PURPLE PAPER
EXEGESIS OF VIRTUAL HOSTS HACKING

BY PETKO PETKOV AND PAGVAC (ADRIAN PASTOR)

VIRTUAL HOST ENUMERATION TECHNIQUES

There is a lot that we can say about finding virtual hosts from a given IP address. Sometimes this task is straightforward, other times a bit of thinking is required. However, in general it is not a mission impossible.

During the last few years, domain name databases have emerged like mushrooms after a rainy day. This has certainly increased the awareness among security professionals about the possibility of using virtual hosts as backdoors when testing the security of a given organization. In reality, a good attacker will try to break into your organization by knocking on the not-so-obvious doors.

The process of getting all valuable virtual hosts *usually* falls into the passive, enumeration gathering practices and it is based on querying databases from the public sector. However, we will also look at some active enumeration techniques for finding virtual hosts.

In the following subsection we will discuss how to find virtual hosts by querying public databases and actively probing the domain name system (DNS) and the HTTP protocol itself.

QUERYING PUBLIC DATABASES

Querying databases designed for the public is probably the easiest approach when searching for virtual hosts. Moreover, it is considered to be the stealthiest since the attacker never sends packets to the target. As such, this is the preferred technique by someone who is trying to break into your organization without being noticed.

There are many public domain databases but just a few worth our attention:

- http://www.domainsdb.net/
- http://www.searchmee.com/web-info/ip-hunt.php
- http://www.whois.sc/reverse-ip/
- http://whois.webhosting.info/
- http://serversniff.net/content.php?do=hostonip

Among them we find that whois.sc and webhosting.info have the most complete set of results, but they are a bit restricted. On the other hand, domainsdb.net, searchmee.com and serversniff.net provide less complete results but more freedom. Of course it is worth mentioning that attackers may choose databases that can be easily integrated into automated scripts and as such, speeding up the information gathering stage.

! **WARNING**

Do not use any automated tools without permission from the data provider as it may break the terms and conditions policy.

You might be searching for the best database that will give you the most complete and accurate results. However, the truth is that none of them are good enough. A proper penetration test should involve collecting data from all available databases. The more diverse your data is, the more results you will get and the better security evaluation you will provide.

The main problems we encountered when querying public domain databases are summarized in the following list:

- some of these services require you to pay a membership fee
- registration is sometimes required
- the domain names returned in the results are sometimes outdated (they no longer resolve to the same IP address)
- a domain name is not necessarily a virtual hosts (i.e. ftp.mydomain.com, smtp.mydomain.com)
- anti-abuse techniques (i.e. entering code included in images) make it hard to write scripts which automate the queries and parse the results
- some of these on-line databases are unreliable (not always available) because some of them belong to small organizations

Fortunately for us, there is another database worth our attention. Surprisingly, this is Microsoft MSN Search engine. During our search for the ultimate virtual host database, we found an extremely nice feature of MSN search engine. This is the search by IP facility.

Normally, this feature works in the same way Google's “site” directive works. For example, if we are interested in finding known virtual hosts of organization X we can fire a query that looks like the following:

```
ip:X.website.ip.address
```

Basically, MSN will give all pages it knows about (indexed pages) that belong to domains that resolve to *X.website.ip.address*. From that point on, all you have to do is to collect all URLs, extract the domain part, *uniq*ue and *sort* the results if you feel like it.

To some extent MSN and Google are probably the best databases to query, mainly because they have a lot of resources that crawl most *public* websites 24/7. Since virtual hosting is a technology that concerns the HTTP/1.1 protocol, making MSN and Google extract information from all websites they know about, is considered as one of the most accurate passive technique for finding virtual hosts.

SUBDOMAIN ENUMERATION

There are two major obstacles that need to be overcome when finding a comprehensive list of virtual hosts. The first one is the completeness of the database results and the second one is finding virtual hosts that are not exposed to the public.

Considering that MSN Search is not as mature as Google (in our humble opinion), the second one should give you better results. Unfortunately, Google does not provide the “ip” directive but it has the “site” directive that can be used in the same way but with a bit of a twist.

```
site:company.com
```

The above statement will find subdomains that belong to company.com (i.e support.company.com, dev.company.com, login.company.com). If you run such a query against a large organization, most of the results that Google returns will be related to the organization's main website, typically www.company.com. In order to start enumerating more interesting subdomains you can use the minus “-” operator:

```
site:company.com -www
```

The following is a real-life query against microsoft.com:

```
http://www.google.com/search?hl=en&q=site%3Amicrosoft.com+-
www&num=100
```

Using search engine tags and operators is a quick and straightforward technique for finding virtual hosts but let's be honest, Google and MSN are not as mighty as they want to be. In fact, sometimes they are not even close. So, there might be other solutions that can bring a bit of a flavor to your findings.

As you might have already figured out, this is the brute-force technique. This technique has been known for quite a while and there is nothing magical about it. All you need is a good dictionary file, the *dig* (or *host*) Unix command and some basic shell scripting skills. The process is not stealth but not too obvious since querying DNS is simple and legitimate, and it happens quite frequently.

! **WARNING**

Subdomain brute-forcing may not be possible in cases where the the target domain uses wildcards '*' in the zone file of the DNS server. Hence, all brute-forced subdomains will resolve to the same valid IP address (“fake” responses).

ACTIVE VIRTUAL HOST PROBING

Last but not least, attackers can actively probe for virtual hosts by using the HTTP/1.1 protocol. This method is quite loud and too obvious for system administrators, but it's still a valid enumeration technique which happens to be the most accurate one.

The idea is very simple: use a large dictionary file of different domain names and probe each of them by using the following template:

```
GET / HTTP/1.1
Host: domain_name_here
<CRLF>
<CRLF>
```

If you get a positive result, then you just found a valid virtual host. If not, continue probing. In order to find what makes a positive result, you'll need to analyze the responses of the server in cases where a valid virtual host is sent in the *Host* header, in comparison to cases where a non-existent virtual host is sent.

Change of web server responses, when sending a non-valid *Host* header, can be found in the HTTP headers (i.e.: “HTTP/1.1 301 Moved Permanently”) and/or in the actual HTML code (i.e.: “Error: This is not a Proxy.”).

In summary, this technique will *only* work against web servers which behave differently depending on whether or not a valid *Host* header is sent.

There are however, a few “dirty tricks” that attackers can use to obfuscate active virtual host probing against your web servers. Some of them are:

- changing the *User-agent* HTTP headers to that of a commonly known search engine spider. The request would then look similar to the following:

  ```
  GET / HTTP/1.1
  Host: www.target.com
  User-Agent: Googlebot/2.1 (+http://www.google.com/bot.html)
  ```

- using HTTP or SOCKS proxies in order to hide the attacker's true identity
- introducing delays between each request
- exploiting web server logging bugs (i.e. IIS web server HTTP TRACK logging failure vulnerability - http://www.securityfocus.com/bid/9313/)

SECURITY THROUGH OBSCURITY

!

You can hide administrative interfaces behind virtual hosts that are not configured on your DNS server. This means that querying DNS servers will not give away your “secret” virtual host. Only your web server will know about it.

Ideally you should use a hard-to-guess virtual host name, so that attackers cannot find it through active virtual host probing.

ATTACK SCENARIOS

We have already discussed the techniques that are involved in finding virtual hosts. However, the real questions is what the bad guys can do with this information. Well, we believe that a lot can be done. Raising awareness of the security industry about these issues is a crucial step to help organizations protect against them. The following scenarios are worth our attention:

- defacing your website by attacking virtual hosts that contain vulnerable scripts which allow attackers to gain access to the remote machine (due to poor filesystem permisions for instance).
- finding administrative interfaces that could be:
  - unauthenticated
  - protected with easy-to-guess or default credentials
  - vulnerable to data injection attack that could be used to bypass the authentication stage completely
- finding information about your organization (i.e.: is your organization serious enough to be willing to spend money in dedicated hosting? who else is sharing resources with you? if other companies are sharing resources with you, are they subdivisions of your company, or maybe political partners/allies?)
- defaming the name of your organization by finding websites that host content of explicit nature (i.e.: pornographic or pedophile websites), or content of non-serious nature (i.e.: a personal homepage of one of your employees for instance)
- performing domain hijacking attacks by compromising the security of a virtual host with higher privileges (cpanel or plesk are good examples).

SURVEY ON UK SECURITY TESTING COMPANIES

Hosting your organization's website on a third-party hosting company and using dedicated web hosting are two security practices that we *do* recommend. The reasons should be obvious after reading this paper. Although these two practices do *not* seem to be widely acceptable in the computer security industry at time of writing, we do hope however that this paper will make security professionals consider the issues involved in violating any them.

We thought it would be interesting to perform a survey on UK computer security companies to find out which ones follow our two recommended security practices. As a reminder these are:

- Avoid hosting your organization's website on your corporate network – use a third party hosting provider instead
- Avoid shared hosting – use dedicated hosting instead

ORGANIZATION MAIN WEBISTE	SHARED HOSTING	ON-SITE HOSTING	APPROXIMATE NUMBER OF VIRTUAL HOSTS
www.boldonjames.com	No	Yes	1
www.btconsulting.com	No	Yes	2
www.contextis.co.uk	No	Yes	1
www.convnet.co.uk	Yes	No	30
www.corsaire.com	No	Yes	1
www.deloitte.com	No	Yes	1
www.dns.co.uk	No	Yes	1
www.echelonltd.com	No	Yes	1
www.eds.com	No	Yes	1
www.insight.co.uk	No	No	1
www.integralis.co.uk	No	Yes	1
www.irmplc.com	No	Yes	2
www.kpmg.co.uk	No	No	5
www.logicacmg.com	No	No	1
www.mwrinfosecurity.com	Yes	Yes	6
www.nccgroup.com	No	No	1
www.verizonbusiness.com	No	Yes	3
www.ngssoftware.com	No	Yes	2
www.nta-monitor.com	Yes	No	159
www.peapod.co.uk	No	Yes	4
www.portcullis-security.com	Yes	No	170
www.qinetiq.com	No	Yes	2
www.sapphire.net	No	Yes	1
www.securetest.com	No	Yes	1
www.seleniacomms.com	No	Yes	3
www.sopranewellandbudge.com	No	Yes	2
www.symantec.com	No	No	1
www.vega-group.com	Yes	No	40

LEGEND

Organization main website - all the websites were obtained from the CESG list of CHECK service providers in the UK (http://www.cesg.gov.uk/site/check/index.cfm?menuSelected=11&displayPage=11)

Shared hosting – third-party organizations' websites are hosted on the same physical server. This is typically found when organizations host their website using a low cost third-party hosting service.

On-site hosting – the organization's website is hosted on their corporate network, that is a public network range registered to them (or any of the organizations belonging to the same group). We found this information through registry queries and web-based research (completely legal enumeration)

Approximate number of virtual hosts – we considered each *different* website a different virtual host. Also, these virtual hosts are located on the same physical web server (they all resolve to the *same* IP address). When the number of virtual hosts equals 1, it probably means the web server is *not* using virtual hosts.

INTERESTING THINGS WE FOUND

We found a number of interesting websites/interfaces using some of the passive virtual hosts enumeration techniques previously described in this paper. All these techniques are absolutely legal. Such websites belong to some of the computer security companies which were included in the previous survey. The sensitive data has been hidden for obvious reasons.

- http://www.<domain_name>.com/ - this is a personal homepage located on the same web server in which www.<security_company_name>.co.uk is located. After some research we found out that this site is run by <name> <surname>, webmaster of <security_company_name> UK's website. This is a good example of the type of content that can damage the image of your organization.
- https://www1.<security_company_name>.com/ - <security_company_name>'s login interface for registered customers to download software.
- http://www.<security_company_name>.com/be/extranet/ - Remote extranet access login interface of <security_company_name>'s Belgium gateway.
- http://<domain_name>.com/ - personal homepage showing wedding pictures of one of <security_company_name>'s employees.
- http://www.<domain_name>.co.uk/ - pornographic website hosted on the same web server where <security_company_name>'s main websites is located. This is another good example of content that can damage the image of your organization.
- http://my<security_company_name>portal.com/ - HTTP basic authentication login interface

POC

We implemented some of the virtual host enumeration techniques discussed in this paper in a tool called *venum* (virtual hosts enumerator).

The command line version of venum is available at met.gnucitizen.org. There is also a web based frontend available at www.ikwt.com

!

WARNING

venum is solely provided as a proof of concept which illustrates some of the techniques described in this paper. If you abuse this tool you will break the terms and conditions policy of the data provider. Additionally, the IP address from which the abusive requests are made will most likely be banned by the data provider.

CONCLUSION

Organizations concerned about their image and security should *not* consider shared hosting (use of virtual hosts) as a solution, despite the reduction of costs. On-site hosting, that is, hosting websites on your organization's corporate network rather than a third party hosting company, should also be avoided in order to mitigate the risks once an attacker gains intranet access to your organization from the outside world (Internet).

Now in plain English - If your organization is serious about security and to keep the trust of your customers, you should NOT:

- publish your websites using shared hosting
- host your website on your corporate network

REFERENCES

ON-LINE DOMAINS DATABASES:

- Domainsdb [http://www.domainsdb.net/]
- IP Hunt [http://www.searchmee.com/web-info/ip-hunt.php]
- Reverse IP [http://www.whois.sc/reverse-ip/]
- Webhosting.info [http://whois.webhosting.info/]
- Serversniff hostnames-on-ip [http://serversniff.net/content.php?do=hostonip]
- Zone-h web defacement archive [http://www.zone-h.com/]

POC TOOLS:

- venum (vhost enumerator) [http://met.gnucitizen.org/], [http://www.ikwt.com/venum]
- revhosts [http://www.revhosts.net/index.php]
- Dmitry – [http://mor-pah.net/code/download.php?file=DMitry-1.2a.tar.gz]

RELEVANT DISCUSSIONS:

- http://lists.grok.org.uk/pipermail/full-disclosure/2005-October/thread.html#38112
- http://archives.neohapsis.com/archives/sf/pentest/2005-08/thread.html#294
- http://www.securitytrap.org/fd/5170

CREDITS

Petko Petkov [pdp a t gnucitizen.org]

pagvac (Adrian Pastor) [unknown.pentester a t gmail.com]

– February 2006.