

The Fundamentals of Pattern Matching in Security

Alex van der Ploeg

Abstract

Computers and networks have the risk of being assaulted everyday. To protect them one needs to provide excellent protection.

To provide this protection one first needs to know where to protect against. The common threats are Malware (viruses, Trojan horses, worms, Spyware), spam and malicious attacks. Each of these threats tries to steal, destroy or deny access to information.

These threats cost a lot of money to the user, companies and ISPs. To solve the money loss one needs to be protected. There is a solution to each type of threat. The antivirus is able to counter the Malware, the anti-spam can block the incoming spam and the intrusion prevention and detection system is able to alarm and block the malicious attacks.

The good security solutions use a pattern matching technology to recognize the threats. The ideal coding language that can be used for recognition is regular expression. This code can be processed quickly, thus it does not add much delay. There are two ways to automate the code: deterministic finite automaton or non-deterministic finite automaton. The deterministic approach is a tree based and the non-deterministic is probabilistic. As updates are important in security solution, the non-deterministic is a better choice, because it is able to update immediately. The deterministic approach needs to rebuild the whole tree for each update, which can cost hours of downtime.

The code can be used in a hardware and/or software solution. Hardware is able to process a lot of traffic, thus a hardware solution is a good choice for a central location. Software is often cheaper, but is not able to handle a lot of traffic at once, thus a better choice for solutions at the hosts or for small companies.

The solutions need to be placed in a network. The best location is near the source of the threat. The locations to choose from are the host, edge and DMZ. A combination of all the locations provides the strongest security. If a malicious user breaks through one line of defence, then he still has a couple to go through. Furthermore, the source is not always from the outside network.

At the end it comes down to what the value of the information is, compared to the price of the solution. One needs to create a good price plan for the product, updates and the services. The solution needs to be of good integrity too. For example, providing a free security solution with Spyware is not an option.

With the combination of faultless pattern matching, the location of the solutions in a network, and good marketing, one can provide an excellent security solution.