Modern Approach to Attributing Hacktivist Groups

research.checkpoint.com/2025/modern-approach-to-attributing-hacktivist-groups/

February 27, 2025

Research by: Itay Cohen (@megabeets_)

Over the past few decades, hacktivism has been, in a lot of cases, characterized by minor website defacements and distributed denial-of-service (DDoS) attacks, which, while making headlines, had minimal lasting impact. However, in recent years, we have observed a significant shift in the nature of these activities. Groups that appear to be state-sponsored, yet masquerade as hacktivists, are now conducting large-scale cyber and influence operations. These actors often use various group names and grassroots facades to maintain anonymity and reduce international backlash. These sophisticated attacks serve as powerful tools for political and social influence, providing plausible deniability and the illusion of legitimacy, thereby evading direct government attribution.

This research explores whether such operations' growing popularity and wide and repeating adoption might also be their Achilles' heel. **We introduce a new approach to attributing hacktivist groups by employing language-based machine-learning models and linguistic analysis.** Our study examines thousands of public messages from dozens of hacktivist groups, integrating traditional cyber threat intelligence methods with modern machine learning technologies. **This combined approach aims to uncover the key topics discussed by these groups, understand their motivations over time, tie some of these groups together, and improve the way we do attribution when it comes to hacktivism.**

Introduction

Hacktivism, a blend of hacking and activism, has evolved dramatically since its inception. Emerging in the 1980s with groups like the Cult of the Dead Cow, which pioneered the concept of hacktivism, the phenomenon has grown from a fringe subculture into a significant force in global cyber politics. Early hacktivists targeted symbolic institutions and networks, aiming to draw attention to various social and political issues. The 1990s and 2000s saw the rise of more organized groups like Anonymous, whose operations, from the 2008 Scientology attacks to the 2010 PayPal protests, showcased the potential of digital activism. Today, hacktivism is not just about defacing websites or DDoS attacks—it has become a sophisticated tool in the arsenal of state and non-state actors alike, influencing geopolitical landscapes worldwide.

One of the most significant developments in hacktivism in recent years is the involvement of nation-state actors. Governments have recognized the strategic value of hacktivist methods, adopting and adapting them for their own purposes. Nations employ cyber operatives to

conduct operations that mimic grassroots hacktivism. These activities often aim to sow discord, influence public opinion, or undermine political adversaries while maintaining plausible deniability. By cloaking their actions in the guise of independent hacktivism, governments can achieve their goals without direct attribution, complicating international responses and accountability.

State-sponsored entities often stage operations to appear as though they are the work of independent hacktivists, thereby muddling the waters of attribution. They often operate multiple, sometimes dozens, of groups that appear to act independently. This strategy not only serves to mislead targets but also to discredit genuine activist movements by associating them with state agendas. Such deceptive tactics complicate the global cyber landscape, making it challenging to distinguish between actual hacktivist actions and state-sponsored subterfuge.

In recent years, ever since this shift in the hacktivism landscape started, Check Point Research has been tracking dozens of hacktivism groups — some are operated by groups of hackers with mutual ideology or interest, and others by nation-state actors masquerade as hacktivists. Our investigations have led to several published articles detailing significant operations we believe were conducted by these state actors. In some cases, we have successfully linked different hacktivist groups to a single operator, piercing their veil of anonymity. Our research has employed traditional threat intelligence and investigative methods, such as binary analysis, code similarity, and the Diamond Model of Intrusion Analysis. For instance, we demonstrated how one hacktivist group used the same unique tools and methods as another group targeting different entities years earlier, suggesting they were the same group. However, since then, nation-state hacktivist groups have refined their techniques, trying even harder to obscure their tracks and act independently.

Tracking these groups has provided us with valuable insights into their motives, attack tactics, and usual targets. Our researchers have developed various theories and hypotheses regarding the operators behind these groups, suspecting that many are run, either directly or indirectly, by the intelligence agencies of different nation-states. The recent major geopolitical events, such as the Russian invasion of Ukraine and the conflict between Israel and Hamas, have highlighted shifts that would have otherwise gone unnoticed. In response to these events, hacking groups that had been inactive for years suddenly resurfaced, often with new focus areas. Independent groups began to echo similar messages, sometimes even referencing one another. These conflicts created a fertile ground for the emergence of new groups and cyber personas at an astonishing pace.

The increased frequency of hacktivist operations and the rise of new groups have added complexity to the cyber threat landscape. The rapid formation and dissolution of these groups, coupled with their ability to adapt quickly to geopolitical changes, make it challenging

for threat researchers to manually keep track of them. Recognizing that relying on gut feeling and intuition is not a solid approach to threat intelligence, we decided to explore more reliable, robust, and consistent methods.

Methodology

In order to address our research questions and tackle the complexities we presented, we adopted a combined approach utilizing Topic Modeling and Stylometric analysis. This decision was driven by the need for better methods to analyze the vast amount of textual data written by hacktivist groups. By integrating language-based machine learning models and linguistic analysis, we aim to systematically uncover the motivations behind hacktivist activities, identify connections between different groups, and enhance the attribution of these operations to specific actors.

Topic Modeling helps us understand the key themes and topics discussed by these groups over time. It allows us to answer questions such as: What are the main goals and targets of these hacktivist groups? How do their focus areas change in response to geopolitical events? By analyzing the content of the groups' social media messages, we can identify clusters of topics and track their evolution, providing insights into the strategic objectives of these actors.

Stylometric analysis, on the other hand, focuses on the unique writing styles of different groups. By examining features like word choice, sentence structure, and other linguistic patterns, we can answer questions such as: Can we tie different hacktivist groups together based on their writing style? Are there groups with similar textual fingerprints that might indicate common authorship or collaboration? This analysis helps us build a network of connections between different groups, offering a deeper understanding of their relationships and potentially revealing the operators behind these cyber personas.

Data Collection

To conduct this research, we collected data from the social media accounts and messages of hacktivist groups believed to be operated by nation-state actors. The primary sources of data were Twitter and Telegram, as these platforms are commonly used by hacktivist groups to disseminate their messages. The collection process involved several steps:

1. **Identification of Target Accounts**: Although there are hundreds of social media accounts associated with cyber-personas of hacktivist groups, many of these accounts are either inactive or not backed by nation-state actors. We chose 35 hacktivist accounts that are active and potentially state-sponsored. These accounts were selected based on their activity, the nature of their messages, and previous reports linking them to nation-state actors. They represent a diverse range of targets and objectives, posting in languages such as English, Russian, Ukrainian, Persian, Hebrew, Hindi, and more. Some of these accounts are active on both Telegram and Twitter, often managing multiple accounts, including backup accounts. For each group, we noted all directly associated accounts.
2. **Data Extraction:** To gather all messages from X, we utilized the official API to retrieve tweets and replies from these accounts. Although retweets might also indicate the topics of interest to these accounts, we chose not to include them due to their lower reliability compared to original tweets and replies. For Telegram, we used the Telegram Desktop application to export all messages published by channels operated by these hacktivist groups. In total, we collected approximately 20,000 tweets and messages from these accounts.
3. **Data Cleaning**: The raw data extracted from social media platforms included various elements such as URLs, hashtags, emojis, and special characters. For the topic modeling algorithm, we cleaned the data by removing these elements to focus on the textual content. Additionally, non-English messages were translated into English to maintain consistency. However, for the stylometric analysis, we retained these elements. Since our stylometric models support multiple languages, we preserved the original language and did not translate these messages.
4. **Data Storage:** The cleaned data was stored in a structured format, making it suitable for further analysis. Each message was tagged with metadata, including the date of publication, the originating account, and the code of the original language.

Having all the data structured and ready for analysis we moved on to the analysis.

Topic Modeling

Topic modeling is a machine-learning technique used to identify themes (topics) within large collections of textual data. In this research, we employed BERTopic, a modular and flexible framework for topic modeling. The process involved the following steps:

1. **Embedding Documents**: The first step is to convert our textual data into numerical representations, which can be processed by machine learning algorithms. Using BERTopic, we use sentence-transformer models optimized for semantic similarity tasks. These models are adept at creating document embeddings that capture the semantic nuances of the text, making them ideal for our clustering tasks. In simpler terms, we transform the text into numbers that reflect their meanings, allowing us to find patterns and similarities.

2. **Dimensionality Reduction**: High-dimensional data can be challenging for clustering algorithms, so we used UMAP (Uniform Manifold Approximation and Projection) to reduce the dimensionality of the embeddings while preserving their essential structure. This step ensures that the clustering process can effectively group semantically similar messages, maintaining both local and global relationships within the data. Essentially, we simplify the data while keeping its important relationships intact, making it easier to analyze.
3. **Clustering Documents**: With reduced-dimensionality embeddings, we applied Agglomerative Clustering and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) to identify clusters within the data. HDBSCAN is particularly suited to our needs because it can find clusters of varying shapes and densities while identifying outliers. This feature helps in improving the quality of the resulting topic representations by reducing noise. In other words, we group similar messages together while filtering out irrelevant ones.
4. **Bag-of-Words Representation**: To create meaningful topic representations, we combined all documents within each cluster into a single document and counted the frequency of each word. This process, known as the bag-of-words representation, helps in identifying the most significant words within each cluster without making assumptions about the cluster structure. Simply put, we focus on the most common words in each group to understand what each topic is about.
5. **Topic Representation**: We then used a modified version of the traditional TF-IDF (Term Frequency-Inverse Document Frequency) approach to work at the cluster level rather than the document level. This modification, called class-based TF-IDF (c-TF-IDF), calculates the importance of words within each bag of words by comparing their frequency within the cluster to their frequency across all clusters. This method allows us to extract the most representative words for each topic, providing clear and meaningful topic descriptions. In essence, we determine the key words that define each topic by comparing their importance within and across groups.

	# Docs	Representation	Documents
Topic 1	**2775**	**attack – russian – cyber – ukrainian – hack**	**[“we announce for all russian hackers, we begin the cyber attack on ukrainian websites …”]**
Topic 2	**1205**	**israel – cyber – leak – zionist – gaza**	**[“#IsraelLeaks – Day Four – Leak One – Israel Innovation Authority – #OpCyberToufan #OpIsrael, …”]**
Topic 3	**978**	**confidential – letter – leak – regime – iran**	**[“We expose a very confidential letter from the cultural and social vice president, …”]**
Topic 4	**490**	**cyber – operation – russia – slava – ukraine**	**[“today we will cut the tongue of russian propaganda and carry out an do attack on rt (tv “russia today”) …”]**
…			
Topic N	**33**	**italy – european – attack – farmers – cyber**	**[“and we continue our joint attack on italy and with the help of our friends from, …”]**
Topic -1	**283**	**NaN**	**[“because you are a dwarf or a small child, you certainly do not know him”, …]**

Figure 1: Table of the topics after the Topic Representation stage

6. **Fine-tuning Topic Representation**: To further refine the topic representations, we split, merged, and dropped some topics. This step ensures that the final set of topics is both comprehensive and precise.
7. **Assigning Topic Names Using LLMs**: To give the topics proper names, we used Large Language Models (LLMs), specifically GPT. We provided the model with a list of keywords and representative documents for each topic and asked it to generate concise and descriptive labels for the topic. This step helps us understand what each topic is about at a glance.

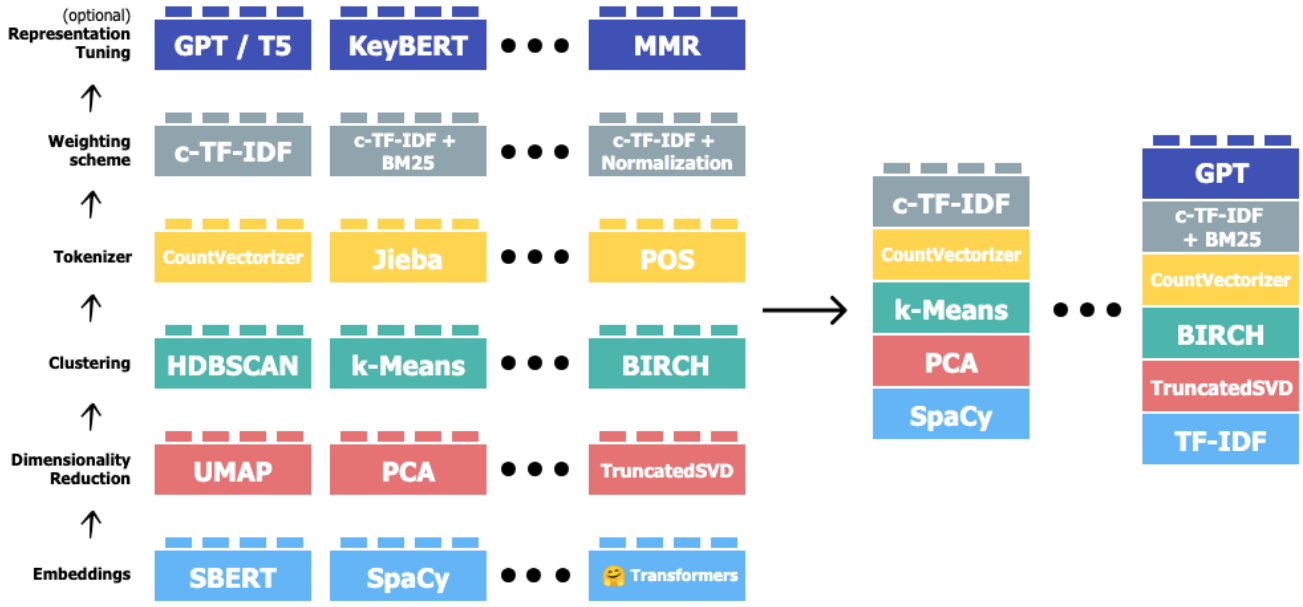

Figure 2*:* The topic-modeling process step by step. BERTopic is modular and flexible.

8. **Visualizing Topics:** We visualized the distribution of topics across different hacktivist accounts and the prevalence of each topic within the accounts. This involved creating charts that show which accounts are discussing each topic and the proportion of messages each account dedicates to different topics. Visualizing topics in this manner helps us identify which groups are focused on particular themes and how their interests align or diverge. For instance, we can see if multiple groups are simultaneously discussing a significant geopolitical event. These visualizations also allow us to track changes in focus over time, providing insights into how hacktivist strategies evolve in response to global events.

By employing topic modeling, we can inspect the topics and themes within vast amounts of textual data — the messages written by the different profiles. It allows us to gain insights into the key topics and trends discussed by these hacktivist groups, track changes in their focus areas over time, and understand their strategic objectives. For example, we can identify how the focus of certain groups shifts in response to geopolitical events, such as an uptick in messages related to cyber attacks following a major conflict. Additionally, topic modeling can reveal unexpected connections between groups discussing similar topics, suggesting potential coordination or shared motivations. Since these groups were most likely created by intelligence agencies, understanding the topics discussed by each of them can help us understand what mission or goal these accounts are serving and what purpose they were created for. Visualizing the topics over time allows us to detect shifts in focus or interest by the entities operating these accounts. Combining these insights together, we can have a better understanding of who might be behind these groups, based on the interest and the change of interest of certain groups.

Stylometry

Stylometry is the study of linguistic style, often used for authorship attribution. It leverages the unique way individuals write, akin to a linguistic fingerprint, to identify authors or establish connections between texts. In this research, we utilized stylometric analysis to compare the writing styles of different hacktivist groups. This method provided us with potential relationships and commonalities between these groups. The process involved the following steps:

1. **Feature Extraction:** We used the StyloMetrix framework to extract nearly 200 features from the collected messages. StyloMetrix is a tool for creating text representations as vectors. Each metric in the vector quantifies a linguistic feature in the text. These features included:
    - **Lexical Features:** These involve word choice, punctuation, vocabulary richness, and the frequency of specific words or phrases. For instance, how often certain words appear.
    - **Syntactic Features:** These include sentence structure, punctuation use, and sentence length. We looked at how often sentences start with capital letters, the average sentence length, and the use of various punctuation marks.
    - **Stylistic Elements:** This covers the use of emojis, hashtags, and special characters. For example, some groups might consistently use certain emojis or have a specific way of formatting their messages.
    - **Functional Words:** These are common words like prepositions, conjunctions, and pronouns, which often unconsciously reflect the author's style.
2. **Aggregation:** The extracted features were aggregated to create a unique stylistic fingerprint for each hacktivist account. This involved summarizing the feature values across all messages from each account. By doing this, we developed a profile for each group that encapsulates their unique writing style.
3. **Comparison:** We used similarity measures to compare the stylistic fingerprints of different accounts. This analysis helped us identify accounts with similar writing styles, suggesting potential connections between them. For example, if two groups frequently use similar phrases, punctuation, or even unique emojis, it might indicate a shared author.
4. **Visualization:** The results of the stylometric analysis were visualized using network graphs. These graphs highlighted the connections between different hacktivist groups based on their writing styles, highlighting potential common authors or coordinated efforts. For instance, a dense cluster in the graph might suggest a single entity operating multiple accounts, while isolated nodes indicate independent operators.

Stylometry can reveal a wealth of information beyond mere authorship. For instance:

- **Attribution:** By comparing the writing styles of various groups, we can attribute specific messages to known hacktivist entities or even uncover previously unknown affiliations. Often, as mentioned above, new groups are created. Stylometric analysis of the messages of the new group can help us understand whether it is a rebranding of a previously known group.
- **Evolution of Writing Style:** Analyzing how the writing style of a single account evolves over time can provide clues about changes in the group's membership or operational strategy. It can help us observe whether a group operated by a single author, started having additional authors, and even spot if they are taking shifts.
- **Detecting Deception:** If an account suddenly changes its writing style, it might indicate that the account has been taken over by a different entity or that the original authors are trying to disguise their identity.

By identifying stylistic consistencies across different messages and accounts, we can:

- **Link seemingly unrelated hacktivist groups:** Discover connections that are not immediately apparent through traditional analysis.
- **Enhance the accuracy of attribution:** Provide additional evidence to support the identification of state-sponsored actors behind certain groups.
- **Uncover coordinated campaigns:** Detect patterns that suggest a larger, organized effort rather than isolated incidents.

By using stylometric analysis in our methodology, we were able to gain a deeper understanding of the relationships and behaviors of hacktivist groups. This approach provided a layer of analysis that complements traditional methods, offering a more nuanced view of these groups.

Results

Topic Modeling Analysis

In our analysis, we utilized the BERTopic framework to identify and categorize the main themes discussed by various hacktivist groups. After clustering the messages, we extracted the most representative words for each cluster using the class-based TF-IDF (c-TF-IDF) method. We used GPT to assign meaningful names to these topics, which provided concise and descriptive labels based on the top keywords and representative documents for each cluster. Some initial topics were too specific or too broad, requiring refinement. For instance, smaller topics related to leaking documents from the Ukrainian Army were merged into the broader "Leaking Documents from Ukrainian organizations" topic, while the general topic of "Cyber Attacks in Europe" was split into more specific topics like "Cyber Attacks Against Spain" and "Cyber Attacks Against UK." This iterative process ensured that the final set of topics was both comprehensive and precise. Eventually, we ended up with a set of topics that represented the main interests of these groups.

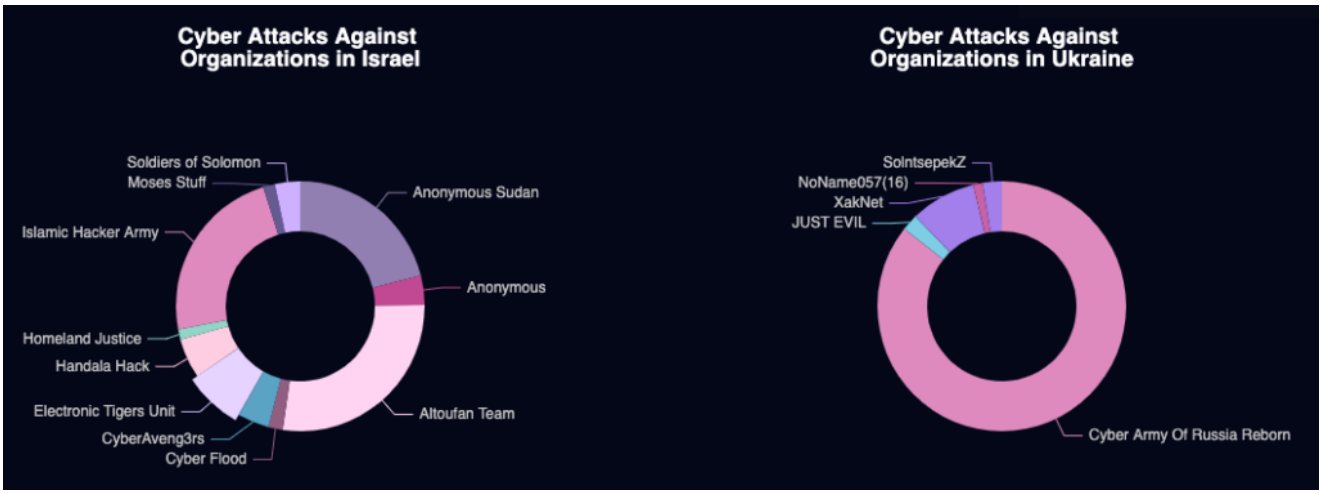

Figure 3: Topic modeling visualization of users-per-topic

There were multiple topics discussing cyber attacks against organizations in different countries such as Israel, Ukraine, Russia, Iran, India, Spain, the USA, and more. Others focused on cyber attacks against militant organizations such as Hezbollah in Lebanon and the Huthis in Yemen. Some topics were related to the leaking of sensitive information from countries such as Ukraine, Israel, Russia, and Iran. Many topics were directly linked to major geopolitical events, such as the Russian invasion of Ukraine and the Israel-Hamas conflict. This highlights the role of hacktivist groups in cyber warfare and propaganda. On some occasions, similar topics discussed by different groups suggest possible coordination or shared objectives.

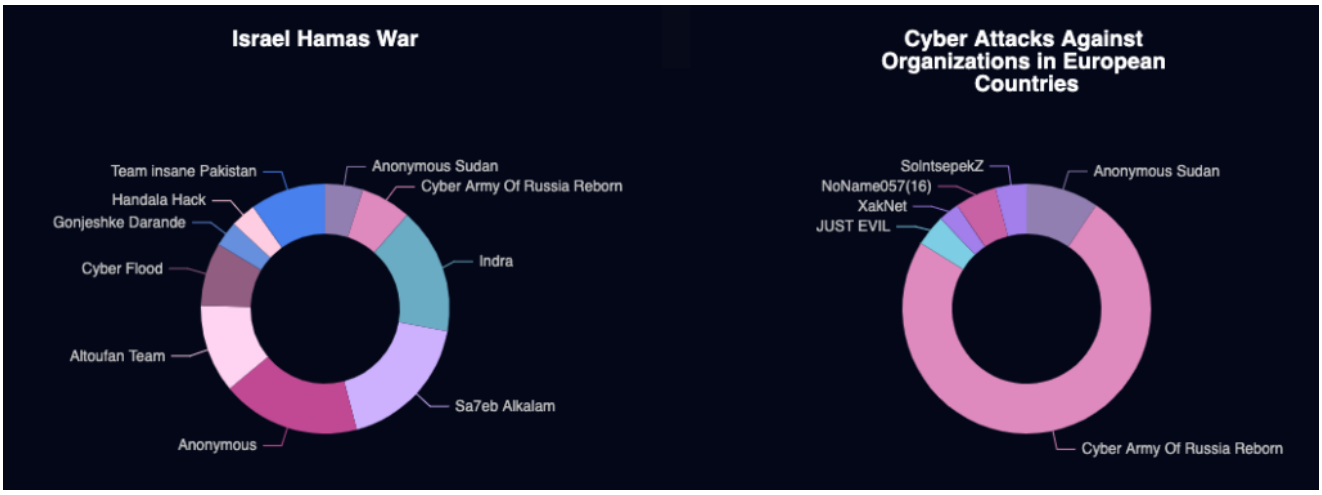

Figure *4*: Topic modeling visualization of users-per-topic

Given that these groups are likely orchestrated by intelligence agencies, analyzing the topics they discuss provides insight into the missions or goals these accounts are intended to serve. By visualizing the evolution of these topics over time, we detected changes in focus or interest by the entities managing these accounts. This combined analysis offers a clearer

understanding of who might be behind these groups, based on their shifting interests and objectives. We observed how groups react to certain geopolitical events and the response time of others. For example, Russia-affiliated hacktivist groups initiated attacks in tandem with the Russian invasion of Ukraine. It took a few months for Ukrainian-related hacktivist groups to consistently retaliate with their own attacks.

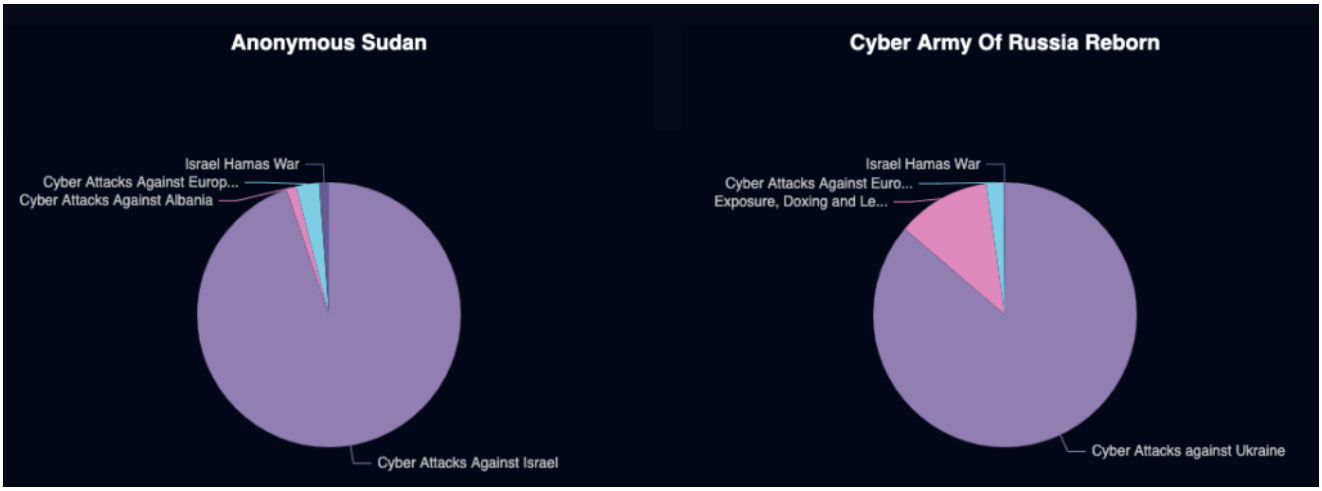

Figure *5:* Topic modeling visualization of topics-per-user

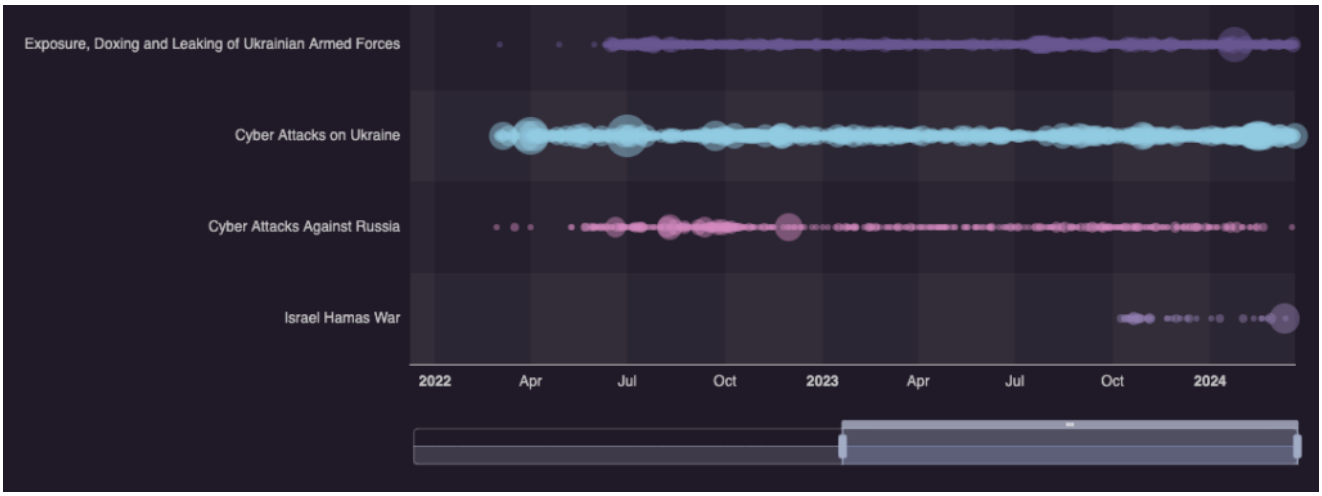

Figure 6: Topic modeling timeline visualization of topics and volume related to certain geopolitical events

Stylometric Analysis

Stylometric analysis was used to compare the writing styles of different hacktivist groups and uncover potential connections. By extracting nearly 200 features from the collected messages, including lexical, syntactic, and stylistic elements, we created unique stylistic fingerprints for each account. This analysis not only identifies authorship but also provides deeper insights into the organizational structure and operational strategies of these groups.

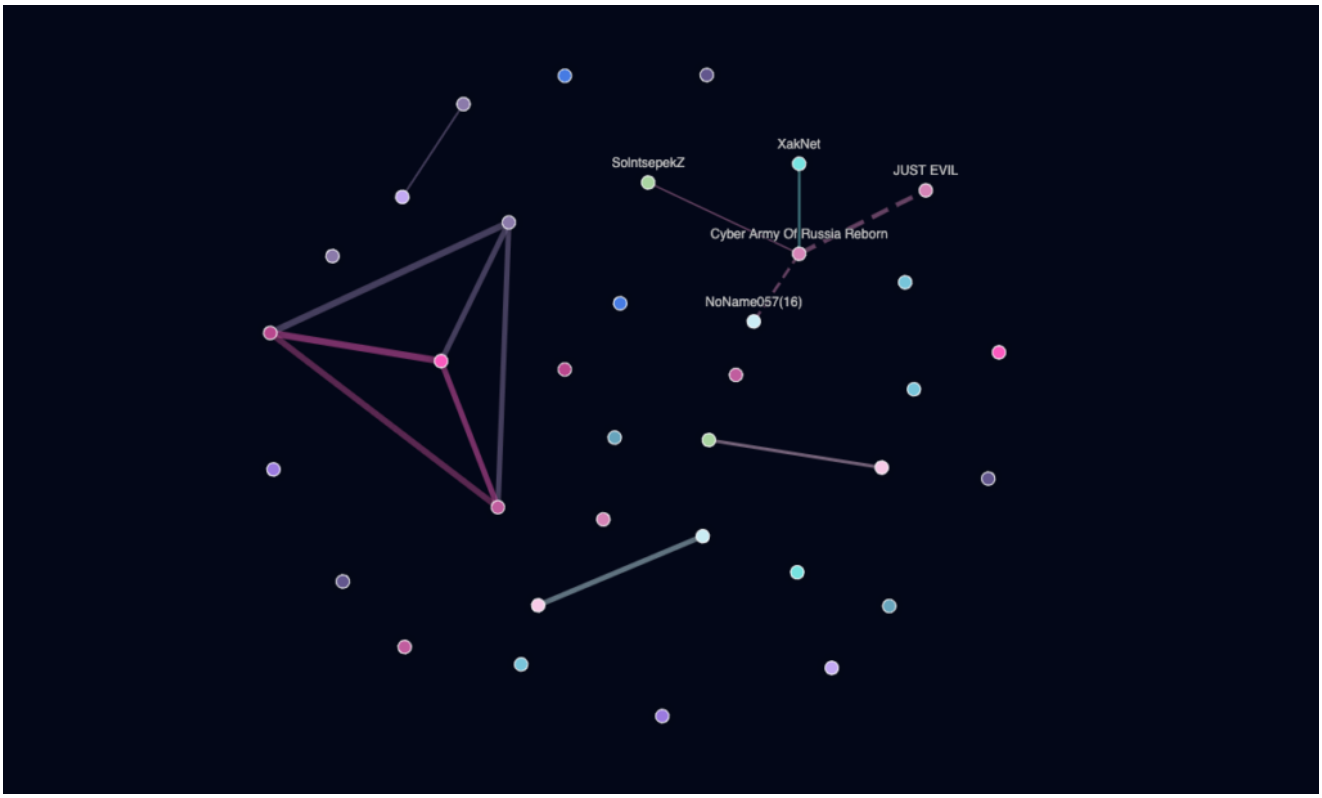

Figure 7*:* Stylometric similarity visualization draws connections between different hacktivist groups

Using similarity measures to compare these stylistic fingerprints, we identified several notable connections. Certain groups exhibited highly similar writing styles, suggesting they might be operated by the same individuals or teams. For instance, the Cyber Army of Russia Reborn, Solntsepek, and XakNet showed significant stylistic overlap. This finding supports Google Mandiant's reports indicating that these groups are cyber personas used by APT44, providing further evidence of their coordinated operations. In addition, two groups not mentioned in the Mandiant report—JustEvil and NoName057—also showed some similarity to this cluster of activity. We also discovered connections between other clusters, including those targeting organizations in Israel and Albania, and another cluster focusing on Iran.

Sudden changes in writing style within a single account indicated possible changes in authorship. For example, the IT Army of Ukraine showed a distinct shift in style around 2022, coinciding with a change in its messaging focus. Initially, the account featured messages in Arabic cheering for the Egyptian football team, but after the Russian invasion of Ukraine, the content shifted dramatically to focus on cyber attacks against Russia. This stark change suggests that the account was purchased, repurposed, or taken over by the group. Other examples demonstrate that more than one person is writing the messages for some groups, while other groups have only a single author who consistently writes in a similar manner.

Implications

The identification of stylistic overlaps between different groups suggests coordinated efforts and shared objectives, possibly under a single directive from intelligence agencies. This enhances our ability to attribute attacks and understand the broader strategy behind them.

Moreover, detecting sudden shifts in writing style within an account can signal changes in the group's control or a strategic redirection. This information is vital for threat intelligence, as it provides early indicators of significant changes in hacktivist activities.

By combining these insights with the results from topic modeling, we gain a better understanding of hacktivist groups' motivations, strategies, and interconnections. This holistic view is essential for understanding and improving the attribution of operations by hacktivist groups.

Discussion

Interpretation of Results

The findings from our topic modeling and stylometric analysis provide insights into the activities and evolution of hacktivist groups. By identifying key themes discussed by these groups, we get a better understanding of their primary targets, which include various countries and organizations. The topics frequently centered around major geopolitical events, such as the Russian invasion of Ukraine and the Israel-Hamas conflict, highlighting the role of hacktivist groups in cyber warfare and propaganda. The evolution of topics over time reveals how these groups adapt their messaging in response to global events, reflecting their strategic objectives.

Stylometric analysis showed the interconnectedness of hacktivist groups, with certain groups exhibiting similar writing styles indicative of common authorship or coordination. This was particularly evident in clusters of groups targeting specific regions, such as Ukraine, Israel, and Iran. Additionally, shifts in writing style within individual accounts, suggest changes in ownership of the accounts and might indicate a shift in strategic direction.

Understanding the topics and writing styles of hacktivist groups improves threat intelligence by organizing the knowledge regarding their motivations and strategies. This is essential for creating effective countermeasures and accurately attributing cyber attacks to specific groups or state-sponsored entities. Continuous monitoring and analysis of hacktivist communications can provide early warnings of emerging threats, enabling proactive defense measures.

Limitations

Despite the valuable insights gained, this research has several limitations. Data collection posed significant challenges, as hacktivist groups frequently change their accounts (the social media platforms often ban their old accounts), limiting the availability of data. Also, it is often hard to understand if an account is legit or some kind of a fan page or a copycat of a hacktivist group. Moreover, the reliance on social media messages means that the analysis may not capture the full scope of hacktivist activities, which might include messages written on their website (often on the Darkweb), forums, and sometimes directly to journalists.

The machine learning models used in this study, while powerful, have inherent limitations. Topic modeling relies on the quality and comprehensiveness of the input data, and small or obscure topics may be overlooked. Stylometric analysis, though effective in identifying writing style similarities, may produce false positives if different authors share similar linguistic habits or use a language that is not properly supported by the models. Also, with enough effort, it is possible to mimic the writing style of certain groups, hence misleading the stylometric model.

Future Research Directions

To build on the research findings, future research should expand the analysis to include more hacktivist accounts. With many groups emerging rapidly, it is essential to set up automated processes for continuous monitoring and analysis. This automation would allow for real-time insights and a more comprehensive understanding of hacktivist activities.

Additionally, researchers should explore other data types, such as the metadata of videos and PDFs, to gain further insights. Analyzing unique hashtags, recurring typos, and mentions across different accounts can also provide valuable information. Applying visual stylometry concepts to videos and graphic images posted by these groups can reveal additional patterns and connections.

Conclusion

Our research's significance lies in its innovative approach to analyzing hacktivist communications. By combining topic modeling and stylometric analysis, we provided a comprehensive view of these groups' activities and their evolution. However, it's important to note that this approach is not infallible. The cyber threat landscape is continuously evolving, and so must our methods. We need to keep experimenting, innovating, and collaborating to stay ahead of these threats.

The blurred lines between state-sponsored activities and traditional grassroots hacktivism highlight the need for adaptive and innovative threat intelligence methods. Our efforts to understand these complex dynamics are driven by the recognition that we do not have all the answers and this fact is what drove us to seek an understanding in the beginning.

Acknowledgments

I would like to extend my gratitude to my colleague, Amit Levin, from Check Point's Data Science team, for their invaluable assistance and guidance throughout this research. Thank you, Amit, for your contributions and collaboration.

GO UP
BACK TO ALL POSTS