

Shedding light on the dark web

[i blog.group-ib.com/ml-in-investigations](https://blog.group-ib.com/ml-in-investigations)



28.01.2022

Cybersecurity analyst's guide on how to use machine learning to show cybercriminals' true colors



Vesta Matveeva

Head of investigation department, APAC



Iaroslav Polianskii

Senior Data Scientist, APAC

Introduction

Data leaks appearing on the dark web are an actual problem of the modern world. Events of the recent past show that even world-renowned enterprise, financial and IT giants are not immune to data leaks. Equifax, British Airways, SingHealth, Marriott International, Sephora, Canva, Zynga, Microsoft, Tokopedia, T-Mobile, LinkedIn, Twitch — these are just a few names that have been spotted in high-profile data breach scandals in the past several years. One would hardly attempt to question if these companies had enough resources to ensure their security, but here is the fact — they all did fall victim to cyberattacks.

Even if the company's data was leaked not as a result of its own actions, but rather the actions of its contractor or a partner with which it shared data, the result will be the same. Data leaks can both disclose sensitive information about the company's internal processes affecting business decisions or reputation and personal data of the company's customers, making them turn their back on a brand or company that used to be their favorite. Both outcomes can bring the business at the verge of collapse, which is why one of the first instincts of breached companies, burning from thirst for revenge, is to find attackers behind the network compromise and data leak to dispense justice.

These data leaks can often end in the shadow part of the Internet. Underground resources, which are closed to unauthorized visitors, are rife with discussions that are directly related to planned or previously committed crimes. These resource regulars are attackers themselves, their accomplices looking for an opportunity to make some money, but also visitors who have their own agenda, though, distinct from others. This is the case for employees of law enforcement agencies and special services, as well as corporate and private security specialists. Under the guise of bad guys, they collect cyber intelligence data and research and analyze criminal activity in order to investigate and prevent cybercrimes.

An average underground forum has huge volumes of daily text messages going through it, which makes analyzing this information manually almost impossible and also ineffective. Algorithm development and machine learning (hereinafter - ML) implementation can break this stalemate, significantly reducing the amount of manual work and enabling analysts to connect the dots in the tremendous amount of data.

This article aims to show the methods cybersecurity analysts who come to the aid of compromised companies in such cases can use to, firstly, determine if an alleged data breach did take place and a database put up for sale in the dark web was authentic, and, secondly, identify the threat actor responsible. It demonstrates how machine learning algorithms can facilitate the processes of cyber intelligence data analysis and cyber investigations, while at the same time further enriching its results. And if it happens that you somehow embarked on the path of investigating a data leak, this guide will give you the ideas from what to begin and how to further proceed.

This guide is intended for:

- cybersecurity greeners who are doing their first steps in the cybersecurity world. Even if not all the terms and techniques described in the text are familiar to you, you can get your first impression of cyber investigation process and strengthen your intention to pursue a cybersecurity career;
- cybersecurity analysts and corporate security team members. From the text you'll learn the methods that can be used to probe into a data leak, even if for the time being you're sure that your customers are reliably protected;
- machine learning algorithm developers who will get a broader perspective of the cybercrime investigation process and be able to apply this knowledge in the future to advance the cybercrime investigation industry, making the process more efficient and prompt.

Thus, we'll focus on two major points:

1. **Real breach or fake.** There are leaks that are purported to be new to the public, but in reality they turn out to be databases that were earlier released somewhere else or are new only to some extent, comprising both old and fresh data.
2. **Skilled threat actor or a newbie.** Many attackers today use multiple accounts on the underground forums in order to better hide their activities. To determine the goal of an attack and proceed with its further probe, one has to know the adversary.

Manual analysis

We decided to demonstrate the methods that can be used for the achievement of the aforesaid goals using a case with **RaidForums** underground platform where a set of databases was released on September 17, 2020 by a user nicknamed **ExpertData**.

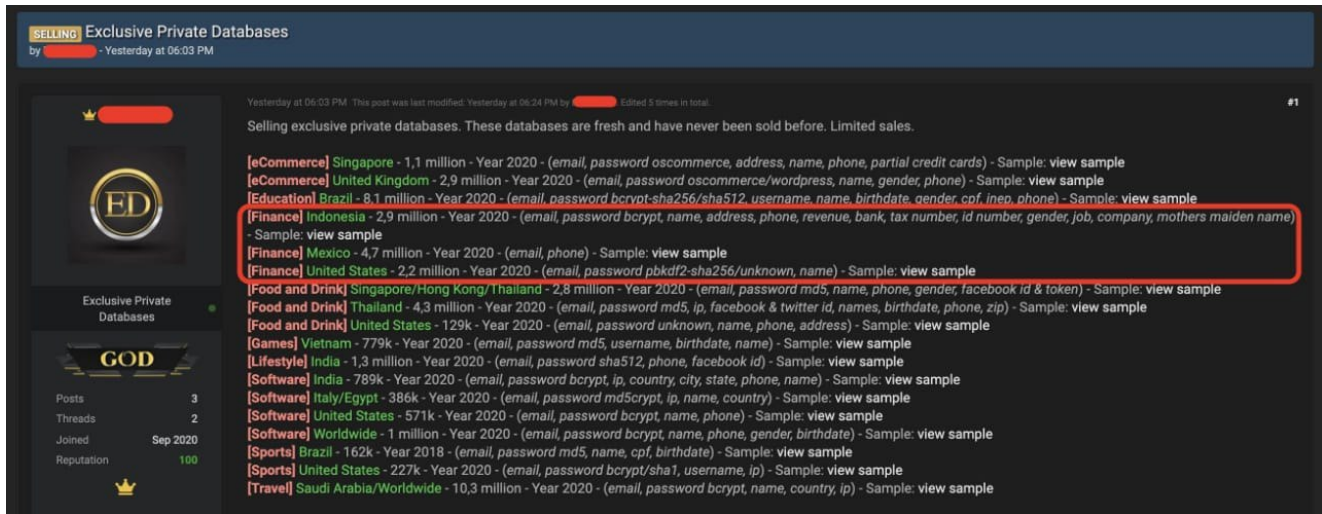


Figure 1 - Screenshot of **ExpertData's** post offering for sale a set of databases posted on **RaidForums** (currently deleted; screenshot taken from https://twitter.com/Bank_Security)

First, Group-IB analysts began to manually analyze the **ExpertData** account. The information publicly available on the forum indicates that this user joined **RaidForums** on September 5, 2020.

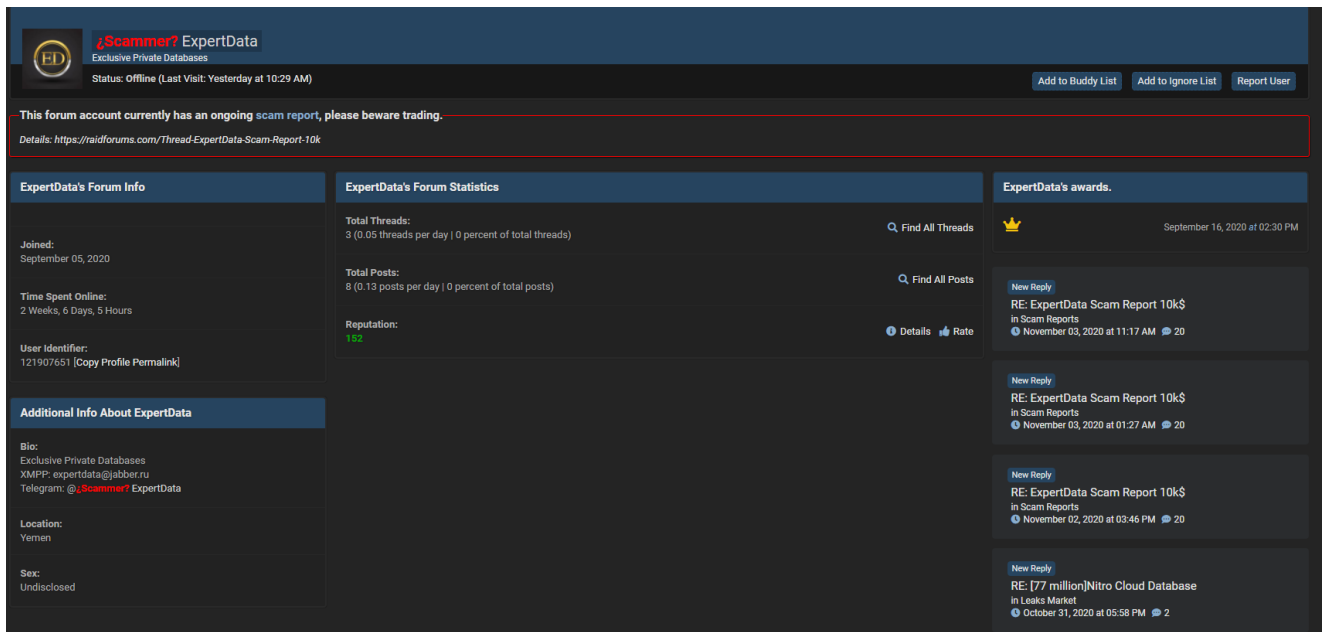


Figure 2 - **ExpertData's** profile on **RaidForums** (captured in November 2020, the label "Scammer" was added by the forum admins after several claims against **ExpertData**. This label was then removed after the disputes were resolved)

The account was then renamed to repeat its user identifier 121907651 (several threads and posts of **ExpertData** were removed by admins, which is why the total number of posts differs from fig.2).

121907651
GOD User
Status: Offline (Last Visit: November 15, 2020 at 08:42 PM)

121907651's Forum Info

GOD

Joined:
September 05, 2020

Time Spent Online:
3 Weeks, 7 Hours

User Identifier:
121907651 [Copy Profile Permalink]

Username Changes:
1

121907651's Forum Statistics

Total Threads:
1 (0 threads per day | 0 percent of total threads)

Total Posts:
1 (0 posts per day | 0 percent of total posts)

Reputation:
212 [Details](#)

Additional Info About 121907651

Sex:
Undisclosed

Figure 3 - **ExpertData**'s profile on RaidForums (captured in September 2021)

So the situation was quite typical — a new user published information about multiple data leaks from companies in different regions. It was reasonable to doubt if the databases were real leaks and if the attacker could have been trusted. Then we started to analyze threads of discussion where this account participated.

On September 17, 2020, **ExpertData** published a thread to sell a leaked database, allegedly containing info on over 100,000 users.

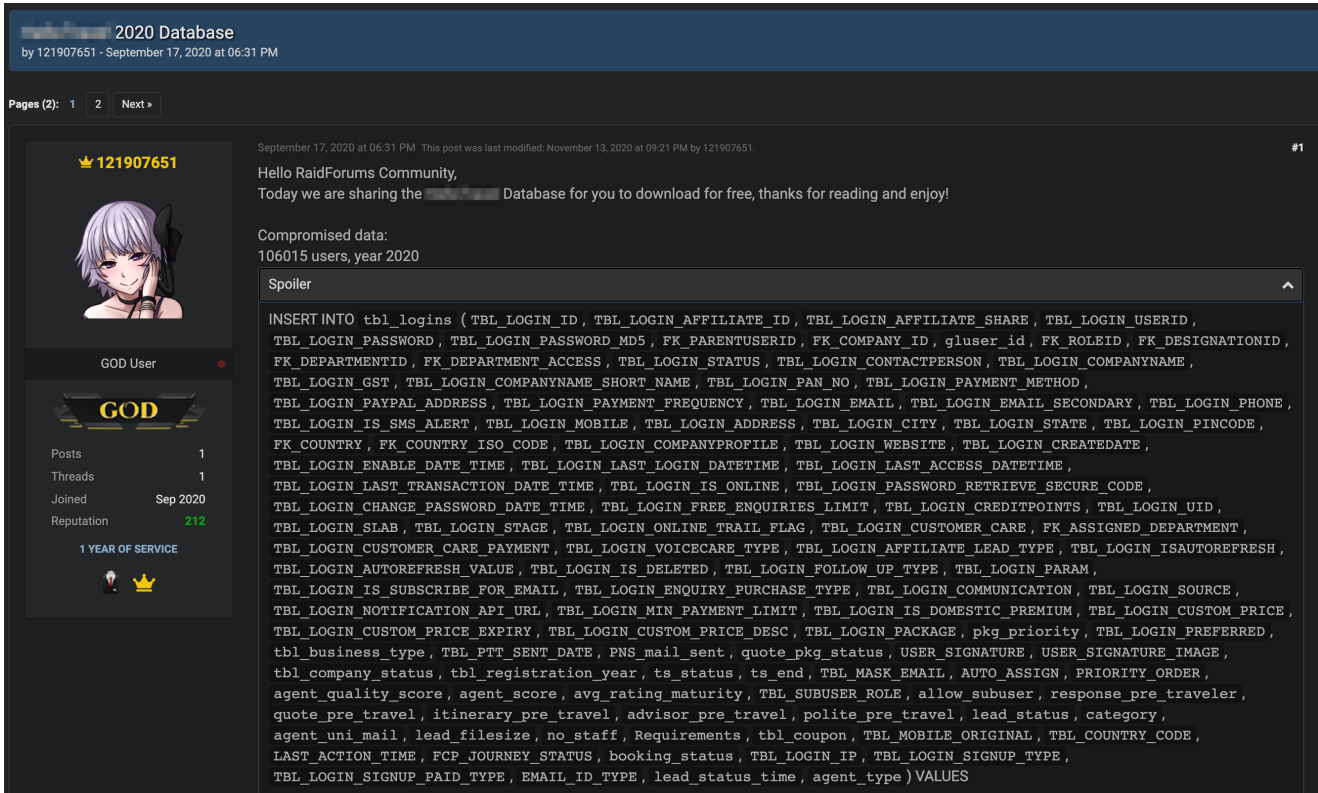


Figure 4 - ExpertData's post offering for sale a database

On the same day, **ExpertData** published a thread to sell 18 databases leaked from companies in Brazil, Egypt, Indonesia, India, Italy, Hong Kong, Mexico, Saudi Arabia, Singapore, Thailand, Vietnam, UK, US, and other countries (fig.1). As it can be seen from the screenshot, the original post didn't mention the names of the companies compromised.

About a month later, on October 28, 2020, **ExpertData** published the same post, this time specifying the names of the companies breached and excluding Saudi Arabian companies in the travel industry from the list.

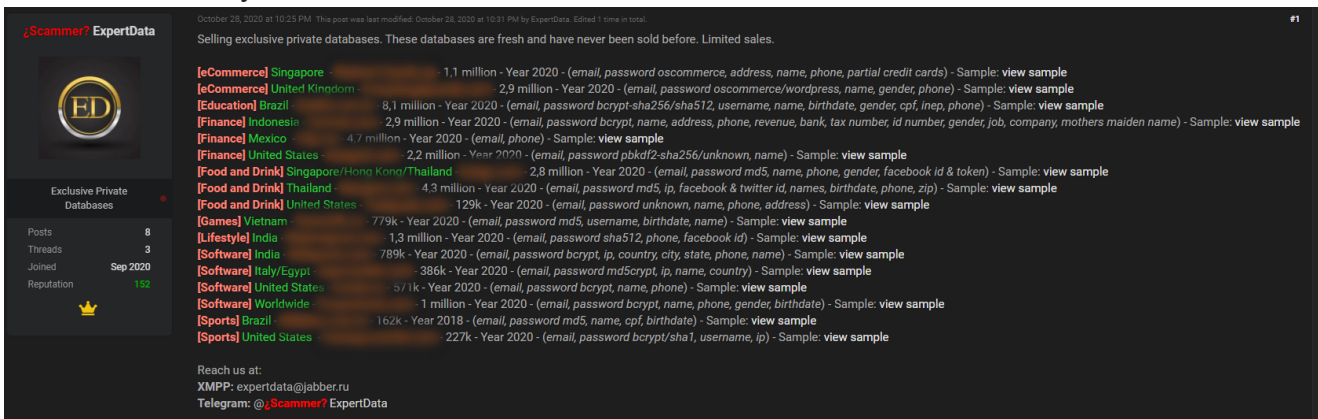


Figure 5 - Screenshot of ExpertData's post offering for sale a set of databases mentioning the names of the companies compromised

In the post **ExpertData** has also specified his contact details: Jabber account expertdata@jabber.ru and Telegram account [@ExpertData](https://t.me/ExpertData).

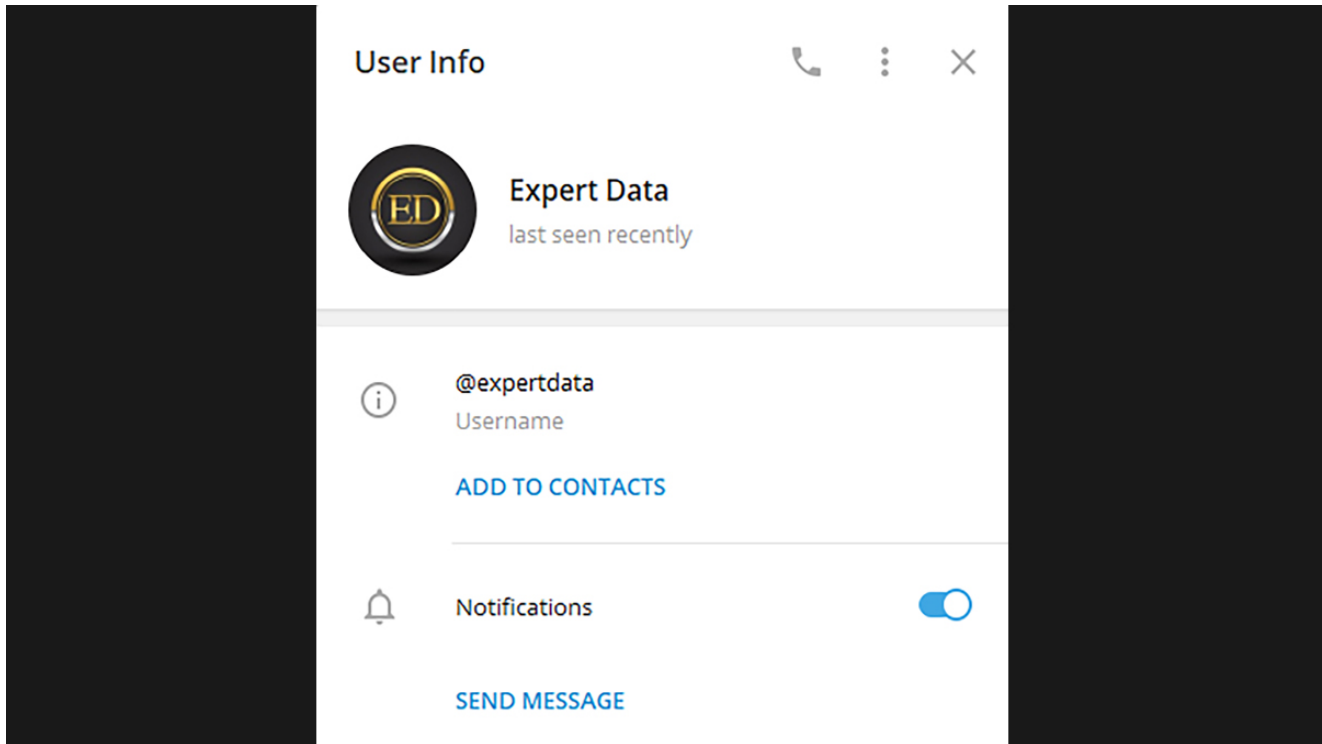


Figure 6 - *ExpertData's* account on Telegram

The databases published by this author seemed to be unique since they had not been seen in the publications of other threat actors before. This was also noted by other forum members, who commented on one of ExpertData's posts.

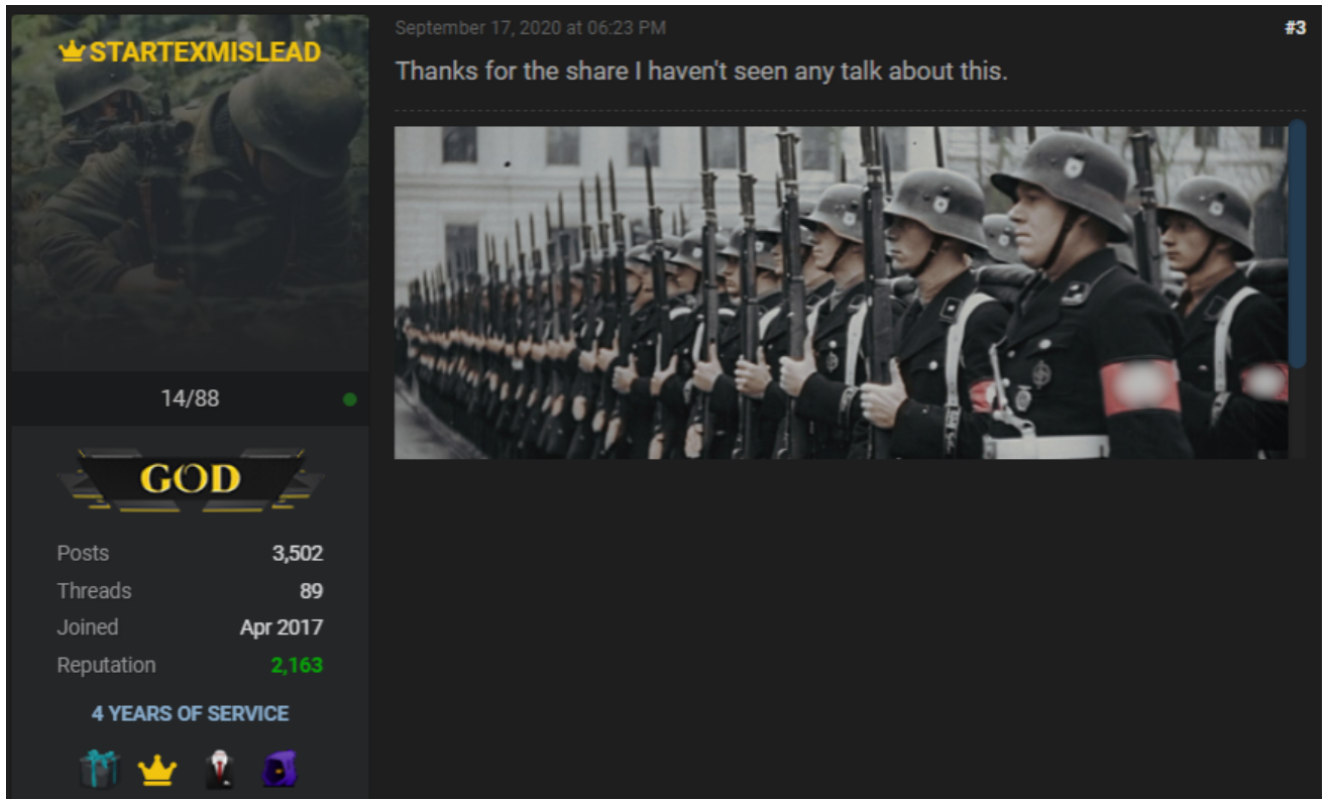


Figure 7 - One of the active forum users with a high reputation commenting on *ExpertData's* post

In early November 2020, a user nicknamed **The Polaris** initiated a dispute against **ExpertData** in a separate thread. **The Polaris**, who attempted to purchase a database from **ExpertData**, then accused the latter of lying about the quality (quantity of correct password hashes) of the NitroGo (GoNitro.com) dump in a bid to get the payment worth \$10,000.

The dispute thread was published on November 2, 2020. At the time of writing, the thread is deleted, but Group-IB analysts managed to retrieve it.

The screenshot shows a forum post on a dark-themed website. The post title is "ExpertData Scam Report 10k\$" and it is attributed to "The Polaris" on November 02, 2020 at 12:10 AM. The post content describes a scam where the user was promised a \$10,000 payment for a database dump but received a file with incorrect password hashes. The user lists a series of sample hashes from the dump, such as "https://lbb.co/hLnJpQj" and "https://lbb.co/GPwwKW6". To the left of the main text is a user profile for "The Polaris" (GOD User) with a crown icon, showing 10 posts, 2 threads, and a join date of August 2020.

Figure 8 - Screenshot of the dispute thread against **ExpertData** initiated by a user nicknamed **The Polaris**

The Polaris joined Raidforums on August 29, 2020, six days before **ExpertData** account was created.

The Polaris
Banned

Status: Offline (Last Visit: November 10, 2020 at 10:24 PM)

This forum account is currently banned.

Ban Reason: Failed to abide by terms of a deal. If you have any questions please contact @Omnipotent. polaris1000@protonmail.com 45.9.236.13 2a0f:df00:0:255::74
Banned By: Jaw — Ban Length: Permanent N/A


The Polaris's Forum Info	The Polaris's Forum Statistics	The Polaris's awards.
<p>Joined: August 29, 2020</p> <p>Time Spent Online: 1 Day, 18 Hours, 15 Minutes</p> <p>User Identifier: 121904915 [Copy Profile Permalink]</p>	<p>Total Threads: 1 (0.01 threads per day 0 percent of total threads)</p> <p>Total Posts: 11 (0.1 posts per day 0 percent of total posts)</p> <p>Reputation: 30 Details</p>	<p> November 05, 2020 at 10:59 PM</p> <p> August 29, 2020 at 11:00 PM</p> <p>New Reply RE: Shinyhunters and ExpertData SCAM 40k\$ in Archives November 10, 2020 at 07:42 PM 6</p> <p>New Reply RE: Shinyhunters and ExpertData SCAM 40k\$</p>
Additional Info About The Polaris		
<p>Sex: Undisclosed</p>		

Figure 9 - The Polaris' account on RaidForums

Surprisingly, **ExpertData** wasn't the first to reply to this dispute: a user with the nickname **ShinyHunters** stood up for him and came up with counter accusations against **The Polaris**.

November 02, 2020 at 12:54 AM #4

ShinyHunters



GOD User

GOD


Posts: 79
Threads: 36
Joined: Apr 2020
Reputation: 1,334

Ok, I don't usually support a rival especially when I don't know him & don't give a fuck, but all I know is that, you're multi accounting and you're a known leaker:
<https://prnt.sc/vbfzb3>
I also see that you extort this seller, so I do support @Scammer? ExpertData, try again Cellbris....
This for having leaked the tokopedia dump.

Reply

November 02, 2020 at 01:32 AM. This post was last modified: November 02, 2020 at 01:37 AM by The Polaris. Edited 1 time in total. #5

The Polaris



GOD User

GOD

Posts: 10
Threads: 2
Joined: Aug 2020
Reputation: 0

ShinyHunters Wrote: → (November 02, 2020 at 12:54 AM)

Ok, I don't usually support a rival especially when I don't know him & don't give a fuck, but all I know is that, you're multi accounting and you're a known leaker:
<https://prnt.sc/vbfzb3>
I also see that you extort this seller, so I do support @Scammer? ExpertData, try again Cellbris....
This for having leaked the tokopedia dump.

it has nothing to do with this report, you can continue to protect whoever bought this from you, shinny

wait for the administrator's response

ShinyHunters Wrote: → (November 02, 2020 at 12:54 AM)

Ok, I don't usually support a rival especially when I don't know him & don't give a fuck, but all I know is that, you're multi accounting and you're a known leaker:
<https://prnt.sc/vbfzb3>
I also see that you extort this seller, so I do support @Scammer? ExpertData, try again Cellbris....
This for having leaked the tokopedia dump.

yesterday I was wondering if you have it, but today it appears on sale, I won't be surprised if at your request from your friend, and we scam me

Reply

Figure 10 - Screenshot of the dispute thread against **ExpertData** showing **ShinyHunters** standing up for the former

In their post, **ShinyHunters** accused **The Polaris** of hiding under multiple accounts and included a link to a screenshot of their earlier conversation in Jabber, in which **ShinyHunters** suspected that it might be cybersecurity researcher Vinny Troia hiding behind the user nicknamed **The Polaris**, and welcomed him in a personal chat as "Hello Vinny."

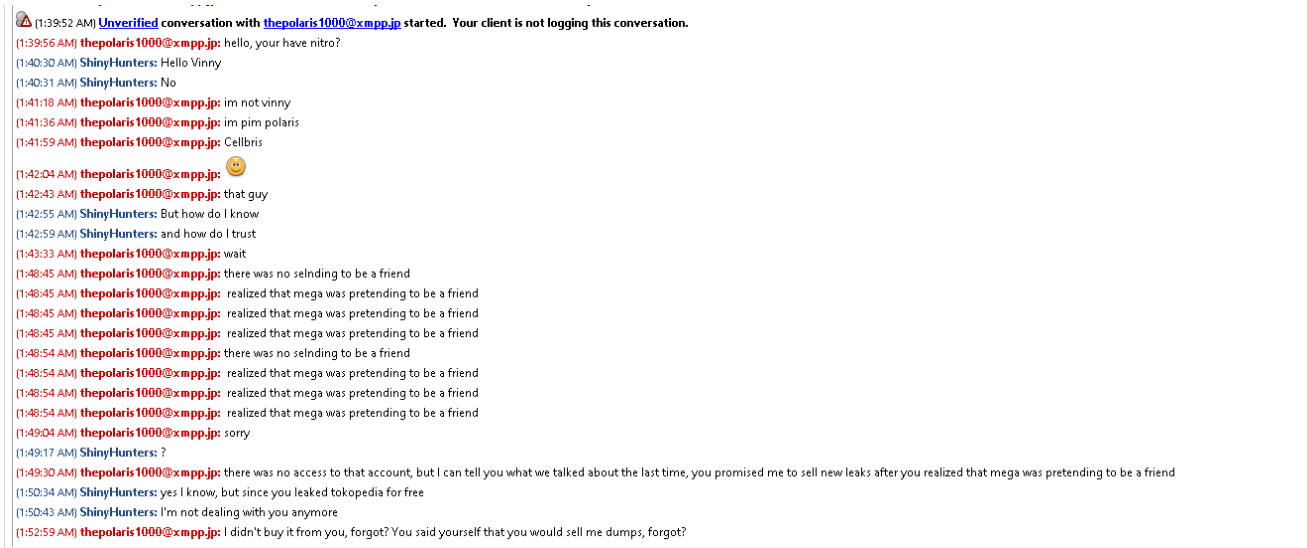


Figure 11 - Screenshot of a chat between **The Polaris** and **ShinyHunters** in Jabber

ShinyHunters are a famous group of hackers who intrude their victim networks and exfiltrate data. The group with this name became famous in the APAC region after Tokopedia leak in May 2020, and is famous for selling unique databases that they exfiltrate during their hacking activity.

Vinny Troia is, in particular, known for researching hacker group **ShinyHunters** and having published in July 2020 a research about them. According to his findings, it is supposed that **ShinyHunters** is a successor to **NSFW** and **Gnostic Players** groups.

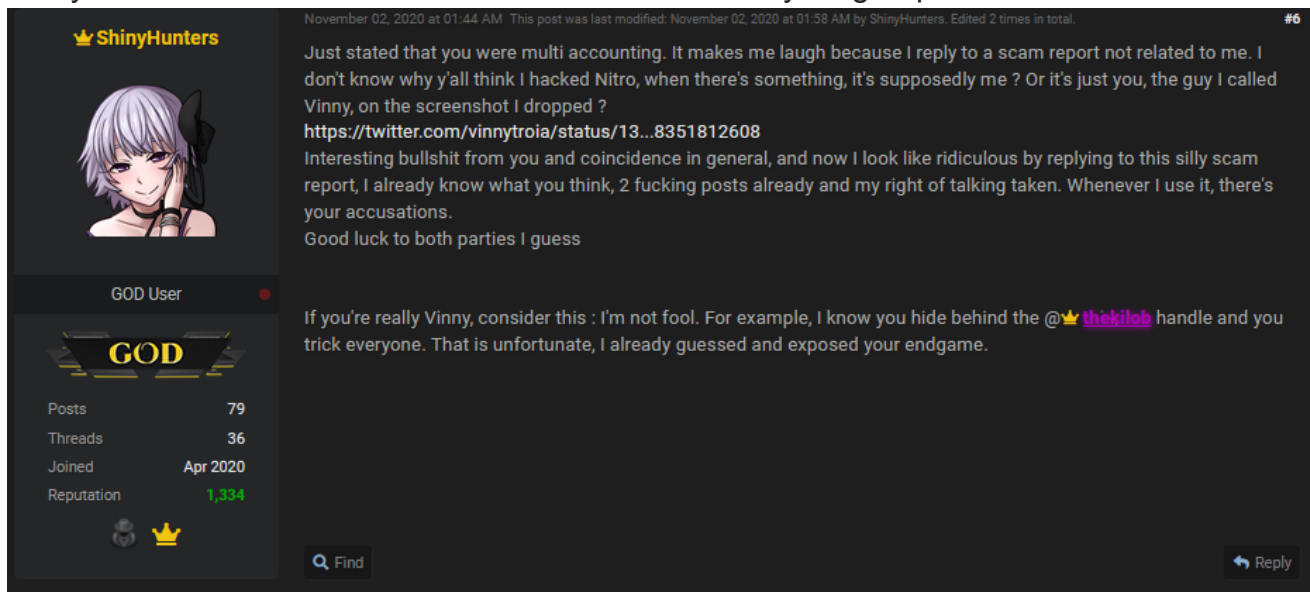


Figure 12 - Screenshot of **ShinyHunters'** post, in which they claim that **The Polaris** is multi accounting and has another profile on the forum

The dispute stopped on November 10, 2020. **The Polaris** was banned. Moreover, the account of **ExpertData** was labeled as a scam one due to several disputes on the forum. After the disputes were resolved, this label was removed from the **ExpertData** account,

which might indicate that the claims of **The Polaris** were not deemed relevant by forum administrators.

Typically, the **ShinyHunters** account doesn't join any disputes not against them, which is why we suggested that there might be a link between **ShinyHunters** and **ExpertData**. To further explore this assumption, we started comparing the threads, in which these two users participated and their posts themselves.

After the post with 18 private databases, **ExpertData** published two more databases (now the posts are deleted from RaidForums): GoNitro (on October 29, 2020) and AnimalJam (on November 11, 2020).

On October 21, 2020, the Nitro PDF (sold as GoNitro dump) service stated that it had been subject to a low impact security incident, however, the media then reported that user and document databases allegedly stolen as a result of the incident were offered for sale in a private auction. **ExpertData** published information about the GoNitro databases on October 29, 2020, which seems to be the first public post offering for sale the databases.

Later, in January 2021, GoNitro was republished again by the user nicknamed **Spiral** with a reference to **ShinyHunters**.

The screenshot shows a forum post by user **Spiral** on a dark-themed forum. The post title is "Nitro PDF / Gonitro.com - Full Breach - [77M]" and it was posted on January 20, 2021, at 02:42 AM. The user's profile picture is an anime-style character with a black cat hat. The post content includes a "Hidden Content" section that is locked and requires 8 credits to unlock. The text of the post states: "More about this breach on: - HavelBeenPwned [link] - BleepingComputer [link]. This is the same database that Troy Hunt has, including all the users, contacts, filenames and so on. There are 77,159,696 unique emails and the archive is around 14GB." Below the main text, it says "Dumped by @👑 ShinyHunters. Enjoy!". The user's profile information on the left shows they are a "MEMBER" with 9 posts, 1 thread, joined in Jan 2021, and a reputation of 152. At the bottom of the post, there are buttons for "PM", "Find", "Reply", "Quote", and "Report".

Figure 13 - Screenshot of **Spiral's** post offering for sale the GoNitro dump

Thus, we can suggest that **ExpertData** had access to the leaked GoNitro databases because of the connection with **ShinyHunters**.

On November 10, 2020, user **Chandler-Bing** published the **AnimalJam** database with a reference to **ShinyHunters** (the original post has been modified by the forum admins, and

now can only be found in the reply of another user).

AnimalJam Database - Leaked, Download!
by Chandler Bing - November 10, 2020 at 07:53 PM

Pages (7): < Previous 1 2 3 4 5 ... 7 Next >

Danny
November 15, 2020 at 03:53 AM #25

Chandler Bing Wrote: → (November 10, 2020 at 07:53 PM)

Animaljam.com

Credits to @**ShinyHunters** and @ThePolaris

```
[04.11.2020 01:45:13] <shinyhunters> Only you so far  
[04.11.2020 01:45:17] <shinyhunters> I won't sell again  
[04.11.2020 01:45:26] <shinyhunters> But I don't want to see these dumps leaked bro  
[04.11.2020 01:45:43] <shinyhunters> You can keep this collection but I think the agreement is on both sides
```

xx,xxx,xxx Entries - Dumped in 2020.

Sample (Username DB):

damn bruh thats awesome

New User

MEMBER

Posts 4
Threads 2
Joined Nov 2020
Reputation 0

Figure 14 - Commentary showing the offer to sell the base by **Chandler-Bing**

ExpertData then published the same database a day later, on November 11, 2020. This could also be a sign of a connection between **ExpertData** and **ShinyHunters**.

However, it is also possible that **ExpertData** just bought the databases and published them by himself. To confirm or reject our hypothesis about affiliation between **ExpertData** and **ShinyHunters**, we decided to analyze **ExpertData** account history.

It was revealed that back in March 2020 the user with the nickname **Expert** used the same Telegram account @expertdata (later in March 2020, the account was banned).

SELLING Selling 112k Spotify Premium Accounts
by Expert - March 14, 2020 at 08:48 PM

Thread Closed

Expert
March 14, 2020 at 08:48 PM #1

Come from high fresh private databases dehashed by my self.

Price: 100\$ BTC

telegram; @expertdata

This forum account is currently banned. Ban Length: Permanent (N/A).
Ban Reason: This user is suspected of scamming as when I tried to middleman a deal he denied to trade, a deal he offered here on site via a thread. <https://i.imgur.com/064SECe.png>

Banned

Posts 22
Threads 9
Joined Mar 2020
Reputation -20

1 YEAR OF SERVICE

PM Find Report

Figure 15 - Screenshot of **Expert's** post offering for sale a database

We compared messages of users **Expert** and **ExpertData** and it seems that they used a similar message template to publish the leaks, except for one case only — the database published on September 17, 2020 (fig.4). The **Expert** account also focused on selling databases. However, the quality of databases posted by **Expert** and **ExpertData** (created in September 2020) accounts is different. Most of **Expert's** databases were not authentic and represented combo lists of previously published data leaks. The changes in the quality of data offered for sale prompted us to think that the user might have started to work with some group in September 2020, presumably **ShinyHunters**.

We moved on with the messages of **ShinyHunters** and found another dispute on RaidForums, in which **ShinyHunters** was previously involved. On May 13, 2020, a user nicknamed **Jumbo** (account created on May 12, 2020) initiated a dispute against **ShinyHunters** and **fs0c131y/whysodank**. Jumbo claimed that he transferred 1.5 BTC to **ShinyHunters** for the databases put up for sale, but didn't get the purchase, with **ShinyHunters** denying the money receipt.

[BANNED] Scam Report against ShinyHunters and fs0c131y/whysodank | 13200\$
by Jumbo - May 13, 2020 at 10:56 AM

Pages (6): 1 2 3 4 5 6 Next »

Jumbo
May 13, 2020 at 10:56 AM

Name: (Of the user who scammed you, also please give a link to his / her profile or just his UID)
<https://raidforums.com/User-ShinyHunters>
<https://raidforums.com/User-fs0c131y>

Product: (Link to the thread / Tell us what product you bought / sold, if you can provide proof of purchase that'd be +1)
All of his databases for 1.5 BTC
<https://www.blockchain.com/btc/tx/fa57e7...f67bddbb1b>

How did you get scammed: (Very important, give a detailed description of how you got scammed, as well as screenshots of conversation / however you got scammed)
We agreed on 1.5 BTC for all of his databases. After I had sent him the money he made excuses that he never received anything and never sent me the databases
<https://ibb.co/HhsNnGn>
<https://ibb.co/8z31trY>
<https://ibb.co/1QMfVcz>
<https://ibb.co/QYQes6K>
<https://ibb.co/2YY16qF>
<https://ibb.co/jZK7D9q>
<https://ibb.co/9YtBMd5>
<https://ibb.co/h1syvSP>
<https://ibb.co/M2wgRtC>
<https://ibb.co/47dV1ZW>
<https://ibb.co/K6WsnFC>

Time of scam: (Tell us approximately when you got scammed)
About 10 hours ago

This forum account is currently banned. Ban Length: Permanent (N/A).
Ban Reason: Submitting a fake scam report.

Figure 16 - Screenshot of a dispute talk initiated by **Jumbo** against **ShinyHunters**

The account **whysodank**, mentioned in a dispute, was created back in April 2020.

whysodankk
M.V.P User
Status: Offline (Last Visit: July 13, 2020 at 12:32 PM)

Joined: April 23, 2020
Time Spent Online: 1 Week, 5 Days, 17 Hours
User Identifier: 121832784 [Copy Profile Permalink]
Username Changes: 2
Members Referred: 36

whysodankk's Contact Details
Private Message: Send whysodankk a private message.

whysodankk's Forum Statistics
Total Threads: 3 (0 threads per day | 0 percent of total threads) Find All Threads
Total Posts: 6 (0.01 posts per day | 0 percent of total posts) Find All Posts
Reputation: 224 Details Rate

whysodankk's awards.
May 16, 2020 at 07:45 PM

New Reply
RE: Corporate Email Lists | Only the Freshest in Leaks Market
July 08, 2020 at 04:32 PM 26

New Reply
RE: [List] 2020 Test and review of popular leak databases in Giveaways & Freebies
July 07, 2020 at 03:40 PM 6

Figure 17 - *Whysodankk* profile on RaidForums

The history of changing usernames in the profile is provided below.

Username History for whysodankk	
Old Username	Date Changed
fs0c131y	May 16, 2020 at 09:13 PM
whysodank	May 06, 2020 at 08:54 PM

Figure 18 - Screenshot of previous nicknames of the *whysodank* account

A user with the nickname **whysodankk** joined the dispute and refuted **Jumbo's** claims saying that the screenshots allegedly confirming the money transfer in fact represented the transfer of money between two wallets belonging to **Jumbo**. **Whysodankk** also confirmed that he knew **ShinyHunters** and supported his position in the dispute.

Jumbo participated in only one thread on the forum — with accusations against **ShinyHunters**. It is noteworthy that it was **whysodankk** who was mainly involved in the discussion, and not **ShinyHunters**, against whom the charges were brought as well.

The dispute ended on May 14, 2020. **Jumbo** was banned from the forum over a fake scam report which shows that he didn't provide evidence in support of his accusations against **ShinyHunters**.

From the above posts it can be concluded that **whysodankk** is either a partner of **ShinyHunters** or just a different account of the latter. To check this, we decided to find accounts on the RaidForums whose contact details overlap with **ShinyHunters'** ones.

It turns out that **ShinyHunters** and **whysodankk** specified the same contact details — shinyhunters@xmpp.jp. Examples of the posts are presented below.

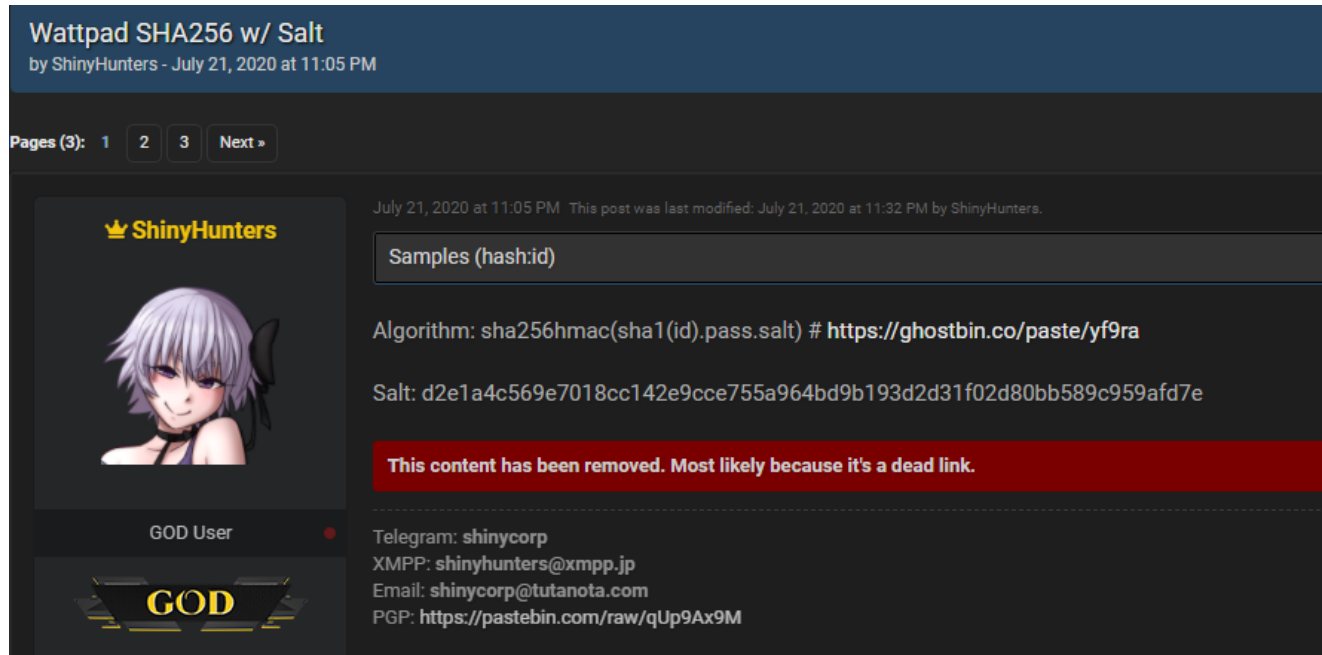


Figure 23 - Screenshot of **ShinyHunters'** post with Jabber account shinyhunters@xmpp.jp

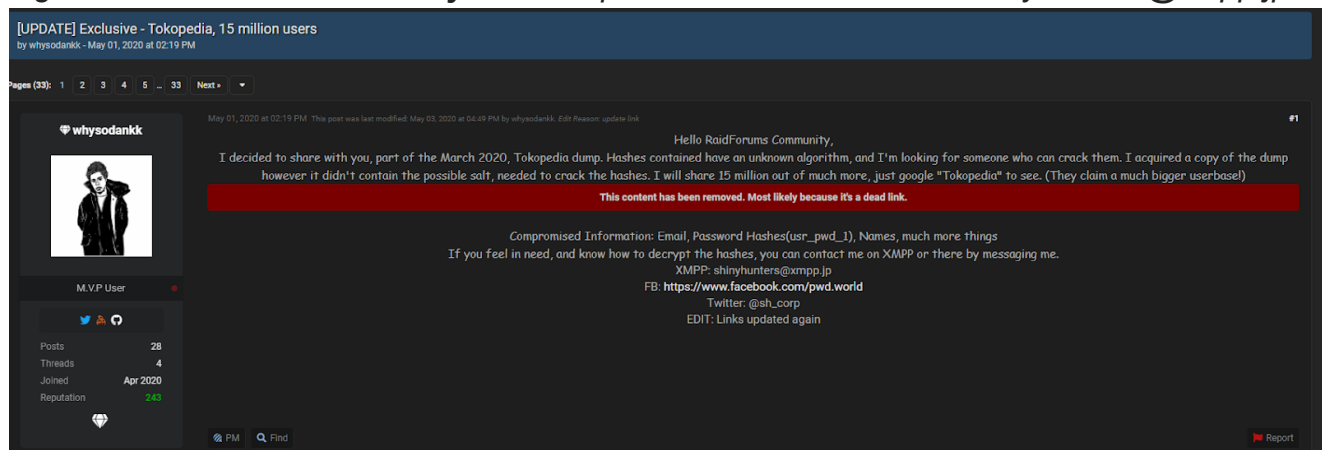


Figure 24 - Screenshot of **whysodankk's** post with Jabber account shinyhunters@xmpp.jp (account **whysodankk** was used to sell the Tokopedia database back in May 2020. The post on the screenshot contains the first part of the dump, while the full dump of Tokopedia was published by this account later)

The same contact details indicate that **ShinyHunters** and **whysodankk** are related to each other. Moreover, these two accounts supported each other in the dispute.

The same ready-to-help strategy was used by **ShinyHunters** and **ExpertData**, which leads us to a conclusion that **ExpertData** might be a member or a data broker of **ShinyHunters**

hacker group, and the databases trafficked by **ExpertData** could have belonged to **ShinyHunters**.

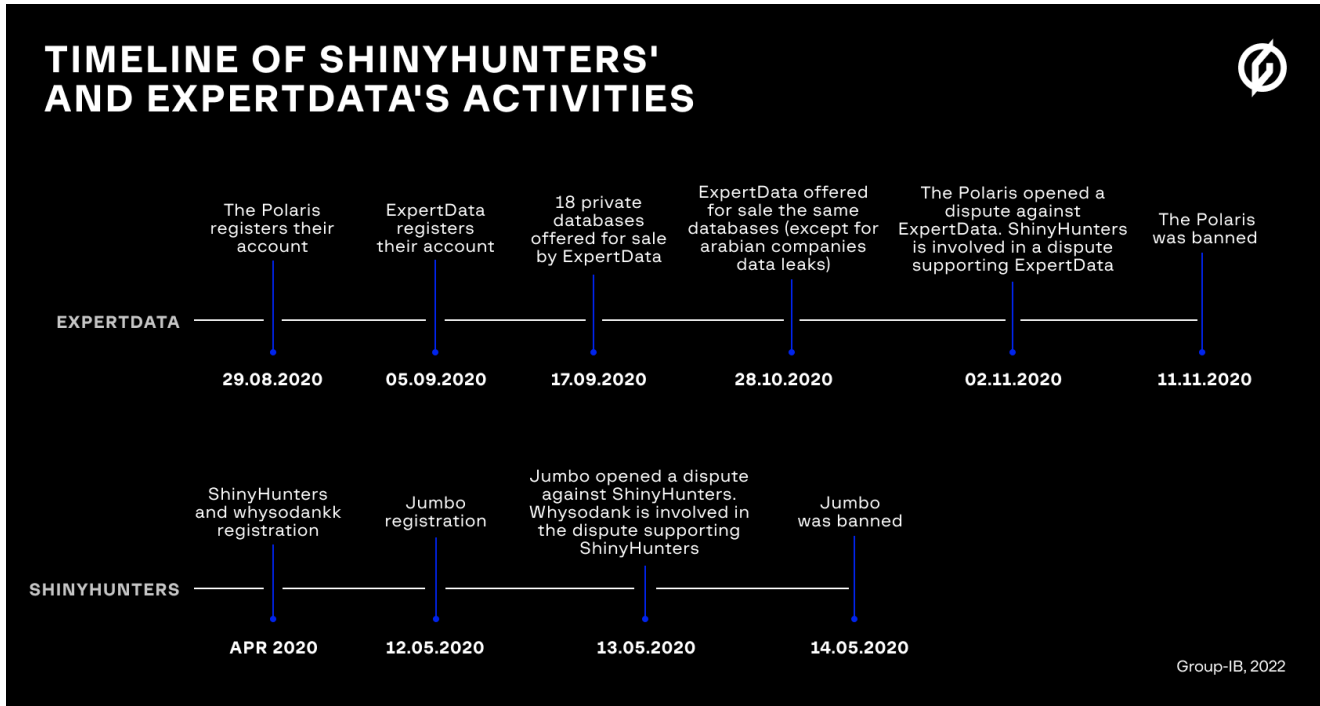



Figure 25 - Timeline of disputes against **ShinyHunters** and **ExpertData**

Additional confirmation of this assumption can be found below.



[BANNED] Shinyhunters and ExpertData SCAM 40k\$
 by The Polaris - November 10, 2020 at 06:29 PM

The Polaris



Banned

Posts: 11
 Threads: 1
 Joined: Aug 2020
 Reputation: 11

November 10, 2020 at 06:29 PM This post was last modified: November 11, 2020 at 01:06 AM by Omnipotent. #1

I contacted Shinyhunters a week ago and offered to buy his dumps from him, I paid him a total of \$ 40k +, I don't remember exactly, not only did each dump contain 2-3 lines of lines than he wrote to me. HE promised NEVER to distribute it, and in a week it manifests itself in the sale from his friend ExpertData, I will also show you messages where he says that ExpertData is his friend, although in the previous scam report shiny wrote that he does not know him, so I demand to return it to me spent \$ 35k for promising never to distribute it, or he should get banned for selling garbage, lying about the number of users, and distributing it when he promised to NEVER do it.

<https://ibb.co/bXMg0Tv>
<https://ibb.co/SxZ7WVQ>
 <Redacted>
<https://ibb.co/dWBtd3Q>
<https://ibb.co/MGwDyVX>
<https://ibb.co/2k8ZBqX>
<https://ibb.co/dgt3yvg>
 <Redacted>
<https://ibb.co/Xy2QJ9T>
<https://ibb.co/p1k2R1G>
<https://ibb.co/MBnZnxw>
<https://ibb.co/ggx0Vsk>
<https://ibb.co/L89jH4Z>

I resigned myself to the fact that he sold me half of the garbage in which there were not even half of the lines, but I will not let this asshole sell it either when he took my \$ 40,000 and promised not to sell it

I demand the blocking of these fagots or a refund of the \$40k that I spent

This forum account is currently banned. Ban Length: Permanent (N/A).
 Ban Reason: Failed to abide by terms of a deal. If you have any questions please contact @Omnipotent.
 polaris1000@protonmail.com 45.9.236.13 2a0f:df00:0:255::74

*Figure 26 - Screenshot of a post by **The Polaris** who alleged friendly links between users **ShinyHunters** and **ExpertData***

The table below shows information about contact details of the accounts revealed during the analysis.

The above analysis revealed ties between accounts **ExpertData**, **ShinyHunters**, and **whysodankk (aka whysodank, fs0c131y)**. The links between **ShinyHunters** and other groups are out of the scope of this article.

Going back to our primary task to understand the quality of databases and the level of threat actor for the post of **ExpertData** on October 28, 2020, we can make the following conclusions:

1. **The databases are most probably authentic.** No other posts with the same databases published earlier by other users were revealed. During the analysis links between **ExpertData** and **ShinyHunters** were uncovered. Several databases published by **ExpertData** were most probably exfiltrated by **ShinyHunters** and subsequently transferred to **ExpertData**. Publishing of unique data leaks is typical for **ShinyHunters** group members.
2. **It is supposed that ShinyHunters was behind these leaks.** **ShinyHunters** is a famous hacker group with a strong background and skills. It is assumed that the group consists of several people, therefore they can publish new leaks using different accounts. If our conclusion is true, then the databases were obtained as a result of the affected companies' compromise. In such situations it is highly recommended to start incident response and investigation as soon as possible.

After the link between **ExpertData** and **ShinyHunters** was uncovered, we decided to further analyze **ShinyHunters'** relation to other members of the forum and find other accounts that can potentially be linked to this threat actor. However, to find other links to **ShinyHunters**, we would have had to analyze around a hundred messages of **ShinyHunters** and a bunch of posts belonging to their potential members. To simplify the research, we decided to use ML-algorithms based on semantic text analysis.

Semantic analysis

The first part of the article represented the manual analysis of the attackers' profiles and posts on the underground forum. Group-IB, however, is trying to automate its methodologies of cyber investigations and resorts to ML-algorithms to analyze dark web and build correlations between various posts in the underground.

We, therefore, decided to check the same range of data utilizing Natural Language Processing (hereinafter - NLP) algorithms to reveal multiaccounts of the same person or various accounts of the same group of people. The analysis was carried out based on the

assumption that the semantics of messages from the same person or sometimes even members of the same group should be similar.

The sequence of data collection and processing is shown below:

- The collection of all posts from the RaidForums (topic, message, nickname, date and time);
- The filtering of messages with a minimum length of less than 100 characters to exclude semantically insignificant texts;
- The vector representation of texts obtained using the BERT neural network model (12-layer, 768-hidden, 12-heads, 110M parameters).

For further analysis, all message vectors (N) belonging to user **ShinyHunters** were selected. For each such vector, the top 10 nearest vectors of other users' messages were found (by the Euclidean distance). As a result, we obtained a matrix of dimension $N*10$ (example of the matrix is in the table below), with the nicknames of published messages as well as its elements. Then we filtered out repeated messages and messages whose Euclidean distance exceeds the specified threshold. The threshold was calculated empirically by comparing the proximity of different messages.

As a result, we received a list of messages semantically close to the messages of **ShinyHunters**. All the messages were grouped by the users who published them (**matches**). Users who have only one semantically close message (1 match) were excluded from the subsequent analysis to decrease false correlations. The last step was averaging the Euclidean distance across all messages for each of the users. The resulting values will be the **scores** we were looking for.

It is important to note that this metric is not symmetric: if, for example, a User A account analysis determined their link to User B, such a link might not be found as a result of the analysis of User B. This is due to the fact that we have limited the search to the top 10 nearest vectors for each user message. Surely, such sets of the nearest vectors, even for semantically close messages, may differ. For this reason, the similar analysis for the found users may not display the link with **ShinyHunters**.



ShinyHunters	Distances to the semantically closest messages (the User who published the message)			
	1	2	...	10
Message 1	0.1 (User_1)	1.6 (User_2)	...	3.3 (User_3)
Message 2	1.1 (User_3)	2.1 (User_4)	...	2.5 (User_5)
Message 3	0.4 (User_2)	1.7 (User_6)	...	2.2 (User_3)
...
Message N	0.25 (User_1)	1.4 (User_6)	...	2.1 (User_4)

Group-IB, 2022

Here is an example of calculation for User_3, based on the results of the first 3 messages in the table above:

$$S = (3.3 + 1.1 + 2.2) / 3 = 2.2$$

Matches (semantically close messages): 3

Average distance (**score**): 2.2

The results of the analysis of the ShinyHunters messages are presented in fig. 27 (the total number of user posts at the time of writing is indicated in square brackets).

```

===== ShinyHunters [118] =====
ShinyHunters      score: 0.0      matches: 118
fs0c131y          score: 2.452   matches: 5
MyBiggyBruteBolt score: 2.564   matches: 3
J4ckd0x           score: 2.794   matches: 2
Megadimarus       score: 2.948   matches: 2
johnlockejrr      score: 2.972   matches: 5
Omnipotent        score: 3.055   matches: 2
Troy Hunt         score: 3.194   matches: 2
Databases          score: 3.228   matches: 4
FluffyBunnyFufu  score: 3.26    matches: 2

```

Figure 27 - The result of the analysis of **ShinyHunters** posts

The analysis showed that accounts **fs0c131y**, **MyBiggyBruteBolt**, **J4ckd0x**, **Megadimarus**, **johnlockejrr**, **Omnipotent**, **Troy Hunt**, **Databases** and **FluffyBunnyFufu** published several messages that are semantically close to **ShinyHunters**' posts.

In fig.27, we see that the number of semantically close messages (matches) between ShinyHunters and other users is small, so it is incorrect to unequivocally state that there is a connection between them. However, the values of this metric give us a reason to take a closer look at possible relationships.

It is noteworthy that the semantic analysis doesn't reveal the strong link between **ExpertData** and **ShinyHunters** accounts. Below is the metric for the **ExpertData** account.

```
===== ExpertData [15] =====
ExpertData      score: 0.0      matches: 15
DrDastan        score: 3.825    matches: 2
The404          score: 3.914    matches: 2
ambins          score: 4.09     matches: 2
killerstr       score: 4.116    matches: 3
Proto           score: 4.399    matches: 2
Momondo         score: 4.468    matches: 2
DonJuji         score: 4.924    matches: 3
ShinyHunters    score: 6.238    matches: 2
```

Figure 28 - The result of the analysis of **ExpertData** posts

The distance to **ShinyHunters** messages is too high, which can be explained by three assumptions:

1. **ExpertData** is a data broker who worked for this group. Big groups usually change data brokers quite often and don't interact with them closely. That is why the format of the messages of a data broker could be different from **ShinyHunters** posts.
2. **ExpertData** has few posts;
3. Semantic analysis may not work for dialogues in which each message carries its own semantic meaning. Let's say that threat actors might discuss the same topic, but each of the messages in this thread will have their unique idea, and semantic analysis will not be able to reveal links between these separate messages.

As you can see from fig.27, there is a link between **ShinyHunters** and **fs0c131y** (aka **whysodankk**) accounts. The link was earlier shown in the part of manual analysis and confirmed by semantic analysis.

===== fs0c131y [17] =====		
fs0c131y	score: 0.0	matches: 17
whysodankk	score: 1.164	matches: 2
ShinyHunters	score: 2.206	matches: 5
johnlockejrr	score: 2.928	matches: 2
MyBiggyBruteBolt	score: 2.984	matches: 3
===== whysodankk [5] =====		
whysodankk	score: 0.0	matches: 5
fs0c131y	score: 1.164	matches: 2

Figure 29 - The result of the semantic analysis of **fs0c131y** and **whysodankk** posts

As you can see different nicknames of the same account are identified by the algorithm.

The link of **ShinyHunters** with users **MyBiggyBruteBolt** and **johnlockejrr** was found, because format of the messages is similar, however most of the posts contain the same note: "This forum account currently has an ongoing scam report, please beware trading. Details: Please respond to this scam report."

ShinyHunters	MyBiggyBruteBolt
It's good, I edited the link This forum account currently has an ongoing scam report, please beware trading.	Yes, really great stuff for sell This forum account currently has an ongoing scam report, please beware trading.
What are you talking about? This forum account currently has an ongoing scam report, please beware trading.	Why did you post this message three times? This forum account currently has an ongoing scam report, please beware trading.
Strange, no replies anymore from the user. Damn photoshop This forum account currently has an ongoing scam report, please beware trading.	try to check, maybe it steel valid material, tnx This forum account currently has an ongoing scam report, please beware trading.

Figure 30 - Semantically close posts of **ShinyHunters/MyBiggyBruteBolt** users

ShinyHunters	johnlockejrr
Now I understand why you wanted a private message on raidforums so bad, to trick me with a transaction to an address of yours This forum account currently has an ongoing scam report, please beware trading.	Here is the real scammer I'm having problems because of him This forum account currently has an ongoing scam report, please beware trading.
Thanks a lot brother, I forgot we could do that, I don't know who you are but you really helped me a lot This forum account currently has an ongoing scam report, please beware trading.	This is a frame up, never thought this could happen. I never scammed anyone in 3 years This forum account currently has an ongoing scam report, please beware trading.
Strange, no replies anymore from the user. Damn photoshop This forum account currently has an ongoing scam report, please beware trading.	Bro, you messed up the hide tag This forum account currently has an ongoing scam report, please beware trading.
You're a liar This forum account currently has an ongoing scam report, please beware trading.	I talked to Omnipotent about this issue, I take full responsibility for the fact my account was hijacked This forum account currently has an ongoing scam report, please beware trading.
What are you talking about? This forum account currently has an ongoing scam report, please beware trading.	What year is this leak? Can you post a sample? This forum account currently has an ongoing scam report, please beware trading.

Figure 31 - Semantically close posts of **ShinyHunters/johnlockejrr** users

Metrics for **MyBiggyBruteBolt** and **johnlockejrr** are presented below.

```

===== MyBiggyBruteBolt [20] =====
MyBiggyBruteBolt    score: 0.0           matches: 20
Leftyy              score: 2.266         matches: 2
Diamondeye          score: 2.322         matches: 4
Asset               score: 2.343         matches: 2
Lvl3-Noob           score: 2.405         matches: 2
DaveCrouse          score: 2.414         matches: 3
BlockDev            score: 2.462         matches: 3
senseye             score: 2.533         matches: 2
Alpaca              score: 2.543         matches: 5
TRGBeast            score: 2.568         matches: 4
abels               score: 2.62          matches: 2

```

Figure 32 - The result of the analysis of *MyBiggyBruteBolt* posts

```

===== johnlockejrr [129] =====
johnlockejrr        score: 0.0           matches: 129
n1js                 score: 1.705         matches: 4
imlesor             score: 2.185         matches: 2
GreenHat77          score: 2.306         matches: 3
noname888           score: 2.367         matches: 2
ONYYXX              score: 2.395         matches: 3
xD1ous77            score: 2.712         matches: 2
brokoli777          score: 2.773         matches: 4
Raid-Kingdom        score: 2.814         matches: 5
megalos66           score: 3.009         matches: 2
2020                score: 3.013         matches: 3

```

Figure 33 - The result of the analysis of *johnlockejrr* posts

In our opinion, these accounts are not relevant to **ShinyHunters** and could be considered as false positives.

According to fig.27, there is a link between **J4ckd0x** and **ShinyHunters** accounts. The account **J4ckd0x** was created back in 2016.

Figure 34 - J4ckd0x account on RaidForum

According to the content of J4ckd0x's messages, his interests lie mostly in selling databases. The metric for this account is presented below.

```

===== J4ckd0x [33] =====
J4ckd0x          score: 0.0          matches: 33
theb3ard         score: 2.447        matches: 2
ShinyHunters    score: 2.794        matches: 2
yametal03       score: 2.894        matches: 2
Malkian         score: 2.999        matches: 2
Omnipotent      score: 3.164        matches: 2
Peas            score: 3.194        matches: 2
irakliss        score: 3.318        matches: 2
  
```

Figure 35 - The result of the analysis of J4ckd0x's posts

On the screenshot below, you can see that the format of J4ckd0x's posts offering databases for sale is similar to those of ShinyHunters.

ShinyHunters	J4ckd0x
Hello RaidForums Community,\nToday I have uploaded the GGumim.co.kr Database for you to download for free, thanks for reading and enjoy!\nNotes\n\n March 2020, the Korean interior decoration website 집꾸미기 (Decorating the House) suffered a data b...	Hello RaidForums Community,\nToday I have uploaded the Dunzo Database for you to download for free, thanks for reading and enjoy!\nNotes\n\n approximately June 2019, the Indian delivery service Dunzo suffered a data breach. Exposing 3.5 million...
Hello RaidForums Community,\nToday I have uploaded the Sonicbids Database for you to download for free, thanks for reading and enjoy!\nNotes\n\n December 2019, the booking website Sonicbids suffered a data breach which they attributed to "a dat...	Hello RaidForums Community,\nToday I have uploaded the [redacted] Database for you to download for free, thanks for reading and enjoy!\nNotes\n\n December 2019 a file allegedly belonging to an online [redacted], surfaced on...

Figure 36 - Semantic proximity in posts of **J4ckd0x** and **ShinyHunters**

There are two matches in messages of **J4ckd0x** and **ShinyHunters** that relate to two databases. **J4ckd0x** reposted (September 2020) one database that was first published by **ShinyHunters** (April 2020). Dunzo database was published by **ShinyHunters** (July 2020) before the post of **J4ckd0x** (September 2020). The format of the message posted by **ShinyHunters** about Sonicbids database is similar to the format of **J4ckd0x**'s post about another leakage. The format of the message posted by **ShinyHunters** about GGumim.com.kr database is similar to the format of the **J4ckd0x**'s post about the Dunzo leakage.

ShinyHunters

```
"Hello RaidForums Community,\nToday I have uploaded the GGumim.co.kr Database for you to download for free, thanks for reading and enjoy!\nNotes|\nIn March 2020, the Korean interior decoration website 집꾸미기 (Decorating the House) suffered a data breach which impacted almost 1.3 million members. Served via the URL ggumim.co.kr, the exposed data included email addresses, names, u sernames and phone numbers, all of which was subsequently shared extensively throughout online hacking communities.\nCompromise d data: Usernames, Email Addresses, Phone Numbers, Socialmedia Links, Passwords, Names\nContentsSpoilerThis download consists o f 1 .SQL file, the passwords in this file are hashed using SHA-256 please look below for data schema. The .7z File's MD5 Hash i s 654A2BD3B2FD820926EF68AEF39F7DA6. Total record count of 1298651.idx,old_idx,grade_idx,special_grade_idx,crema_user_id,point,n ame,nickname,password,password_secure,email,phone_number,access_permission,kakao_token,facebook_token,google_token,naver_token, account_type,is_old,website,signup_platform,recent_profile,recent_shout,recent_cover,push_setting_star,push_setting_comment,pus h_setting_news,check_privacy_policy,check_term,first_buy_date,buy_count,is_active,force_logout,is_confirm,last_login_date_time, created_date_time,status,api_token\nDownloads\nHidden Content\nUnlock for 8 credits\nDatabase Index <> How To Get Credits"
```

...

J4ckd0x

```
"Hello RaidForums Community,\nToday I have uploaded the Dunzo Database for you to download for free, thanks for reading and enjoy!\nNotes|\nIn approximately June 2019, the Indian delivery service Dunzo suffered a data breach. Exposing 3.5 million unique email addresses, the Dunzo breach also included names, phone numbers and IP addresses which were all broadly distributed online via a hacking forum.\nCompromised data: Device information, Email addresses, Geographic locations, IP addresses, Names, Phone n umbers\nContentsSpoilerThis download consists of 1 .SQL file, this dump does not contain any passwords and we have included the data schema below. The .7z File's MD5 Hash is AB56EB71F6AD5209DE41C2B877605AE9. Total record count of 3465259.id,password,last_ login,is_superuser,uuid,first_name,last_name,email,phone,country_code,type,status,device_token,phone_type,phone_make,date_joine d,last_updated,secret_key,app_version,registered_on,registered_platform,send_logistics_pricing,send_logistics_pricing_image_for mat,last_pricing_version_shared,preferred_mode_of_payment,credit_amount,credit_score,maximum_retries_count,profile_data_updated _on_firebase,merchant_id,permission_role,user_status,flow_version,extra_data_json,city_id,current_runner_task_id,source,first_k nown_location,last_known_location,referral_code,referred_by_code,advertising_id,device_id,bucket_id\nDownloads\nHidden Content\nUnlock for 8 credits\nDatabase Index <> How To Get Credits"
```

Figure 37 - Posts published by **J4ckd0x** and **ShinyHunters** and written using the same format

Based on the foregoing, we concluded that accounts **ShinyHunters** and **J4ckd0x** could belong to the same group or at least be partners.

According to fig.27, there is a link between **ShinyHunters** and **Megadimarus** accounts. **Megadimarus** account was created on May 4, 2020 on the RaidForums and is banned at the time of writing.

Nassim Benhaddou
Prosox
Status: Offline (Last Visit: July 27, 2020 at 05:26 PM)

This forum account is currently banned.
Ban Reason: self-ban
Banned By: Jaw – Ban Length: Permanent (N/A remaining)

Nassim Benhaddou's Forum Info
Joined: May 04, 2020
Time Spent Online: 2 Weeks, 4 Days, 6 Hours
User Identifier: 121843961 [Copy Profile Permalink]
Username Changes: 1
Members Referred: 5

Nassim Benhaddou's Contact Details
Skype Username: Nassim Benhaddou
Discord Handle: Nassim Benhaddou

Nassim Benhaddou's Forum Statistics
Total Threads: 1 (0 threads per day | 0 percent of total threads) Find All Threads
Total Posts: 1 (0 posts per day | 0 percent of total posts) Find All Posts
Reputation: 3,220 Details Rate

Nassim Benhaddou's awards.
Apple Award: June 01, 2020 at 11:15 PM
Crown Award: May 04, 2020 at 10:45 PM

Additional Info About Nassim Benhaddou
Bio: My name is Nassim Benhaddou. But people call me Prosox.

New Thread
Petflow Database - Leaked, Download!
in Official
May 05, 2020 at 06:41 PM 24

Figure 38 - *Megadimarus* profile on RaidForums

The metric for **Megadimarius** is presented below and is symmetric with **ShinyHunters**.

===== Megadimarus [157] =====		
Megadimarus	score: 0.0	matches: 157
Databases	score: 2.892	matches: 2
ShinyHunters	score: 2.948	matches: 2

Figure 39 - The result of the analysis of *Megadimarus* posts

Most of his messages were removed from the visible content of the forums however thanks to the capabilities of Group-IB's proprietary products the removed contents are preserved and analyzed as well. Most of the **Megadimarus** posts are related to selling databases. Moreover, some of them are associated with the activities of **ShinyHunters** predecessors. In the report of cybersecurity researcher Vinny Troia the relationship between **ShinyHunters** and **Megadimarus** was also shown.

The metric shows a similarity of posts' format for **ShinyHunters** and **Megadimarus**. Moreover, in this particular case every match is related to the same database (**Appen.com** and **HomeChef** leakages).

```

ShinyHunters
'Description: HomeChef is a unique lifestyle network that connects viewers to the power and joy of food.\nDate: 02/2020\nFieldsSpoile
r'id","email","encrypted_password","reset_password_token","reset_password_sent_at","remember_created_at","sign_in_count","current_sign_
in_at","last_sign_in_at","current_sign_in_ip","last_sign_in_ip","created_at","updated_at","name","customer_id","last_4_digits","provide
r","uid","status","servings","confirmation_token","confirmed_at","confirmation_sent_at","unconfirmed_email","phone","delivery_day","cul
inary_level","agreed_to_terms","discovery","weekly_meals","age","gender","region","relationship","campaign_id","promotion_type","min_ag
e","max_age","behavior","interest","optional_1","optional_2","optional_3","optional_4","optional_5","active","signup_redemption_id","zi
p_code_id","completed_signup_at","current_sign_in_platform","last_sign_in_platform","web_sign_in_at","mobile_sign_in_at","experiment_da
ta","preference_data","meal_plan_id","vendor","paypal_email","terms_accepted_at","shipping_cost_cents","accepted_agreements","uuid","br
and_id","monthly_credit"\nAlgorithm: BCRYPT ($2a$10$)\nRow count: 8717763\n[Download]\nHidden Content\nUnlock for 8 credits\nPrivate da
tabases for sell.\nXMPP: shinyhunters@xmpp.jp'
...
Megadimarus
'Selling HomeChef.com database. HomeChef is a meal delivery website.\nDate: 2020\nUsers: 8 million\nData: Email addresses, IP addresse
s, Names, Social media profiles, Phone numbers, Passwords\nPasswords: BCRYPT\nTable structure:\n'id","email","encrypted_password","rese
t_password_token","reset_password_sent_at","remember_created_at","sign_in_count","current_sign_in_at","last_sign_in_at","current_sign_i
n_ip","last_sign_in_ip","created_at","updated_at","name","customer_id","last_4_digits","provider","uid","status","servings","confirmati
on_token","confirmed_at","confirmation_sent_at","unconfirmed_email","phone","delivery_day","culinary_level","agreed_to_terms","discover
y","weekly_meals","age","gender","region","relationship","campaign_id","promotion_type","min_age","max_age","behavior","interest","opti
onal_1","optional_2","optional_3","optional_4","optional_5","active","signup_redemption_id","zip_code_id","completed_signup_at","curren
t_sign_in_platform","last_sign_in_platform","web_sign_in_at","mobile_sign_in_at","experiment_data","preference_data","meal_plan_id","ve
ndor","paypal_email","terms_accepted_at","shipping_cost_cents","accepted_agreements","uuid","brand_id","monthly_credit"\nData sample:\n
2623210,soldierwife075@aol.com,$2a$10$zc2zIrJxzijt.Vh8je0INeKxZqnk0.g.qfts0bkk4jQNccvkT03yq,,2,2018-06-25 17:22:22,2017-06-10 00:35:5
0,"2606:a000:140c:4271:0:3ef0:a93a:b743","2601:0603:0080:995c:a053:fe95:db0d:6382",2017-06-10 00:35:50,2018-06-25 17:22:35,Mary Coble,c

```

Figure 40 - Semantically close posts of **ShinyHunters** and **Megadimarus**

Based on the analysis of **Megadimarius's** posts and interests we concluded that there could be a link between this account and **ShinyHunters**. It could be members of the same group or they could be partners.

Another account for the analysis from fig.27 is **Omnipotent**. It is one of the administrator accounts created back in 2015.

Figure 41 - **Omnipotent** profile on **RaidForums**

One of the matches between **ShinyHunters** and **Omnipotent** is due to the following post.

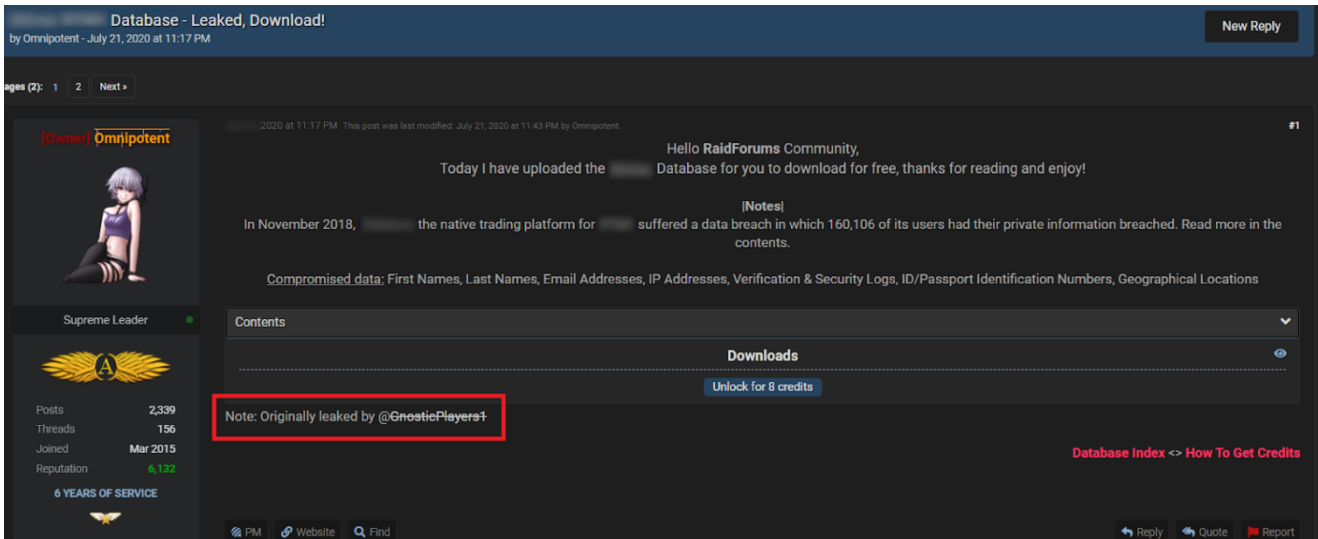


Figure 42 - Post of **Omnipotent** about an uploaded database

It was reposted by **Omnipotent** from **GnosticPlayers1** account which is now banned. The format is similar to a format of **ShinyHunters** post about selling databases.

GnosticPlayers1 interest lies in selling databases and this nickname is similar to the name of the group Gnostic Players.

Two messages of **Omnipotent** have similar format with the format of **ShinyHunter's** posts. Since **Omnipotent** account has thousands of messages, and only one (excluding the reposted one) having the format similar to **ShinyHunters**, as well as knowing that **Omnipotent** is the forum administrator, we consider this match as a false positive.

The metric for **Omnipotent** is presented below.

Username	score	matches
Omnipotent	0.0	710
Mimi	0.0	4
BoringApe	0.0	3
RawCat	0.128	6
Istrator	0.154	5
raidflacs2k	0.193	4
Mr-Clair	0.385	2
September	0.5	7
2a-45	0.516	5
September-11-2001	0.546	5
2a45	0.558	6

Figure 43 - The result of the analysis of **Omnipotent** posts

Another account from fig.27 is **Troy Hunt**. This account was created in June 2019 and impersonated the name and photo of the famous cybersecurity researcher Troy Hunt.

Figure 44 - Screenshot of **Troy Hunt** (aka **Jnx3cx**) account

The history of the account's nickname change is presented below.

Username History for Troy Hunt	
Old Username	Date Changed
Jnx3cx	November 17, 2019 at 09:32 PM

Figure 45 - Username history for **Troy Hunt/Jnx3cx** account

The main interest of the account is selling databases some of which are associated with the activities of **ShinyHunters** predecessors. The format of the messages is also correlated with the one of **ShinyHunters**.

ShinyHunters

"Hello RaidForums Community,\nToday I have uploaded the Swvl Database for you to download for free, thanks for reading and enjoy!\n\nNotes|\nIn June 2020, the Egyptian bus operator Swvl suffered a data breach which impacted over 4 million members of the service. The exposed data included names, email addresses, phone numbers, profile photos, partial credit card data (type and last 4 digits) and passwords stored as bcrypt hashes, all of which was subsequently shared extensively throughout online hacking communities.\nCompromised data: Email addresses, Names, Partial credit card data, Passwords, Phone numbers, Profile photos\nContentsSpoilerThis download consists of 1 .JSON file, this dump's passwords are hashed using bcrypt. The .7z File's MD5 Hash is EB868C273A5E4794209D7840E8F95589. Total record count of 4195918.\nDownloads\nHidden Content\nUnlock for 8 credits\nDatabase Index <> How To Get Credits"

...

Troy Hunt

"Hello RaidForums Community,\nToday I have uploaded the Zynga Database for you to download for free, thanks for reading and enjoy!\n\nNotes|\nIn September 2019, one of the world's most successful social game developer Zynga suffered a data breach that exposed 217M accounts. The compromised data included email addresses, names, usernames, phone numbers, and passwords stored as salted SHA1 hashes.\nCompromised data: Email addresses, Names, Usernames, Phone Numbers and Passwords\nContentsSpoilerThis dump uses SHA1DASH hashing furthermore the data is stored in full SQL format. The .7z File's MD5 Hash is C0773FF2C6B8ED0AB7CD8ABA1C2067A2. Total record count of 216595850.\nHidden Content\nUnlock for 8 credits\nDatabase Index <> How To Get Credits"

Figure 46 - Posts published by **Troy Hunt** and **ShinyHunters** and written using the same format

The algorithm shows the similarity of messages of **ShinyHunters** with the posts of **Troy Hunt** about selling Canva and Zynga (now deleted from the visible contents of RaidForums). Moreover, **Troy Hunt** posted a message on behalf of Gnostic Players (is supposed to be a predecessor of **ShinyHunters** according to the research of Vinny Troia).

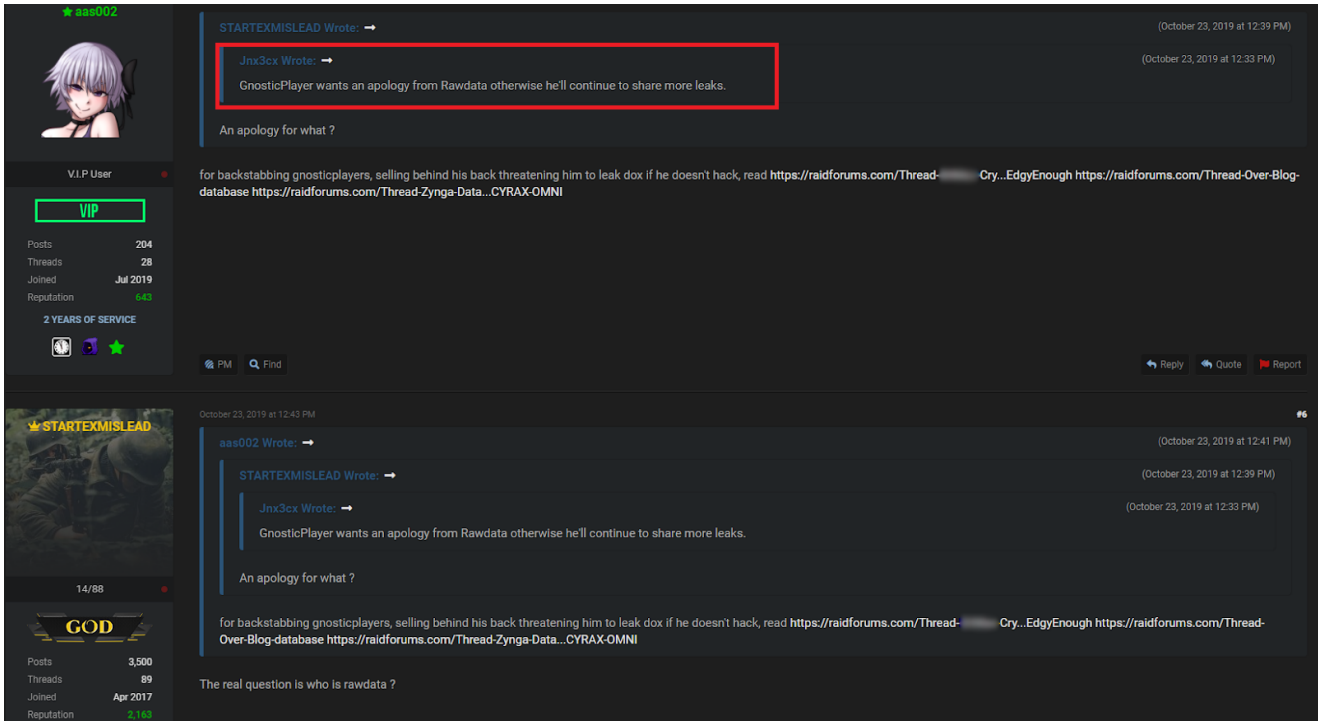


Figure 47 - Replies to the deleted post of **Troy Hunt** with a reference to **GnosticPlayers**

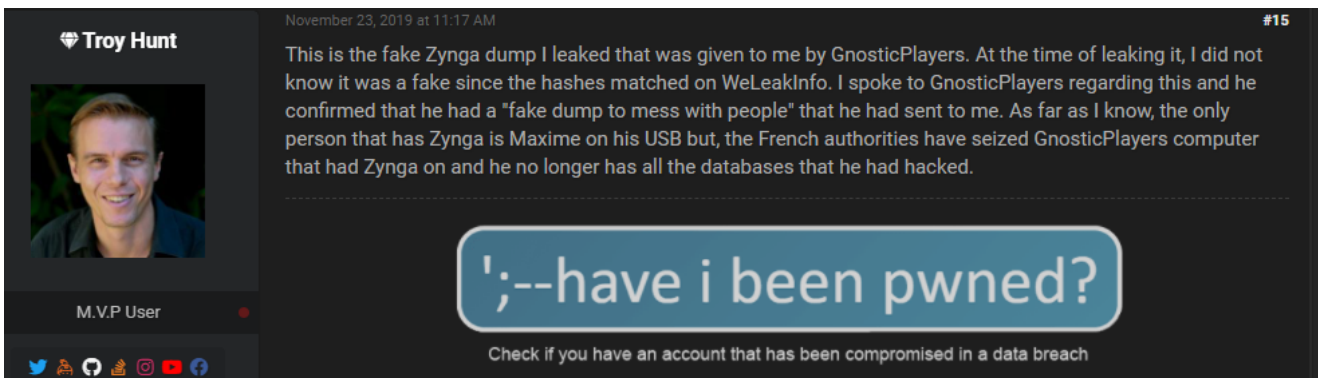


Figure 48 - Post of **Troy Hunt** with a reference to **GnosticPlayers** and **Zynga** database

The metric for the **Troy Hunt** account is presented below.

```

===== Troy Hunt [104] =====
Troy Hunt          score: 0.0         matches: 104
Jnx3cx             score: 0.437      matches: 39
0day               score: 2.413      matches: 4
albialbi          score: 2.509      matches: 4
canifani           score: 2.51        matches: 6
DonJuji           score: 2.903      matches: 5
Databases          score: 2.981      matches: 10
valentin0         score: 3.131      matches: 3
Omnipotent        score: 3.194      matches: 3
ShinyHunters      score: 3.22        matches: 2
Yaccin            score: 3.226      matches: 2

```

Figure 49 - The result of the analysis of **Troy Hunt** posts

```

===== Jnx3cx [106] =====
Jnx3cx            score: 0.0         matches: 106
Troy Hunt         score: 0.437      matches: 39
DonJuji           score: 2.955      matches: 4
Databases          score: 3.106      matches: 8
valentin0         score: 3.122      matches: 2
Omnipotent        score: 3.194      matches: 3

```

Figure 50 - The result of the analysis of **Jnx3cx** (aka **Troy Hunt**) posts

After analyzing the **TroyHunt** account, his posts and interests, we concluded that its link with **ShinyHunters** could be relevant meaning the accounts could belong to the same group or at least be partners.

Next account from fig.27 is **Databases** created back in 2015.

Databases
GOD User
Status: Offline (Last Visit: October 25, 2021 at 03:54 AM)

[Add to Buddy List](#) [Add to Ignore List](#) [Report User](#)

Databases's Forum Info

GOD

Joined:
March 22, 2015

Time Spent Online:
5 Months, 2 Weeks, 4 Days

User Identifier:
666 [Copy Profile Permalink]

Referred by:
[\(Owner\) Omnipotent](#)

Username Changes:
3

Members Referred:
4

Additional Info About Databases

Sex:
Undisclosed

Databases's Contact Details

Private Message:
Send Databases a private message.

Databases's Forum Statistics

Total Threads:
51 (0.02 threads per day | 0.04 percent of total threads) [Find All Threads](#)

Total Posts:
50 (0.02 posts per day | 0 percent of total posts) [Find All Posts](#)

Reputation:
1,197 [Details](#) [Rate](#)

Databases's awards.

- [Award](#) September 21, 2019 at 12:15 AM
- [Award](#) August 21, 2019 at 11:09 AM
- [Award](#) August 21, 2019 at 08:59 AM
- [Award](#) August 19, 2019 at 05:44 AM

New Thread
buying twitter scraped emails
in Leaks Market
September 13, 2020 at 01:50 AM 0

New Thread
forums.r2games.com 2017 [PARSED]
in Databases Removed Content
July 04, 2020 at 11:06 PM 1

New Thread
creative-scape.net DB! [186k]
in Databases Removed Content
July 04, 2020 at 10:50 PM 6

Figure 51 - **Databases** profile on RaidForums

The history of name changes is presented below.

Username History for Databases	
Old Username	Date Changed
0 12	June 13, 2019 at 09:02 PM
012	June 11, 2019 at 07:01 PM
Eutropios	June 11, 2019 at 09:58 AM

Figure 52 - Username history for **Databases**

The account started to post databases only in 2019. In August 2021, the format of their messages about selling databases started to resemble **ShinyHunter's** posts. That is why we suppose that **Databases** probably became a partner of **ShinyHunters** in 2021. **ShinyHunters** themselves have used this format of messages since 2020.

The metric for **Databases** is presented below.

==== Databases [182] =====		
Databases	score: 0.0	matches: 182
0-12	score: 0.0	matches: 2
B342384	score: 2.631	matches: 2
Troy Hunt	score: 2.927	matches: 9
lex205	score: 2.969	matches: 2
DonJuji	score: 3.018	matches: 5
Omnipotent	score: 3.044	matches: 9
Jnx3cx	score: 3.096	matches: 8
Essaye	score: 3.134	matches: 2
chmod	score: 3.304	matches: 2

Figure 53 - The result of the analysis of **Databases'** posts

As you can see from the metric **Troy Hunt** (aka **Jnx3cx**) and **Databases** are close to each other as well.

The last account for the analysis is **FluffyBunnyFufu**, created back in 2018.

Ceech
M.V.P User
Status: (Hidden) (Last Visit: Hidden)

[Add to Buddy List](#) [Add to Ignore List](#) [Report User](#)

Ceech's Forum Info	Ceech's Contact Details	Ceech's awards.
<p>Joined: January 16, 2018</p> <p>Time Spent Online: (Hidden)</p> <p>Username Changes: 2</p> <p>User Identifier: 121401828 [Copy Profile Permalink]</p>	<p>Private Message: Send Ceech a private message.</p> <p>Ceech's Forum Statistics</p> <p>Total Threads: 9 (0.01 threads per day 0.01 percent of total threads) Find All Threads</p> <p>Total Posts: 588 (0.51 posts per day 0.02 percent of total posts) Find All Posts</p> <p>Reputation: 820 Details Rate</p>	<p> November 17, 2020 at 07:12 PM</p> <p> May 27, 2020 at 02:15 AM</p> <p> May 07, 2020 at 10:18 PM</p> <p> May 05, 2020 at 02:34 PM</p> <p>Pages (2): 1 2 Next »</p>
<p>Additional Info About Ceech</p> <p>Sex: Undisclosed</p>	<p>Ceech's Signature</p> <p>Bye poor fellows 😊</p>	<p>New Reply</p> <p>RE: Leaks.sh, Free & Unlimited Data Breach Search Engine (Beta) in The Lounge 🕒 February 13, 2021 at 08:46 AM 🗨️ 33</p>

Figure 54 - **FluffyBunnyFufu** (aka **Ceech**, **FunnyBunnyHere**) profile on RaidForums

Username History for Ceech	
Old Username	Date Changed
FluffyBunnyFufu	November 17, 2020 at 07:49 PM
FunnyBunnyHere	July 11, 2018 at 11:03 PM

Figure 55 - Username history for **Ceech/FluffyBunnyFufu/FluffyBunnyHere** account

Almost all mutual threads of the users relate to databases. One of **FluffyBunnyFufu's** posts with a leak publication is similar in format to one of **ShinyHunters'** posts. Additionally, **FluffyBunnyFufu** joined a dispute initiated against **ShinyHunters** with **fs0c131y/whysodankk**. **FluffyBunnyFufu** joined the talk to defend the position of **ShinyHunters**. This tactic is common for the members of **ShinyHunters** group as we noted during the Manual analysis part of this research.

ShinyHunters

"Hello RaidForums Community,\nToday I have uploaded the Ulmon Database for you to download for free, thanks for reading and enjoy!\n\nNotes|\nIn January 2020, the travel app creator Ulmon suffered a data breach. The service had almost 1.3M records with 777k unique email addresses, names, passwords stored as bcrypt hashes and in some cases, social media profile IDs, telephone numbers and bios.\n\nCompromised data: Bios, Email addresses, Names, Passwords, Phone numbers, Social media profiles\n\nContentsSpoilerThis download consists of 1 .SQL file, this dump's passwords are hashed using bcrypt. The .7z File's MD5 Hash is 9717BB410611FB2D31CA62922D4C009C. Total record count of 777769.\n\nDownloads\n\nHidden Content\n\nUnlock for 8 credits\n\nDatabase Index <> How To Get Credits"

...

FluffyBunnyFufu

"Hello RaidForums Community,\nToday I have uploaded the CryptoStar.Asia Database for you to download for free, thanks for reading and enjoy!\n\nNotes|\nIn 2020, the cloud miming website CryptoStar.Asia was hacked and 22.3k accounts were exposed. The data included email, IP addresses and hashed passwords(md5).\n\nCompromised data: Email addresses, Passwords, IP Addresses\n\nContentsSpoilerThis download consists of 1 TXT containing colon separated columns, hashing method used seems to be MD5. The .7z File's MD5 Hash is A2E37748C097AC7B9AA17CFCAF414DBE. Total record count of 22327.\n\nDownloads\n\nHidden Content\n\nUnlock for 8 credits\n\nDatabase Index <> How To Get Credits"

Figure 56 - Posts published by **FluffyBunnyFufu** and **ShinyHunters** and written using the same format

According to the available data, it cannot be stated for sure if there is a clear connection between these two accounts, and it may be just a coincidence. However, it is clear that **FluffyBunnyFufu** keeps an eye on **ShinyHunters's** activities.

We can conclude that accounts **J4ckd0x**, **Megadimarus**, **Troy Hunt**, and **Databases** either occasionally used the same message format to sell databases, or they are related to each other and **ShinyHunters**. However, the majority of **ShinyHunters'** posts about selling databases have been using the same format since 2020, so we tend to think that this format is typical for the group.

Hello RaidForums Community,\nToday I have uploaded the Mathway Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn January 2020, the math solving website Mathway suffered a data breach that exposed over 25M records. ...

Hello RaidForums Community,\nToday I have uploaded the Appen Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the AI training data company Appen suffered a data breach exposing the details of almost 5.9...

Hello RaidForums Community,\nToday I have uploaded the Heavenly Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the interior design website Heavenly suffered a data breach which impacted almost 1.4 milli...

Hello RaidForums Community,\nToday I have uploaded the Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn July 2020, the Database was breached leading to 2.7 million users being affect...

Hello RaidForums Community,\nToday I have uploaded the Drizly Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn approximately July 2020, the US-based online alcohol delivery service Drizly suffered a data breach. T...

Hello RaidForums Community,\nToday I have uploaded the OrderSnapp Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020 the online POS platform OrderSnapp.com suffered a data breach causing 1.3 Million of its...

Hello RaidForums Community,\nToday I have uploaded the Swvl Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the Egyptian bus operator Swvl suffered a data breach which impacted over 4 million members o...

Hello RaidForums Community,\nToday I have uploaded the Hurb Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn approximately March 2019, the online Brazilian travel agency Hurb (formerly Hotel Urbano) suffered a dat...

Hello RaidForums Community,\nToday I have uploaded the GGumim.co.kr Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn March 2020, the Korean interior decoration website 집꾸미기 (Decorating the House) suffered a data b...

Hello RaidForums Community,\nToday I have uploaded the TrueFire Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn February 2020, the guitar tuition website TrueFire suffered a data breach which impacted 600k member...

Hello RaidForums Community,\nToday I have uploaded the Dave Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the digital banking app Dave suffered a data breach which exposed 7.5 million rows of data an...

Hello RaidForums Community,\nToday I have uploaded the GGumim.co.kr Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn March 2020, the Korean interior decoration website 집꾸미기 (Decorating the House) suffered a data b...

Hello RaidForums Community,\nToday I have uploaded the ProctorU Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the online exam service ProctorU suffered a data breach which was subsequently shared ext...

Hello RaidForums Community,\nToday I have uploaded the Yotepresto Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, 1,444,676 login credentials from Yotepresto loaning service were exposed online. To thi...

Hello RaidForums Community,\nToday I have uploaded the RewardStyle Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn February 2020, the influencer marketing consultation agency RewardStyle seem to have been breache...

Hello RaidForums Community,\nToday I have uploaded the Vakinha Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the Brazilian fund raising service Vakinha.com.br suffered a data breach which impacted al...

Hello RaidForums Community,\nToday I have uploaded the Bhinneka Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn May 2020, login credentials from a massive breach of Bhinneka Indonesia-based technology shop were f...

Hello RaidForums Community,\nToday I have uploaded the Minted Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn February 2020, the online marketplace for independent artists Minted suffered a data breach that expos...

Hello RaidForums Community,\nToday I have uploaded the Kreditplus Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the Indonesian credit service Kreditplus suffered a data breach which exposed 896k reco...

Hello RaidForums Community,\nToday I have uploaded the Promo Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn July 2020, the self-proclaimed "World's #1 Marketing Video Maker" Promo suffered a data breach which wa...

Hello RaidForums Community,\nToday I have uploaded the Sonicbids Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn December 2019, the booking website Sonicbids suffered a data breach which they attributed to "a da...

Hello RaidForums Community,\nToday I have uploaded the Mashable Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn 2020 the website Mashable was allegedly breached. The subsequent leaked files are very difficult to ...

Hello RaidForums Community,\nToday I have uploaded the Scentbird Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the online fragrance service Scentbird suffered a data breach that exposed the personal ...

Hello RaidForums Community,\nToday I have uploaded the Wishbone Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn January 2020 the app Wishbone had their database leaked by a hacker known as "ShinyHunters", this da...

Hello RaidForums Community,\nToday I have uploaded the Ulmon Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn January 2020, the travel app creator Ulmon suffered a data breach. The service had almost 1.3M records ...

Hello RaidForums Community,\nToday I have uploaded the Chatbooks Database for you to download for free, thanks for reading and enjoy!\nNotes\nOn the 26th of March 2020, the photo print service Chatbooks suffered a data breach. The breach contain...

Hello RaidForums Community,\nToday I have uploaded the LiveAuctioneers Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn June 2020, the online antiques marketplace LiveAuctioneers suffered a data breach which was s...

Hello RaidForums Community,\nToday I have uploaded the StarTribune Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn October 2019, the Minnesota-based news service StarTribune suffered a data breach which was subse...

Hello RaidForums Community,\nToday I have uploaded the Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn July 2020, Database suffered...

Hello RaidForums Community,\nToday I have uploaded the JamesDelivery Database for you to download for free, thanks for reading and enjoy!\nNotes\nIn March 2020 the Brazilian delivery service "James", suffered a databreach. The breach exposed 1...

Figure 57 - The result of the analysis of ShinyHunters' posts

Another argument is that the databases which were posted by these accounts are associated with **ShinyHunters** or their predecessors. According to Vinny Troia's research, **ShinyHunters** is a successor to NSFW and Gnostic Players groups.

The semantic analysis along with our manual research has also shown that the higher the distance between the messages of two different accounts is, the weaker their connection is. This sounds reasonable from a theoretical point of view and our analysis has confirmed it.

Conclusions

In this article, we wanted to show several methods of analysis using which researchers can establish possible links between various accounts in the underground. To showcase them, we used the example of infamous group **ShinyHunters** known for their unique and resonant leaks.

The starting point of our research was a message from **ExpertData**, a relatively new account with several unique leaks that appeared in September 2020. The investigation could have stopped there if we treated **ExpertData** as just another new actor of RaidForums. However, our manual analysis suggested a link between this account and **ShinyHunters** based on their conversational tactics and databases posted for sale.

The suggested NLP algorithm seems to provide a relevant metric to find connected accounts although with some false positives. Combined manual and automatic analysis showed a potential link of **ShinyHunters** with the following accounts: **fs0c131y** (aka **whysodankk**, **whysodank**), **J4ckd0x**, **Megadimarus**, **Databases** (aka **Eutropius**, **012**, **0 12**), **Troy Hunt** (aka **Jnx3cx**). Moreover, accounts belonging to the same user can be automatically

detected, like in cases of **fs0c131y**, **Databases**, and **Troy Hunt**. It is also noteworthy that accounts **fs0c131y**, **J4ckd0x** and **Jnx3cx** have similar format of the name.

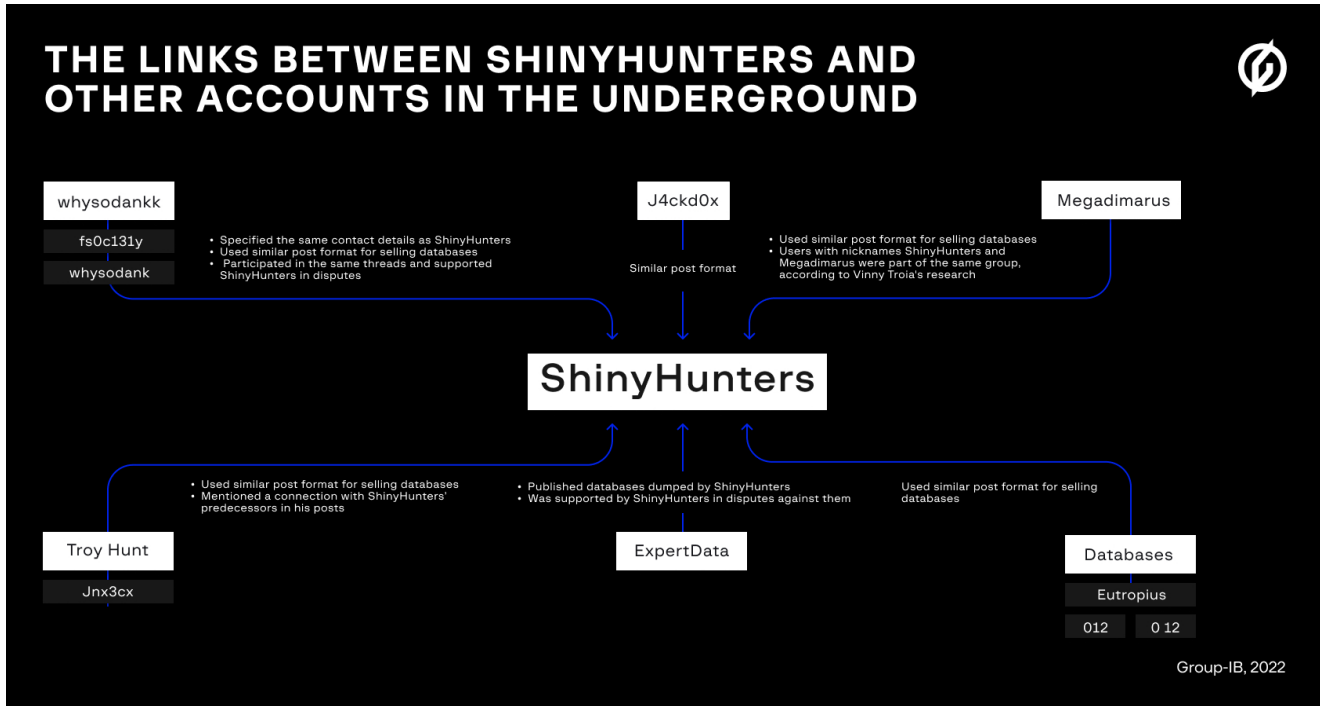


Figure 58 - The links between **ShinyHunters** and other accounts in the underground

This article is not about the **ShinyHunters** activity, the group's active behavior in the dark web is just a good example for demonstrating automatic techniques that can be used for analyzing discussions. In this case, we deliberately employ a well-known approach for searching semantically close messages to demonstrate the advantages of even common ML methods in cybersecurity investigations.

As you can see, the results of the manual method and the method based on the semantics of the messages both overlap and complement each other. The examination of behavior patterns of accounts in the dark web can show the correlation between accounts used by the same group, with automatic analysis based on the proposed algorithms allowing to increase the depth of research, while reducing the list of suspects. In the article we've examined the following connections:

- between **ShinyHunters** and **fs0c131y** (aka **whysodankk**, **whysodank**), whose link with **ShinyHunters** is quite obvious due to the common contact details;
- between **ShinyHunters** and **J4ckd0x**, **Megadimarus**, **Databases**, **Troy Hunt** (aka **Jnx3cx**) where the link was uncovered using semantic analysis;

- between **ShinyHunters** and **ExpertData**, whose link with **ShinyHunters** was not obvious, and could't be revealed by the semantic analysis alone and required some manual work.

False positives were related to standard phrases added by the forum administrators to the messages of other members. Moreover, admins reposted the messages of other members which are semantically close to what is searched. Depending on the goal of the analysis these results can be improved by applying more specific filters. The techniques proposed can be used to speed up and widen the results of the cybercrime investigations related to dark web activities. Once automated, it can even help in the deanonymization process.

Try Group-IB Threat Intelligence & Attribution right now

Detect public leaks, identify affected accounts in breached databases and try advanced threat actor profiling with best-in-class threat intelligence

Group-IB Threat Intelligence & Attribution