Security Breaches in the Healthcare Domain: A Spatiotemporal Analysis

Mohammed Al Kinoon[1(✉)], Marwan Omar[1], Manar Mohaisen[2], and David Mohaisen[1]

[1] University of Central Florida, Orlando, USA
malkinoon@knights.ucf.edu
[2] Northeastern Illinois University, Chicago, USA

Abstract. Over the past several years, data breaches have grown and become more expensive in the healthcare sector. Healthcare organizations are the main target of cybercriminals due to the sensitive and valuable data, such as patient demographics, SSNs, and personal treatment records. Data breaches are costly to breached organizations and affected individuals; hospitals can suffer substantial damage after the breach, while losing customer trust. Attackers often use breached data maliciously, e.g., demanding ransom or selling patient's information on the dark web. To this end, this paper investigates data breaches incidents in the healthcare sector, including community, federal, and non-federal hospitals. Our analysis focuses on the reasoning and vulnerabilities that lead to data breaches, including the compromised information assets, geographical distribution of incidents, size of healthcare providers, the timeline discovery of incidents, and the discovery tools for external and internal incidents. We use correlation to examine the impact of several dimensions on data breaches. Among other interesting findings, our in-depth analysis and measurements revealed that the average number of data breaches in the United States is significantly higher than in the rest of the world, and the size of the health provider, accounting for factors such as the population and number of adults in a region, highly influences the level of exposure to data breaches in each state.

Keywords: Healthcare data breaches · Confidentiality · Data security

1 Introduction

Electronic health records (EHR) can be described as "a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting. Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data, and radiology reports" [15]. The adoption of EHR improves the healthcare industry and patients alike, and the transformation of healthcare organizations from paper-based to digital has increased healthcare quality by improving patient care and participation, care coordination, diagnostics and patient outcomes, and practice efficiency. However, despite the numerous benefits of EHR, this transformation has led to numerous privacy and security issues which may arise from vulnerabilities (e.g. software vulnerabilities, insider threats, human error, etc.) increasing the possibility of cyber-attacks [11]. The alarming

D. Mohaisen and R. Jin (Eds.): CSoNet 2021, LNCS 13116, pp. 171–183, 2021.
https://doi.org/10.1007/978-3-030-91434-9_16

surge in healthcare data breaches has caused huge concerns in the healthcare sector due to the illegitimate and unauthorized disclosure of private healthcare data [2,20].

Healthcare Data breaches can be classified as either internal or external, and they can occur as a result of theft of private health records, hacking, loss of sensitive patient data, and unauthorized access to patient's private information [27]. External cybersecurity incidents are typically committed by cybercriminals operating in the dark web, while internal data breaches result from something internal to an organization such as disgruntled employees, malicious insiders, employee negligence, and human error. Patient medical records and personal information are often targeted in healthcare data breaches due to their sensitivity and value. External attacks aim to steal those records and demand a ransom or sell those records for hundreds of dollars per single patient on the dark web [22].

Data breaches are devastating and can cause significant damage to healthcare organizations; all the research in this domain demonstrates that the healthcare industry is the most targeted sector due to the attractive financial return of selling sensitive patient records on the dark web [26]. Additionally, the lenient security controls deployed by healthcare organizations further complicate matters and make the healthcare domain a favorite target for hackers. The cost of recovering from such breaches varies greatly by the nature of the incident and number of compromised health records. To better understand the cost aspect, we can break down the cost of data breaches for healthcare entities into two categories: direct costs and indirect costs. Direct expenses include activating incident response teams, engaging forensic experts, outsourcing hotline support, and providing free credit monitoring subscriptions and discounts for future products and services. On the other hand, indirect costs include in-house investigations and communication, as well as the extrapolated value of customer loss resulting from turnover or diminished customer acquisition rates [10]. Given these facts, it's compelling to conduct extensive research studies into the causes, effects, and consequences of healthcare data security incidents. Perhaps more importantly, gaining insights into the different trends and the landscape, and understanding, analyzing, and measuring the statistics in data breaches is crucial for combating such incidents. This is the motivation of this paper and we wish to also motivate the research community in this space to extend the body of knowledge by conducting more studies to be able to better understand data breach and propose solutions in the fight against cybercrimes.

Contributions. To understand the landscape of healthcare data breaches against several attributing characteristics, we provide a detailed measurement-based study of the VERIS (Vocabulary for Event Recording and Incident Sharing) and the Office of Civil Rights (OCR) datasets. To understand attackers' intents and motives, we analyze the type of assets targeted during breaches over various characteristics to investigate their effect. We also analyzed data breaches considering multiple views looking at their distribution, affected entities, breached information, location of the breach, etc.

2 Data Sources

One of the challenges with analyzing cybersecurity incidents, in general, and in the healthcare sector, in particular, is that most datasets are proprietary [25]. Additionally, most breached healthcare organizations shy away from disclosing their vulnerabilities after a breach due to a variety of concerns, including public image, reputation, and patient-trust. The other challenge lies in the fact that each victim healthcare entity tends to take a different approach in analyzing and documenting a data breach [26]. This, in turn, complicates research efforts because data breach statistics are not stored in a central online repository and thus inaccessible to the broader research community. To address the above challenges and conduct our measurements and analysis of data breaches, we turn to the largest publicly available datasets of cybersecurity incidents, namely, the VERIS dataset, and the OCR dataset, which we describe below.

VERIS. We obtained a reliable data source to conduct our research, namely, the Vocabulary for Event Recording and Incident Sharing (VERIS). Veris provides a common language for reporting data breaches incidents in an organized and repeatable manner [13]. Thus, Veris plays a significant role in providing a solution to one of the most critical and persistent challenges in the security industry; lack of quality information. Veris contributes to the solution of this problem by helping organizations collect helpful incident-related details and share them anonymously and responsibly with others. Veris's primary goal is to lay a foundation to constructively and cooperatively learn from our experiences to ensure the proper measurements and managing risk [3].

Office of Civil Rights (OCR). Our second dataset is obtained from the U.S. Department of Health and Human Services Office of Civil Rights. The U.S. Department of Health and Human Services (HHS) Office for Civil Rights (OCR) enforces federal civil rights laws, conscience and religious freedom laws, the Health Insurance Portability and Accountability Act (HIPAA) Privacy, Security, Breach Notification Rules, and the Patient Safety Act and Rule, which together protect your fundamental rights of nondiscrimination, conscience, religious freedom, and health information privacy [16]. The OCR has its breach portal, where data breaches are reported. The website contains data breaches that are currently under investigation within the last 24 months by the OCR. There is also an archived dataset, where resolved data breaches and/or those older than 24 months are archived. All the data breaches reported by the OCR are in the U.S. only. Additionally, all records in the subsequent data breaches affect 500 or more individuals as minor data breaches that affect less than 500 individuals are not reported by the OCR.

3 Studied Dimensions and Variables

This study aims to examine healthcare data breaches considering different aspects of threat characterization and modeling.

- **Geographical mapping:** Section 4.1 analyzes the geographical mapping and distribution of incidents around the world. Analyzing the geographical mapping of the incidents is necessary for several purposes: (i) it provides us with an understanding of the areas most targeted by adversaries for an affinity characterization, (ii) identifying locations around the world where the number of incidents varies due to valuable medical information, particular age group, banking details, etc. We can use this analysis for correlation and prediction capabilities.
- **State-level distribution:** Section 4.2 measures the state distribution of incidents in the U.S. This analysis is necessary for (i) identifying the hot spots targeted by attackers and (ii) conducting correlation analysis between states.

- **Compromised assets**: Section 4.3 details the targeted assets by breaches such as media, server, terminal, etc. Alongside, we will categorize the assets into groups, then dive into their varieties by an individual group against the number of incidents.
- **State-level correlation:** Section 4.4 carries a correlation analysis of the number of incidents within the top ten states with characteristics such as population, Gross Domestic Product (GDP), number of adults, etc. This correlation provides us with essential insights into the reasoning and bearings for each state.
- **Healthcare provider size:** Section 4.5 analyzes the number of breaches versus the size of organizations in terms of the number of employees. We intend to discover if the number of employees influences the frequency of data breach incidents.
- **Timeline discovery:** Section 4.6 examines the response time for incidents affecting healthcare organizations. We will measure the amount of taken time until the discovery of incidents. This analysis helps us determine the organization's security level, and whether more extended discoveries cause more damage.
- **Discovery methods:** Section 4.7 aims to identify the discovery mechanisms used by healthcare entities. Then, we will measure the reported tools and their use in data breaches in our dataset. This analysis can help with determining the appropriate tools needed to be implemented in organizations
- **Adversary demography—The threat intent:** Section 4.8 measures the intention of attackers during data breaches. We intend to acknowledge whether the incidents are targeted or opportunistic.

4 Measurement Results and Discussions

4.1 The Global Distribution of Incidents

Mapping incidents is explicitly provided in our dataset. The dataset uses the ISO 3,166 country codes for each country variable [7], where the codes are generated based on the physical location of the hospital targeted by the attack. Based upon this analysis, we discovered that 1,955 incidents out of the total incidents (2,407) had taken place in the United States, representing 81% of the total incidents. The United Kingdom comes in second, with 157 incidents, representing 7%, and Canada comes in third with 152 incidents, representing only (6%). Figure 1 presents the results for the remaining highest ten countries, while the rest of the world represents (2%) comprising 58 incidents.

As a result of the geographical mapping analysis, we decided to conduct our in-depth analysis study on the United States since most incidents occurred in this country. Several reasons explain why the majority of the incidents are in the United States. First, the Health Insurance Portability and Accountability Act (HIPPA) requires healthcare entities to notify the Department of Health and Human Services (DHHS) whenever a data breach occurs. Second, covered entities must notify affected individuals following the discovery of a breach of unsecured protected health information [16]. In addition to that, covered entities must notify the Secretary of breaches of unsecured protected health information if the affected individuals are 500 or more [16]. Third, covered entities that experience a breach affecting more than 500 residents of a State or jurisdiction are, in addition to notifying the affected individuals, required to provide notice to prominent media outlets serving the State or jurisdiction [16]. Moreover, breach notification is

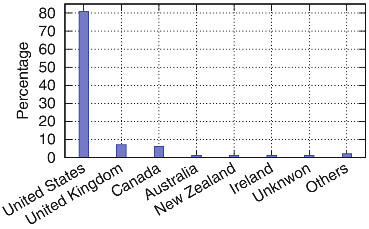


Fig. 1. Incidents by country.

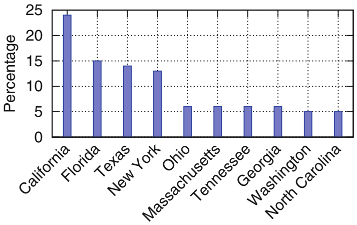


Fig. 2. Incidents by state.

also required for vendors and third-party service providers under the Health Information Technology for Economic and Clinical Health Act (HITECH) [14]. Finally, the HIPPA Security Rule requires healthcare organizations to create a risk management plan protecting all personal health data against security incidents (Office of Civil Rights 2015), which may explain the significant number of reported incidents in the United States [1].

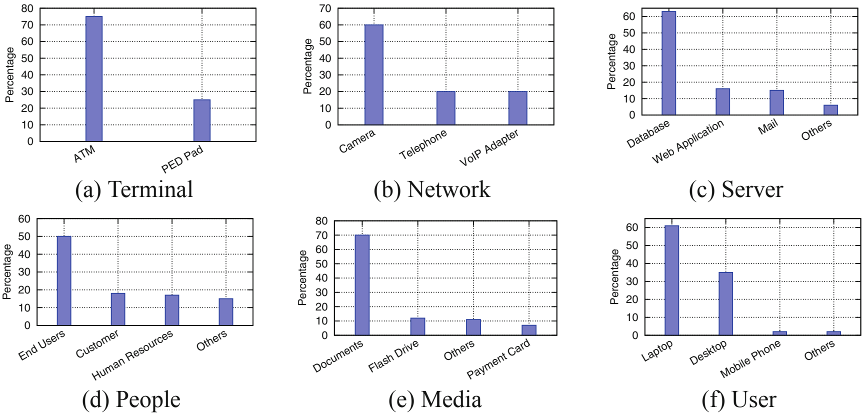
4.2 Number of Incidents by State

Following the global distribution of incidents, we moved into the mapping of incidents on the state level. We analyzed the number of incidents by state. As a result of this analysis, we noticed that California is the highest state with the number of incidents comprising 241 incidents, representing 24% of the overall. Florida comes in second with 147 incidents, representing 15%, and Texas with 145 incidents, representing 14%. Figure 2 shows the remaining results of this analysis.

4.3 Analyzing the Compromised Assets

This section investigates the compromised information assets in the Veris dataset. We harnessed the power of Natural Language Processing (NLP) models to help with analyzing the data gathered from breaches. Information assets fall into six main groups: media, server, terminal, network, user, and people. Each group comprises different varieties [18]. First, the network group includes access control readers such as badge and biometrics, camera or surveillance system, firewall, intrusion detection system (IDS) or intrusion prevention systems, and others. Second, the media group comprises disk media such as CDs or DVDs, flash drives or cards, hard disk drives, identity smart cards, and others. Third, the people group includes administrator, auditor, cashier, customer, former employee, guard, and others. Fourth, the server includes authentication, backup, database, Dynamic Host Configuration Protocol (DHCP), DNS, mail, and others. Fifth, the terminal group includes an automated Teller Machine (ATM), detached PIN pad or card reader, gas "pay-at-the-pump" terminal, self-service kiosk, and others. Finally, the user group includes an authentication token or device, desktop or laptop, media player or recorder, mobile phone or smartphones, and many others.

The existence of assets depends on several reasons and conditions during each incident. We will measure each asset group based on their occurrences in the incidents, and then, we get into the measurement of their varieties to look into the most targeted type of each asset group. This analysis is essential, and its primary purpose is to adequately



(a) Terminal (b) Network (c) Server

(d) People (e) Media (f) User

Fig. 3. Information asset groups and their varieties.

describe the incidents, assess control weaknesses and vulnerabilities, determine impact, and identify mitigation strategies.

Usually, during a data breach incident, one or more assets get compromised by hackers [9]. A compromised asset refers to any loss of confidentiality, integrity, availability during or after the incidents. In the following section, we seek to analyze and measure the asset groups and the total incidents for each group; then, we move to their different asset groups. Based on this analysis, we noticed that media assets are the clear leader comprising 564 incidents out of the overall, representing 33.97%, and server comes in second, comprising 560 incidents, representing 33.73%. Table 1 shows the remaining asset categories and their number of incidents.

After measuring the number of incidents for each asset group as a whole, we moved into measuring their varieties. Based on the analysis done, we found that 61% of the incidents in the user group are through laptops, followed by the terminal group with 75% of the incidents through ATMs. In the server asset group, we found out that 63% of the incidents happened through exploiting the database. While for the people asset group, 50% of the incidents are because of the end-user. Most of the incidents that happen in the network are throughout cameras, with represent 60%. Lastly, 70% of the incidents in the media group are through documents. In Fig. 3, we present the remaining results for the other asset groups and their varieties.

Table 1. Assets varieties with the number of incidents during data breaches.

Asset Group Type	# Incidents	Percentage
Media	564	33.97%
Server	560	33.73%
User	493	29.69%
People	34	2.04%
Network	5	0.30%
Terminal	4	0.24%
Overall	1660	100 %

4.4 State Level Correlation

This section will conduct a state-level correlation between the number of reported incidents and hospitals, staffed beds, population, and gross domestic product (GDP) for the top 10 states. GDP is the gross domestic product and is represented in billion U.S. dollars. To address the following question, we conducted a state-level analysis considering these factors related to the reported incidents in our dataset. We decided to run this analysis on the highest 10 states in terms of the number of reported incidents. We started by collecting the specified statistics for each state, including population, GDP, staffed beds, and hospitals. The relationship between two variables can be a positive relationship (1), no relationship (0), and an inverse relationship (−1). Upon this analysis, we discovered that the population and adults are highly correlated with the number of incidents (0.96). Followed by the GDP (0.95). The remaining results of the correlation are shown in Table 2.

4.5 Organizations Size

The following section investigates the size of healthcare entities and how organization's size might contribute to a data breach. Using Veris, we performed the analysis by looking into the scope of healthcare organizations at the time of the incident. We classified healthcare organizations into two main groups: small and large. A small group includes a size of up to 1,000 employees, while a large organization would be over 1,000 employees. Upon this analysis, there were a total of 1,361 incidents divided into two groups. Our analysis revealed that 57% of the incidents are in the small group, while 43% are in large groups.

Table 2. State level correlation. Numbers of incidents (I), hospitals (H), employees (E), staffed beds (B), GDP (G), population (P), and adults (A) are considered.

	I	**H**	**E**	**B**	**G**	**P**	**A**
I	1.00						
H	0.88	1.00					
E	0.92	0.91	1.00				
B	0.94	0.92	0.97	1.00			
G	0.95	0.86	0.92	0.89	1.00		
P	**0.96**	0.95	0.94	0.94	0.95	1.00	
A	**0.96**	0.88	0.94	0.96	0.89	0.90	1.00

4.6 Timeline Discovery

Timeline discovery of data breaches varies depending on the type of industry, geography, and level of security of an organization. According to a recent study conducted by the IBM security team in the healthcare sector, the average time to discover a data breach is 329 days, and 93 days are required to regain control. Unfortunately, prior work fails to provide in-depth analysis on the timeline discovery of the data breaches, including discovery tools for external and internal incidents. To fill this gap, we analyzed the timeline discovery of the reported incidents and went over the tools used for incident discovery for both internal and external discovery methods. This analysis is essential to address the lessons learned during the incidents and remediation process and provide organizations with insights and corrective actions to improve their detection and defensive capabilities. Our analysis found out that organizations fail to identify data breaches early enough, resulting in more damage. From the reported incidents, we discovered that 3% of the incidents took minutes until discovery, 9% took hours, 15% took days, 6% took weeks, 52% took months, and 15% took years. In the coming section, we will address different discovery methods and whether there is a difference between internal attacks and external attacks.

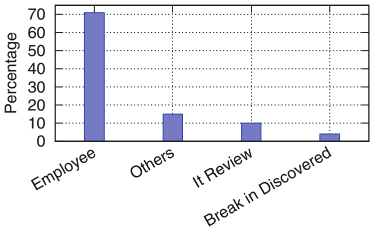


Fig. 4. Internal discovery

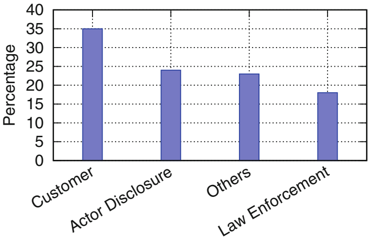


Fig. 5. External discovery

4.7 Internal and External Discovery Methods

Discovery methods fall into two main categories; internal and external. Organizations use several tools to discover an incident depending on the type of data breach. External and internal data breaches are different, and each one of them requires special discovery tools. First, healthcare organizations use numerous tools to discover incidents for internal incidents, such as Host IDS or file integrity monitoring, network IDS, and IPS alerts. In contrast, practices including law enforcement, actor disclosure, and customer notifications can help discover external incidents. Our analysis found out that most of the internal incidents are discovered by employees, representing 71% of the total incidents. In contrast, customers discover 35% of the external incidents, and actor disclosure comes in second, representing 24%. The remaining results of this analysis are shown in Figs. 4 and 5.

4.8 Targeted vs Opportunistic

To understand the nature of the data breach incidents and whether they are intentional or non-intentional, we conducted a measurement analysis to investigate the number of targeted incidents and opportunistic ones. This classification is uniquely relevant to deliberate and malicious actions. There are two main categories: targeted and opportunistic. First, opportunistic incidents occur when the victim exhibits a weakness that the actor has the knowledge to exploit. Second, targeted incidents happen when the adversary chooses the victim as a target, and then the actor will investigate possible vulnerabilities to exploit. Using our exclusively given records in our dataset, we found that more than half of healthcare data breaches are opportunistic, representing 80%, while, on the other hand, 20% are targeted.

5 Analysis of the OCR Dataset

Type of Breach. We analyzed the causes of healthcare data breaches based on the reported incidents and observed that most incidents occur due to hacking or IT-related disclosure comprising 1,069 incidents, representing 31% of the overall incidents. Unauthorized access and disclosure came in second, holding 934 incidents overall, representing 27%. Finally, the theft category came in the third place, comprising 909 incidents, accounting for 26% of the total incidents.

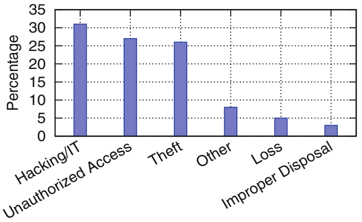


Fig. 6. Type of breach.

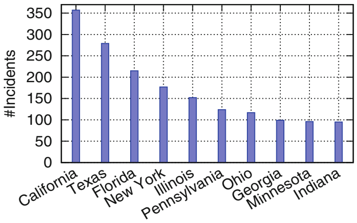


Fig. 7. Distribution by state.

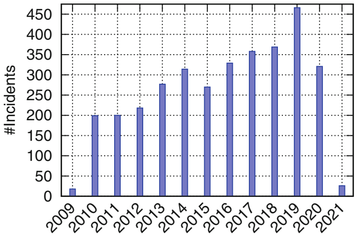


Fig. 8. Incidents by year.

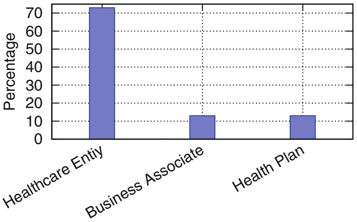


Fig. 9. Covered entities.

State Distribution. The following section addresses the distribution of the incidents for the U.S states. Using the OCR data, we measured the incidents for each state; this analysis is essential for trends and comparison. Following this analysis, we have observed that states with large population, high Gross Domestic Product (GDP), and large adult population are more targeted than others, as shown in Sect. 4.4. California was the most affected, totalling 357 incidents, followed by Texas with 279 incidents, while Florida was the third largest with 215 incidents (Figs. 6 and fig:UsspsStateshhs).

Distribution of Incidents by Year. Using the ORC dataset, and over the period between 2009 and the time of conducting this study in 2021, we measured the reported incidents in the dataset affecting 500 or more victims and reported to the HHS OCR. Following this analysis, we notice that the number of incidents surged over time, indicating a lack in implementing stringent security controls by organizations in the healthcare industry. As shown in Fig. 8, there is a massive increase in the number of incidents in 2019, as it was the year with the highest number of breaches in the whole dataset.

Covered Entity. We analyzed the distribution of incidents by organization type. According to the OCR dataset, there are three main targeted entities. First, healthcare entities that provide health care services and engages in professional review activity through a formal peer review process for the purpose of furthering quality health care, a committee of that entity, a professional society, a committee or agent thereof, including those at the national, state, or local level, physicians, dentists, or other health care practitioners that engage in professional review activity through a formal peer review process to further quality health care [17]. Second, a business associate, which is a person or entity that performs certain functions or activities that involve the use or disclosure of protected health information on behalf of or provides services to a covered entity [24]. Third, health plan, which constitutes individual or group health plans that provide or

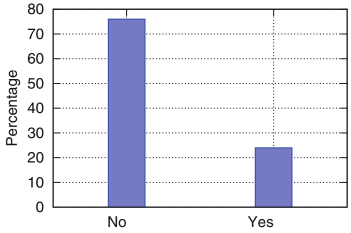


Fig. 10. Business associate.

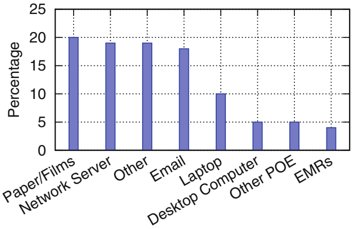


Fig. 11. Information breached.

pay the cost of medical care [5]. Following this analysis, we observed that healthcare entities are most targeted during the incidents, having 2,450 incidents which represents 73% of the total incidents, business associate and healthcare plan came in second and third comprising 451 and 439 incidents, and representing 14% and 13%, respectively. Figure 9 depicts the results of this analysis.

Business Associates. We further analyzed the existence of incidents when a business associate is present or not. According to HIPPA, any covered entities and business associates enter into a contract to ensure the safety of protected healthcare information. A business associate may use or disclose protected health information only as permitted or required by its business associate contract or as required by law [23]. Our analysis revealed that 2,532 incidents had no business associates included, representing 76%, while only 819 incidents had a business associate, representing 24% of the incidents as shown in Fig. 10.

Location of Breached Information. When a data breach occurs, private and confidential patient information gets disclosed due to either unauthorized access or human error. Healthcare system keeps record of valuable information and medical records, containing sensitive personally identifiable information (PII) such as address history, financial information, social security numbers, and patient medical treatment records. This sensitive information is often targeted by hackers due to its outstanding value. Hackers can easily use that data to set up a line of credit or take out a loan under patients' names. Unfortunately, healthcare organizations often lack the stringent security measures (e.g., encryption, robust anti-virus software, multi-factor authentication, etc.) required to secure medical records. To this end, we analyzed the most targeted information to gain insight into the type of medical and personal data prioritized by hackers in healthcare data breaches. We observed that paper/films are the most breached information comprising 662 of the overall incidents, representing 20%. Closely, the network server came in second, comprising 643 incidents, accounting for 19%. The other category came in third, comprising 641 incidents, representing 19% as well. The remaining attributes and results of this analysis are presented in Fig. 11.

6 Related Work

In the past few years, numerous studies have analyzed data breaches in the healthcare sector. Choi *et al.* [4] estimate the relationship between data breaches and hospital advertising expenditures. They concluded that teaching hospitals were associated with significantly higher advertising expenditures two years after the breach. Another study [12] investigated the privacy-protected data collection and access in IoT-based healthcare applications and proposed a new framework called PrivacyProtector to preserve the privacy of patients' data. Another study [6] found that the healthcare industry was being targeted for two main reasons: being a rich source of valuable data and its weak defenses. Other study [28] suggested a framework to examine the accuracy of automatic privacy auditing tools. Siddartha *et al.* [21] suggested that current healthcare security techniques miss data analysis improvements, e.g., data format-preserving, data size preserving, and other factors. Most related to our work, the 2021 Data Breach Investigations Report [8] summarized the findings and determined that external actors are behind 61% of data breaches while 39% of data breaches involved internal actors. According to the same report, personal information is the most compromised, comprising 66% of data breaches. In contrast to our work, authors of [19] conducted a comprehensive analysis of HIPPA data breach reports. They found that the main disclosure types of protected healthcare information were hacking incidents, unauthorized access (internal), theft or loss, and improper disposal of unnecessary data. The authors used the Simple Moving Average (SMA) and Simple Exponential Smoothing (SES) time series methods. They applied them to the data to determine the trend of healthcare data breaches and their cost to the healthcare industry. Our comprehensive study comprises but is not limited to analyzing compromised assets, internal and external discovery methods, discovery timeline of data breaches, distribution of the incident globally and in the united states, and breached information. In addition, we used correlation as a mathematical tool to determine healthcare data breaches and quantify the effects of different factors like GDP, population, number of hospitals, and their sizes in terms of the staffed beds on data breaches.

7 Conclusion

Our study revealed that the number of adults and the state population highly influence the exposure to data breach incidents, with California, Florida, and Texas being the lead targets. We show that the media group was the most breached asset, followed by the Server and User group. Interestingly, we found that the majority of incidents occur in small size organization – 57%. In contrast, 43% of the incidents occur in large organizations, suggesting that large healthcare organizations tend to have better security systems. Our timeline discovery revealed that most of the incidents, approximately 52%, were discovered within months, while 15% of the incidents took years to be discovered. Employees discovered the majority of the incidents for internal incidents. Based on a long-term dataset analysis, most of the incidents, 80%, tend to be opportunistic, while 20% are targeted. In the future, it would be interesting to conduct research harnessing the power of machine learning to enable information sharing on data breaches.

Acknowledgement. This work was supported by NRF-2016K1A1A2912757.

References

1. Adebayo, A.O.: A foundation for breach data analysis. J. Inf. Eng. Appl. **2**(4), 17–23 (2012)
2. Alkinoon, M., Choi, S.J., Mohaisen, D.: Measuring healthcare data breaches. In: Kim, H. (ed.) WISA 2021. LNCS, vol. 13009, pp. 265–277. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-89432-0_22
3. Chernyshev, M., Zeadally, S., Baig, Z.: Healthcare data breaches: implications for digital forensic readiness. J. Med. Syst. **43**(1), 1–12 (2019)
4. Choi, S.J., Johnson, M.E.: Understanding the relationship between data breaches and hospital advertising expenditures. Am. J. Manag. Care **25**(5), e14–e20 (2019)
5. Employers Council: What is the definition of a health plan under HIPAA? (2015). https://bit.ly/3Aherpb
6. Coventry, L., Branley, D.: Cybersecurity in healthcare: a narrative review of trends, threats and ways forward. PubMed, April 2018. https://doi.org/10.1016/j.maturitas.2018.04.008
7. Developers: International organization for standardization: 3,166 country codes (2021). https://bit.ly/3Eoem5J
8. Verizon Enterprise: Verizon data breach investigations report (2021). https://vz.to/3AvCNfn
9. Gwebu, K., Barrows, C.W.: Data breaches in hospitality: is the industry different? J. Hosp. Tour. Technol. (2020)
10. (HC3), Health Sector Cybersecurity Coordination Center: A cost analysis of healthcare sector data breaches (2019). https://bit.ly/3hHpJMj
11. Kamoun, F., Nicho, M.: Human and organizational factors of healthcare data breaches: the swiss cheese model of data breach causation and prevention. Int. J. Healthc. Inf. Syst. Inform. (IJHISI) **9**(1), 42–60 (2014)
12. Luo, E., Bhuiyan, M.Z.A., Wang, G., Rahman, M.A., Wu, J., Atiquzzaman, M.: PrivacyProtector: privacy-protected patient data collection in IoT-based healthcare systems. IEEE Commun. Mag. **56**(2), 163–168 (2018)
13. Makridis, C., Dean, B.: Measuring the economic effects of data breaches on firm outcomes. J. Econ. Soc. Meas. **43**(1–2), 59–83 (2018)
14. McLeod, A., Dolezel, D.: Cyber-analytics: modeling factors associated with healthcare data breaches. Decis. Support Syst. **108**, 57–68 (2018). https://doi.org/10.1016/j.dss.2018.02.007
15. Menachemi, N., Collum, T.H.: Benefits and drawbacks of electronic health record systems (2011). https://bit.ly/3EscQ2k
16. Office for Civil Rights: Breach notification rule (2013). https://bit.ly/3jCpHXI
17. Rank, N.P.D.: NPDB guide book (2021). https://bit.ly/2XwgTdw
18. Sarabi, A., Naghizadeh, P., Liu, Y., Liu, M.: Risky business: fine-grained data breach prediction using business profiles. J. Cybersecur. **2**(1), 15–28 (2016)
19. Seh, A.H., et al.: Healthcare data breaches: insights and implications. Healthcare **8**, 133 (2020). https://doi.org/10.3390/healthcare8020133
20. Seh, A.H., et al.: Healthcare data breaches: insights and implications. In: Healthcare. vol. 8, p. 133. Multidisciplinary Digital Publishing Institute (2020)
21. Siddartha, B.K., Ravikumar, G.K.: Analysis of masking techniques to find out security and other efficiency issues in healthcare domain. In: Third International conference on I-SMAC, pp. 660–666 (2019). https://doi.org/10.1109/I-SMAC47947.2019.9032431
22. Smith, T.: Examining data privacy breaches in healthcare. Ph.D. thesis, Walden U. (2016)
23. U.S. HHS: Business associate contracts (2013). https://bit.ly/3ChsJH9
24. U.S. HHS: Business associates (2019). https://bit.ly/3tM4PQV
25. Walker-Roberts, S., Hammoudeh, M., Aldabbas, O., Aydin, M., Dehghantanha, A.: Threats on the horizon: understanding security threats in the era of cyber-physical systems. J. Supercomput. **76**(4), 2643–2664 (2020). https://doi.org/10.1007/s11227-019-03028-9

26. Walker-Roberts, S., Hammoudeh, M., Dehghantanha, A.: A systematic review of the availability and efficacy of countermeasures to internal threats in healthcare critical infrastructure. IEEE Access **6**, 25167–25177 (2018)
27. Wikina, S.B.: What caused the breach? An examination of use of information technology and health data breaches. Perspect. Health Inf. Manag. **11**(Fall), 1–16 (2014)
28. Yesmin, T., Carter, M.W.: Evaluation framework for automatic privacy auditing tools for hospital data breach detections: a case study. Int. J. Med. Inform. **138**, 104123 (2020)