(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 October 2013 (03.10.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/149178 A3

(51) International Patent Classification:
G06F 12/14 (2006.01)

(21) International Application Number:
PCT/US2013/034660

(22) International Filing Date:
29 March 2013 (29.03.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13/434,311 29 March 2012 (29.03.2012) US
13/607,789 9 September 2012 (09.09.2012) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 13/607,789 (CON)
Filed on 9 September 2012 (09.09.2012)

(71) Applicant: **CYBER ENGINEERING SERVICES, INC.** [US/US]; Suite 170, 8825 Stanford Blvd., Columbia, Maryland 21045 (US).

(72) Inventors: **BOJAXHI, Hermes**; Suite 170, 8825 Stanford Blvd., Columbia, Maryland 21045 (US). **DRISSEL, Joseph**; Suite 170, 8825 Stanford Blvd., Columbia, Maryland 21045 (US). **RAYGOZA, Daniel**; Suite 170, 8825 Stanford Blvd., Columbia, Maryland 21045 (US).

(74) Agent: **DIENWIEBEL, Thomas**; Dienwiebel Transatlantic Intellectual Property, 2479 East Bayshore Road, Suite 800, Palo Alto, California 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(88) Date of publication of the international search report:
21 November 2013

(54) Title: SYSTEMS AND METHODS FOR AUTOMATED MALWARE ARTIFACT RETRIEVAL AND ANALYSIS

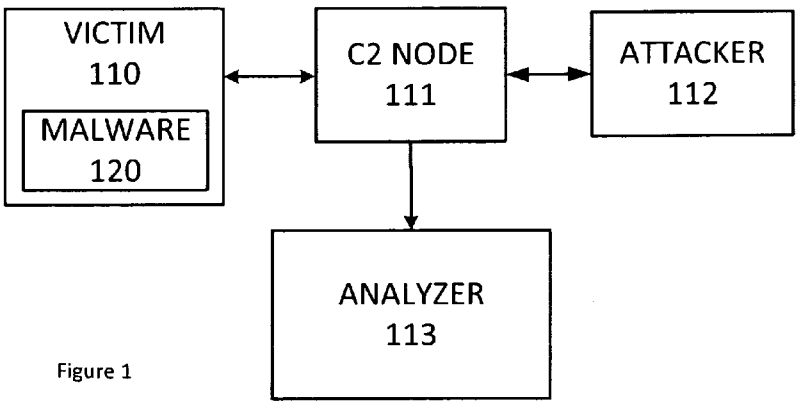

Figure 1

(57) Abstract: An automated malware analysis method is disclosed which can perform receiving a first universal resource locator identifying a first intermediate network node (403), accessing the first intermediate network node (403) to retrieve a first malware artifact file, storing the malware artifact file in a data storage device (245), analyzing the malware artifact file to identify a second universal resource locator (404) within the malware artifact file, and accessing a second intermediate network node to retrieve a second malware artifact file.

WO 2013/149178 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US13/34660

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 12/14 (2013.01)
USPC - 726/23
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) Classification(s): G06F 12/14 (2013.01)
USPC Classification(s): 726/24, 22, 12, 23

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MicroPatent (US-G, US-A, EP-A, EP-B, WO, JP-bib, DE-C,B, DE-A, DE-T, DE-U, GB-A, FR-A); DialogPRO; IEEE/IEEEXplore; Google/Google Scholar; IP.com; Search Terms Used: malware, malicious, phish, pharm, attacker, control command, resource locator, URL, DNS, domain, predict, probability, suspect, likelihood, redirect)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2006/0075500 A1 (BERTMAN, J. et al.) April 6, 2006; figure 1; paragraphs [0009], [0017]-[0019], [0021]-[0026], [0028]-[0030], [0036], [0046], [0055], [0056]	1, 2, 4/1, 4/2, 13-18, 19/17, 19/18, 41, 42, 44/41, 44/42, 51
-- Y		3, 4/3, 10, 43, 44/43, 52
Y	US 2008/0209552 A1 (WILLIAMS, J. et al.) August 28, 2008; paragraphs [0029]-[0032], [0045]	3, 4/3, 10, 43, 44/43, 52
A	US 2011/0030060 A1 (KEJRIWAL, N.) February 3, 2011; entire document	1-3, 4/1-4/3, 10, 13-18, 19/17, 19/18, 41-43, 44/41-44/43, 51, 52

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 September 2013 (05.09.2013)	13 SEP 2013
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

-***-Continued from Box No. III - Observations where unity of invention is lacking -***-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fee must be paid.

Group I: Claims 1-4, 10, 13-19, 41-44, 51, and 52 are directed toward a computerized method and a system for automatically processing a plurality of files, comprising: receiving a universal resource locator (URL), the URL identifying a malware artifact file at a command and control node; receiving the malware artifact file; determining whether the malware artifact file is at least partially obfuscated; decoding the malware artifact file to reverse at least one obfuscating transformation if the malware artifact file is at least partially obfuscated; storing the malware artifact file in an electronic data store; and analyzing the decoded malware artifact file.

Group II: Claims 21-24 are directed toward a computerized method for automatically processing a plurality of files, comprising: receiving user input comprising a first universal resource locator, the first universal resource locator identifying a first malware artifact file at a first intermediate network node; retrieving the first malware artifact file stored at the first intermediate network node; determining whether the first malware artifact file is at least partially obfuscated; decoding the first malware artifact file to reverse at least one obfuscating transformation; and storing the decoded first malware artifact in an electronic data store.

Group III: Claims 33-34, 53, and 54 are directed toward a computer method for automatically processing a plurality of files, comprising: receiving a URL identifying a first intermediate network node; receiving a fetch schedule identifying a monitoring schedule for attempting to access a malware artifact file at the first intermediate network node; and repeatedly accessing the first intermediate network node according to the fetch schedule and attempting to retrieve the malware artifact file.

Group IV: Claims 35-37 and 55-57 are directed toward an electronic framework for automatically processing a plurality of files, comprising: an electronic data store configured to store an identification of a first URL; a task manager configured to: insert into a queue a request for a fetch attempt for the target source object at the first network node; and automatically execute the fetch attempt for the target resource object at the first network node; and a processor module configured to: store a fetched target resource object; decode the fetched target resource object to reverse an obfuscation transformation; and analyze the decoded target resource object to determine if a new URL is identified in the fetched target resource object.

Group V: Claims 38-40 and 58-60 are directed toward a method comprising: receiving user input comprising a URL, the URL identifying a malware artifact at an intermediate network node; decoding the malware artifact file to reverse at least one obfuscating transformation; interpreting the decoded malware artifact to determine whether it contains: a) a command to malware to perform a function; or b) a command to a victim computing device to perform a function; or c) a data file containing exfiltrated data; or d) a universal resource locator; storing the decoded malware artifact file in an electronic data store; and executing a next instruction based on the interpretation of the decoded mal ware artifact file.

The common technical feature shared by Groups I, II, III, IV and V is a system for automatically processing a plurality of files, comprising: receiving a URL, the URL identifying a malware artifact file; retrieving the malware artifact file; determining whether the malware artifact is at least partially obfuscated; decoding the malware artifact file to reverse at least one obfuscating transformation if the malware artifact is at least partially obfuscated; and storing the malware artifact file in an electronic data store.

However, this common feature is previously disclosed by US 2006/0075500 A1 (Bertman). Bertman discloses a system for automatically processing a plurality of files (content; Abstract), comprising: receiving a URL (receive an initial URL; paragraph [0009]), the URL identifying a malware artifact file (URL is associated with a web site having malware content (artifact file); paragraph [0009]); retrieving the malware artifact file (download (retrieve) Web pages (artifact file) containing malware content; paragraph [0009]); determining whether the malware artifact is at least partially obfuscated (identify any obfuscation techniques used to hide malware or pointers to malware; paragraphs [0009] and [0029]); decoding the malware artifact file to reverse at least one obfuscating transformation if the malware artifact is at least partially obfuscated (parsing (decoding) downloaded Web pages to undo (reverse) the tricks used by malware programmers to hide their malware; paragraphs [0029]-[0030]); and storing the malware artifact file in an electronic data store (downloaded page is stored in the database 105; paragraph [0040]).

Since the common technical feature is previously disclosed by the Bertman reference, this common feature is not special and so Groups I, II, III, IV and V lack unity.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US13/34660

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-9, 11, 12, 20, 25-32, 45-50
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Group I: Claims 1-4, 10, 13-19, 41-44, 51, and 52; Group II: Claims 21-24; Group III: Claims 33-34, 53, and 54; Group IV: Claims 35-37 and 55-57; Group V: Claims 38-40 and 58-60

-***-Please see Supplemental Page-***-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4, 10, 13-19, 41-44, 51, and 52

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.