

Semester Thesis:

Detecting Bots in Internet Relay Chat Systems

for Jonas Bolliger <jonasbo@ee.ethz.ch>
and Thomas Kaufmann <thomaska@ee.ethz.ch>

1 Introduction

DDoS in general

Distributed Denial of Service (DDoS) attacks are a threat to Internet services ever since the widely published attacks on ebay.com and amazon.com in 2000. ETH itself was the target of such an attack six months before these commercial sites were hit. ETH suffered repeated complete loss of Internet connectivity ranging from minutes to hours in duration. Massively distributed DDoS attacks have the potential to cause major disruption of Internet functionality up to severely decreasing backbone availability.

Internet Relay Chat and DDoS

It is well known that Internet Relay Chat (IRC) is used not only by humans for chatting but can also serve as a means to send commands to malicious programs (the “bots”) running on compromised hosts (the “zombies”). A person (the “master”) can log into a specific IRC channel, which hundreds or even thousands of bots are listening to, and issue a command such as e.g. *attack <IP address>* that is received and executed by the bots. In this way, the IRC service can be abused to coordinate and launch DDoS attacks.

The DDoSVax Project

In the joint ETH/SWITCH research project “DDoSVax”¹ abstract Internet traffic data (Cisco Netflow) is collected at all border gateway routers operated by SWITCH. This data contains information about which Internet hosts were connected to which others and how much data was

¹See <http://www.tik.ee.ethz.ch/~ddosvax/>

exchanged over which protocols.

For this thesis the DDoSVax research team has established a contact to an administrator of a frequently used IRC system that is temporarily located in the SWITCH network.

2 The Task

Based on tests with real IRC bots, literature research, the results of a previous thesis and traffic measurements on a real IRC server and on routers in the Internet backbone, algorithms that detect bots abusing an IRC system will be developed and validated.

The three following approaches to detect bots and botnets will be considered. Further approaches are optional to this thesis.

- Bots that are installed by a worm will join shortly after each other into the same IRC network. Such *fast joining bots* should be detected in the DDoSVax Netflow data.
- Bots normally stay in an IRC system for a long time. Long standing connections to an IRC server should be detected in the DDoSVax Netflow data.
- Bots are usually not talkative in an IRC channel. Therefore, IRC connections that consists mostly of IRC ping-pong traffic and no real conversation should be detected.

The thesis is divided into four main parts, namely information gathering, IRC traffic measurements, algorithm development and validation.

2.1 Information gathering

Studying the thesis “Analysis of Internet Relay Chat Usage by DDoS Zombies” written by Stéphane Racine is the first step to familiarize with IRC and its possible misuse by bots.

Further literature research on bots and botnets using IRC, studying real bot code, and setting up an own IRC server will give further insights.

2.2 IRC traffic measurements

By using the worldwide distributed computers of Planet-lab, a setup with our own bots connecting to a (test) IRC server will be installed and its traffic measured on the server (tcpdump) and in the backbone (DDoSVax Netflow data). The attack part of these bots (e.g. for denial of service attacks) will be disabled to prevent any misuse. The focus of interest lies on the use of IRC to communicate between a master and the bots.

Other IRC traffic measurements which provide bot-like traffic patterns, that can be used to design and validate bot detection algorithms will be conceived and executed.

Time consuming analysis, especially that of large amounts of DDoSVax Netflow data, will be done on the TIK experimental cluster “Scylla”.

2.3 Algorithm development

With the information gathered and the measurements done, various algorithms to detect characteristic IRC bot traffic (based on the three approaches) will be developed.

Offline algorithms will be run on the cluster “Scylla” or on a workstation. Online algorithms will run as plugins in the UPFrame UDP processing framework provided by the DDoSVax project.

2.4 Validation

The last step will be to thoroughly test the online and offline detection algorithms and to adjust important parameters to reduce false positives.

3 Deliverables

1. *IRC Bot traffic signatures* The bot signatures obtained from literature research, own measurements and tests that describe which characteristics specific bot traffic has.
2. *Offline IRC bot detection algorithms* Design decisions and implementation of the various offline IRC bot detection algorithms.
3. *Online IRC bot detection algorithms* Design decisions and UPFrame plugin implementation of the various online IRC bot detection algorithms.
4. *Thesis Documentation* A concise description of the work conducted in this thesis (task, related work, environment, measurements, results and outlook).

Documentation and Presentation

A documentation that states the steps conducted, lessons learnt, major results and an outlook on future work and unsolved problems has to be written. The code should be documented well enough such that it can be extended by another developer within reasonable time. At the end of the thesis, a presentation will have to be given at TIK that states the core tasks and results of this semester thesis. If important new research results are found, a paper might be written as an extract of the thesis and submitted to a computer network and security conference.

Dates

This semester thesis starts on April 18th, 2004 and is finished by July 9th, 2004.

The milestones of this thesis are:

05/17/2004 Information gathering mostly completed; measurement setups defined; test IRC server is running; outline of detection signatures and algorithms defined

05/24/2004 Intermediate presentation

06/21/2004 Core algorithms are prototypically implemented

07/02/2004 Algorithm tuning and validation finished

07/08/2004 Final presentation

07/09/2004 Delivery of semester thesis documentation and code

Tutors and Supervisor

Tutor: Thomas Dübendorfer, duebendorfer@tik.ee.ethz.ch, +41 1 632 71 96, ETZ G64.1

Co-Tutor: Arno Wagner, wagner@tik.ee.ethz.ch, +41 1 632 70 04, ETZ G64.1

Supervisor: Prof. B. Plattner, plattner@tik.ee.ethz.ch