APT detection Methods with tracking file access and comparing file access logs

Makoto Sonoda (1310037)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 10, 2016

Keywords: APT, Security, Tree Structure, Anomaly Detection, File System.

Advanced Persistent Threat（APT）is a kind of Targeted Attack. APT aim at classified information. Large scale information compromise by APT lead to a reduction of tangible assets by the compensation for damages, and that lead to a reduction of intangible assets by the decrease of the corporate brand. characteristics of APT are perform persistent attack as a target for specific person, and select the intrusion method according to the opponent, and it makes a long-term search after the intrusion. Long-term search is carried out in order to avoid the anomaly detected in a short period of time.

Intrusion prevention of APT is signature type of detection system, such as URL filter and Spam filter. but, the attacker utilize Zero-day Attacks performing invade until it is vulnerability fixed after being reported vulnerability. Or with the human error and human psychological weakness to use Social Engineering to carry out the intrusion. Therefore, it is impossible to completely prevent intrusion. Data exfiltration detection of APT is anomaly type of detection system, such as network traffic anomaly detection. but, data exfiltration detection is difficult due to counterfeiting of communications protocol and encryption of the transmitted data. and it is assumed that situation of attacker can be seen classified information is the information compromise. Therefore, in order to prevent information compromise by APT it is required to be detected during the period from after the intrusion until the attacker to obtain classified information.

I propose file system peek（“fspeek”）that provides anomaly detection by building Tree Structured Log (TSL) from the file access log, tracking and comparison of the TSLs. I propose File Access Scope T-test（FAST）that performs anomaly detection method as a measure of the access range of file access. FAST utilize activities in the file access trend difference between legitimate user and attacker. File access trend can be represented by a range of access in the file access log. FAST quantify the access range of file access as the area of the TSL, and perform a paired T-test the area of TSLs. If there is a period in which the attacker operating a long-term within the system, the area of TSL in the period is increased. Therefore, it is possible anomaly detection by FAST. ”fspeek” is assumed to operate at all times on the system in order to compare the long-term TSL in FAST. The data amount of TSL is kept smaller than the FAL to be recorded in the long term TSL.

“fspeek” implementation conduct comparison experiment of the TSL. A term of TSL is 10 months from February 2015 to November 2015. Read only TSL from February

2015 to June 2015 (TSL A) is referred to as a group A. Read only TSL from July 2015 to November 2015 (TSL B) is referred to as a group B. TSL A include exploratory file access is referred to as a group A_n. TSL B include exploratory file access is referred to as a group B_n. "n" of A_n and B_n represents the number of times per day of file access that mix to read only TSL. I putting "n" =(1,2,3,4,5,6,8,12), calculating the area of each TSL. I conduct paired T-test to define null hypothesis as "no difference between the groups". As an experimental result, in case of significance level α=10%, paired T test(B,A_3),(B,A_4),(B,A_6),(B,A_8),(B,A_{12}),(A,B_6),(A,B_8),(A,B_{12}) are significantly different.