Chapter-3
Reasons and Remedies of False Positive

In this chapter, we have started with few real life examples of false positive and their effects. We have discussed about few basic terminologies about normal and attack traffic and represented equations for false positive rate and detection rate. As part of the research work we have identified major causes of false positive. Finally we have discussed about how we can overcome such false positives.

3.1 False Positive Basics

In modern information era, Internet is essential service for our society. But Internet represents an insecure channel for exchanging information leading to a high risk of intrusion. We need to make use of security mechanisms and measures to protect against attacks from the Internet. Internet Security mechanisms like VPN, SSL, TLS, IPSec, PGP, antivirus, firewall, Intrusion Detection System can help to protect our network. Still there is possibility of attack from Internet or insider from our computer network. It is very critical to detect such life-threatening attacks. Basic responsibility of Intrusion Detection System is to identify intrusion/attack and immediately send alert to concern security staff. Over the years researchers have invented many techniques to detect intrusion. Intrusion Detection System needs to differentiate normal traffic and intrusion. If Intrusion Detection System fails to differentiate normal traffic and intrusion with 100% accuracy, it generates false positive and false negative. An event signaling IDS to produce an alarm when no attack has taken place is called false positive. In general, false positive is any regular traffic which is identified as an attack or anomaly. And such false alarms will eventually direct the victim to ignore such alerts whether it is genuine or false alarm.
3.1 False Positive Basics

Few examples of such cases are email notifications sent by email service provider and bank. For providing increased security, email service provider identifies the device, IP address, and the browser from which user is accessing email. They build the profile of the user and keep track of how user accesses email regularly. Once the profile is built, any deviation is sent as email notification to the user. For example, initially user is using Ubuntu 12.04 LTS and Firefox 30.0.1 browser. After sometime user is promoted to the new department and now she uses multiple devices to access her email. She has also procured Internet connectivity at home and from multiple devices she access Internet from her home also. As a result each time she login to the email service, she receives an email claiming that your account is used from different device and browser. For initial few alert notification user read the alert email and verifies the content. After verifying such inappropriate alerts, user starts ignoring alerts sent by the email service provider. Eventually user feels that all alerts sent by the email service provider are useless and deletes it without even reading. This is the worst scenario. Now the email service provider has lost their credibility. In the case of actual fraud, when the email service provider sends the email alert notification, user simply deletes the alert without reading it. This is due to lack of trust developed by large unnecessary/false email notifications.

Another such example is email alerts sent by the bank. To increase the security, bank sends the email notification for each transaction performed by the customer. For normal savings account customer who is performing about 10 transactions per month, these alerts might not be a problem. But the customer using current account might be receiving large number of such email notifications. Initially the customer may verify the email notifications sent by the bank. She may spend some time to cross check with all the transactions performed on the yesterday. Eventually most customers unable to spend such time to verify large email notifications daily. As a result customers start ignoring such daily email notifications sent by the bank and deletes all without reading. Now in the case of actual incident also customer simple delete the alert notification sent by the bank. Once again this happened due to lack of trust in the email alerts sent by the bank.
From such examples we can understand that large number of false alerts makes the user inconvenient. If Intrusion Detection System generates large number of alerts and out of all these alerts most are false positives, the same result may be observed. Initially security person goes through all such alerts generated by Intrusion Detection System and verify the incidents. After spending lots of time and resources also nothing useful is found due to large number of false alerts. Eventually administrator of the Intrusion Detection System starts ignoring such large number of alerts. So the basic objective of Intrusion Detection System is completely washed out.

Here, we need to understand the detection terms used in the context of Intrusion Detection System. We can understand the normal traffic and attack bifurcation as under:

<table>
<thead>
<tr>
<th>Actual</th>
<th>Identified</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (Normal)</td>
<td>TN</td>
<td>normal traffic</td>
<td>correctly identified</td>
</tr>
<tr>
<td>Positive (Attack)</td>
<td>FN</td>
<td>attack identified as normal traffic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FP</td>
<td>normal traffic</td>
<td>identified as attack</td>
</tr>
<tr>
<td></td>
<td>TP</td>
<td>correctly identified</td>
<td>attack</td>
</tr>
</tbody>
</table>

Table – 3.1: Attack and Normal Traffic

As indicated in Table-3.1, actual traffic can be seen by Intrusion Detection System correctly or incorrectly. Here, normal traffic is indicated as negative while attack indicates as positive. If actual normal traffic is identified as normal by Intrusion Detection System it is called True
3.2 Reasons of False Positive

Negative (TN). True Positive (TP) indicates attack traffic is identified as attack by Intrusion Detection System. So TP and TN are correct identification done by Intrusion Detection System. The other two square represents incorrect identification by Intrusion Detection System. Undetected attack is called False Negative (FN). This indicates attack is identified as normal traffic by Intrusion Detection System. Normal traffic identified as attack by Intrusion Detection System is called False Positive (FP).

Out of total normal traffic how much is incorrectly identified as attack is called false positive rate. Out of all correctly identified traffic (normal and attack both) how much traffic is attack is called the detection rate. For any Intrusion Detection System detection rate must high and false positive rate must be very low. Formal equations for false positive rate and detection rate can be written as under:

\[
\text{False Positive Rate FPR} = \frac{\text{incorrectly identified normal traffic}}{\text{total normal traffic}}
\]

\[
\text{False Positive Rate FPR} = \frac{FP}{(FP + TN)} \quad \text{................. (3.1)}
\]

\[
\text{Detection Rate DR} = \frac{\text{correctly identified attacks}}{\text{total attacks}}
\]

\[
\text{Detection Rate DR} = \frac{TP}{(TP + FN)} \quad \text{................. (3.2)}
\]

Ideally we don’t want any of the false positive but in practical environment false positives are bound to happen. At least we can strive to minimize the false positive rate. There are many challenges which results in higher false positive rate.

3.2 Reasons of False Positive

Lack of sufficient training may result in false positives. An application not seen in the training stage of an anomaly detection system will likely trigger an alert when the application attempts to run. For anomaly based Intrusion Detection System new application is deviation from the normal behavior. So the fundamental rule for success of anomaly based Intrusion Detection system is sufficient and effective training. While signature based IDS will likely to generate false alerts if too broad signature is developed. A broadly written signature can include both legitimate and illegitimate traffic.
The second major reason for high false positive rate is probability of attack within a normal traffic flaw is very low. For example, let say 0.001 false positive rate is estimated. If one attack is present in a million sessions, the ratio of false to real alarm will be 100. This problem is not just restricted to Intrusion Detection system. In other areas like cyber forensic, medical tests, chemical reaction also this problem persists. If probability of symptom occurrence is very low compare to error rate of test, large number of false positives can be observed.

The third major reason for false positive is the Internet. Today, most networks are connected to Internet. And the Internet is full of noise and incomplete data. Many corrupted, incomplete, or escaped packets roaming over the Internet. For example, router is designed to drop the IP packet once TTL value reaches to zero. Due to temporary hardware or software bug at the router such LAN IP packet is escaped to the Internet. Some corrupted packets are result of communication error at network or MAC layer.

Spurious and duplicate data passed on to the Intrusion Detection System is one more reason of false positive. Most modern Intrusion Detection System comes with preprocessing unit. The primary focus of this preprocessing is to provide clean data to the Intrusion Detection System. Such clean data not only reduces false positive rate but also improves the overall performance of the Intrusion Detection System.

Next reason for false positive is network peculiarity. Some behavior is expected in one environment while prohibited in another situation. As an example NetBIOS over TCP/IP is normal in a LAN environment with Windows workstations but not typically seen on the Internet. Many legacy applications still use NetBIOS for name resolution/registration in LAN environment. Few genuine applications do not exactly follow RFCs. If signature is written as per the RFC, it may generate false positive when such applications encountered. The problem is signature is as per the RFC but application deviates significantly. We cannot eliminate incorrectly written (not as per standard) client applications from network. Thus Intrusion Detection System needs to understand such deviated applications.
3.2 Reasons of False Positive

Another reason for false positive is network packets generate in reaction of some network event. For example, failure of a network component may raise number of ICMP destination unreachable packets. All of sudden Intrusion Detection System observes large number of ICMP messages passing through the network. Sudden increase of ICMP error messages can be seen by Intrusion Detection System as an attack (ICMP flooding). But the traffic that is caused by such network event is often normal.

Some outdated or odd equipments can be one more reason for false positive. Such odd equipment generates packets which are not recognized by the Intrusion Detection System. For example some load balancing devices generates odd packets which are seen as an attack by Intrusion Detection System. Again we cannot eliminate such devices from network so Intrusion Detection System needs to learn about it.

In a specific scenario, may be as part of the experiment, genuine user is transmitting network packet which resembles some network attack. For example network administrator has received alert from Intrusion Detection System. To verify this alert network administrator tries to communicate on a prohibited port. Such genuine act which is not malicious will result in further alerts from Intrusion Detection System. In this type of scenario we cannot blame Intrusion Detection System for false positive.

Even it is possible that Intrusion Detection System generates false positives without any significant reason. It may be result of efforts to optimize some parameters of Intrusion Detection System itself or some unidentified bug within Intrusion Detection System encountered in a specific situation.

Here we have discussed only few sample scenarios. It is notable that too many things can affect the decision of IDS which may result in false alerts. In practice IDS generate large number of alerts daily and out of those many are false positive [4][15][17][20]. This makes it very impractical if not impossible for security staff to respond.
3.3 False Positive Remedies

After understanding various causes of false positive, we can think about how to curb it. One cannot eliminate all the causes discussed earlier but certainly effect of it can be minimized. Each reason of false positive is having different characteristics and remedies should be designed accordingly. Based on the literature review and exhaustive study of false positives generated by Intrusion Detection System, we have summarized major solutions.

Anomaly based Intrusion Detection System will work efficiently if adequate training is provided. So the very first step towards reducing false positive rate is to teach the anomaly based Intrusion Detection System about normal behavior. The major challenge is each network is having vast variety of network components like routers, layer-2 and layer-3 switches, bridges, repeaters, access points, patch panels, fiber management systems, modules, media convertors. Various devices like servers, workstations, notebooks, PDA, handheld devices, printers, scanners, automation instruments, RFID, biometric, sensors may also connected to the network. Computing devices, network components, and other devices may be connected with communication media like fiber optic cable, CAT6 cable, or radio signals. One of the important things for any network is the services provided by that network. Typically network provides Internet access, DNS, local Web portals, database services based on the authentication, authorization, and access control. Most modern networks use LDAP, active directory, RADIUS, or AAA servers. The organization which supports BYOD (Bring Your Own Device) suffers from additional network security challenges. During the training phase anomaly based Intrusion Detection System needs to learn about normal behavior imposed by all the devices, services, and communication.

In addition to training Intrusion Detection System for normal behavior, most prevailing attack behavior training should be provided. Any attack scenario will generate large number of alerts, if it is identified correctly. Based on the altered behavior and the number of alerts generated by detection engine, we can further suppress false alerts. For example, in the beginning of SYN flooding attack, N number of alerts is generated but in the case of false positive, less than 10% of alerts generated. Reverse alert can be sent notifying false alerts
already sent to output interface. Another approach can be delay the alert notification till detection engine generates significant number of alerts expected depending on the attack scenario. But it will reduce the prevention window. Security person will get the alert after attack is already executed successfully.

Many researchers have suggested alert aggregation to reduce large number of alerts generated by Intrusion Detection System [10][26][66][84][99]. After intrusion detection engine generates an alert but before sending it to output interface or user, post processing is required. Such simple alert clustering technique can reduce large number of alerts and provides unified view of alerts. Practically when attacker starts attacking, Intrusion Detection System identifies deviated behavior. For all such deviations, Intrusion Detection System will generate alerts. So a single attack scenario is represented by large number of alerts. By using some mechanism, if we can combine these alerts indicating a single attack, we can reduce the number of alerts drastically.

Alert correlation is another important tool used to reduce the false positive rate of Intrusion Detection System. For example, Intrusion Detection System has observed significant deviation from the normal behavior. In place of immediately raising an alert Intrusion Detection System communicates with other security measures. Other audits and logs from host, firewall, and honeypot are verified for any related attack information. If other sources also conform than the alarm is raised otherwise this may be assumed as false positive or sent for still further verification. This can be implemented as alert processing filters designed by intrusion detection analyst. Such filters should be consistently monitored by intrusion detection analyst. If required new filtered can by created or old can be dropped by intrusion detection analyst. Success of this alert filtering technique mainly depends on the knowledge and skill of intrusion detection analyst.

Up to date status of network devices and systems can be used for reduction of false positive rate. For example maintain configuration database and store above mentioned status in it. By using this configuration database as attack confirmation, false positives can be reduced. For example, as per the organization’s security policy every Sunday morning 7.00 AM
vulnerability scanner executes. Intrusion Detection System may raise large number of alerts like port scanning, injection attacks, cross site scripting. From the configuration database Intrusion Detection System confirms that these alerts are result of scheduled vulnerability scanner and not the real attacks. If Intrusion Detection System has identified an attack against one of the server but as per the configuration database such attack is not applicable to the operating system installed in the server. So clearly this alert is also false positive, a deviation from normal behavior which is not malicious.

Reputation enabled mechanism can also use for reduction of false positive rate. Intrusion Detection System maintains a reputation database based on various parameters like IP address, MAC address, application, and user/group. While alert generation, Intrusion Detection System considers the reputation parameters. These reputation values can be static fixed (provided by intrusion detection analyst) or dynamic. In the absence of static fixed reputation, initial value is neutral. With each attack incident reputation value is decreased while long absence of attack increases the reputation value.

In depth analysis up to the higher layer in protocol stack can further help in reducing false positive rate of Intrusion Detection System. Just like Next Generation Firewall (NGFW), which claims to be capable of deep protocol inspection, intrusion detection can be benefited. Most modern Next Generation Firewalls are still in the early days and we have to wait for good results. Rather, signification performance degradation is observed in some of the Next Generation Firewall due to such attempts in this direction.

One of the important reason for higher false positive is Internet traffic is very noisy. Preprocessing can remove noise from the traffic coming from Internet and provide clean data to Intrusion Detection System for attack identification. Quality input often results in quality output. Such preprocessed data reduces the possibility of false alert and improves overall performance of Intrusion Detection System.