US007487544B2

(12) **United States Patent**
**Schultz et al.**

(10) **Patent No.: US 7,487,544 B2**
(45) **Date of Patent: Feb. 3, 2009**

(54) **SYSTEM AND METHODS FOR DETECTION OF NEW MALICIOUS EXECUTABLES**

(75) Inventors: **Matthew G. Schultz**, Ithaca, NY (US); **Eleazar Eskin**, Santa Monica, CA (US); **Erez Zadok**, Middle Island, NY (US); **Manasi Bhattacharyya**, Flushing, NY (US); **Stolfo J. Salvatore**, Ridgewood, NJ (US)

(73) Assignee: **The Trustees of Columbia University in the city of New York**, NY, NY (US)

( * ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1122 days.

(21) Appl. No.: **10/208,432**

(22) Filed: **Jul. 30, 2002**

(65) **Prior Publication Data**

US 2003/0065926 A1 Apr. 3, 2003

**Related U.S. Application Data**

(60) Provisional application No. 60/308,622, filed on Jul. 30, 2001, provisional application No. 60/308,623, filed on Jul. 30, 2001.

(51) **Int. Cl.**
***G06F 21/00*** (2006.01)

(52) **U.S. Cl.** ..... **726/24**; 713/188

(58) **Field of Classification Search** ..... 726/13, 726/22–25; 713/156, 188; 709/206, 207, 709/225

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,452,442 A 9/1995 Kephart et al.
5,485,575 A * 1/1996 Chess et al. ..... 714/38
5,675,711 A 10/1997 Kephart et al.
5,765,170 A * 6/1998 Morikawa ..... 707/200
5,832,208 A * 11/1998 Chen et al. ..... 726/24
6,016,546 A * 1/2000 Kephart et al. ..... 726/24
6,161,130 A * 12/2000 Horvitz et al. ..... 709/206
6,275,850 B1 * 8/2001 Beyda et al. ..... 709/206

(Continued)

OTHER PUBLICATIONS

Jeffrey O. Kephart and William C. Arnold, "Automatic Extraction of Computer Virus Signatures," *4th Virus Bulletin International Conference*, pp. 178-184, 1994.

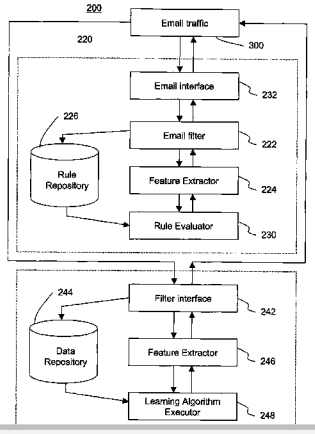
(Continued)

*Primary Examiner*—Gilberto Barron, Jr.
*Assistant Examiner*—Abdulhakim Nobahar
(74) *Attorney, Agent, or Firm*—Baker Botts LLP

(57) **ABSTRACT**

A system and methods for detecting malicious executable attachments at an email processing application of a computer system using data mining techniques. The email processing application may be located at the server or at the client or host. The executable attachments are filtered from said email, and byte sequence features are extracted from the executable attachment. The executable attachments are classified by comparing the byte sequence feature of the executable attachment to a classification rule set derived from byte sequence features of a data set of known executables having a predetermined class in a set of classes, e.g., malicious or benign. The system is also able to classify executable attachments as borderline when the difference between the probability that the executable is malicious and the probability that the executable is benign are within a predetermined threshold. The system can notify the user when the number of borderline attachments exceeds the threshold in order to refine the classification rule set.

**43 Claims, 7 Drawing Sheets**



U.S. PATENT DOCUMENTS

| | | | | |
|---|---|---|---|---|
| 6,598,076 | B1 * | 7/2003 | Chang et al. | 709/206 |
| 6,778,995 | B1 | 8/2004 | Gallivan | |
| 6,820,081 | B1 | 11/2004 | Kawai et al. | |
| 6,826,609 | B1 * | 11/2004 | Smith et al. | 709/225 |
| 6,888,548 | B1 | 5/2005 | Gallivan | |
| 6,978,274 | B1 | 12/2005 | Gallivan et al. | |
| 7,035,876 | B2 | 4/2006 | Kawai et al. | |
| 7,080,076 | B1 | 7/2006 | Williamson et al. | |
| 2002/0059383 | A1 * | 5/2002 | Katsuda | 709/206 |
| 2002/0065892 | A1 * | 5/2002 | Malik | 709/206 |

OTHER PUBLICATIONS

R Kohavi, "A study of cross-validation and boot-strap for accuracy estimation and model selection," *IJCAI*, 1995.

Ronald L. Rivest. "The MD5 Message Digest Algorithm." published as Internet RFC 1321, Apr. 1992. http://www.freesoft.org/CIE/RFC/1321/.

Stephen R. van den Berg and Philip Guenther, "Procmail." online publication, 2001. http://www.procmail.org.

Steve R. White, Morton Swimmer, Edward J. Pring, William C. Arnold, David M. Chess, and John F. Morar, "Anatomy of a Commercial-Grade Immune System," IBM Research White Paper, 1999.

Eleazar Eskin et al. "System and Methods for Intrusion Detection with Dynamic Window Sizes," filed Jul. 30, 2000, U.S. Appl. No. 10/208,402.

U.S. Appl. No. 10/352,343, filed Jan. 27, 2003 claiming priority to P34981 (070050.1936) U.S. Appl. No. 60/351,857, filed Jan. 25, 2001, entitled "Behavior Based Anomaly Detection For Host-Based Systems For Detection Of Intrusion In Computer Systems," of Frank Apap, Andrew Honig, Shlomo Hershkop, Eleazar Eskin and Salvatore J. Stolfo.

U.S. Appl. No. 10/352,342, filed Jan. 27, 2003 claiming priority to U.S. Appl. No. 60/351,913, filed Jan. 25, 2002, entitled "Data Warehouse Architecture For Adaptive Model Generation Capability In Systems For Detecting Intrusion In Computer Systems," of Salvatore J. Stolfo, Eleazar Eskin, Matthew Miller, Juxin Zhang and Zhi-Da Zhong.

U.S. Appl. No. 10/327,811, filed Dec. 19, 2002 claiming priority to U.S. Appl. No. 60/342,872, filed Dec. 20, 2001, entitled "System And Methods for Detecting A Denial-Of-Service Attack On A Computer System" of Salvatore J. Stolfo, Shlomo Hershkop, Rahul Bhan, Suhail Mohiuddin and Eleazar Eskin.

U.S. Appl. No. 10/320,259, filed Dec. 16, 2002 claiming priority to U.S. Appl. No. 60/328,682, filed Oct. 11, 2001 and U.S. Appl. No. 60/352,894, filed Jan. 29, 2002, entitled "Methods of Unsupervised Anomaly Detection Using A Geometric Framework" of Eleazar Eskin, Salvatore J. Stolfo and Leonid Portnoy.

U.S. Appl. No. 10/269,718, filed Oct. 11, 2002 claiming priority to U.S. Appl. No. 60/328,682, filed Oct. 11, 2001 and U.S. Appl. No. 60/340,198, filed Dec. 14, 2001, entitled "Methods For Cost-Sensitive Modeling For Intrusion Detection" of Dec. 14, 2001, entitled "Methods For Cost-Sensitive Modeling For Intrusion Detection" of Salvatore J. Stolfo, Wenke Lee, Wei Fan and Matthew Miller.

U.S. Appl. No. 10/269,694, filed Oct. 11, 2002 claiming priority to U.S. Appl. No. 60/328,682, filed Oct. 11, 2001 and U.S. Appl. No. 60/339,952, filed Dec. 13, 2001, entitled "System And Methods For Anomaly Detection And Adaptive Learning" of Wei Fan, Salvatore J. Stolfo.

U.S. Appl. No. 10/222,632, filed Aug. 16, 2002 claiming priority to U.S. Appl. No. 60/312,703, filed Aug. 16, 2001 and U.S. Appl. No. 60/340,197, filed Dec. 14, 2001, entitled "System And Methods For Detecting Malicious Email Transmission" of Salvatore J. Stolfo, Eleazar Eskin, Manasi Bhattacharyya and Matthew G. Schultz.

* cited by examiner

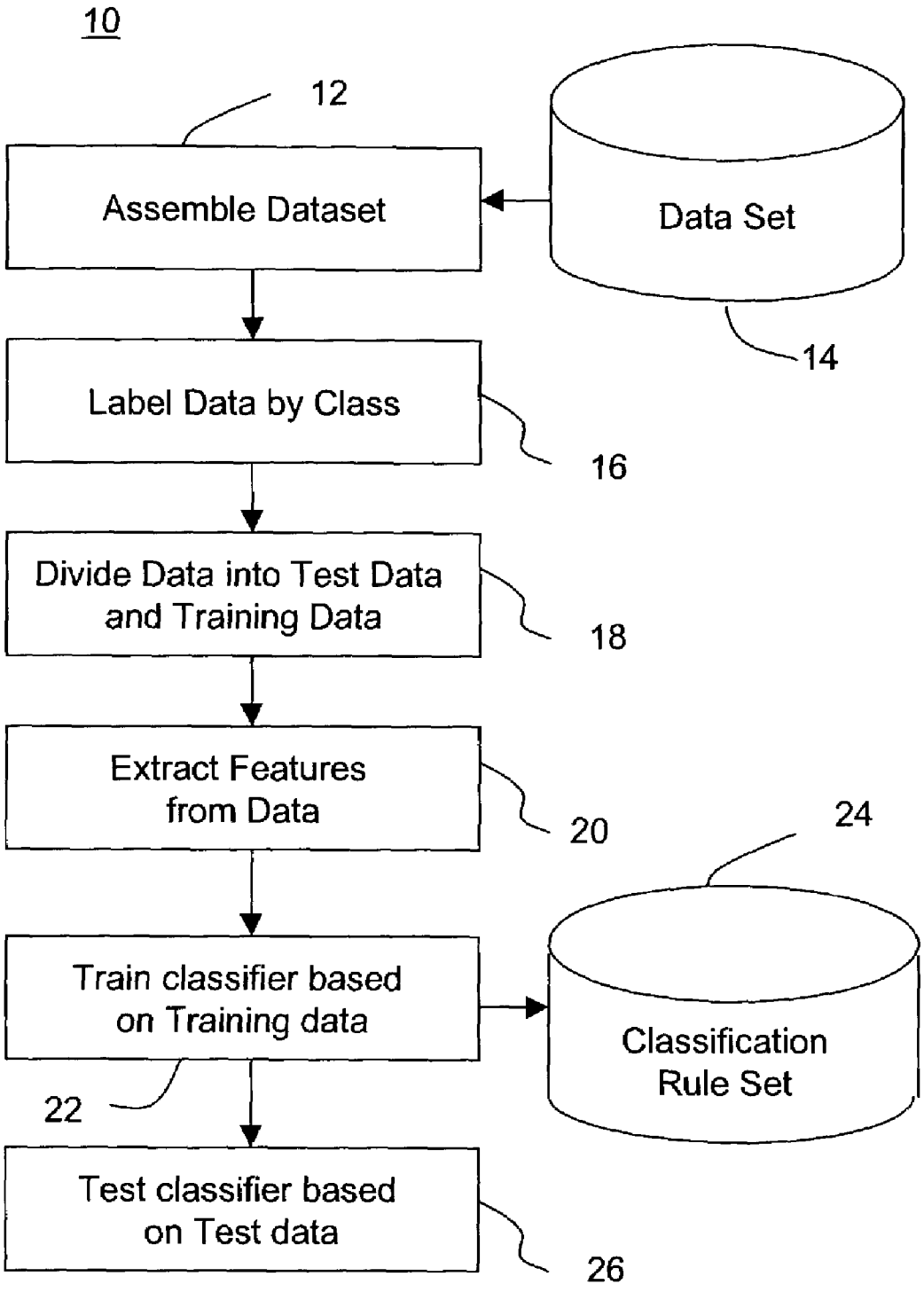



FIG. 1

```
646e 776f 2e73 0a0d 0024 0000 0000 0000
454e 3c0f 026c 0009 0000 0000 0302 0004
0400 2800 3924 0001 0000 0004 0004 0006
000c 0040 0060 021e 0238 0244 02f5 0000
0001 0004 0000 0802 0032 1304 0000 030a
```

FIG. 2

$$\neg advapi32 \wedge avicap32 \wedge \cdots \wedge winmm \wedge \neg wsock32$$

FIG. 3

$$advapi32.AdjustTokenPrivileges()$$
$$\wedge\ advapi32.GetFileSecurityA() \wedge \cdots$$
$$\wedge\ wsock32.recv() \wedge wsock32.send()$$

FIG. 4

$$advapi32 = 2 \wedge avicap32 = 10 \wedge \cdots$$
$$\wedge \quad winmm = 8 \wedge wsock32 = 2$$

FIG. 5

$$
\begin{aligned}
malicious &:= \neg user32.EndDialog() \wedge \\
&\quad kernel32.EnumCalendarInfoA() \\
malicious &:= \neg user32.LoadIconA() \wedge \\
&\quad \neg kernel32.GetTempPathA() \wedge \neg advapi32. \\
malicious &:= shell32.ExtractAssociatedIconA() \\
malicious &:= msvbvm. \\
Benign &:- otherwise
\end{aligned}
$$

FIG. 6

$$
\begin{aligned}
P(\text{" windows"}\backslash benign) &= 45/47 \\
P(\text{" windows"}\backslash malicious) &= 2/47 \\
P(\text{"*.COM"}\backslash benign) &= 1/12 \\
P(\text{"*.COM"}\backslash malicious) &= 11/12
\end{aligned}
$$

FIG. 7

DOCKET ALARM

# Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

## Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

## API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

**LAW FIRMS**
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

**FINANCIAL INSTITUTIONS**
Litigation and bankruptcy checks for companies and debtors.

**E-DISCOVERY AND LEGAL VENDORS**
Sync your system to PACER to automate legal marketing.