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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 August 2022 (25.08.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/177633 A1

(51) International Patent Classification:
H04L 29/06 (2006.01) *G06F 21/00* (2013.01)
G06N 20/00 (2019.01)

(21) International Application Number:
PCT/US2021/063944

(22) International Filing Date:
17 December 2021 (17.12.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
17/178,386 18 February 2021 (18.02.2021) US

(71) Applicant: **SECUREWORKS CORP.** [US/US]; 2711 Centerville Road, Suite 400, Wilmington, Delaware 19808 (US).

(72) Inventors: **BARGNESI, Nicholas**; 263 Peeptoad Road, Scituate, Rhode Island 02857 (US). **VERSTEEG, Steve**; 63 Beatty Steet, Ivanhoe, Victoria 3079 (AU). **SUN, Li**; 5 Bolton Street, Beaumaris, Victoria, 3193 (AU).

(74) Agent: **SUDDERTH, D., Scott**; Womble Bond Dickinson (US) LLP, P.O. Box 7037, Atlanta, GA 30357-0037 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG).

(54) Title: SYSTEMS AND METHODS FOR AUTOMATED THREAT DETECTION

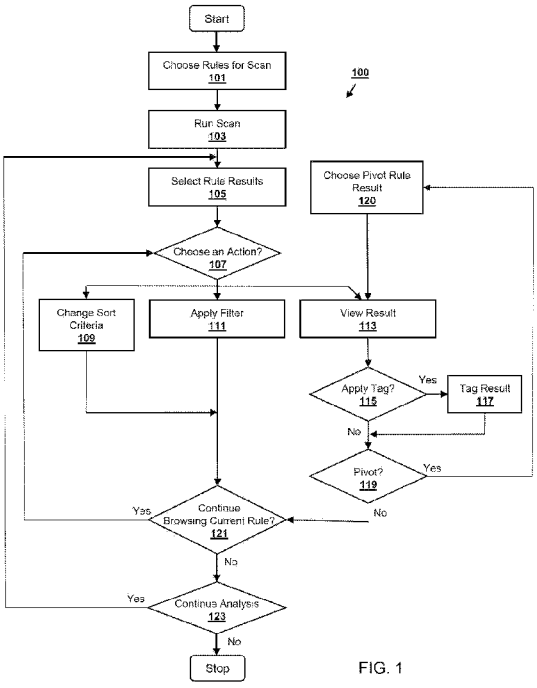

FIG. 1

(57) Abstract: Systems and methods for dynamically training a threat detection system include monitoring security analyst workflow data from security analysts analyzing scans of security logs. The workflow data includes rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or actions associated with a pivot by security analysts. A tagging classifier is then trained based on tags assigned to scan results. A review classifier is trained based on scan results previously reviewed by security analysts. A filter and ranking method is trained based on filters and rankings applied to the scan results. An automated threat hunting playbook is generated including the tagging classifier, the review classifier, and the filter and ranking method. The automated threat hunting playbook generates one or more scripts to automatically analyze incoming security data.

WO 2022/177633 A1

[Continued on next page]

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))*

SYSTEMS AND METHODS FOR AUTOMATED THREAT DETECTION

Background

[0001] Identifying attack patterns and suspicious activity from malicious actors is an important aspect of computer network security. Highly trained security analysts spend many hours reviewing scans of security logs in order to identify and investigate potential compromise events. The sheer volume of security logs that could potentially be reviewed can overwhelm the resources of security analysts. As malicious actors become more sophisticated and adjust their attack strategies, it becomes more and more difficult to identify attack patterns and suspicious activity, and the limited resources of trained analysts are increasingly spread thin.

[0002] Accordingly, it can be seen that a need exists for systems and methods that can automatically detect potential compromise events and suspicious activities, as well as organize and prioritize scan results for more efficient review by security analysts.

[0003] The present disclosure is directed to the foregoing and other related, and unrelated, problems or issues in the relevant art.

Summary

[0004] Briefly described, according to various aspects, the present disclosure includes systems and methods for dynamically training a security threat detection system. According to one aspect, a method for dynamically training a security threat detection system is disclosed. The method includes monitoring security analyst workflow data from one or more security analysts analyzing scans of security logs. The workflow data includes one or more rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof. The method also includes training a tagging classifier based on the tags assigned to rule results from the workflow data; training a review classifier based on the rule results selected for further analysis; training a filter and ranking method based on filters and rankings applied to rule results from one or more security analysts; generating an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method; and generating one or more scripts for automatically analyzing incoming security data using the automated threat hunting playbook. In one embodiment, the method also includes training a pivot sequence model based on actions executed by one or more security analysts,

and the automated threat hunting playbook also includes the pivot sequence model. In one embodiment, the tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts. In one embodiment, the scripts for automatically analyzing incoming security data generate a number of tags, and each tag is an indicator of compromise within a computer network. In one embodiment, the method also includes receiving a tag update from one or more security analysts; and dynamically updating the tagging classifier based on the tag update. In one embodiment, the scripts for automatically analyzing incoming security data generate a selection of results for review. In one embodiment, the method also includes receiving analyst feedback regarding the selection of results for review; and dynamically updating the review classifier based on the analyst feedback regarding the selection of results for review. In one embodiment, the scripts for automatically analyzing incoming security data generate a selection of prioritized results. In one embodiment, the method also includes receiving analyst feedback regarding the selection of prioritized results; and dynamically updating the filter and ranking method based on the analyst feedback. In one embodiment, the scripts for automatically analyzing incoming security data generate one or more pivot chains, wherein a pivot chain is a series of rule results that trace a potential attack. In one embodiment, the method also includes receiving pivot chain feedback from one or more security analysts; and dynamically updating the pivot sequence model based on the pivot chain feedback.

[0005] According to another aspect, a dynamically trained threat detection system, includes a computing system for monitoring and storing security analyst workflow data from one or more security analysts analyzing scans of security logs. The workflow data includes rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof. The system also includes a tagging classifier trained based on the tags assigned to rule results from the workflow data; a review classifier trained based on the rule results selected for further analysis; a filter and ranking method trained based on the filters and rankings applied to rule results from one or more security analysts; and an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method. The automated threat hunting playbook is configured to generate one or more scripts for automatically analyzing incoming security data. In one embodiment, the system also includes a pivot sequence model based on actions executed by one or more security analysts, wherein a pivot chain is a series of rule results that trace a potential attack, and the automated threat hunting

playbook also includes the pivot sequence model. In one embodiment, the tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts. In one embodiment, the pivot sequence model generates pivot chains when applied to raw scan data from a security log. In one embodiment, the tagging classifier generates tags when applied to raw scan data from a security log, each tag being an indicator of compromise within a computer network. In one embodiment, generates a selection of results for review when applied to raw scan data from a security log. In one embodiment, the filter and ranking method generate a selection of prioritized results when applied to raw scan data from a security log.

[0006] According to another aspect, a system for dynamically training a security threat detection system includes one or more processors and at least one memory having stored instruction. When executed, the instructions cause the system to monitor and record workflow data from one or more security analysts analyzing security logs within a computer network. The workflow data includes rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof. The instructions also cause the system to train a tagging classifier based on the tags applied to rule results from the workflow data; train a review classifier based on the rule results selected for further analysis by one or more security analysts; train a filter and ranking method based on the filters and rankings applied to rule results from security analysts; and train a pivot sequence model based on actions executed by one or more security analysts. The tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts. The instructions also cause the system to generate an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method. In one embodiment, the instructions also cause the system to analyze incoming security data using the one or more scripts to generate a number of tags, a selection of results for review, a selection of prioritized results, and one or more pivot chains. Each tag is an indicator of compromise within a computer network, and a pivot chain is a series of rule results that trace a potential attack. The instructions also cause the system to receive analyst feedback regarding the tags, the selection of results for review, the selection of prioritized results, and the one or more pivot chains; and dynamically update the tagging classifier, review classifier, filter and ranking method, and pivot sequence model based on the analyst feedback.

[0007] For example, in some aspects, a process for threat detection and training of an automated threat detection system can include a first step in which a scan is run. For example, for an organization with a thousand computers, the detection rules will run on all thousand computers and collect evidence for selected events, actions, indicators, etc. These logs then can be loaded for the analyst to review. In some embodiments, a log can include a raw log, aggregated information from multiple logs, information about a file, or any other information collected from a host, system, or service. However, given the substantial volume thereof, potentially millions of logs may be available for the analyst to review. The analyst then needs to decide which logs they are actually going to review, because they can't review all of the incoming logs. The system can initiate a scan applying a selected collection of rules on one or more computing devices to collect security logs that are relevant. Such detection rules may initiate processing of the security logs by a host computer or server, and then resend the results. For example, some of the rules can be configured to scan for known types of malware, other detection rules may just collect all of the logon events for a certain time or location, etc. The analyst can create a search query or initiate the scan search to evaluate if the logs can be triaged and/or sorted to narrow them down to certain hits or particular scans that should be evaluated closer to see if there has been an attack.

[0008] In one embodiment, the analyst starts off by selecting a series of rules to run on the investigation, or the organization that's being investigated for compromise, initiating or generating the scan. A scan might be a collection of 10 rules, 20 rules, 100 rules, 1000 rules, etc. Then the scan can run on those computers of a selected set or networked group, and return the results to a central location for the analyst to review. Thus, the analyst has a collection of rule results. In some embodiments, the results can be organized by which rule they came from. For example, the analyst can be shown 20 rules results, and once accepted, one or more particular rule numbers can be selected, e.g. a rule to collect, service, and review all the various services that have been installed on a host. As the analyst is presented with more results, they may filter the results. So, they can either apply a filter, for example, they may look for a specific host, specific user name, an internal, external IP address, or they may order it based on like criteria, such as ordering the results by rarity or by whether there are any annotations. Some of the rules may return all the files that are available on a host, which files can be scanned by a collection of virus scanners, resulting in annotations such as whether there is any known malware. Alternatively, the analyst may focus on one or more of the rule results in more detail.

[0009] These rule results can be organized to enable the analyst to click on one of the rule results for more detail to make a determination of whether there is evidence of compromise and tag the result or not. If they find that it's evidence of compromise, they'll give it a malware tag. If they find that it's just a commercial software, they can apply an appropriate tag for that. In addition, if the analyst does find evidence of compromise, they may execute a pivot. For example, if the rule result relates to a particular computer, they may get more information about that computer. In other embodiments, the analyst may pivot by time, by the user name, toward a different host that was or is connected to the relevant computer, or toward other relatable results.

[0010] The system can record the analyst's actions, also described as workflow data, and these actions can be organized into playbooks or collections analyzed actions. A playbook can be made up of multiple rules, including any rules that are selected for a scan, as well as one or more pivot sequences. The playbook can include a tagging classifier trained based on the analyst's actions observed when they are tagging. In one embodiment, a developed tagging classifier can be trained with a group of labeled or tagged results using a supervised learning technique. Starting with on a binary classification that goes between malicious or not, the system can look at all the rule results that were tagged as malicious and all the rule results that were tagged as clean, and then provide the tagged rule results as the supervised learning input. If given an unknown rule result, or the rule result hasn't been tagged, the classifier can tag it automatically.

[0011] In some embodiments, the playbook can filter down the number of results an analyst looks at. This filtering can be ordered by what the analyst is most likely to click on. For example, the system can have a list of all the rule results that the analyst clicked on, and all the ones that they didn't really use, as well as a record of what filters they used. For example, if an analyst received a million results, the system can monitor what filters were applied, whether any type any regular expression was used in filtering, whether any rankings were applied, what results did the analyst pivot on, whether a pivot was by host name or username, and what pivot sequence was followed.

[0012] To execute the playbook, a scan can be run to get a set of rule results. The rule results are all the logs that come from the scan. Those rule results can be fed through a tagging classifier that exists in the playbook to give a collection of tags. The resultant tags provide a list of malicious results and a list of evidence of compromise. Another output of the playbook is the review classifier that can predict what rule results are more likely to get looked at by an analyst. For example, from a million scan results, the review classifier might provide a thousand results that are more likely to be looked at. Still another output of the playbook can be a ranking of the criteria on which the analyst

might want to review the results, e.g. based on what filters and ranking methods were applied. For example, a million results can be ranked in order of predicted importance to an analyst. Still another output of the playbook can be pivot sequence, which when run on the results can provide a list of pivot chains. Thus, from the playbooks, some of the outputs or the scans are fully automated, such as processing rule results through a selected tag classifier for a collection of evidence of compromise.

[0013] In some embodiments, each of the operations described herein can be configured to run in parallel to generate a series of results that can be used to create scripts based on the learned behaviors/actions from the observed analysts, which scripts can be used and applied to future incoming security information/data/logs as part of a threat hunting playbook or set of rules to detect security threats at earlier stages.

[0014] Various objects, features and advantages of the present disclosure will become apparent to those skilled in the art upon a review of the following detail description, when taken in conjunction with the accompanying drawings.

Brief Description of the Drawings

[0015] It will be appreciated that for simplicity and clarity of illustration, elements illustrated in the Figures are not necessarily drawn to scale. For example, the dimensions of some elements may be exaggerated relative to other elements. Embodiments incorporating teachings of the present disclosure are shown and described with respect to the drawings herein, in which:

[0016] FIG. 1 is a schematic illustration of a workflow for identifying attack patterns or suspicious activity, according to one aspect of the present disclosure.

[0017] FIG. 2 is a block diagram of a method for training a threat detection system, according to one aspect of the present disclosure.

[0018] FIG. 3 is a block diagram of a method of applying a threat detection system to raw scan results, according to one aspect of the present disclosure.

[0019] FIG. 4 is a schematic illustration of a workflow for dynamically training a threat detection system, according to one aspect of the present disclosure.

[0020] FIG. 5 is a schematic illustration of a networked system of information handling systems, according to one example of the present disclosure.

Detailed Description

[0021] The following description in combination with the figures is provided to assist in understanding the teachings disclosed herein. The description is focused on specific implementations and embodiments of the teachings, and is provided to assist in describing the teachings. This focus should not be interpreted as a limitation on the scope or applicability of the teachings.

[0022] In one embodiment, the present disclosure relates to a system for automated threat detection that learns threat hunt playbooks, or threat detection sequences, based on analyzing the behavior of human security experts during threat hunts. In some embodiments, the system can include a supervised machine learning (ML) algorithm that is trained on the behavior and workflows of trained security analysts, and which can automatically discover malicious attacks. Such threat detection systems can also increase the efficiency of analysts during threat hunts by allowing them to concentrate their time on information most likely to be associated with malicious activity, or by discovering new attack techniques or new suspicious behavior.

[0023] As used herein, a host describes one or more computers in an organization or group that is scanned as part of a threat hunt.

[0024] As used herein, a threat hunt describes a process for examination of forensic information to search for evidence of malicious attacks.

[0025] As used herein, a rule describes a detection rule that including logic or programming that can be stored and/or executed on a host machine, and which is configured to look for or identify patterns associated with potentially malicious behavior.

[0026] As used herein, a rule result describes information collected from a host machine when a detection rule finds at least one indicator of a potentially malicious activity.

[0027] As used herein, a scan describes the execution of a series of detection rules on hosts to collect information about potentially malicious activities as part of a threat hunt.

[0028] As used herein, a false positive describes a rule result from a detection rule that is determined to not be associated with malicious activity.

[0029] As used herein, a true positive describes a rule result from a detection rule that is verified to be associated with an actual instance of malicious activity.

[0030] Threat hunting is the process by which one or more security analysts review available evidence, such as security logs and the outputs of threat detection rules, to determine if, where, and how a security breach has occurred. Threat hunting often requires many hours of attention from a highly skilled security analyst. According to an embodiment of the present disclosure, an automated threat detection system can learn from a security expert's threat hunting behavior in order to automatically detect potential compromise events, and also to generate prioritized lists of potential suspicious activity in order to streamline the threat detection process.

[0031] FIG. 1 shows a schematic illustration of a workflow 100 for identifying attack patterns or suspicious activity, according to one aspect of the present disclosure. According to this example, a threat hunt starts with choosing the rules for a scan 101, and running a scan 103. The detection rules can be selected to run on a host computing system and can include, for example, tactical rules designed to detect specific kinds of attacks, and strategic rules designed to pick up generalized patterns. The scan results can be uploaded to a database, in some embodiments, and then they can be evaluated by analysts conducting a threat hunt.

[0032] Not every rule result is a true positive for a malicious attack. Strategic rules may pick up a wide range of events or file artifacts, designed to find traces left behind by a malicious attack, but may also include a large number of results relating to legitimate use. For large organizations, a scan may return tens of millions of results. Searching through these results to look for attacks generally requires both time and skill.

[0033] The workflow 100 may continue at operation 105 with selecting rule results. For each rule result, there is a vector of columns, in some embodiments. The columns differ for different types of rules. For example, a rule that returns installed services may include columns such as "Service Name", "Service type, "Service Start Type", "Service Path", "Service Account", "Host Name", "IP Address", "Detection Timestamp", "Installation Timestamp", "Log", "Record Number", "Category" and "Event Description". Rule results may be annotated with additional information, such as how many times similar results have been seen before (on how many hosts and in how many organisations). If the result is from a file, it may be annotated with virus scan information. These annotations are added as additional columns appended to the rule results.

[0034] In an embodiment, threat hunting can be done on a platform designed to search, view and tag rule results. Once rule results are selected, a number of different actions can be chosen at operation 107. A non-exclusive list of the types of actions that analysts can perform on the threat

hunting platform include, for example, changing sort criteria 109, applying a filter 111, and viewing a result 113. Selecting a rule can include selecting a rule to browse the returned scan results. Applying a filter can include applying a regular expression to one or more columns. Changing sort criteria can include sorting and/or ordering results based on a particular column. Viewing rule results can include viewing a result in more detail to see the original log information returned by the detection rule.

[0035] In an embodiment, upon reviewing a result at operation 113, an analyst can decide at operation 115 to apply a tag to a result to record whether that result indicates a truly malicious attack or a false positive. If a tag is to be applied, the workflow 100 can include tagging the result 117. In some embodiments, tagging schemes can be binary (e.g. "malicious" or "clean"), or list of categories, (e.g. "commercial", "potentially unwanted", "adware", "potentially malicious", "commodity malware", "APT malware"). If no tag is to be applied, or after a tag is successfully applied, the workflow 100 may continue with pivoting 119 to other results. For example, a threat hunt may pivot at operation 119 and continue to choose a pivot rule result 120, and then view rule results again at 113. In one embodiment, a threat hunt may pivot to results from other rules that have the same or similar value for a certain attribute (such as host, user, or source IP address).

[0036] The workflow 100 can continue with deciding at operation 121 whether to continue browsing the current rule results. If yes, the workflow 100 can return to choose a different action at 107. If no, the workflow 100 can continue with deciding at operation 123 whether to continue analysis. If yes, the workflow 100 can return to select different or new rule results at operation 105. If no, the workflow 100 ends.

[0037] In one embodiment, this threat hunting workflow 100 can be performed by one or more trained security analysts, and can generate workflow data that can be used to train an automated threat detection system. Workflow data can include, for example, a listing of the security log scan results selected for further analysis. Workflow data can also include types of filters, rankings, sort criteria, or tags, applied to different types of results. Workflow data can also include one or more security log items or actions associated with a pivot by one or more security analysts, as well as what pivot sequence was executed. For example, if a rule result that may be related to a compromise event relates to a particular computer, the pivot may involve getting more information about that computer. A pivot may also include recovering additional information about the time when the suspicious activity occurred, or a username or different host connected to the particular computer.

[0038] FIG. 2 is a block diagram of a method for training a threat detection system, according to one aspect of the present disclosure. In one embodiment, the threat detection system can execute an automated threat hunting playbook 201. This threat hunting playbook 201 can include a pivot sequence model 205 and rules 203, as well as a number of trained elements including a tagging classifier 207, a review classifier 209, a filter 211, and a ranking method 213, in some embodiments. This playbook 201 can be constructed or taught based on workflow data gathered by monitoring the actions of trained security analysts, such as the actions described above in reference to FIG. 1.

[0039] According to the embodiment shown in FIG. 2, an automated threat hunting playbook 201 includes rule sets 203, tagging classifiers 207, review classifiers 209, filters 211, ranking methods 213, and a pivot sequence model 205.

[0040] The rule sets 203 can include a collection of detection rules to be run in a scan. In one embodiment, clustering algorithms may be applied to get the rule sets that security analysts select on threat hunts. In this way, sets of rules can be automatically maintained based on the workflow data discussed above.

[0041] The tagging classifiers 207 can include classifiers that analyze the rule results and automatically tag whether the result is malicious (a true positive) or a false positive. Each rule can have one associated tagging classifier. In one embodiment, for each rule, a set of tagging results can be used to train an automated tagging classifier 207. This may be achieve using, for example, a supervised ML algorithm. The rule column values, plus the annotation columns, can be used as the input feature vector for such an algorithm. The tags can be used as the class label. Example classification algorithms that may be used include, but are not limited to: instance based classifiers, decision trees, decision forests, support vector machines, or neural networks. To ensure that classifiers are not over-fitted, they may be trained on data obtained from multiple organizations, in some embodiments. In some embodiments, the tags can be indicators of compromise within a computer network.

[0042] The review classifiers 207 can include classifiers that filter the rule results to recommend a subset of results that should be manually reviewed by a security analyst. Each rule can have one associated review classifier. In some embodiments, for each rule a review classifier 209 can be trained to automatically identify results that are of high interest to analysts and which should be reviewed manually. Since analysts only have time to view a small subset of the entire set of rule results, the recorded logs of which results were viewed by analysts can be used as the training set for

the review classifier 209. In one embodiment, the feature vector for the training set can be the rule column data and annotations. The class label can be binary, denoting whether the result was viewed or not viewed by an analyst. Any class based supervised learning algorithm can be used to train the review classifier 209, such as those algorithms listed above in reference to training the tagging classifier 207.

[0043] The filters 211 can include filters that are automatically applied to the rule results to reduce the set of results to be reviewed. A rule may have zero or multiple filters, in various embodiments. Multiple filters can be applied in an AND combination (where only results that satisfy all filters are retained) or an OR combination (where results that satisfy any filter are retained). The ranking methods 213 can include the order in which rule results should be viewed in order from results of highest priority to lowest priority. Each rule can have one associated ranking method. In one embodiment, the filters and ranking methods can be associated with rules using the recorded filters and sorting methods and columns from the workflow data of security analysts as a training set.

[0044] The pivot sequence model 205 can include an automated sequence of results to be viewed tracing a possible attack. In one embodiment, a pivot sequence model can be constructed and trained using the actions of analysts in threat hunting as a training set. A pivot sequence can include a series of actions taken in investigating a potential compromise event, as recorded in the workflow data discussed above.

[0045] FIG. 3 is a block diagram of a method of applying a threat detection system to raw scan results, according to one aspect of the present disclosure. In this embodiment, the raw scan results 301 can be applied to a trained tagging classifier 303, review classier 305, filter 307 and ranking method 309, and pivot sequence model 311. The tagging classifier 303, review classier 305, filter 307 and ranking method 309, and pivot sequence model 311 can be trained as discussed above in reference to FIG. 2 using one or more ML algorithms, and based on the recorded actions and workflow data of trained security analysts. The outputs of the threat detection system can include, for example, automated tags 313, results for manual review 315, prioritized results 317, and pivot chains 319.

[0046] In one embodiment, executing the automated threat detection system can include inputting the raw results 301 (i.e. results from the scan 103 described in reference to FIG. 1) into the trained tagging classifier 303 to generate automated tags 313. These automated tags 313 can be automatically generated according to an assigned maliciousness level using the tagging classifier 303.

In one embodiment, the tagging classifier might review 10 million results and might automatically generate 20 malicious tags. In such an example, this provides a limited number of potential compromise events from the 10 million results reviewed.

[0047] In one embodiment, executing the automated threat detection system can also include inputting the raw results 301 into a review classifier 305 to generate a selection of results for review 315. These results can be a subset of the full scan results, and can include a series of results that should be manually reviewed by a security analyst, as these results are more likely to be related to a malicious attack. In one embodiment, the review classifier may automatically analyze 10 million results and provide a list of 10 thousand results that are selected for further manual review. In such an example, these 10 thousand results are the ones that the review classifier 305 predicts are the most important for an analyst to review.

[0048] In one embodiment, executing the automated threat detection system can also include inputting the raw results 301 into a filter 307 and ranking method 309 to generate a list of prioritized results 317. In some embodiments, this list of prioritized results 317 can include the full set of scan results that are filtered and ranked, should a security analyst wish to inspect the results manually.

[0049] In one embodiment, executing the automated threat detection system can also include inputting the raw results 301 into a pivot sequence model 311 to generate pivot chains. The pivot chains 319 can include a series of rule results that trace a potential attack. In some embodiments, these results may be used for automated resolutions, or be reviewed manually by a security analyst.

[0050] As discussed herein, the recorded actions of analysts can be used as a basis for creating automated threat hunt scripts and training the various models and classifiers. In some embodiments, each of these processes described in FIG. 3 can be running in parallel to generate the tags 313, results for manual review 315, prioritized results 317, and pivot chains 319.

[0051] FIG. 4 is a schematic illustration of a workflow 400 for dynamically training a threat detection system, according to one aspect of the present disclosure. In some embodiments, the accuracy of the automated threat detection system can be continuously improved by retraining the models according to how analysts review and correct the results.

[0052] In one embodiment, the workflow 400 begins with monitoring workflow data 401. As discussed above, workflow data can include, for example, a listing of the security log scan results selected for further analysis. Workflow data can also include types of filters, rankings, sort criteria,

or tags applied to different types of results. Workflow data can also include one or more security log items or pivot chains associated with a pivot by security analysts, as well as actions executed by security analysts. In some embodiments, a pivot chain includes a series of rule results that trace a potential attack.

[0053] This workflow data can then be used to train the ML models 403 discussed above, including the tagging classifier, review classifier, filter and ranking methods, and pivot sequence model. The training of these ML models is discussed in more detail in reference to FIG. 2.

[0054] Once the models have been trained at 403, the automated threat hunting playbook can be generated at 405, including the tagging classifier, the review classifier, and the filter and ranking method. The playbook can also include a pivot sequence model, as discussed above.

[0055] The workflow 400 continues at 407 with generating one or more scripts for automatically analyzing incoming security data using the automated threat hunting playbook. Once the playbook has been generated, the trained models within the playbook can be applied to raw scan data at 409 to generate tags, a selection of results for review, prioritized results, and/or pivot chains. These operations, and the generation of the outputs of the ML models, are discussed in detail in reference to FIG. 3.

[0056] The workflow 400 continues at 411 with receiving analyst feedback on the outputs of the ML models generated at 409. In some embodiments, the analyst feedback can include edits or changes to automated tags generated by the tagging classifier. For example, when an analyst corrects a tag, or finds other results to assign tags, this information can be used to further train the tagging classifier.

[0057] The analyst feedback can also include a list of the actual results reviewed by the analyst from the results for review generated by the review classifier. For example, if the review classifier generated a focused or curated list of two thousand results for review, and the analyst only reviewed a subset of 800 results, this information can be stored for further training of the review classifier.

[0058] In some embodiments, the analyst feedback can include a list of the results reviewed by the analyst from the prioritized results generated by the filter and ranking methods. For example, if the filter and ranking methods organized results in a particular ranking, or applied a particular filter,

and the analyst reviewed the results in a different order than the automatically-generated ranking, this information can be stored for further training of the filter and ranking methods.

[0059] In some embodiments, the analyst feedback can include an alternative pivot sequence executed by the analyst that is different from the pivot chain generated by the pivot sequence model. If the analyst executed a different pivot than the one recommended by the pivot sequence model, this can be used as an additional input for further training of the pivot sequence model.

[0060] The workflow can continue by applying the analyst feedback to the training of the ML methods at 403 in order to dynamically update the models and increase the accuracy of the automated threat detection system.

[0061] For purposes of this disclosure, an information handling system 80 (FIG. 5) may include any instrumentality or aggregate of instrumentalities operable to compute, calculate, determine, classify, process, transmit, receive, retrieve, originate, switch, store, display, communicate, manifest, detect, record, reproduce, handle, or utilize any form of information, intelligence, or data for business, scientific, control, or other purposes. For example, an information handling system may be a personal computer (e.g., desktop or laptop), tablet computer, mobile device (e.g., personal digital assistant (PDA) or smart phone), server (e.g., blade server or rack server), a network storage device, or any other suitable device and may vary in size, shape, performance, functionality, and price. The information handling system may include random access memory (RAM), one or more processing resources such as a central processing unit (CPU) or hardware or software control logic, read only memory (ROM), and/or other types of nonvolatile memory. Additional components of the information handling system may include one or more disk drives, one or more network ports for communicating with external devices as well as various input and output (I/O) devices, such as a keyboard, a mouse, a touchscreen and/or a video display. The information handling system also may include one or more buses operable to transmit communications between the various hardware components.

[0062] As shown in FIG. 5, in some embodiments, the clients 12 can manage or otherwise include one or more networked system 82 of information handling systems/devices 80 or other communicable systems/devices. A network 84 may provide data communications among the information handling systems/devices 80, which can include workstations, personal computers, smart cellular telephones, personal digital assistants, laptop computers, servers, and other suitable devices. The network 84 can include a private or public network, such as a local area network, or other suitable

network(s), such as the Internet or another wide area network, virtual personal network, peer to peer filing sharing system, and/or other suitable communication lines, or combinations thereof. FIG. 5 also indicates that the linked or networked information handling systems 80 may include one or more monitoring devices 86 communicatively coupled to the network 84. The monitoring device(s) 86 can be managed by a managed security service provider (MSSP).

[0063] In one embodiment, the monitoring device(s) 86 may include a server or sequence analyzer or other client suitable computing device that has a processor and a memory or other suitable storage. The memory can include a random access memory (RAM), read only memory (ROM), and/or other non-transitory computer readable medium. The monitoring device(s) 86 further typically will be operable to store and execute computer readable instructions to continuously monitor, in real-time, activity at each networked system, for example, activity of the information handling systems 80 connected to network 84. The monitoring device(s) 86 can ingest or aggregate information or data logs related to activities of the information handling systems 80 and can provide these ingested/aggregate data logs or information or data related thereto to by the automated threat detection system described herein. In addition, or in the alternative, the automated threat detection system described herein can include a data center 88, such as a data center 88 management by an MSSP, with a plurality of networked information handling systems 80, e.g., including one or more servers 90 with at least one memory 92 and one or more processors 94 for receiving information or data logs related to activities of the information handling systems 80 of system 82. These information/data logs can be a part of the raw logs 14 provided to the automated threat detection system described herein.

[0064] One or more components of the systems described herein can be resident on or accessed by the devices 80, the server(s) 90, or other devices or information handling systems in communication therewith. One or more processors of the device 80 of the one or more processors 94 can process or execute instructions, workflows, etc., stored in at least one memory (e.g., a memory of the devices 90 or memory 92) to facilitate performance of various processes, functions, etc. of the automated threat detection system described herein.

[0065] The foregoing description generally illustrates and describes various embodiments of the present disclosure. It will, however, be understood by those skilled in the art that various changes and modifications can be made to the above-discussed construction of the present disclosure without departing from the spirit and scope of the disclosure as disclosed herein, and that it is intended that all matter contained in the above description or shown in the accompanying drawings shall be

interpreted as being illustrative, and not to be taken in a limiting sense. Furthermore, the scope of the present disclosure shall be construed to cover various modifications, combinations, additions, alterations, etc., above and to the above-described embodiments, which shall be considered to be within the scope of the present disclosure. Accordingly, various features and characteristics of the present disclosure as discussed herein may be selectively interchanged and applied to other illustrated and non-illustrated embodiments of the disclosure, and numerous variations, modifications, and additions further can be made thereto without departing from the spirit and scope of the present invention as set forth in the appended claims.

Claims

What is claimed is:

1. A method for dynamically training a security threat detection system, comprising:

monitoring security analyst workflow data from one or more security analysts analyzing scans of security logs, wherein the workflow data includes one or more rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof;

training a tagging classifier based on the tags assigned to rule results from the workflow data;

training a review classifier based on the rule results selected for further analysis;

training a filter and ranking method based on filters and rankings applied to rule results from one or more security analysts;

generating an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method; and

generating one or more scripts for automatically analyzing incoming security data using the automated threat hunting playbook.

2. The method of claim 1, further comprising:

training a pivot sequence model based on actions executed by the one or more security analysts, and the automated threat hunting playbook also includes the pivot sequence model.

3. The method of claim 2, wherein the tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts.

4. The method of claim 3, wherein the one or more scripts for automatically analyzing incoming security data generate a plurality of tags, wherein each tag is an indicator of compromise within a computer network.

5. The method of claim 4, further comprising:

receiving a tag update from one or more security analysts; and

dynamically updating the tagging classifier based on the tag update.

6. The method of claim 3, wherein the one or more scripts for automatically analyzing incoming security data generate a selection of results for review.

7. The method of claim 6, further comprising:

receiving analyst feedback regarding the selection of results for review; and

dynamically updating the review classifier based on the analyst feedback regarding the selection of results for review.

8. The method of claim 3, wherein the one or more scripts for automatically analyzing incoming security data generate a selection of prioritized results.

9. The method of claim 8, further comprising:

receiving analyst feedback regarding the selection of prioritized results; and

dynamically updating the filter and ranking method based on the analyst feedback regarding the selection of prioritized results.

10. The method of claim 3, wherein the one or more scripts for automatically analyzing incoming security data generate one or more pivot chains, wherein a pivot chain is a series of rule results that trace a potential attack.

11. The method of claim 10, further comprising:

receiving pivot chain feedback from one or more security analysts; and

dynamically updating the pivot sequence model based on the pivot chain feedback.

12. A dynamically trained threat detection system, comprising:

one or more computing systems configured to monitor and store security analyst workflow data from one or more security analysts analyzing scans of security logs, wherein the workflow data includes rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results, filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof;

a tagging classifier trained based on the tags assigned to rule results from the workflow data;

a review classifier trained based on the rule results selected for further analysis;

a filter and ranking method trained based on the filters and rankings applied to rule results from one or more security analysts; and

an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method, wherein the automated threat hunting playbook is configured to generate one or more scripts for automatically analyzing incoming security data.

13. The system of claim 12, further comprising:

a pivot sequence model based on actions executed by one or more security analysts, wherein a pivot chain is a series of rule results that trace a potential attack, and the automated threat hunting playbook also includes the pivot sequence model.

14. The system of claim 13, wherein the tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts.

15. The system of claim 13, wherein the pivot sequence model, when applied to raw scan data from a security log, is configured to generate one or more pivot chains.

16. The system of claim 12, wherein the tagging classifier, when applied to raw scan data from a security log, is configured to generate a plurality of tags, each tag being an indicator of compromise within a computer network.

17. The system of claim 12, wherein the review classifier, when applied to raw scan data from a security log, is configured to generate a selection of results for review.

18. The system of claim 12, wherein the filter and ranking method, when applied to raw scan data from a security log, is configured to generate a selection of prioritized results.

19. A system for dynamically training a security threat detection system, comprising:

one or more processors and at least one memory having stored therein instructions that when executed by the one or more processors, cause the system to:

monitor and record workflow data from one or more security analysts analyzing security logs within a computer network, wherein the workflow data includes rules applied to security log scan results, rule results selected for further analysis, tags applied to rule results,

filters applied to rule results, rankings applied to rule results, or one or more actions associated with a pivot by the one or more security analysts, and/or combinations thereof;

train a tagging classifier based on the tags applied to rule results from the workflow data;

train a review classifier based on the rule results selected for further analysis by one or more security analysts;

train a filter and ranking method based on the filters and rankings applied to rule results from one or more security analysts;

train a pivot sequence model based on actions executed by one or more security analysts, wherein the tagging classifier, review classifier, filter and ranking method, and pivot sequence are each supervised machine learning models trained based on the workflow data of one or more security analysts; and

generate an automated threat hunting playbook including the tagging classifier, the review classifier, and the filter and ranking method.

20. The system of claim 19, wherein the instructions further cause the system to:

analyze incoming security data using the one or more scripts to generate a plurality of tags, a selection of results for review, a selection of prioritized results, and one or more pivot chains, wherein each tag is an indicator of compromise within a computer network, and a pivot chain is a series of rule results that trace a potential attack;

receive analyst feedback regarding the plurality of tags, the selection of results for review, the selection of prioritized results, and the one or more pivot chains; and

dynamically update the tagging classifier, review classifier, filter and ranking method, and pivot sequence model based on the analyst feedback.

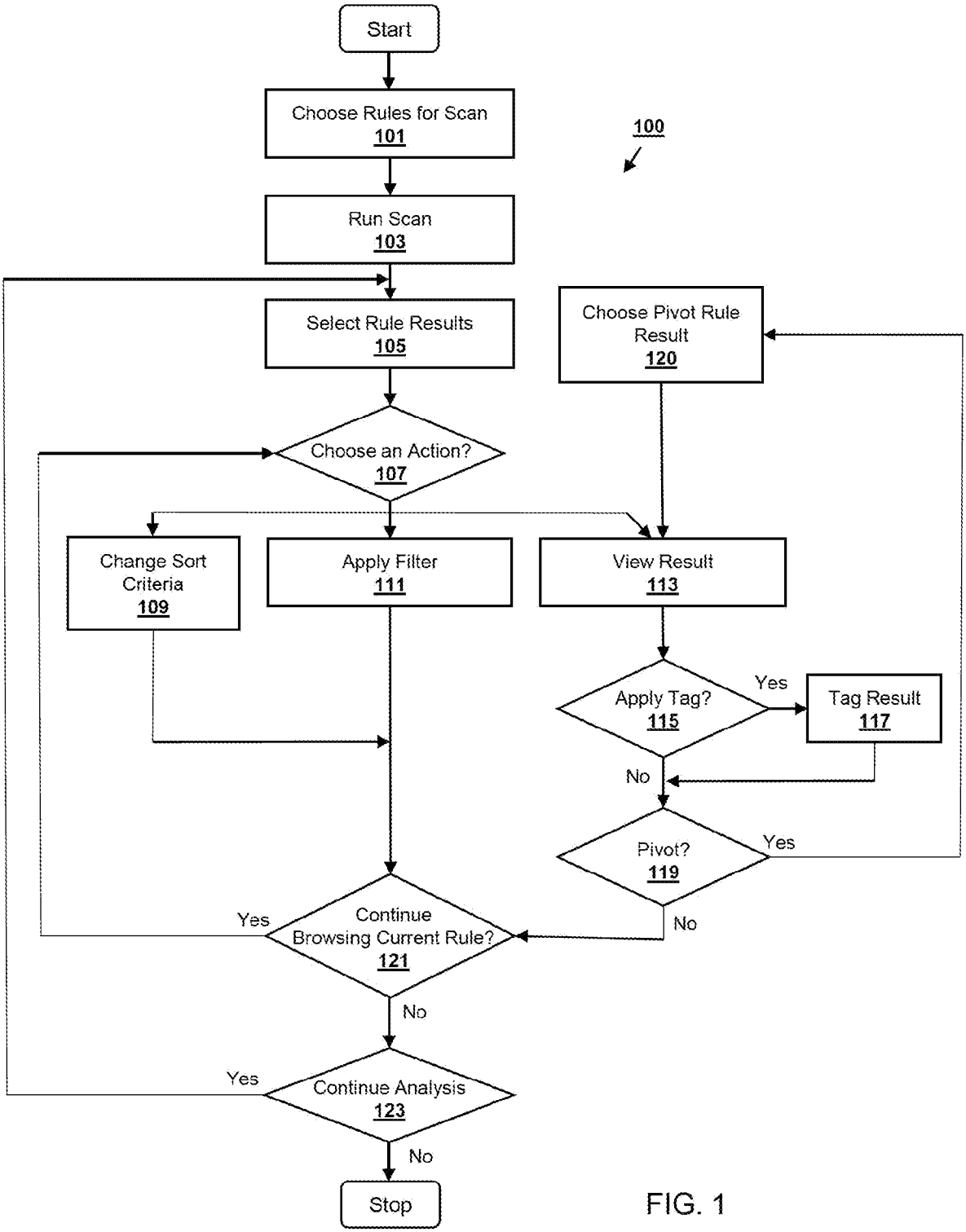
Start
Choose Rules for Scan
101
100
Run Scan
103
Select Rule Results
105
Choose Pivot Rule Result
120
Choose an Action?
107
Change Sort Criteria
109
Apply Filter
111
View Result
113
Apply Tag?
115
Yes
Tag Result
117
No
Pivot?
119
Yes
No
Continue Browsing Current Rule?
121
Yes
No
Continue Analysis
123
Yes
No
Stop

FIG. 1

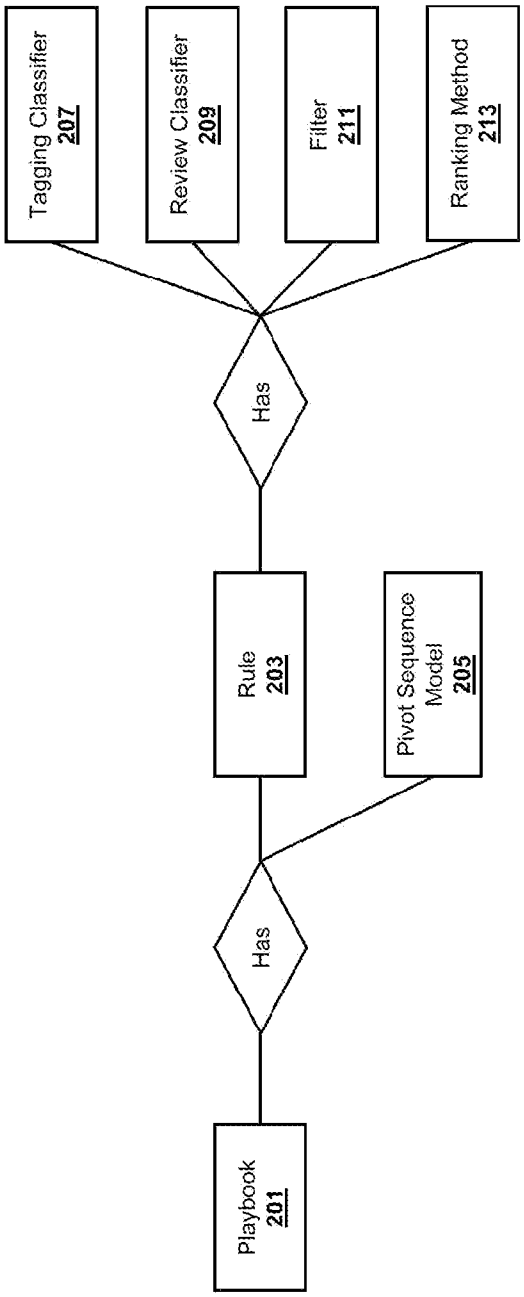
Tagging Classifier
207
Review Classifier
209
Filter
211
Ranking Method
213
Has
Rule
203
Pivot Sequence
Model
205
Has
Playbook
201

FIG. 2

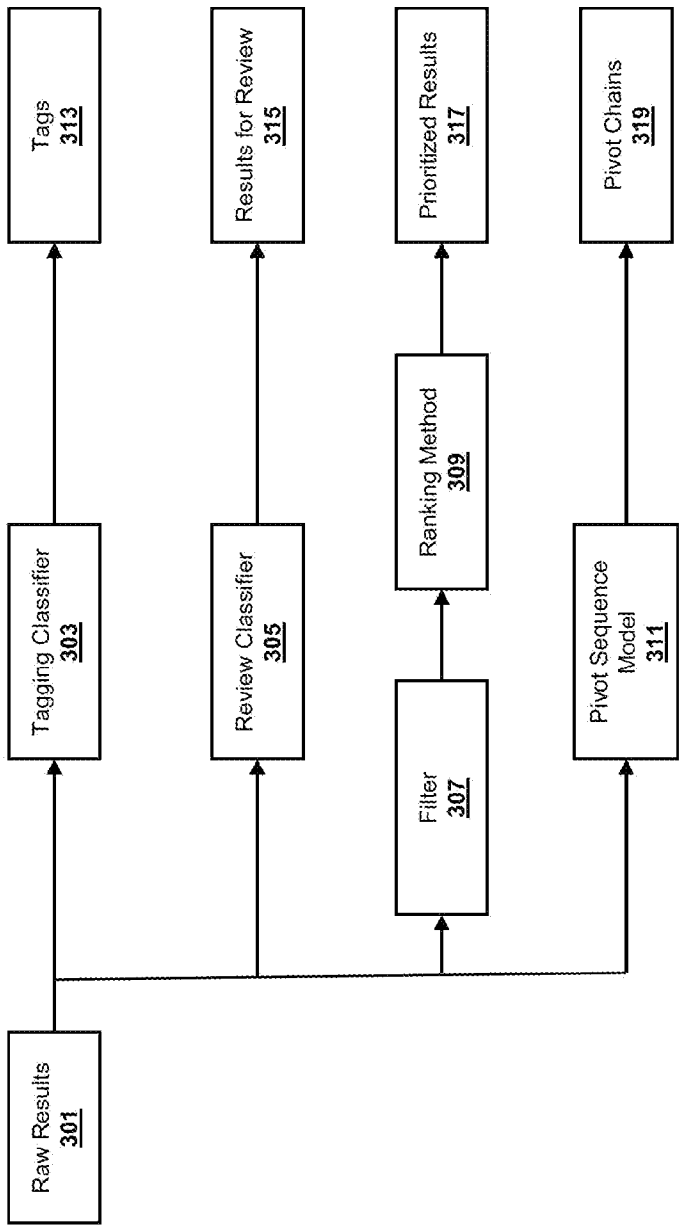
Raw Results
301
Tagging Classifier
303
Review Classifier
305
Filter
307
Ranking Method
309
Pivot Sequence Model
311
Tags
313
Results for Review
315
Prioritized Results
317
Pivot Chains
319

FIG. 3

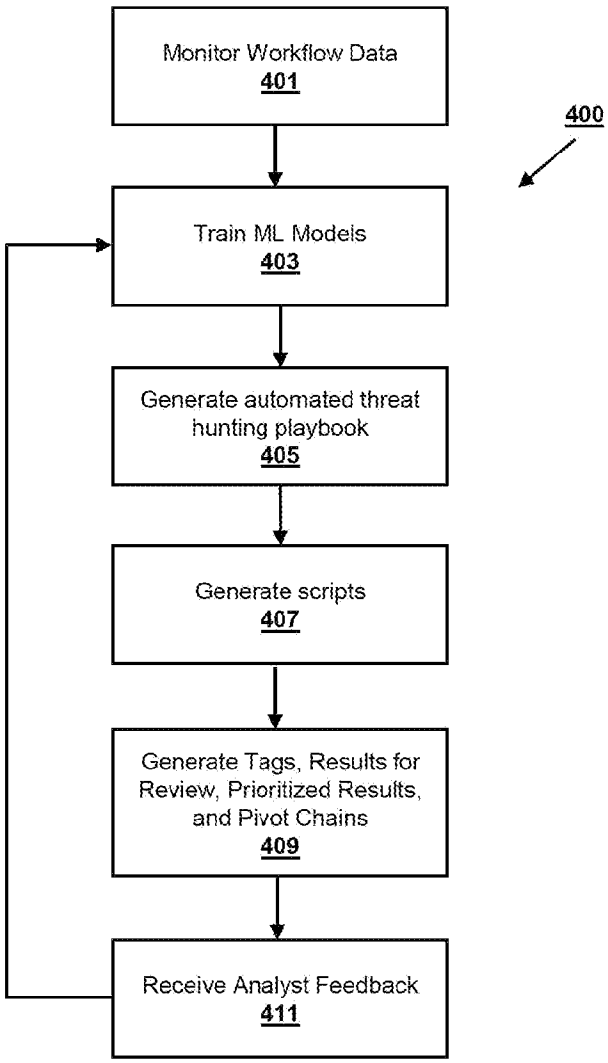
Monitor Workflow Data
401
400
Train ML Models
403
Generate automated threat hunting playbook
405
Generate scripts
407
Generate Tags, Results for Review, Prioritized Results, and Pivot Chains
409
Receive Analyst Feedback
411

FIG. 4

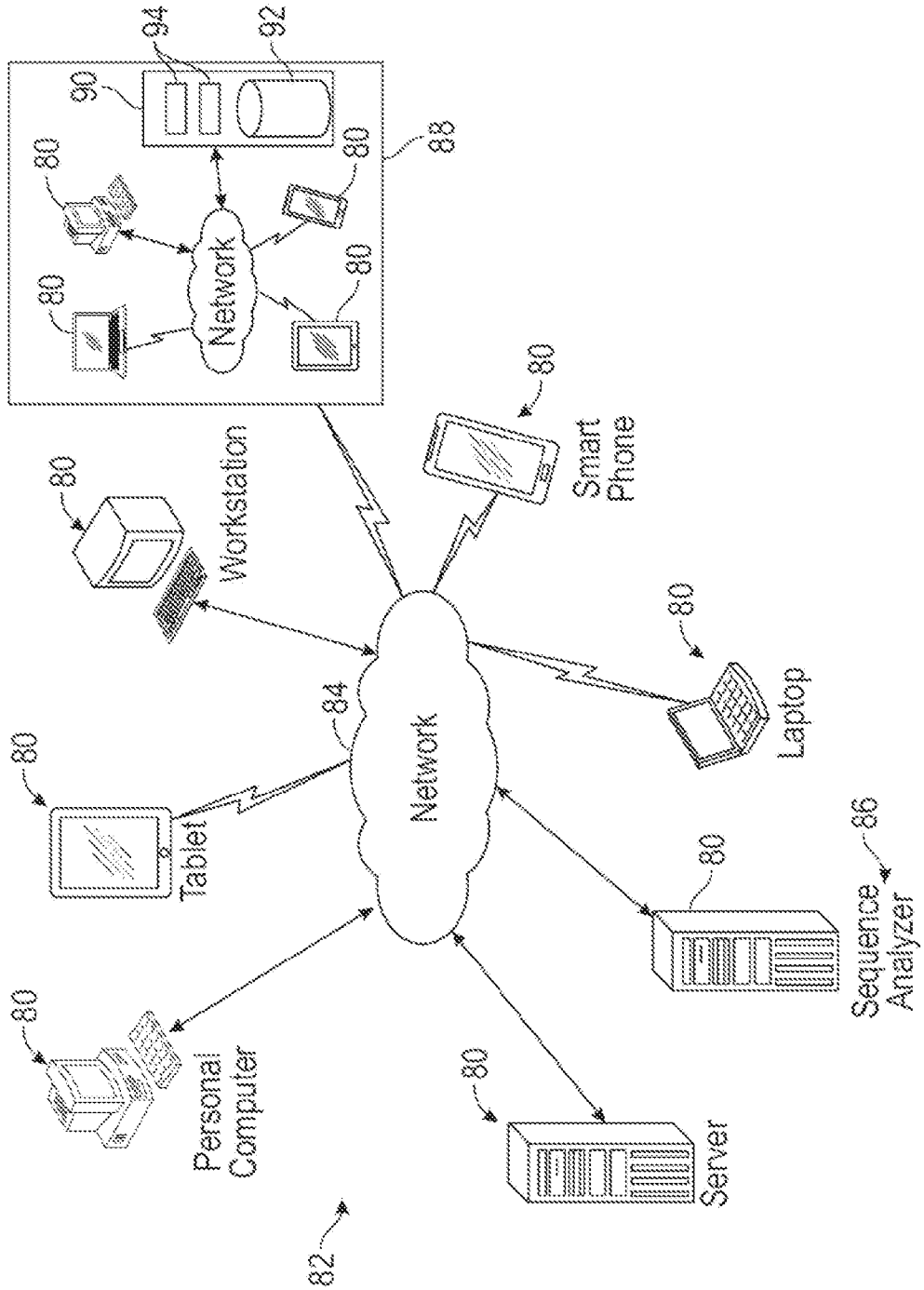
94
92
90
88
80
Network
Workstation
Smart Phone
Laptop
Tablet
Personal Computer
Server
Sequence Analyzer
86
84
82

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 21/63944

A. CLASSIFICATION OF SUBJECT MATTER

IPC - H04L 29/06; G06N 20/00; G06F 21/00 (2022.01)

CPC - H04L 63/1416; G06F 21/564; G06N 20/00

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)
See Search History document

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2017/0063905 A1 (Splunk Inc.) 02 March 2017 (02.03.2017) Entire document, especially para [0151], [0171], [0220], [0236], [0270], [0271], [0279], [0298].	1-20

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"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
08 February 2022 (08.02.2022)	MAR 16 2022
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. 571-273-8300	Authorized officer Kari Rodriquez Telephone No. PCT Helpdesk: 571-272-4300