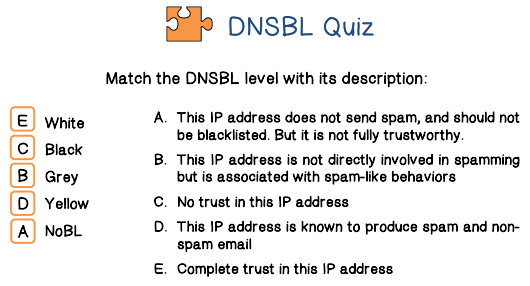
The recommended reading for this lesson:

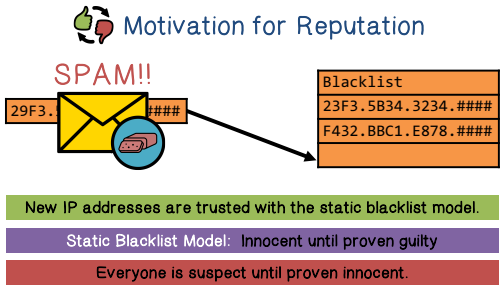
1. Manos Antonakakis, Roberto Perdisci, David Dagon, Wenke Lee, and Nick Feamster. Building a Dynamic Reputation System for DNS. In Proceedings of the 19th USENIX Security Symposium, Washington, DC, August 2010.
2. Charles Lever, Manos Antonakakis, Bradley Reaves, Patrick Traynor, and Wenke Lee. The Core of the Matter: Analyzing Malicious Traffic in Cellular Carriers. In Proceedings of the 20th Annual Network and Distributed System Security Symposium (NDSS), San Diego, CA, February 2013.

In this lesson, we will discuss domain and network reputation. We will learn who to trust and, more importantly, who not to trust. Hopefully, we can figure out who is disreputable as soon as they are on the Internet. In other words, we're going to do all those things that we're told not to do: prejudge, profile and stereotype in order to stop hackers and protect the Internet.

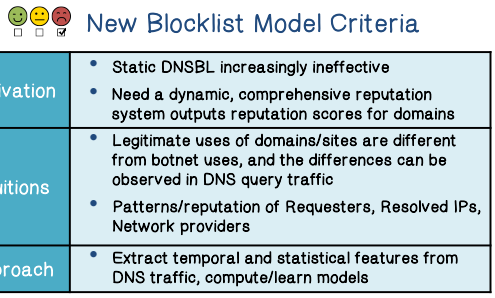


Before we discuss domain name reputations, let's do a quiz on DNBSL. DNSBL stands for DNS black list. Match the DNSBL level with its description.

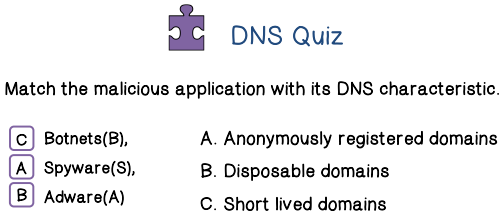
White means that this IP address is completely trusted. Black means that there's no trust in this IP address. Gray means that this IP address is not directly involved in spamming but is associated with spam-like behaviors. In other words, it's in the gray area. Yellow means that this IP address is known to produce spam and non-spam e-mail. NoBL means that this IP address does not send spam and should not be blacklisted, but it is not fully trustworthy.



Traditionally, when an IP address is discovered to have been used to send spam, it is added to a blacklist. This is called a static blacklist. If an IP address is in a blacklist, then emails coming from this IP address can be blocked. This is a great idea except that attackers also know about the blacklist. So spammers can circumvent the blacklist by using new IP addresses, and by the time an IP address is discovered to have been used to send spam and edit to blacklist, the spammers can then move on to use a new IP address. We need to change this model to be in line with our philosophy about network security, and that is an IP address should not be trusted by default.

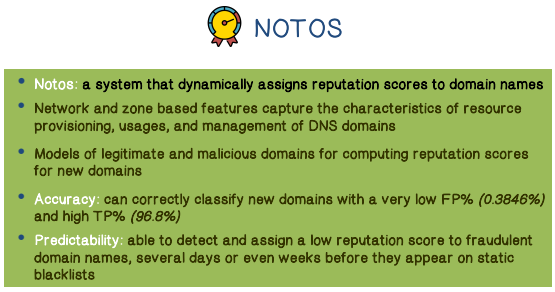


Now, let's discuss a new blocklist model. We know that static blocklist are increasingly ineffective. What we need is a dynamic reputation system that outputs the reputation score or the trustworthiness of a domain. The intuition of this new model is that legitimate use of domains and sites are different for malicious use. The question is, how do we observe such differences? Our intuition is that such differences can be observed in DNS traffic. For example, we can look at the patterns of requests and the reputation of the requester, the resolved IPs, and the network providers for these domains. Therefore, our approach is to analyze DNS traffic, extract temporal and statistical features and then apply machine learning algorithms to learn models that can provide the dynamic score of a domain.

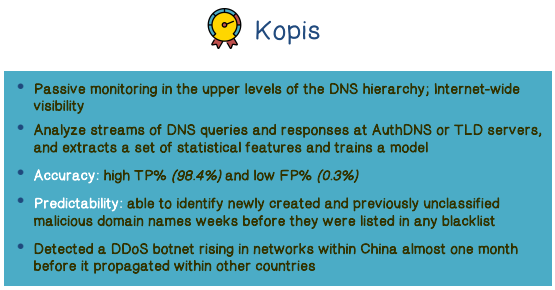


Now, let's do a quiz on how malicious applications use DNS domains. Match the malicious application with its DNS characteristics.

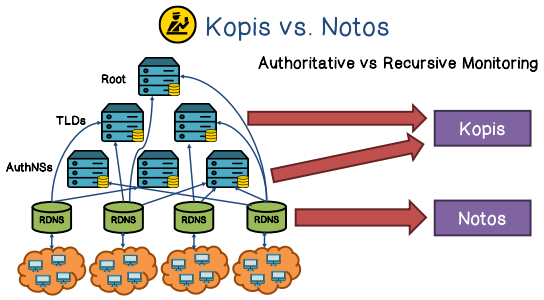
A botnet typically has a set of domains at his disposal. So, each domain is only used for short period of time. A spyware is used to steal information, and it needs to upload information to a site. This site is typically registered anonymously. Adaware uses domains. They are not associated with legitimate businesses. Therefore, these domains are disposable.



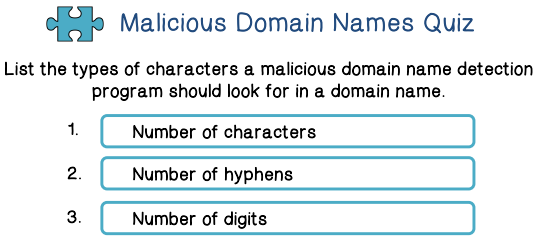
Now, let's discuss two DNS repetition systems. The first one is NOTOS. NOTOS is a system that dynamically assigns repetition score to a domain name. It uses features that capture the characteristics of resource provisioning, usages, and management of DNS domains. Given examples of legitimate and malicious domains, NOTOS uses machine learning algorithm to compute a scoring function based on these features. This scoring function can output a repetition score for new domain. In our study, we are shown that NOTOS has very high accuracy. That is, it has a very low false positive rate and very high true positive rate. That is, NOTOS can detect a fraudulent or malicious domain weeks or days before it is widely used. NOTOS can be applied to a large network, for example, a particular ISP.



Whereas another system called Kopis has Internet-wide visibility because it performs monitoring at the upper levels of the DNS hierarchy. That is, Kopis uses DNS traffics at authoritative DNS or top-level domain DNS service. Similar to NOTOS, Kopis also extract a set of statistical and temporal features from the traffic and then uses machine learning algorithm to chain a scoring function. Similar to NOTOS, Kopis also has very high accuracy rate. Similar to NOTOS, Kopis can also detect malicious domains weeks before they were listed in any blacklist. Since Kopis has the Internet-wide visibility, it can detect that a domain is being used by malware within one country months or weeks before the malware begins to spread to other countries. For example, in our study, Kopis was able to detect a DDoS botnet rising in networks within China almost one month before it propagated to other countries.



To summarize what we have discussed so far, Notos can be deployed at Large Local Network, such as an ISP. In other words, Notos use traffic to do recursive DNS servers whereas Kopis is deployed at the upper level of the DNS hierarchy. In particular, the authoritative name servers or the top-level domain servers. In other words, Notos has local views and Kopis can have a global Internet view.



One of the methods used to detect malicious domain names involves name analysis. Domain names are analyzed to determine the likelihood that a name is used or created for not legitimate purposes. Lists of types of characters a malicious domain name detection program should look for in a domain name.

We can look for the number of characters. Malicious domain names tend to be long. We can also look for the number of hyphens. Again, malicious domain names tend to have a lot more hyphens. We can look the number of digits. Again, malicious domain names tend to have a lot more digits. Similarly, with numbers.

Notation and Terminology

RR	Resource Record	www.example.com 192.0.32.10
2LD, 3LD	2nd and 3rd level domain	2LD = example.com 3LD = www.example.com
RHIPs	Related Historic IPs	All "routable" IPs historically mapped with the domain name in the RR or any domain name under the 2LD and 3LD
RHDNs	Related Historic Domains	All fully qualified domain names (FQDN) that historically have been linked with the IP in the RR, its corresponding CIDR and AS
ADNT	Authoritative domain name tuple	The requester (or RDNS), the domain name and the RDATA

Before we discuss the details of notice and coppice, let's go over some of the notations and terminologies. We use RR to represent resource record. It's a tuple of domain name and its reserved IP address. For domain, www.example.com. 2LD is example.com, and 3LD is www.example.com. The related historic IPs of a domain are all the routable IPs historically mapped with this domain name, and any other domain name, within the same 2LD and 3LD. Related historic domains of an IP address are all the fully qualified domain names that historically had been linked with this IP, and also, all the IPs with this address box, and the autonomous systems. The authoritative domain name tuple is the requester for example, that because of DNS, the domain name and RDATA, which includes all information about this domain.

Data Used in Research

Passive DNS *(pDNS)* data collection is the harvesting of successful DNS resolutions that can be observed in a given network

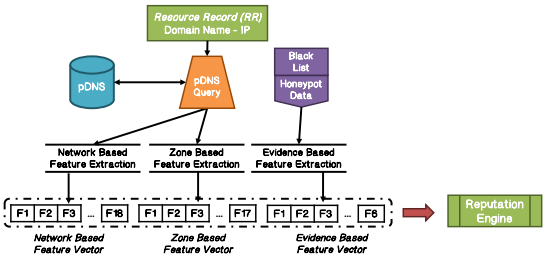
Passive DNS database contains traffic from several ISP sensors and data repositories

- Observed that different classes of zones demonstrate different passive DNS behaviors

Obtained authoritative DNS traffic from 2 large authoritative DNS servers (AuthNS) and the Canadian TLD

Notos uses passive DNS data collected at a recursive DNS server. For example, this can be your ISP. In our study, we use data from multiple ISPs and data repositories. For coppice, we use data from two large authoritative DNS servers, and the Canadian top-level domain server.

Statistical Features of Notos



Now, let's discuss the temporal and statistical features of Notos. Given a resource record, and that is a tuple of domain name and its resolve IP, Notos uses the passive DNS data to extract network based features. These features are based on the related historic IPs. It also extracts the zone features. These features are based on the related historic domains. Notos also constructs the so-called evidence features by analyzing the blacklist and honeypot data, and look for evidence associated with these IPs and domains. These features are then combined and forwarded to the reputation engine which computes a reputation score for this domain.

DNS Database Quiz

The information extracted from the pDNS database can be grouped into three categories. Match the category to its definition.

C Network-based features(N),
B Zone-based features(Z),
A Evidence-based features (E)

A. The number of distinct malware samples that connected to any of the IPs.
B. The average length of domain names, the occurrence frequency of different characters, etc.
C. Quantities such as the total number of IPs historically associated with the diversity of their geographical locations, the number of distinct autonomous systems, etc.

Now, let's do a quiz. We have just discussed that the information extracted from the passive DNS database can be grouped into three categories of features. Match the category to its definition.

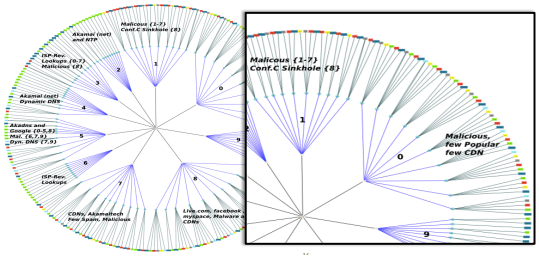
Network-based features. These include, the total number of IPs historically associated with a domain, or the number of distinct autonomous systems. Zone-based features. These include, the average length of domain names, the occurrence frequency of different characters, etcetera. Evidence-based features. These include, the number of distinct malware samples.

Notos Statistical Features

Network-Based Features:	• Extracted from the set RHIPs • E.g., the total number of IPs historically associated with a domain, the diversity of their geographical location, the number of distinct autonomous systems (ASs) in which they reside, etc.
Zone-Based Features:	• Extracted from the set RHDNs. • E.g., the average length of domain names in RHDNs, the number of distinct TLDs, the occurrence frequency of different characters, etc
Evidence-Based Features:	• E.g., the number of distinct malware samples that contacted the domain, and the same for any of the resolved IPs, etc.

Here's a summary of Notos statistical and temporal features. The network-based features are extracted from the set of related historic IPs. The features include, the total number of IPs historically associated with a domain, the diversity of their geographical locations, the number of distinct autonomous systems, and so on. The zone-based features are extracted from the set of related historic domains. The features include, the average length of domain names, the number of distinct top-level domains, the occurrence frequency of different characters, and so on. The evidence-based features include, the number of distinct malware samples that contacted the domain, and the same for any of its resolved IPs.

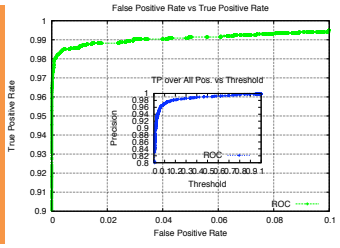
Clusters of DNS Domains



Here we show the result of clustering 250,000 domains using the Noto features. As we can see, the clusters don't overlap. This tells us that the Notos features are very good at separating domains of different types into different clusters.

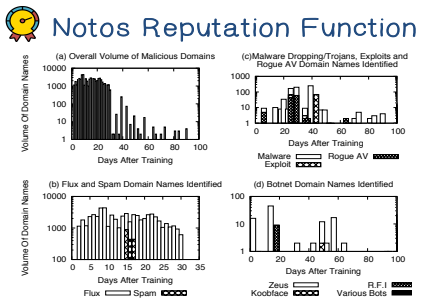
Notos Reputation Function

Given domains known to be legitimate and malicious:
- Gather Notos features for each domain
- Malicious domain label: 1
- Legitimate domain label: 0
- Learn a function that given the Notos feature vector for a domain, outputs a label (0 or 1)
- Reputation score is the "confidence" of the label (or, the probability that the domain is malicious)

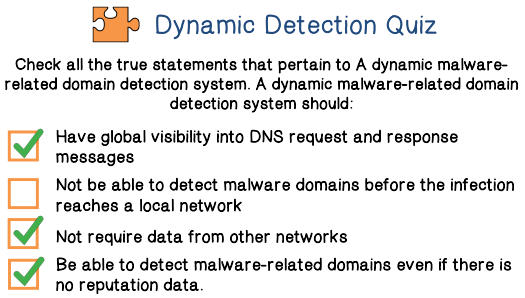


Now, let's discuss how Notos computes reputation of a domain. First, Notos needs to chain a repetition function, given a set of domains known to be legitimate or malicious. Notos extract features for each of these domains and give label 1 to malicious domains, and label 0 to legitimate domains. Then, given this training data set, Notos

can use a machine learning algorithm to learn a function that, given the Notos feature vector of a domain, it will output a label, meaning 0 or 1, and the reputation score of this domain is simply the confidence of this label. That is, the probability that this domain is malicious. As we can see here, this function is very accurate. That is, it has a very high true positive array and very low false positive rate.

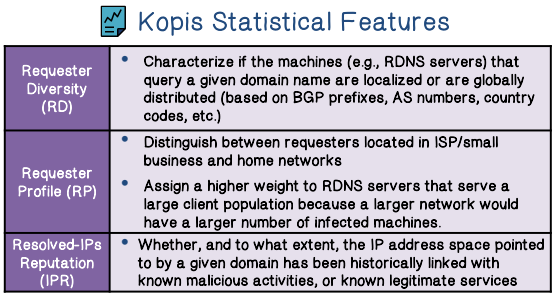


Here, we show that Notos can detect many malicious domains days or weeks before they show up in any blacklist. Sometimes that's even months before they show up in any blacklist. This is true for all the malicious domains and also for different types of malicious domains.



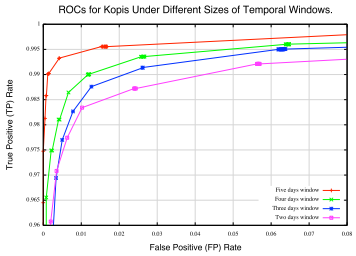
Let's do a quiz. Check all true statements, that pertain to a dynamic, malware-related domain detection system.

A dynamic malware-related domain detection system should: have global visibility into DNS request and response messages, it should not require data from other networks, should be able to detect malware-related domains even if there's no repetition data. Such a system should also be able to detect malware domains before the infection which is a local network. Therefore, the second statement is false.



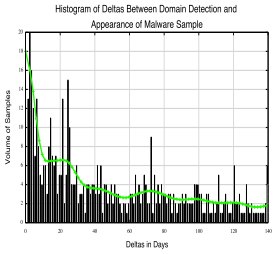
Now, let's take a look at Kopis. Recall that Kopis is similar to Notos except that Kopis is deployed at the upper level of the DNS hierarchy. Therefore, the main difference between Kopis and Notos is in their features. In particular, Kopis analyzes requester diversity. This requester diversity features characterize if the machines that query a given domain name are localized or are globally distributed. The requester profile features distinguish request from a small network versus a large network because a larger network would have a larger number of infected machines. The resolved IP reputation futures look at whether the IP address space pointed to by a given domain name has been historically linked with known malicious activities.

Kopis Detection Performance



Similar to Notos, given examples of malicious and legitimate domains, Kopis can use machine learning algorithm to learn a scoring function based on the features. Here, we show the detection performance of the scoring function. As we can see here, if we use more data, meaning using a longer time window, the accuracy is higher. In particular, with a five-day time window, we can achieve very high detection rate and very low false positive rate.

Predictability



Here, we show that Kopis can detect many malicious domains days, weeks or even months, before they show up in any blacklist.

Study of Mobile Malware Prevalence

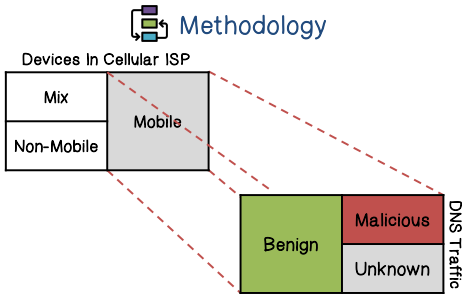
Motivation	• Much work on mobile malware has been on analysis of (malicious) mobile apps • But, how prevalence are infections on mobile devices?
Intuitions	• The (malicious) mobile web is a part of the (malicious) web • Mobile malware uses similar infrastructure (C&C) techniques as non-mobile/Internet malware
Approach	• Obtain DNS traffic in cellular network and identify domains looked up by mobile apps • Analyze information related to the Internet hosts pointed by these domains.

Now, let's discuss a study of mobile malware prevalence using domain amputation. The motivation of our study is that most research has been focusing on analyzing malicious mobile apps, but the question remains, how prevalence are the infections on mobile devices? Our intuition is that mobile web is actually part of the regular web and therefore mobile malware will use similar command control infrastructure and therefore our approach is to analyze DNS traffic obtained from cellular network providers and identify domains looked up by mobile apps. Then for the internet machines that host these domains, we analyze their reputation.

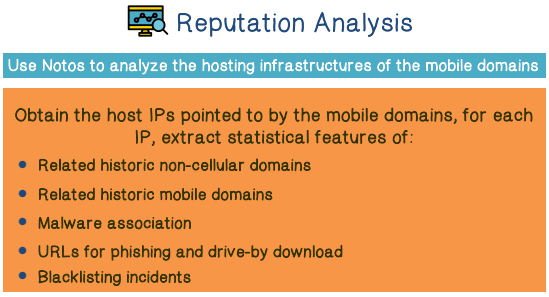
Key Data and Findings

- Three months of data from a major US cellular provider and a major US non-cellular ISP
- Known mobile malware samples remain rare in US: only 6,585 out of 380,537,128 devices, or 0.002%
- iOS vs. Android and other devices: equally likely to connect to suspicious domains

In our study, we used data from a major US cellular provider and a major non-cellular ISP. We find that, at least in the US, the number of machines that are infected with no malware remains to be very small, and iOS and Android devices are equally likely to look up these suspicious domains.



In our study, we first identified the mobile devices, and then we attribute each DNS query to a device. We then analyze the repetition of the resource records associated with the DNS query.



We use the Notos features to analyze the hosting infrastructures of mobile domains and compute repetition scores. The difference is that we use different features. In particular, for mobile domain, we first find the IP address of the machine that hosts this domain. Then for each of these IP addresses, we extract the following features: the historic non-cellular domains, the related historic mobile domains, any evidence of malware association with these domains, as well as URLs for phishing and drive-by download associated with these domains, and evidence that these domains have been blacklisted before. With these features and examples of malicious and legitimate mobile domains, Notos can then use machine learning algorithm to learn a scoring function. This scoring function can tell us how likely a mobile domain is malicious.

Tainted Hosts and Platforms

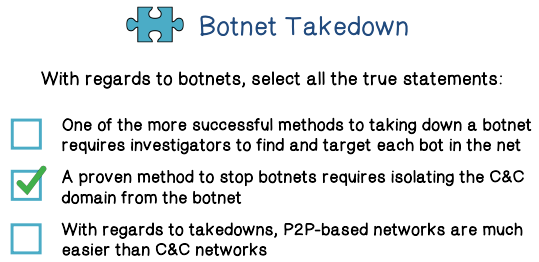
Device platform	% Total Requests by mobile device	% Population requesting tainted hosts	% Total tainted host requests
iOS	31.6%	8.8%	33.2%
All others (Android, etc.)	68.4%	8.2%	66.8%

Here are some results from our study. This column is the breakdown of total number of DNS requests for mobile devices. As you can see, there are lot more Android devices than iOS devices. When a machine on Internet has filed a petition, we call it as machine tainted. Some mobile domains are hosted on these tainted machines. This column shows the percentage of requests to this tainted hosts. As we can see, the iOS and Android devices are equally likely to connect to these tainted hosts.

Mobile Malware Prevalence

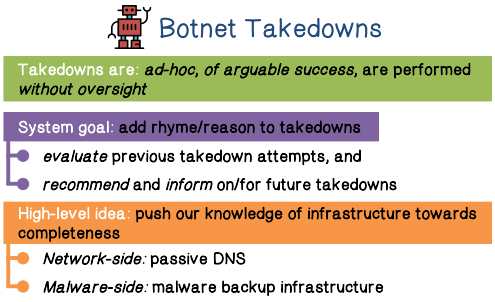
Malware Family	# Associated Domains	# Devices
DroidDreamLight	3	44
DroidKungFu	1	6
FakeDoc	1	2145
Fatakr	1	151
GGTrackers	3	1
NotCompatible	3	762
Planton	4	286
Malware β	1	1
WalkInWat	1	95
Gone60	1	1

We also measure the number of devices, their lookup domains that are associated with known malware families. As you can see, the number of devices we've known, malware infection is small.

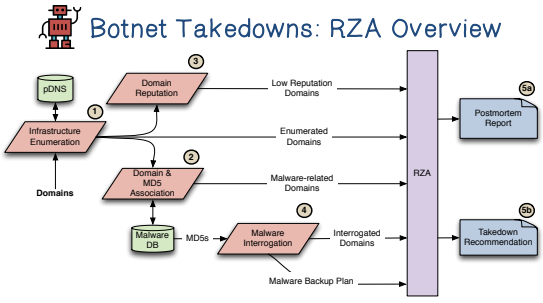


Now, let's discuss botnet takedown using DNS reputation. But, before we discuss the details, let's do a quiz. With regards to botnets, select all the true statements.

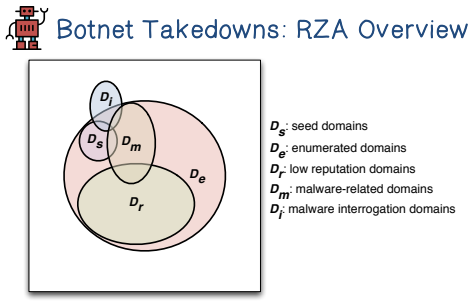
The first statement, one of the most successful methods to taking down a botnet requires investigators to find and target each bot in the net. This is false because typically, it is not possible to enumerate all paths in the net. The second statement, a proven method to stop botnets require isolating the C&C domain from the bot net. This is true. If you take down the command control infrastructure, the bot net will cease to function. The third statement, with regards to take downs, P2P-based networks are much easier than centralized C&C networks. This is false because P2P-based networks use distributed C&C and that's much harder to take down than centralized C&C.



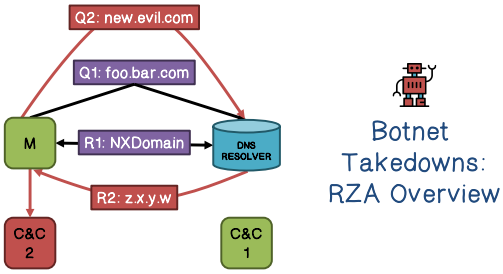
To motivate our study, we observed that currently Botnet Takedowns are ad hoc, and performed without oversight and sometimes are not successful. Our goal is to develop a system and framework to reason about Takedowns and recommend the best Takedown strategies. The goal of our study is to develop a system to reason about Takedowns, evaluate previous takedown attempts, and recommend future directions. The high-level idea is that we need to have as complete knowledge of the Botnet infrastructure as possible. We use both passive DNS analysis and malware analysis to expand our knowledge of the Botnet infrastructure.



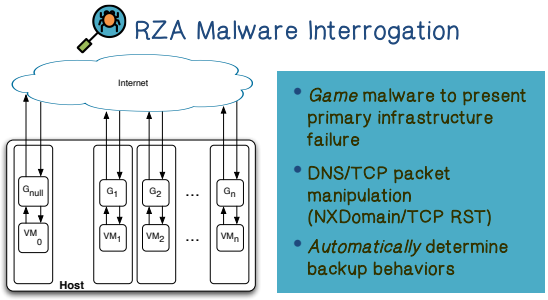
Here's an overview of our system called RZA. We start with a set of seed domains. These domains are known to be associated with Botnet infrastructure. They may use passive DNS data to analyze domains related to these seed domains including the reputation and malware samples associated with these domains. For the malware samples, we perform malware analysis to find out more domains, and that is how our system now has a much more complete knowledge of the Botnet infrastructure. Based on this knowledge, we can perform analysis of previous takedowns, and also recommend takedown strategies of a current Botnet.



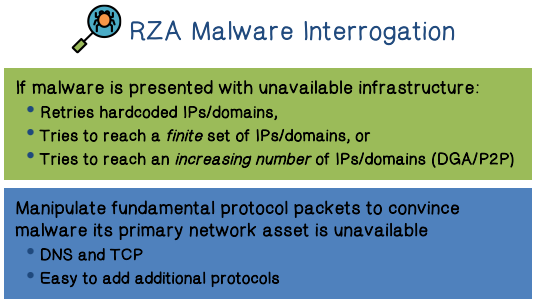
To summarize, we start with a set of seed domains they are known to be associated with Botnet infrastructure, then we enumerate all related domains. Some of these domains have low reputation, and some are associated with malware. For the malware samples, we further interrogate to find out more domains.



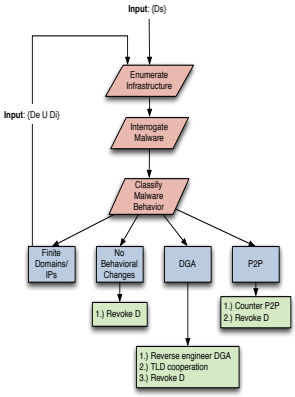
Here's a high level idea of how we interrogate a malware. Suppose this is domain that are malware would use by default. So, the malware is going to perform a lookup of this domain, and then connect with CNC server. To force the malware to tell us more domains that you may use, we can intercept the DNS query and respond with no such domain. Then, if there's a backup domain, the malware is going to look it up. Therefore, by analyzing the malware DNS traffic we can find out more domains that a malware might use.



Here is our infrastructure to interrogate malware. You run the malware image machine and have a gateway between the malware and Internet. Therefore, we can control how the malware can connect to its infrastructure. For example, we can play with DNS, no such domain, or TCP reset to force the malware to exhibit backup behaviors. Such a backup behaviors can include using additional domains or switching to a Peer-2-Peer.



We observe that a malware typically would use hard-coded domains or when these domains are becoming not available, it may we try again and then it may try the next set of domains and IPs as a backup. When those are not available, they may switch to Peer-2-Peer, or the so-called, domain generation algorithm, to use randomly generated domains. The intuition behind our approach is that malware will use an increasing number of domains and IPs when the infrastructure becomes now available. The easiest manipulation is by DNS and TCP. Of course, we can easily add more protocols.



RZA Takedown

Now, let's discuss an ideal takedown procedure. Again, we start with a set of seed domains. These domains are known to be associated with partner infrastructure. We enumerate the infrastructure using passive DNS. We get the malware associated with the domains and interrogate a malware. If the malware tells us additional domains, we look back. On the other hand, if there is no addition no domains, then we take the set of domains that we know so far and revoke all of them. That's how we take down the botnet. If you find that the malware is using dynamically-generated domains, then we need to first reverse engineer the algorithm, then we need to tell the TLD operators and DNS registrars to stop hosting or adjusting these new domains. We will also revoke the set of non CNC domains. If the malware exhibits behavior of using peer-to-peer network, then we need to take down the peer-to-peer protocol. For example, by penetrating into the peers. We also need to revoke the set of non CNC domains.

RZA Studies

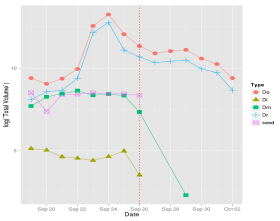
Postmortem study: analysis of *Kelihos*, *ZeuS*, and *3322.org/Nitol* takedowns

- Use lookup volume to show activity to infrastructure

Takedown study: analysis of 45 active botnet C&Cs

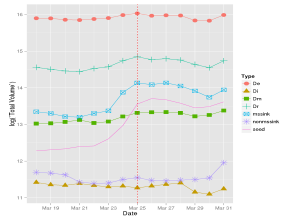
- Can we take them down?

Now, let's discuss some case studies. We perform analysis of previous takedowns. We also analyze how we can take down some of the active botnets.



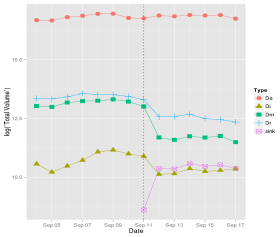
RZA Studies: Postmortem Kelihos

Here we showed a log scale of lookup volumes to different set of domains. This is the date of takedown. As we can see, lookups to the seed domains, the intogrey domains, and malware domains stop immediately at takedown or soon after. This means that, this takedown is very effective.



RZA Studies: Postmortem Zues

In the case of Zeus, although there are multiple groups including Microsoft and several research groups taking place in takedowns, their efforts were not coordinated. More importantly, none of them had complete knowledge of the botnet infrastructure. Therefore, even after the takedown effort, the botnet continues to connect to its infrastructure.



RZA Studies: Postmortem 3322.org/Nitol

This takedown was accomplished by transferring the entire 3322.org name server authority to Microsoft and domains the malicious result to a set of long single IP addresses. That is, domains were sunk on a day of takedown and were limited to the 3322.org domain names. Unfortunately, this only accounted for a fraction of domains associated with malware. Therefore, lookups to this malware associated domains remain frequent. In other words, this is not an effective takedown.

RZA Takedown Study

Of the 45 botnets:
- 2 had DGA-based backup mechanism
- 1 had P2P-based backup mechanism
- 42 susceptible to DNS-only takedown

Current drawbacks to takedown
- Ad-hoc
- Little oversight
- Arguable success
- All point to need for central authority

Of the 45 active Botnet, we find that too had TGA based backup mechanisms, and one had a peer-to-peer based backup mechanism. On the other hand, 42 of them are subject to DNS only Takedown. To summarize the drawbacks of current Takedown efforts, there ad hoc meaning that they don't have complete knowledge of the Botnet infrastructure. There's a valid oversight, meaning that different groups may step on each other, and the success rate is not high. Therefore, we believe that a central authority to coordinate Takedown efforts is necessary.

RZA Takedown Study

ICANN's UDRP/URS as example frameworks
- Criteria for takedown
- More eyes = more successes
- Test with new TLDs (much like w/ URS)

Icons policies can provide an example framework. For example, there is policy for domain name dispute resolution. There's also system for rapid suspension. Therefore, we should study policy criteria for Takedowns, and have the community to review such criterias. We can test the policy with the newly created top-level domains.

The Core of the Matter: Analyzing Malicious Traffic in Cellular Carriers

Much of the attention surrounding mobile malware has focused on the in-depth analysis of malicious applications. While bringing the community valuable information about the methods used and data targeted by malware writers, such work has not yet been able to quantify the prevalence with which mobile devices are actually infected. In this paper, we present the first such attempt through a study of the hosting infrastructure used by mobile applications. Using DNS traffic collected over the course of three months from a major US cellular provider as well as a major US non-cellular Internet service provider, we identify the DNS do- mains looked up by mobile applications, and analyze information related to the Internet hosts pointed to by these domains. We make several important observations. The mobile malware found by the research community thus far appears in a minuscule number of devices in the network: 3,492 out of over 380 million (less than 0.0009%) observed during the course of our analysis. This result lends credence to the argument that, while not perfect, mobile application markets are currently providing adequate security for the majority of mobile device users. Second, we find that users of iOS devices are virtually identically as likely to communicate with known low reputation domains as the owners of other mobile platforms, calling into question the conventional wisdom of one platform demonstrably providing greater security than another. Finally, we observe two malware campaigns from the upper levels of the DNS hierarchy and analyze the lifetimes and network properties of these threats. We also note that one of these campaigns ceases to operate long before the malware associated with it is discovered suggesting that network-based countermeasures may be useful in the identification and mitigation of future threats.

In this paper, we presented a study of traffic obtained from a major US cellular provider as well as a major US non-cellular Internet service provider. Our work provides an in-depth understanding of the Internet infrastructure used for mobile malware. In particular, we showed that the net- work infrastructure used by mobile applications is part of the core Internet infrastructure used by applications in the non-cellular world; in other words, the mobile web is part of the Internet. We presented evidence showing that the mobile malware discovered by the research community ap- pears in a minuscule number of devices in the network; this suggests that mobile application markets are already providing adequate security for a majority of mobile devices. We compared traffic to suspicious hosts between different mobile device platforms and demonstrated that iOS devices are no less likely than other platforms to reach out to such devices. Finally, we analyzed two major mobile threats and found that their network characteristics are similar to those of non-cellular botnets. Overall, these findings suggest that there are commonalities, in terms of both network infrastructure and characteristics, between malicious mobile applications and non-cellular malware. Therefore, we should leverage our successful experiences with DNS monitoring and reputation systems for non-cellular ISPs to develop a similar system for cellular carriers to

identify (emerging) mobile threats. We leave this as a future work.

Building a Dynamic Reputation System for DNS (NOTOS)

This paper introduces a new system – termed Notos – aimed at addressing the problem of dynamically assigning reputation scores to DNS systems. Normally, malicious domains are placed on blacklists after intensive investigation and thus filtered against by IDS. To evade such security attackers use tactics like using new domains daily to evade static blacklists and firewalls. The premise of the Notos system is that malicious, agile use of DNS has unique characteristics and can be distinguished from legitimate, professionally provisioned DNS services. When evaluated in a large ISP's network with DNS traffic from 1.4 million users, Notos showed the ability to identify malicious domains weeks or even months in advance of them being blacklisted with high accuracy (true positive rate of 96.8%) and low false positive rate (0.38%).

The primary source of data for Notos is a passive DNS Database which is kept up to date with historical DNS data aggregated from diverse geographical locations. In addition to the pDNS Database is a list of known bad domains typically contacted by malware at one point or the other. Finally, a pool of known good domains is aggregated from alexa, all combining into the knowledge base upon which Notos is founded. For each domain being considered, the researches extract a set of statistical features which can be broadly classified into three: network-based, zone-based, and evidence-based features.

Network-based features themselves are 18 in number and broken down into BGP features, AS features and Reputation features and all serve to characterize the specific agility of a given domain; domain names and IPs that are used for malicious purposes are often short-lived and are characterized by a high churn rate. This agility avoids simple blacklisting or removals by law enforcement. Zone-based features are 17 in number and divide down into two broad categories of TLD features and String features. Together, these two sub-groups aim to characterize the diversity of the domains in the RHDNs set around a given domain because while legitimate Internet services may be associated with many different domain names, these domain names usually have strong similarities. For example, google.com, googlesyndication.com, googlewave.com, etc., are all related to Internet services provided by Google, and contain the string "google" in their name. On the other hand, malicious domain names related to the same spam campaign, for example, often look randomly generated and share few common characteristics. Finally, Evidence-based features, of which there are six measured, are split into two groups of Blacklist features and Honeypot features with both serving as a way to measure the extent to which a given domain is tainted by associated with other known bad domains.

Once extracted, these features can then be fed into a model to be trained offline, before being used in online mode to assign reputations to new domains. With data sets collected from DNS traffic from two ISP-based sensors, one located on the US east coast (Atlanta) and one located on the US west coast (San Jose), in addition to the aggregated DNS traffic from the different networks covered by the SIE, the Notos system was tried and tested on a total of 27,377,461 unique resolutions from all these sources over a period of 68 days, from 19th of July 2009 to 24th September 2009. The final result, as mentioned previously, was Notos is highly accurate in identifying new malicious domains in the monitored DNS query traffic, with a true positive rate of 96.8% and false positive rate of 0.38%. In addition, Notos is capable of identifying these malicious domain weeks or even months before they appear in public blacklists, thus enabling proactive security countermeasures against cyber attacks.