

# Behavioral Analysis of Zombie Armies

Olivier Thonnard  
Royal Military Academy  
Belgium

Van-Hau Pham  
Eurecom  
Sophia Antipolis, France

Marc Dacier  
Symantec Research  
Sophia Antipolis, France

March 23, 2009

## Abstract

In this paper, we introduce a method to identify, observe and characterize zombie armies on the Internet, based on the traffic collected with very simple, low-cost and easily deployable low-interaction sensors. A zombie army can be a single botnet or a group of botnets the actions of which are coordinated during a given time interval. Still today, zombie armies and botnets constitute, admittedly, one of the main threats on the Internet. Indeed, botnets and zombie machines are the “Swiss army knives” used by cybercriminals to perform various kinds of illegal activities: bulk sending of spams, fraud and extortion, hosting of phishing websites, DDoS, etc. More importantly, the extensive analysis of recent cyber conflicts (such as the presumed cases related to Estonia, Georgia or more recently Gaza) have lead experts to the conclusion that botnets can be easily turned into digital weapons, which can be used by cybercriminals to attack the network resources of a whole country by performing very simple DDoS attacks against critical web services (DNS servers, network routers, government or financial websites, and more). A deep understanding of the long-term behavior of botnet armies and their evolution, is thus a vital requirement to be able to combat effectively those latent threats.

Most previous studies related to botnets have focused on techniques for detecting bots on a network, or analyzed by which infection vectors botnets successfully propagate. This problem is obviously far from being trivial, due to the highly sophisticated structure of today’s botnets: new types of C&C channels, P2P architecture, fast-flux domain hosting, etc. Nevertheless, despite a lot of attention and efforts devoted by researchers to the analysis of botnets, we still lack long-term, strategic information on the global behavior of those large-scale zombie armies that are plaguing the Internet.

In this work, we present a general method that enables us to identify and characterize armies of zombie machines that are very likely controlled by a same entity. In contrast with previous work, we are not interested in studying a particular botnet in details or in the analysis of the various protocols used by bots to communicate with their C&C server. Instead, we want to perform

a long-term strategic analysis of those armies from a behavioral point of view, i.e.: how long do they stay active on the Internet, what is their size, which kind of scanning do they perform, how fast do they seem to move from one set of IP blocks to another, etc. Our approach is based on an appropriate combination of different knowledge discovery and data mining techniques, which consists of the following main components:

1. detection of attack events and characterization of those events (e.g., geolocation, type of activity and targeted subnets),
2. unsupervised clique-based clustering, so as to discover correlations among all attack events,
3. dimensionality reduction techniques, which allow us to visualize and to assess the cliques correlations,
4. a simple fuzzy inference system that leverages the results obtained in the previous steps, in order to identify sequences of attack events that are very likely attributed to the same zombie army.

We present also some preliminary results obtained from a proof-of-concept analysis framework in which we implemented the techniques described here above. The experiments have been performed on a dataset collected with a worldwide distributed honeynet. Our validation dataset comprises about 500 attack events collected on the Internet in a time period spanning from Aug 2006 until November 2008. The key findings we would like to present deal with the behavioral properties of the zombie armies found in our dataset, i.e.: what is their average lifetime and size, how do they evolve over time with respect to their origins or the type of activities they perform. Finally, we provide also some global statistics on the characteristics of the zombie armies observed in our honeynet.