The Gh0st in the Shell: Network Security in the Himalayas

Matthias Vallentin
vallentin@icir.org

Jon Whiteaker
jbw@berkeley.edu

Yahel Ben-David
yahel@airjaldi.net

Abstract

The town of Dharamsala in the Himalayas of India harbors not only the Tibetan government in-exile, but also a very unique Internet community operated by AirJaldi. The combination of high-profile clientele and naive users makes for a very interesting setting from a network security standpoint. Using packet capture and network intrusion detection systems (NIDS), we analyze the security of the network. Given the sensitive history between China and Tibet, and the general public's penchant to support the freedom of Tibet, it would not be surprising for the Chinese government to be interested in the activities of the community in-exile. Therefore, we also look for evidence of malware targeted at this unique user-base. In our work, we find significant amounts of malicious activity in the traffic, including a solid link to a previously discovered high-profile spy network operated in China.

1 Introduction

The town of Dharamsala, in the rural Indian state of Himachal-Pradesh, has become the headquarters of the Tibetan Community-in-Exile and the home for its spiritual leader H.H. the Dalai Lama. Since the Dalai Lama fled Chinese-occupied Tibet in 1959, this little Himalayan town grew to host a large number of pro-Tibetan NGOs and many related non-profit organizations supporting the community and its struggle to regain its land and freedom.

In recent years, the Tibetan community has learned to harness the Internet as its key communications medium, which is effectively connecting them with the rest of the world. Enabling affordable Internet access to this mountainous and rural area was no simple challenge — the AirJaldi wireless network [1] which spans over a radius of 80km in and around Dharamsala plays a key role in overcoming these constraints and has quickly grown to connect more than 10,000 users to the Internet.

The intense political tension between the Community-in-exile and the Chinese government sets the backdrop for our quest — the Chinese view the Dalai Lama as a serious threat to their regime, while the international empathy towards the Tibetan struggle is likely top on the list of China's concerns. Juicy spy stories and intrigues are the predominant subject of the day-to-day gossip in Dharamsala, occasionally fueled by indications of early knowledge the Chinese had regarding Tibetan activities, further indicating some unwanted flow of information from Dharamsala to Beijing must exist. While surely there are non-electronic and non-computerized forms of information flow, anecdotal evidence and disorganized reports about specific incidents, do provide strong indications that the Chinese are harnessing the Internet and the growing usage of computers in Dharamsala as a valuable vehicle for their intelligence gathering.

The contributions of this paper work are twofold. On the one hand, we perform a traffic analysis over two months of the AirJaldi network in Dharamsala, serving the Tibetan community in-exile, and the associated server-farm in San Jose, CA. On the other hand, we were eager to verify the speculations regarding targeted attacks in Dharamsala. In particular, we set out to confirm the existence of Ghost-

Net [16, 15], identify non-mainstream malware that performs activities of intrigue, and develop a picture of the threat landscape in the AirJaldi network.

The remainder of this paper is structured as follows. We begin with summarizing related work in §2. After explaining our methodology and infrastructure in §3, we present our findings in §4. We turn then in §5 to the limitations of our study and give promising directions for future work in §6. Finally we conclude in §7.

2 Related Work

Since the late 1990's, politically motivated cyber-attacks have been observed in the wild, usually involving defacement of websites with messages, as opposed to debilitating attacks [12]. However, the cyberattacks against Estonia in 2007 were clearly meant to impose harm. Thousands of machines flooded important websites and services of Estonia, essentially crippling its network [12].

Although it is not known if any governments perpetrated any of the attacks, it is suspected that the attacks originated from individuals involved in the issue. Recently, however, a group linked to the Russian government, Nashe, claimed responsibility for the attacks [8]. This link to the Kremlin, indirect as it may be, breaks new ground for government supported cyberattacks.

Since the attacks against Estonia, politically motivated cyber-attacks occurred in Georgia [4, 14] in 2008. These attacks drew a lot of public attention, as it was coupled with actual military action, sparking further suspicion of government involvement.

Active traffic intervention is not uncommon today. The "Great Firewall of China" strictly censors Internet content deemed as inappropriate by injecting forged TCP reset packets into the traffic to shutdown the undesired connection [6]. As an evasion strategy, Clayton et al. suggest to ignore RST packets at both endpoints [2] to prevent the connection teardown. Not only the Chinese government employs this technique, but also network intrusion detection systems (NIDS) make use of it to terminate malicious connections [17, 22].

Weaver et al. develop a reliable detector for RST injection and confirm that ISPs also employ this technique to manage P2P traffic, thwart spam, and counter virus spreading [26]. The authors further fingerprint different types of injectors and show that anomalous artifacts, such as non-RFC compliant TCP implementations, pose an inherent limitation in the detection process. The conducted measurements also include connections terminated by the Great Firewall of China.

The People's Liberation Army (PLA) of China is believed to have been practicing "Information Warfare" (IW) as early as the 1950's [33]. Initially IW consisted of gathering information to increase the potency of psychological warfare attacks. However, with the rise of the Internet age, the PLA is believed to have expanded IW to the Internet as well [33, 13]. A recent US DoD report states not only does the PLA have defensive measure in place to protect against Internet-related threats, but that they are actively developing malware for use on their enemies [33].

Indeed, the vast majority of malware observed today appears in China [23, 18, 19]. A recent analysis of web-based attacks finds that the primary goal of malware that compromises web-servers is to infect its visitors in order to exfiltrate Personally Identifiable Information (PII) and online game account information by leveraging Internet Explorer 7 0-day exploits [21].

In addition to the prevalence of malware in China, there is significant actual and empirical evidence of targeted attacks against pro-Tibetan organizations originating from computers in China [27, 9]. The attacks are well coordinated, suggesting the people behind them may be more than just individuals with a vendetta, but rather an organized group with access to significant resources for planning and preparation [25].

Establishing that the Tibetan community in particular is being targeted for malware is no easy task, even when using other lower profile networks as ground truth. Past studies have shown that attack traffic is not homogeneous from location to location [32, 31, 3], and the unique setup of the network in Dharamsala will likely only accentuate these observations.

Besides the targeted attacks against specific organizations and countries, security analysts have also recently observed malware directed at *single* individuals [28, 11].

2.1 GhostNet

The closest work to ours was released during the middle of our investigations. In March, two related reports were released, one from the InfoWar Monitor [15], and the other from Cambridge University [16]. The reports collectively uncovered a network of infected machines reporting back to machines in China, dubbed "GhostNet", named after one of the offending pieces of malware — Gh0stRat [16, 15].

The network consisted of a number of high profile machines inside embassies and government offices of countries around the globe [15]. In particular the report from Cambridge investigated evidence from the private office of the Dalai Lama being compromised [16].

As it turns out, GhostNet came up in our own investigations as well, and we discuss what we found further in section §4.2.2.

3 Methodology

We begin with a high-level analysis of traffic patterns to distill characteristic patterns of security-related incidents. By augmenting the connection records with geographic information, we obtain per-country breakdowns of activity which is particularly helpful to separate distinct events. As complementary low-level angle, we use signature-based detection to identify known malware, which is otherwise difficult to pin-point in the aggregated traffic analysis. In combination, these approaches constitute a powerful means to find a needle in the haystack.

After introducing the two environments and sketching our monitoring infrastructure in §3.1, we turn to the details of our trace files in §3.2.

3.1 Network Topology

During our study, we analyzed two networks operated by AirJaldi that complement each other: a server farm in San Jose, California, and the community network in Dharamsala, India. There exists a mutual relation between these two networks, as the machines in San Jose provide services for users in Dharamsala. However, the topology of the sites is quite different.

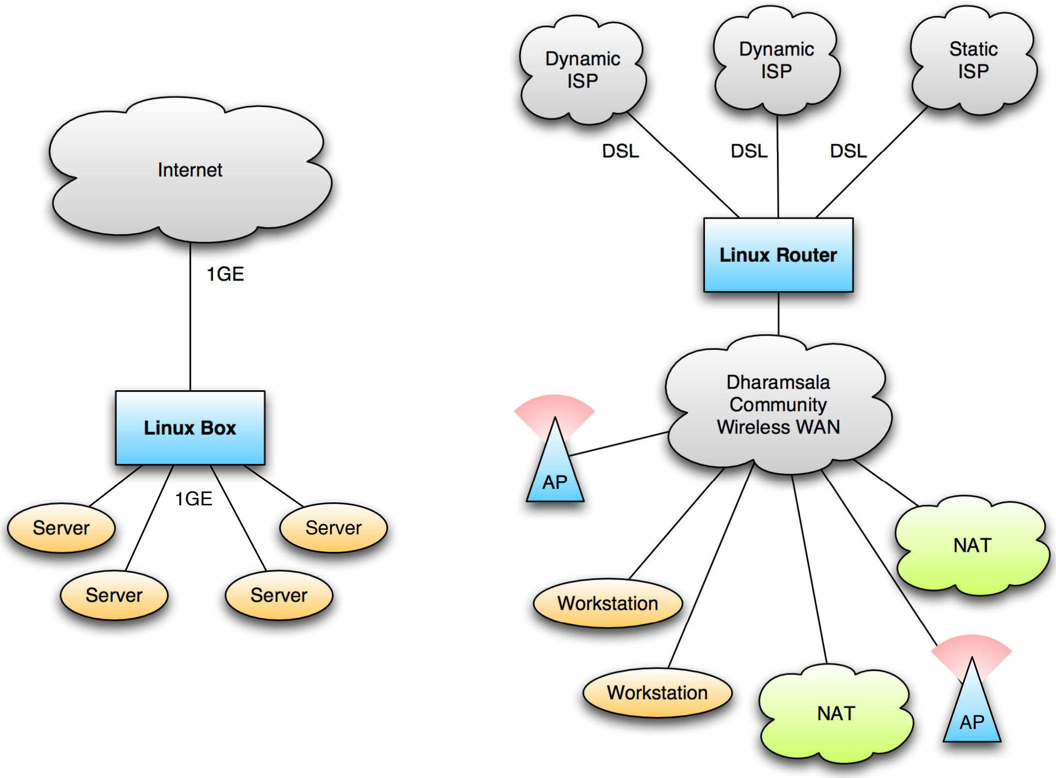
As shown in Figure 1a, servers in San Jose have Gigabit connectivity to the Internet and we introduced a new Linux-based bridge in the traffic-path for our monitoring and analysis. The vast majority of machines are Linux boxes (e.g., web, VPN and VoIP servers) and are carefully maintained by the AirJaldi operators. We conducted our experiments on an AMD Opteron with two 2.6 GHz cores. The operating system runs a Linux 2.6.18 SMP kernel on Cent OS 5.2.

Figure 1b illustrates the network in Dharamsala, which exhibits a higher degree of heterogeneity. Internet connectivity is enabled through load-sharing of multiple connections to multiple ISPs, namely four ADSL lines to BSNL and two leased-lines to Relience and AirTel. Some of the uplink connections (such as the ADSL lines) use dynamic IP addresses which change over time, while others offer a block of static IPs. The Linux router load-balances outgoing flows over the various uplinks based on load and a pre-defined routing policy. All IP addresses in Dharamsala are private and are translated by the router. While the network was initially designed without NAT devices and allowed complete bi-directional connectivity among all peers within the network, it experienced uncontrolled growth. Local operators tend to overlook the complex routing and addressing issues that support the above design goal, yielding large isolated islands behind NAT devices that further complicate our ability to map local IP addresses to a single host at the Linux router.

The Linux router in Dharamsala, dubbed the Bandwidth Maximizer (BWM), runs on a dual-core 3Ghz server, with 4Gb of RAM and two Gigabit Ethernet interfaces. Using a VLAN supported switch, we provide the virtual port-density to the BwM for the multiple upstream connections that are some PPPoE, Ethernet, and wireless LAN. The router runs on CentOS 5.0 with a 2.6.18 SMP Linux kernel.

To monitor the network traffic, we employ two popular open-source NIDS available today: Bro [17] and Snort [20].[1]. While we can use a

[1] We use the most recent development version of Bro from



(a) The San Jose server farm.

(b) The network in Dharamsala.

Figure 1: Network topology of San Jose and Dharamsala.

dedicated Linux bridge in San Jose (see Figure 1a), Dharamsala has less infrastructure in place and we had to install the NIDS directly on the BWM. All our analyses were conducted offline on `pcap` trace files that we characterize below.

3.2 Datasets

The packet trace in San Jose was recorded over 47 days, from February 28 to April 15. It contains 12.4 million connections and the top 6 services in terms of number of connections are DNS (65.3%), HTTPS (14.8%), HTTP (5.6%), SMTP (5.3%), IDENT (2.2%), and SIP (1.3%). 84.8% of the connections were established and shutdown successfully, 8.5% of the connections were comprised of an unanswered SYN packet, and 1.0% of the connections were rejected.

The packet trace of Dharamsala was recorded over 59 days, from March 1 to April 28, containing 57.0 million connections. For the largest share of the connections (35.4%), Bro could not determine the application protocol. The top 6 services in terms of number of connections are HTTP (31.4%), Windows RPC (11.4%), DNS (7.1%), ICMP echo (6.6%), SMTP (2.0%), and HTTPS (1.8%). We only saw a SYN packet for 23.0% of the connections, 4.3% were rejected and 11.2% reset by the connection originator. In contrast to San Jose, a much smaller percentage of connections (43.3%) were established and shutdown successfully.

4 Results

This section presents the results of our security examination of the AirJaldi network. After discussing our findings in San Jose (§4.1), we present our results for the network in Dharamsala (§4.2).

the Subversion repository and Snort in version 2.8.3.2 (Build 22) with subscription signatures from May 20, 2009.

4.1 San Jose

The focus of our analysis in San Jose is on inbound traffic because outgoing traffic can only come from operators and a limited set of known services. Figure 2 shows failed inbound and total inbound traffic during our observation period. In the following, we investigate the two remarkable spikes in both figures that occurred from April 6 17:00 (UTC) to April 8 17:00. When mentioning a spike in the text below, we refer to the connections during this time interval.

4.1.1 Failed Inbound Traffic

Figure 2a displays failed inbound connection attempts. These are connections that were either rejected or for which we only saw a SYN packet. We continue to use this terminology throughout the paper. There is a constant noise of failed connections from China and the USA. Note that the number of failed connections per day are a magnitude lower than the total inbound connections in Figure 2b, which also contains a spike around the same time. The majority of failed inbound connections originate from Taiwan and China during the spike. 93% (23,164) of all connections originate from TCP port 6005 and stem from a single scanning IP address in Taiwan (202.39.49.10). The targets of this scan are AirJaldi machines in the address range from 72.13.87.164 to 72.13.87.189. Each machine is contacted 927 times on average (sd = 29.13).

The spike from China in the same Figure represents scanning activity as well: 64% (11,598) of all inbound connections failed. 30% (3,154) of these failed attempts also contained the source port 6005 and originate from the IP 222.141.223.190, which belongs to a dynamic DSL connection in Beijing, China. The scan covered the AirJaldi network ranges 72.13.87.162 to 72.13.87.172 and 72.13.87.177 to 72.13.87.189. Unlike the scanner from Taiwan, the addresses from .173 to .176 were excluded. As these address ranges are not associated with Tibetan content hosted in San Jose, we do not believe that the scans constitute a targeted attack.

Another 28% (2,973) of failed Chinese inbound connections originate from TCP port 6000, but from 83 different addresses. Among these scan sources, a reverse DNS lookup succeeded for 10 IPs. One particular IP (132.201.18.119) resolved to www.zhaoyangbook.cn which appears to be an online shop for books and magazines. We suspect the site is infected with malware scanning the AirJaldi network.

Finally, 14% (1,533) of failed connection attempts from China have TCP source port 12,200 and come from 5 different IPs with no reverse DNS entries. The remaining scans are scattered across different high-level source ports and do not have an salient characteristic.

4.1.2 DNS Amplification Attacks

Our trace in San Jose contains 55,335 connections on port 445, which is a port used for the Server Message Block (SMB) file-sharing protocol on Windows machines. Since the majority of machines in San Jose run Linux, we were curious and investigated them further. 99.37% of are failed inbound connections and all 14 outbound connections were unsuccessful as well. [2]

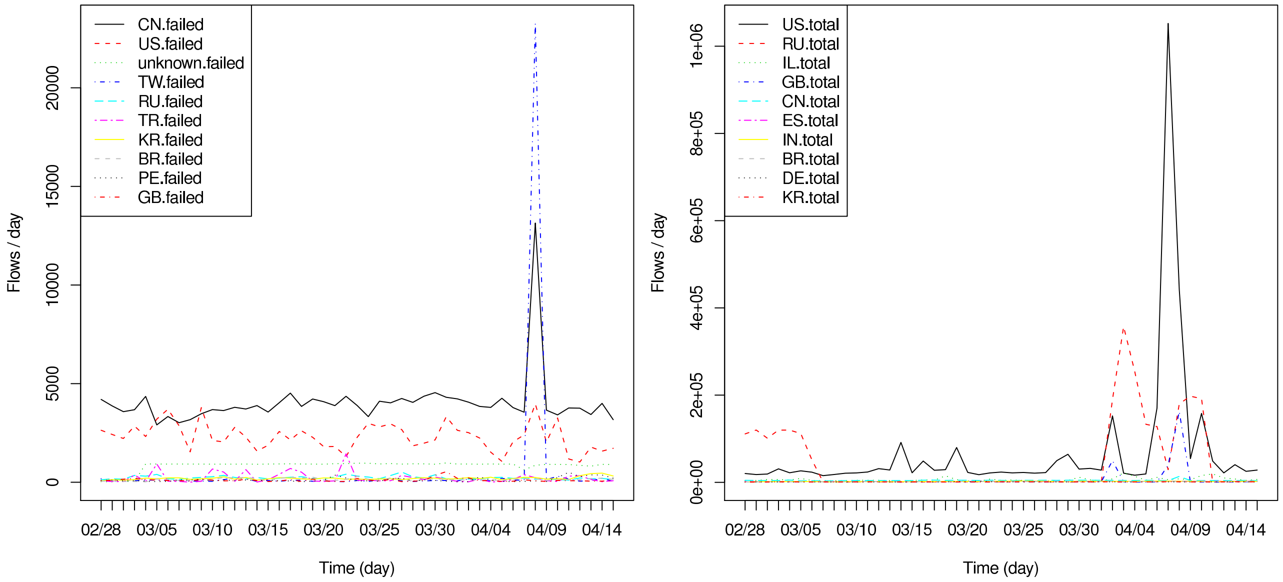
The remaining interesting inbound traffic consists of 66 connections from port 445 to dns1.airjaldi.net and dns2.airjaldi.net. These are not SMB connections, but rather spoofed DNS requests with 5 of the 13 IPs resolving to hosts in Russia:

```
162.223.218.207 (ns2.theplanet.com)
198.230.193.212 (kaztoday.nichost.ru)
17.224.189.213 (respublika-kz.info)
89.4.109.62 (invest-pool.ru)
24.51.20.72 (gm-gen.ru)
```

Upon closer examination, we found out that these hosts are asking the AirJaldi name servers to resolve NS . to return the list of root name servers. This very short request entails a long reply and is a known technique to use name servers as amplifiers in DoS attacks. The AirJaldi network was not the only network experiencing this attack [29].

We also have now an explanation for the prominent spike in Figure 2b that comprises 2,043,052

[2] Upon closer examination, we discovered that all outbound connections on port 445 constitute unsuccessful attempts to connect a VPN network or represent manual scans initiated by the network operators.



(a) Failed inbound traffic.

(b) Total inbound traffic.

Figure 2: Total and failed inbound traffic per country in San Jose.

connections during the time of the spike, which alone accounts for 28% of the all inbound connections in our trace. 88% of connections in this spike are DNS connections. Looking beyond just the spike, we observe that 98.8% (2,409,498) of the total connections from Russia, and 90.2% (275,986) total connections from Great Britain constitute spoofed DNS queries. 19.4% of all DNS replies observed returned a list of root name servers.

We believe that the majority of these queries is malicious. To prevent further exploitation of this vector that causes participation in DoS attacks, we recommend to reconfigure the AirJaldi name servers. This issue can be mitigated by ignoring recursive DNS queries from addresses for which the name servers are not authoritative.

4.2 Dharamsala

In Dharamsala, we observe a much higher traffic volume. The distribution was what we expected for the network given its size and location. Looking at the breakdown of all traffic is not necessarily insightful, as the majority of the traffic is web traffic, so we filtered out low-level ports. Low-level ports are far from immune to malware, but the majority of the traffic on these ports is harmless web surfing. We visualized the results on a map of the world in figure Figure 3. The center of each circle represents a city and the radius scales logarithmically with the number of connections with a destination IP in that city. The circles are semi-transparent so overlapping circles can be seen in a more opaque red.

Some of the results are quite striking. The US and India represent strong centers of activity, as they did with the lower-level port traffic. There are two notable takeaways from this map. First, the high-level port traffic to Moscow is unusually large. Second there is proportionately more high-level port traffic to China than low-level port traffic. These two results are particularly interesting, as we assume low-level port traffic like HTTP comprises the majority of the traffic.

4.2.1 Traffic Analysis

We discovered several suspicious issues during our traffic analysis. First, when we compare the number of connections per port rather than by service identified by Bro,[3] we find that 38.1% of the connections are on port 80. However, the identified HTTP con-

[3] Bro does not rely solely on ports to determine an application protocol, but rather uses *dynamic protocol detection* to reliably identify the protocol in use [5].

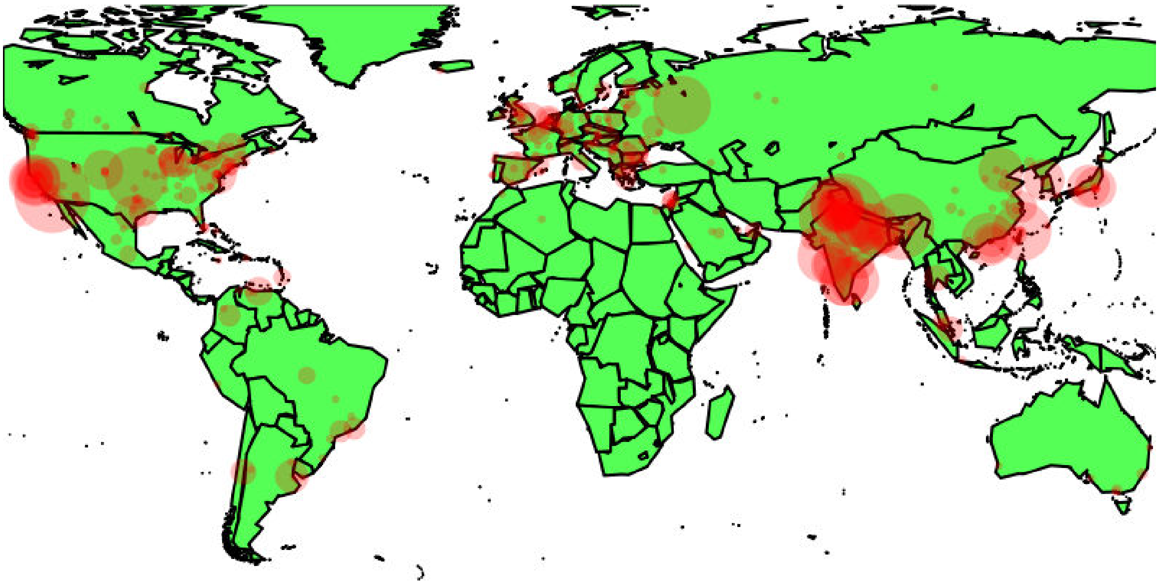


Figure 3: Geographic destinations of high-level port traffic in Dharamsala. The center of each circle represents a city and its radius increases logarithmically with the number of connections to that city.

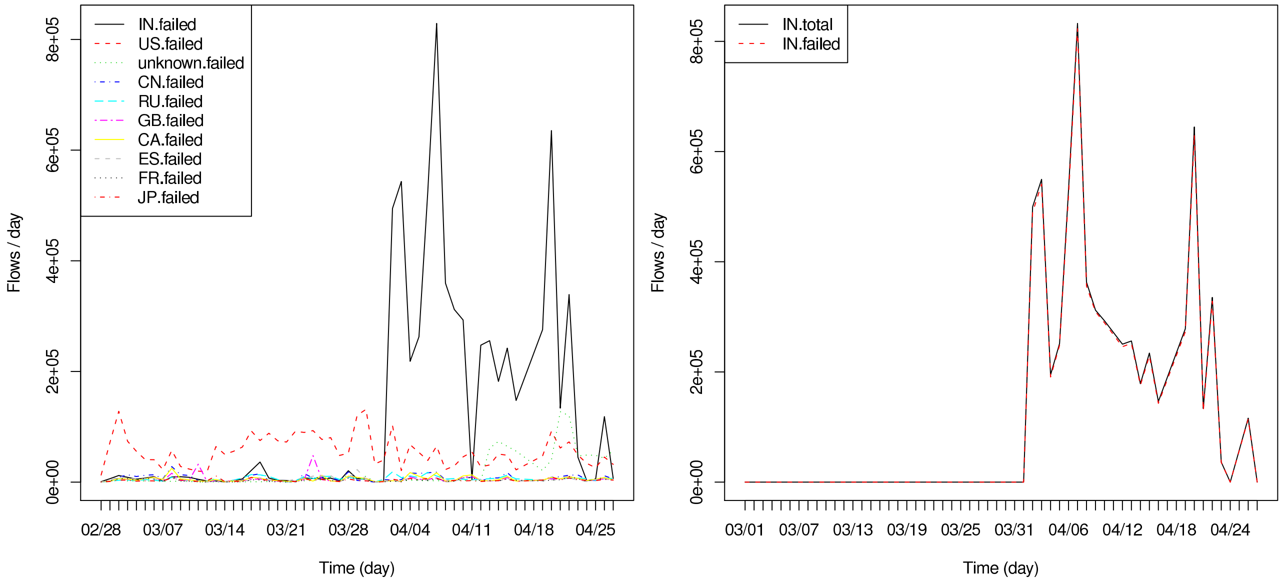
nections account only for 31.4%. Even when adding ports 443 (1.8%), 8000 (0.5%), and 8080 (0.2%), we have a remaining difference of 4.2% of port 80 traffic that is potentially not HTTP.[4] All other ports account for less than 0.2%. Consequently, this observation suggests that roughly 520,800 connections used port 80 for non-HTTP connections. Given that malware often tries to conceal its communication by using high-volume ports, like port 80, it is an indicator that these non-HTTP connections are perhaps not benign.

Furthermore, we examined failed outbound traffic which is illustrated in Figure 4a. To our surprise, a significant share of all connections are Windows RPC connections. Looking closer, we see that all outbound traffic on port 135 failed and went *only* to India, as shown in Figure 4b. The remarkable spike in both Figure 4a and Figure 4b at April 7 represents 912,000 failed connections destined to port 135, which is roughly half of the connection volume the entire network faced that day. Three internal addresses generated the traffic: `172.28.1.152` (1,989), `192.168.11.2` (34,606), and `10.2.5.102` (872,546).

Turning to Figure 5a which displays the top 10 flow contributors in number of flows per day, we see a enormous spike at the beginning that represents traffic to New York, USA. Coincidentally, we further observed that 6.1% of the total traffic went to a single IP address in the US: `64.34.164.84` with a reverse DNS entry of `onair2.billydonair.com` that has no A record. We plot the activity of this address in Figure 5b. Comparing the two Figures, we clearly see that the big spike relates to this address. In fact a total of 3,466,786 connection attempts on port 80 were made to this single IP. Our attempts contact this machine to check for the existence of a HTTP web server were unsuccessful.

Digging further, we found out that this address appeared in the context of the malware `Trojan-Spy.Agent.ENP` according to ThreatExpert [24], an automated threat analysis system which encountered this sample at the beginning of April. The report mentions that this piece of malware installs a keystroke logger, contains its own SMTP engine to presumably send spam or spread, opens local TCP ports 1033 – 1035, and tries to contact `64.34.164.84` on TCP port 2211. Indeed, examining port 2211 separately, we see both outbound (Figure 6a) and (Figure 6b) inbound activity. At the same time, port 2211 is used by the Na-

[4] Among the top 20 connections by port number, there were no other clear port numbers that suggest obvious HTTP usage.



(a) Failed outbound traffic.

(b) Outbound traffic for port 135 in Dharamsala.

Figure 4: Outbound traffic characteristics in Dharamsala.

tional Weather Service and MikroTik Secure management for "The Dude" [30]. Furthermore, there is an irregular ratio between successful and failed connections and the general traffic patterns in Figure 6a and Figure 6b do not correspond with Figure 5a. A more detailed analysis on the full packet trace could have provided more insight, but was unfortunately not possible due to a disk failure of our drive with the full trace files.

4.2.2 GhostNet

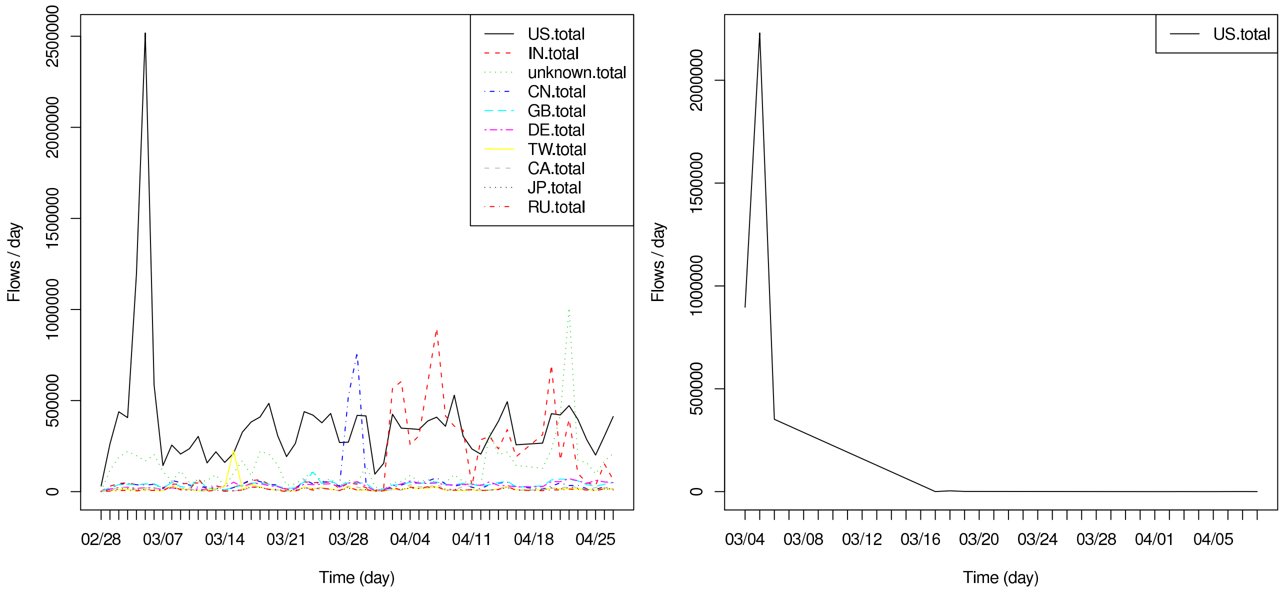
Shortly after the reports exposing GhostNet were released, the IPs associated with the network became inactive. Fortunately, our monitoring infrastructure was already established, so traffic was being recorded in Dharamsala prior to the reports. This meant that we could check if our network contained any instances of GhostNet.

We identified all of the IPs associated with GhostNet in the two reports, and searched for activity involving said IPs. Indeed we found traffic to two IPs mentioned in the report: `61.188.87.58` and `210.51.7.155`. However, traffic to these IPs is not necessarily indicative of a GhostNet infection, particularly if the IPs were on a shared webserver. Thus, to verify the activity as GhostNet, we isolated the traffic to these IPs for a closer look.

Investigating the traffic with Wireshark, the traffic to the IPs consists largely of HTTP GET and POST requests, particularly involving a script named `Owpq4.cgi`. This file and behavior was specifically mentioned in one of the GhostNet reports [15], increasing our suspicion of GhostNet activity.

In addition to this traffic, we saw two binaries being transfered on the wire multiple times to the infected machines – `timesvc.dll` and `ActiveX_dx9.14_plugin.icx`. Reconstructing these binaries and taking a closer look at them would have been ideal, but unfortunately due to an error in our packet capture configuration the packets were cut off.

All in all, five unique IPs within the Dharamsala network communicated to these two hosts. It does appear that there was a single host behind each IP in the communications, however, we are unable to identify the individual machines for several reasons. First, each of the internal IPs we see actually represent a whole separate NATs, some of which serve entire villages. This could be remedied by monitoring traffic at each of the routers serving the NATs we see. Unfortunately, as mentioned earlier, the IPs became inactive and traffic to them has stopped since



(a) Top-10 traffic by country.

(b) Total traffic for `62.34.164.84`.

Figure 5: Traffic breakdown per country and a specific IP address in Dharamsala.

then.

It is important to note that a number of the IPs in the GhostNet reports were redacted. We could only verify traffic from the publicly available IPs in the reports. We tried to attain the redacted IPs, but we were not granted access. This means that there still could be active GhostNet activity that we cannot uncover due to these restrictions. But since we are still actively recording traffic, we should be able to identify any remaining GhostNet infections should the remaining IPs be released to the public.

4.2.3 Locksky

We also encountered other malware specimen during our study. One particular instance of malware we found is Locksky,[5] an email worm that spreads both via SMTP, HTTP, and IRC [7]. We detected Locksky with Bro's builtin IRC-based botnet detector. Below is an excerpt of the of C&C communication that uses the channel topic to assign spreading instructions to an infected machine.

```
Matching NICK [00|USA|XP|466993]
Matching TOPIC \
  !asc -S -s|
  !patch|
  !ip.wget -S s|
  !http http://bojifun.com/hlio|
  !asc s 20 3 0 -a -r s|
  !asc s 60 3 0 -b s|
  !asc s 40 3 0 -c s|
  !ip.wget http://bojifun.com/ep.exe \
     C:\msr32.exe 1 s
```

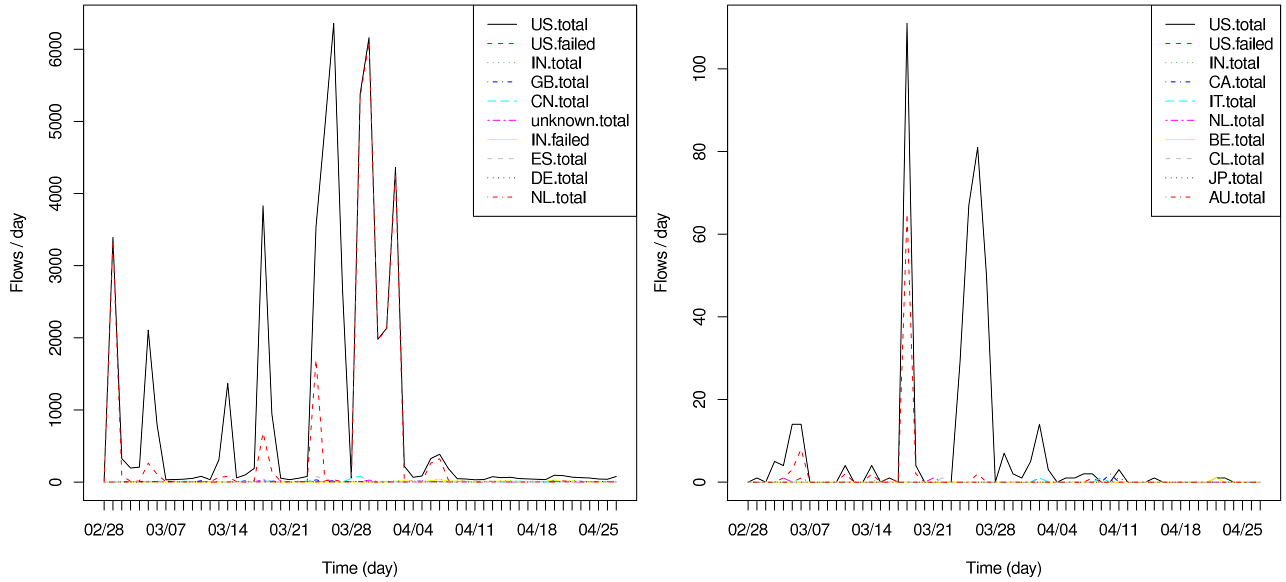
Although this bot ships with its own SMTP engine, we did not observe a significant amount of outbound connections. We stay in contact with network operators to clean up the infected machines.

5 Limitations

We acknowledge that our study has some limitations. Although we took great care to avoid measurement outages, the rural and harsh weather conditions regularly cause power loss in Dharamsala. These outages not only interrupt our packet capturing, but also disconnect the entire community WAN from the Internet.

The remote location and conditions also made for some headaches during our analysis. Due to the slow DSL uplink in Dharamsala and our massive packet traces (1+ TB), much of our analysis had

[5] Locksky, also known as Loosky, is often mentioned in the context of the Nucrypt botnet, which is estimated to consist of 20,000 compromised machines and sending 5 billion spam messages per day [10].



(a) Outbound traffic on TCP port 2211.

(b) Inbound traffic on TCP port 2211.

Figure 6: Outbound and inbound traffic for port 2211 in Dharamsala.

to be performed in Dharamsala. The only transfers of data from Dharamsala to Berkeley were of low-volume pre-processed logs from Bro and Snort. Furthermore, these transfers had to be performed at night in Dharamsala so as not to bog down the connection of the entire community during peak traffic hours. Also, the external USB hard drive that we used for storing our packet traces had a propensity for disk failures during our analysis, adding to our headaches.

More fundamentally, the scope of our analysis is restricted to what we see on the network. Host-based context would have been beneficial in many situations, in particular during our analysis of GhostNet. Had we been able to isolate infections on individual machines instead of just at the granularity of NATs, we would have both been able to help the network operators clean the network, as well as analyze the actual piece of offending malware.

The sheer volume of the data we collected also posed some limitations, particularly in regards to manual analysis. Despite dealing with the truncated output from Snort and Bro, the data was still unwieldy and hard to manually inspect. This forced us to hone in on anomalies in the traffic pattern, essentially performing manual analysis in slivers of the overall data - usually by traffic spikes, destination IP, or port. Unfortunately, this means we may have missed some of the interesting facets of the network if they did not stand out against the overall traffic.

One of our first thoughts when we began this project was to compare our findings in Dharamsala with that of another network in an effort to provide ground truth to our results. Initially we had hoped the San Jose network could provide this, but we really could not compare a server-farm to an ISP that serves thousands of users. As an alternative, we considered a comparison between traffic at the International Computer Science Institute or at Lawrence Berkeley National Labs, as both of these networks have Bro running continuously. However, we ultimately decided against this because even though the comparison was more aligned since both the networks have users, the differences still outweighed the similarities given the other unique factors in Dharamsala network. Thus, due to the unique circumstances of the network, we were fairly limited in the ground truth we could provide for the Dharamsala network.

6 Future Work

However, we do envision two possible networks that could put the results in better perspective as part of future work. The first leverages the location of the Dharamsala network – an ISP in India, preferably one located as close as possible to Dharamsala. This network could show us the typical traffic patterns and malware seen for users in a similar locale. In this case, the harsh conditions would be absent, but there might be more of a likeness in the users. While this would be a great comparison, it is highly unlikely that an ISP would grant us the access we are looking for. A more realistic alternative lies within AirJaldi.

Right now, Dharamsala is the flagship network operated by AirJaldi and the first major installation of its kind outside of the initial testing site. Nevertheless, this is just the beginning, as AirJaldi plans to provide networks to areas with similarly harsh conditions. As soon as a second network with a similar setup to Dharamsala is constructed and established, we envision to transfer our gained security expertise into this new environment. In particular, a network with users who are likely new to the Internet are not familiar with the prevalent malice and may be more susceptible to attacks.

We will continue to collaborate with AirJaldi, install Bro and Snort permanently, and provide the operators with advice on security incidents. Furthermore, we are in contact with FireEye, a company specializing in zero-day attacks. FireEye agreed to donate one of their proprietary network analysis machines, but between legal and transportation issues, the actual installation of the machine has yet to happen. Nonetheless, the process is moving forward and the installation will hopefully occur in the near future. Once this process is complete, the FireEye machine promises to detect zero-day attacks as they appear.

7 Conclusion

Many of the users in Dharamsala are new to the Internet, and may be more susceptible to malware as a result. This, combined with the tense political situation involving China and Tibet make the security of the AirJaldi network in Dharamsala a complex task.

After setting up the monitoring infrastructure in San Jose and Dharamsala, we analyzed the security of the networks. We inspected the traffic from a high level perspective using Bro and Snort and also at the granularity of individual connection contents. Our analysis of the network showcased an interesting variety of malware and traffic pattern, both at the AirJaldi server-farm in San Jose and the Dharamsala network itself. In San Jose, we witnessed many attempted attacks, with most failing, and we even saw the DNS servers used in a possible DoS attack as a traffic reflector and amplifier.

In Dharamsala we were able to identify many different types of malware that had infected machines on the network. We also identified and confirmed instances of GhostNet in the network, lending credence to suspicion of espionage against the community. However, even though we found some intriguing results, we were limited in our manual analysis of the traces, so further interesting results may appear as we continue to analyze the logs.

We believe that our continuing collaboration with AirJaldi will increase the security of the network and provide its users with a safer Internet experience. With the tools we put in place, and with our continued partnership with AirJaldi, we hope we can aid in the detection of any threats against the security and privacy of the users in Dharamsala.

References

[1] AirJaldi. `http://www.airjaldi.org`.

[2] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson. Ignoring the great firewall of china. In *Proceedings of the 6th Workshop on Privacy Enhancing Technologies*, June 2006.

[3] Evan Cooke, Michael Bailey, Z. Morley Mao, David Watson, Farnam Jahanian, and Danny McPherson. Toward understanding distributed blackhole placement. In *WORM '04: Proceedings of the 2004 ACM workshop on Rapid malcode*, pages 54–64, New York, NY, USA, 2004.

[4] Dancho Danchev. Coordinated Russia vs Georgia cyber attack in progress.

http://blogs.zdnet.com/security/?p=1670, August 2008.

[5] Holger Dreger, Anja Feldmann, Michael Mai, Vern Paxson, and Robin Sommer. Dynamic Application-Layer Protocol Analysis for Network Intrusion Detection. In *USENIX-SS'06: Proceedings of the 15th conference on USENIX Security Symposium*, Berkeley, CA, USA, 2006. USENIX Association.

[6] James Fallows. The Connection Has Been Reset. *Atlantic Monthly*, March 2008.

[7] Jan Goebel. A short visit to worm Locksky. http://zeroq.kulando.de/gallery/5567/Locksky.pdf, 2008.

[8] Dan Goodin. Kremlin-backed youths launched Estonian cyberwar, says Russian official. *The Register*, March 2009.

[9] Maarten Van Horenbeeck. Overview of cyber attacks against Tibetan communities. http://isc.sans.org/diary.html?storyid=4177, March 2008.

[10] SecureWorks Joe Stewart. Top Spam Botnets Exposed. http://www.secureworks.com/research/threats/%topbotnets/, April 2008.

[11] Websense Security Labs. Targeted Attacks Use "Recession Relief" Theme. http://securitylabs.websense.com/content/Blogs/%3340.aspx, April 2009.

[12] Michael Lesk. The New Front Line: Estonia under Cyberassault. *IEEE Security & Privacy*, 5(4):76–79, 2007.

[13] Qiao Liang and Wang Xiangsui. *Unrestricted Warfare*. Unknown Panamanian publisher, 1999.

[14] John Markoff. Before the Gunfire, Cyberattacks. *The New York Times*, August 2008.

[15] The Information Warfare Monitor. Tracking GhostNet: Investigating a Cyber Espionage Network. Technical report, SecDev Group (Ottawa) and Citizen Lab (University of Toronto), March 2009.

[16] Shishir Nagaraja and Ross Anderson. The snooping dragon: social-malware surveillance of the Tibetan movement. Technical Report UCAM-CL-TR-746, University of Cambridge, Computer Laboratory, March 2009.

[17] Vern Paxson. Bro: A System for Detecting Network Intruders in Real-Time. *Computer Networks*, 31(23–24):2435–2463, 1999.

[18] Phillip Porras, Hassen Saidi, and Vinod Yegneswaran. An Analysis of Conficker's Logic and Rendezvous Points. Technical report, SRI International, February 2009.

[19] N. Provos, P. Mavrommatis, M.A. Rajab, and F. Monrose. All your iFRAMEs point to us. In *Proceedings of the 17th USENIX Security Symposium*, pages 1–15, 2008.

[20] Martin Roesch. Snort: Lightweight Intrusion Detection for Networks. In *Proceedings of the Systems Administration Conference*, 1999.

[21] Val Smith, Colin Ames, and Delchi. Dissecting Web Attacks. In *BlackHat DC Briefings*, Washington DC, February 2009.

[22] The Snort NIDS. http://www.snort.org.

[23] StopBadware.org. Badware Websites Report. http://www.stopbadware.org/pdfs/StopBadware_Infected_Sites_Report_062408.pdf, May 2008.

[24] ThreatExpert. Report for Trojan-Spy.Agent.ENP. http://www.threatexpert.com/report.aspx?md5=6920fc03428298ef6df37dca71f52e71, April 2009.

[25] Greg Walton. Cyber attacks target pro-Tibetan groups. http://www.infowar-monitor.net/modules.php?op=modload&name=News&file=article&sid=1571, 2008.

[26] Nicholas Weaver, Robin Sommer, and Vern Paxson. Detecting Forged TCP Reset Packets. In *Proceedings of the 16th Annual Network & Distributed System Security Symposium*, February 2009.

[27] F-Secure Weblog. Targeted Malware Attacks Against Pro-Tibet Groups. `http://www.f-secure.com/weblog/archives/00001406.html`, March 2008.

[28] F-Secure Weblog. Spying via XLS files. `http://www.f-secure.com/weblog/archives/00001649.html`, April 2009.

[29] Daniel Wesemann. DNS queries for. `http://isc.sans.org/diary.html?storyid=5713`, January 2009.

[30] Wikipedia. List of TCP and UDP port numbers. `http://en.wikipedia.org/wiki/List_of_TCP_and_UDP_port_numbers`, May 2009.

[31] Vinod Yegneswaran, Paul Barford, and Somesh Jha. Global Intrusion Detection in the DOMINO Overlay System. In *In Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2004.

[32] Vinod Yegneswaran, Paul Barford, and Johannes Ullrich. Internet intrusions: global characteristics and prevalence. *SIGMETRICS Perform. Eval. Rev.*, 31(1):138–147, 2003.

[33] J. Yin and P.M. Taylor. Information Operations from an Asian Perspective: A Comparative Analysis. *Journal of Information Warfare*, 7(1):1–23, 2008.