Security and Privacy Challenges of Big Data Adoption: A Qualitative Study in Telecommunication Industry

https://doi.org/10.3991/ijim.v16i19.32093

Syarulnaziah Anawar[1](✉), Nur Fadzilah Othman[1], Siti Rahayu Selamat[1], Zakiah Ayop[1], Norharyati Harum[1], Fiza Abdul Rahim[2]
[1] Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2] Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
syarulnaziah@utem.edu.my

Abstract—The telecommunication industry is the leading industry in big data trends as this industry has the most capable infrastructure for big data. However, the adoption of big data in telecommunication services also raises important security and privacy challenges due to the high volume, velocity, and variety of big data characteristics. To address the issue, this study shed light on the security and privacy challenges of big data adoption in the telecommunication industry. This study focuses on investigating the security and privacy challenges for data users in telecommunication services from the lens of the technological, organisational, and environmental (TOE) framework and examines the mitigation strategies to address the privacy and security challenges. This study is conducted using a focus group qualitative methodology. From the perspectives of data users (telecommunication providers), it could be concluded that data management, data privacy, data compliance, and regulatory orchestration challenges are the most pressing concerns in big data adoption. This study offers contributions in presenting a thematic classification of security and privacy challenges and their mitigation strategies for big data adoption in the telecommunication industry. The thematic classification highlights potential gaps for future research in the big data security domain. This study is significant in that it provides empirical evidence for the perspectives of telecommunication data users in addressing privacy and security issues that are related to big data adoption.

Keywords—big data, security, privacy, telecommunication, TOE framework

1 Introduction

Big data analytics has the capacity to extrapolate trends and patterns to predict the behaviour of a given population or even of individuals [1]. The telecommunication industry is the leading industry in big data trends as the industry has the most capable infrastructure for big data [2]. With the roll-out of 5G technology, numerous emerging big data technologies take place to benefit from improved connectivity. The collection of geolocation data of telecommunication subscribers has opened many opportunities

in collecting continuous and real-time data [3] as the service provider is able to obtain geolocation data without internet services through the cellular network protocol once the subscriber turns on their mobile devices. The capacity may greatly benefit most areas of government services by enabling surveillance systems, cybersecurity, and public safety and defence.

The emergence of big data through telecommunications has led to a large amount of sensitive data being generated, particularly geolocation information. Despite the potential advantages of big data, automated data collection by telecommunication service providers is not without scrutiny as it may pose privacy and security challenges. In telecommunication services, data privacy is usually defined through privacy policies that describe the purpose of data usage and how the data is handled and disclosed to protect the identity of the users. However, the availability of the privacy policy in a subscription contract does not automatically induce a subscriber's willingness to share information.

Privacy and security risks may vary depending on the purpose and types of collected data in the big data application, and the type of framework used in developing the application. Many big data applications in Malaysia are considered privacy-invasive because these applications adopted a centralised architecture, where all collected data is stored on a central server. Data breaches are the main threat in big data applications. Therefore, the telecommunication service application must adopt an open-source framework that allows system transparency for the public to test and suggest measures to correct vulnerabilities in the big data application. However, the lack of an open framework is to be expected as the requirement for the telecommunication service provider to protect personal information becomes complicated due to the uncertain reliability of data de-identification. The data user's best efforts in de-identifying personal identifiable information (PII) may not prevent the re-identifying of an individual because data could be combined with other sources [4].

On the other hand, the use of predictive analytics in big data will also lead to significant concerns and controversial overreach by the telecommunication service provider. The technologies can be used to collect massive amounts of data about entire populations that include demographic information like gender, race, and living location (urban vs rural). The demographic information might lead to marginalisation and stigmatisation of a particular demographic segment [5] due to the sensitivity of the information. For example, predictive analytics may lead to racial profiling by public safety authorities to target certain groups of people that are believed to pose a greater risk of criminal behaviour [1].

Currently, research that focuses on security and privacy challenges in the telecommunication industry is scarce. Specifically, there is a general lack of research that empirically analyse the security and privacy challenges for data users or telecommunication providers and how it is related to big data adoption in telecommunication services. Therefore, the aim of this study is to empirically analyse the security and privacy challenges from the perspectives of data users in the telecommunication industry and examine the mitigation strategies to address the privacy and security challenges. Guided by the technological, organisational, and environmental (TOE) framework, this study seeks to address the following research questions: (RQ1) What are the perceived

security and privacy risks by the telecommunication provider for big data adoption? (RQ2) What are the specific mitigation strategies to address the security and privacy challenges?

This study offers contributions in presenting a thematic classification of security and privacy challenges and their mitigation strategies for big data adoption in the telecommunication industry. The thematic classification highlights potential gaps for future research in the big data security domain. This study is significant in that it provides empirical evidence for the perspectives of telecommunication data users in addressing privacy and security issues that are related to big data adoption. The outcome of this study may serve as guidelines for regulators, telecommunication providers, and stakeholders for securing big data systems and promoting security best practices within telecommunication industry operations.

The rest of the paper is organized as follows. Section 2 presents related work that focuses on the security and privacy challenges in big data. In Section 3, the research details of the focus group qualitative methodology are presented. Finally, the results of data analysis and discussion on the results are presented in Sections 4 and 5. This paper is concluded in the last section.

2 Literature review

Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse [6]. Gartner et al. [7] further define big data as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making". Based on the definition by Gartner, [8] highlights the main characteristics of big data namely Volume (data size), Velocity (data speed), Variety (various data formats), Value (benefits), and Veracity (integrity and trustworthiness).

Privacy and security challenges in big data adoption have been discussed in many pieces of literature. However, there is a general lack of research that investigate security and privacy challenges specifically in the telecommunication industry. Among a few are the works done by [9]-[11]. Apart from [10], there is a general lack of research that empirically investigates the security and privacy challenges for data users and how it is related to big data adoption in telecommunication services. This review will analyse previous work on privacy and security challenges from the lens of the TOE framework [30]. The framework is usually used to explain the influence of technology, organisational, and external task context on IT adoption.

In the technological context, most issues revolve around the need for strong security and privacy solutions to protect the high volume of data that is collected in a distributed manner. The potential multipoint infrastructure failure in a distributed framework [11]-[13] will easily lead to security risks in the big data system. The main challenge that appears in all studies is the granular access control issues [9-10] due to the large scale and diverse data characteristics in big data. As data owners have no absolute control over their data, an effective method of controlling big data access is required to prevent unauthorised access and maintain the confidentiality of the data. In addition to this, the

separation of data subjects and data users requires secure data storage and transaction logs [12]-[13]. The uncertainty on a specific location of the data due to massive transactions of the data will generate data security risks and be subject to the hacker's target. The issue of ineffective scalable privacy-preserving mechanisms [14]-[15] is also a big challenge, particularly during data mining and data analysis, where the analysed data could easily be re-identified and exploited by malicious data users.

The organisational context can be referred to as 'organisational security practice and culture, security planning, and risk mitigation strategies' [15]. Addressing organisational culture [1, 15] is very important to shape and define an organisation's security practices. A lack of leadership [1] to drive the organisation and promote a security culture will dampen the process of big data adoption among employees and deter security technology resources from functioning effectively. The skill shortage is also cited in [1],[11],[15] as another organisational-related issue. [11] reveal a direct correlation between employees' skills and the adoption of big data in an organisation. The lack of a correct security mindset to execute and operate big data is often attributed to a lack of training and exposure.

In the environmental context, the most widely cited issues are the lack of relevant laws and regulations [9],[13],[16]. The responsibility for ensuring the mitigation of security and privacy risks relating to big data require international collaboration across governments and international organisations. Currently, there are at least two main regulatory frameworks that apply to the security and privacy risk of big data, including data protection regulations and the security of essential services [17]. Due to geopolitical uncertainty, many countries have imposed laws and regulations that tighten cross-border data flow and telecommunication service equipment. Legal inconsistency [10] between countries may hinder big data adoption on a global scale across nations. Such decisions are driven by data localisation requirements, enforcing how data can be collected, processed, and stored within a country. For instance, the EU General Data Protection Regulation (GDPR) places conditions on permitting EU residents' data to be transferred only when an adequate protection level is met. Under GDPR, geolocation data that is usually collected and stored by the telecommunication service provider is also protected. With the increasing number of cyberattacks against critical infrastructures in telecommunication services, a common security framework is required to prevent and minimise the impacts of cyberattacks on telecommunication service providers' interconnected infrastructure. The government must define a minimum obligation required from telecommunication service operators and digital service providers to share information on cyberattacks.

The presented issues are categorised into three main contexts, namely technological, organisational, and environmental, to facilitate the extraction of important concepts that emerged from the reviewed works. Figure 1 illustrates the classification of related concepts and emerging themes found from the literature on security and privacy challenges of big data adoption.

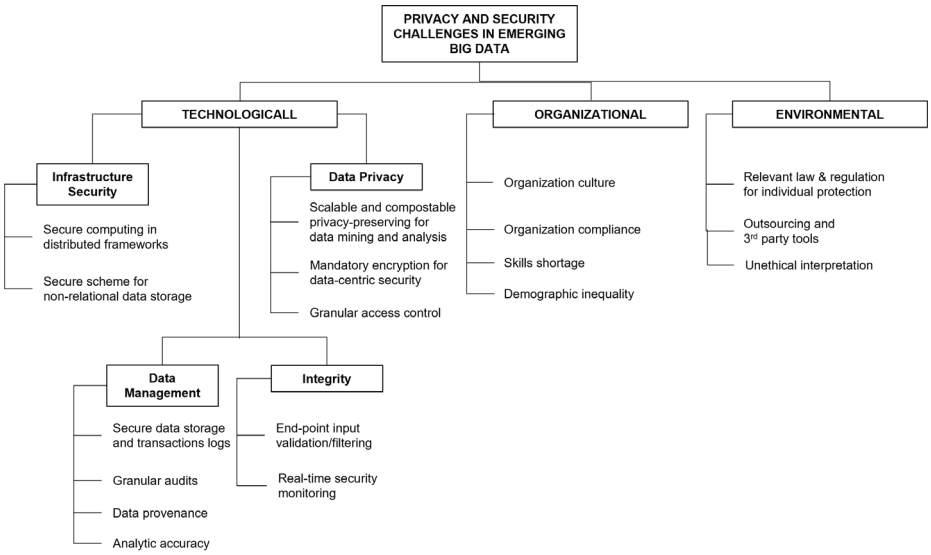

Fig. 1. Emerging themes from literature review on privacy and security challenges in big data adoption

3 Research method

3.1 Instrument design

The focus group instrument has been designed according to the technological, organisational, and environmental (TOE) framework. A focus group protocol and questions have been developed based on the research questions. Focus group protocol is extremely important in this study to serve as a general guideline and procedure for the researcher to conduct the focus group session. The protocol is vital in keeping the researcher focused on the targeted topic in the focus group. Part of the focus group topics that address each research question are shown in Table 1.

Table 1. Focus group topics

Research Question	Topics
What are the perceived security and privacy risks by the telecommunication provider for big data adoption?	• General understanding of security and privacy challenges • Specific technological challenges that affect big data adoption • Specific organisational factors that affect secure big data adoption • Specific security culture factors that affect big data adoption • Specific environmental factors that affect big data adoption • Specific geopolitical uncertainty factors that affect big data adoption globally
What are the specific mitigation strategies to address the security and privacy challenges?	• Familiarity with specific mitigation strategies for technological, organisational, and environmental challenges • Specific industry and/or internal standards that are being applied for the mitigation strategies

3.2 Sampling

The requirements for participant recruitment are outlined in Table 2. Field expert screening is done according to age (35-45 years old), occupation (upper management and senior positions), and experience in cybersecurity and big data projects in Malaysia's telecommunication operators. Convenient sampling is done by sending an invitation to the representative of the telecommunication providers.

Table 2. Requirements for participant recruitment

Requirements	Details
General requirements	• Minimum 5 years experience in cybersecurity
Specific requirements	• Minimum 3 years experience in the telecommunication industry • Minimum 2 years experience in a managerial post • Minimum 1-year experience in big data/cloud computing project

The total number of focus group participants is 8. The sample size of the focus group is sufficient following the recommendation by [30] where the recommended focus group size is between 6 to 12 participants. Following their recommendation, the focus group session is conducted in two rounds, where the first round consists of five participants, and the second round consists of three participants. The data analysis is performed after each round to reach data saturation. Data saturation has been achieved after the second round when no new themes emerged from the collected data. In order to establish validity in the study, focus group interview transcript analysis is done by two coders to ensure data saturation has been reached.

3.3 Data collection

The participants will be informed prior to their involvement about the focus group arrangements; it was communicated to and agreed with them that the sessions would be video-recorded and transcribed, but the data would be anonymised and used only for research purposes. All focus group sessions will be video-recorded and then, transcribed verbatim. The focus group session will be conducted in English. The data is analysed using NVivo 1.6 software. The focus group session is conducted in two modes: a virtual mode and an email mode. The virtual session is approximately 1 hour and 30 minutes in length and is recorded for analysis purposes. The email mode is the follow-up question from the focus group session.

Ethics approval is obtained from the research ethics committee at Universiti Teknikal Malaysia Melaka. Although the chances are very small, there is a risk that someone could get access to the data being stored. The risk may include reputational harm, losing customers, fears of misuse of the information, and strong emotional relatedness to the organisational data. The researcher protects personal identifiable information that would allow any participant to be identified. Their participation in this study is voluntary and participants may withdraw from the study even after signing the consent form, or up until one week after the focus group session.

4 Data analysis

4.1 Data reduction

In this study, the first cycle coding used a mix of in vivo, process, and descriptive coding approaches. To ease the coding process, a deductive coding method is applied, whereby a 'start list' of codes was first developed based on the emerging themes and concepts that were initially derived during the literature review. The start list codes represent three different challenge contexts from the lens of the TOE framework, namely technological, organisational, and environmental challenges. The start list was then applied to the first focus group data. Figure 2 shows how the start list is applied to the focus group session excerpts. For example, we conceptualised the excerpt "… data can provide us true value?" to derive useful information from data in the first cycle coding. The first cycle code is checked against the start list and is grouped as data usability during patterned coding. If a new code emerges, it will be added to the start list.

Interview excerpts	1st Cycle Coding	Patterned Coding
..data collection, as you know, because big data itself requires a lot of data from... from many different sources. It collects whatever and everything that you do, all the data. So we are concerned in terms of what is the data that we are actually collecting whether it's the lawful basis, can we actually collect the data? Have we informed our data subjects or our customers? Do we have a legitimate interest to actually collect this data? Is it too much data that we are collecting, basically, and whether the data can provide us true value?	Many different source Legal requirements of data collection Collecting too much data Derive useful information from data	Multiple data source Regulatory Compliance Data over collection Data usability

Fig. 2. Example of start list development from the focus group data

In the pattern coding stage, the large number of coding in the start list was revised again into a smaller analytical unit to see whether it possessed a structural unity. All first cycle codes were transferred into nodes in NVIVO 1.6 to generate the Pattern codes. A series of related codes were clustered together into a smaller number of themes. For codes that did not belong to the construct in a preliminary conceptual framework, a general pattern code was used, which was helpful in creating sub-codes for easy retrieval. Next, the developed patterned codes were applied again to the data source in NVIVO 1.6. The validity of the developed code is established by having two researchers code the focus group data independently.

4.2 Data display

To present and organise the data display, a group of prominent themes emerged from each of the sub-research questions. In this study, data display was done through the lens of the TOE framework. Owing to different topics in the sub-research questions, the following discussions will show the example of how data display is conducted in the context of technological challenges in big data adoption. In the technological context, initially, 22 themes emerged from the data source as shown in Figure 3.

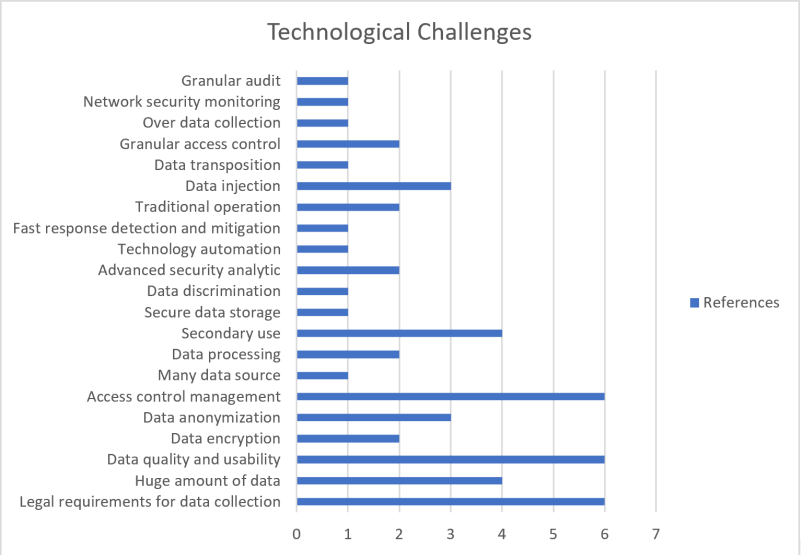

Fig. 3. Initial themes emerged from the focus group data

4.3 Conclusion drawing

After the process of data reduction and data display, 14 themes were found to address the main research questions: What are the perceived security and privacy risks and mitigation strategies by the telecommunication provider for big data adoption? The themes are categorised under the following context:

1. Technological challenge (4 themes)
2. Organisational challenge (2 themes)
3. Environmental challenge (3 themes)
4. Mitigation strategies (5 themes)

The challenge and mitigation strategies for big data adoption in the telecommunication industry are represented in our revised TOE framework as explained in Section 5. From the findings, it can be seen that Data Management is seen as the highest concern under technological challenges, followed by Data Privacy, Regulatory Compliance, Integrity, and Reactive Security, respectively.

5 Results and discussion

In this study, the researchers have concurred that the three TOE challenges do influence big data adoption, and the findings show that there are distinctive challenges pertaining to the telecommunication industry in Malaysia. 14 salient themes emerged from the collective opinions of the participants in the focus group session. Among the 14 challenges identified, four of them concurred with the original TOE framework [18].

These differences can be explained due to the national and industry type influences on the adoption [19].

On the other hand, when we compare the emerging themes found in this study with the literature review, our findings slightly differ from the initial security and privacy challenges categorisation done in the literature review for the technological, organisational, and environmental (TOE) contexts. New themes have been added to reflect the insight gained from the analysis of the focus group session. Regulatory compliance categories have been added to the technological challenge, while infrastructure security themes have been removed from the category. Data governance has emerged as a new theme in the organisational context; while under the environmental context, competitor, and market structure themes have been included.

There are some themes from our initial study in the literature that has been removed from the final findings. Although big data requires organisation-wide adoption, the findings show that very few themes were extracted under organisational challenges. On the other hand, considering the uncertainty of the political landscape in Malaysia and the global COVID-19 pandemic situation, it is expected that geopolitical factors may have a significant effect on big data adoption in the telecommunication industry. Surprisingly, this study found that the telecommunication industry does not regard geopolitical factors as a challenge. One participant explained that "... even if any changes in the political baseline occur, our national strategy on the cybersecurity the principle remains the same".

In the following subsection, all main themes that represent security and privacy challenges and their mitigation strategies for big data adoption in the telecommunication industry will be discussed.

5.1 Technological challenges

Theme 1- data management. The theme covers big data management which includes data collection and data processing. The telecommunication industry's concerns on data collection revolve around the huge volume of collected data from multiple data sources and excessive data collection. Under data processing, the participants mostly agree that data usability is the most significant challenge in big data management. The participants have expressed concerns that the increase of data volume from multiple data sources will significantly degrade the data quality while increasing the risk of data misinterpretation. The availability of usable and quality data will improve the agility and efficiency of the business process. Therefore, the provider needs to ensure that the data filtering is done accurately based on the business case. One of the participants states that "…that's the thing about big data, you don't know what's the end outcome. So, at the end of the day, the challenge is to get the information from the business side whether that information was required in the first place".

Theme 2- data privacy. Under the technological challenge, data privacy [20]-[21] is one of the biggest concerns for telecommunication providers that adopt big data. The the challenge requires the telecommunication provider to explore the techniques and mechanisms to protect the privacy of the data while catering to the business needs and profit. The most prominent dimension under data privacy themes is granular access

control. One of the participants states that "… [challenge for] our big data team is basically to provide the right access to the right people. So that will be ongoing operational processes and approval and so on which need to be enforced along the way". The challenge for securing massive heterogeneous data by applying adequate, cost-effective data anonymisation and encryption techniques are also crucial to ensure sensitive information is not personally identifiable in order to avoid a data breach.

Theme 3- integrity and reactive security. As one of the national critical infrastructure sectors, the high number of cyber threats targeting the telecommunication industry requires the provider to position themselves to be more analytics-driven, rather than following the traditional security operation. With more sophisticated cyberattacks, many participants agree that mitigating them has developed into a challenging task that requires advanced security analytics and security automation to ensure the integrity of the big data environment. One of the participants explained that "…we are moving into more advanced security analytics, right, so that has become so essential in finding the malicious activity in the network, right, to detect a potential threat." Another participant further explained that advanced security analytics and security automation is important for fast detection of threats in real-time and crucial to remediate cyber threats by performing automated incident response.

Theme 4- big data compliance. With the constant updates of the legal compliance requirements due to the growth of data volumes, big data compliance is an ongoing challenge for big data adoption in the telecommunication industry. One participant states that "…when we want to deploy any changes or you know, to the big data, and also collect the information we have to continuously, you know, stay up-to-date and see what the changes from the regulator standpoint, to make sure that we comply with the requirement". The participants highlighted the three aspects of big data management that were subject to regulatory compliance, namely data collection, data injection, and secondary use. From the telecommunication providers' perspectives, the challenge is whether they can provide an adequate level of data security and privacy that meets business needs and profits while adhering to the law, making them especially vulnerable to data breaches [22]. Therefore, the security and privacy of big data must be addressed not just in terms of static regulatory requirements, but also in terms of developing best practices for the industry [23].

5.2 Organisational challenges

Theme 5- data governance. Under the organisational context, data governance is a new theme that emerged in this study. As the telecommunication provider manages a huge volume of customer data from multiple sources, data governance has been cited by the participants as an important organisational challenge. Under data governance, data stewardship is found as the most significant concern because the provider is responsible to provide the right dataset to the right party at the right time, and is compliant with the regulations. One participant said that "…there are challenges, everybody wants something from big data. The challenges are how much time, how much detail that they have to go through until we... so-called validate their claim that they require those type of data, especially the one that actually could be sensitive to the

company". Additionally, this study also found the challenge of data transposition in big data. Our findings corroborate the study by Al-Badi et al. [24] who highlighted that data governance challenge is often ignored by big data operators, and thus the employees face problems in the transition process from existing data sets due to a lack of governance.

Theme 6- subject matter expertise. The focus group data shows that subject matter expertise is a concern within big data operation, but is not deemed a pressing issue. The consensus of the focus group was that the industry has no shortage of talent, and can be developed through the company's talent development programme and up-skilling training. However, they admitted that the market is currently lacking subject matter experts in big data security. Some of the subject areas that are said to be lacking at the moment are data analytics and scripting skills for security automation.

5.3 Environmental challenges

Theme 7- competition intensity and market structure. Competition intensity and market structure are common environmental challenges in any oligopoly market structure, as seen in the telecommunication industry in Malaysia. The growth of the market share between service providers was the result of intense competition in terms of price, telecommunication service quality, and customer service quality [25]. The findings from the focus group data are supported by a study by [26], which states that the sustainability of Malaysia's telecommunication provider is determined by the individual provider's efforts and investments in big data technologies to enable service innovation. Owing to fierce competition among providers, the service providers must offer innovative services at a competitive price that matches customers' requirements and expectations to attract new customers as well as retain existing customers.

Theme 8- regulatory orchestration. Changes and updates in regulatory requirements pose a significant challenge for big data deployment in the telecommunication industry. As regulatory requirement changes, technological compliance needs are becoming more complex, and the challenges for data management are increasing as well. The obtained findings indicate that the telecommunication industry currently suffers from a significant orchestration deficit from all relevant regulators and law authorities. Therefore, the national telecommunication industry requires one party to act as orchestrators of the regulatory systems to improve and better achieve the national goals. The importance of regulatory orchestration is highlighted by one of the participants, "…when we have this type of engagement as one centralised forum, that'd be more I would say, productive and more meaningful when you have the bigger group drive this engage".

Theme 9- technological support. Under technological support themes, vendor support and open-source tools were the most frequent topic extracted from the focus group data. One of the participants explained that "…in terms of operational security, the security hardening of the big data platform itself requires in-depth knowledge. And then we do require the vendors to support us in, in terms of security hardening for the platform itself". The finding is not surprising and corroborates a systematic review by Elgh-

dban et al. [27] who stated that vendor support is the most prominent factor that positively affects the adoption of any advanced IT technologies. In addition, the reliance on third-party open-source software or tools emerges as an ongoing risk to the telecommunication providers if the open-source software development is stopped or discontinued in the future. Adversely, it will slow down the adoption of big data in the industry.

5.4 Mitigation strategies

Theme 10- advanced security tools. The inadequacy of traditional security solutions requires advanced security tools to mitigate more sophisticated threats and attacks. One of the participants explained that "…what we are planning is to also leverage the company's big data to have a bigger scale of analytics for the whole organisation. So, it basically means for the enrichment of security monitoring, right? And so of course that's part of our pipeline...". The telecommunication provider must invest in advanced dynamic security analysis tools to improve big data security. The aim is to detect and analyse security events and related user actions in real-time or near real-time in order to mitigate security and privacy risk as well as prevent illegal attacks.

Theme 11- continuous security assessment. A continuous security assessment is an important mitigating strategy performed by telecommunication service providers in Malaysia to mitigate security and privacy risks. The security assessment is done by evaluating associated security and privacy risk, current security policies, and regulatory compliance. From the focus group data, security assessment is especially important when the provider would like to introduce new initiatives involving customers' personal data such as for new services or new marketing campaigns that may introduce hidden security and privacy risk as well as unforeseen security threats.

Theme 12- security culture promotion. Security culture promotion is seen as an effective strategy to inculcate security awareness among employees. The consensus of the focus group data shows that telecommunication providers in Malaysia have a strong security culture and practise good security behaviour. Security culture is promoted through various awareness programmes, awareness training, and periodical communication through email and message boards. Additionally, support from the top management is important to drive the organisation and advocate for the best security practices. This is explained by of the participants, "…it's important for top management to support the security initiative, so that the bottom, you know, the people will follow because they can always see the how the management you know, shows the importance of the security."

Theme 13- security talent development. In the previous section, we have highlighted that the focus group participants agree that the telecommunication industry has no shortage of talent for big data adoption. Many telecommunications providers offer their talent development programme and offer continuous security training for upskilling. This might be a possible explanation of why the subject matter expertise challenge is not deemed a pressing issue in the industry. One participant commented that "… I won't say it's a bottleneck because if you give them the rightful training and

send them for you know, the relevant exposure that we can build up the competency within one to two years down the road".

Theme 14- strategic plan. Strategic planning is another mitigation strategy to guarantee secure big data adoption is realised. The findings from the focus group data show that a long- and short-term strategic roadmap and the key performance indicator (KPI) for each strategy speed up the big data adoption in the organisation. From the telecommunication provider's perspective, the participants highlighted that the the strategic plan is guided by a responsible business need, which includes the requirements of privacy and security compliance as well as the requirements of new technology. Strategic planning is also an effective mitigation strategy to establish strategic priorities in big data security management to guarantee agility across security systems and solutions [28]- [29].

6 Conclusion and future work

The growing importance and need for big data services in the telecommunication industry have led to a huge amount of sensitive data being generated. Despite the growing importance and need for big data adoption, there is a need for a comprehensive assessment of how security and privacy concerns may affect big data adoption from both data users' and data subjects' perspectives. The main objective of this study is to investigate the perspectives of telecommunication data users in addressing privacy and security issues. Guided by the TOE framework using a focus group qualitative method, this study seeks to address to following research questions: (RQ1) what are the perceived security and privacy risks by the telecommunication provider for big data adoption are, and (RQ2) what are the mitigation strategies to address the security and privacy challenges.

Based on the focus group qualitative analysis, 14 salient themes have emerged where this study found 9 challenges related to RQ1 and 5 mitigation strategies related to RQ2. Under research question 1, it can be concluded that out of the themes that emerged in this study, data management, data privacy, big data compliance, and regulatory orchestration challenges are the most pressing concerns in big data adoption in the telecommunication industry. On the other hand, the telecommunication provider has implemented several strategies to mitigate security and privacy risks, which include the deployment of advanced security tools, performing continuous security assessments, advocating strong security culture, and in-house security talent and development programmes.

This study contributes a thematic classification of security and privacy risks and their mitigation strategies for big data adoption in the telecommunication industry. The thematic classification is important to identify potential gaps for future research and provide general direction in terms of the impact of security and privacy concerns on data users' and data subjects' adoption decisions for big data services in the telecommunications industry. The limitation of the study is that the findings are based on data collected from the participants working in the major telecommunication providers in Malaysia.

Therefore, the findings cannot be generalized and represent the whole telecommunication industry. Additional research and data are certainly needed to validate the findings. The follow-up investigation should concentrate on the specific security and privacy issues raised in our findings. The highlighted security and privacy challenges and their mitigation strategies captured in the qualitative study are interrelated across the TOE context. Therefore, future work must carefully consider challenges or solutions across themes as well as their application to the domain.

7 Acknowledgment

This research is funded by the Malaysian Communications and Multimedia Commission through the 2021 Digital Society Research Grant. Next, we would like to thank the Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka (UTeM), who partially funded the publication of this paper.

8 References

[1] K. Hardy, and A. Maurushat. "Opening up government data for Big Data analysis and public benefit", *Computer law & security review*, vol. 33, no. 1, pp. 30-37, 2017. https://doi.org/10.1016/j.clsr.2016.11.003

[2] H. N. Chua, A. Herbland, S. F. Wong, and Y. Chang. "Compliance to personal data protection principles: A study of how organizations frame privacy policy notices". *Telematics and Informatics*, vol. 34, no. 4, pp. 157-170, 2017. https://doi.org/10.1016/j.tele.2017.01.008

[3] Z. Wang, G. Wei, Y. Zhan, and Y. Sun, "Big data in telecommunication operators: data, platform, and practices". *Journal of Communications and Information Networks*, vol. 2, no. 3, pp. 78-91, 2017. https://doi.org/10.1007/s41650-017-0010-1

[4] A. Narayanan, and V. Shmatikov, "Robust de-anonymization of large sparse datasets". *In 2008 IEEE Symposium on Security and Privacy*, pp. 111-125, 2008. https://doi.org/10.1109/SP.2008.33

[5] U. Gasser, M. Ienca, J. Scheibner, J. Sleigh, and E. Vayena, "Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid", *The Lancet Digital Health*, vol. 2, no. 8, pp. e425-e434, 2020. https://doi.org/10.1016/S2589-7500(20)30137-0

[6] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Hung Byers,. "Big data: The next frontier for innovation, competition, and productivity". McKinsey Global Institute, 2011.

[7] B. Gärtner and R.W. Hiebl,. Issues with big data. The Routledge companion to accounting information systems, pp. 161-172, 2018. https://doi.org/10.4324/9781315647210-13

[8] R. Patgiri, and A. Ahmed, "Big data: The v's of the game changer paradigm", *In 2016 IEEE 18th international conference on high performance computing and communications; IEEE 14th international conference on smart city; IEEE 2nd international conference on data science and systems (HPCC/SmartCity/DSS)*, pp. 17-24, 2016. https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0014

[9] A. M. Leonardo, Z. Albin, B. Smeets, M. H. Sheikh, T. Johansson, and S. Nahid, Privacy, "Security and Trust in Cloud Computing: The Perspective of the Telecommunication Industry", *In The Third International Symposium on Multidisciplinary Emerging Networks and Systems (MENS 2012)*, pp. 627-632, 2012.

[10] H. N. Chua, Y. Chang, S. F. Wong, and C. M. Tan, "Privacy protection policy for big data analytics in the Malaysian telecommunications sector", *In Proceeding of the 26th European Regional Conference of the International Telecommunications Society (ITS): "What Next for European Telecommunications?"*, pp. 1-14, 2015.

[11] I. Malaka and I. Brown, "Challenges to the organisational adoption of big data analytics: A case study in the South African telecommunications industry", *In Proceedings of the 2015 annual research conference on South African institute of computer scientists and information technologists*, pp. 1-9, 2015. https://doi.org/10.1145/2815782.2815793

[12] A. D. Mishra and Y. B. Singh, "Big data analytics for security and privacy challenges", *In 2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 50-53, 2016. https://doi.org/10.1109/CCAA.2016.7813688

[13] W. Fang, X. Z. Wen, Y. Zheng and M. Zhou, "A survey of big data security and privacy preserving", *IETE Technical Review*, vol. 34, no. 5, pp. 544-560, 2017. https://doi.org/10.1080/02564602.2016.1215269

[14] A. Cuzzocrea, "Privacy and security of big data: current challenges and future research perspectives", *In Proceedings of the first international workshop on privacy and security of big data*, pp. 45-47, 2014. https://doi.org/10.1145/2663715.2669614

[15] K. A. Salleh and L. Janczewski, "Technological, organizational and environmental security and privacy issues of big data: A literature review", *Procedia computer science*, vol. 100, pp. 19-28, 2016. https://doi.org/10.1016/j.procs.2016.09.119

[16] C. A. Ardagna, P. Ceravolo and E. Damiani, "Big data analytics as-a-service: Issues and challenges", *In 2016 IEEE international conference on big data (big data)*, pp. 3638-3644 2016. https://doi.org/10.1109/BigData.2016.7841029

[17] L. Urquhart and D. McAuley, "Avoiding the internet of insecure industrial things", *Computer law & security review*, vol. 34, no. 3, pp. 450-466, 2018. https://doi.org/10.1016/j.clsr.2017.12.004

[18] L. G. Tornatzky, M. Fleischer and A. K. Chakrabarti, "Processes of technological innovation". Lexington books, 1990.

[19] J. Baker, "The technology–organization–environment framework", *Information systems theory*, pp. 231-245, 2012. https://doi.org/10.1007/978-1-4419-6108-2_12

[20] M. S. Albulayhi and S. E. Khediri, "A Comprehensive Study on Privacy and Security on Social Media", *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 1, pp. 4–21, 2022. https://doi.org/10.3991/ijim.v16i01.27761

[21] R. S. Almogbel and A. A. Alkhalifah, "User Behavior in Social Networks Toward Privacy and Trust: Literature Review", *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 1, pp. 38–51, 2022. https://doi.org/10.3991/ijim.v16i01.27763

[22] P. Mulgund, B. P. Mulgund, R. Sharman and R. Singh, "The implications of the California Consumer Privacy Act (CCPA) on healthcare organizations: Lessons learned from early compliance experiences", *Health Policy and Technology*, vol. 10, no. 3, pp. 100543, 2021. https://doi.org/10.1016/j.hlpt.2021.100543

[23] R. Thorburn, A. Margheri and F. Paci, "Towards an integrated privacy protection framework for IoT: contextualising regulatory requirements with industry best practices", *In Proceeding of Living in the Internet of Things (IoT 2019)*, 2019. https://doi.org/10.1049/cp.2019.0170

[24] A. Al-Badi, A. Tarhini and A. I. Khan, "Exploring big data governance frameworks", *Procedia computer science*, vol. 141, pp. 271-277, 2018. https://doi.org/10.1016/j.procs.2018.10.181

[25] P. Jain and V. Sridhar, "Analysis of competition and market structure of basic telecommunication services in India", *Communications & Strategies*, vol. 52, no. 4, pp. 271-293, 2003.

[26] S. K. Taghizadeh, The Relationship Between Service Innovation Management Practices On Performance Within Telecommunications Industry In Malaysia (Doctoral dissertation, Universiti Sains Malaysia), 2015.

[27] M. G. Elghdban, N. B. Azmy, A. B. Zulkiple and M. A. Al-Sharafi, "Factors Affecting the Adoption of Advanced IT with Specific Emphasis on Building Information Modeling Based on TOE Framework: A Systematic Review", *International Journal of Advanced Science and Technology*, vol. 29, no.4, pp. 3314 – 3333, 2020. https://doi.org/10.1007/978-3-030-47411-9_2

[28] F. Z. Benjelloun and A. A. Lahcen, "Big data security: Challenges, recommendations and solutions", *In Web Services: Concepts, Methodologies, Tools, and Applications*, pp. 25-38, 2019. https://doi.org/10.4018/978-1-5225-7501-6.ch003

[29] S. A. El-Seoud, H. F. El-Sofany, M. A. F. Abdelfattah and R. Mohamed, "Big Data and Cloud Computing: Trends and Challenges", *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 11, no., pp. 34–52, 2017. https://doi.org/10.3991/ijim.v11i2.6561

[30] P. I. Fusch and L. R. Ness, "Are we there yet? Data saturation in qualitative research", The qualitative report, vol. 20, no. 9, pp. 1408, 2015. https://doi.org/10.46743/2160-3715/2015.2281

9 Authors

Syarulnaziah Anawar is currently a senior lecturer at the Fakulti Teknologi Maklumat dan Komunikasi, UTeM. She is a member of the Information Security, Digital Forensic, and Computer Networking (INSFORNET) research group. Her research interests include human-centered computing, information system, health informatics, usable privacy and security, digital ethics and societal impact of IoT (email: syarulnaziah@utem.edu.my).

Nur Fadzilah Othman received a degree in Computer Engineering in 2008 and master's in educational technology in 2011 at Universiti Teknologi Malaysia (UTM). In 2017, she obtained her PhD in Information Security at Universiti Teknikal Malaysia Melaka (UTeM). She started her career as a senior lecturer at the Faculty of Information Technology and Communication, UTeM from March 2018. Her research interests include information security, usable privacy and security and Internet of Things (IoT) (email: fadzilah.othman@utem.edu.my).

Siti Rahayu Selamat is currently a senior lecturer at the Universiti Teknikal Malaysia Melaka, Malaysia. She received her Doctor of Philosophy in Computer Science (Digital Forensics). Her research interests include network forensic, cyber terrorism,

cyber violence extremism, intrusion detection, network security and penetration testing. She is also a member of Information Security, Forensics and Networking (INSFORNET) research group and actively doing research in malware, criminal behavior and cyber violence extremism profiling (email: sitirahayu@utem.edu.my).

Zakiah Ayop holds BSc. in Computer Science (2000) from UTM and MSc in Computer Science (2006) from UPM. Currently, she is a senior lecturer in Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka (UTeM). She is a member of the Information Security, Digital Forensic, and Computer Networking research group. Her research interest is Information System, Internet of Things (IoT) and Digital Ethics (email: zakiah@utem.edu.my).

Norharyati Harum holds her bachelor's in engineering (2003), MSc. in Engineering (2005) and PhD in Engineering (2012) from Keio University, Japan. She is currently a senior lecturer at Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Her interests in research area are Internet of Things (IoT), Smart Applications, Embedded System, Wireless Sensor Network, Next Generation Mobile Communication, Radio Frequency Planning and Signal Processing. She is an accomplished inventor, holding patents to radio access technology, copyrights of products using IoT devices (email: norharyati@utem.edu.my).

Fiza Abdul Rahim is a senior lecturer at Universiti Teknologi Malaysia (UTM). Her research interests in cybersecurity, advanced metering infrastructure, cryptography, digital forensics, information privacy, and machine learning algorithm. She is a research member of Multimedia Content Protection Technology (MProtec), a research associate of the Institute of Informatics and Computing in Energy (IICE), Universiti Tenaga Nasional (UNITEN) and an associate member of Green Safe Cities (GreSafe), Universiti Teknologi Mara (UiTM) (email: fiza.abdulrahim@utm.my).

Article submitted 2022-04-29. Resubmitted 2022-08-22. Final acceptance 2022-09-15. Final version published as submitted by the authors.