

Análise da Recorrência de Técnicas e Táticas do MITRE ATT&CK em Relatórios Públicos de *Threat Intelligence*

Charllynson Carvalho Caxias¹, Brenda Salenave Santana¹

¹Centro de Desenvolvimento Tecnológico – Universidade Federal de Pelotas (UFPel)
Caixa Postal 354 – 96.010-610 – Pelotas – RS – Brasil

{charllynson, bssalenave}@inf.ufpel.edu.br

Abstract. *The growing need to understand adversary behaviors has been driven by the automated analysis of Threat Intelligence reports. This work proposes an approach for identifying and analyzing the recurrence of MITRE ATT&CK techniques and tactics in public CTI reports. The methodology combines textual constraints, technique identification, and mapping to the framework's official catalogs, enabling the structuring of non-standardized information for frequency and temporal distribution analysis. The approach was applied to a corpus of public reports to characterize recurring patterns of adversary behavior.*

Resumo. *A crescente necessidade de compreender comportamentos adversários tem impulsionado a análise automatizada de relatórios de Threat Intelligence. Este trabalho propõe uma abordagem para identificar e analisar a recorrência de técnicas e táticas do MITRE ATT&CK em relatórios públicos de CTI. A metodologia combina extração textual, identificação de técnicas e associação aos catálogos oficiais do framework, permitindo estruturar informações não padronizadas para análise de frequência e de distribuição temporal. A abordagem foi aplicada a um corpus de relatórios públicos, com o objetivo de caracterizar padrões recorrentes de comportamento adversário.*

1. Introdução

A *Cyber Threat Intelligence* (CTI) reúne informações sobre ameaças cibernéticas que podem apoiar atividades de detecção, investigação e resposta a incidentes. Entre suas fontes estão relatórios públicos produzidos por organizações de segurança, equipes de resposta e empresas especializadas, que descrevem campanhas, vulnerabilidades, grupos, *malwares* e comportamentos adversários (Johnson et al. 2016). Esses documentos apresentam diferentes níveis de detalhamento e terminologias, dificultando a identificação sistemática de padrões recorrentes (Legoy et al. 2020; Caxias and Santana 2026).

Nesse contexto, o MITRE ATT&CK¹ constitui uma base de conhecimento para representação de comportamentos adversários, organizando-os em táticas, técnicas e subtécnicas (Al-Sada et al. 2023). Enquanto as táticas representam objetivos adversários, técnicas e subtécnicas descrevem os mecanismos empregados para alcançá-los. O framework disponibiliza matrizes para diferentes contextos tecnológicos, incluindo *Enterprise*, *Mobile* e *ICS*.

¹<https://attack.mitre.org/>

A utilização de TTPs (*Tactics, Techniques and Procedures*) permite representar características comportamentais dos adversários, complementando abordagens baseadas em indicadores específicos, como endereços IP, domínios e hashes (Johnson et al. 2016). A associação de comportamentos descritos em relatórios de CTI a uma taxonomia comum permite, assim, estruturar informações heterogêneas e viabilizar análises comparativas.

Apesar da disponibilidade desses documentos, sua heterogeneidade dificulta a análise sistemática. Nesse contexto, este trabalho investiga a recorrência de técnicas e táticas do MITRE ATT&CK em relatórios públicos de CTI, considerando sua frequência e distribuição temporal.

Para isso, são extraídos identificadores ATT&CK explicitamente presentes nos relatórios e associados aos catálogos oficiais das matrizes *Enterprise*, *Mobile* e *ICS*, permitindo caracterizar as técnicas identificadas e analisar sua recorrência no corpus. Diante disso, o trabalho busca responder às seguintes questões de pesquisa:

- **RQ1:** Quais técnicas e subtécnicas do MITRE ATT&CK apresentam maior frequência nos relatórios públicos analisados?
- **RQ2:** Quais táticas apresentam maior recorrência no período analisado?
- **RQ3:** Quais técnicas permanecem recorrentes ao longo do tempo?

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 descreve a metodologia; a Seção 4 apresenta e discute os resultados; e a Seção 5 apresenta as conclusões e possibilidades de trabalhos futuros.

2. Trabalhos Relacionados

A identificação automatizada de comportamentos adversários em textos não estruturados tem sido explorada para estruturar relatórios de CTI com base no MITRE ATT&CK. (Legoy et al. 2020) propuseram o rcATT, utilizando aprendizado de máquina multirrótulo e TF-IDF para classificar táticas e técnicas. (Rani et al. 2023) desenvolveram o TTPHunter, baseado em embeddings BERT e RoBERTa, alcançando F1 de 88% em sentenças e 75% em relatórios completos. (Rahman et al. 2024) combinaram aprendizagem ativa e SciBERT para reduzir a quantidade de dados necessários à anotação, enquanto (Rahman et al. 2025) exploraram relações temporais entre técnicas do ATT&CK em relatórios de CTI.

Em conjunto, esses trabalhos avançam principalmente na extração e classificação automatizada de TTPs. Entretanto, a recorrência dos comportamentos já identificados e sua distribuição ao longo do tempo permanecem menos exploradas. Nesse contexto, este trabalho adota uma perspectiva exploratória e quantitativa, utilizando expressões regulares para estruturar o corpus a partir dos identificadores ATT&CK explicitamente registrados nos documentos. A abordagem busca caracterizar a frequência das técnicas, sua distribuição entre táticas e sua recorrência em diferentes períodos.

3. Metodologia

O processo metodológico desta pesquisa foi estruturado nas etapas ilustradas na Figura 1.

3.1. Caracterização do Corpus

O estudo foi realizado com base em relatórios públicos de CTI, disponibilizados em formato PDF por organizações da área de segurança cibernética. O corpus inicial foi

composto por 18 relatórios da CISA, 20 da Unit 42, 1 relatório do CERT.PL e 1 relatório da Microsoft Defense. Após a verificação de documentos equivalentes, obteve-se um corpus final de 39 relatórios distintos.

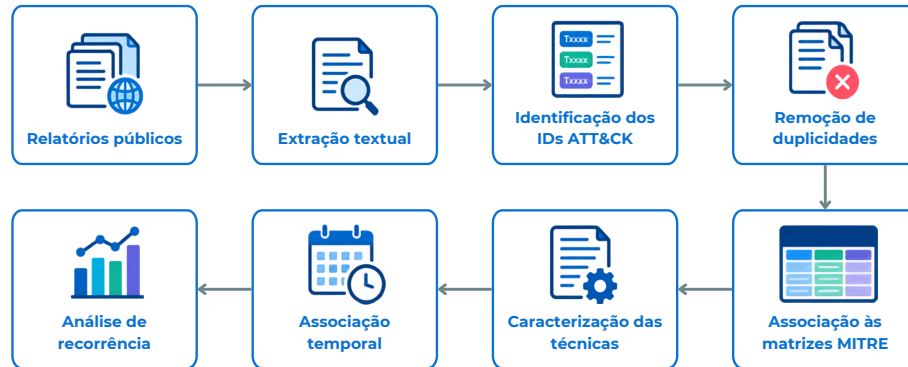


Figura 1. Metodologia empregada no pipeline

3.2. Extração, identificação e caracterização

Os arquivos foram processados automaticamente em Python, utilizando a biblioteca PyMuPDF² para extração textual. A identificação das técnicas e subtécnicas foi realizada por expressões regulares capazes de localizar identificadores no padrão MITRE ATT&CK, como T1059 e T1059.001. As ocorrências foram associadas aos respectivos relatórios e deduplicadas pela combinação entre relatório e técnica, de modo que múltiplas menções de uma mesma técnica em um documento fossem contabilizadas como uma única ocorrência, resultando em 112 relações únicas.

Os identificadores extraídos foram confrontados com as bases oficiais STIX do MITRE ATT&CK v18.0, considerando as matrizes *Enterprise*, *Mobile* e *ICS*. Para cada identificador válido foram obtidos nome, descrição, matriz e táticas associadas, permitindo a caracterização padronizada das técnicas encontradas.

A estratégia apresenta como vantagem a simplicidade e a reprodutibilidade, mas possui uma limitação: técnicas mencionadas sem identificador ATT&CK explícito não são capturadas. Assim, os resultados representam a recorrência dos identificadores registrados nos relatórios, e não a ocorrência exhaustiva de todos os comportamentos adversários descritos.

3.3. Análise de recorrência

A data de publicação de cada relatório foi associada às respectivas técnicas, possibilitando sua organização cronológica. A RQ1 foi analisada pela frequência de relatórios distintos associados a cada técnica; a RQ2, pela distribuição das técnicas segundo suas táticas; e a RQ3, pela identificação das técnicas presentes em diferentes anos do período estudado.

4. Resultados e Discussão

A análise dos 39 relatórios distintos resultou em 112 relações únicas entre relatórios e técnicas e 106 entre técnicas e subtécnicas.

²Ferramenta *open-source* em Python para extração e manipulação de documentos PDF.

4.1. Frequência das técnicas e subtécnicas

Observou-se uma distribuição dispersa: seis técnicas apresentaram a maior frequência, com duas ocorrências em relatórios distintos cada: T1036 (*Masquerading*), T1003.003 (*NTDS*), T1190 (*Exploit Public-Facing Application*), T1105 (*Ingress Tool Transfer*), T1083 (*File and Directory Discovery*) e T1071.001 (*Web Protocols*). As demais 100 técnicas ou subtécnicas foram identificadas em apenas um relatório.

A baixa concentração indica elevada heterogeneidade nos comportamentos descritos. As técnicas mais recorrentes abrangem diferentes etapas da intrusão: T1190 está relacionada ao acesso inicial por exploração de aplicações expostas; T1105, à transferência de ferramentas; T1083, à descoberta de arquivos e diretórios; T1003.003, à obtenção de credenciais a partir do NTDS; e T1071.001, à comunicação de comando e controle por protocolos web. Assim, as técnicas recorrentes não se concentram em uma única etapa do ciclo de ataque.

Embora essas seis técnicas apresentem a maior frequência no corpus, nenhuma foi identificada em mais de dois relatórios. Portanto, os resultados caracterizam a amostra analisada e não permitem inferir quais técnicas são mais utilizadas no cenário geral de ameaças cibernéticas.

4.2. Recorrência das táticas

A distribuição das técnicas segundo as táticas do MITRE ATT&CK mostrou diferenças na recorrência entre os relatórios analisados. As táticas *Execution* e *Initial Access* apresentaram a maior recorrência, estando presentes em cinco relatórios distintos cada. *Credential Access*, *Command and Control*, *Collection*, *Exfiltration* e *Stealth* foram identificadas em quatro relatórios cada; *Lateral Movement* e *Discovery*, em três; *Persistence*, *Impact* e *Privilege Escalation*, em dois; e *Defense Impairment*, em um.

A maior recorrência de *Execution* e *Initial Access* indica sua presença em parcela relevante dos documentos. Uma possível explicação está na natureza dos relatórios de CTI, que frequentemente descrevem o início do comprometimento e as ações executadas após o acesso inicial. Essa interpretação, contudo, constitui uma hipótese, pois o corpus não foi construído para representar estatisticamente a distribuição global das atividades adversárias.

A presença de *Credential Access*, *Command and Control*, *Collection* e *Exfiltration* em quatro relatórios indica que parte dos documentos também descreve atividades posteriores ao comprometimento inicial. A frequência das táticas foi calculada pela presença de suas técnicas em relatórios distintos. Assim, múltiplas técnicas associadas à mesma tática em um relatório foram contabilizadas uma única vez, evitando que documentos mais extensos ou detalhados tivessem peso excessivo.

4.3. Recorrência temporal das técnicas

A análise temporal considerou relatórios publicados entre 2023 e 2026. As 112 relações únicas entre relatórios e técnicas distribuíram-se em 25 relações em 2023, 78 em 2025 e 9 em 2026, sem registros em 2024. Como a quantidade e a composição dos relatórios variam entre os anos, esses valores caracterizam a composição temporal do corpus e não são utilizados como medida direta de crescimento ou redução da atividade adversária.

Das 106 técnicas e subtécnicas identificadas, 105 ocorreram em apenas um ano. A exceção foi T1190 (*Exploit Public-Facing Application*), identificada em 2025 e 2026. A Tabela 1 apresenta a distribuição das seis técnicas mais recorrentes ao longo dos anos.

Tabela 1. Distribuição das técnicas ao longo dos anos			
Técnica	2023	2025	2026
T1036 (<i>Masquerading</i>)	•		
T1003.003 (<i>NTDS</i>)		•	
T1190 (<i>Exploit Public-Facing Application</i>)		•	•
T1105 (<i>Ingress Tool Transfer</i>)		•	
T1083 (<i>File and Directory Discovery</i>)	•		
T1071.001 (<i>Web Protocols</i>)		•	

A presença de T1190 em dois anos distintos constitui o principal caso de persistência temporal identificado. Entretanto, essa recorrência não permite determinar aumento real desse tipo de atividade no cenário de ameaças, pois a quantidade e a composição dos relatórios variaram entre os anos.

A recorrência de uma técnica em diferentes períodos pode estar relacionada à repetição de comportamentos em campanhas distintas, à permanência de determinados vetores de comprometimento ou à própria seleção dos documentos que compõem o corpus. Dessa forma, a análise temporal identifica comportamentos que podem ser investigados posteriormente, mas não permite estabelecer relações causais a partir dos dados disponíveis.

A concentração de 105 das 106 técnicas em apenas um ano também evidencia a heterogeneidade temporal do corpus e pode decorrer da diversidade dos incidentes e relatórios selecionados, bem como da limitação da abordagem, que considera apenas identificadores ATT&CK explicitamente presentes nos documentos.

5. Conclusão

Este trabalho analisou a recorrência de técnicas e táticas do MITRE ATT&CK em 39 relatórios públicos de Threat Intelligence, totalizando 112 relações únicas entre relatórios e técnicas ou sub técnicas. Foram identificadas 106 técnicas ou sub técnicas distintas, das quais 100 ocorreram em apenas um relatório. As técnicas T1036, T1003.003, T1190, T1105, T1083 e T1071.001 apresentaram maior frequência, enquanto Execution e Initial Access foram as táticas mais recorrentes. Temporalmente, T1190 foi a única técnica identificada em mais de um ano, ocorrendo em 2025 e 2026.

Os resultados complementam abordagens anteriores de identificação de TTPs em textos, como (Legoy et al. 2020), (Rani et al. 2023) e (Rahman et al. 2024), ao analisar a recorrência das técnicas explicitamente identificadas em diferentes relatórios e períodos. A elevada dispersão das técnicas evidencia a heterogeneidade do corpus, enquanto a baixa recorrência temporal indica que poucos comportamentos foram observados em mais de um ano. Entretanto, a composição desigual do corpus entre os anos e a dependência de identificadores explicitamente registrados limitam a generalização dos resultados.

Como trabalhos futuros, pretende-se ampliar o corpus e incorporar métodos semânticos capazes de identificar técnicas mencionadas sem seus respectivos identificadores ATT&CK. Também será investigada a relação entre técnicas recorrentes e atributos como grupos, *malwares*, vulnerabilidades e campanhas.

Referências

- [Al-Sada et al. 2023] Al-Sada, B., Sadighian, A., and Oligeri, G. (2023). Mitre att&ck: State of the art and way forward.
- [Caxias and Santana 2026] Caxias, C. and Santana, B. (2026). Mapeamento e detecção de indicadores de comprometimento em relatórios de threat intelligence. In *Anais Estendidos do XXVI Simpósio Brasileiro de Cibersegurança*, pages 532–542, Porto Alegre, RS, Brasil. SBC.
- [Huang et al. 2024] Huang, Y.-T., Vaitheeshwari, R., Chen, M.-C., Lin, Y.-D., Hwang, R.-H., Lin, P.-C., Lai, Y.-C., Wu, E. H.-K., Chen, C.-H., Liao, Z.-J., and Chen, C.-K. (2024). Mitretrieval: Retrieving mitre techniques from unstructured threat reports by fusion of deep learning and ontology. *IEEE Transactions on Network and Service Management*, 21(4):4871–4887.
- [Johnson et al. 2016] Johnson, C., Badger, M., Waltermire, D., Snyder, J., and Skorupka, C. (2016). Guide to cyber threat information sharing. Technical Report NIST Special Publication 800-150, National Institute of Standards and Technology, Gaithersburg, MD.
- [Legoy et al. 2020] Legoy, V., Caselli, M., Seifert, C., and Peter, A. (2020). Automated retrieval of att&ck tactics and techniques for cyber threat reports.
- [Liu et al. 2022] Liu, C., Wang, J., and Chen, X. (2022). Threat intelligence att&ck extraction based on the attention transformer hierarchical recurrent neural network. *Applied Soft Computing*, 122:108826.
- [Rahman et al. 2024] Rahman, F. I., Halim, S. M., Singhal, A., and Khan, L. (2024). Alert: A framework for efficient extraction of attack techniques from cyber threat intelligence reports using active learning. In Ferrara, A. L. and Krishnan, R., editors, *Data and Applications Security and Privacy XXXVIII*, pages 203–220, Cham. Springer Nature Switzerland.
- [Rahman et al. 2025] Rahman, M. R., Wroblewski, B., Matthews, Q., Morgan, B., Menzies, T., and Williams, L. (2025). Mining temporal attack patterns from cyberthreat intelligence reports. *Knowledge and Information Systems*, 67(10):8941–8981.
- [Rani et al. 2023] Rani, N., Saha, B., Maurya, V., and Shukla, S. K. (2023). Ttphunter: Automated extraction of actionable intelligence as ttps from narrative threat reports. In *Proceedings of the 2023 Australasian Computer Science Week, ACSW '23*, page 126–134, New York, NY, USA. Association for Computing Machinery.
- [Strom et al. 2020] Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., and Thomas, C. B. (2020). Mitre att&ck®: Design and philosophy. Technical Report MP180360R1, The MITRE Corporation, McLean, VA.
- [Tamanna et al. 2026] Tamanna, M., Mitra, S., Erfan, M., Ryan, A., Mittal, S., Williams, L., and Rahman, M. R. (2026). What are adversaries doing? automating tactics, techniques, and procedures extraction: A systematic review.