Measuring Healthcare Data Breaches

Mohammed Alkinoon, Sung J. Choi, and David Mohaisen(✉)

University of Central Florida, Orlando, USA
mohaisen@ucf.edu

Abstract. Over the past few years, healthcare data breaches have grown rapidly. Moreover, throughout the COVID-19 pandemic, the level of exposure to security threats increased as the frequency of patient visits to hospitals has also increased. During the COVID-19 crisis, circumstances and constraints such as curfew imposed on the public have resulted in a noticeable increase in Internet usage for healthcare services, employing intelligent devices such as smartphones. The Healthcare sector is being targeted by criminals internally and externally; healthcare data breaches impact hospitals and patients alike. To examine issues and discover insights, a comprehensive study of health data breaches is necessary. To this end, this paper investigates healthcare data breach incidents by conducting measurements and analysis recognizing different viewpoints, including temporal analysis, attack discovery, security attributes of the breached data, attack actors, and threat actions. Based on the analysis, we found the number of attacks is decreasing, although not precluding an increasing severity, the time of attack discovery is long across all targets, breached data does not employ basic security functions, threat actions are attributed to various vectors, e.g., malware, hacking, and misuse, and could be caused by internal actors. Our study provides a cautionary tale of medical security in light of confirmed incidents through measurements.

Keywords: Healthcare data breaches · Data confidentiality · Data security · Data analysis

1 Introduction

The United States Department of Health and Human Services defines a data breach as an intentional or non-intentional use or disclosure of confidential health information. A data breach compromises privacy and security, resulting in a sufficient risk of reputation, financial, and other harm to the affected individuals [16]. Over the past few years, concerns related to healthcare data privacy have been mounting, since healthcare information has become more digitized, distributed, and mobile [7]. The medical records have transformed from paper-based into Electronic Health Records (EHR) to facilitate various digital system possesses. Medical EHR can be described as "a longitudinal electronic record of patient health information generated by one or more encounters in any care

H. Kim (Ed.): WISA 2021, LNCS 13009, pp. 265–277, 2021.
https://doi.org/10.1007/978-3-030-89432-0_22

delivery setting. Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data, and radiology reports" [10]. EHR enhances patient care by enhancing diagnostics and patient outcomes, improving patient participation, enhancing care coordination, practicing efficiencies, and cost savings [5]. Despite the numerous benefits of EHR, the transformation has inflated the security and privacy concerns regarding patients' information. The growing usage of Internet of Things (IoT) and intelligent devices affects the methods of communication in hospitals and helps patients quickly access their medical treatment whenever necessitated.

Nevertheless, the usage of such technologies is a fundamental factor that can cause security risks and lead to data breaches [13]. Broadly, healthcare data breaches are external and internal. External breaches are malicious, including at least one or more threat actions from cyber criminals such as hacking, malware, and social attacks. On the other hand, internal data breaches typically occur due to malfeasance by insiders, human errors, and negligence from employees. Data breaches have increased in the past decade. In comparison with other industries, healthcare is the worst affected [8]. Cybercriminals are targeting healthcare for two fundamental reasons: it is a rich source of valuable data, and its defenses are weak [3]. Medical records contain valuable information such as victims' home addresses and Social Security Numbers (SSNs). Adversaries utilize such information for malicious activities and identity theft, or exchange those medical records for financial profit on the dark web.

Contributions. For a better understanding of the landscape of healthcare data breaches against various attributing characteristics, we provide a detailed measurement-based study of the VERIS (Vocabulary for Event Recording and Incident Sharing) dataset. Among other characteristics, we temporally analyze data breaches and their growth over time. To understand attacks' intent, we analyze the type of breaches over various security attributes and characterize the threat actions, highlighting the attack vector employed for the breach. We hope that those characterizations will shed light on the trend and the attack vectors, thus providing directions for mitigating those breaches.

2 Data Source and Temporal Analysis

The object of this paper is to conduct a measurement of healthcare data breaches to understand trends and motives. To accomplish that, we used a trusted and reliable data called VERIS. In the past, there were numerous initiatives to accumulate and share security incidents. Nonetheless, commitment and participation have been minimal. Reasons behind that are many, including (i) the difficulty of categorization, (ii) the uncertainty of what to measure [15]. To facilitate data collection and sharing, VERIS is established as a nonprofit community designed to accommodate a free source of a common language for describing security incidents in a structured and repeatable way [15]. Due to the prevailing lack

of helpful information, the VERIS dataset is an effective solution to the most critical and persistent challenges in the security industry. VERIS tackles this problem by offering organizations the ability to collect relevant information and share them responsibly and anonymously.

VERIS and Incident Attributes. VERIS's primary purpose is to create an open-source database to design a foundation that constructively and cooperatively learns from their experience to assure a more reliable measurement and managing risk system. VERIS is a central hub whereby information and resources are shared to maximize the benefits of contributing organizations. During the incident collection process, the VERIS community focuses on successfully implementing an intersection, namely the 4A's, which indicate the following: who is behind the incidents (actors), the action used by the adversary (actions), devices affected (assets), and how are they effected (attributes). An example of the 4A's for an incident can be as follows: internal (actor), hacking (action), network (asset), and confidentiality (attribute). VERIS designers estimate the needed information to be collected about an incident based on the level of threat, asset, impact, and control. If organizations understand the complete image of those risk aspects, they can learn to improve their management system to make the correct decisions. The power of VERIS is the collection of evidence during and after the incidents, besides providing helpful metrics to maximize risk management.

2.1 Distribution of the Incident's Timeline

We analyzed the timeline mapping of the incidents across the years. The VERIS dataset contains incidents that took place between the year 1971 until 2020. While a long period of time is considered, the time frame from 1971 until 2010 seemed to contain a low number of incidents, with only 272 (total) of them, per the VERIS dataset. Thus, to understand the actual trend in the active region, we limit ourselves to the year 2010 onward. This analysis is essential because it provides us with insights into the active period of breaches and attacks, and could hint on the underlying ecosystem. To this end, and upon this analysis, we found out the following (1) the per year number of incidents follows a normal distribution, with the peak at 2013. (2) 2013 was the highest year in the number of incidents, with 395 (16%), followed by 2015 with 317 (13%), and 2014 with 310 (≈13%). (3) Contrary to the common belief that the number of attacks is increasing, we found that the number of breaches has been decreasing since 2013, per VERIS reporting, as shown in Fig. 1.

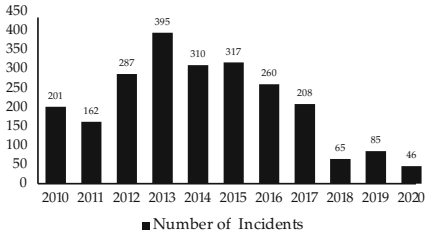


Fig. 1. The distribution of incidents

Takeaway. *There is a decrease in the number of incidents per year, possibly due to the lax reporting. This decrease, however, does not preclude the possibility that each of those breaches is getting more severe than past breaches.*

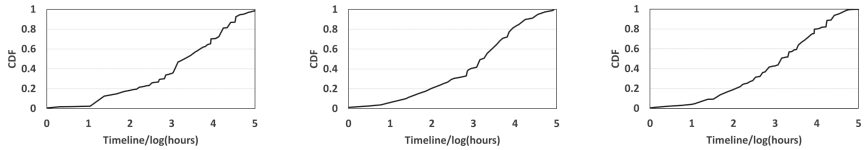
2.2 Timeline Discovery for Data Breaches

Health organizations encounter various difficulties in attempting to keep patients' medical records safe. The *timeline discovery* affects both the patient and the hospital. The longer it takes for an organization to discover a data breach, the more significant harm it can cause. The damage cannot only result in data loss or the disclosure of information but also includes businesses. In the literature, it was shown that organizations take 197 days to identify a data breach and 69 days to contain it, on average [6]. That amount of time to detect a data breach is considered long and costs organizations millions of dollars. An organization containing the data breach incident in less than 30 days from the date it happened can save up to $1 million compared to others who fail to do so, per the same study. In healthcare, hospitals and organizations can suffer many consequences due to a data breach, including lawsuits from the affected individuals, as well as reputation and trust loss. In addition, healthcare organizations incur significant costs fixing the problem and protecting patients from additional harm.

We examined the response time for incidents affecting victims; the patients, customers, and employees. In the following, we present the results and contrast.

Results. We began by converting the timeline discovery into one unit (hours). Then, we calculate the cumulative distribution function (CDF) of incidents. Due to the extensive range of timelines, we used the logarithmic function to the discovery time range for simplicity and visibility, as shown in Fig. 2. Based on this analysis, we noticed that the discovery time of incidents for employees is significantly faster than for customers and patients. As we can see in Fig. 2, we discovered that 20% of the incidents for employees were discovered within four days or less. It took five days or less to discover the same percentage for customers, and up to six days to discover that for patients. Such results indicate the difference between the different categories breach discovery time, and perhaps the priorities associated with their discovery and protection, although all are relatively high. To further establish that, for 50% of the incidents, the discovery time was 2, 2.5, and 3 months for customers, employees, and patients, respectively. The patients represent most victims with 41%, and the discovery time for their data breaches extends to years (14 years to discover 100% of all incidents). While discovering 100% of incidents for customers require a longer time: up to 21 years. On the contrary, the discovery time of incidents for employees is much less because discovering 100% of the incidents for this category is about ten years.

Takeaway. *Incidents discovery, even for most protected victims, can take many years, highlighting the lax security posture of healthcare organizations.*



(a) Patients (b) Employees (c) Customers

Fig. 2. CDF for the timeline discovery of different types of victims.

3 Security Attributes

The VERIS dataset uses pairs of the six primary security: confidentiality/possession, integrity/authenticity, and availability/utility as an extension of the CIA triad.

In this section, we attempt to investigate the compromised security attributes during the incidents by conducting the following: (i) analyzing the confidentiality leakage that occurred during data breaches, (ii) present the different data types, and note which is the most targeted by adversaries, (iii) determine the state of the compromised data at the time of the incidents.

Data Confidentiality. Confidentiality refers to the limit of observations and disclosure of data [15]. We start by examining the data confidentiality leakage that occurred during data breaches. This analysis is necessary because it examines the amount of compromised data and their varieties throughout the incidents. Using the VERIS dataset, we found that 1,045 out of total data 1,937 incidents had *information disclosure*, representing 54% of the total incidents, 882 had a *potential information disclosure*, representing 46%, while only two incidents that had *no information disclosure* at all and eight incidents are *unknown*.

We analyzed data that attackers often target. Based upon this analysis, we discovered the following: medical information exposed to higher disclosure compared to the other types of information, encompassing 1,413 incidents, representing 73%, while personal information appeared in second place, with 345 incidents, representing 18%. Lastly, payment information appeared in third place, having 61 incidents, representing 3%. Other targeted information include *unknown* (44; 2%), *banking* (33; 2%), *credentials* (23; 1%), and *others* (18; 1%).

Takeaway. *Despite their variety in breaches, medical and personal information are the most targeted, with 91% of the incidents combined.*

Status of Breached Data. During the exposure or compromise process, we investigated the state of the data and whether it was encrypted, transmitted, or stored unencrypted during the attack. This categorization aims to understand the security controls while the data is at rest or in motion due to transformation. As a result of this investigation, we noticed 36% of the data was *stored unencrypted*, 30% *stored*, 25% *unknown*, 3% *printed*, 2% *transmitted unencrypted*, and 4% with other attributes.

Takeaway. *The majority of breached data does not employ basic security functions, making it an easy target to adversaries for exploitation at rest or in transit.*

Data Integrity and Authenticity. Integrity refers to an asset or data to be complete and unchanged from the original state, content, and function [15]. Example of loss to integrity includes but is not limited to unauthorized insertion, modification, and manipulation. We wanted to discover the varied nature of integrity loss. Each time incidents occur, there can be at least one integrity attack. However, many losses can be associated with a single incident. Following the analysis, we noticed that most data integrity losses are due to altering behaviors containing 93 incidents, representing 31% of the overall. Software installation comes in second with 91 incidents, representing 30% of the known reasons. Other integrity related attacks include *fraudulent transmission* (18%), *data modification* (11%), *re-purposing* (3%), and *others* (6%).

Authenticity refers to the validity, conformance, correspondence to intent, and genuineness of the asset (or data). Losses of authenticity include misrepresentation, repudiation, misappropriation, and others. Short definition: Valid, genuine, and conforms to intent [15]. Based upon this analysis, we observed that the authenticity state was poorly reported at the time of the incidents.

Data Availability. Availability refers to an asset or data being present, accessible, and ready for use when needed [15]. A loss to availability includes deletion, destruction, and performance impacts such as delay or acceleration. We will show the varieties of the data available that might happen during the incidents. This analysis is necessary to understand the nature or type of availability or utility loss. Based on this analysis, we found that 769 incidents contained a loss of data regarding their effect on availability, representing 90% of the total incidents with the reported attribute. *Obfuscation*, and *interruption* are reported as remaining causes affecting availability, with 9% and 1% of all incidents, respectively.

Takeaway. *Despite limited reporting, more than 20% of all the studied incidents suffer from integrity and authenticity attacks, due to a range of factors, magnifying the potential of attacks without data leaving the organization.*

4 Analyzing the Threat Actors

Threat actors are entities that can cause or contribute to an incident [15]. Each time an incident happens, there can be at least one of the three threat actors involved, but on some occasions, there can be more than one actor involved in a particular incident. Threat actor's actions can be malicious, or non-malicious, intentional or unintentional, causal or contributory [15]. VERIS classifies threat actors into three main categories, namely: external, internal, and partner. This

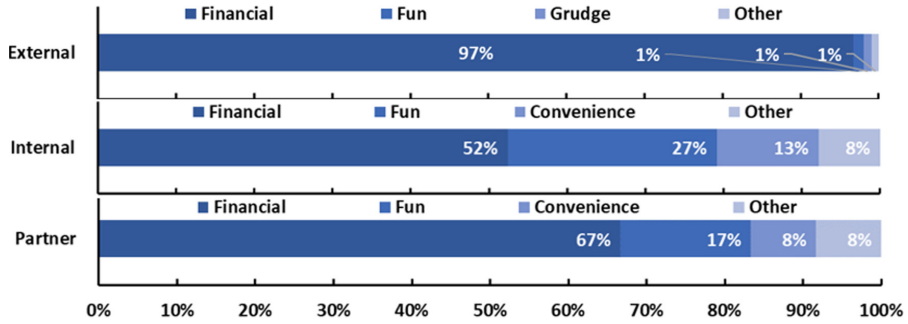


Fig. 3. Threat actors motives for external, internal, and partner actors.

classification excludes the contributory error that unintentionally occurs. For instance, if an insider unintentionally misconfigures an application and left it vulnerable to an attack. The insider would be not be considered as a threat actor if the applications were successfully breached by another actor [15].

On the other hand, an insider who deliberately steals data or whose inappropriate behavior (e.g., policy violations) facilitated the breach would be considered a threat actor in the breach [15]. This section will explain and analyze each category of the threat actors with their presence in incidents from our dataset. This analysis is essential because of the following reasons: (i) it provides us with an understanding of the reasons or motives that can leads actors to act, (ii) the analysis can provide knowledge for organizations to consider proper precautions to defend against how threat actors operate. Several motives can be a reason for a data breach, such as fear, ideology, grudge, espionage, convenience, fun, and financial. Based upon this analysis, we noticed that the financial motive is the primary motive for adversaries, followed by looking for fun.

External Actors. External threats originate from outside of an organization and its third-party partners [15]. Examples include criminal groups, lone hackers, former employees, and government entities. It is also comprised of God (as in "acts of"), "Mother Nature," and random chance. Typically, no trust or privilege is implied for external entities. We found out that 97% of the external actor motives are financial, 1% are for fun. Figure 3 shows the different motives of the actor's external motives.

Internal Actors. Internal threats originate from within the organization, which encompasses full-time company employees, independent contractors, interns, and other staff. Insiders are trusted and privileged (some more than others). Upon further analysis, we found that 52% of the internal motives for adversaries are financial, while 27% are for fun. Figure 3 presents the distribution of motives for internal motives.

Partner Actors. Partners include any third party sharing a business relationship with the organization, including suppliers, vendors, hosting providers, out-

sourced IT support, and others. Some level of trust and privilege is usually implied between business partners [15]. Based on this analysis, we found out that most of the motives behind the incidents are financial 67%; fun and convenience are 17% and 8% respectively. The remaining results for the internal motives distribution are shown in Fig. 3.

Results: Data Breaches Victims. We analyzed the most targeted victims from adversaries according to the number of incidents. Reasons often differ as to why these victims have been targeted, and it also depends on several other aspects such as location, specific personal information, or a high number of patients in a hospital. We found that most of the targeted victims are patients (88%), the customer came in second (5%), and 5% for employees. Other types of victims include students (interns) working inside healthcare organizations or third-party companies that share data with a specific entity. Figure 4 shows the most targeted victims in the incidents.

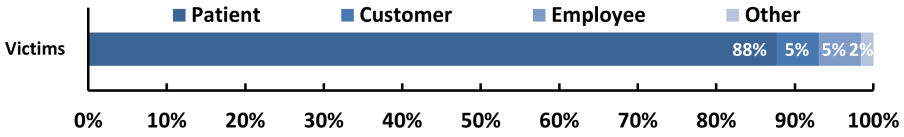


Fig. 4. Distribution of incidents by the different targeted victims from attackers.

Takeaway. *For the three threat actors combined, financial gain is the primary motive for adversaries to launch their attacks.*

5 Analyzing Threat Actions

In this section, we introduce our measurement and analysis of the threat actions used by adversaries during a data breach. This investigation intends to provide insight and the causes of threat actions and their occurrences in our dataset. The following section discusses the two types of threat actions: the different action varieties and the most used vectors by adversaries during an attack. The VERIS dataset classifies threat actions into seven primary categories: malware, social, hacking, misuse, physical, error, and environmental. Analyzing threat actions is essential due to the amount of risk associated with each of them every time an incident occurs. Generally, an incident usually contains a least one of the threat actions; however, most of the incidents will comprise multiple actions that often come with numerous categories.

Terminology Definitions. Below, we define several types of threat actions.

Malware Malicious software or malware is a computer code designed to disable, disrupt, or take control of the computer system by altering its state or function without the owner's informed consent [15]. Malware exploits technical flaws or vulnerabilities in hardware or software.

Hacking Refers to all attempts to intentionally access or harm information assets without (or exceeding) authorization by circumventing or thwarting logical security mechanisms. It includes brute force, SQL injection, cryptanalysis, denial of service attacks, etc. [15].

Social Social engineering criminals strive to exploit the users of these technologies by pretending to be something they are not to persuade others. Attackers utilize the trust to their advantage by misleading users into disclosing information that compromises data security. Social engineering tactics employ deception, manipulation, intimidation, and other techniques to exploit the human element, or users, of information assets, includes pretexting, phishing, blackmail, threats, scams, etc. [15].

Misuse The use of entrusted organizational resources or privileges for any purpose or manner contrary to intended is considered misuse. It includes administrative abuse, use policy violations, use of non-approved assets, etc. [15]. These actions can be malicious or non-malicious.

Physical Encompass deliberate threats that involve proximity, possession, or force. These include theft, tampering, snooping, sabotage, local device access, assault, etc. [15]. Natural hazards and power failures are classified into physical actions. However, VERIS restricts these events to intentional incidents only caused by human actors.

Error Error broadly encompasses anything done (or left undone) incorrectly or inadvertently. It includes omissions, misconfigurations, programming errors, malfunctions, etc. [15]. It does not include any intentional incidents.

Environmental The environmental category includes natural events such as earthquakes and floods and hazards associated with the immediate environment or infrastructure in which assets are located. The latter encompasses power failures, electrical interference, pipe leaks, and atmospheric conditions.

Results: Threat Actions Analysis. We measured the existence of each threat action category by calculating their varieties and vectors used in the incidents. We observed that ransomware represents 82% of the malware threat followed by others 8%. VERIS "other" to define any enumeration not represented by one of the categories in the data set. For the social threat actions category, with a percentage of 69%, phishing plays a large part in threat actions. The use of stolen credentials represents 80% of the hacking threat actions. With an increase in the number of employees, errors increased. Loss errors represent the main factor in this threat actions category representing 28%, followed by a disposal error of 27%. It is worth noticing that theft in the physical threat actions category with a percentage of 96%. Finally, privilege abuse in the misuse category with a rate of 59% is behind most of the threat actions in these two categories.

On the other hand, when we analyze the threat action vectors as shown in Fig. 6, we found out that the direct install represents 45% of the malware threat

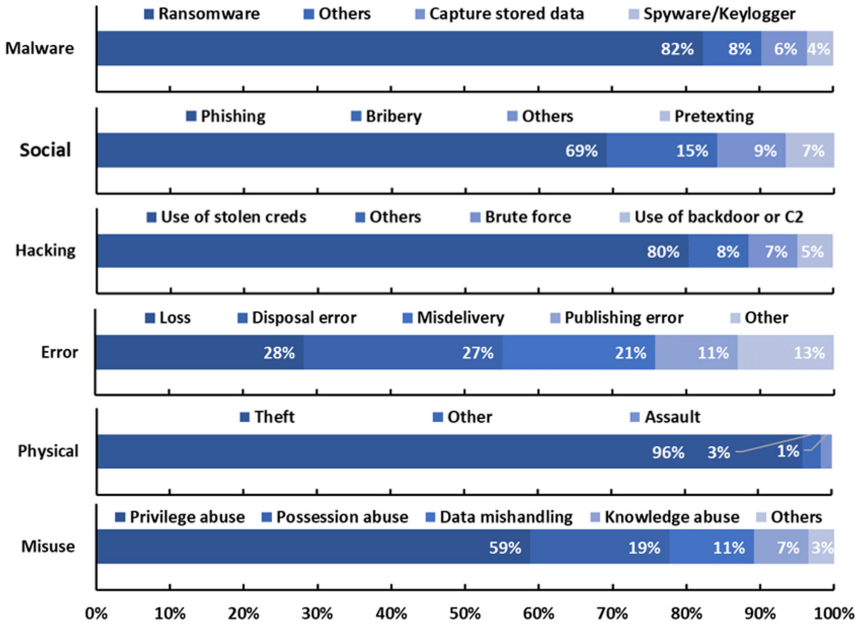


Fig. 5. Threat actions varieties.

actions. Email attachment is the second most common malware breach vector with 32%. The email vector represents 81% in the social category for other categories of threat actions, and web applications represent the primary vector with 81% of all hacking threat actions.

Carelessness is the primary vector with 92% of the error category. Although most data breaches using hacking by threat actors involve brute force, or the use of lost or stolen credentials [1], At the same time, LAN access is the most effective vector in the misuse category with 65%. It is clear that email and web application vectors represent the highest percentages among other vectors, and this is associated with the shift of valuable data to the cloud, including email accounts and business-related processes [1].

Takeaway. *Despite the variety of vectors, ransomware is still the leading malware method involving 82% of the incidents.*

6 Related Work

Recently, several studies have been conducted aimed at analyzing data breaches in the healthcare sector. Choi *et al.* [2] estimate the relationship between data breaches and hospital advertising expenditures. They concluded that teaching hospitals were associated with significantly higher advertising expenditures in

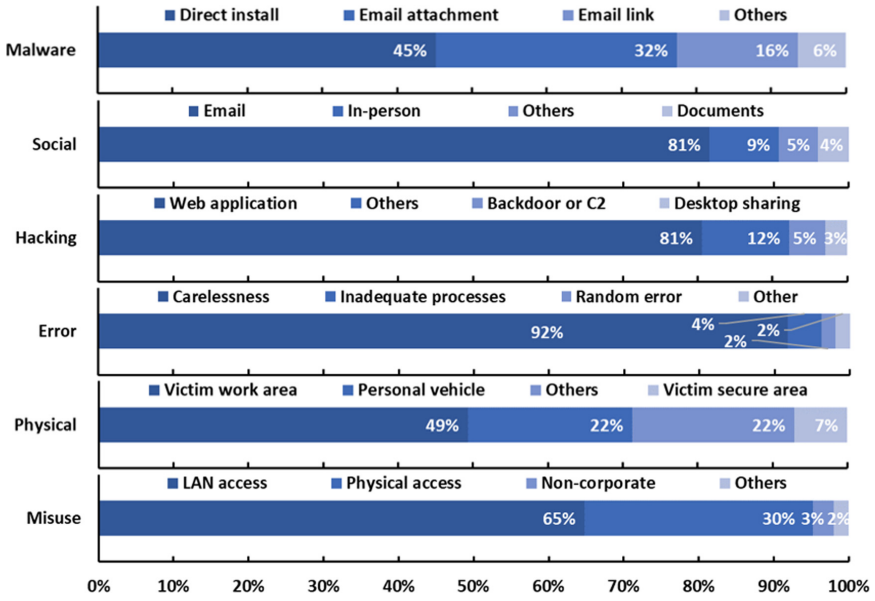


Fig. 6. Threat actions vectors.

the two years after the breach. Another study [9] 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. Siddartha *et al.* [12] found that the healthcare industry is being targeted for two main reasons: being a rich source of valuable data, and its weak defenses. Another study [17] suggested a framework to examine the accuracy of automatic privacy auditing tools. Another study [12] suggested that current healthcare security techniques miss data analysis improvements, e.g., data format-preserving, data size preserving, and other factors.

Using a collection of two years of data, [14] characterized DNS requests and TLS/SSL communications to understand the threats faced within the hospital environment without disturbing the operational network. Another analysis in [11] shows that attackers mostly use hacking/IT incidents, and the email and network servers are the primary locations for confidential health data breach.

Most relevant to our work, the 2020 Data Breach Investigations Report [4] summarized the findings and determined that external actors are behind 80% of data breaches while 20% of data breaches involved internal actors. According to the same report, hacking is the action that was used in 45% of data breaches, followed by errors that were causal events in 22% of breaches. The rest of the actions used in data breaches are social attacks, malware, misuse by authorized users, and physical actions presented in data breaches. The report also shows that financially motivated breaches are more common than espionage by a wide margin. In contrast to our work, a study by [11] presents information on data

breach incidents by sector, and they focused on the data breaches that occurred in the healthcare industry in the last five years because they account for 61.55% of the total data breaches. Authors of the following study have also compiled the data of healthcare breaches published by the HIPAA journal from 2010 to 2019 to authenticate their data.

7 Conclusion

While analyzing the timeline of the data set that comprised all the data breaches, the results showed that the highest number of incidents occurred from 2010–2020. Moreover, this long-term study revealed that health organizations are exposed to internal, external, and partner attacks. The financial is the primary motivation for the external, internal, and partner attackers. Without a doubt, there is a high cost associated with data breaches, the price for each stolen health record increases with time. Moreover, the victims include different types, but the primary victims are the patients, followed by the customers and the employees. Based on a long-term analysis of the data set, the actions used by the threat actors are classified into seven categories: malware, hacking, social, misuse, physical, error, and environmental. Ransomware motivated 82% of malware threat actors, and 45% of malware threat actions are directly installed. In the future, it would be worthwhile examining the correlation between security breaches and other indicators, including GDP, hospital size, etc.

Acknowledgement. This work was supported by NRF-2016K1A1A2912757, the UCF ORC Graduate Fellowship Program and Faculty Mentorship Program.

References

1. Introduction to the 2020 DBIR: Verizon Enterprise Solutions. Verizon Enterprise (2020). https://vz.to/3h5rva1
2. Choi, S.J., Johnson, M.E.: Understanding the relationship between data breaches and hospital advertising expenditures. Am. J. Managed Care **25**(5), 14–20 (2019)
3. 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
4. Enterprise, V.: Verizon data breach investigations report (2020). https://vz.to/3w66fpa
5. HealthIT.gov: Benefits of ehrs (2017). https://bit.ly/3qA5X8S
6. IBM: Cost of a data breach report 2020 (2020). https://ibm.co/2TlDKX7
7. Kamoun, F., Nicho, M.: Human and organizational factors of healthcare data breaches: the swiss cheese model of data breach causation and prevention. Int. J. Heal. Inf. Syst. Inform. **9**(1), 42–60 (2014)
8. Liu, V., Musen, M.A., Chou, T.: Data breaches of protected health information in the United States. JAMA **313**(14), 1471–1473 (2015)
9. 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)

10. Menachemi1, N., Collum, T.H.: Benefits and drawbacks of electronic health record systems (2011). https://bit.ly/3x7XNXJ
11. Seh, A.H., et al.: Healthcare data breaches: insights and implications. Healthcare **8**, 133 (2020). https://doi.org/10.3390/healthcare8020133
12. 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
13. Smith, T.T.: Examining data privacy breaches in healthcare. Walden University, Technical report (2016)
14. Vargas, L., et al.: Digital healthcare-associated infection: a case study on the security of a major multi-campus hospital system (2019). https://doi.org/10.14722/ndss.2019.23444
15. Veris community: Veris (2021). https://bit.ly/3jrTUbZ
16. Wikina, S.: What caused the breach? An examination of use of information technology and health data breaches. Perspect. Health Inf. Manag. **11**, 1h (2014)
17. 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)