Selangor Science &Technology Review
Special Issue: Science and Technology for Society
Vol 5 No 5 (2021)

Preliminary Literature Review on Data Breach

Nur Azlan Abdul Rahim, Norazmir Mohd Nordin

nurazlan@raudah.usim.edu.my

norazmir@usim.edu.my

Abstract: In Malaysia, the digital economic revolution is seen as significant especially when the country is hit by the Covid-19 disaster. Various products, especially food, are ordered through digital systems. This is due to the movement control order (MCO) imposed by the government which limits the travel of the people. Indirectly, this also causes an increase in data in the virtual world as a result of ordering goods earlier. What is worrying is that this scenario could attract data rogue or cybercrime, especially data breaches. Worried about being too passionate about filling out personal data on a website, not realizing that it is a fake website. Thus, the users or customers were exposed to his or her data to an authorized party. This certainly threatens consumers when their personal information can be misused for personal or criminal purposes. Thus, this paper will discuss the importance of implication issues and challenges in the data breach.

Keywords: Cybercrime, Jabatan Perlindungan Data Peribadi, Personal Data, Personal Data Protection Act

1. Introduction

In Malaysia, there are so many websites related to shopping online or E-Commerce sites. Among them are *Lazada, Shopee, Mudah, Lelong, Ezbuy* and *Zalora*. These websites are the top 10 websites for E-Commerce (Lab, 2020). This proves that the trend of Malaysians now prefers to shop online, especially when Covid-19 or movement control order (MCO). In addition to saving time and convenience, it can even limit the spread of the Covid-19 virus among locals. This will cause the data volume to increase and the network to become dense. This is followed by various advertisements that captivate customers provided by the relevant parties. Business in cyberspace is very fun because without the need for physical business space.

When a company has a database of customers, this means they already have targets and potential customers. So, it is easy for them to lure customers when they already know their age, location, phone number and occupation. All products or services will be customized for that group. What if these data leak and get to an unauthorized party ? The data breach is a very serious disaster or threat to the organization. This data involves a large and complex amount. It is not the only threat to the customer, in fact, it also will damage the company's reputation and cost of finance (Cheng et al., 2017). The definition of a data breach is *the intentional or inadvertent exposure of confidential information to unauthorized parties* (Cheng et al., 2017).

National Cyber Security Centre (NCSC), based in London, England defined data breach as *occurs when information held by an organisation is stolen or accessed without authorisation* (NCSC, 2022). Also included in this category are information leakage, data leak and unintentional information disclosure. While anti-virus software is popular, Trend Micro defined it as *an incident where information is stolen or taken from a system without the knowledge or authorization of the system's owner* (Micro, 2020). This sensitive data is included as well as trade secret, proprietary and credit card numbers. National Initiative for Cybersecurity Careers and Studies (NICCS) mentioned that cybersecurity is described as *the activity or process, ability or capability or state whereby information and communications* *systems and the information contained therein are protected from and/or defended against damage, unauthorized use or modification or exploitation* (NICCS, 2017).

Meanwhile, according to IBM Security via their 2019 Cost of Data Breach Report, defined data breach as *an event in which an individual's name and a medical record and/or a financial record or debit card is potentially put at risk either in electronic or paper format* (IBM, 2019). In Malaysia, the department responsible for handling this affair is *Jabatan Perlindungan Data Peribadi* (JPDP), located at Putrajaya, Federal Territory This department is responsible for implementing the Personal Data Protection Act, 2010 (PDPA). This act is defined as *the processing of personal data in commercial transactions and to provide for matters connected therewith and incidental thereto* (Personal Data Protection Act 2010, 2010). In website JPDP, it defined Personal Data as *any information used in a commercial transaction that is directly or indirectly related to a data subject identified from that information* (JPDP, 2020a).

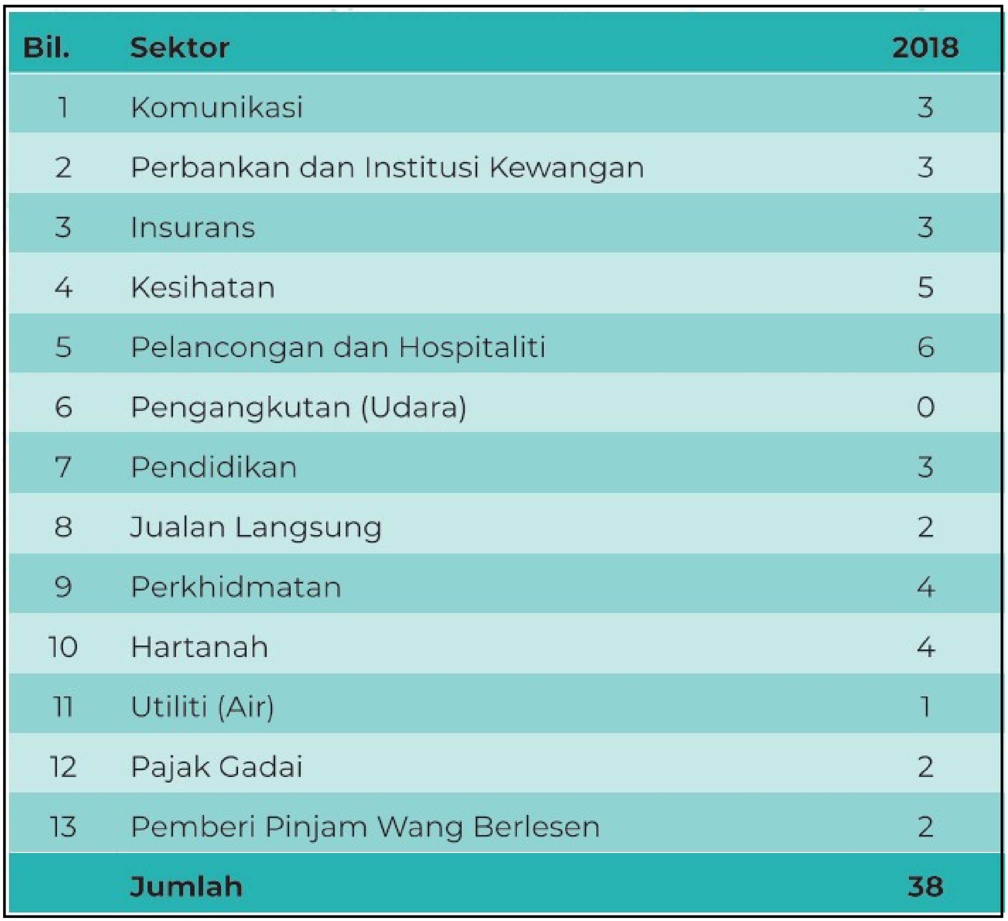
Also, this personal data can be recorded either via manual or electronic whether involving objective or subjective, regardless of where the information is obtained from. JPDP has outlined that the following information is "personal data". There are:-

- Name and address
- Identity card number
- Passport number
- Health information
- E-mail address
- Picture
- An image in closed circuit recording (CCTV)
- Information in personal files
- Bank account details
- Credit card details

JPDP also defined Open Data as *data that can be used freely, can be shared and reused by citizens, Public or Private Sector agencies* (JPDP, 2020b). The purpose is to increase the transparency of public service delivery and the quality of government information. Besides, the goal is also a condition of global evaluation in E-Government services. Therefore, based on all this information, the data breach should not be taken lightly as it is considered a criminal act.

In 2018, JPDP has done 38 series of inspection visits were conducted on registered Data Users processing personal data. There was an increase of six series of inspection visits for 2018 compared to 2017 (Jabatan Perlindungan Data Peribadi Malaysia, 2018). These inspection activities are followed under Section 101 of 2018 law. Table 1 below is the list of these activities.

Table 1. Number of Inspection Visits under Section 101- By Group of Data Users



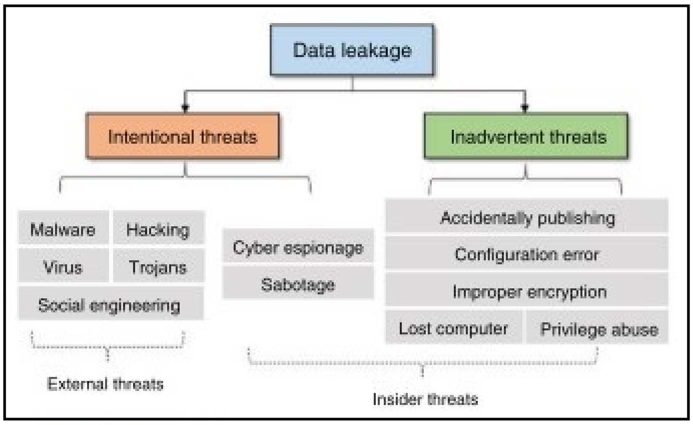
Bil.	Sektor	2018
1	Komunikasi	3
2	Perbankan dan Institusi Kewangan	3
3	Insurans	3
4	Kesihatan	5
5	Pelancongan dan Hospitaliti	6
6	Pengangkutan (Udara)	0
7	Pendidikan	3
8	Jualan Langsung	2
9	Perkhidmatan	4
10	Hartanah	4
11	Utiliti (Air)	1
12	Pajak Gadai	2
13	Pemberi Pinjam Wang Berlesen	2
	Jumlah	**38**

Source: Jabatan Perlindungan Data Peribadi Malaysia, (2018)

The organization that receives and processes customer data, is at great risk as it is its responsibility for the security of the data. According to the PDPA Act (Personal Data Protection Act 2010, 2010), those who transfer personal data abroad without a valid authority will be fined RM 300,000 or imprisoned for a term not exceeding 2 years or both. While for those illegal collections of personal data, a fine not exceeding RM 500,000 or/and imprisoned for a term not exceeding 3 years can be imposed.

2. Issues and Implication

This study examined a preliminary systematic literature review from various sources of data collections such as Al Quran, journals, proceedings, reports, newspapers and websites. Review and compilation of these data then from here generate the contexts such as issues, implication, challenges and recommendations. In the information technology world, data is an important component of the enterprise. When the data tend to be large scale and growing, it needs a system that can detect and prevent loss or hacking. Due to the nature of the data itself being sensitive and the source of channels for marketing, the position of data in the field of business has been considered "noble". From the study of Tanriverdi et al., (2019), the data breach was ranked third after extreme weather & natural disaster. The mediums of data leakage are many. Among them are email, file transfer protocol (FTP), social media, cloud file sharing, website, social network, camera, firewall system weakness, poor database maintenance, laptop theft and the like (Cheng et al., 2017). This data leakage can be divided into two categories, intentional and inadvertent threats, as shown in Figure 1 below.



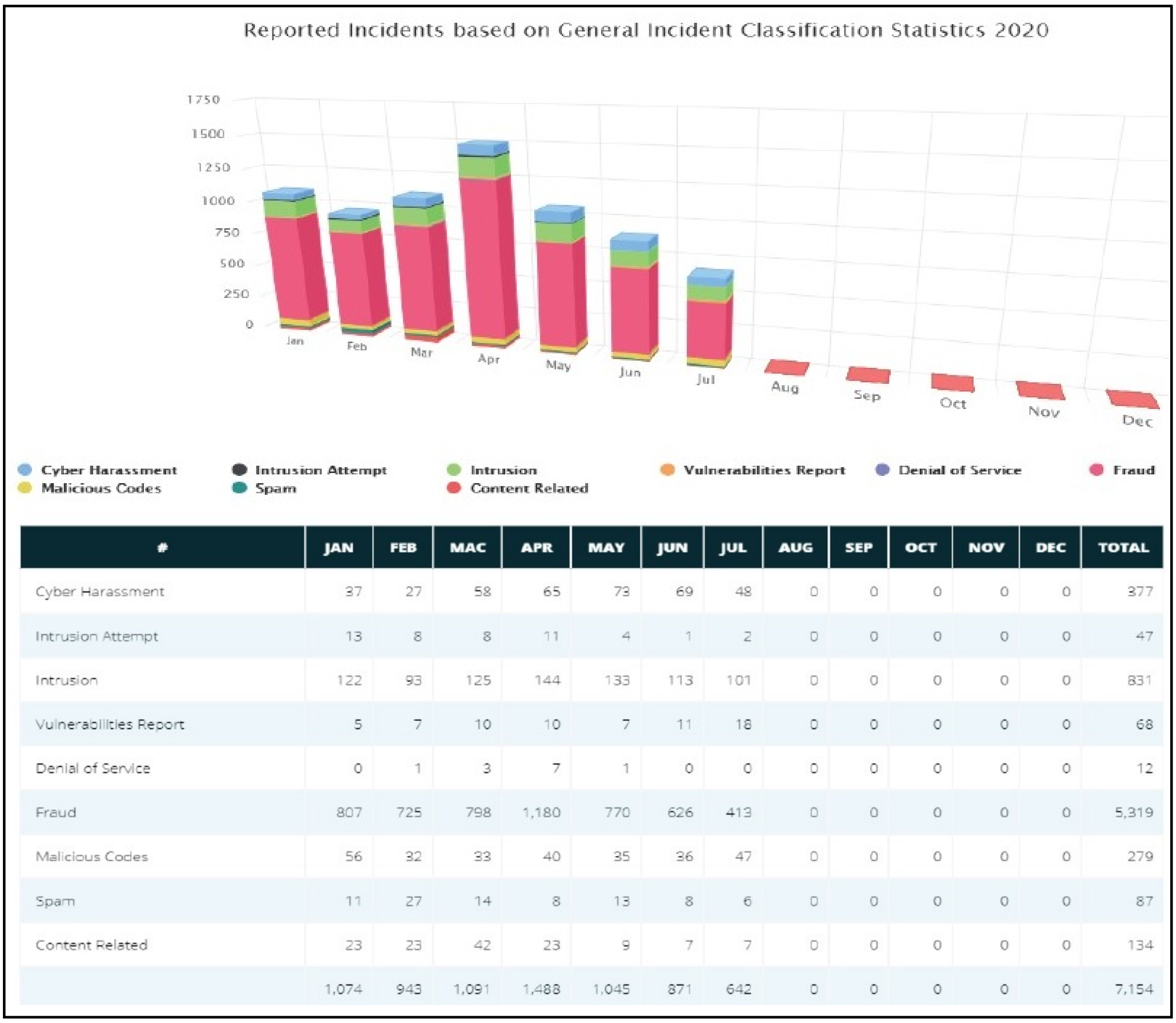
Source: Cheng et al., (2017)

Fig. 1. Classification of enterprise data leak threats

From Figure 1, these are the elements of an enterprise or organization that always needs to be vigilant and caring. Data leakage also can be from an outsider or insider threat. It also needs to be judged on one's intentions whether intentionally or not. Normally, external parties such as a hacker, malware, virus and social engineering are interested in leaking this data. Social engineering such as phishing is one easy way to get sensitive information when cybercrime creates a website that is very similar to the authorities or companies related to the company. Due to staff negligence, the data breach occurred.

The cause of internal data leakage can be inadvertently or deliberate actions. An example inadvertently is when staff transmitted important data with low encryption technology, whereas an example of deliberate actions are staff grievances or sabotage. Phishing and data breach crimes have two things in common, fraud. The goal is to take sensitive information by deceiving the public for their illegal benefit. Only in terms of its implementation or methodology is different. The rising trend of cybercrime especially in terms of fraud is somewhat worrying. These statistics of CyberSecurity Malaysia shows that fraud cases have increased significantly compared to other cybercrimes. A total of 5,319 cases were recorded from January to July 2020. Followed by intrusion crimes, a total of 831 cases (Malaysia, 2020).

All data above has been shown in Figure 2 as below:-

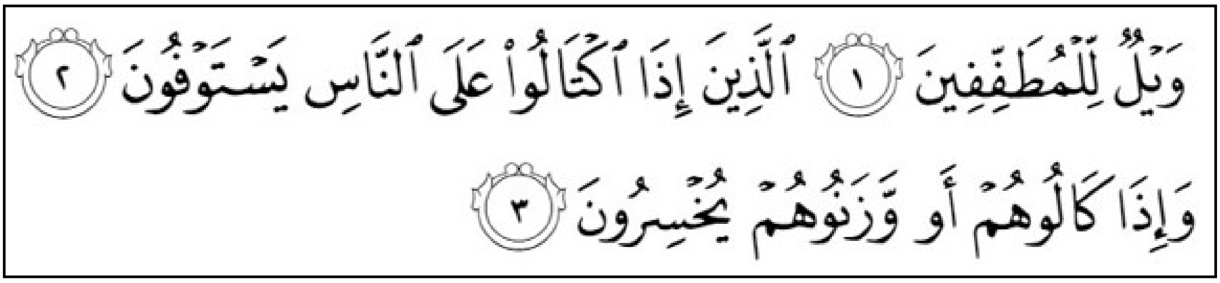


#	JAN	FEB	MAC	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
Cyber Harassment	37	27	58	65	73	69	48	0	0	0	0	0	377
Intrusion Attempt	13	8	8	11	4	1	2	0	0	0	0	0	47
Intrusion	122	93	125	144	133	113	101	0	0	0	0	0	831
Vulnerabilities Report	5	7	10	10	7	11	18	0	0	0	0	0	68
Denial of Service	0	1	3	7	1	0	0	0	0	0	0	0	12
Fraud	807	725	798	1,180	770	626	413	0	0	0	0	0	5,319
Malicious Codes	56	32	33	40	35	36	47	0	0	0	0	0	279
Spam	11	27	14	8	13	8	6	0	0	0	0	0	87
Content Related	23	23	42	23	9	7	7	0	0	0	0	0	134
	1,074	943	1,091	1,488	1,045	871	642	0	0	0	0	0	7,154

Source: Malaysia, (2020)

Fig. 2. Reported Incidents based on General Incidents Classification Statistic 2020

From the Islamic perspective, Allah SWT has mentioned fraud in Al Quran via Surah Al Muthaffifin, verses 1-3 as below:-



وَيْلٌ لِّلْمُطَفِّفِينَ ۝١ ٱلَّذِينَ إِذَا ٱكْتَالُوا۟ عَلَى ٱلنَّاسِ يَسْتَوْفُونَ ۝٢
وَإِذَا كَالُوهُمْ أَو وَّزَنُوهُمْ يُخْسِرُونَ ۝٣

Source: Al Quran

Fig. 3. Surah Al Muthaffifin, verses 1-3

“Woe to those who give less [than due]. Who, when they take a measure from people, take in full. But if they give by measure or by weight to them, they cause loss”. These verses clearly stated that in Islam, we cannot tolerate fraud and it is not ethical for the Muslims. The effects of a data breach and even fraud are terrible and can even threaten one's dignity. A company can face civil cases by the government as well as lawsuits by victims. Even worse, a person can be depressed as a result of his personal information being disseminated or becoming a public focus. Not surprisingly, it can eventually lead to suicide.

Refer to the case Equifax in 2017. There are over 145 million customers' data have been reported stolen. This resulted from breach-related costs of close to $90 million, whereby 240 customers' lawsuits actions have been taken (Khan et al., 2019). A report from IBM Security and Ponemon Institute stated that in 2019, the global total cost of a data breach is $3.92 million with an average size of a data breach of 25,575 records. The data was collected between July 2018 and April 2019. The cost per lost record is $150.00 (IBM, 2019). This report consists of more than 500 companies in-depth interviews around the world. Due to the cost being very impressive, therefore, it is very important for any organization needs to implement high-security mechanisms to prevent a data breach. In Malaysia, the well-known airline Malindo Airways (OD) has been the victim of a data breach for its passengers. 30 million customer data has been stolen in September last year (Augustin, 2020).

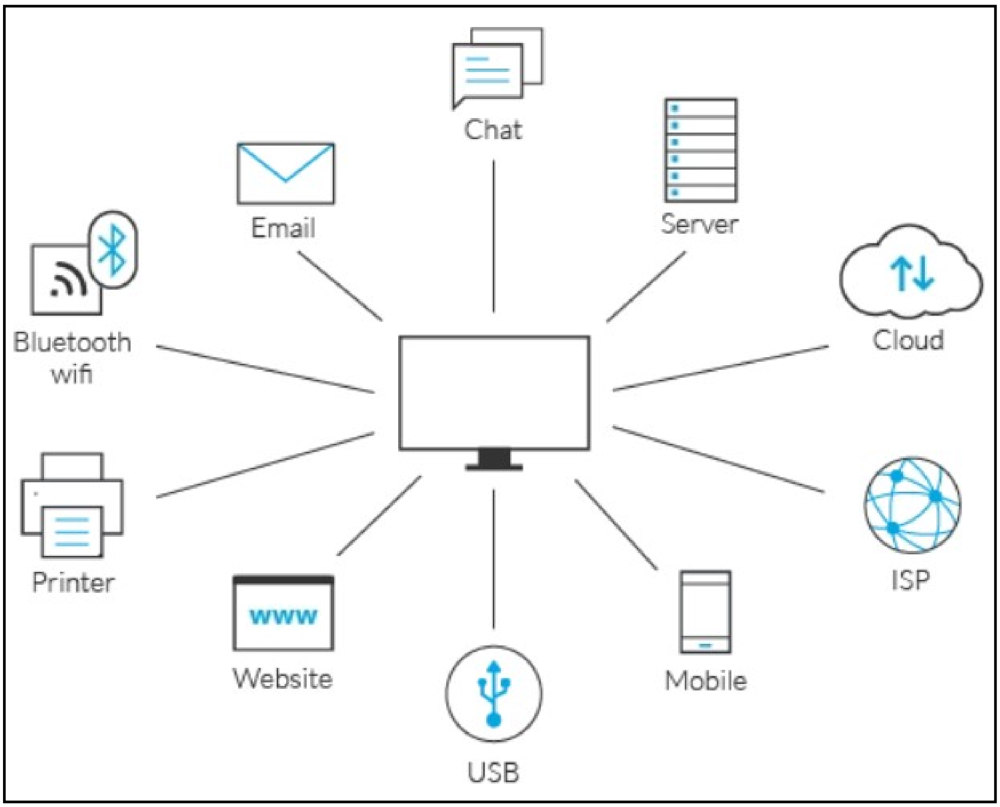
This action has been done by two former employees of the company's e-commerce service provider from GoQuo (M) Sdn Bhd at its development centre in India. According to the CEO, Captain Mushafiz Mustafa Al-Bakri, police reports were filed in Malaysia and India. He added, this case also has been referred to the National Cyber Security Agency (NCSA), JPDP and other authorities abroad (Bernama, 2019). In another report from Astro Awani, Kementerian Komunikasi dan Multimedia (MCMC) is still waiting for this report from Malindo Air. It does not set any time limit but hopes to get it as soon as possible (Awani, 2019).

Recently a piece of news from Singapore’s The Straits Times (ST) reported that our Navy, Tentera Laut Diraja Malaysia (TLDM), had leaked the sensitive document via the website, namely Dark Web (Hakim, 2020). According to ST, the source of hacked come from TLDM personnel emails. About 70 documents were released that have been stolen. This data breach included a conversation between a United States (US) Navy vessel and TLDM, which the Malaysian Army Forces has denied such allegation. Recognizing the dangers of a data breach and its impact on consumers, companies and countries, the government has taken proactive steps. Former Minister of Communications and Multimedia, Gobind Singh Deo said that the government is ready to enact Act 709 to be stricter. This includes convicting the party who obtained the data illegally. He added that this is in line with the development of the digital economy and the importance of e-commerce (Mutalib, 2019).

In line with the sophistication of the digital technology world today, the software has been created to detect and prevent data from being invaded by irresponsible parties. Generally, this software is named Data Loss Prevention software (DLPS). Cisco System Inc (Cisco), a multinational conglomerate corporation specialist in network, hardware, software and telecommunication equipment defined DLPS as *a set of technologies, products and techniques that are designed to stop sensitive information from leaving an organization* (Cisco, 2022). Meanwhile, anti-virus famous software, Norton described DLPS as *the software tools and processes used to protect sensitive data and detect the presence of malicious actors looking to get their hands on your data* (Norton, 2021). Today, there are plenty of companies that develop this software. Such as SolarWinds Data Loss Prevention with ARM, CoSoSys Endpoint Protector, Symantec Data Loss Prevention, SecureTrust Data Loss Prevention, Clearswift Adaptive DLP, McAfee Data Loss Preventing Tool and many more (Comparitech, 2021).

The structure of DLPS how it works from various sources, as described in Figure 4. Typically this system or software functioning as below (Imperva, 2021):-

- Achieve data visibility in large organizations.
- Protect Intellectual Property critical for the organization.
- Secure data on remote cloud systems.
- Protect Personally Identifiable Information (PII) and comply with relevant regulations.
- Secure mobile workforce and enforce security in Bring Your Own Device (BYOD) environments.



Source: Imperva, (2021)

Fig. 4. DLPS: Preventing data loss from various sources

3. Challenges

Handing big data is not easy. From one angle, it is good to have customer data, so the enterprise or company has its loyal customers or existing customers. Therefore, it will reduce the cost and time of advertising. However, it is too risky to handle it. Nowadays, due to a disease outbreak, people are now working from home (WFH), as a new norm. Another terminology similar to this is the Site Office Home Office (SOHO). Thus, the tendency for data leaks is very high. Not many organizations develop their standard operating procedure (SOP) while dealing with sensitive data at home during a pandemic.

Therefore, this is the best time for cybercrime to implement phishing or fraud activities. The pandemic outbreak is now a new challenge for a data breach environment. Below are among the technical challenges in data breach detection (Cheng et al., 2017):-

- Privacy Concern

As we know, a major concern is to preserve sensitive data from leaking. The capability to preserve privacy will be questioned when Data Leak Prevention and Detection (DLPD) software comes from outsourcing. There is no 100% guarantee when outsourcing DLPD from a third party that data will not be leaked.

- Scalability
  Scalability is the crux to accurately deal with the massive enterprise-scale amount of growth of data. From megabytes to gigabytes then to terabytes. When data is too big and complex, it is then distributed via Cloud-based. And the cloud-based is owned by a third party. There is a possibility of data leaking in the Cloud. Reducing the scalable will be easier to manage and control the data during the process of prevention.

- Timeliness
  When handling big data that contains variety and velocity, it is not easy to detect and respond as immediately when it has a data breach activity. It probably will take time for detection. Thus, this data will be harmed by cybercrime before can be tracked by the system.

4. Data Breach Resolutions

In order to minimize, control or avoid a data breach, there are a few suggestions or recommendations. A few are taken from journals or report on this subject or field. There are:-

- *Install and run DLPS*. Is not deniable some of the data leakings are accidentally or from internal staff. Therefore the organization must have some security mechanism to prevent the outbreak so fast. The usage of DLPS is recommended even the technology is changing so fast. Especially for firms that have large data to be stored and manipulated like supermarkets and hospitals.
- *Service recovery*. When data was leaking, nothing much we can do to control 100%. It is like undercooked steaks at a restaurant. Therefore the firms can use numerous restoration plans, ranging from a transparent apology to the product or service replacement. This issue has related to positively encouraging customer perceptions (Kude et al., 2017). Firms have to give compensation to the affected customers. For example case Sony PlayStation network. When its network has been breached in 2011, Sony offer free downloadable content to their customers.
- *Enhanced the law*. Given that there has been a significant increase in data breach cases in addition to adverse effects on the economy, security, public and image, it is time for the government to improve the existing laws to be more comprehensive. The government is aware of this existing law when the minister was talking during the parliament conference (Mutalib, 2019). Personal Data Protection Act 2010 needs to be given a new lease of life as our country and the world are rapidly heading into the era of digitalization,
- *Awareness and cultivating personal data security*. The government via its agency, JPDP keep moving these activities to the communities. It has been implemented via electronic advertising, website, social media, speech and conference. These can be referred to as their Annual Report 2018 (Jabatan Perlindungan Data Peribadi Malaysia, 2018).

- *Educate the young generation*. As parents, they have the responsibility to send their children to schooling. But at the same time, the parents must monitor their children gadgets and educate them will jeopardize the security of personal data via social media such as Tik Tok, YouTube, Messenger and others.

5. Conclusion

A data breach is a criminal act under Cybercrime. It should not be underestimated by all parties. The impact is significant whether it is economic, social, political, military or business. Loss of money and dignity is also an effect of the data breach. From a religious point of view, it is illegal (*haram*) and the perpetrator will be punished by Allah SWT. No party likes its data being used by third parties for personal gain or as a ransom. Therefore, the strictness and enforcement of the Cybercrime law should be the priority of the relevant authorities to ensure that the issue of a data breach could be dealt with accordingly.

References

Al Quran.

Augustin, S. (2020). *Malindo Air hauled to court over Data Breach*. FMT News. https://www.freemalaysiatoday.com/category/nation/2020/02/20/malindo-air-hauled-to-court-over-data-breach/

Awani, A. (2019). *KKMM masih tunggu laporan isu pencerobohan data penumpang*. Astro Awani. http://www.astroawani.com/video-malaysia/kkmm-masih-tunggu-laporan-isu-pencerobohan-data-penumpang-1810075

Bernama. (2019). *Kebocoran data pelanggan telah ditangani, kata Malindo Air*. Astro Awani. http://www.astroawani.com/berita-malaysia/kebocoran-data-pelanggan-telah-ditangani-kata-malindo-air-218664

Cheng, L., Liu, F., & Yao, D. D. (2017). Enterprise Data Breach: Causes, Challenges, Prevention & Future Directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *7*(5), 1–14. https://doi.org/10.1002/widm.1211

Cisco. (2022). *What Is Data Loss Prevention (DLP) ?* https://www.cisco.com/c/en/us/products/security/email-security-appliance/data-loss-prevention-dlp.html

Comparitech. (2021). *13 Best Data Loss Prevention Software Tools*. Comparitech Limited. https://www.comparitech.com/data-recovery-software/data-loss-prevention-tools-software/

Hakim, A. (2020). *Sensitive TLDM Documents Hacked & Leaked On Dark Web*. New Strait Times. https://www.msn.com/en-my/news/national/sensitive-tldm-documents-hacked-and-leaked-on-dark-web/ar-BB182UxE?li=BBr8Hnu&ocid=mailsignout

IBM. (2019). Cost of a Data Breach Report 2019. In *IBM Security*.

Imperva. (2021). *Data Loss Prevention (DLP)*. Imperva Website. https://www.imperva.com/learn/data-security/data-loss-prevention-dlp/

Personal Data Protection Act 2010, 1 (2010).

Jabatan Perlindungan Data Peribadi Malaysia. (2018). *Laporan Tahunan Jabatan Perlindungan Data Peribadi Malaysia 2018*.

JPDP. (2020a). *Data Peribadi*. JPDP. https://www.pdp.gov.my/jpdpv2/awam/data-peribadi/

JPDP. (2020b). *Data Terbuka*. JPDP. https://www.pdp.gov.my/jpdpv2/pusat-media/data-terbuka/

Khan, F., Kim, J. H., Moore, R., & Mathiassen, L. (2019). Data breach risks and resolutions: A literature synthesis. *25th Americas Conference on Information Systems, AMCIS 2019*, 1–10.

Kude, T., Hoehle, H., & Sykes, T. A. (2017). Big Data Breaches & Customer Compensation Strategies: Personality traits and social influence as antecedents of Perceived Compensation. *International Journal of Operations and Production Management*, *37*(1), 1–20. https://doi.org/10.1108/IJOPM-03-2015-0156

Lab, M. S. (2020). *Top 10 E-Commerce Site in Malaysia 2020*. Marketing Signal Lab. https://marketingsignallab.com/top-ecommerce-sites-in-malaysia/

Malaysia, C. (2020). *General Incident Statistics 2020* (p. 1). MyCert Malaysia.

Micro, T. (2020). *Data Breach*. Trend Micro. https://www.trendmicro.com/vinfo/us/security/definition/data-breach

Mutalib, Z. A. (2019). *Kerajaan kaji perketat Akta Perlindungan Data Peribadi 2010*. BH Online. https://www.bharian.com.my/berita/nasional/2019/11/628206/kerajaan-kaji-perketat-akta-perlindungan-data-peribadi-2010

NCSC. (2022). *Data Breaches: Guidance for individuals and families. How to protect* *yourself from the impact of data breaches*. NCSC Website. https://www.ncsc.gov.uk/guidance/data-breaches#section_1

NICCS. (2017). *Cybersecurity Glossary*. NICCS. https://niccs.us-cert.gov/about-niccs/cybersecurity-glossary#C

Norton. (2021). *Data loss prevention: What is DLP + how does it work?* https://us.norton.com/internetsecurity-emerging-threats-data-loss-prevention.html

Tanriverdi, H., Nwankpa, J., & Hall, E. (2019). Structural Complexity & Data Breach Risk. *Fortieth International Conference on Information Systems, Munich*, 1–18.