Finding the hidden zombie in your network: Statistical approach to unraveling computer botnets

4 February 2014

How do you detect a "botnet", a network of computers infected with malware -so-called zombies - that allow a third party to take control of those machines? The answer may lie in a statistical tool first published in 1966 and brought into the digital age researchers writing this month in the *International Journal of Electronic Security and Digital Forensics*.

Millions of computers across the globe are infected with malware, despite the best efforts of public awareness campaigns about phishing attacks and antivirus software. Much of the infection is directed towards allowing a third party to take control of a given machine or indeed a network of machines and exploiting them unbeknownst the legitimate users in malicious and criminal activity. Security and software companies do monitor internet activity and there have been many well-publicized successes in destroying such botnets. However, malware writers are always developing new tools and techniques that allow them to infect unprotected computers and rebuild botnets.

Botnets are widely used in organized crime to attempt breaches on security systems by mounting distributed denial of service (dDOS) attacks, among other techniques, on corporate, banking and government systems. Such attacks can open up "backdoors" into a private computer network that lets the botnet controller access proprietary and other sensitive information, passwords or even voting systems. Botnets have also been used for simply malicious purposes to force websites and other services offline, occasionally in an act of protest or rebellion.

Now, R. Anitha and colleagues at PSG College of Technology, Coimbatore, India, have turned to a statistical tool known as the hidden semi-Markov model (HsMM) to help them develop monitoring software that can detect the telltale signs of botnet activity on a computer and so disable the offending malware. In probability theory and statistics, a Markov process is one in which someone can predict the next state of a process based on its current state without knowing the full history of the process. An example in gambling would be that if you have chip now and the odds of winning or losing on the next bet are even then we can predict without knowing how many chips you had earlier that you will either have none or two after the next bet.

A hidden-Markov model would thus include variables of which the observer has no sight but can infer and so predict an outcome. Predicting whether it rained on a given day based on whether a fair-weather-only walker was out on a given day without you having a weather report for their area involves a hidden-Markov process. A hidden semi-Markov model then involves a process of this sort but where the time-elapsed into the current state affects the prediction. For example, one might predict the rainfall pattern based on how long it is since our fair-weather walk last ventured out.

The team has applied the statistical logic of the hidden semi-Markov model to forecast the characteristics of internet activity on a given computer suspected of being a "zombie computer" in a botnet based on management information base (MIB) variables. These variables are the components used to control the flow of data packets in and out of the computer via the internet protocol. Their approach can model the "normal" behavior and then highlight botnet activity as being a deviation from the normal without the specific variables that are altered by the malware being in plain sight.

The team points out that botnet and malware

developers have focused recently on web-based, http, type activity, which is easier to disguise among the myriad packets of data moving to and fro across a network and in and out of a particular computer. Their tests on a small zombie computer network shows that the hidden semi-Markov model they have developed as a lightweight and real-time detection system can see through this disguise easily. If implemented widely such as system could lock down this kind of botnet very quickly and slow the assimilation of zombie computers by criminals and others with malicious intent.

More information: "HTTP botnet detection using hidden semi-Markov model with SNMP MIB variables" in *Int. J. Electronic Security and Digital Forensics*, 2014, 5, 188-200.

Provided by Inderscience Publishers