Troll Hunting

written by Mert SARICA | 1 March 2024

Table of Contents

Introduction

In recent years, we have seen the increasing importance of cyber threat intelligence for organizations. As a result, the number of products and services used for this purpose within organizations has been growing rapidly. Firstly, cyber threat intelligence provides a significant advantage to organizations by helping them identify threat actors targeting their industry, understand the tactics, techniques, and procedures they employ, and prepare for potential cyber attacks. Furthermore, when organizations are exposed to a cyber attack, cyber threat intelligence can shed light on various aspects of the incident, from enriching the data collected during the Incident Response process to establishing connections with threat actors.

Not only organizations but also end users who are closely intertwined with technology, individuals like us, need to remember that we can benefit from cyber threat intelligence and platforms in some cases (such as Troll account investigation).

What is Troll, Trolling and Disinformation?

As in all over the world, we witness the proliferation of messages shared by Troll accounts on social media and in the media, with the aim of manipulating the masses with disinformation attack. Sometimes, these false pieces of

information can be shared from individuals' own accounts, as well as through fake, anonymous accounts.

In slang, a troll is a person who posts or makes inflammatory, insincere, digressive, extraneous, or off-topic messages online (such as in social media, a newsgroup, a forum, a chat room, an online video game) or in real life, with the intent of provoking others into displaying emotional responses, or manipulating others' perception, thus acting as a bully or a provocateur. The behavior is typically for the troll's amusement, or to achieve a specific result such as disrupting a rival's online activities or purposefully causing confusion or harm to other people. (Source: Wikipedia)

Disinformation attacks involve the intentional dissemination of false information, with an end goal of misleading, confusing, or manipulating an audience. False information that is not intentionally deceptive is referred to as misinformation, although that has also been used as a catch-all term. Disinformation attacks may be executed by political, economic or individual actors to influence state or non-state entities and domestic or foreign populations. These attacks are commonly employed to reshape attitudes and beliefs, drive a particular agenda, or elicit certain actions from a target audience. Tactics include the presentation of incorrect or misleading information, the creation of uncertainty, and the undermining of both correct information and the credibility of information sources. (Source: Wikipedia)

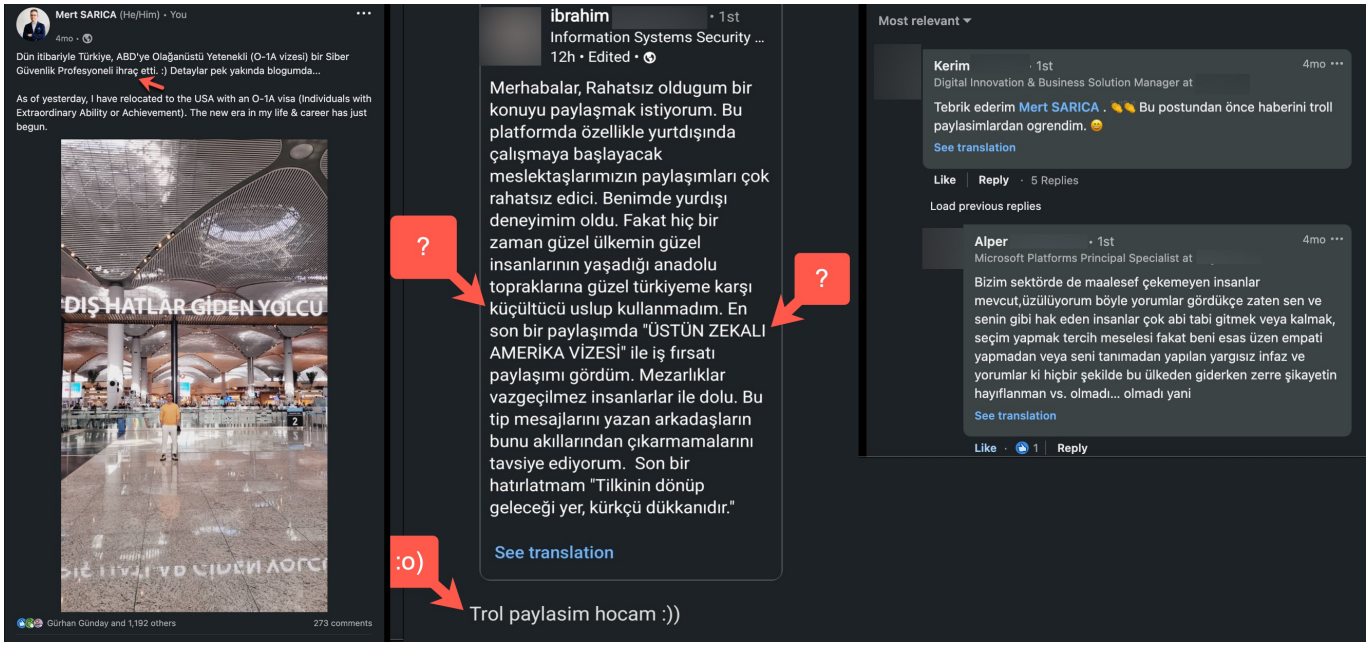
From 2009 until now:

1. I have shared over 200 technical cybersecurity articles on my blog in both English and Turkish, with the motto 'Knowledge is power and grows as it is shared'
2. Since 2015, I have organized the 'Pi Hediyem Var' cybersecurity game, through which I have gifted more than 15 Raspberry Pi devices to university students with the support of my sponsors.
3. I have been a speaker at nearly 30 cybersecurity events.
4. With my presentations titled 'Ethical Hacking and Career,' I have guided thousands of students interested in the field of cybersecurity in almost 40 universities.

Despite all my hard work, I am rarely targeted by troll accounts and in some cases the reason is just because I moved to the United States in 2022.

Examples of Trolling

Most of the time, troll accounts have already closed their accounts, deleted their messages, or received responses before I can even understand what's happening, thanks to the reactions and criticisms from followers who don't doubt my good intentions even for a moment. In such cases, there is often no need for me to rely on cyber threat intelligence or platforms. However, in situations like my blog posts titled "WhatsApp Scammers," "Exposing Pig Butchering Scam," I personally benefit greatly from cyber threat intelligence and platforms when dealing with scammers.



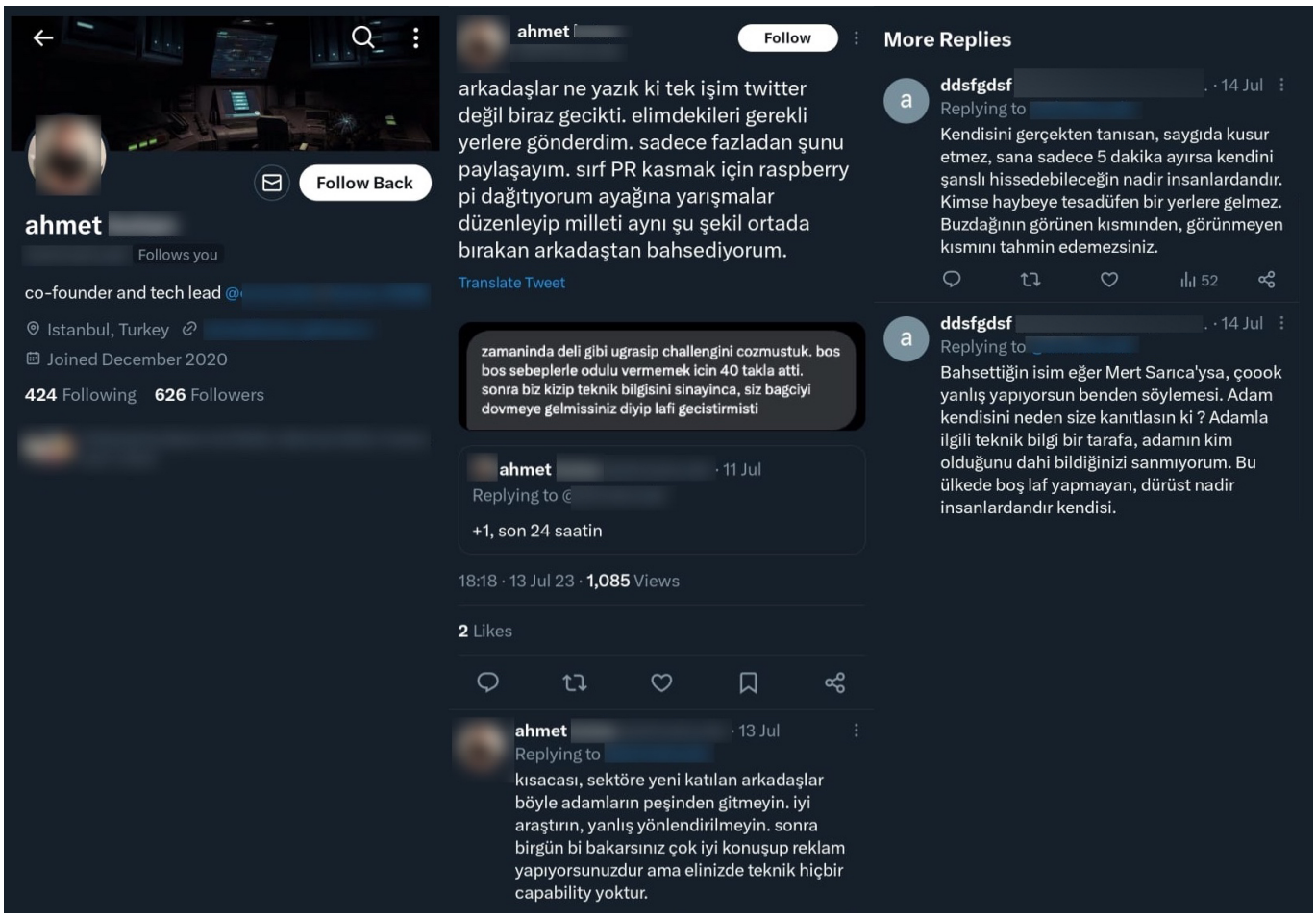
Message deleted

In this screenshot, someone shared a message regarding my LinkedIn message and tried to mislead, confuse, or manipulate people. On the left side, you see my post in English, and on the middle, you see his post in Turkish. After he received reactions and criticisms from my beloved connections/followers, he had to delete his message.

Here is the English translation of his message;

"Hello, I'd like to share a concern that has been bothering me. On this platform, especially the posts of our fellow members who are about to start working abroad can be quite disturbing. I've had international experience myself, but I have never used a belittling tone when referring to the beautiful lands of Anatolia and my beloved Turkey, where wonderful people live. In a recent post, I saw a job opportunity shared with the title "GENIUS

AMERICA VISA.” Cemeteries are filled with irreplaceable individuals. I recommend that those who write such messages do not forget this. One final reminder: “No matter where the fox goes, he shall end in the furrier’s shop.”



Message deleted

In another screenshot, someone acted as a bully or a provocateur with falsified claims about me. After he received a reaction and criticism from one of my beloved followers, he had to delete his message.

Here is the English translation of his messages;

“I have sent what I had to the necessary places. I just want to share one more thing. I’m talking about the guy who just gifts Raspberry Pi for the sake of building public relations (PR), organizes competitions, and then leaves people hanging in the same way.”

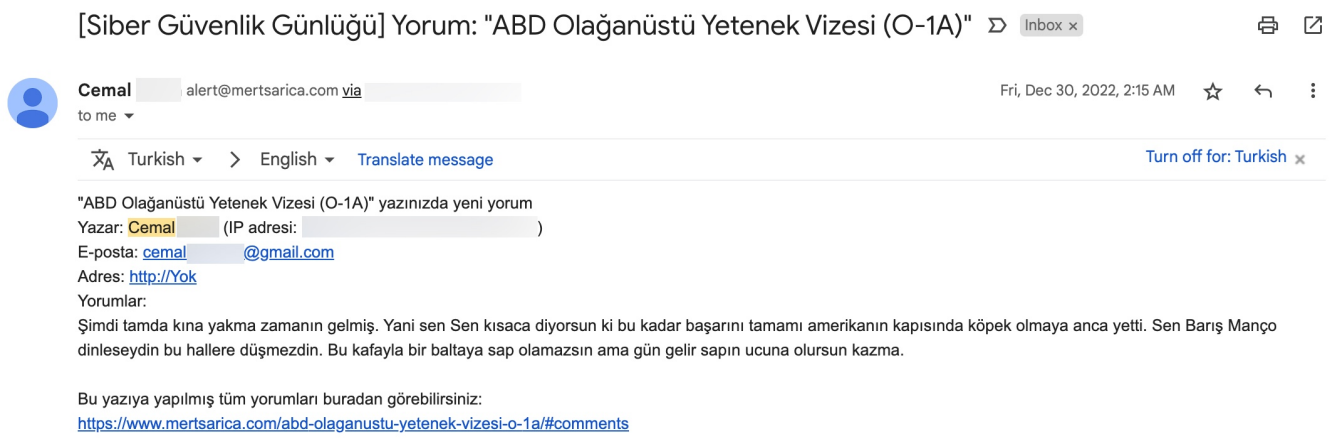
“In short, newcomers to the industry, don’t follow such people. Research well, don’t be misled. Someday, you may find yourself speaking well and advertising, but you may lack the technical capability.”

Especially when you insult, use profanity, engage in character assassination,

spread false accusations, threaten, or criticize someone's patriotism directly or indirectly on social media, you should always remember that sooner or later, this will come back to haunt you. Even if you regret your actions and delete your messages for a reason, it's not something that can be easily erased from records and memories.

At times, when you receive a suspicious, malicious or evil message via email, social media, or a network, you can also benefit from cyber threat intelligence platforms to understand the intentions and identity of the person behind it.

In this example, someone first insulted me through my website via email, and then, after not being responded to for a month, attempted to make contact via social media with a different approach.


[Siber Güvenlik Günlüğü] Yorum: "ABD Olağanüstü Yetenek Vizesi (O-1A)" Inbox ×

Cemal alert@mertsarica.com via Fri, Dec 30, 2022, 2:15 AM

to me

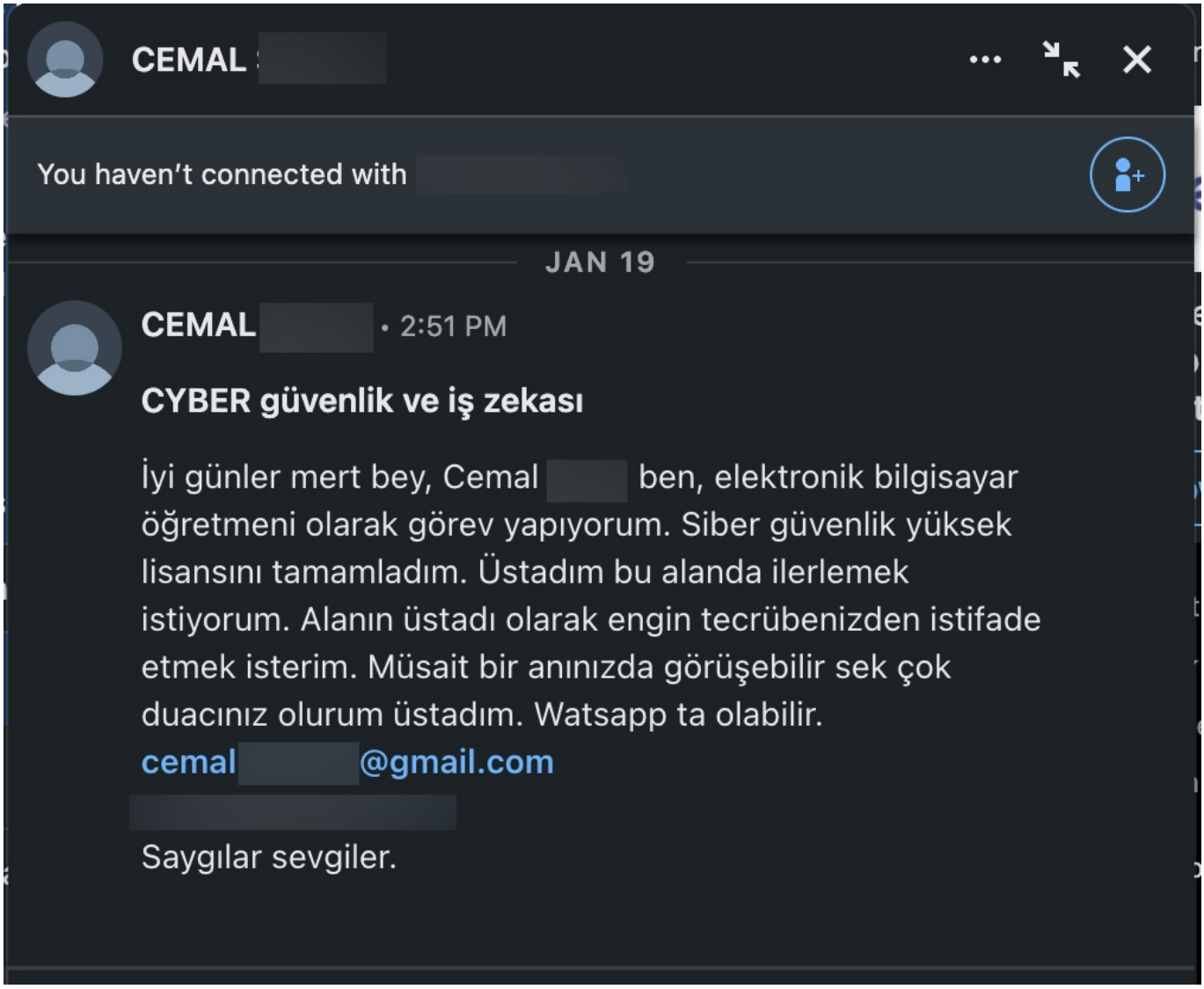
Turkish English Translate message Turn off for: Turkish

"ABD Olağanüstü Yetenek Vizesi (O-1A)" yazınızda yeni yorum
Yazar: Cemal (IP adresi:)
E-posta: cemal @gmail.com
Adres: http://Yok
Yorumlar:
Şimdi tamda kına yakma zamanın gelmiş. Yani sen Sen kısaca diyorsun ki bu kadar başarını tamamı amerikanın kapısında köpek olmaya anca yetti. Sen Barış Manço dinleseydin bu hallere düşmezdin. Bu kafayla bir baltaya sap olamazsın ama gün gelir sapın ucuna olursun kazma.

Bu yazıya yapılmış tüm yorumları buradan görebilirsiniz:
https://www.mertsarica.com/abd-olaganustu-yetenek-vizesi-o-1a/#comments

Here is the English translation of his e-mail;

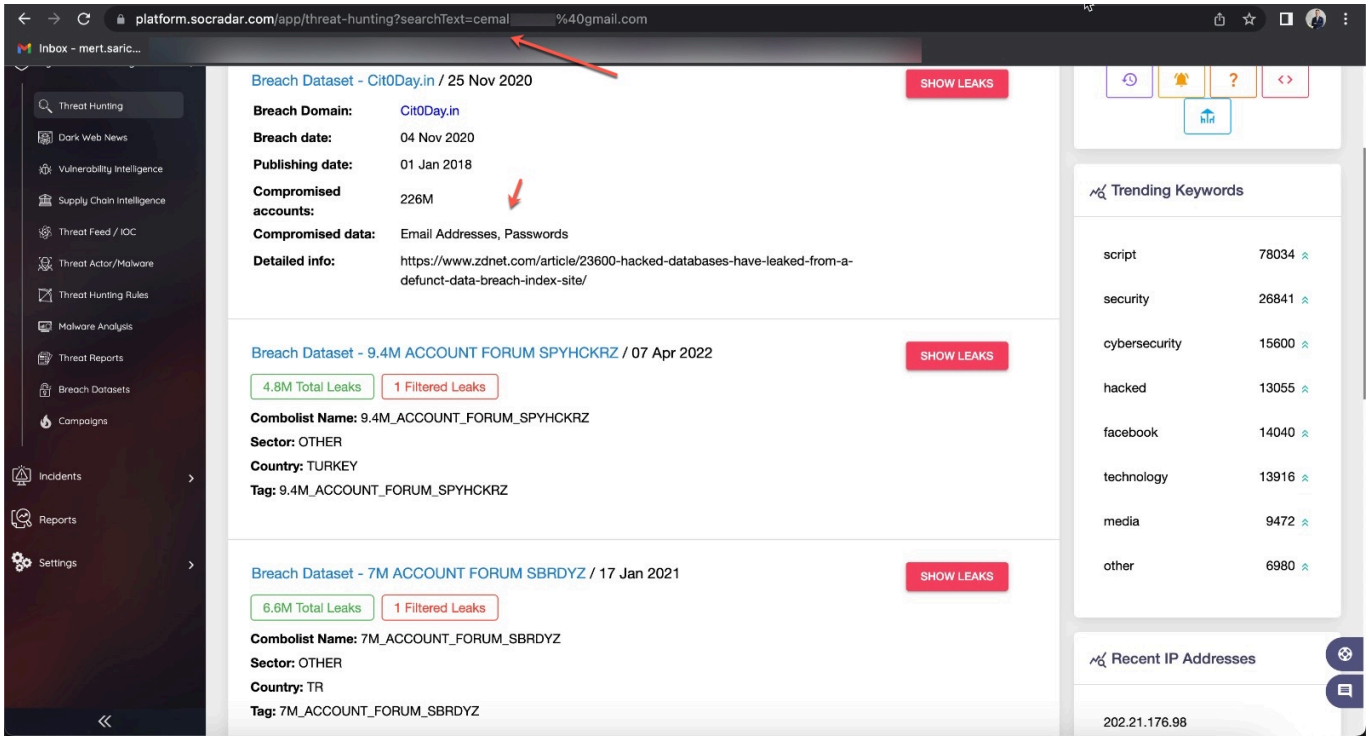
"Now it's time to light the henna, you're basically saying that all your success only managed to get you as far as being a doorman of the United States. If you had listened the songs of Barış Manço, you wouldn't have ended up like this. With this mindset, you won't achieve much, but one day, you might just become the handle of a hoe."



Here is the English translation of his LinkedIn message;

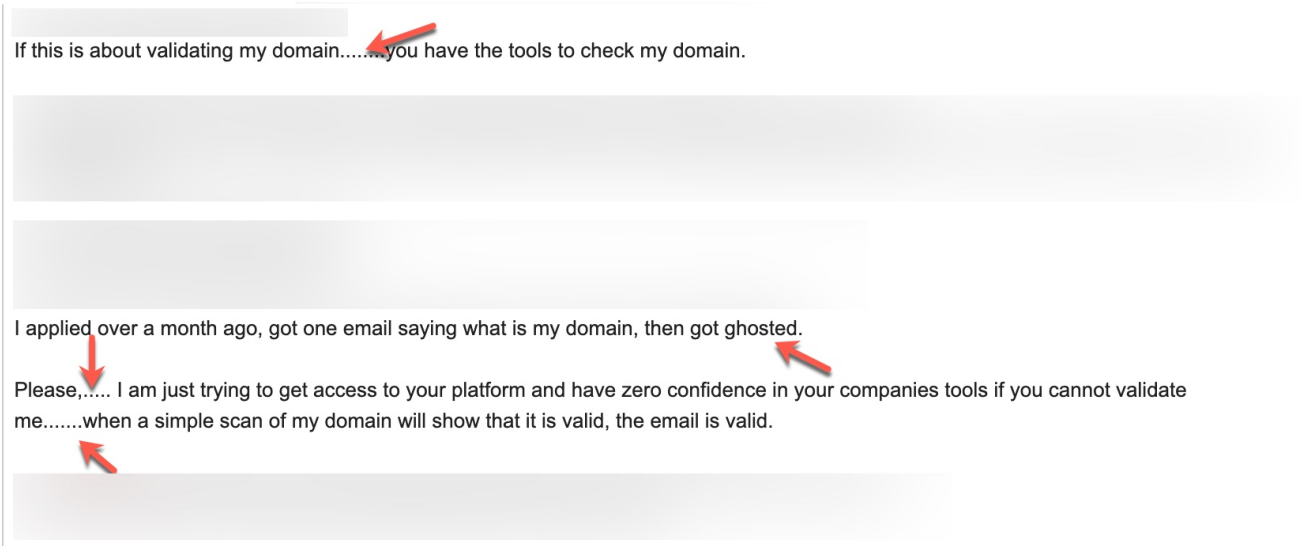
"Good day, Mr. SARICA. I'm Cemal, and I work as a computer science teacher. I have completed my master's degree in cybersecurity. I would like to advance in this field and benefit from your extensive experience as a master in the field. I would greatly appreciate it if we could have a discussion at a convenient time. You can also reach me on WhatsApp."

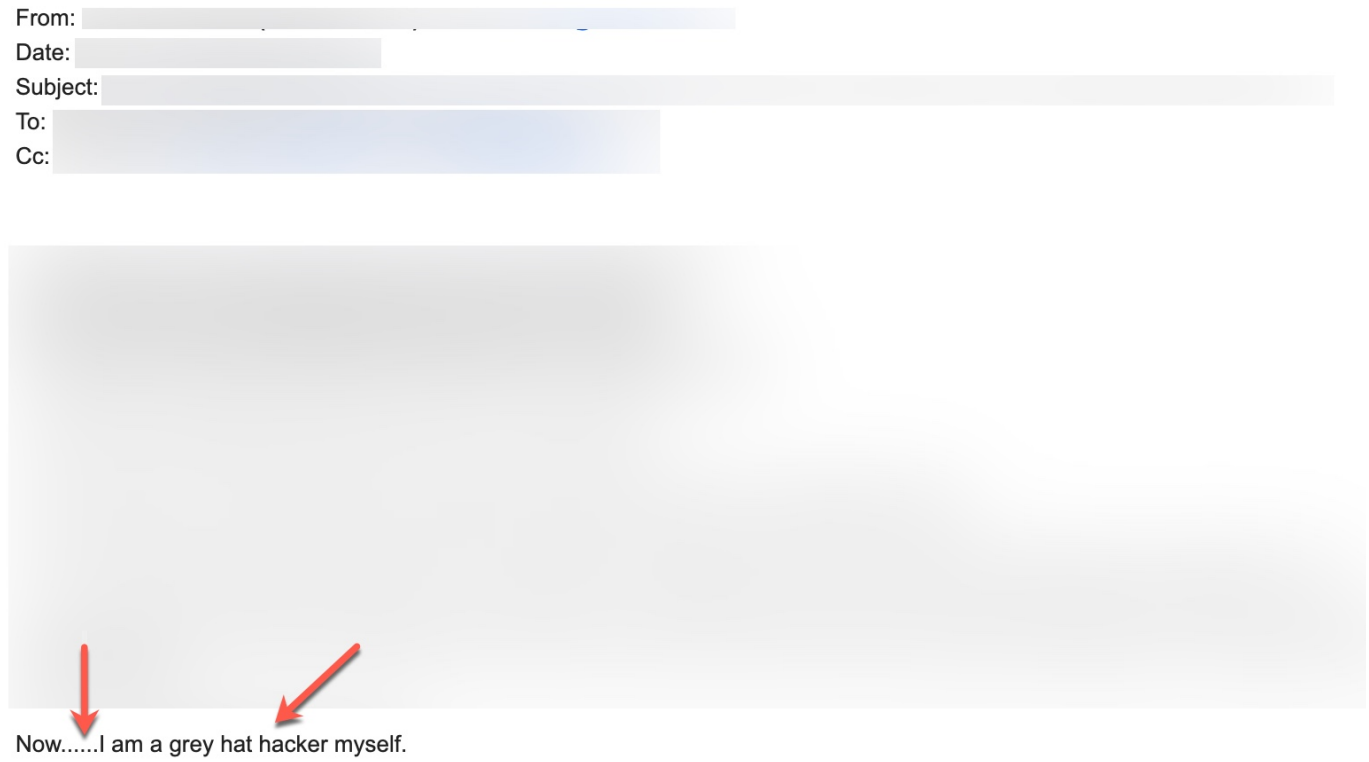
When I searched the email address in the SOCRadar Cyber Threat Intelligence Platform, I easily gained insights and information about his intentions and motivations, such as learning that the individual has been active in hacker forums for years.

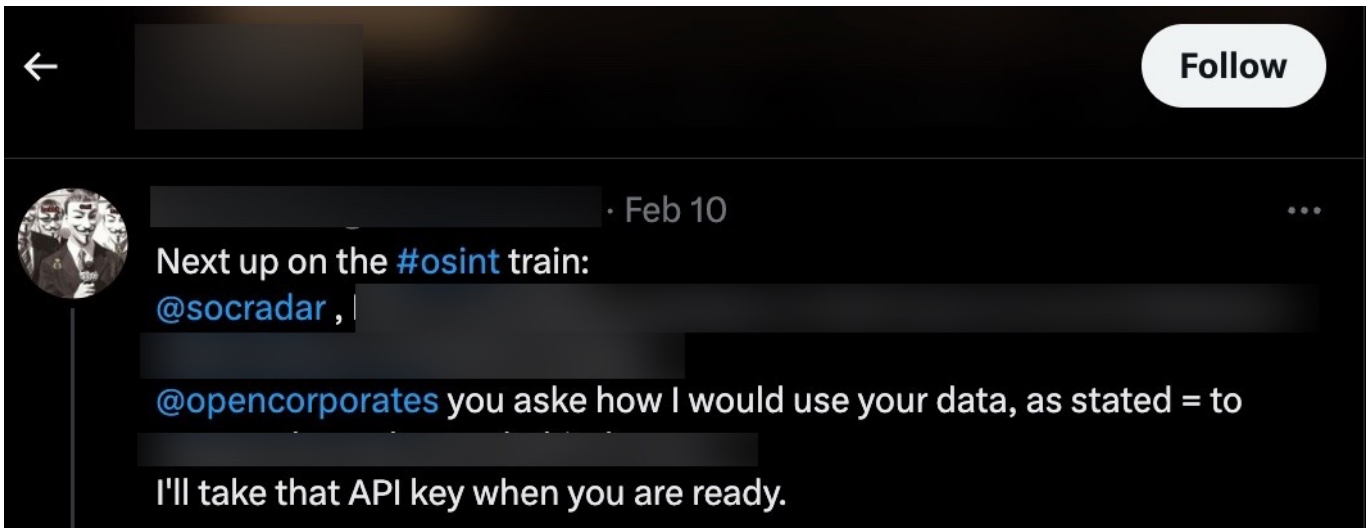


If you work for a cyber threat intelligence firm that tracks threat actors, cybercriminals, and scammers, and shares intelligence related to their operations, there may be times when you encounter threatening messages targeting your organization through anonymous accounts on social media. In such situations, you can utilize your own platform as well as employ different methods.

In this example, an individual who was not allowed to register for the cyber threat intelligence platform due to failed security checks first sends threatening emails, and then starts trolling through an anonymous Twitter account.







Account disabled

In order to determine that a person who sends emails using their real identity and another person who shares messages through an anonymous Twitter account behind a Guy Fawkes mask are highly likely to be the same individual, we can rely on stylometric methods to identify the author based on the samples we have from the suspected person's emails and the messages shared on the anonymous Twitter account. These stylometric methods involve analyzing various elements such as punctuation, spelling mistakes, emphasis, foreign

words, slang and jargon, conjunctions, abbreviations, numbers, subject tags, and symbols.

By examining and comparing these elements, we can observe similarities or differences between the emails and the anonymous Twitter account, which can help us determine if they are likely authored by the same person. However, stylometric analysis alone cannot provide a definitive conclusion, as it relies on statistical patterns and probabilities. Therefore, it is important to consider additional factors and evidence when making a determination.

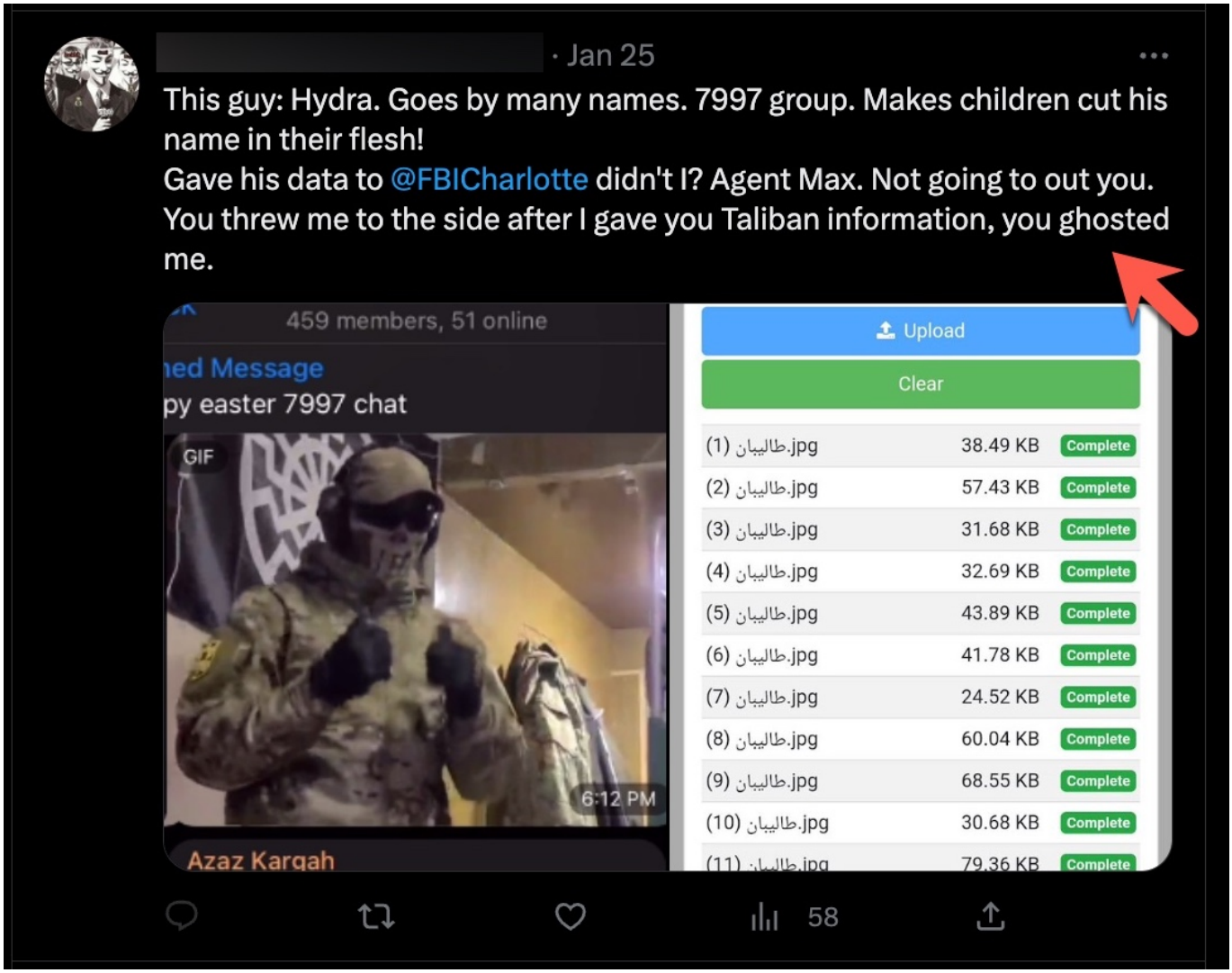
What is Stylometry?

Stylometry is a style identification method used primarily in written literature, but also in other artistic disciplines such as visual arts and music, as well as in fields like history, religion, law, and forensic sciences. Stylometric analysis is a method based on the examination of style markers as variable factors using statistical and computational techniques.

For approximately two centuries, stylometry has been applied to compare and analyze the literary styles of authors, particularly in the context of authorship attribution problems. The methods used in stylometry range from basic statistical calculations and tests to artificial neural networks. Stylometric studies have been conducted on various subjects, ranging from religious texts and historical documents to the examination of plagiarism in scientific works and the analysis of literary works and author styles. (Source: A Bridge From Statistics To Literature: The Stylometry Analysis – Ayşe İŞİ, Fatih ÇEMREK, Zeki YILDIZ)

Troll Hunting

At this stage, rather than conducting an extensive stylometric analysis, I decided to have a brief look at the 109 messages on the anonymous Twitter account and focus on common words and punctuation marks that overlap with the emails. Based on the notes I took, I noticed that the messages frequently used ellipsis (…..) punctuation marks and included the term “ghosted” which is not commonly encountered in English correspondence. When I examined the emails from the suspected individual, I found that my observations closely matched, significantly increasing the likelihood of similarity between the two individuals.



When I started thinking about what I would do if the number of messages shared on the anonymous Twitter account was not 109 but 10,009, I decided to seek the help of data science.

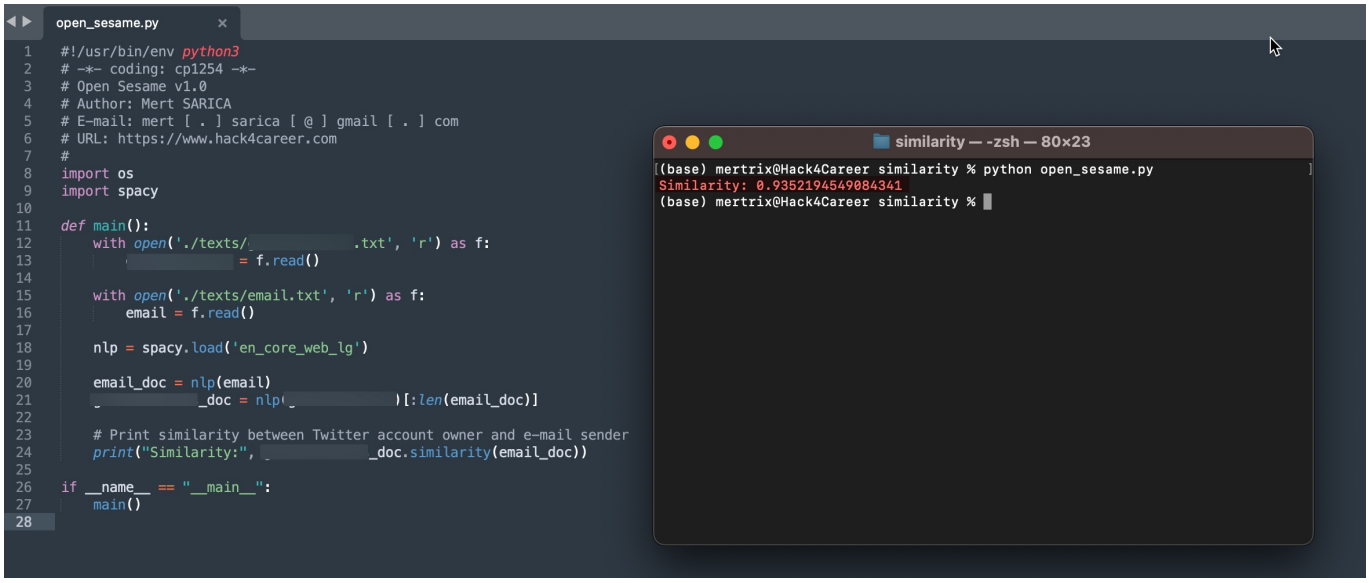
Data science is the field of study that involves working with data to extract

meaningful insights and make predictions for business purposes. It is an interdisciplinary approach that combines principles and practices from mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data. This analysis helps data scientists ask questions about what the data is, why it is the way it is, what it could become, and what can be done with the results. By employing data science techniques, professionals can uncover patterns, trends, and correlations within the data to drive informed decision-making and solve complex problems.

After conducting some research, I learned that I can leverage the similarity method in the Natural Language Processing (NLP) library called SpaCy, which measures the similarity between texts using Cosine Similarity.

Cosine Similarity is a measure that quantifies the similarity between texts in a vector space. It calculates how many times words appear in the texts. Then, each text is represented as a vector, with 1s and 0s indicating the presence or absence of words. When these vectors are placed in a three-dimensional space, the smaller the cosine angle between them, the closer the texts are to each other. For completely unrelated vectors, the cosine value is 0, while for completely opposite documents, the cosine value will be -1. (Source: A Content Recommendation System Application with TF-IDF Algorithm and Cosine Similarity on Netflix Data – Özlem GELEMET, Hakan AYDIN, Ali ÇETİNKAYA)

After I coded a tiny Python tool to examine the similarity between the emails from the person I suspected and the Twitter messages, with the help of the SpaCy library I concluded that they were highly likely sent by the same person. :)



Hope to see you in the following articles.

Note: For those who want to learn more about the use of stylometry, I recommend reading the free section of the book “Real-World Python: A Hacker’s Guide to Solving Problems with Code“.