A Study on WHOIS Privacy or Proxy Abuse on Domains associated with Harmful Internet Communications.

Nonso Emmanuel Ejeana, Pavol Zavarsky, Ron Ruhl, Dale Lindskog

Information Systems Security Department

Concordia University College of Alberta

7128 Ada Boulevard, Edmonton, AB T5B 4E4, Canada

Phone: 1.866.479.5200

nonsoe@yahoo.com, {pavol.zavarsky, ron.ruhl, dale.lindskog}@concordia.ab.ca

Abstract – **A concern was raised that some domains associated with harmful internet communication may use the means of Privacy/Proxy registration to obscure their identity, making it more difficult to investigate and possibly shut down. Our goal is to investigate the degree/scope of this abuse. To conduct our study, we collected a number of confirmed malicious domains on the internet and investigated to determine whether they were registered under privacy/proxy services or not. We did this by conducting WHOIS query on each domain and analyzing the "registrant" and "registrar" sections of the returned WHOIS data, based on the assumptions in the following paragraphs. We contacted about fifty (50) domain registrars to confirm our findings; though none of them were willing to share their subscriber registration data with us for privacy reasons, one of them confirmed that they do not support proxy registration in their country of operation. This information was confirmed in our findings in Fig. 3 below. Our result show the same percentage for malicious domains investigated was registered under privacy/proxy and non-privacy/non-proxy services respectively.**

Keywords: Malicious domains; Proxy/privacy services; WHOIS registration; ICANN community.

I. INTRODUCTION

A concern was raised by the Internet Corporation for Assigned Names and Numbers (ICAAN) [1] that some domains associated with harmful internet communication may use the means of Privacy/Proxy registration to obscure their identity; the use of privacy/proxy services makes these domains more difficult to investigate and shut down [2]. These domains are often used to perpetrate harmful internet communications including phishing, cyber-squatting, intellectual property theft, media laundering, advanced fee fraud, identity theft, child pornography, harassment and stalking [3].

This research answers the following questions regarding the malware domains investigated namely:-

- Are these actually malware domains?
- If yes, did they hide their WHOIS registration data by using proxy/privacy methods?
- Do the majority of malware domains investigated use privacy/proxy services to hide their identity?

The sections that follow outline our research methodology, and summarize our most significant findings.

In the first phase of our analysis, we collected data of about nine hundred (900) malicious domain URLs, out of which Six hundred and fifty-eight (658) unique domain names were extracted, between 5th March to 20th March 2011; from "malwareurl.com". Malwareurl.com [4] is a website of an organization that maintains daily updated list of reported and blacklisted malware domain URLs from other organizations such as "Google Diagnostic Page/Google Safe Browsing"[5], "My WOT/WOT Score Card"[6], "hp Hosts/HP Host Listings"[7] and "MalwareDomainList/MDL listings"[8]. Malwareurl.com further processed these malicious domain URLs by running them through tools from "Virus Total" [9], "Anubis" [10], "Wepawet" [11] and "ThreatExpert" [12] to eliminate false positives.

The list of the malware domains were arranged in a table showing the malicious domain host Internet Service Provider (ISP) IP address, threats posed by each domain, the domain registrar's name if returned by WHOIS query, the registrant's name and contact if available, the country where web server is located, and whether from the WHOIS investigation, we consider the domain registered under privacy/proxy services or not; based on our assumptions in the following paragraphs below.

A. Background on Privacy/proxy Registration

Registrars of domains are required by the ICANN to collect and give free public access to information about of registered domain names and its name servers, date of domain creation, expiry, registered name holder's contact, technical and administrative contacts, to ensure that community identifies the person responsible for a particular domain name [3].

However, registered name holders have the option to limit the amount of personal information returned through public WHOIS queries of registered domain name data bases [13], by the use of privacy and proxy services.

The difference between "privacy" and "proxy" services is that, privacy service providers offer a contact address for use by the registered name holder as an alternative to the registrant's actual addresses and other contact information.

A "proxy" service provider on the other hand actually acts as registered name holder, and grants the use of the registered domain name to its customers or beneficial users for the domain [13]. The contact information returned by a WHOIS query for this type of domain name is therefore that of the proxy service provider or registered name holder and not of the person using the domain name. Both services have the primary aim of limiting the amount of personal information displayed by WHOIS query about the actual domain registrant.

Since the aim of a WHOIS query is to provide information about a domain registrant, it appears that privacy and proxy registration at times may work directly against this goal.

It has been a subject of interest to the ICANN community as to what use domains registered under privacy or proxy services are put to [13]. The anonymity of this registration could possibly be of great attraction to "cyber criminals", and has lead to concerns as to whether "bad actors" in the internet community resort to its use as a way of hiding or obscuring their identity [14].

B. *Related Research*

Other researchers have examined aspects of policies regulating registration of country code Top Level Domains (ccTLDs) and their impact on domains from which malicious activities are perpetrated. Hyacintho [15] conducted a comparative analysis of ccTLDs administration. He used the scores assigned by the Anti-phishing Work Group (APWG) as a metric to estimate the level of malicious activities originating from different ccTLDs. His analysis dealt with the relationship between the incidence of malicious activities and governance of ccTLDs. He found that there is no significant correlation between the policies regulating the registration of ccTLDs and phishing/malicious activities. He suggested that rather enforcement of these policies is the more important factor that determines the percentage of domains registered in each ccTLD involved in malicious activities. He noted however that this was a subject for further investigation.

Another paper by Collins [16], specifically examined the security-related components of ccTLDs administration policies to determine their impact on the level of malicious activities originating from domains under their jurisdiction. He found that easily enforceable and strong security-related policies are necessary to prevent abuse in Internet domains. Further he suggested a role for Internet Service Providers (ISPs) in policing Internet domains and the auditing of registrars by agencies responsible for regulation of each ccTLD to ensure they are accountable for domain names they register.

This research extends the above studies and provides empirical evidence of their findings.

The ICAAN's study, "Prevalence of domain names registered using proxy or privacy services on the top 5 gTLDs", September 2009, indicates that 18% of domain names registered under the top 5 gTLD s are most likely registered under proxy or privacy services [13]. This study however did not determine the percentage of those domains that were associated with malicious internet traffic.

National Opinion Research Center (NORC) at the University of Chicago [17], proposed a design for a study of the accuracy of WHOIS registrant contact information relative to the total population of domain names registered in five generic Top Level Domains (gTLDs) namely; .com, .net, .org, .info, and .biz, which represents about 98.4 % of the 15 global Top-Level Domains [17].

In its work, NORC listed four steps that could be used to verify the accuracy of registrant information by:

- Checking the mailing address of the registrant; this includes the registrants address type and the registrant's address deliverability.
- Classifying the type of registrant; for example, whether name completely missing or patently false (99999), registrant a natural person or registered business/ organization and so on.
- Finding an independent name/address association; for example through phone listing.
- Contacting the registrant to verify if given names are same as the registrant's.

There are daunting tasks and hurdles which limit it's practicability, for instance, since Internet encompasses the whole globe, making international calls to people across the globe to verify their identity could prove particularly difficult bearing in mind differences in privacy laws in different countries.

"Prevalence of Private Registration among malicious domains hosted at 3FN" [18], found that Privacy protection services are used by registrants who hosted commercial contents at 3FN; these findings contradicts beliefs regarding who use privacy protection – the common assumption would be that they are individuals who does not want their contact information exposed on the internet. This study also found that 49% of 3FN hosted domains that use privacy protection services where reported for more than one malicious activity [19].

"Registration Abuse Policy Final Report" [20], wrote that in December 2009, the Generic Names Supporting Organization (GNSO) council agreed to charter a working group to investigate the loop holes identified in their registration policies [21], to decide on whether or not to initiate a Policy Development Process. Her research was to:

- Determine if and how proxy registration abuses are dealt with by those registries that do not have specific abuse policies in place.
- Establish if proxy registration abuse can be curtailed if consistent proxy abuse policy were in place.
- Identify how these proxy abuses are implemented in practice or deemed effective in addressing proxy registration abuse.

She concluded that about 85% proxy registration abused domains used for Phishing are compromised or hacked, therefore, it is unproductive for these domains to be suspended, and therefore mitigation must be performed by the hosting provider [20].

II. ANALYSIS OF MALICIOUS DOMAINS

The following assumptions were made in classifying the domains associated with malicious activities into privacy/proxy registered or otherwise. For all domains classified as registered under "privacy/proxy services", at least one of the following must be true on the returned WHOIS query:

- Registrars name is same as the Registrant's name in the WHOIS record;
- Registrant has a real name but the Administrative or Technical contact is same as the Registrar's, then it is assumed that the name is probably that of the registrar staff assigned the responsibility of managing registrants' accounts;
- Registrant's name not returned in WHOIS query;
- Registrant's name and contact provided appears fictitious, deceptive or unrealistic for example, if the registrant's name is given as "Spam Master", at "2nd hackers Heaven Street, no-where";
- Registrar's name not returned in the WHOIS query.

If none of the above is true, and we were able to extract personally identifying information such as name, street address and other contact information of both the registrar and the registrant of these domains, we classify as non-privacy/non-proxy.

Information extracted using the above outlined process was used to compile Table 1 below.

In Table 1, the nine hundred malicious domain URLs investigated between 5th March 2011 and 20th March 2011 were arranged into various countries under two different titles: (1) "privacy/proxy" and (2) "non-privacy/non-proxy" based on our earlier stated assumptions.

We were able to deduce from this table that out of the Six hundred and fifty-eight malicious domains investigated, 50.1% were found to be "privacy/proxy" registered, while 49.9% were "non-privacy/non-proxy" registered. This information was represented in a pie chart on Fig.1 below.

Table 1 also showed a breakdown of "privacy/proxy" registered and "non-privacy/non-proxy registered malware domains into various countries on the six continents, the predominance of one form of registration against the other in various countries can easily be deduced from this table.

When we converted Table 1 into a bar chart of Fig. 2, we were able not only to compare the privacy/proxy and non-privacy/non-proxy registered malicious domains hosted within a country, but could also compare among all the countries of the six continents.

From Fig. 2, analysis of data on the Six hundred and fifty-eight malicious domains investigated leads to the conclusion that these five host countries have the highest rate of occurrence; China topped the list with 252 reported cases, followed by Czech Republic at 144, USA 105, Poland 88 and other EU countries at 34 combined; Canada has 2 reported cases.

When we broke this statistics down into malicious domains that are "privacy/proxy" registered as shown in Fig. 4, Czech Republic has 47%, followed by Poland at 27%, USA 14%, Ukraine 5%, EU 3% and China 2%; while Canada has less than 1% of the privacy/proxy registered malware domains. This result may indicate that authorities in these countries are more likely to investigate reported malicious internet activities, thereby forcing "bad actors" to hide under the cloak of "privacy/proxy" registration to evade detection.

It may be that country's population has effect on the number of registrants in a particular country; assuming people like to register close to where they live.

It is also a fair assumption to think that, since these domains are use for malicious purposes, people may prefer to use registrars farther away from where they live, to make identification and prosecution more difficult. Therefore, malicious traffics from say China, may have originated from registrants anywhere in the world.

However, there is no confirmation that this is true since our analysis is based on our assumptions and incidences such as WHOIS record omissions, deliberate or otherwise, by the registrars and the registrants may have been included in the category of "privacy/proxy" services in our earlier assumptions.

In Fig. 5, we show a pie chart of all investigated malicious domains under non-privacy/non-proxy registered. China leads in the group at 78%, followed by the USA at 20%, Ukraine and Russia 1% respectively. Other countries are at less than 1%. Apart from China and USA, registrations of malicious domains by non-privacy/non-proxy services seem to be less popular in other countries of the six continents.

China has 78% of all investigated malicious domains registration under "non-privacy/non-proxy" registration; this may be attributed to the following points:

- There may be lack of investigation or prosecution against reported malicious domains, so "bad actors" are not afraid to use their identity to register for domains.
- Domains previously registered under "non-privacy/non-proxy" for legitimate use were hijacked by hackers and subsequently used to propagate malicious internet traffics.
- Identities of real persons were stolen by "bad actors" and used to register a domain and were never verified by the registrars.
- From our previous assumptions, it may be that "Bad actors" are taking advantage of "domain tasting", generating a number of URLs under a temporary malicious domain name and using other person's identity, in order to use such domain URLs to propagate malicious internet traffic.

The above points, especially from the 2nd to the last may also apply to USA. Our analysis show that Canada has less than 1% "non-privacy/non-proxy" registered malicious domains probably because of greater degree of prosecution for personally identifiable malicious domains.

In Fig 3 we show the distribution of the investigated malicious domains against the registrars. Notice the predominance of one form of Malware registration against the other among the registrars. REGRU-REG-RIPN registrar has the highest number of combined malware domain registration, about 90% of all its registration is “non-privacy/non-proxy” registered. This may be the company policy or because of the prevailing laws in its country of operation. The “unspecified” was used to categorize registrars whose name do not appear on WHOIS query, in some cases both the registrars’ and the registrants’ name were either hidden or missing from WHOIS record. Naunet.Reg-fid registrar has 99% malware registration under “Privacy/proxy”, this could be a result of prevailing laws in the country of operation or; since WHOIS data omission and falsification were possibly included in the “privacy/proxy” category, it could be that the registrar simply do not ask or verify the identity of their customers thereby allowing registrants to enter blank, incomplete or invalid contacts. This assertion is based on our earlier assumptions above and could be incorrect.

#	**Country**	**Privacy/ proxy registered domains**	**Non-privacy/ non-proxy registered domains**	**Total domains**
1	Brazil	2	0	2
2	Canada	1	1	2
3	China	7	245	252
4	Czech	144	0	144
5	EU	10	0	10
6	France	1	0	1
7	Germany	1	1	2
8	Latvia	5	0	5
9	Lithuania	1	1	2
10	Netherlands	3	0	3
11	Poland	88	0	88
12	Romania	1	1	2
13	Russia	3	2	5
14	Slovenia	1	0	1
15	South Korea	4	1	5
16	Spain	0	3	3
17	Sweden	1	0	1
18	UK	0	4	4
19	Ukraine	15	5	20
20	USA	42	63	105
21	Vietnam	0	1	1
Total		330	328	658
Percentage		**50.1%**	49.9%	

Table 1 Breakdown of the Six hundred and fifty-eight investigated malicious domains by country and “privacy/proxy” and “non-privacy/non-proxy” registrations.

We arranged the malicious domains of Table 1 by the continent and made a bar chart in Fig.6 to show the distribution of malicious domains on each continent. Whereas the majority of investigated malicious domains in Asia were registered as non-privacy/non-proxy, the reverse is the case in Europe. United States and Canada has almost an equal number for privacy/proxy and non-privacy/non-proxy registrations respectively. This is probably a reflection of various domain registration policies in each continent or region.

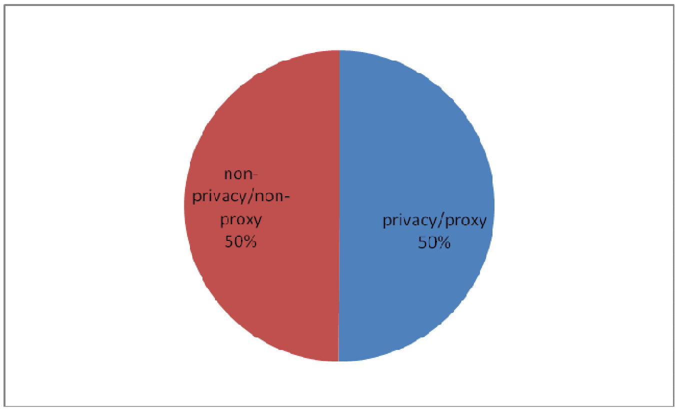

Fig. 1 Pie Chart Representation of the analysis of Six hundred and fifty-eight Malicious Domain URLs; 50% represents both “privacy/ proxy” registered domains and “non-privacy/non-proxy” registered domains respectively.

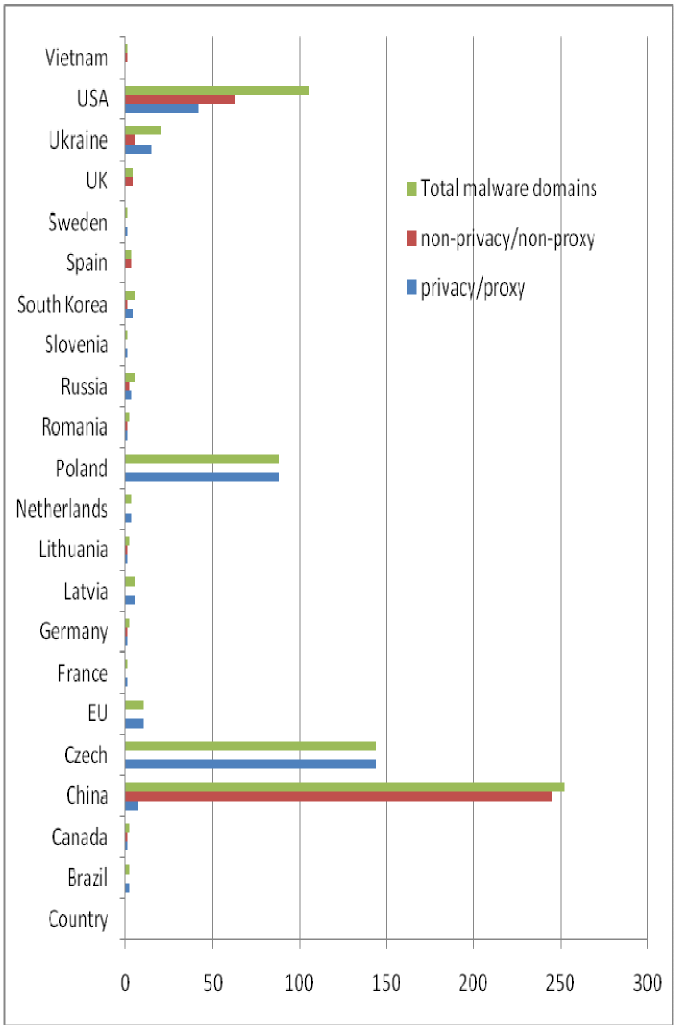

Fig. 2 Bar Chart of the location of Malicious Domains invested. Blue Bars depicts “privacy/proxy” Red Bars depicts “non-privacy/non-proxy” registrations respectively.

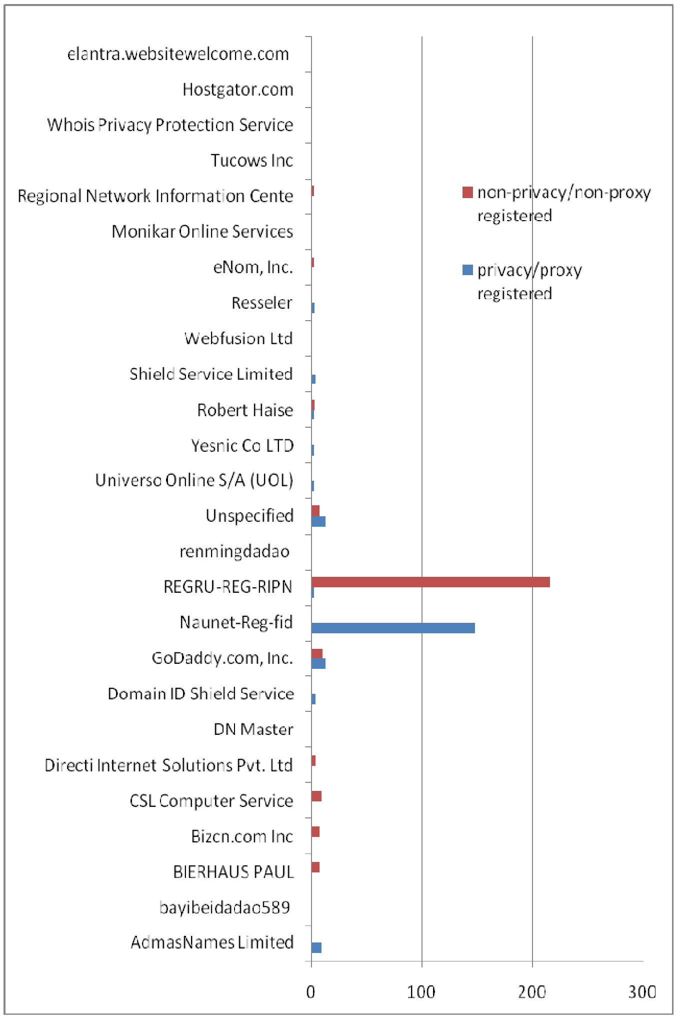

Fig. 3 Bar Chart of Registrars across the six continents of malicious domains investigated. Blue bars depicts “privacy/proxy”, Red bars depicts “non-privacy/non-proxy” registered domains.

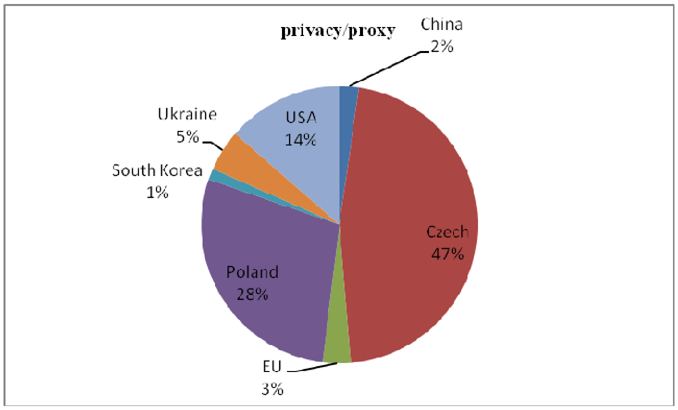

Fig. 4 Pie Chart representing countries of the investigated privacy/proxy registered malicious domains.

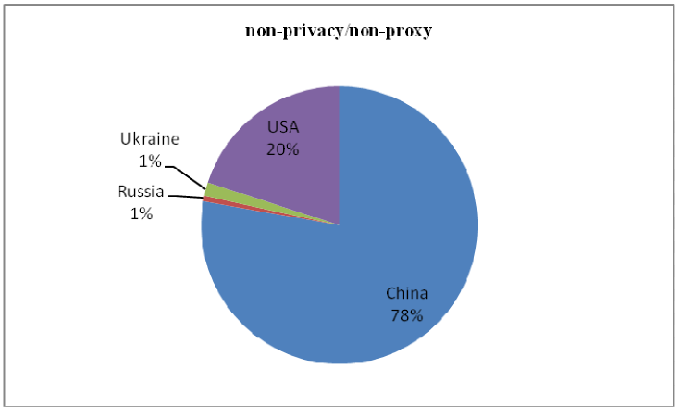

Fig. 5 Pie Chart representing countries of the investigated non-privacy/non proxy registered malware domains.

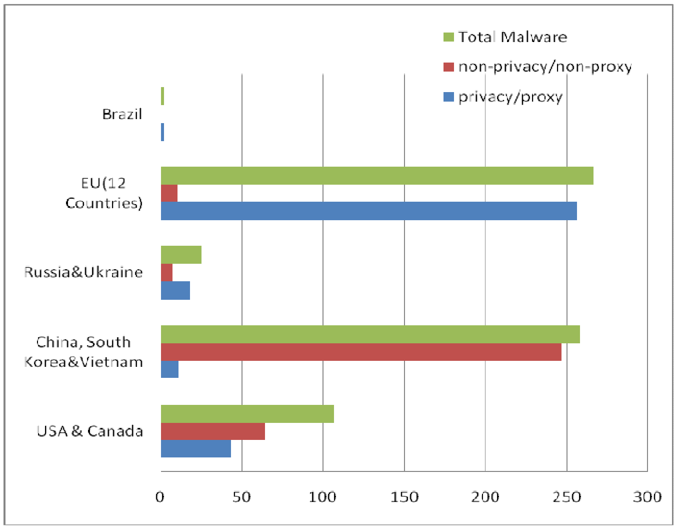

Fig. 6 Bar chart representing privacy/proxy registered malicious domains against non-privacy/non-proxy registered malicious domains around the continents.

III. CONCLUSION

In this research we have been able to analyse malware domains over a period of time based on our earlier stated assumptions, and produced a result which indicate that overall, there is no predominant use of one form of domain registration over the other for domains associated with malicious or harmful internet communications however, when we analyse the data based on geographical locations, we show a clear preference of on form of domain registration over the other for domains associated with malicious internet activities.

This research has also shown the probable distribution of this abuse among countries in the six continents.

We have also shown the probable distribution of this abuse among all registrars, this could probably serve as starting points for ICANN to further investigate the practices of these registrars.

Finally, we were able to discuss the reluctance of registrars to cooperate with researchers or investigators in obtaining registration records for malicious domains, probably due to privacy concerns. ICANN should therefore, probably work to amend relevant sections of their policy to facilitate faster investigation and takedown of malicious domains.

IV. ACKNOWLEDGEMENT

This research was supported by the Faculty of Professional Education and the Faculty of Graduate Studies of Concordia University Collage of Alberta. The first author is very grateful to his supervisors: Pavol Zavarsky, Ron Ruhl and Dale Lindskog for their support and guidance during this research, and to my colleagues, Ayonote Clement and Oludele Ogundele for their help in the investigation of malicious domains.

V. REFERENCES

[1] Request for Proposals, WHOIS Privacy and Proxy Abuse Studies, 18 May 2010.

[2] NORC University of Chicago “Draft Report for the Study of Accuracy of WHOIS Registrant Contact Information”, Project Reference: 6558, 6636, 17 January 2010.

[3] WHOIS Proxy / Privacy Abuse Study, Working Draft 18 May 2010.

[4] Malware URL. [Online] Available: http://www. Malware url.com/index.php.

[5] Google safe browsing [Online] Available http://www. grape thinking . com/ google-safe- browsing- diagnostic#

[6] My WOT [Online] Available: http://www.mywot.com/

[7] Hp Hosts. [Online] Available: http:// hosts- file. net

[8] MalwareDomains. [Online] Available: http://www.mal ware domains.com.

[9] Virus Total [Online] Available: http://www.virustotal.com

[10] Anubis. [Online] Available: http:// anubis. iseclab.org/

[11] Wepawet. [Online] Available: http://wepawet.cs.ucsb.edu/

[12] ThreatExpert[Online]Available:http://www.threatexpert.c om

[13] ICANN’s Study on the Prevalence of Domain Names registered using a Privacy or Proxy Service among the top 5 gTLDs, 28 September 2009.

[14] Rod R et al. “Advisory on Utilization of WHOIS Data for Phishing Site Take Down”, APWG, March 2008.

[15] M.Hyacintho,“Internet Security Governance: Comparative Analysis of Country Code Top Level Domain (ccTLD)Analysis,2008.[Online]Available:http://www.info sec.concordia.ab.ca/system/files/Hyacintho2009.pdf

[16] C. Umana, “Comparative Analysis of ccTLD Security Policies” 2009, [Online] Available: http://www.infosec. concordia.ab.ca /system/files/Umana2010.pdf

[17] NORC,” Proposed Design for a Study of the Accuracy of WHOIS Registrant Contact Information”, Project Reference: 6558, 6636, 3 June 2009.

[18] D. Piscitello et al. “Domain Name Privacy Misuse Studies”; ICANN INET Asia, Hong Kong, 13 April 2010.

[19] D. Piscitello, “Privacy Domain Registrations: examining the relationship between private domain registrations and malicious domains at 3FN”, 19 October 2009.

[20] M. Konings, “Registration Abuse Policy Working Group Final Report” RAPWG, May 29, 2010.

[21] GNSO Drafting Team, “Definition for Key Terms that may be used in Future WHOIS Studies”, 18 February 2009.