Detecting and Measuring Security Risks of Hosting-Based Dangling Domains

Mingming Zhang
zmm18@mails.tsinghua.edu.cn
Tsinghua University
Beijing, China

Xiang Li
x-l19@mails.tsinghua.edu.cn
Tsinghua University
Beijing, China

Baojun Liu*
lbj@tsinghua.edu.cn
Tsinghua University and Quan Cheng Laboratory
Beijing, China

Jianyu Lu
QI-ANXIN Technology Research Institute
Beijing, China

Yiming Zhang*
zhangyiming@tsinghua.edu.cn
Tsinghua University
Beijing, China

Jianjun Chen
Tsinghua University and Zhongguancun Laboratory
Beijing, China

Haixin Duan
Tsinghua University and Quan Cheng Laboratory
Beijing, China

Shuang Hao
University of Texas at Dallas
Richardson, Texas, USA

Xiaofeng Zheng
Tsinghua University and QI-ANXIN Technology Research Institute
Beijing, China

ABSTRACT

Public hosting services offer a convenient and secure option for creating web applications. However, adversaries can take over a domain by exploiting released service endpoints, leading to *hosting-based domain takeover*. This threat has affected numerous popular websites, including the subdomains of *microsoft.com*. However, no effective detection system for identifying vulnerable domains at scale exists to date. This paper fills the research gap by presenting a novel framework, HostingChecker, for detecting domain takeovers. HostingChecker expands detection scope and improves efficiency compared to previous work by: (i) identifying vulnerable hosting services using a semi-automated method; and (ii) detecting vulnerable domains through passive reconstruction of domain dependency chains. The framework enables us to detect the subdomains of Tranco sites on a daily basis. It discovers 10,351 vulnerable subdomains under Tranco Top-1M apex domains, which is over 8× more than previous findings, demonstrating its effectiveness. Furthermore, we conduct an in-depth security analysis on the affected vendors (e.g., Amazon, Alibaba) and gain a suite of new insights, including flawed domain ownership validation implementation. In the end, we have reported the issues to the security response centers of affected vendors, and some (e.g., Baidu and Tencent) have adopted our mitigation. The full paper is provided in [2].

CCS CONCEPTS

• **Security and privacy → Network security**.

KEYWORDS

public hosting service; domain takeover

ACM Reference Format:
Mingming Zhang, Xiang Li, Baojun Liu, Jianyu Lu, Yiming Zhang, Jianjun Chen, Haixin Duan, Shuang Hao, and Xiaofeng Zheng. 2023. Detecting and Measuring Security Risks of Hosting-Based Dangling Domains. In *Abstract Proceedings of the 2023 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '23 Abstracts), June 19–23, 2023, Orlando, FL, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3578338.3593534

1 INTRODUCTION

In this paper, we focus on the threat model that attackers try to take over victim domains by exploiting vulnerable public hosting services. In the following, we refer to this threat as "hosting-based domain takeover" and relevant dangling domains as $D_{vulhost}$. Attackers can manipulate the endpoints of the victim domains by deploying a new service on vulnerable platforms.

We introduce HostingChecker, an effective framework to spot hosting services and detect vulnerable domains hosted on discontinued services. Figure 1 depicts the architecture of HostingChecker. First, we design a semi-automated service discoverer (Part 1) based on the observation that public hosting services share common features that can be mined from passive DNS (PDNS) data. For example, endpoints of the same service use identical naming conventions, and service endpoint names are highly depended by customers' domain names. Second, we propose a passive method to efficiently detect vulnerable domains (Part 2). Rather than active DNS queries, we reconstruct DNS resolution chains for detected domains from the PDNS logs.

2 HOSTINGCHECKER DESIGN

Our high-level idea for detecting hosting-based domain takeover threats includes: (i) identifying hosting services and (ii) discovering vulnerable domains hosted on discontinued services.

[*] Corresponding Authors: Baojun Liu and Yiming Zhang.

SIGMETRICS '23 Abstracts, June 19–23, 2023, Orlando, FL, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0074-3/23/06.
https://doi.org/10.1145/3578338.3593534

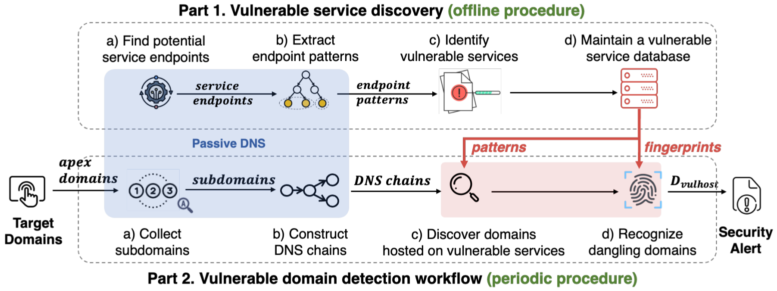
Part 1. Vulnerable service discovery (offline procedure)
a) Find potential service endpoints
b) Extract endpoint patterns
c) Identify vulnerable services
d) Maintain a vulnerable service database
service endpoints
endpoint patterns
Passive DNS
patterns
fingerprints
apex domains
subdomains
DNS chains
$D_{vulhost}$
Target Domains
a) Collect subdomains
b) Construct DNS chains
c) Discover domains hosted on vulnerable services
d) Recognize dangling domains
Security Alert
Part 2. Vulnerable domain detection workflow (periodic procedure)

Figure 1: Overview of HostingChecker modules.

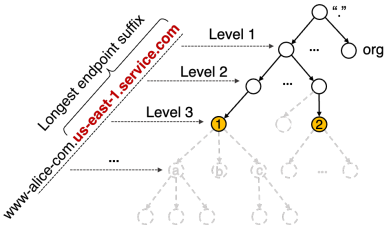

Figure 2: Domain suffix tree.

Empirical Observation. We design HostingChecker based on the following observations:

O1 Public hosting services adopt similar endpoint patterns, as illustrated in Figure 3.

O2 Custom domains heavily rely on public hosting services' endpoint domains, namely, a large number of custom domains are resolved to the endpoint domains.

O3 Reconstructing DNS resolution chains from PDNS data can improve detection efficiency and reduce network resource consumption compared to active DNS resolution.

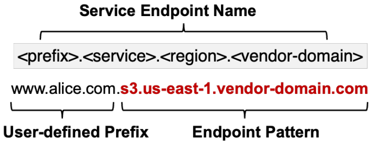
Service Endpoint Name
<prefix>.<service>.<region>.<vendor-domain>
www.alice.com.s3.us-east-1.vendor-domain.com
User-defined Prefix
Endpoint Pattern

Figure 3: Service endpoint name pattern.

System workflow. HostingChecker utilizes PDNS for service discovery and vulnerable domain detection by exploiting the above characteristics of hosting services and DNS chains.

Part 1: a semi-automated service discoverer. (i) The discoverer automatically gathers potential endpoint names that may belong to public hosting services using domain dependency (O2). (ii) It extracts endpoint patterns from endpoint candidates using a novel *Domain Suffix Tree* (O1), as shown in Figure 3. (iii) We detect vulnerabilities by identifying service types and examining hosting policies, while also collecting HTTP (e.g., response headers) and DNS (e.g., resolution answers) fingerprints that indicate discontinued vulnerable services. (iv) Vulnerable endpoint patterns and fingerprints are fed into Part 2.

Part 2: an efficient $D_{vulhost}$ detector. HostingChecker detects $D_{vulhost}$ by examining if a domain depends on a vulnerable and discontinued service. It (i) gathers subdomains and DNS RRs from PDNS logs by their query volume, (ii) reconstructs the subdomains' DNS resolution chains by linking the extracted CNAME RRs, (iii) checks them against endpoint patterns obtained in Part 1, and (iv) inspects their HTTP and DNS fingerprints to detect a discontinued service. In the end, it generates a security alert if it finds a $D_{vulhost}$.

3 IMPLEMENTATION AND EVALUATION

We implement HostingChecker in Golang and deploy it on a PDNS dataset that is collected by the largest DNS provider (114DNS[1]) in China. The dataset contains 101B DNS responses and covers 99.9% domains in the Tranco Top 1M list. Compared to previous studies, HostingChecker has a higher detection efficiency. Experiments show that HostingChecker takes about one day in each measurement round, and takes only 13.9 hours for DNS chain reconstruction, which outperforms active DNS resolution methods. In addition, the detection accuracy of the system is about 95%, among which 46.4% have already eliminated the threats by removing dangling DNS records, and the 5% false positives are mainly caused by the domain whitelist maintained by the platforms.

4 MEASUREMENT FINDINGS

We conduct over 101 rounds of measurements from Dec. 16, 2021, to Jul. 28, 2022, examining 11,446,359 subdomains under Tranco Top-1M, 9,808 .edu and 7,198 .gov apex domains. Our key contributions and findings are summarized below.

(i) A holistic characterization of public hosting services. We identify 65 hosting services, including 34 new ones, across 52 public vendors (e.g., Alibaba, Amazon), which can be exploited for domain takeover. We find public service providers employ diverse domain connection methods, and 7 out of 9 methods are vulnerable to domain takeover. We also identify 4 flawed implementations that enable bypassing domain ownership validation (DOV) and affect the top 20 hosting vendors. We report the vulnerabilities to the affected vendors and receive confirmation from ten, including Amazon, Tencent, and Huawei.

(ii) A longitudinal measurement of $D_{vulhost}$ among high-profile domain names. We detect that 114,063 (1.0%) of tested domains are hosted on vulnerable services, with 10,351 $D_{vulhost}$ (8× more than previous findings) from 2,096 popular apex domains, including reputable universities (e.g., Stanford and Rice) and companies (e.g., Baidu, Huawei, and Marriott). We find such threats appear frequently and are long-lasting. For example, 270 new $D_{vulhost}$ emerge per week, and 60% remain vulnerable for over 5 days, leaving a substantial attack time window. Moreover, 45.5% of $D_{vulhost}$ continually receive queries from 70 million client IPs, indicating the need for timely detection to reduce attack surfaces.

REFERENCES

[1] 2022. *114 DNS.* https://www.114dns.com/

[2] Mingming Zhang, Xiang Li, Baojun Liu, Jianyu Lu, Yiming Zhang, Jianjun Chen, Haixin Duan, Shuang Hao, and Xiaofeng Zheng. 2023. Detecting and Measuring Security Risks of Hosting-Based Dangling Domains. *Proc. ACM Meas. Anal. Comput. Syst.* 7, 1, Article 9 (mar 2023), 28 pages. https://doi.org/10.1145/3579440