KOVRR
Cyber Decisions. Financially Quantified.

www.kovrr.com

Achieving Accurate Likelihood Probabilities for Cyber Events

SEPTEMBER 2022

Abstract

As a consequence of the rise of cyber attacks and the immense damages they cause, the need for cyber risk quantification and assessments emerged to help enterprises understand and decrease the risk to their business, both in terms of identifying potential areas where risk-reduction can take place, and where risk-transfer (through insurance) would be appropriate. One of the key components of cyber risk quantification is the ability to predict the frequency of cyber events. According to Kovrr's simulations, even a slight change of 0.02 in frequency of cyber events can cause up to 30-40% increase in the annual overall loss. The challenges of predicting frequency of cyber events includes: lack of data or very sparse data stemming from enterprises not willing to expose themselves, the natural chaotic dynamic of cyber events, understanding the natural target population of certain events, the fast pace of the technology landscape changing etc. The aim of this paper is to shed light on this issue and create an accurate output that reflects the frequencies of different types of cyber events.

In this paper, the team has established a methodology of aggregating and smoothing the data while addressing the challenges mentioned above. Kovrr has investigated the usage of different time-series state-of-the-art algorithms and created a hybrid approach achieving an average RMSE (Root Mean Square Error) of ~ 0.03 on the smoothed test data. Comparing this method with other approaches proved to be significantly better in capturing long term trends, cleaning noise perturbations and fitting the curve of raw observations. Moreover, this method possesses the mathematical robustness needed to balance over fitting issues, providing the flexibility to better generalize..

Introduction

The first distinction made regarding cyber events is that there are different types of cyber events, which by nature occur at different rates and naturally target different types of entities. In this paper we focus on Data Breach events and Ransomware events. First let's provide a clear definition of these two types of events:

Data Breach

"A data breach is an incident where information is stolen or taken from a system without the knowledge or authorization of the system's owner. A small or large organization may suffer a data breach."

Ransomware

"Ransomware is a type of malware that prevents or limits users from accessing their system, either by locking the system's screen or by locking the user's files until a ransom is paid."

Data sources

Kovrr's cyber incidents database, which contains both threat intelligence and financial data on a vast collection of cyber incidents was used in this work to derive frequencies of cyber events. The database contains a curated collection of both paid and open sources, from which various cyber intelligence and cyber attack data is collected. The collected data includes data on attacks (targets, attack vectors etc.), and financial data on the impact of the attacks.

Our solution

Data Preprocessing

First thing to note is that different industries and different sized companies experience different likelihoods to suffer from certain attacks. Based on our preliminary research regarding delay between events starting and being disclosed, we found many events in 2022 still emerging. We therefore used 1/1/2022 as an upper bound for our data and 1/1/2013 as a lower bound for our subset. We believe events that occurred prior to this date are obsolete. Moreover, we encountered a substantial increase in cyber events from 2013 and onwards. This data better reflects today's cyber risk landscape. Instead of using a strict Gregorian calendar, we attached a sliding window (January 2013 to December 2013, February 2013 to January 2014 and so on) of accumulated events per each time frame normalized by the compatible population size. This allowed for smoothing out of spikes in the data and created a better base to capture long term trends.

Data Breach

Our objective was to predict the frequency of data breach events per year per company. Different industries and different sized companies have different likelihoods to suffer from an attack (for example, companies in financial services naturally have more data that can be breached compared to agriculture or construction companies).

We found the right granularity for data breach events was separating the data into industry divisions. One of the main challenges was to determine the appropriate population size for each division (our "denominator"). Even more challenging was addressing the under-reporting of cyber incidents by firms (estimating the "numerator"). Large companies are less likely to suffer from under-reporting because they are subject to more regulations, clients and issues surrounding public relations. Therefore, we chose the Fortune1000 companies list as a representative of our population and derived our vanilla rates for these kinds of companies. We apply certain factors to rescale our rates for different revenue ranges, which we elaborate on in the next section.

We attached an aggregated sliding window to each division as mentioned above and the outcome can be seen in the next figure:

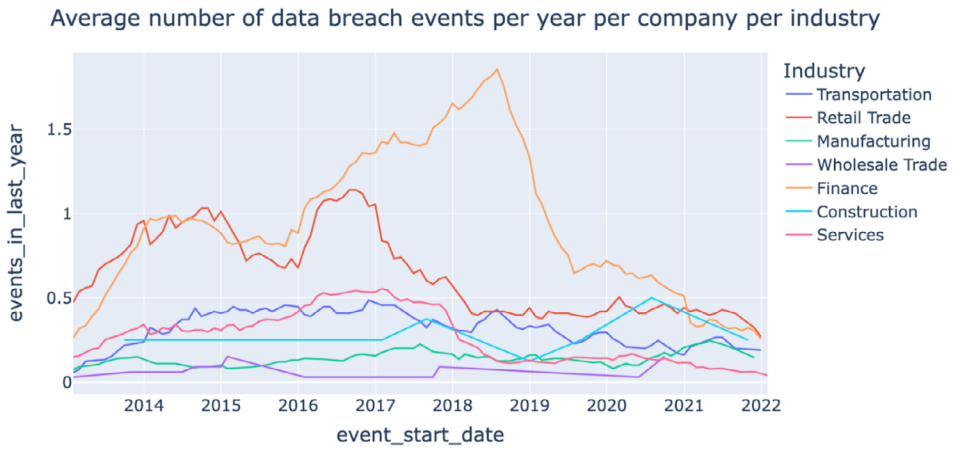

Figure 1

Ransomware

We found the right granularity for ransomware events was to separate the data into revenue ranges. We took Kovrr's database of companies as the representative for the population\denominator. We have taken into account that a vast amount of small companies are obviously missing from our data, therefore our "denominator" is not precise for small companies. Nevertheless, small companies suffer significantly more from the under-reporting issue (they are less likely to report due to a lack of regulatory requirements) and as a consequence our "numerator" decreases which provides compensation for the above.

Finally, our aggregated sliding window can be seen in the following figure:

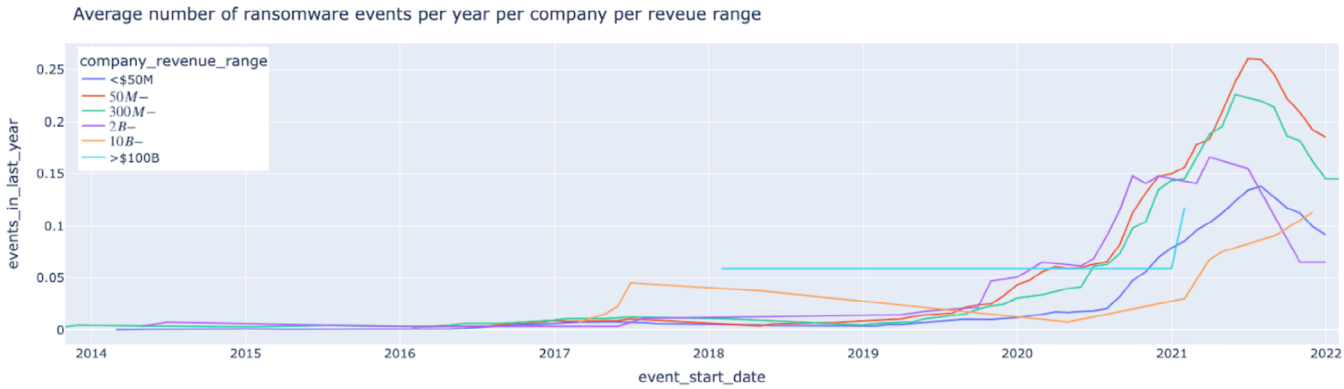

Figure 2

Methodology

Preliminary trials

Using our smoothed, aggregated data we have investigated different time-series algorithms. One approach was to use observation driven models. This approach leverages a dynamic process in which the current observation is seen as a function of past values. We also tried the ARIMA (Autoregressive Integrated Moving Average) model and LSTM (Long Short-Term Memory), both observation driven models. Alternatively, parameter driven models assume that the dynamic process is a function of the latent parameters of the model. In these kinds of models, the latent parameters of the distribution are a function of the unobserved dynamic process. Negative binomial regression belongs to the second approach and is the natural regression technique to infer a poisson-distributed response variable (count-based data or rates). The pros of this approach is that it allows us to directly model the relations between revenue and industry division to the outcome rates. Nevertheless, we have achieved higher errors using this method. Additionally, this approach assumes non-correlation between observations which contradict the way we have constructed our data (we have addressed this issue with some further preprocessing of the data, dealing with auto-correlated residuals).

In the following figure we can see predicted rates using negative binomial regression (after auto-correlation fixes) vs. real rates:

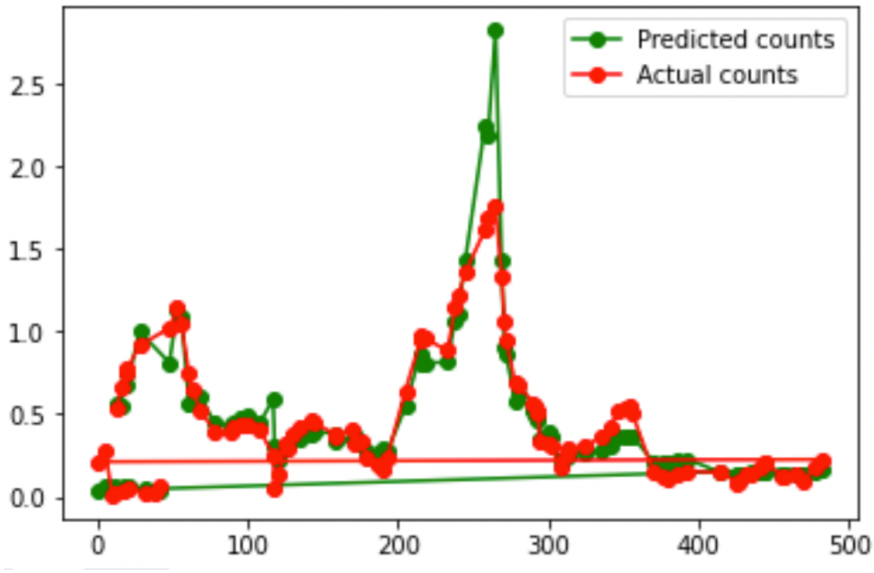

Figure 3

Chosen method

LSTM provided the lowest residuals, however, the differences between ARIMA and LSTM encouraged us to use a hybrid model as our final predictions for data breach events (for ransomware we have used a model based solely on LSTM). Consequently, Data Breach predictions benefit from the advantages of both of the methods (ARIMA is better in explaining linear tendencies of correlation, while the LSTM component explains non-linear tendencies).

For the ARIMA model, we have made simple transformations (shifting and log) to transform the observations into a stationary series (and confirmed it using an augmented Dickey-Fuller test). We then performed a grid search to identify the appropriate p, d and q values for each dataset.

For the LSTM model, we used a simple network which has a visible layer of 1 input, a hidden layer with 4 LSTM neurons, and an output layer that makes a single value prediction. The default sigmoid activation function was used for the LSTM blocks. The network was trained for 100 epochs, a batch size of 1 with ADAM used as an optimizer.

Proportion factors

As mentioned above, for ransomware, we used a model to predict an appropriate rate for each revenue range. On the contrary, for data breach events we used a model providing predictions per industry division based on Fortune1000 companies. As a result, the need for correction factors for different revenue ranges arose.

We created a separated negative binomial model for each industry division to analytically estimate the relation between revenue and the outcome rates (note that we can interpret the negative binomial regression coefficient as follows: for a one unit change in the predictor variable, the difference in the logs of expected counts of the response variable is expected to change by the respective regression coefficient, given the other predictor variables in the model are held constant). This approach did not prove itself and a simpler approach was used to estimate the factors. We created a smoothed sliding window of data breach rates based on revenue ranges that can be seen in the following figure:

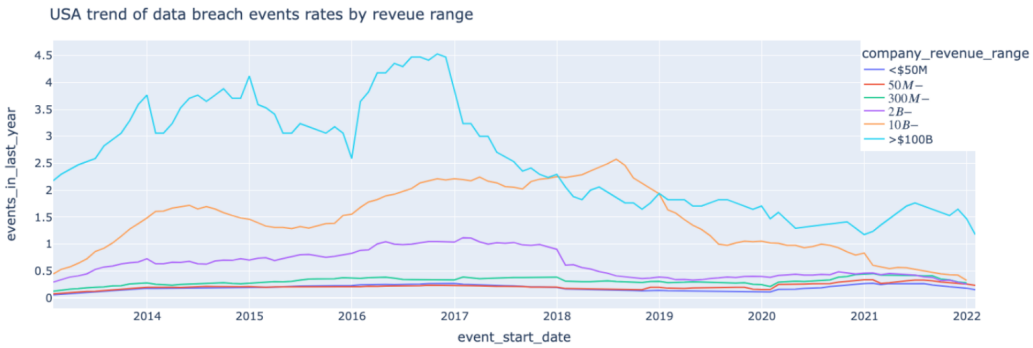

Figure 4

Based on this data we have calculated the average ratio between the different revenue ranges that resulted with the desired factors.

Results

Findings

As seen in figure 1, data breach rates for finance and retail trade industries are steadily higher than the rest of the industries throughout our entire time frame (2013-2022). In both industries, there are remarkably high rates between 2016-2019 (especially extreme peaks can be found in mid 2018 for the finance industry with almost 2 events per company). It should come as no surprise that our predicting rate for data breach events for both finance and retail trade industries (approximately 1 event every 3 years) is almost 2 times higher than our average predicted rate for all other industries. Additionally, figure 1 shows a sharp decrease in data breach events for the services industry in 2018 after which it plateaued at around ~ 0.07 (7 events in 100 years). In general,

data breach rates seem to be a bit noisy, but with moderate fluctuations providing a reasonable (not too chaotic) signal for predictions.

As seen in figure 2, there has been a rapid increase in ransomware rates from mid 2020 until mid 2021 (from mid 2021 we observe a drop down which we believe is partially due to delayed disclosures of events). Moreover, we can observe an inverse correlation between revenue range and ransomware rates - for higher revenues it is less probable to suffer a ransomware attack (extra small companies < 50M do not follow this rule as we can see they are likely to suffer a ransomware attack approximately once every 10 years which is about half of the rate we have predicted for small-medium companies - 50M-300M).

Errors

Our hybrid model for data breach events resulted with an average RMSE of ~ 0.04 on our test data while LSTM provided an average RMSE of ~ 0.02 for ransomware events. In the following figure we can see an example of the LSTM fit (the blue line is the raw data, the orange line is the train fitted model and the green line is the test fitted model).

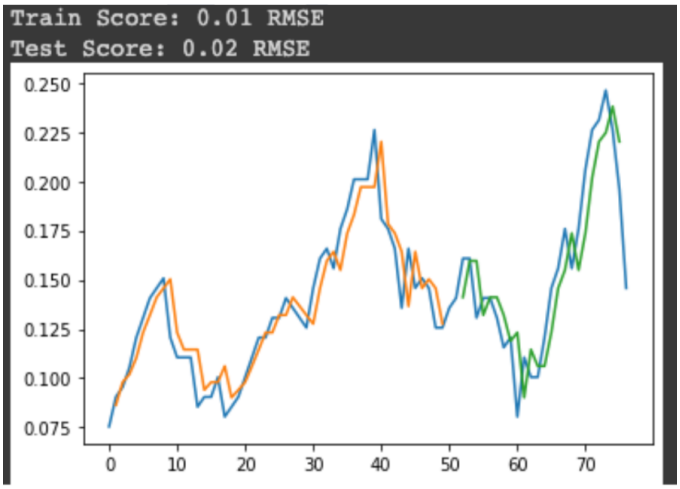

Figure 5

Remarks

Small companies (in terms of revenue) suffer significantly more from the under-reporting issue. As a result, our “numerator” (number of incidents) is significantly smaller than reality. Therefore, we have chosen our own dataset as a representative of our population from each division which is also smaller than reality (especially for small companies). As a consequence, our “denominator” is also smaller, compensating in a sense for the under-reporting issue.

We observe a drop down in ransomware raw data beginning in mid 2021. According to preliminary research this drop down might be due to the delayed disclosure issue. Fortunately, our LSTM predictions do not coincide with the bottom of this hill, but rather fall a bit beneath the middle of the slope. This way we can also compensate for this drop down (we believe this drop down is partially true and partially an artifact of the delayed disclosures).

Conclusion

We believe that we have achieved reasonable and robust results that can serve as solid priors for further estimations.

Future work

Several other algorithms can be examined as well. That includes: Prophet, PEWMA (based on Kalman filter which actually combines both approaches - observation based models and parameter based models), and an attention mechanism to try to have better control on weighting of recent observations.

Using a larger population based on the Russell 3000 index.

The Author

Amit Sarel

Data Scientist

About Kovrr

Kovrr's cyber risk modeling platform delivers global enterprises and (re)insurers transparent data-driven insights into their cyber risk exposures. The Kovrr platform is designed to help chief risk officers, chief information security officers, underwriters, exposure managers, risk professionals and catastrophe modelers understand, financially quantify and manage cyber risk by utilizing AI-powered risk models.

To learn more please contact the Kovrr team: contact@kovrr.com