Attack distinction

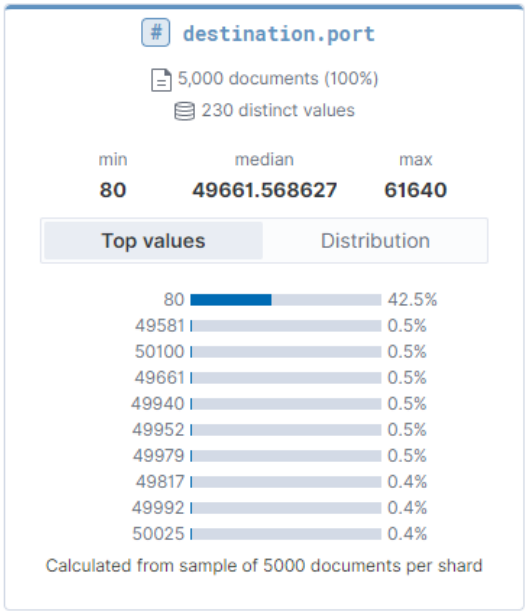
A. Attack 1: Brute-Force Attack

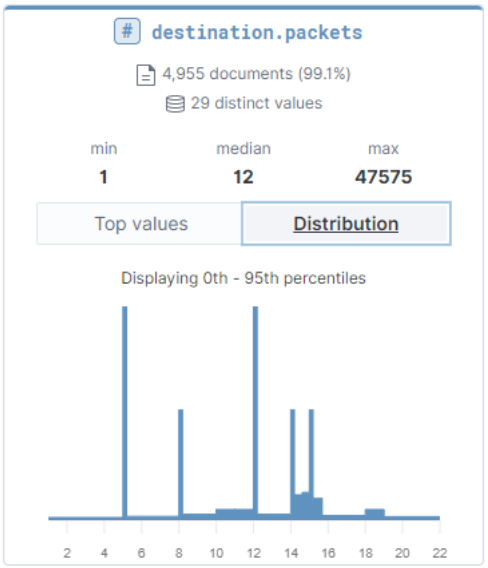
A lot of **url.query** which include different passwords and same username.

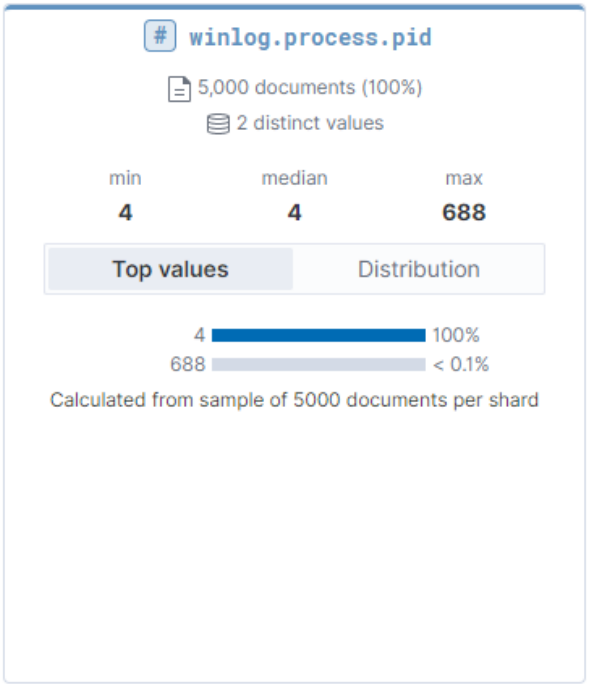
Time	url.query
Oct 28, 2020 @ 13:46:20.053	Login=Login&password=flower1&username=aaliyah
Oct 28, 2020 @ 13:46:20.040	Login=Login&password=forall&username=aaliyah
Oct 28, 2020 @ 13:46:19.995	Login=Login&password=flyguy&username=aaliyah
Oct 28, 2020 @ 13:46:19.833	Login=Login&password=fordf350&username=aaliyah
Oct 28, 2020 @ 13:46:19.741	Login=Login&password=franci&username=aaliyah
Oct 28, 2020 @ 13:46:19.711	Login=Login&password=Fuck1&username=aaliyah
Oct 28, 2020 @ 13:46:19.355	Login=Login&password=fucker1&username=aaliyah
Oct 28, 2020 @ 13:46:19.298	Login=Login&password=fuckshit&username=aaliyah
Oct 28, 2020 @ 13:46:19.271	Login=Login&password=funny1&username=aaliyah
Oct 28, 2020 @ 13:46:19.216	Login=Login&password=gates&username=aaliyah
Oct 28, 2020 @ 13:46:19.211	Login=Login&password=gatito&username=aaliyah
Oct 28, 2020 @ 13:46:19.188	Login=Login&password=geibcnbr&username=aaliyah
Oct 28, 2020 @ 13:46:19.165	Login=Login&password=Ginger&username=aaliyah
Oct 28, 2020 @ 13:46:19.137	Login=Login&password=glennwei&username=aaliyah
Oct 28, 2020 @ 13:46:19.111	Login=Login&password=goethe&username=aaliyah
Oct 28, 2020 @ 13:46:19.096	Login=Login&password=golfman&username=aaliyah
Oct 28, 2020 @ 13:46:19.062	Login=Login&password=google1&username=aaliyah
Oct 28, 2020 @ 13:46:19.037	Login=Login&password=gretta&username=aaliyah
Oct 28, 2020 @ 13:46:18.992	Login=Login&password=halcyon&username=aaliyah
Oct 28, 2020 @ 13:46:18.932	Login=Login&password=heathe&username=aaliyah

B. Attack 2: DDoS

42.5% **destination port** is 80, and there are many packets sending to server in a very short time. Most **winlog.process.pid** is 4.

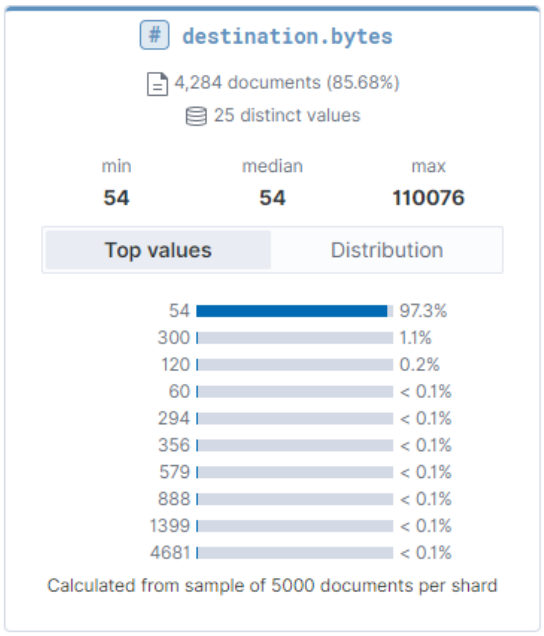


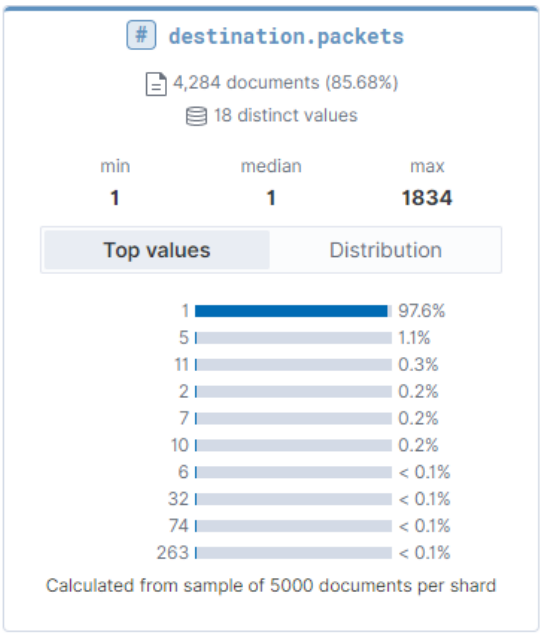


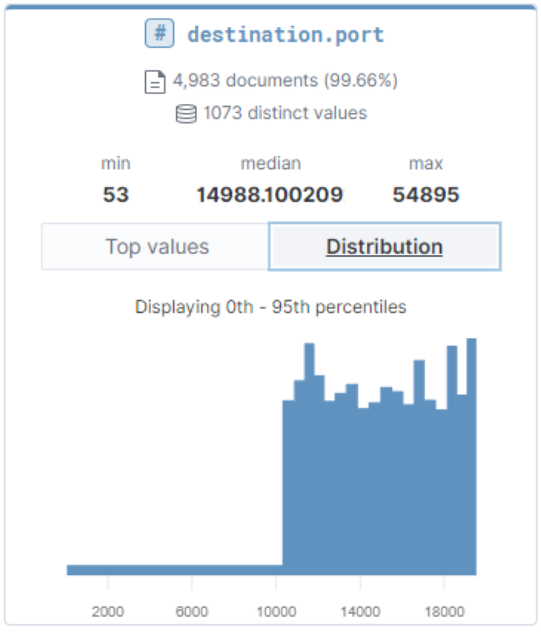


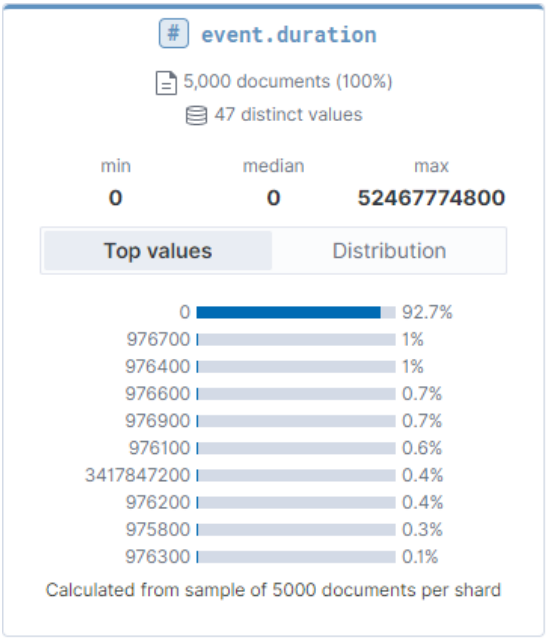
C. Attack 3: Port Scanning

97.6% sends only a packet to a big range of ports, and the packet size is very small. Most **event.duration** is 0.









D. Attack 4: Phishing Email

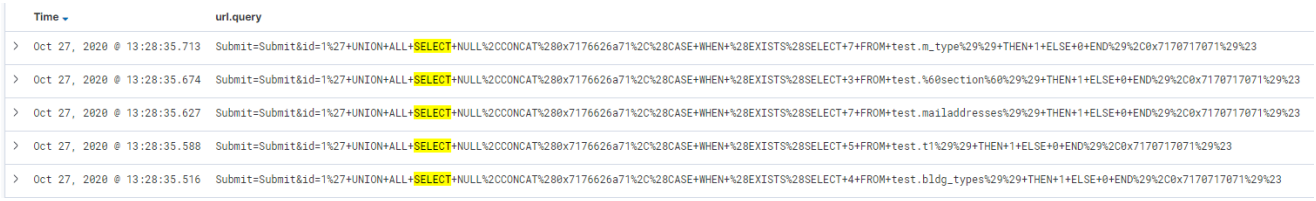
There are many strange path of **winlog.event_data.Application** like below: “\device\harddiskvolume2\windows\system32\svchost.exe”. It means the system might be attacked.

Time	winlog.event_data.Application
Oct 30, 2020 @ 13:29:20.311	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.311	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.311	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.311	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.310	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.309	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.307	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.307	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.305	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.305	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.200	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.197	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.072	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:20.072	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:29:07.154	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:28:28.015	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:28:28.015	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:27:38.053	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:27:38.053	\device\harddiskvolume2\windows\system32\svchost.exe
Oct 30, 2020 @ 13:27:24.082	\device\harddiskvolume2\windows\system32\svchost.exe

E. Attack 5: SQL Injection

There are many **url.query** like below:

“Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+7+FROM+test.m_type%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23” using SQL syntax.



Time	url.query
Oct 27, 2020 @ 13:28:35.713	Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+7+FROM+test.m_type%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23
Oct 27, 2020 @ 13:28:35.674	Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+3+FROM+test.%60section%60%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23
Oct 27, 2020 @ 13:28:35.627	Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+7+FROM+test.mailaddresses%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23
Oct 27, 2020 @ 13:28:35.588	Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+5+FROM+test.t1%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23
Oct 27, 2020 @ 13:28:35.516	Submit=Submit&id=1%27+UNION+ALL+SELECT+NULL%2CCONCAT%280x7176626a71%2C%28CASE+WHEN+%28EXISTS%28SELECT+4+FROM+test.bldg_types%29%29+THEN+1+ELSE+0+END%29%2C0x7170717071%29%23

Algorithm

1. Load the training data
2. Calculate each type of attack score

   Below is my function used to calculate score:

   check_Port_Scan(beat_file, attack_score, file)

   check_SQL_Injection(beat_file, attack_score, file)

   check_DDoS(beat_file, attack_score, file, DDoS_winlog_process_pid)

   check_Brute_Force(beat_file, attack_score, file)

   check_Phishing_Email(beat_file, attack_score, file)

   They will check related features I mentioned in previous part of different attack types to calculate associated score.
3. Determine attack type

   Below is my function used to determine attack type:

   determine_attack(attack_score, dir)

   It will find which attack type has the highest score, and determine this attack belongs to the attack type.
4. Load the testing data
5. Run 2~3 to know testcase belongs to which attack type

Conclusion

Due to less data, I chose rule-based algorithm to determine the attack type. I thought it didn't need to do machine learning or even deep learning. Through ELK I can easily find some important features, so I will use this tool next time I need to do some similar tasks.