










# Spatiotemporal Patterns, Typology and Survival Analysis of Cyber Incidents in Brazil

Gabriel Arquelau Pimenta Rodrigues<sup>1</sup> , Matheus Noschang de Oliveira<sup>1</sup> ,  
André Luiz Marques Serrano<sup>1</sup> , Amanda Nunes Lopes Espiñeira Lemos<sup>2</sup> ,  
Evanei Gomes dos Santos<sup>1</sup> , Geraldo Pereira Rocha Filho<sup>3</sup> ,  
Fábio Lúcio Lopes de Mendonça<sup>1</sup> , Daniel Alves da Silva<sup>1</sup> ,  
Vinícius Pereira Gonçalves<sup>1</sup> 

<sup>1</sup>Electrical Engineering Department (ENE) – University of Brasilia, Brasília, DF, Brazil

<sup>2</sup> Law School – University of Brasilia, Brasília, DF, Brazil

<sup>3</sup> Department of Exact and Technological Sciences (DCET) – State University of Southwest Bahia, Vitória da Conquista, BA, Brazil

{gabriel.arquelau,matheus.oliveira,evanei.santos,fabio.mendonca,daniel.alves}@redes.unb.br, {andrelms,vpgvinicius}@unb.br, amandaespiñeira@ccom.unb.br, geraldo.rocha@uesb.edu.br

**Abstract.** *Cyber incidents, such as data breaches, ransomware attacks and vulnerability exploitations, are consistently impacting Brazil. This study analyzes 1249 cyber incident reported to the Brazilian National Data Protection Authority from 2021 to 2025. We identify São Paulo and the Federal District as the most targeted states, with ransomware being the most prevalent attack type. The Kaplan-Meier survival analysis reveals rapid incident recurrence in major regions and of certain types of incidents, such as ransomware, vulnerable systems and credential theft. This study provides a foundational overview of Brazil's cybersecurity scenario. The findings emphasize the importance of enhancing incident reporting and of improving proactive defenses.*

## 1. Introduction

Brazil ranks among the most targeted countries for cybercrime, particularly in botnet infections, banking fraud, and financial malware [Bhardwaj et al. 2021], such as Astaroth. As these incidents escalate, investigating Brazil's cyber threat landscape becomes relevant for national cybersecurity strategies and for international cooperation. Cyber incidents, such as data breaches and ransomware attacks, have been found to reduce the public trust in the government and to cause anxiety and anger in the affected people [Shandler and Gomez 2023]. In addition to that, these events may also result in reputational damage and financial loss [Perera et al. 2022].

Given the relevance of this matter, Article 48 of the Brazilian General Data Protection Law (LGPD) establishes that data controllers must report cyber incidents that could cause significant harm to data subjects, which includes notifying both the affected individuals and the National Data Protection Authority (ANPD). Similar requirements are found in data protection laws around the world [Pimenta Rodrigues et al. 2024]. While the cybersecurity patterns in Brazil are well documented [Hurel and Lobato 2021], to the best of our knowledge, no study analyzes the ANPD's incident reports. Hence, this work

analyzes the security events documented by the ANPD to evaluate the temporal and geographical patterns of occurrence and the type of the event. It also conducts a survival analysis to assess the elapsed time between different cyber incidents in Brazil. By these means, this work seeks to quantify regional disparities in incident frequency and in typology, and to model recurrence intervals between cyber incidents.

This study, however, is subject to some limitations. The ANPD's dataset may underrepresent the true scope of cyber incidents in Brazil, as non-compliance with the legal obligations could result in unreported breaches. Also, the analysis is limited to a relatively short timeframe, which may not capture long-term trends. Ultimately, the dataset provides few features for each incident, especially when comparing to similar datasets for other countries, such as the Privacy Rights Clearinghouse for the USA. These limitations hinder a deeper analysis, but the work still establishes a foundational study and a benchmark for future research.

The remainder of this paper is structured as follows. Section 2 reviews some of the relevant similar works. Section 3 describes the methodology of the work and Section 4 presents the obtained results. Section 5 concludes the paper and proposes future works.

## **2. Related works**

To enable a proactive defense, [Almahmoud et al. 2023] propose a machine learning-based framework for long-term cyber threat forecasting, achieving a modified symmetric Mean Absolute Percentage Error. The work categorizes attack lifecycles into five phases, finding that password and vulnerability-related attacks are rapidly increasing threats. Their findings indicate the importance of cybersecurity measures to address such threats, to reduce the occurrence of incidents. This is relevant, for instance, due to password breaches, with weak or reused credentials being exploited in credential stuffing and brute-force attacks, leading to unauthorized access and to new incidents [Rodrigues et al. 2025].

An analysis of cybersecurity incident patterns is fundamental for improving threat detection, resource allocation and proactive defense strategies. As an example of this, [Rodrigues et al. 2024] have provided an exploratory data analysis of data breaches registered in the United States, promoting a discussion of security controls that could mitigate them. Our work aims to provide a similar analysis, with a focus on Brazil.

Brazil is a relevant case for studying emerging economies' vulnerabilities, as it has been found that there are several databases in the public Internet with critical vulnerabilities that violate its data protection law principles [Ponce et al. 2023]. The work identified the presence of malware like Meow Attack, RedisWannaMine, Mars and Balada Injector. These exposures indicate existent challenges in implementing data protection frameworks despite progressive legislation.

## **3. Materials and Methods**

This study utilizes official cyber incident records from the National Data Protection Authority of Brazil, covering the period from 11th February 2021 to 25th April 2025, which were obtained through a Freedom of Information request to the agency.

The dataset contains information about 1249 reports that, cumulatively, (i) have been confirmed by the responsible agent; (ii) involve personal data subject to Brazil's

LGPD; and (iii) may result in relevant risk or harm to data subjects. Out of these, 13 refer to other type of process (complaint or data subject request), thus resulting in 1236 cyber incidents within the time span. Among these, 86 are of foreign origin. The cyber incidents records are complete, with no null or missing values across all fields.

For each record, only three features are provided, namely the entry date, the state of occurrence and the incident type. A brief summary of these columns is presented in Table 1. The dataset is analyzed with Python version 3.11.12.

**Table 1. Statistical summary of the dataset columns**

Column	Count	Unique	Top	Freq
Entry date	1249	697	2022-01-24	13
State	1236	26	SP	434
Incident type	1249	19	Ransomware, no data transfer	243

The Hurst exponent ( $H$ ) quantifies the long-term predictability of a time series, indicating whether it exhibits persistent trends ( $0.5 < H \leq 1.0$ ), mean-reverting behavior ( $0.0 \leq H < 0.5$ ), or random noise ( $H = 0.5$ ). It is used in this work to assess the forecastability of the cyber incidents in Brazil, which may inform predictive models and guide the allocation of resources for proactive defense.

To compare the distributions of the types of incidents per state, the chi-square test is used. It evaluates whether observed categorical data deviates significantly from expected values under a null hypothesis. A p-value  $\leq 0.05$  suggests significant deviation, while p-value  $> 0.05$  indicates consistency with expectations. All p-values were adjusted for multiple comparisons using the Benjamini-Hochberg procedure to control the false discovery rate. We further quantified the strength of these differences using Cramér’s V effect size, where values approaching 0 indicate negligible association between variables, and values close to 1 signify stronger associations.

Ultimately, the Kaplan-Meier estimator is employed to model the time-dependent probability of a cyber incident occurring following an initial event. This method calculates the conditional survival probability at each observed incident time point, generating a step-function that represents the cumulative likelihood of remaining incident-free over successive days.

## 4. Results and Discussion

This section presents the results of the analysis of the reported incidents in Brazil, with a special focus on spatiotemporal patterns (Section 4.1), on its types (Section 4.2) and on the survival analysis (Section 4.3).

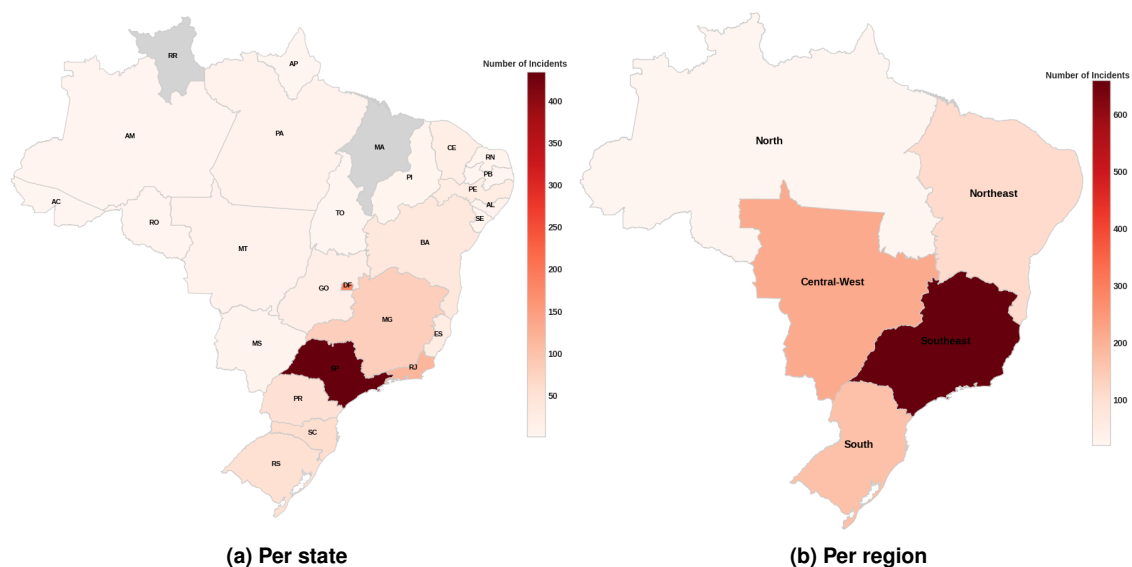
### 4.1. Spatiotemporal patterns

As seen in Figure 1a, São Paulo (SP) is the most affected Brazilian state, potentially due to its economic and technological importance in the country, with more businesses and residents than other areas. The Federal District (DF) ranks as the second most impacted federative unit, which is significant as it hosts the nation’s capital and concentrates most government services and public institutions.

The Southeast region has consistently reported the highest number of incidents throughout the observed period, as evidenced by both spatial (Figure 1b) and temporal (Figure 2) analyses. This pattern shows a temporal stability, with the sole exception occurring in January 2022, when the Central-West region temporarily surpassed all others in incident frequency.

It is important to note that variations in incident reporting across Brazilian states may reflect, beyond differences in cyberattack frequencies, their digital maturity, regulatory compliance, and incident response capacity. For instance, states with more developed digital economies and mature cybersecurity infrastructure may be more compliant with LGPD reporting requirements. This may skew the analysis by underrepresenting less digitally connected regions, where cyberattacks may go undetected or unreported.

The Hurst exponent analysis of the time series presented in Figure 2 provides preliminary knowledge regarding the predictability patterns. However, given the limited sample size, these results should be interpreted as indicative rather than definitive quantitative measures. The achieved Hurst exponents are: Central-West (0.5069), North (0.4015), Northeast (0.4017), South (0.4120), and Southeast (0.4566), with the aggregated series showing a value of 0.4567.



**Figure 1. Map of reported incident counts.**

These values, being close to 0.5, indicate a pattern close to a random walk, thus suggesting difficulty in forecasting. Some regions, such as North and Northeast, are more mean-reverting, but still present a noisy configuration.

## 4.2. Incident types

Ransomware incidents with and without data transfer or publication together account for 32.6% of the reported incidents (407 in total), as seen in Figure 3. Mitigation strategies for this type of malware are being adopted by the Brazilian government [Ferreira 2022].

Exploitation of systems vulnerability (195 incidents, 15.6%) and credential thefts and brute force (160 incidents, 12.8%) also represent a significant portion

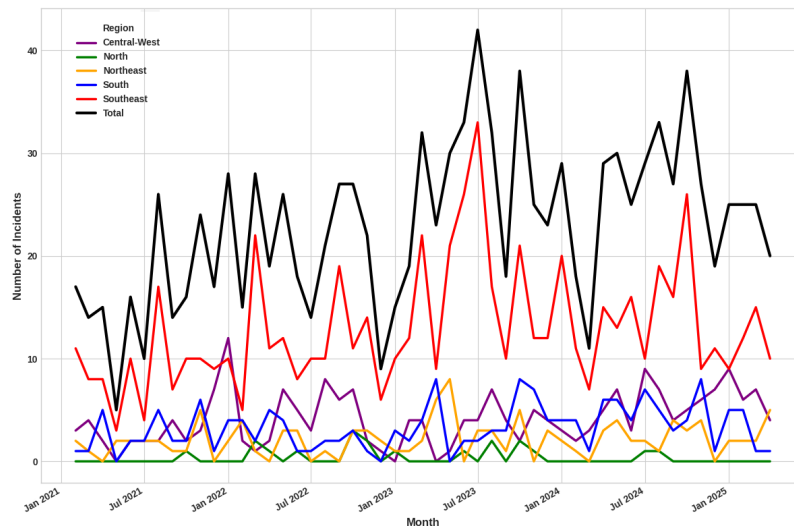


Figure 2. Time series of incidents per region

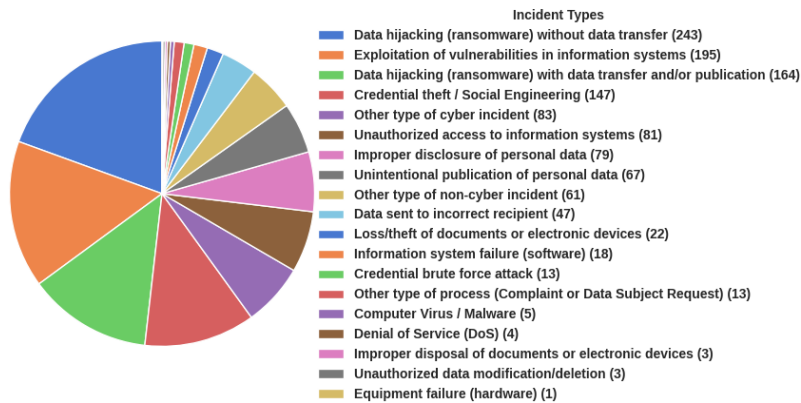


Figure 3. Frequency of incident types

of the security events reported to ANPD. These findings converge with those of [Almahmoud et al. 2023], that categorized these as rapidly increasing.

To compare the similarity in the distributions of the incident types per states, a  $\chi^2$  test is conducted, with the results presented in Table 2.

Table 2. Pairwise comparison of incident types distributions between states with  $\chi^2$  p-value < 0.5

State 1	State 2	Chi <sup>2</sup>	Cramér's V	p-value
Distrito Federal	São Paulo	83.57	0.33	2.87e-08
Distrito Federal	Minas Gerais	46.42	0.32	0.016
Bahia	São Paulo	46.38	0.25	0.016
Piauí	São Paulo	43.62	0.24	0.032

Pairwise comparisons revealed statistically significant differences in four pairs, namely DF–SP, DF–MG, BA–SP and PI–SP. The significant  $\chi^2$  result and moderate Cramér's V for the DF–SP and DF–MG comparison reveal that there are disparities in which types of incidents dominate each region, but with some structural similarity. The

greater p-values in other pairwise comparisons implies relative homogeneity in incident distributions among most states.

### 4.3. Survival analysis

To analyze the expected duration until a cyber incident occurs in each state or geographical region, we use the Kaplan-Meier curve. It models the time until an event happens, helping estimate the probability of remaining incident-free over time. The Kaplan-Meier method is used to estimate the duration before a system experiences a cyber incident [Bradley et al. 2023]. Plotting these curves for different regions enables a comparison of their resilience, identifying areas with higher or lower risk based on how quickly their survival probabilities decline.

Figure 4a, that presents the Kaplan-Meier curve for the 10 Brazilian states with highest incident counts, indicate that SP and DF have the greatest incident frequencies, as they have the steepest curves. São Paulo's curve drops to zero within nearly 30 days, indicating an almost certain likelihood of experiencing another cyber incident within a month, whereas seven out of the 10 states would reach this 0% chance after 150 days. All these states reach a 50% chance of new incident after 60 days after an event.

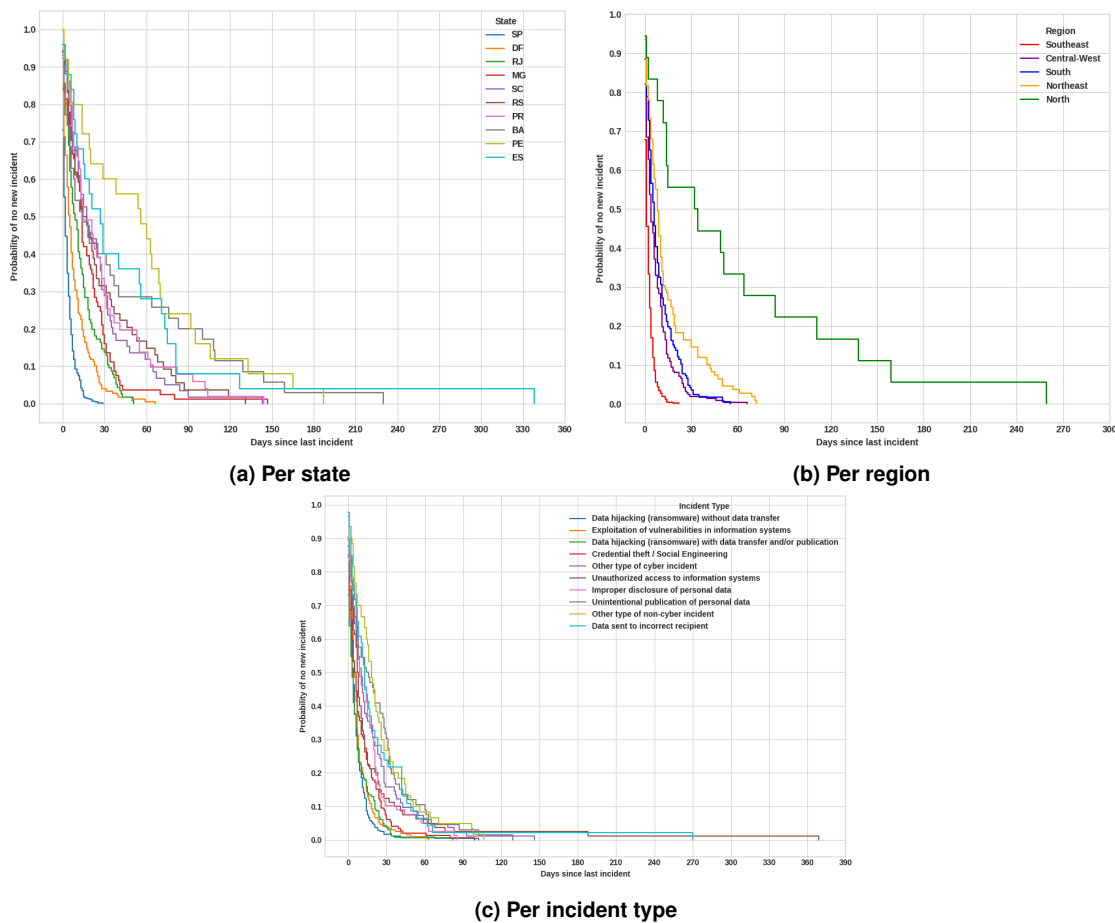


Figure 4. Kaplan-Meier curve for incident occurrence.

Similarly, Figure 4b shows that all geographical regions, with the exception of North, have 0% chance of not registering new incidents after 90 days. After around 30

days, there is less than 10% of chance of not suffering incidents related to ransomware, vulnerability or credential thefts, as depicted in Figure 4c.

## 5. Conclusions and Future Works

This study analyzed cyber incidents reported to the Brazilian National Data Protection Authority between February 2021 and April 2025, focusing on their spatiotemporal distribution, typology, and survival patterns. Despite the data limitations, such as the narrow time frame, our findings show that São Paulo and the Federal District are the most impacted federative units, likely reflecting their economic and governmental importance to the country. Ransomware remains the predominant type of cyber incident, and the survival analysis revealed a rapid recurrence of incidents.

These findings enable targeted mitigation strategies, as organizations in high-risk states should prioritize defenses against common incident types in their area, and they also indicate the need for region-specific cybersecurity policies. For example, the high incident recurrence rate in SP and DF emphasizes the urgency for an improvement in cybersecurity maturity. These states, due to their economic and governmental significance, should receive focused investment in incident detection, response capabilities, and awareness programs. Additionally, the survival thresholds suggest minimum response-time benchmarks for critical sectors.

For future work, we propose expanding the analysis to longer periods as more data becomes available, and integrating more features, such as organization sector and incident severity, and employing forecasting models to estimate future incident counts in the country. Comparative studies with other countries' datasets, like the U.S. Privacy Rights Clearinghouse, would also enhance the understanding of global cybersecurity trends and Brazil's position within them.

## References

- Almahmoud, Z., Yoo, P. D., Alhussein, O., Farhat, I., and Damiani, E. (2023). A holistic and proactive approach to forecasting cyber threats. *Scientific Reports*, 13(1):8049.
- Bhardwaj, G., Gupta, R., Srivastava, A. P., and Singh, S. V. (2021). Cyber threat landscape of G4 nations: Analysis of threat incidents & response strategies. In *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*, pages 75–79. IEEE.
- Bradley, T., Alhajjar, E., and Bastian, N. D. (2023). Novelty detection in network traffic: Using survival analysis for feature identification. In *2023 IEEE International Conference on Assured Autonomy (ICAA)*, pages 11–18. IEEE.
- Ferreira, L. V. (2022). Cybersecurity and ransomware in the brazilian government. *Revista InterAção*, 13(1):58–65.
- Hurel, L. M. and Lobato, L. C. (2021). Cyber security governance in brazil: Keeping silos or building bridges? In *Routledge companion to global cyber-security strategy*, pages 504–518. Routledge.
- Perera, S., Jin, X., Maurushat, A., and Opoku, D.-G. J. (2022). Factors affecting reputational damage to organisations due to cyberattacks. In *Informatics*, volume 9, page 28. Multidisciplinary Digital Publishing Institute.

- Pimenta Rodrigues, G. A., Marques Serrano, A. L., Lopes Espiñeira Lemos, A. N., Canedo, E. D., Mendonça, F. L. L. d., de Oliveira Albuquerque, R., Sandoval Orozco, A. L., and García Villalba, L. J. (2024). Understanding data breach from a global perspective: Incident visualization and data protection law review. *Data*, 9(2):27.
- Ponce, L. M., Gimpel, M., Ribeiro, I., Oliveira, E., Cunha, Í., Hoepers, C., Steding-Jessen, K., Chaves, M. H., Guedes, D., and Meira Jr, W. (2023). Um arcabouço para processamento escalável de vulnerabilidades e caracterização de riscos à conformidade da LGPD. In *Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg)*, pages 15–28. SBC.
- Rodrigues, G. A. P., Fernandes, P. A. G., Serrano, A. L. M., Rocha Filho, G. P., Vergara, G. F., Bispo, G. D., de Oliveira Albuquerque, R., and Gonçalves, V. P. (2025). From RockYou to RockYou2024: Analyzing password patterns across generations, their use in industrial systems and vulnerability to password guessing attacks. *Journal of Internet Services and Applications*, 16(1):69–86.
- Rodrigues, G. A. P., Serrano, A. L. M., Vergara, G. F., Albuquerque, R. d. O., and Nze, G. D. A. (2024). Impact, compliance, and countermeasures in relation to data breaches in publicly traded US companies. *Future Internet*, 16(6):201.
- Shandler, R. and Gomez, M. A. (2023). The hidden threat of cyber-attacks—undermining public confidence in government. *Journal of Information Technology & Politics*, 20(4):359–374.