Passive Monitoring of DNS Anomalies
(Extended Abstract)

Bojan Zdrnja¹, Nevil Brownlee¹, and Duane Wessels²

¹ University of Auckland, New Zealand
{b.zdrnja,nevil}@auckland.ac.nz
² The Measurement Factory, Inc.
wessels@packet-pushers.com

Abstract. We collected DNS responses at the University of Auckland Internet gateway in an SQL database, and analyzed them to detect unusual behaviour. Our DNS response data have included typo squatter domains, fast flux domains and domains being (ab)used by spammers. We observe that current attempts to reduce spam have greatly increased the number of A records being resolved. We also observe that the data locality of DNS requests diminishes because of domains advertised in spam.

1 Introduction

The Domain Name System (DNS) service is critical for the normal functioning of almost all Internet services. Although the Internet Protocol (IP) does not need DNS for operation, users need to distinguish machines by their names so the DNS protocol is needed to resolve names to IP addresses (and vice versa).

The main requirements on the DNS are scalability and availability. The DNS name space is divided into multiple zones, which are a “variable depth tree” [1]. This way, a particular DNS server is authoritative only for its (own) zone, and each organization is given a specific zone in the DNS hierarchy. A complete domain name for a node is called a Fully Qualified Domain Name (FQDN). An FQDN defines a complete path for a domain name starting on the leaf (the host name) all the way to the root of the tree. Each node in the tree has its label that defines the zone. An example of an FQDN is “www.auckland.ac.nz.”. A domain is a subdomain when it is contained in another domain; in the previous example “auckland.ac.nz” is a subdomain of “ac.nz”.

As DNS is not centrally controlled, the domain names can be abused by attackers outside any organization. Besides domain name trading, attackers can shift domain name records quickly, making access blocking difficult. Another advantage for attackers is that from the client point of view security is often relaxed around DNS traffic, even in tightly controlled organizational networks. Most organizations have strict firewall policies at least on their perimeter firewalls, but DNS traffic is usually unrestricted because it is used by many other protocols. Attackers are commonly abusing this fact, not only to covertly send data over DNS, but also to deploy rogue DNS servers that can be used to completely control victim’s Internet behavior.
This paper describes a passive DNS anomaly detection project based on data captured at the University of Auckland Internet gateway. Our original motivation for deploying the passive DNS monitor was to detect and correlate domains used for botnet controls. We quickly realized that the database is also a rich source of information about spam, anti-spamming tools, typosquatting, and other anomalies.

2 Related Work

Florian Weimer presented a passive DNS replication project at the FIRST 2005 conference [17]. As a result of his project a web site was established by RUS CERT [2] that allows public access to data collected “from the public Domain Name Service (DNS) system.” Weimer’s software, dnslogger consists of sensors deployed around a network that send captured DNS responses to a central collection service. Sensors encapsulate captured DNS responses in new UDP packets which are then relayed (in real time) to the collector. The collector analyzes received UDP packets and imports them into a database. Weimer’s passive DNS replication project is very similar to the one deployed at the University of Auckland, however, our setup is simpler and our database stores more information, for a longer period of time.

The University of Amsterdam [3] based their DNS capture project on Weimer’s work. Schonewille et al modified Weimer’s program to capture outgoing DNS queries in order to identify machines in the local network that have been compromised. Malware-infected machines tend to emit DNS queries that allow them to be easily identified.

John Kristoff’s DNSwatch [4] software can be used in a similar manner, as described by Elton et al [5], but it requires an external black list of well known malicious IP addresses (servers used to spread malware or contacted by malware).

3 Data Capture Methodology

DNS traffic uses either UDP or TCP on port 53 for communication [7]. Most DNS communication happens over UDP, which is the default protocol used by resolvers, i.e. applications that communicate with DNS servers on behalf of other applications when they need to resolve a DNS query. TCP was originally used only for zone transfers, but RFC 1123 [18] expanded the use of TCP as a backup communication protocol when the answer needs to be larger than 512 octets. In cases like this, the first UDP DNS response contains only partial answers. The truncation bit is set so that the resolver can repeat the query over TCP. However, RFC 2671, “EDNS0” [19], defined a new opcode field/pseudo resource record that allows UDP DNS traffic to be bigger than 512 octets. Because almost all of today’s DNS traffic uses UDP as its transport protocol, the deployment at the University of Auckland ignores TCP traffic.