Economics of Cybersecurity

Databreaches

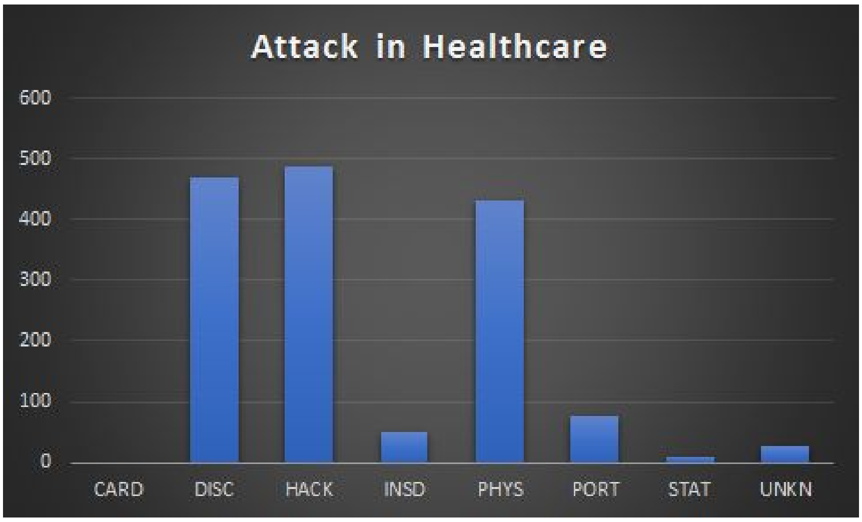
Group 5

Hasan Abdullah - 4614097
Seu Man To - 4064976
Elsa Turcios Rodriguez - 4597818
Tim van Rossum - 4246306

1. What security issue does the data speak to?

We have analyzed the data from Privacy Rights Clearinghouse which provide information of data breaches records from 2005 to 2017. The data breach record is also classified based on the type of organization type and based on the type of attack. We used all records from 2014 onwards.

Our first analysis of the data sets provide several findings. First, based on the type of organization, we found that 70% of the total data breaches happened in the Healthcare sector. Second, the type of attack that contribute the most to data breaches are Hacking and Unintended Disclosure.



Based on those two findings, we will try to narrow down the security issues that will be our main discussion topic furthermore. The issue that potentially caused the most of the data breaches, based on our data analysis finding, could be Unprofessionalism or unethical behavior of personnel. Especially in the healthcare sector where IT is not the main focus of the business, the allocation of IT expert and IT awareness training might be limited. Unprofessionalism can lead to the lack of security level in the organization's IT infrastructure resulting in it being more prone to hacking attacks. The lack of ethical or IT awareness could be the reason why there are a lot of data breaches caused by unintended disclosure.

2. What would be the ideal metrics for security decision makers?

According to Aghari, van Eten & Bauer (2016) getting trustable metrics is not an easy task, but this is a way of producing markets with a more efficient security. The ideal security metrics should include security cost, security levels and security benefits.

First of all, measuring the security cost allocated to improve Unprofessionalism or unethical behavior makes sense because the money allocated to avoid it will not be invested in other areas of the healthcare sector (Böhme, R. 2010). The first metric proposed in an ideal situation is ROSI (Return of Security Investment) including the cost that was avoided due to prevented losses which is really hard to measure in reality. In addition, the budget allocated to training to avoid this type of behavior, recurrent cost of training, cost of deployment of awareness campaigns, and sunk cost due to the turnover of personnel trained.

Secondly, security levels can be measured counting the incidents that happened due to unprofessionalism or unethical behavior, mean time to mitigate an unintended disclosure, mean cost to mitigate an unintended disclosure.

Finally, to measure the security benefits metrics such as the numbers of incidents because of unintended disclosure prevented, cost of liability avoided, and recovery cost avoided have to be performed.

3. What are the metrics that exist in practice?

On the site of the VERIS Community Database VCDB (VCDB, 2017), a couple of metrics are used. In fact, VERIS is a set of metrics that looks at the following:

- Incident counts by year
- Incident counts by actor (external, internal, partner, unknown)
- Incident counts by attribute (confidentiality, integrity, availability)
- Incident counts by action (type of breach, e.g. malware, hacking etc.)
- Incident counts by asset (e.g. server, network, person etc.)
- Incidents by industry (e.g. healthcare, real estate, finance etc.)

Metrics that can be used to recover from a cyber event include cost, time, damage assessment, and number of incidents (NIST, 2016). These can be used to assess the damage and cost of an incident.

The Center for Internet Security (CIS, 2010) has defined twenty-eight metrics for seven business functions. Some that can be used for our issue are cost of incidents, number of incidents, mean-time to mitigate vulnerabilities, and mean cost to mitigate vulnerabilities.

4. A definition of the metrics you can design from the dataset

The main metrics that can be designed using the current dataset are:

- Amount of breaches per year, to check if the amount of data breaches has gone up or down (this can then partly be used to conclude whether or not anti-breach measures are better)
- Amount of breaches per type of attacks, to check if some type of attack are more prone to data breaches than others
- Amount of breaches per type of organization, to see if some type of organizations are either more prone to being attacked than others or if some organizations are less secure

5. An evaluation of the metrics you have defined. This should include graphical representations of the metrics (e.g., histograms, scatter plots, time series, bar charts)

Before we start, a couple of abbreviations we use are listed below.

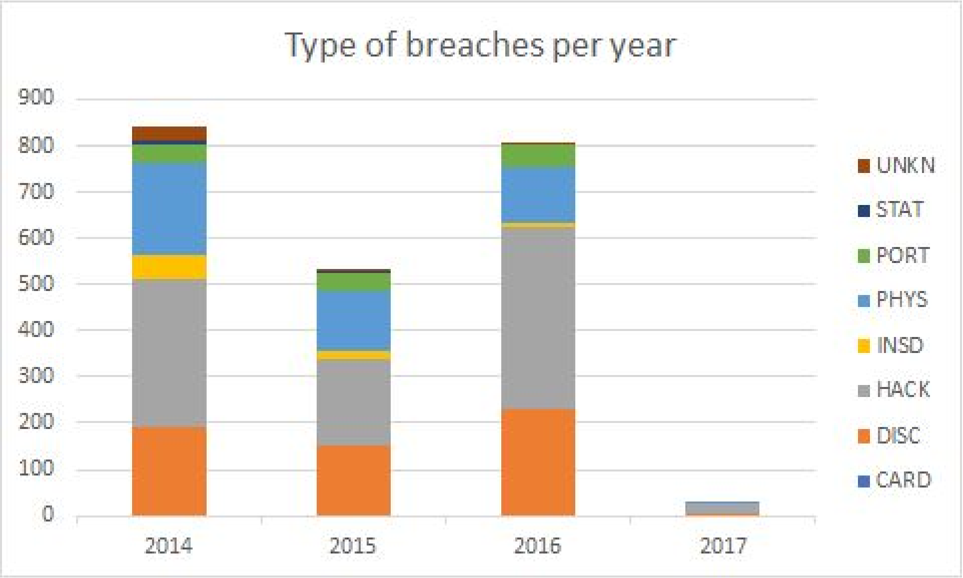
For the types of breaches we have:

- Payment Card Fraud (CARD)
- Hacking or Malware (HACK)
- Insider (INSD)
- Physical Loss (PHYS)
- Portable Device (PORT)
- Stationary Device (STAT)
- Unintended Disclosure (DISC)
- Unknown (UNKN)

For the types of organizations we have:

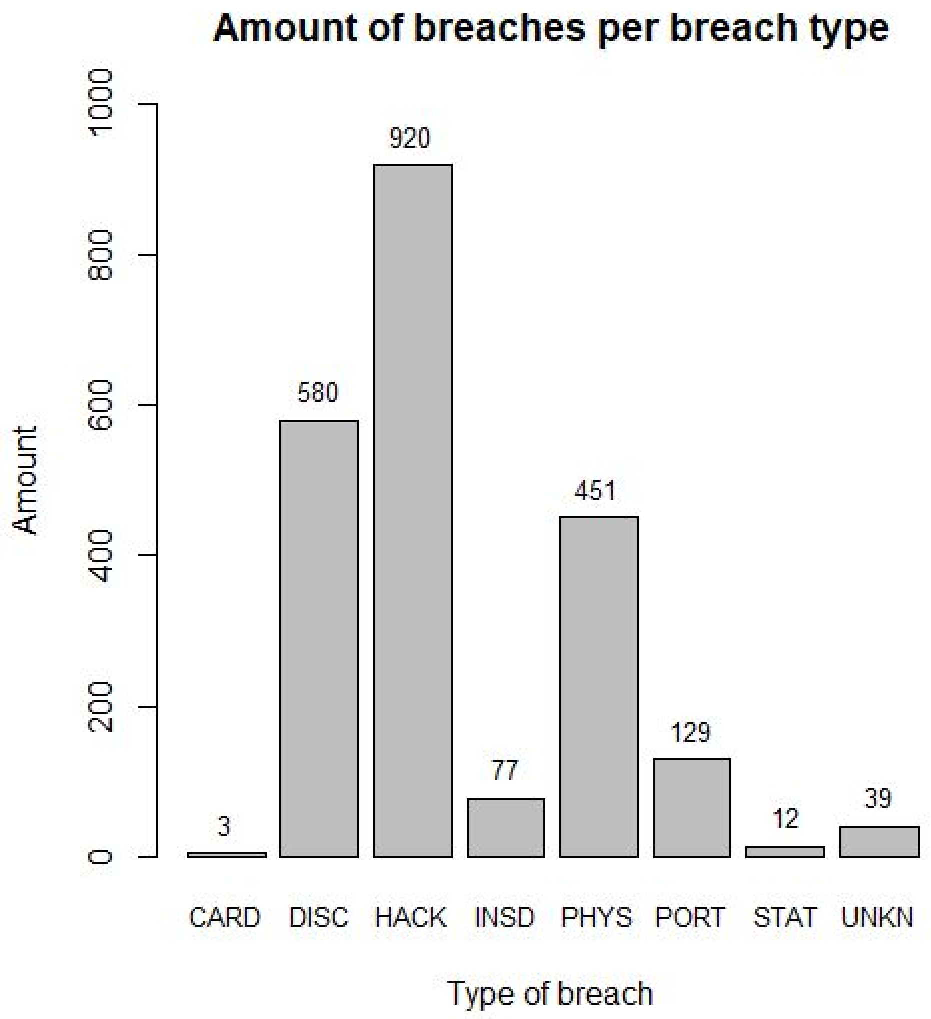
- Businesses-Financial and Insurance Services (BSF)
- Businesses-Other (BSO)
- Businesses-Retail/Merchant - Including Online Retail (BSR)
- Educational Institutions (EDU)
- Government and Military (GOV)
- Healthcare, Medical Providers and Medical Insurance Services (MED)
- Nonprofits (NGO)

Data breaches per year



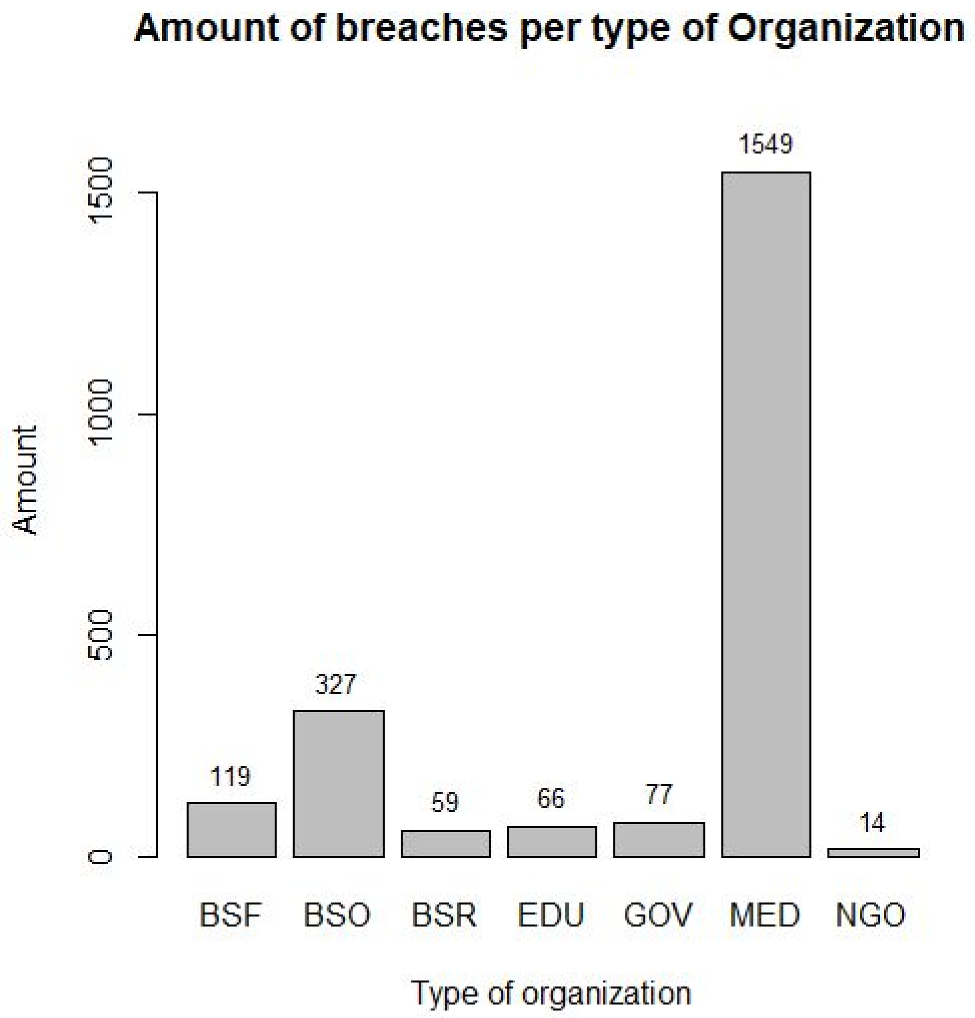
As can be seen in this histogram, most attacks took place in either 2014 or 2016. We expected the amount of attacks to gradually fall over the years, but it rose up again in 2016. This might have to do with attackers becoming more professional and general attacking behaviour changing to take advantage of new things.

Data breaches based on the type of attack



As can be seen, the main breaches are hacks, unintended disclosures, and physical losses. Also, payment card fraud is almost non-existent. We think that the latter has to do with possibly a higher security investment in credit card transactions (because a lot of people use payment cards for just about anything nowadays, especially here in The Netherlands), and that the former has to do with general network vulnerabilities and staff being poorly educated on hacks.

Data breaches based on the type of organization



As can be seen in this histogram, the amount of data breaches in medical institutions is far larger than the amount of data breaches in any other institution. We think this has to do with medical data being very sensitive and therefore it could be a popular target for attackers to obtain.

References

VCDB (2017). Retrieved from http://vcdb.org/explore.html

NIST (National Institute of Standards and Technology) (2016). Retrieved from
http://csrc.nist.gov/publications/drafts/800-184/sp800_184_draft.pdf

Asghari, Hadi, Michel van Eeten, Johannes M. Bauer. “Economics of Cybersecurity”, In Handbook on the Economics of the Internet (2016), Edgar Elgar, pp. 262–287, 2016.

Böhme, R. (2010, November), “Security Metrics and Security Investment Models”. In *IWSEC* (pp. 10-24).

CIS (Center for Internet Security) (2010), “The Center for Internet Security. The CIS Security Metrics”.