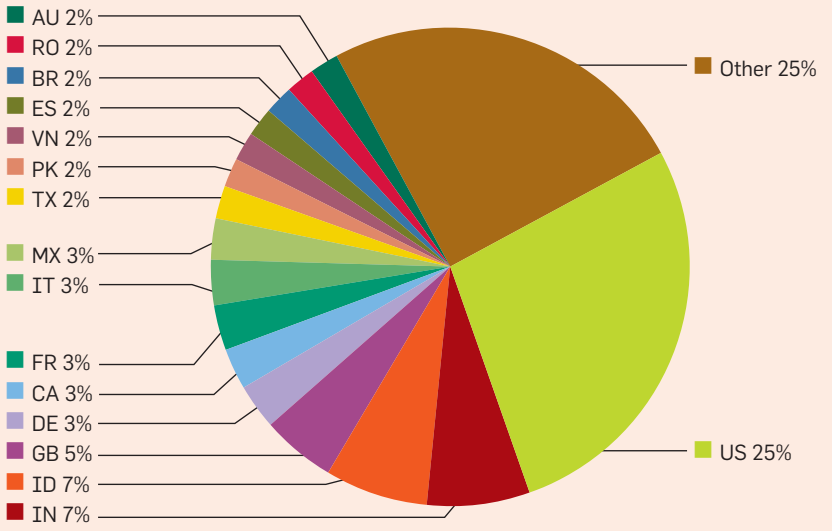
Society | DOI:10.1145/2076450.2076456 Samuel Greengard

The War Against Botnets

Increasingly sophisticated botnets have emerged during the last several years. However, security researchers, businesses, and governments are attacking botnets from a number of different angles—and sometimes winning.

Distribution of TDL4-infected computers by country.

The TDL4 botnet had infected 4,524,488 computers in the first three months of 2011, according to an analysis by Kaspersky Lab. No Russian users appear in the statistics because affiliate marketing programs do not pay for infecting computers located in Russia.



AT THIS MOMENT, millions of people around the world are happily using their computers without realizing their system has been hijacked. While they are answering email or browsing Web sites, cybercriminals are surreptitiously using their computer to wreak havoc across the Internet—and beyond. The computer may be unleashing a torrent of spam, inflicting denial-of-service (DoS) attacks, or engaging in other malicious acts, such as stealing passwords, trade secrets, and personal information.

Welcome to the nefarious world of botnets. "The sophistication level is getting ratcheted up," states Merrick Furst, Distinguished Professor at Georgia Tech's College of Computing. Botnet code, which usually ranges from a few hundred to a few thousand bytes, infiltrates computer systems and provides a means of hijacking or controlling those machines to perform illegal tasks. Between 5%–10% of all Internet Protocol addresses have devices infected with some form of malware. Bots generate the vast majority of the spam that clogs computers and networks, Furst says.

Last June, TDL4—the fourth generation of a Windows PC botnet that first appeared in 2008—surfaced. It is difficult to detect and, so far, impossible to stop.

But the stakes have suddenly grown. Last June, TDL4—the fourth generation of a Windows PC botnet that first appeared in 2008—surfaced. Researchers noticed immediately that the TDL4 botnet, also referred to as TDL-4 and TDSS (a string of characters that the malware generated when it dropped component files and registry entries in earlier versions), is difficult to detect and, so far, impossible to stop. The code hijacks a personal computer, eludes anti-malware applications, and uses custom encryption to control computers. Antivirus company Kaspersky Lab has described TDL4 as the "most sophisticated threat today." It is believed to have infected five million or more computers.

Under Attack

Although viruses, worms, Trojan horses, and rootkits have all made their mark over the last decade, security analysts are increasingly concerned about the impact of botnets. Many use social engineering techniques—including promises of photos of naked celebrities or free movies or music—that trick individuals into clicking a link or downloading a file. Furst says botmasters also use Google AdWords and banner ads to entice surfers to follow links or visit URLs that download bot code to their computer. As quickly as organizations like Google snuff out the fake ads, new ones appear.

Roel Schouwenberg, a senior researcher at Kaspersky Lab, says that while there is no way to know the exact scope of the problem, it is likely that hundreds of millions of computers worldwide belong to botnets. Making matters worse is that the perpetrators are not just hackers looking to take over computers and generate spam. "The bots are created and spread by crimi-

nals looking to monetize machines," says Schouwenberg. "They're increasingly looking to steal identities, credit card numbers, and trade secrets."

These computers—referred to as zombie systems—lie in a dormant state until the perpetrator of the attacks—a.k.a. the "botmaster"—decides to unleash them. Then, with stealth abandon, they pursue their mission without users of infected computers knowing. "These types of programs surreptitiously spread themselves and create networks that are far more powerful than groups of independently infected systems," observes Ira Winkler, a computer consultant and former U.S. National Security Agency analyst who is author of *Spies Among Us: How to Stop the Spies, Terrorists, Hackers, and Criminals You Don't Even Know You Encounter Every Day*.

But TDL4 is raising the stakes and ushering in a new era of complexity and risk. "TDL4 is an evolution in the design of botnets," says A.L. Narasimha Reddy, J.W. Runyon Professor in the Department of Electrical and Computer Engineering at Texas A&M University. "It combines rootkit and botnet features. TDL4 seems to communicate directly with command and control servers [that crooks use to manage the bot network], but it also communicates through a peer-to-peer network."

TDL4 relies on an array of methods to evade the signature, heuristic, and proactive detection that antivirus programs typically use. It loads before the operating system starts. Communication between the command center and the bots is encrypted, and when TDL4 detects that a command and control server has been taken out, it automatically locates another server. In addition, a rootkit hides the presence of other types of malware in the code. Kaspersky Lab security analysts Sergey Golovanov and Igor Soumenkov say they have found evidence of a malicious SHIZ program embedded in the TDL4 code. It causes search engine redirects to sites that download malware and infect other systems.

In fact, Kaspersky Lab describes TDL4 as an attempt to design indestructible malware. Although the antivirus company offers an anti-rootkit utility named TDSSKiller that cleans an infected system, eliminating TDL4 is almost impossible. "The botnet uses a hybrid approach that makes it next to impossible to take down," Schouwenberg says. "It is extremely stealthy—you don't see any of its files on a system—and you don't see any [typical] traffic on an infected system. This, combined with peer-to-peer capabilities, changes the stakes."

"The bots are created and spread by criminals looking to monetize machines," says Roel Schouwenberg. "They're increasingly looking to steal identities, credit card numbers, and trade secrets."

The TDL4 plague appears to be growing worse. In October, researchers discovered the malware was being rewritten and modified to make it more resilient to antivirus detection and removal. Some components, including its kernel-mode driver and user-mode payload, had changed. Moreover, the rootkit now creates a rogue partition on a computer to ensure its components are intact.

"The damage TDL4 can do is the same as what other botnets could do in the past," says Leyla Bilge, a researcher at International Secure System Lab. "They could steal information, they could attack another system, and they could send spam." The difference, she says, is that it is becoming more difficult to protect systems against botnets and put malware out of operation using conventional detection and mitigation tools. "We are engaged in an arms race," she concludes.

Defense Mechanisms

Battling botnets is an increasingly difficult proposition. As a result, researchers are attacking the problem from a number of different angles. At Texas A&M University, Reddy and a team of researchers have developed a technique that detects botnets—including more common varieties like Conficker, Kraken, and Torpig—that rely on so-called domain-fluxing or polymorphic Domain Name System (DNS) methods to evade detection. Domain-fluxing bots typically generate random domain names, but only one of them is real. For instance, Conficker-C generates up to 50,000 random domain names per hour. This makes it extraordinarily difficult to find, identify, and eradicate the bot.

Researchers typically reverse-engineer bot malware to understand the domain generation process and reach the command and control server. However, it is a slow and tedious task. Instead, Reddy's approach examines the pattern and distribution of alphabetic characters in a domain name to determine whether it is malicious or legitimate. The researchers analyze DNS traffic to spot domain names the bot generates through an algorithm. "It can detect previously unknown botnets by analyzing a small fraction of the network traffic," Reddy says. In fact, Carnegie Mellon University's CERT research lab plans to distribute a tool based on this approach.

Bilge and other researchers at International Secure System Lab are taking a different approach. They have developed a tool called Exposure that identifies malicious domains by analyzing traffic patterns and abnormalities within the DNS. Telltale signs include domains that suddenly appear and then disappear immediately following an attack and domains comprised of many numbers but few meaningful words. Exposure sorts through data from known malicious domains and, once trained, uses the patterns to recognize botnet sites. In a test with a French Internet service provider (ISP), it ferreted out more than 3,000 malicious domains.

Meanwhile, Zhi-Li Zhang, Quest Chair Professor in the Department of Computer Science and Engineering at the University of Minnesota, is focusing on alternating and changing domain names in a different way. In order to spot botnets, Zhang's research team examines failed DNS queries.

When a large number of failures occur it is frequently a sign that a botnet exists. The challenge, he admits, is identifying which failures are caused by bots rather than other factors. Nevertheless, "the more we can disrupt the command and control channel, the more difficult it is for the bot to succeed," Zhang says.

Over the last few years, law enforcement agencies and a handful of companies, most notably Microsoft, have successfully taken down bots. In September, for example, Microsoft turned to the U.S. federal court system to shut down command and control servers running the Kelihos botnet. The company obtained a court order to pull the plug on 21 domains associated with the botnet, which is suspected of controlling 50,000 or more zombie machines. Microsoft was also instrumental in a takedown of the Rustock botnet earlier in 2011 and the Waledac botnet in 2010.

However, the botnet scourge is not going away anytime soon. Bilge believes future botnets will be even more sophisticated and problematic. Indeed, "the infection medium is likely to switch from computers to mobile phones. Malware authors are always attracted to the systems that are most widely used," she notes. Another potential problem is cloud computing environments, which, lacking adequate protection, have put powerful resources at the fingertips of cybercrooks, enabling them to spread malware even more quickly.

In order to spot botnets, Zhi-Li Zhang's research team examines failed DNS queries. When a large number of failures occur, it is frequently a sign that a botnet exists.

Zhang says better DNS security is required, including name authentication. "It would make things easier if we could detect unsavory Web sites, or other domain names associated with command and control servers." In addition, "we have to figure out better ways to monitor and filter suspicious traffic," he says. "The difficult thing is that the Internet is diverse and traffic travels all over the world. Getting people to protect machines and ISPs to monitor sites is extraordinarily difficult." Unfortunately, in many cases, there is little or no economic incentive to protect these devices and sites.

In the end, a coordinated approach—using technology tools such as honeypots, traffic analysis, and binary analysis combined with increased law enforcement and user awareness—offers the best chance for success. Yet as long as humans can be duped into clicking a malicious link or opening an infected file, the problem will persist. Perhaps the most important question is how big will the problem become before governments, ISPs, and the computer science community takes the problem seriously.

Further Reading

Bhatia, J.S., Sehgal, R.K., and Kumar, S.
Honeynet based botnet detection using command signatures, *Advances in Wireless, Mobile Networks and Applications*, Al-Majeed, S.S., Hu, C.-L., and Nagamalai, D. (Eds.), Springer-Verlag, Berlin and Heidelberg, Germany, 2011.

Yadav, S., Reddy, A.K.K., Reddy, A.L.N., and Ranja, S.
Detecting algorithmically generated malicious domain names, 2010 Internet Measurement Conference, Melbourne, Australia, Nov. 1–3, 2010.

Yin, C.Y., Ghorbani, A.A., and Sun, R.X.
Research on new botnet detection strategy based on information materials, *Advanced Materials Research 282–283*, July 2011.

Zhang, J., Luo, X., Perdisci, R., Gu, G., Lee, W., and Feamster, N.
Boosting the scalability of botnet detection using adaptive traffic sampling, *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, Hong Kong, March 22–24, 2011.

Samuel Greengard is an author and journalist based in West Linn, OR.

Government

European Union's Open Data Initiative

While a few European Union (EU) countries—including France, the Netherlands, and the United Kingdom—currently make their government data available to the public via the Web, most EU members do not. That may soon change.

If approved by the member states, new open data proposals by Neelie Kroes, the EU's digital agenda commissioner, would force all EU countries to make their government data digitally available instead of on paper as most do now.

"We are sending a strong signal to administrations that your data is worth more if you give it away ... so start releasing it now," Kroes announced at the initiative's launch. "Taxpayers have already paid for this information; the least we can do is give it back to those who want to use it in new ways that help people and create jobs and growth."

The data represents a huge opportunity for tech companies, which could be worth $53.5 billion annually to the EU's economy, she added. It is expected tech startups could benefit most from the new rules by turning the raw data into smartphone apps, such as maps, real-time traffic and weather information, price comparison tools, and more.

In an email interview, Ryan Heath, an EU Commission spokesperson, reports that "the proposals received favorable initial reactions from member states in the Telecoms Council meeting of Dec. 13, but of course the real challenge is in the details and implementing the new rules once there is agreement. So we are at the start of a long journey."

In terms of timing, Heath says "we hope to achieve agreement within a year from the European Parliament and Council, and then each member state needs time to 'transpose' agreement into all their relevant national laws. So that process will still be going on in 2013.

"We hope various public authorities or national governments take the hint and get on with releasing data voluntarily in parallel with making the formal legal changes," he adds.

—Paul Hyman