):556–563.

D. Yates and S. Paquette. 2011. Emergency knowledge management and social media technologies: A case study of the 2010 haitian earthquake. *International Journal of Information Management*, 31(1):6–13.

L. Youseff, M. Butrico, and D. Silva. 2008. Toward a unified ontology of cloud computing. In *Proc. Grid Computing Environments Workshop (GCE 2008)*.

X. Zhang, N. Wuwong, H. Li, and X. Zhang. 2010. Information security risk management framework for the cloud computing environments. In *Proc. IEEE 10th International Conference on Computer and Information Technology (CIT 2010)*.

M. Zhou, R. Zhang, W. Xie, W. Qian, and A. Zhou. 2010. Security and privacy in cloud computing: A survey. In *Proc. 6th International Conference on Semantics, Knowledge and Grids*.