Proactive management of offensive profiles: detecting trends in cyberattacks on institutions in Brazil through the analysis of hacker communities using complex networks and machine learning algorithms.

Claudio H. M. de Oliveira[1], Marcelo Ladeira[1], Flavio Q. Guimarães[2]

[1]Departamento de Ciências da Computação (CIC) – Universidade de Brasília (UnB)
Brasília – Distrito Federal – Brasil

[2]Instituto de Computação (IC) – Universidade Federal Fluminense (UFF)
Niterói – RJ– Brasil

claudio.oliveira@aluno.unb.br, mladeira@unb.br,
flavioqueiroz2004@gmail.com

Abstract. *"X"(Twitter) has established itself as an influential platform for the exchange of ideas and information, but it also attracts hackers for illegal activities. This study proposes an enhanced approach to detect offensive profiles linked to hacktivism on "X", utilizing complex networks and machine learning algorithms, focusing on notifiers from the Zone-H platform reporting hacktivist actions in Brazil. Key users were identified based on network metrics and keywords, analyzing their posts using clustering. The main contribution lies in identifying accounts aligned with hacktivism and assessing their threat potential to prevent cyberattacks, generating accurate and timely alerts.*

Resumo. *O Twitter, atual "X", é uma das maiores plataformas digitais para a troca de ideias e informações que atrai hackers com intuito de atividades ilegais e ações danosas. Este estudo propõe uma abordagem aprimorada para detectar perfis ofensivos ligados ao hacktivismo, utilizando redes complexas e algoritmos de aprendizado de máquina, com foco em notificadores da plataforma Zone-H que relatam ações hacktivistas no Brasil. Foram identificados usuários mais atuantes com base em métricas de rede e palavras-chave e clusterização. Esta é a principal contribuição na avaliação de ameaças para prevenir ataques cibernéticos, gerando alertas precisos e oportunos.*

1. Introdução

A partir da construção de uma rede de usuários baseada em notificadores listados no Zone-H, site de monitoramento de atividades hacker [Zone-H 2023], é possível identificar e categorizar usuários com comportamentos típicos de hackers atuando em redes no Brasil. A pesquisa destaca conexões entre usuários através de métricas de rede como centralidade, proximidade e intermediação, revelando atores centrais e influentes na comunidade hacker. Postagens desses atores centrais foram analisadas quanto à positividade ou negatividade e submetidas a técnicas de processamento de linguagem natural (NLP) para pré-processamento das postagens, e aprendizado de máquina.

O estudo inova ao explorar um conjunto de dados brasileiro autêntico de perfis de hackers no "X" com foco em *defacements*, um tipo de ataque cibernético onde um invasor modifica a aparência visual de um site ou página da web. Embora a eficácia dos métodos

careça de aprofundamento, espera-se que o modelo identifique hackers com maior precisão e avalie proativamente a gravidade de suas intenções [Zhang et al. 2022]. A proposta de estudo é um sistema aprimorado que alerte sobre ameaças cibernéticas iminentes, contribuindo para a proteção contra invasões cibernéticas [Hernandez et al. 2016].

Este artigo está estruturado da seguinte forma: a Seção 2 revisa a literatura relacionada; a Seção 3 detalha a metodologia; a Seção 4 aborda a coleta e pré-processamento de dados; a Seção 5 analisa os resultados; e a Seção 6 conclui com as principais descobertas e implicações para futuras pesquisas.

2. Trabalhos relacionados

O hacktivismo é um fenômeno cultural e social complexo, caracterizado por princípios como liberdade de informação, desconfiança da autoridade e defesa da descentralização. Apresenta características de mérito e competição, valorizando a demonstração de habilidades técnicas e sociais [Himanen 2001]. Desafia as noções tradicionais de política digital, combinando elementos de individualismo com experimentos de coletivismo não hierárquico [Coleman 2014]. Alsaffar et al. (2019) avaliam o desempenho de vários algoritmos de aprendizado de máquina e de aprendizado profundo na detecção de spam no "X". Benjamin e Chen (2015) utilizam modelos de linguagem baseados em redes neurais recorrentes (sigla em inglês RNNLMs) para aprender relações semânticas entre termos usados por hackers, sugerindo que esses modelos podem ser ferramentas promissoras na modelagem da linguagem hacker. Khandpur (2018) propõe uma abordagem para detectar ataques cibernéticos usando mídias sociais como fonte de dados, destacando a importância de identificar atividades maliciosas em estágios iniciais e apresentando uma metodologia proativa para identificar ameaças cibernéticas. Le Sceller et al. (2017) introduzem o SONAR para detectar eventos de segurança cibernética em tempo real no "X", enfatizando a importância de monitorar eventos de segurança cibernética de maneira prévia.

A origem das postagens utilizadas não foi abordada em profundidade nestes estudos, focados no conteúdo delas, sem analisar as características de contas de hackers e usuários comuns. Esse estudo busca propor o preenchimento dessa lacuna, ao incorporar metodologia para identificar perfis de hackers e atividades correlatas no Brasil, contribuindo para compreensão mais abrangente das ameaças cibernéticas divulgadas nesta plataforma "X" de mídia social nesse país.

3. Método

Este estudo adota uma abordagem multifásica para melhorar a detecção de atividades de hacker no "X", selecionando-as a partir de informações coletadas do Zone-H e utilizando análise de redes complexas e aprendizado de máquina.

Fase 1 – Coleta inicial dos dados. No repositório Zone-H são extraídas informações de notificadores e registros de *defacement*, a partir da documentação de incidentes correlatos na web, com ênfase no contexto brasileiro. Com base nessas informações, busca-se identificar no "X" e registrar perfis de usuários associados a atividades ilícitas e perfis ativos na comunidade hacker ativa. Os dados das postagens coletados incluem informações tais como data, texto, nome de usuário, interações e conteúdo multimídia. Eles são processados e armazenados em banco de dados SQLite. Tal técnica é amplamente estabelecida e documentada na literatura acadêmica [Cogburn & Espinoza-Vasquez 2011]. Este método permite a obtenção de informações públicas

disponíveis na plataforma “X”, seguindo diretrizes éticas e legais vigentes. Foram coletadas 455.735 postagens no período de 20/10 a 01/12/2023.

Fase 2 – Identificação de atividades de hackers. Técnicas de Processamento de Linguagem Natural (NLP) são usadas para pré-processar o texto contido nas postagens, padronizando o uso de letras minúsculas ou maiúsculas, retirar *stop-words* e identificar menções de um usuário a outro(s), URLs e outros elementos relevantes [Manning et al. 2008]. As postagens pré-processadas são filtradas utilizando 17 *hashtags*, consideradas as principais relacionados a *defacement*, resultando em 23.672 postagens. As *hashtags* utilizadas são: *Leaked, deface, Anonymous, OwneD, hack, deface, breach, cyberattack, nofields, hacked, hacking, defacing, owned, leak,* hackeada*, Leaked, cyberteam, zoneh, BrazilianCyberArmy, InvasãoEspecial* e *Invasão*. Essas postagens são armazenados em banco de dados SQLite. O uso de filtros de *hashtags* é uma abordagem estabelecida na literatura [Morstatter et al. 2013], e permite focar a coleta em tópicos de interesse específico, proporcionando contexto e compreensão mais profunda das atividades relacionadas a esses tópicos.

Fase 3 – Identificação das comunidades. A partir da base de postagens filtrada é realizada uma análise de clusterização. O algoritmo de clusterização utilizado é o k-Means. O número de clusters é determinado com o método do cotovelo, devido à sua comprovada eficácia na determinação do número de clusters [Fortunato, 2010; Rousseeuw 1987]. Para cada um dos clusters são identificados os termos mais frequentes relacionados a *defacement* e realizada uma análise de rede complexa com o objetivo de identificar os principais usuários. As postagens dos usuários que mencionam usuários identificados como potenciais atores de ameaças cibernéticas são coletadas se não o tiverem sido. Métricas de rede como centralidade, proximidade e intermediação são aplicadas com o objetivo de identificar e destacar atores de ameaça. A identificação de usuários-chave desempenha papel crucial na análise de mídias sociais e na compreensão da dinâmica das comunidades online [Wasserman & Faust 1994]. A etapa de análise de rede envolve a aplicação de ferramenta de análise de redes para calcular métricas, tais como proximidade e intermediação, que fornecem insights valiosos sobre a estrutura e a influência dentro da rede [Newman 2010].

Fase 4 – Construção e análise da rede de interações. O cluster mais apropriado (com maior quantidade de usuários e maior frequência de termos relacionados à cíbersegurança) é selecionado para análise detalhada, seguido pela construção de uma rede de interações que conecta os diversos usuários abrangidos por aquele cluster. Chouchani & Abed (2020) apresentam uma revisão comparativa abrangente das abordagens para agrupamento de atores de redes sociais em comunidades de interesse que ajuda a contextualizar a construção de redes de interações. Destaca-se a relevância da segmentação de comunidades e grupos de interesse em ambientes online [Chouchani & Abed 2020]. Bellaby (2021) contribui propondo um framework ético para operações de *hacking*. A construção de redes de interações é baseada em menções entre usuários e desempenha um papel crucial na revelação de padrões e compreensão da estrutura das interações dos usuários envolvidos em discussões relacionadas ao *hacking* e atividades cibernéticas. A construção da rede de interações requer as seguintes tarefas:

i. Coleta de dados de menções. A partir do banco de dados coletado, as menções direcionadas ou recebidas por usuários identificados são extraídas. Inclui postagens que incorporam o símbolo "@" seguido pelo nome de usuário do indivíduo mencionado;

ii. Construção da Rede de Menções. Com os dados de menções acima, procede-se à construção da rede. Nesta rede, os nós representam os usuários do “X”, enquanto as

arestas representam as menções, estabelecendo uma conexão direcional entre o usuário que fez a menção e o usuário mencionado. A direção é importante para entender quem está iniciando a comunicação e quem está recebendo atenção [Maharani et al. 2018].

iii. Análise da Rede de Menções. As análises são conduzidas para identificar usuários com alta centralidade de grau, ou seja, aqueles que são frequentemente mencionados e/ou mencionam muitos outros usuários. Esta identificação pode indicar influência ou importância dentro da rede [Barabási 2016; Freeman 1979]. Além disso, técnicas de detecção de comunidades podem ser aplicadas para identificar grupos densamente conectados dentro da rede, que podem representar grupos de hackers ou colaboradores [Claustet et al. 2004; Fortunato 2010]. As representações gráficas ajudam a identificar padrões visuais proeminentes, grupos e usuários de forma intuitiva [Hansen et al. 2011]. Esta metodologia baseada na análise de redes sociais encontra respaldo na literatura acadêmica, onde estudos anteriores reconheceram a utilidade da análise de redes para entender comunidades online [Knoke & Yang 2008; Scott 2017].

4. Implementação e Avaliação

Essa seção apresenta detalhes sobre como a metodologia foi implementada e testada, a discussão dos desafios enfrentados durante a implementação e como foram superados.

4.1 Coleta de Dados do Zone-H

Os dados foram coletados do site Zone-H, conhecido por registrar atividades de *defacement*. Os dados extraídos incluíam informações como data e hora, notificador, tipos de *defacement*, domínio, sistema operacional e URL do *defacer*. Eles foram armazenados em um data frame para análise, conforme ilustrado na Fig. 1.

	Time	Notifier	H	M	R	L	Special	Domain	OS	View URL
0	23:41:08	Gab	1	0	0	Brazil	0	esthosting.com.br	Linux	/mirror/id/40979686
1	00:23:56	Clash Hackers	0	0	0	Brazil	0	silvam.pompeumg.com.br/cl.html	Linux	/mirror/id/40979416
2	23:02:40	Clash Hackers	0	0	0	Brazil	0	rdev.epimaringa.com.br/cl.html	Linux	/mirror/id/40979370
3	22:07:39	Clash Hackers	0	0	0	Brazil	0	teste.epimaringa.com.br/cl.html	Linux	/mirror/id/40979350
4	09:01:53	Plastyne	0	0	0	Brazil	0	palinialves.com.br/play.txt	Linux	/mirror/id/40974065
...	...	...	...	...	...	...	...	...	...	...
995	2023/05/22	VandaTheGod	1	0	1	Brazil	0	diaseproenca.com.br	Linux	/mirror/id/40609325
996	2023/05/21	B1G0D1N	1	0	0	Brazil	0	tatianacapanema.com.br	Linux	/mirror/id/40608842
997	2023/05/14	diparis	0	0	0	Brazil	0	glpi.cenciseg.com.br/glpi/	Linux	/mirror/id/40591452
998	2023/05/13	Rxc404	1	0	0	Brazil	0	cetri.com.br	Linux	/mirror/id/40591087
999	2023/05/13	Rxc404	1	0	0	Brazil	0	explosaodeleads.com.br	Linux	/mirror/id/40591085

1000 rows × 10 columns

Fig. 1 - Data frame do Zone-H

4.2 Extração de Notificações e Busca Correspondente no “X”

A partir dos dados do Zone-H coletados, notificadores foram extraídos (Fig. 2) e usados como base para buscar postagens ou nomes de usuários relacionados no “X”. As postagens que apresentam semelhanças com o nome do notificador são baixadas e armazenadas no banco para filtragem.

	Notifier	Counts
0	VandaTheGod	63
1	ProtoWave Reloaded	40
2	Junin-CLS	32
3	B1G0D1N	30
4	Tux Society	29
...	...	...
195	B0yzTeam	2
196	terezinha security	2
197	rtax	2
198	Finistro	2
199	abeille23	2

200 rows × 2 columns

```
hashtags
#CyberAttack                4214
#cyberattack                3685
#cybersecurity              3552
#Hacked                     3296
#Anonymous                  2698
                            ...
#OSINT                       113
#aleistercrowley             113
#CyberSecurityAwareness      113
#cloudsecurity               112
#Israeli                     111
Name: count, Length: 200, dtype: int64
```

Fig. 2 - Notificadores e número de *defaces* e *Hashtags* frequentes.

4.3 Coleta de Dados do "X"

A partir da coleta de dados no "X", postagens relevantes foram identificadas e extraídos dados como data, texto, nome de usuário, nome exibido, número de comentários, repostagens, curtidas, links e outros metadados relevantes.

4.4 Filtragem dos Dados do "X" e Coleta de *Hashtags* Comuns

Os dados foram filtrados para identificar hashtags comuns e palavras-chave associadas ao *hacking* e atividades cibernéticas. Uma nova coleta de dados foi realizada com base nessas *hashtags*, permitindo expandir o escopo da busca já realizada (Fig. 3).

	text	hashtags
0	Join us! #OpNewBlood #Anonymous #ExpectUs	[#OpNewBlood, #Anonymous, #ExpectUs]
1	Ghosts of Palestine targeted major websites of...	[#Cti, #Threatintel, #Israel]
2	Ghosts of Palestine is targeting Rafael's Iron...	[#Threatintel, #Israel]
3	Several #Zionist websites were taken offline b...	[#Zionist, #OpIsrael, #OpIsraelv2, #AlaqsaStorm]
4	Malek Team Hacked Ono Academic College! Hebre...	[]
...	...	...
95	Do you remember when you joined X? I do! #MyXA...	[#MyXAnniversary]
96	Israeli government educational portal has been...	[#TangoDown, #OpIsrahell]
97	Ministry Of Industries: https:///industry./go...	[#MTB]
98	Sri Lanka Government websites dropped. Nation...	[#MTB]
99	Websites belonging to the Israeli government #...	[#TangoDown, #OpIsrahell]

100 rows × 2 columns

Fig. 3 - Extração de *hashtags* e filtro com *hashtags* relevantes.

4.5 Identificação de Usuários Relevantes e Coleta de Postagens

Usando os dados coletados, uma análise foi realizada para identificar os principais usuários mencionados em relação às 21 *hashtags* predeterminadas. Posteriormente, as postagens desses usuários foram coletadas para análise adicional, particularmente aquelas que mostraram uma alta incidência de termos relacionados ao *hacking*.

4.6. Análise de Clusters

Após o pré-processamento das postagens, os dados foram clusterizados com o algoritmo K-Means para auxiliar a identificar padrões e agrupamentos. O número de clusters foi determinado com a aplicação dos métodos cotovelo (Fig. 4) e *Silhouette* (Fig. 5).

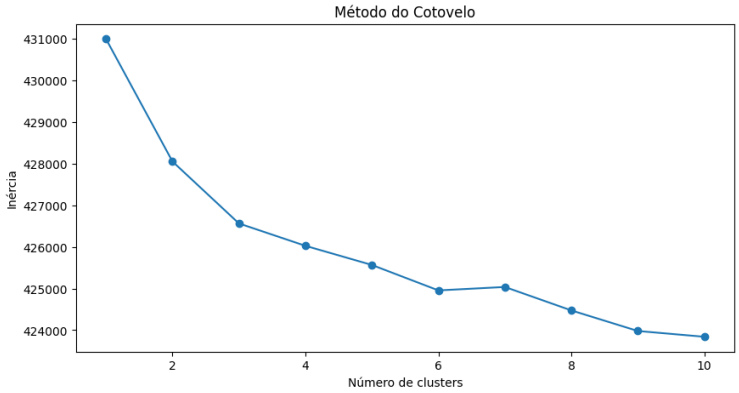

Fig. 4 - Gráfico de variância intra-cluster em relação ao número de clusters.

O valor do *Silhouette Score* varia de -1 a 1. Um valor alto indica que o objeto está bem combinado com seu próprio cluster e mal combinado com clusters vizinhos. Se a maioria dos objetos tiver um valor alto, então a configuração dos clusters é apropriada. Se muitos pontos mostrarem um valor baixo ou negativo, a configuração dos clusters pode ter espaço para melhoria. Foram realizados três testes com 3, 4 e 5 clusters, como mostrado na Fig. 5, onde o *Silhouette Score* para 3 clusters é maior que para as demais configurações de clusters.

```
Para n_clusters = 3, o silhouette score médio é: 0.010558376277538322

Para n_clusters = 4, o silhouette score médio é: 0.010416846371843945

Para n_clusters = 7, o silhouette score médio é: 0.0352063713265274
```

Fig. 5 - Silhouette Score.

O valor mais alto de *Silhouette Score* médio entre essas configurações é para 3 clusters, sugerindo que a configuração com 3 clusters é a mais apropriada. Contudo, foi escolhida a configuração de 4 clusters, pois, em uma análise manual, essa configuração mostrou uma melhor distribuição de grupos hackers. Essa decisão foi baseada na observação qualitativa dos dados, onde a configuração de 4 clusters proporcionou uma separação mais intuitiva e relevante dos diferentes grupos de hackers, permitindo uma melhor interpretação e análise das atividades desses grupos, essa escolha de 4 clusters facilitou a identificação de padrões e comportamentos específicos dentro de cada grupo, o que é crucial para a análise de segurança cibernética.

4.7. Análise Temporal, Identificação de Termos Comuns e Sentimento de Cluster

Foi realizada uma análise dos termos das postagens em cada cluster com o objetivo de identificar tendências e padrões temporais relevantes ao longo do período de estudo. Observa-se flutuações no volume de postagens, o que é essencial para entender o impacto de eventos ou atividades específicos em determinados períodos [Lin et al. 2009]. Os termos relacionados à cibersegurança mais comuns nos clusters foram *cyber*, *attack*, *website*, *russian*, *cybersecurity*, *ciberattack*, *israel*, *security* e *family* (Fig. 6). Isso sugere uma concentração significativa de discussões sobre questões de cibersegurança e possíveis ataques cibernéticos, refletindo os interesses predominantes na comunidade de hackers e hacktivistas. A análise da distribuição de postagens por cluster revela que certos clusters têm um número significativamente maior de postagens, indicando tópicos de maior interesse durante o período de estudo.

Finalmente, a avaliação do sentimento médio das postagens de cluster revelou nuances sobre o tom emocional das conversas. Por exemplo, os clusters 1 e 0 mostram sentimentos médios de 0.03 e 0.04, respectivamente. Embora esses valores sejam relativamente neutros, é importante notar que qualquer desvio significativo de 0 pode indicar uma tendência emocional nas discussões do cluster. Está métrica, além de ser uma ferramenta crucial para perceber o ambiente emocional predominante, pode ajudar na categorização e identificação dos temas que geram respostas mais positivas ou negativas na comunidade [Hernandez-Suarez et al. 2018; Hernandez et al. 2016].

4.8 Construção e Análise das Redes de Menções do Cluster

Nesta etapa do estudo, a análise foca na construção e análise de redes de menções, com foco nas interações de usuário para usuário. Este processo é crucial para entender a estrutura e os padrões de comunicação entre os participantes nas discussões relacionadas ao hacktivismo e atividades cibernéticas. Para melhor processamento, os usuários mais ativos de cada cluster foram selecionados e extraíram-se os usuários mais mencionados por eles. O objetivo é identificar possíveis canais de disseminação de informações relevantes para esses contextos. A título de exemplo, a Fig. 7 apresenta os usuários mais ativos no cluster 1 e a Fig. 8 os usuários mais mencionados pelos usuários mais ativos no cluster 1.

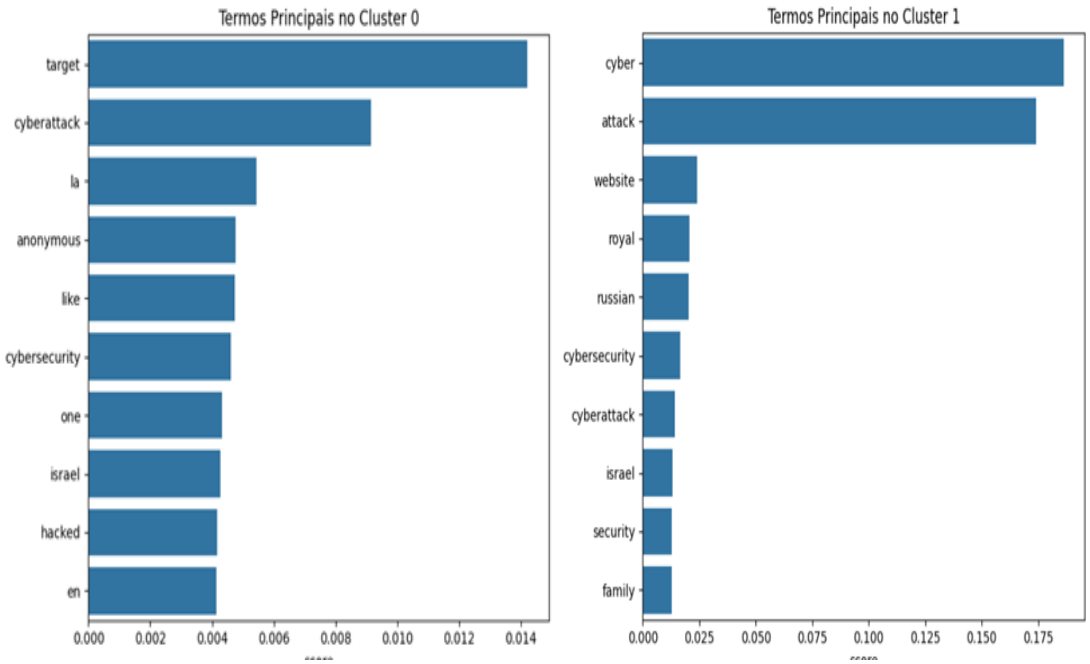

Fig. 6 - Gráfico de distribuição de postagens dos clusters 0 e 1

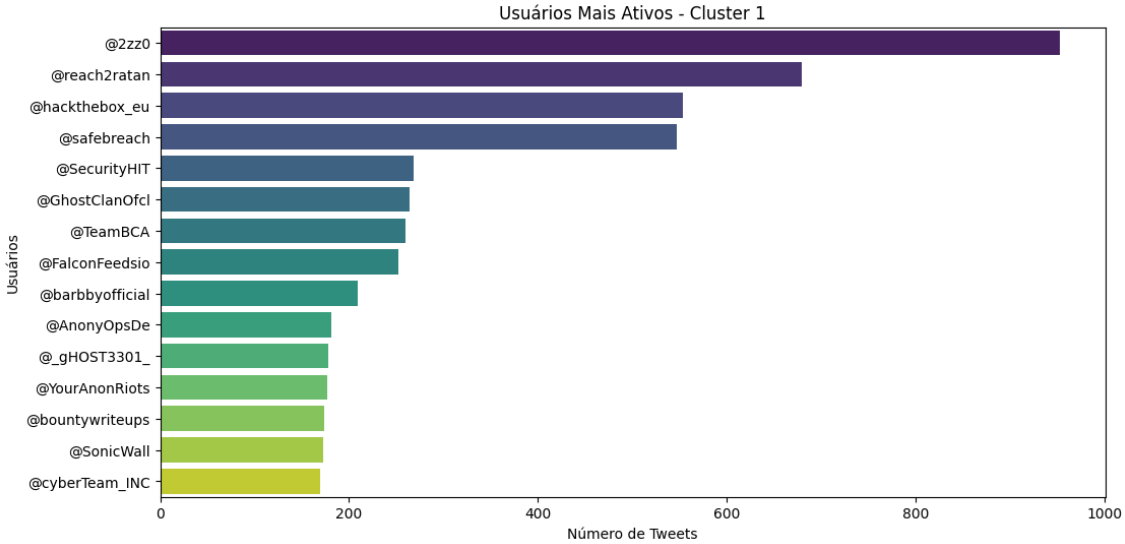

Fig. 7 - Usuários mais ativos por número de postagens no cluster 1

```
========================================
Cluster 1:
@BleepinComputer: 47
@MysteriousTeam0: 44
@TheHackersNews: 43
@Land2Cyber: 34
@MichelleRagusa: 29
@_barbby: 25
@OstermanRsch: 24
@MysteriousT34m0: 21
@DarkReading: 20
@aavivi: 19
@AWS: 18
@defcon: 17
@LulzSecSL: 16
@KamikazeJapan5: 16
@YourAnonNews: 16
@Sprek3rsSec: 14
@anonbarbby: 14
@Hornetsecurity: 14
@TweetBrookcourt: 14
@ArmisSecurity: 13
----------------------------------------
Menções em string para o Cluster 1: @reach2ratan, @BleepinComputer,
========================================
```

Fig. 8 - Usuários mais mencionados pelos usuários mais ativos no Cluster 1

Durante esta análise, foram identificados atores de ameaça ativos que não haviam sido coletados inicialmente (Fig. 9). São usuários que tinham sido coletados na Fase 1 mas que foram excluídos na Fase 2 ou usuários antigos, com alta probabilidade de serem hacktivistas, mencionados em postagens recentes, mas que não postaram durante o período de coleta inicial das postagens, não foram descobertos novos usuários a partir do cluster 3.

```
Cluster 0 - Usuários não existente na coleta inicial (Primeiros 40):
EPSLinares, 0xWORD, CyberHunterSec, 0xWord, rootedcon, DrGiammattei, Singularity_Ex, FundacionINCYDE, hack, gmail
NavajaNegra_AB, eldpit, geeks_academy, zendalibros, ivoox, Gwalrock, inetum_es, AlightSolutions, MiolnirST, WatchGuardSpain
Women4Cyber_SP, 123emprende, AntonioCortesB, _CARITAS, chema_garabito, ssantosv, perezreverte, C1b3rWall, Spreaker, elpais_tec
HazzimIO, MPguatemala, GuatemalaGob, PlexusTech_, oricio_org, fundacionfulgenciomeseguer, martrudix, mikiminoru, TwitchES, c1b3rwall
----------------------------------------
Cluster 1 - Usuários não existente na coleta inicial (Primeiros 40):
Land2Cyber, aavivi, anonbarbby, KamikazeJapan5, cybersaiyanIT, darkstar7471, RoadRunnerHacks, Sprek3rsSec, hacktivistlink, JTSEC13
hpylarinos, _leHACK_, itzikkotler, AWS, zekeriufunet, GITEX_GLOBAL, CertBros, RecordedFuture, Jenny_Radcliffe, Microsoft
TAG_Cyber, securelink, GhostCodin, AnonDragonNeb, MysteriousTeam0, _TheGhostSquad, dilagrafie_, Google, szymex73, 21y4d
mrb3n813, HagueSabastian, snyksec, Infosecurity, TheCyberGeek19, idekCTF, _kavigihan, EuroInformation, KatzcyPlayCyber, Cero_On3
----------------------------------------
Cluster 2 - Usuários não existente na coleta inicial (Primeiros 40):
eric_jeanjean, RollingStones, officialKeef, MysteriousTeamO, AfricaCyberMag, hackyourjob, OVHcloud_FR, ovh_support_fr, icann_fr, ICANN
----------------------------------------
Cluster 3 - Usuários não existente na coleta inicial (Primeiros 40):

----------------------------------------
```

Fig. 9 - Usuários ativos coletados após análise da rede de menções

Finalmente, são geradas redes de menções a partir das postagens no “X” contidas em cada um dos clusters. A estrutura de cada rede de menções sugere a possibilidade da existência de diferentes subgrupos ou de tópicos sendo discutidos, com alguns indivíduos atuando como pontes entre diferentes subcomunidades. A identificação de atores-chave e a análise de redes são componentes essenciais na análise de ameaças cibernéticas [Romagna 2020]. A importância de identificar atores de ameaça em atividades hacktivistas também é destacada na literatura [Benjamin & Chen 2012]. A Fig. 10 apresenta a rede de menções obtida para o cluster 1 sendo o grafo gerado a partir dos dados das postagens devida às interações dos usuários. Observa-se uma densidade significativa de conexões, indicando um alto grau de interação entre os usuários, bem como a presença de um núcleo central com vários nós altamente interconectados, sugerindo a existência de indivíduos ou entidades influentes dentro do grupo. Esta estrutura densa pode indicar um grupo de interesse, onde os membros compartilham informações e interagem frequentemente.

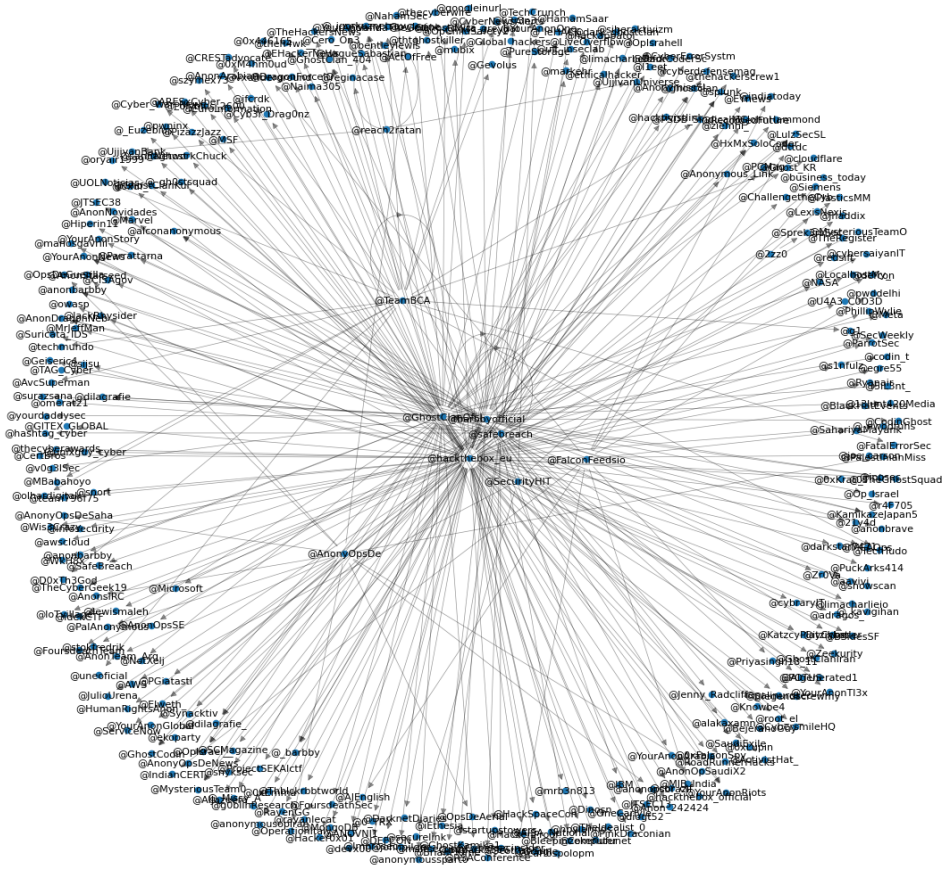

Fig. 10 - Gráfico da rede de interações entre usuários no Cluster 1

6. Conclusões

Neste estudo, a intricada rede de hacktivismo e atividades cibernéticas no “X” foi investigada com o emprego de análise de redes complexas e aprendizado de máquina, que permitiu uma incursão nas camadas prevalecentes da comunicação online entre indivíduos envolvidos em hacking, evidenciando uma complexa rede de interações, influências e intenções.

A segmentação dos dados revelada pela análise de clusters, especialmente em nuvens de palavras e distribuição de sentimentos, não apenas destacou a frequência da

linguagem hacker, mas também expôs nuances significativas nas comunicações dentro desta comunidade. A coexistência de sentimentos tanto negativos quanto positivos sugere uma comunidade heterogênea e multidimensional, a ser pesquisada e detalhada no futuro. Outra evidência foi a visualização da rede de menções, que destacou a vasta e densa rede de indivíduos envolvidos em práticas de hacking, com padrões de interconexão que refletem um robusto ecossistema de colaboração e compartilhamento de informações.

O presente estudo foi eficaz em encontrar novos atores de ameaça a partir dos dados analisados e mostrou a importância crucial da coleta contínua de novos usuários potencialmente maliciosos antes que o perfil do usuário seja apagado na rede social. Para trabalho futuro é pertinente desenvolver uma rotina de coleta com palavras-chave dinâmicas criadas a partir da análise de atores de ameaça já identificados em coleções anteriores, bem como a coleta na rede social de novos usuários detectados como potenciais ameaças e sua análise.

Referências

A.-L. Barabási, Network Science, Cambridge University Press, 2016.

V. Benjamin and H. Chen, “Developing understanding of hacker language through the use of lexical semantics,” In IEEE International Conference on Intelligence and Security Informatics (ISI), pp. 79–84, 2015.

V. Benjamin and H. Chen, “Securing cyberspace: Identifying key actors in hacker communities,” 2012 IEEE International Conference on Intelligence and Security Informatics, Washington, DC, USA, 2012, pp. 24-29, doi: 10.1109/ISI.2012.6283296.

J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” Journal of Computational Science, vol. 2, no. 1, pp. 1-8, 2011.

V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM Computing Surveys (CSUR), vol. 41, no. 3, pp. 1-58, 2009.

N. Chouchani and M. Abed, “Online social network analysis: Detection of communities of interest,” Social Network Analysis and Mining, vol. 10, no. 1, pp. 1-19, 2020.

A. Clauset, M. E. Newman, and C. Moore, “Finding community structure in very large networks,” Physical Review E, vol. 70, no. 6, 066111, 2004.

D. L. Cogburn and F. K. Espinoza-Vasquez, “From Networked Nominee to Networked Nation: Examining the Impact of Web 2.0 and Social Media on Political Participation and Civic Engagement in the 2008 Obama Campaign,” Journal of Political Marketing, vol. 10, nos. 1-2, pp. 189-213, 2011. Available at SSRN: https://ssrn.com/abstract=2854273

G. Coleman, Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous, Verso Books, 2014.

D. Alsaffar, A. Alfahhad, B. Alqhtani, L. Alamri, S. Alansari, N. Alqahtani, and D. A. Alboaneen, “Machine and deep learning algorithms for Twitter spam detection,” In International Conference on Advanced Intelligent Systems and Informatics, Springer, Cham, pp. 483–491, 2019.

D. Grewal, D. Herhausen, S. Ludwig, and F. Villarroel Ordenes, “The future of digital communication research: considering dynamics and multimodality,” Journal of Retailing, vol. 98, no. 2, pp. 224-240, 2022.

H. L. Gururaj, U. Tanuja, V. Janhavi, and B. Ramesh, “Detecting malicious users in the social networks using machine learning approach,” International Journal of Social Computing and Cyber-Physical Systems, vol. 2, no. 3, pp. 229-243, 2021.

D. Hansen, B. Shneiderman, and M. A. Smith, Analyzing Social Media Networks with NodeXL: Insights from a Connected World, Morgan Kaufmann, 2011.

T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2009.

P. Himanen, The Hacker Ethic and the Spirit of the Information Age, Random House, 2001.

R. P. Khandpur, “Augmenting Dynamic Query Expansion in Microblog Texts,” 2018.

R. P. Khandpur, T. Ji, S. Jan, G. Wang, C.-T. Lu, and N. Ramakrishnan, “Crowdsourcing cybersecurity: Cyber attack detection using social media,” In Association for Computing Machinery Conference on Information and Knowledge Management (ACM), pp. 1049–1057, 2017.

D. Knoke and S. Yang, Social Network Analysis (2nd ed.), SAGE Publications, 2008.

Q. Le Sceller, E. B. Karbab, M. Debbabi, and F. Iqbal, “Sonar: Automatic detection of cyber security events over the Twitter stream,” In 12th International Conference on Availability, Reliability and Security (ACM), pp. 23–34, 2017.

C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, Cambridge University Press, 2008.

F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley, “Is the sample good enough? Comparing data from Twitter’s streaming API with Twitter’s firehose,” In Seventh international conference on weblogs and social media (ICWSM 2013), pp. 400-408, 2013.

M. E. J. Newman, Networks: An Introduction, Oxford University Press, 2010.

L. C. Freeman, “Centrality in social networks: Conceptual clarification,” Social Networks, vol. 1, no. 3, pp. 215-239, 1979.

F. Santa, R. Henriques, J. Torres-Sospedra, and E. Pebesma, “A Statistical Approach for Studying the Spatio-Temporal Distribution of Geolocated Tweets in Urban Environments,” Sustainability, vol. 11, no. 3, 595, 2019.

J. Scott, Social Network Analysis, SAGE Publications, 2017.

M. Romagna, “Hacktivism: Conceptualization, techniques, and historical view,” 2020.

S. Fortunato, “Community detection in graphs,” Physics Reports, vol. 486, nos. 3-5, pp. 75-174, 2010.

P. Gonçalves et al., “Machine Learning for Hacker Detection: A Comprehensive Review,” International Journal of Machine Learning and Cybersecurity, vol. 8, no. 1, pp. 32-46, 2020.

A. Hernandez-Suarez, G. Sanchez-Perez, K. Toscano-Medina, V. Martinez-Hernandez, H. Perez-Meana, J. Olivares-Mercado, and V. Sanchez, “Social sentiment sensor in Twitter for predicting cyber-attacks using regularization,” Sensors Journal, vol. 18, no. 5, pp. 1–17, 2018.

A. Hernandez, V. Sanchez, G. Sanchez, H. Perez, J. Olivares, K. Toscano, and V. Martinez, “Security attack prediction based on user sentiment analysis of Twitter data,” In IEEE International Conference on Industrial Technology (ICIT), pp. 610–617, 2016.

P. J. Rousseeuw, “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis,” Journal of Computational and Applied Mathematics, vol. 20, pp. 53-65, 1987.

A. P. Rodrigues, R. Fernandes, A. Bhandary, A. C. Shenoy, A. Shetty, and M. Anisha, “Real-Time Twitter Trend Analysis Using Big Data Analytics and Machine Learning Techniques,” Wireless Communications and Mobile Computing, 2021, Article ID 39203252.

B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Foundations and Trends in Information Retrieval, vol. 2, nos. 1–2, pp. 1-135, 2008.

Z. Zhang, H. Ning, F. Shi et al., “Artificial intelligence in cyber security: research advances, challenges, and opportunities,” Artif Intell Rev, vol. 55, pp. 1029–1053, 2022. https://doi.org/10.1007/s10462-021-09976-0

Y.-R. Lin, H. Sundaram, M. De Choudhury, and A. Kelliher, “Temporal patterns in social media streams: Theme discovery and evolution using joint analysis of content and context,” 2009 IEEE International Conference on Multimedia and Expo, New York, NY, USA, pp. 1456-1459, doi: 10.1109/ICME.2009.5202777, 2009.

S. Wasserman and K. Faust, Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press, 1994.

D. L. Alsaffar, A. Alfahhad, B. Alqhtani, L. Alamri, S. Alansari, N. Alqahtani, and D. A. Alboaneen, “Machine and deep learning algorithms for Twitter spam detection,” In International Conference on Advanced Intelligent Systems and Informatics, Springer, Cham, pp. 483–491, 2019.

Zone-H, "Zone-H – Unrestricted information," disponível em: http://www.zone-h.org/, acessado em 20 outubro 2023