

Assessing the Attack Threat due to IRC Channels

Robert Meyer* and Michel Cukier**

*Department of Electrical
and Computer Engineering
University of Maryland, College Park
rmeyer@umd.edu

**Center for Risk and Reliability
Department of Mechanical Engineering
University of Maryland, College Park
mcukier@umd.edu

Abstract

This practical experience report presents the results of an investigation into the threat of attacks associated with the chat medium IRC. A combination of simulated users (i.e., bots), some configured with scripts that simulated conversations, and regular users were used. The average number of attacks per day a user on IRC can expect, the effect of channel activity, gender based on the name, and network type on the number of attacks were determined. The social structure of IRC channels and the types of users that use it were analyzed. The results indicate that attacks through IRC channels come from human users selecting targets rather than automated scripts targeting every user in a channel.

1. Introduction and Motivation

Among the chat programs widely used today, the vast majority (e.g., AIM [1], MSN messenger [12], Yahoo [17], ICQ [9]) focus on two-person conversations and require distinct steps to be taken to allow a multi-person chat. IRC [10] is based upon the opposite philosophy, consisting primarily of chat rooms containing as many as several thousand people, and requiring users to preface messages with /msg <user name> to initiate a private conversation. This approach allows for easier communication between large groups of people, but offers the opportunity for attackers to reach large numbers of people quickly. The types of attacks and overall threat of attacks associated with IRC are not well documented yet. The experiment described in this paper was designed to investigate that threat and determine what, if any, factors affected it. The average daily attack frequency was compared among a combination of bots

(i.e., programs that simulate users in a channel) and regular users to determine whether channel activity, gender of username, or the security rules of the network, specifically whether or not the network allowed bots, increased the threat of attack.

The paper is organized as follows. Section 2 provides background information on the IRC protocol and the types of networks and channels that implement it. Section 3 details the setup of the experiment. Section 4 details the methods used to collect and analyze data and the results garnered. Section 5 concludes the paper.

2. Background on IRC

The IRC protocol is implemented by hundreds of networks, each of which operates independently. Each network consists of several dozen servers linked together. Some networks primarily focus on local issues, allowing people in that area to connect with each other. Others provide channels devoted to one specific subject, such as gaming or sports. The largest networks provide some channels for all of these topics, as well as channels with no particular topic intended for general chat. One main difference between the networks is whether or not they provide services to keep registered channels from being taken over when no one is using them. EFNet [5] and IRCNet [11] are the largest networks that do not provide channel services, requiring users to protect their own channels. DALNet [3] provides “full services” including registration of channels of any size and nicknames, while other large networks, such as UnderNet [15], QuakeNet [13], and GalaxyNet [8], fall somewhere in the middle. While most of the large networks provide channels for all subjects, QuakeNet is devoted to the discussion of online games.

Because of the fast connection speeds possible using the XDCC protocol (Xabi Direct Client-to-Client) [16], channels devoted to file sharing (i.e., around 80% of all channels), especially of large files such as movies and pornography, are the most heavily populated by far, with several averaging more than 3,000 users at any given time [14]. A study of the effects of channel activity on attack threat would be rather useless in these types of channels, since the only communication is the constant spamming of the files available. The next most popular channel types are channels used by guilds (i.e., around 10% of all channels), organized groups of players of one or more online games. These channels are restricted to members of the guild and password protected, and without access to the channel, users cannot be placed in it to run the experiment. The remaining channels, those allowing any connection and used for chat, consist of about 10% [14] of the channels of IRC, but were the only ones that could provide meaningful results.

IRC's security status is made more ambiguous by several extra features not directly related to chatting. Through XDCC, IRC users can download files hosted by other users in addition to being able to send files from one person to another. This carries the risk of the file not being what the user intended to download, as well as the possibility of receiving copyrighted material illegally. DCC (Direct Client to Client) [4] allows users to have direct conversations with each other, without the traffic passing through the IRC server. It is also the only way to issue commands to a bot while it is running. Bots are simulated users that are frequently used to act as administrators in channels and also as placeholders, keeping a presence in the channel even when all the human users disconnect to ensure that ownership of the channel does not change. Bots can, however, be used in a variety of attacks against IRC networks and users, flooding a channel with repeated connections and disconnections to make conversation impossible and to increase server load. IRC also uses CTCP (Client to Client Protocol) [2], a protocol that allows users to get information about others, including the version of the client they are running. Note that a simple "version" sent against a bot will provide the name of the software used to create the bot and thus help differentiating simulated users (i.e., bots) from real users.

Starting in 2000, nearly all of the large IRC networks have slowed to a halt at one point or another by massive Denial of Service (DoS) attacks. EFNet was hit hard enough in 2000 and again in 2001 to cause widespread rumor that the network was about to shut down. DALNet had similarly serious problems in 2003. The attacks were conducted mostly through the use of large numbers of bots flooding first channels, then the servers themselves. Once order was restored, the large networks implemented various security policies to prevent future attacks. Some full and partial service networks decided there was no legitimate need for bots, and automatically

closed any connection by a bot (UnderNet, DALNet). Others limit the number of connections allowed per IP, usually to five (QuakeNet, GalaxyNet). Note that during the study, no evidence of DoS attacks was observed (i.e., no connection problem, no massive number of users leaving the network simultaneously).

3. Experimental Setup

The channels chosen for the experiments described in this paper all met the criteria of being used primarily for chat, not requiring passwords for entry and allowing bots. They were #teens (in the GalaxyNet network), #guildwars and #wow (both in QuakeNet), #usa and #allnightcafe (both in UnderNet), #chat and #poker (both in EFNet). The experiments include real and simulated users.

Several different groups have developed open source IRC bots for the purpose of channel administration. Eggdrop [6] and EnergyMech [7] are the two most popular, and they were both given serious consideration because of the large support groups available in case help was needed with the bots. Eggdrop was chosen over EnergyMech because it includes several dozen tcl commands and bindings designed to make scripting for Eggdrop bots easier. The bots used in this experiment had all of the features designed for channel administration disabled.

For the experiments described in this paper, three types of bots were developed: one silent bot, one slightly talkative bot, and one very talkative bot. Each type of bot interacts with another bot of the same type in pairs. Both types of talkative bots ran a simple tcl script that simulated a simple conversation. Each bot listened for a specific private message, and if it received that message, it posted a public message to the channel and then sent a private message to its partner. This message triggered a public message and a private message, and the conversation continued. To implement this the msg bind was used, which upon reception of a private message ran a procedure. In this case the procedure sent the corresponding public message and the private message that continued the conversation. The very talkative bots sent one public message every two minutes for a period of three hours, at which point the conversation looped after a one-hour break. The conversation consisted of the bots telling each other about their activities the previous day, which included on-line gaming and playing basketball. The slightly talkative bots sent messages every two minutes in short bursts that lasted about 10 minutes, with 30 minute breaks between bursts. The conversation involved exchanging greetings and making plans to meet up in real life, which is presumably what happened during the breaks. The bots had names using different letters of the alphabet to keep track of them.

4. Data Collection and Data Analysis

For the purposes of the experiments conducted in this paper, the definition of an attack was restricted to malicious behavior that could occur through IRC. This way, it would be clear that the attack was a result of IRC activity and directed towards a particular bot or user. The set of behaviors defined as attacks included *attempts to send a file to the user, attempted DCC chats with a user, malicious private messages sent to a user, and links sent to a user*. File send attempts and links were considered malicious because both were not solicited and could very easily be viruses, trojans, commands, or other harmful programs. DCC chat attempts were considered malicious because commands can be sent without traveling through the IRC server, and since DCC chat is the method used to send a bot commands during runtime, such connections could indicate attempts to take over bots. A private message was considered malicious if it contained sexually explicit or threatening language, and these were considered attacks for the purpose of this experiment because they indicate unwanted and possibly dangerous attention towards the user from others in the channel. These attacks were collected and separated from one another by the logging capability of the Eggdrop bots. Eggdrop provides several flags that can be used to log different types of messages and commands. For each bot, five logs were kept, including the public text in the channel, private messages to the bot, user joins/kicks, server commands, and file transfers. Public text was logged to see if there was any discussion in the channel generated by the bots' conversations with each other. Joins/kicks were logged to see if attacks were accompanied by the attackers being removed from the channel by channel administrators. The other categories represented the possible types of attacks, with links being logged with private messages. Using this approach, it was clear when an attack occurred, which bot was attacked, and what channel they were in. In the case of a human connection, log files kept by the IRC client used for connection to the server were used.

4.1 Experiment One: Impact of User Activity on the Attack Threat

The first experiment was designed to determine if increased activity in an IRC channel increased the threat of attack. Three pairs of bots were deployed in three different channels (i.e., #teens, #guildwars and #wow), one pair used the very talkative script, one pair used the slightly talkative script, and one pair was completely silent. All three channels had around 400 users at any given time. The silent pair was the control, and the average number of attacks per day they received was compared to the slightly and very talkative bots to see if there was a link between activity and attacks. All bots had male names (i.e., Andy, Brad, Dan, Gregg and

Kevin). The name of the bot was changed from one week to another but was selected among these five names. The logs for each bot were collected and the average number of each type of attack per day, per channel and for each kind of bot was calculated. Since all three channels contained a similar number of users, these results do not need to be normalized by the number of users per channel. The number of different types of attacks per day, shown in Table 1, is based on four weeks of collected data. Table 1 contains the 95% confidence interval around the mean assuming a normal distribution of the different attack types.

Table 1: Number of Attacks per Day for First Experiment

Type of Attack	Type of Bot		
	Silent	Slightly Talkative	Very Talkative
Files Sent	0.10 +/- 0.04	0.83 +/- 0.08	0.98 +/- 0.08
DCC Connections	0.10 +/- 0.04	0.31 +/- 0.05	0.01 +/- 0.01
Malicious Private Messages	3.7 +/- 0.1	2.0 +/- 0.08	1.20 +/- 0.08
Links	2.0 +/- 0.1	2.0 +/- 0.1	2.0 +/- 0.09

All bots received on average one or fewer file connection attempts a day. Note that the average number of attempts increases from silent bots to slightly talkative and to very talkative ones. The number of DCC connections was low for all three types of bots. The silent bots received the highest number of private messages: on average about 2.5 more malicious private messages than the very talkative bots and 1.7 more malicious messages than slightly talkative bots. This result is particularly interesting since intuitively a higher number of attacks would be launched against more talkative bots. The number of links sent to a bot was almost equal for the three types of bots. The results indicated that there was no huge difference in attack frequency between the very talkative, slightly talkative, and silent bots. This experiment showed that increased activity in an IRC channel does not significantly increase the threat of attack.

4.2 Experiment Two: Impact of User Gender on the Attack Threat

The second experiment investigated whether or not the gender of the username had an affect on the number of attacks received. The layout in the channels changed to three silent bots, one with a female name, one with a male name, and one with an ambiguous name. The female names consisted of Cathy, Elyse, Irene, Melissa

and Stephanie. The male names were Andy, Brad, Dan, Gregg and Kevin. The ambiguous names consisted of Nightwolf, Orgoth, Redwings and Stargazer. The name of the bot was changed from one week to another but was selected among these five female, male names and four ambiguous names.

The bots were all silent because, based on the conclusion of the first experiment, channel activity was no longer considered a factor. The three identical channels (i.e., #teens, #guildwars and #wow) were selected. For the bots with female names, six bots (i.e., two per channel) were run for two weeks and three bots (i.e., one per channel) were run for four weeks. For the bots with male names, six bots (i.e., two per channel) were run for four weeks and three bots (i.e., one per channel) were run for another four weeks. For the bots with ambiguous names, three bots were run during four weeks (i.e., one per channel). The number of different types of attacks per day, per channel and for each kind of bot is shown in Table 2. It also contains the 95% confidence interval around the mean assuming a normal distribution of the different attack types. As mentioned for the first experiment, all three channels contained a similar number of users, thus the results do not need to be normalized by the number of users per channel.

Table 2: Number of Attacks per Day for Second Experiment

		Type of Bot		
		Silent Female	Silent Male	Silent Ambiguous
Data Collection Length		6 weeks	8 weeks	4 weeks
Type of Attack	Files Sent	0.40 +/- 0.05	0.10 +/- 0.03	0.38 +/- 0.08
	DCC Connections	0.09 +/- 0.04	0.11 +/- 0.04	0.01 +/- 0.02
	Malicious Private Messages	100.0 +/- 0.5	3.7 +/- 0.1	24.9 +/- 0.7
	Links	2.0 +/- 0.1	1.97 +/- 0.09	2.0 +/- 0.2

The female bots received on average 100 malicious private messages a day, exceeding by far the totals of any of the other bots, with the other attack types being roughly equal. It is interesting to note that the bots with ambiguous names received significantly more malicious private messages (on average 25) than the male bots (on average 3.7), but less than the average between the male and female bots (which is around 52). This experiment shows that the user gender has a significant impact on one component of the attack threat (i.e., the number of malicious private messages received for which the female bots received more than 25 times more private

messages than the male bots and 4 more times than the bots with an ambiguous name) and no significant impact on the other components on the attack threat. Indeed, for each of the three types of bots, on average, less than 0.5 files were sent per day and 0.1 DCC connections and 2 links sent to a bot were observed per day.

4.3 Experiment Three: Comparison of Simulated User and Real User on the Attack Threat

Since the conversation scripts were no longer being used for the bots, there was no longer a definitive reason why their role could not be filled by open human connections. This allowed for the placement of users in channels in networks that did not allow bots, in order to see if those networks had more attacks than the networks already explored with the bots. For two weeks, channels in UnderNet (i.e., #usa and #allnightcafe) and EFNet (i.e., #chat and #poker) were set up similarly to the channels with the bots, one male, one female and one ambiguous, except that all the connections were humans instead of bots. The number of users in each channel was around 300. The name of the users was changed each week and was picked among the set of female/male/ambiguous names listed in Section 4.2. Three users (one female, one male and one ambiguous) were placed in each of the four channels. The number of different types of attacks per day, per channel and for each type of user, shown in Table 3, is based on two weeks of collected data. Table 3 also contains the 95% confidence interval around the mean assuming a normal distribution of the different attack types. For some attack types, no attack was observed for both weeks. For these cases, the confidence interval could not be calculated. These cases are indicated with "N/A" in Table 3.

Table 3: Number of Attacks per Day for Third Experiment

Type of Attack	Type of User		
	Silent Female	Silent Male	Silent Ambiguous
Files Sent	2.5 +/- 0.5	1.8 +/- 0.2	1.2 +/- 0.2
DCC Connections	0 +/- N/A	0.14 +/- 0.09	0 +/- N/A
Malicious Private Messages	163.0 +/- 0.7	27.5 +/- 0.5	65.0 +/- 0.3
Links	6.5 +/- 0.3	5.2 +/- 0.4	5.0 +/- 0.4

The human connections received a lot more attacks than the bots did in the previous experiment. But the differences between male, female and ambiguous

connections remained the same. One difference from the other experiments was that the female usernames were receiving more files and links than the male and ambiguous usernames, though the gap was not nearly as wide as it was for private messages. Indeed, between 1.2 and 2.5 files were sent per day, when, for bots, between 0.1 and 1 files were sent. Moreover, when the number of files sent to female users is higher than the number sent to male users, the lowest was sent to users with ambiguous names. As for experiments 1 and 2, almost no DCC connections were observed. Depending on the user type, between 5.0 and 6.5 links were sent to the user. This is also higher than for the bots where 2 links were observed. When the highest number of links was observed for female users, the users with an ambiguous name received fewer links than the users with a male name. All three types of human users also received more malicious private messages than the associated bots (i.e., 163 versus 100 for female names, 28 versus 4 for male names, and 65 versus 25 for ambiguous names). For human users, female users received about 6 times more private messages than the male users and about 3 times more than the users with an ambiguous name. Overall, the networks that did not allow bots (i.e., UnderNet and EFNNet) seemed to produce more attacks than the networks that did allow bots (i.e., GalaxyNet and QuakeNet). The fact that the two networks used in this experiment produced a difference between the male and female users like that of the other two networks reinforces the findings of the second experiment (Section 4.2).

4.4 Analysis of Results

In addition to helping to assess the attack threat, the experiments provided some insight into the social structure of IRC. Even in channels containing hundreds of people, only a small group of people actually participated in the conversations. The rest of the connections remained idle or broadcast spam into the channel before eventually being removed. When two bots had conversations, they seemed to generate attention for a few minutes, as people were trying to figure out whom the bots were talking to, but after a short while they were ignored. This exposed a flaw in the setup: the bots were unable to communicate with the rest of the channel, and would not be able to do so effectively without more complicated scripts. The extra attention the female usernames received and the nature of the messages (i.e., sexually explicit or threatening language) they were bombarded with suggests that male users outnumber females, as it would be difficult for an automated script to filter usernames based on gender when sending messages. This indicates the male human users specifically targeted female users.

One rather surprising result was the fact that networks that banned bots seemed to have a higher threat

of attack. This goes directly against the hypothesis that the majority of attacks were conducted by automated bots. It is unclear why the networks without bots seem to be less prone to attacks on users. Analysis of the user kicks (i.e., on average 20 kicks/hour on GalaxyNet, 10.7 kicks/hour on QuakeNet, 16.4 kicks/hour on EFNNet, and 14.2 kicks/hour on UnderNet) indicated a higher rate of users who committed attacks getting removed from channels in the networks that allowed bots. Since the bots can be used as administrators, it is possible that they assist in cutting down on spam and other forms of harassment. However, the networks that did not allow bots did provide administration bots to registered channels.

One constant throughout the results of all three setups was the relative frequencies of the different types of attacks. DCC connections were the least common, probably because they can only be effectively used as attacks against bots. File connection attempts were also rare, and significantly less common than suspicious links. This is probably because both can perform the function of sending a virus to another computer, and links have the benefits of not requiring the other user to accept the transaction and the ability to send commands as well as files. Note that file sent, DCC connections, and links were not investigated further (i.e., by opening them) because we did not have the necessary apparatus to contain the potential infection. The most common attack by far was malicious private messages (i.e., sexually explicit or threatening language). This is probably the most prevalent attack because it does not require programming knowledge, malicious code, or, for that matter, anything but an IRC client to send. Among the private messages, on average, we found 30% of malicious ones for the female bots, 24% for the male bots, 23% for the ambiguous bots, 28% for the female human users, 26% for the male human users, and 25% for the ambiguous human users. Most other private messages include “hello” or “how are you doing?” messages that typically would be followed by a malicious private message (i.e., sexually explicit or threatening language).

5. Conclusions

In summary, the threat of attack on IRC seems to be rather low. The only type of attack that occurs consistently daily is malicious private messages, and in and of themselves they pose no threat to computer security. This threat does not seem to depend on whether or not a user is active in a channel. Users with female names are, however, far more likely to receive malicious private messages, slightly more likely to receive files and links, and equally likely to be attacked in other ways. This implies that the attacks are carried out by humans selecting targets rather than automated scripts sending attacks to everyone in the channel. Users with

ambiguous names are far less likely to receive malicious private messages than female users, but more likely to receive them than male users. Users in channels that do not allow bots at all are more likely to receive attacks than users in channels that allow a minimal number of bots.

References

- [1] <http://www.aim.com/>
- [2] <http://en.wikipedia.org/wiki/CTCP>
- [3] <http://irchelp.org/irchelp/networks/servers/dalnet.html>
- [4] <http://en.wikipedia.org/wiki/DCC>
- [5] www.efnet.org
- [6] www.eggheads.org
- [7] www.energymech.net
- [8] www.galaxynet.org
- [9] www.icq.com
- [10] <http://www.irchelp.org/irchelp/new2irc.html>
- [11] www.ircnet.org
- [12] <http://messenger.msn.com/>
- [13] www.quakenet.org
- [14] <http://searchirc.com/top100.php>
- [15] www.undernet.org
- [16] <http://en.wikipedia.org/wiki/XDCC>
- [17] <http://messenger.yahoo.com/>