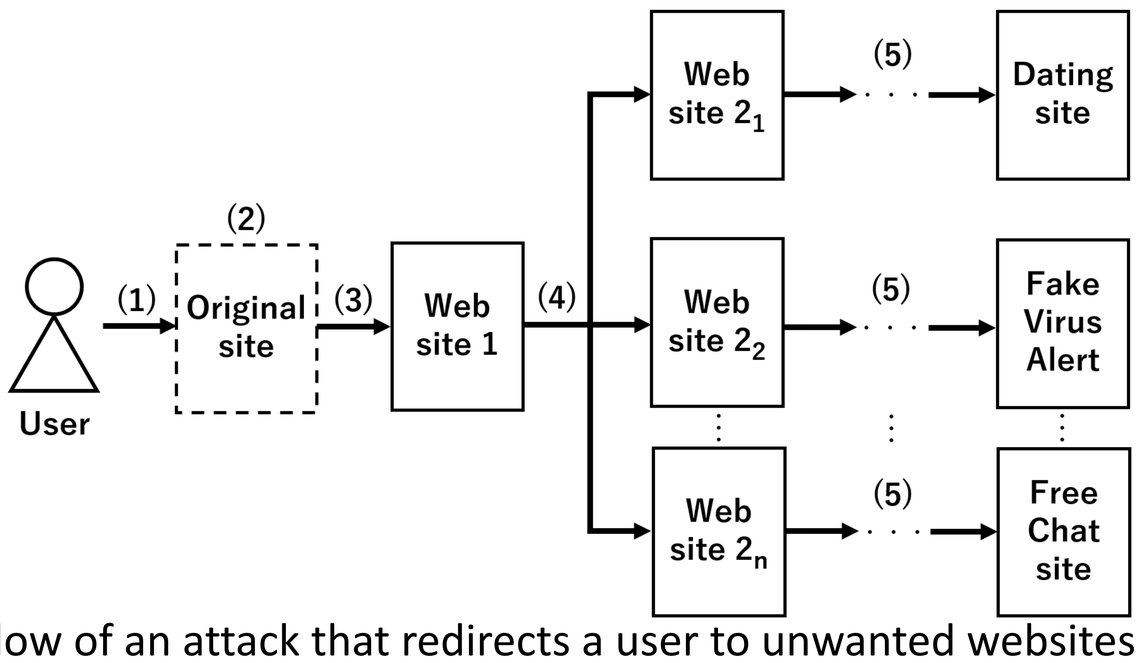
Method of Generating a Blacklist for Mobile Devices by Searching Malicious Websites

Takashi Ishihara, Masaya Sato, Toshihiro Yamauchi (Okayama University)

1. Introduction

- Mobile devices are more frequently targeted in cyberattacks and web-based attacks have been reported.
- We proposed a method for searching malicious websites and generating a blacklist for mobile devices.

2. Attacks of Redirecting a User to Unwanted Websites



Flow of an attack that redirects a user to unwanted websites [1]

(1) Visiting the malicious landing site.
(2) Tapping anywhere on the screen.
(3) Redirecting to intermediate site 1 (Website 1).
(4) Redirecting to intermediate site 2 (Website 2).
(5) Redirecting to the unwanted website.

- The URL of the intermediate site is generated from the specified URL and a randomly created character string [1].
- The URL of the unwanted website includes the user's device information [1].

3. Methods of Generating Blacklists

Data Collection Unit

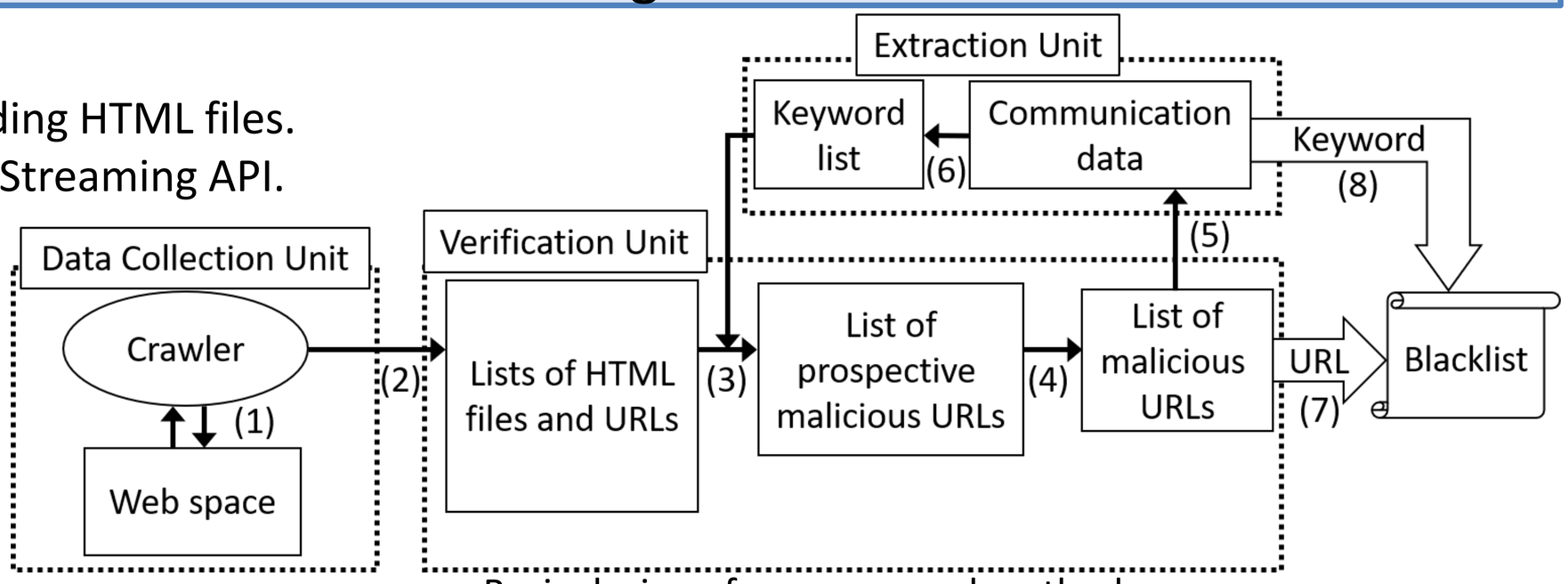
- Collecting URLs and corresponding HTML files.
- To collect URLs, using Twitter's Streaming API.

Verification Unit

- Searching the HTML files using keywords extracted from known malicious websites and finding URLs that are highly likely to be malicious.
- Checking a malicious URL by manual access.

Extraction Unit

- Extracting the keywords used for keyword searches and blacklist.



Basic design of our proposed method

keyword to extract for keyword searches and blacklist

Target	Keywords to extract
HTML files of the landing site	The filename (e.g., example.js) that causes the redirection
	FQDN that provides the file that causes the redirection
URL of the intermediate site	FQDN
URL of the unwanted website	FQDN

4. Evaluation

1. Number of malicious websites discovered

- 122,350 websites that collected between July 23 and December 16, 2019.
- 200 landing sites were discovered and 182 landing sites could not be detected by Google Safe Browsing.
- As the keywords, 3 filenames and 108 FQDNs were extracted.

2. Detection rate of malicious websites using a blacklist

- Using the blacklist generated by the proposed method in Evaluation 1.
- The five malicious websites collected using the proposed method from December 20 to December 30, 2019 were used.
- The blacklist detected all 10 accesses at four websites.

5. Conclusions

- We found 182 landing sites that could not be detected by Google Safe Browsing.
- The blacklist using keywords has a sufficiently high detection rate for specific malicious websites.
- In future work, we will evaluate the detection rate of our blacklist for common malicious URL lists.

Acknowledgement : The research results have been achieved by "WarpDrive: Web-based Attack Response with Practical and Deployable Research InitiatiVE," the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan.

[1] Imamura, Y., Orito, R., Chaikaew, K., Manardo, C., Leelaprute, P., Sato, M., Yamauchi, T: Threat Analysis of Fake Virus Alerts Using WebView Monitor, In: 2019 Seventh International Symposium on Computing and Networking (CANDAR), pp.28-36 (2019).