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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 December 2024 (26.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/263172 A1

(51) International Patent Classification:
G06F 21/55 (2013.01)

(21) International Application Number:
PCT/US2023/026162

(22) International Filing Date:
23 June 2023 (23.06.2023)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **VISA INTERNATIONAL SERVICE ASSOCIATION** [US/US]; P. O. Box 8999, San Francisco, California 94128 (US).

(72) Inventors: **BRAHMKSTRI, Krupa**; 3655 Colegrove Street, Apt 20, San Mateo, California 94403 (US). **SINGH, Birender**; 945 Foxchase Dr, Apt 415, San Jose, California 95123 (US).

(74) Agent: **BUCKLEY, Timothy**; P.O. Box 158, Austin, Texas 78767 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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,

(54) Title: MACHINE LEARNING BASED UNAUTHORIZED SERVICE ACCOUNT ACCESS DETECTION SYSTEM AND METHOD THEREFOR

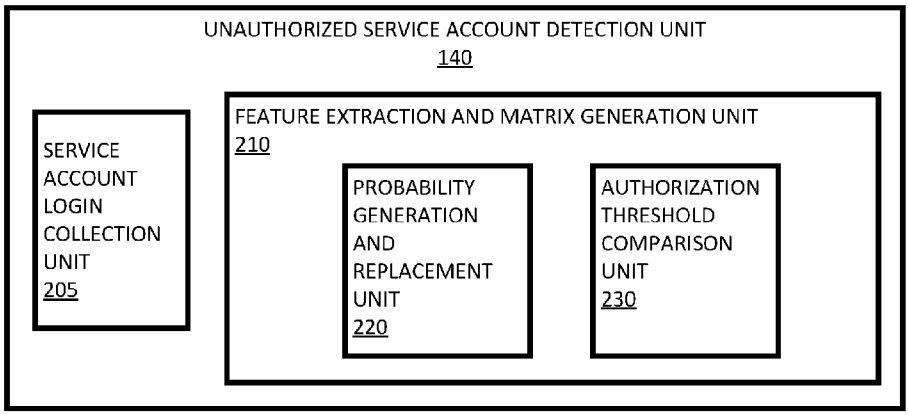

FIG. 2

(57) Abstract: In some embodiments, a system, includes a processor; and a non-transitory computer readable medium coupled to the processor, the non-transitory computer readable medium including code that: receives a service account login history matrix associated with users of a service account, the service account login history matrix being transformed to a service account login history probability matrix; generates a source-machine service-account matrix from the service account login history probability matrix; generates a service-account destination matrix from the service account login history probability matrix; and utilizes the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

WO 2024/263172 A1

[Continued on next page]

DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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).

Published:

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

MACHINE LEARNING BASED UNAUTHORIZED SERVICE ACCOUNT ACCESS DETECTION SYSTEM AND METHOD THEREFOR

BACKGROUND

[0001] The background description provided herein is for the purpose of generally presenting the context of the disclosure. Work of the presently named inventor(s), to the extent it is described in this background section, as well as aspects of the description that may not otherwise qualify as prior art at the time of filing, are neither expressly nor impliedly admitted as prior art against the present disclosure.

[0002] Service accounts are highly privileged accounts created to perform automation of various services and applications. The automation of service accounts may be achieved by running a programming script to maintain the security of various downstream applications. However, use of the programming scripts may create a technical dependency on the configuration of the programming scripts and, if, for example, passwords are changed, backward compatibility issues may arise resulting in authentication failures and unnecessary system backup. As a result, service accounts are highly vulnerable to hackers and other nefarious actors. Therefore, a need exists to continuously monitor users that attempt to login to service accounts to prevent unauthorized access to the service accounts.

SUMMARY

[0003] The Summary provided herein is utilized to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. The Summary is not intended to identify key features or essential features of the claimed subject matter, nor is the Summary intended to be used to limit the scope of the claimed subject matter.

[0004] Methods, systems, and computer readable mediums that store code for performing methods are described herein. In one aspect, a computer-implemented method includes receiving a service account login history associated with users of a service account, the service account login history being transformed to a service account login history probability matrix; generating a source-machine service-account matrix from the service account login history probability matrix; generating a service-account destination matrix from the service account login history probability matrix; and utilizing the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

[0005] Further features and advantages of embodiments, as well as the structure and operation of various embodiments, are described in detail below with reference to the accompanying drawings. It is noted that the methods and systems are not limited to the specific embodiments described herein. Such embodiments are presented herein for illustrative purposes only. Additional embodiments will be apparent to persons skilled in the relevant art(s) based on the teachings contained herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] FIG. 1 illustrates a block diagram of a system in accordance with some embodiments.

[0007] FIG. 2 illustrates a block diagram of an unauthorized service account detection unit of the system in FIG. 1 in accordance with some embodiments.

[0008] FIG. 3A illustrates an example three-dimensional coordinate system utilized to describe an example service account login history probability matrix in accordance with some embodiments.

[0009] FIG. 3B illustrates example 2D plane representations of an example service account login history probability matrix in accordance with some embodiments.

[0010] FIG. 3C illustrates a source-machine service-account matrix of the service account login history probability matrix of FIG. 3B with example probability values in accordance with some embodiments.

[0011] FIG. 3D illustrates a service-account destination matrix of the service account login history probability matrix of FIG. 3B with example probability values in accordance with some embodiments.

[0012] FIG. 4 is a flow diagram illustrating an unauthorized service account detection method in accordance with some embodiments.

[0013] FIG. 5 illustrates an example utilization of the unauthorized service account detection method of FIG. 4 in accordance with some embodiments.

DETAILED DESCRIPTION

[0014] FIG. 1 illustrates a block diagram of an exemplary system 100 for implementing embodiments consistent with the present disclosure. In some embodiments, system 100 includes an input/output (IO) interface 101, processor/s 102, a storage interface 104, a network interface 103, and memory 105. In some embodiments, system 100 may be, for example, a server, such as an authentication server or source machine utilized to implement the embodiments described herein. In some embodiments, memory 105 may include an operating system (OS) 107, processes 120, and an unauthorized service account detection unit 140. In some nonlimiting embodiments or aspects, the unauthorized service account detection unit 140 is configured to implement an unauthorized service account detection method configured to detect an unauthorized service account login as described further herein.

[0015] In some embodiments, processors 102 may comprise at least one data processor for executing program components for dynamic resource allocation at run time. The processors 102 may include specialized processing units such as integrated system (bus) controllers, memory management control units, floating point units, graphics processing units, digital signal processing units, etc. In some embodiments, the processors 102 may be disposed in communication with one or more input/output (I/O) devices (not shown) via an I/O interface 101. The I/O interface 101 may employ communication protocols/methods such as, without limitation, audio, analog, digital, monoaural, RCA, stereo, IEEE-1394, serial bus, universal serial bus (USB), infrared, PS/2, BNC, coaxial, component, composite, digital visual interface (DVI), high-definition multimedia interface (HDMi), RF antennas, S-Video, VGA, IEEE 802.1 n /b/g/n/x, Bluetooth®, cellular (e.g., code-division multiple access (CDMA), high-speed packet access (HSPA+), global system for mobile communications (GSM), long-term evolution (LTE), WiMax®, or the like), etc.

[0016] In some embodiments, using the I/O interface 101, the system 100 may communicate with one or more I/O devices. For example, an input device (not shown) may be an antenna, keyboard, mouse, joystick, (infrared) remote control, camera, card reader, fax machine, dongle, biometric reader, microphone, touch screen, touchpad, trackball, stylus, scanner, storage device, transceiver, video device/source, etc. An output device (not shown) may be a printer, fax machine, video display (e.g., cathode ray tube (CRT), liquid crystal display (LCD), light-emitting diode (LED), plasma, Plasma display panel (PDP), Organic light-emitting diode display (OLED) or the like), audio speaker, etc.

[0017] In some embodiments, the processors 102 may be disposed in communication with a communication network or other type of network via a network interface 103. The network

interface 103 may communicate with the communication network. The network may include the system, a source machine, and a destination machine. In some embodiments, the source machine may be a computer system configured to allow a user of the source machine to login or attempt to login to a service account. In some embodiments, the destination machine may be a computer system that the user of the source machine is attempting to login into via system 100. In some embodiments, each source machine and destination machine may be identified by a source machine identification (ID) or destination machine ID, respectively. In some embodiments, the service account may be identified by a service account ID. The network interface 103 may employ connection protocols including, without limitation, direct connect, Ethernet (e.g., twisted pair 10/100/1000 Base T), transmission control protocol/Internet protocol (TCP/IP), token ring, IEEE 802.11a/b/g/n/x, etc. The communication network may include, without limitation, a direct interconnection, e-commerce network, a peer to peer (P2P) network, local area network (LAN), wide area network (WAN), wireless network (e.g., using Wireless Application Protocol), the internet, Wi-Fi®, etc. Using the network interface 103 and the communication network, the system 100 may communicate with the one or more service operators, service machines, and/or destination machines.

[0018] In some non-limiting embodiments or aspects, the processors 102 may be disposed in communication with a memory 105 (e.g., RAM, ROM, etc.) via a storage interface 104. In some embodiments, the storage interface 104 may connect to memory 105 including, without limitation, memory drives, removable disc drives, etc., employing connection protocols such as serial advanced technology attachment (SATA), Integrated Drive Electronics (IDE), IEEE-1394, Universal Serial Bus (USB), fiber channel, Small Computer Systems interface (SCSI), etc. The memory drives may further include a drum, magnetic disc drive, magneto-optical drive, optical

drive, Redundant Array of Independent Discs (RAID), solid-state memory devices, solid- state drives, etc.

[0019] In some embodiments, the memory 105 may store a collection of program or database components, including, without limitation, a user interface, an operating system 107, a web server, etc. In some non-limiting embodiments or aspects, the system 100 may store user/application data, such as the data, variables, records, etc. as described in this disclosure. Such databases may be implemented as fault-tolerant, relational, scalable, secure databases such as Oracle or Sybase.

[0020] In some embodiments, the operating system 107 may facilitate resource management and operation of the system 100. Examples of operating systems include, without limitation, APPLE® MACINTOSH® OS X®, UNIX®, UNIX-like system distributions (E.G., BERKELEY SOFTWARE DISTRIBUTION® (BSD), FREEBSD®, NETBSD®, OPENBSD, etc.), LINUX® DISTRIBUTIONS (E.G., RED HAT®, UBUNTU®, KUBUNTU®, etc.), IBM®OS/2®, MICROSOFT® WINDOWS® (XP®, VISTA®/7/8, 10 etc.), APPLE® OS®, GOOGLE™ ANDROID™, BLACKBERRY® OS, or the like.

[0021] In some non-limiting embodiments or aspects, the system 100 may implement a web browser (not shown in the figures) stored program component. The web browser (not shown in the figures) may be a hypertext viewing application, such as MICROSOFT® INTERNET EXPLORER®, GOOGLE™ CHROME™, MOZILLA® FIREFOX®, APPLE® SAFARI®, etc. Secure web browsing may be provided using Secure Hypertext Transport Protocol (HTTPS), Secure Sockets Layer (SSL), Transport Layer Security (TLS), etc. Web browsers may utilize facilities such as AJAX, DHTML, ADOBE® FLASH®, JAVASCRIPT®, JAVA®, Application Programming Interfaces (APIs), etc.

[0022] Furthermore, one or more computer-readable storage media may be utilized in implementing embodiments consistent with the present disclosure. In some embodiments, a computer- readable storage medium refers to any type of physical memory on which information or data readable by a processor may be stored. Thus, a computer-readable storage medium may store instructions for execution by one or more processors, including instructions for causing the processor(s) to perform steps or stages consistent with the embodiments described herein. The term "computer-readable medium" should be understood to include tangible items and exclude carrier waves and transient signals, e.g., non-transitory. Examples include Random Access Memory (RAM), Read-Only Memory (ROM), volatile memory, non-volatile memory, hard drives, Compact Disc (CD) ROMs, Digital Video Disc (DVDs), flash drives, disks, and any other known physical storage media.

[0023] FIG. 2 illustrates a block diagram of an unauthorized service account detection unit 140 of FIG. 1 in accordance with some embodiments. In some embodiments, the unauthorized service account detection unit 140 is executable code configured to detect an unauthorized service account login to a service account on system 100. In some embodiments, a service account is a type of user account that is used by services or applications on system 100 to interact with other systems or resources within a network. In some embodiments, a service account login refers to the process or action of accessing or logging into a service account from, for example, a source machine or other computer system within a computer network or service account system. In some embodiments, an authorized service account login allows a user of system 100 to gain access to the functionalities, resources, or permissions of the service account. In some embodiments, detection of an unauthorized service account login is performed by the unauthorized service account detection unit 140 of system 100 in order to restrict access to

service accounts by an unauthorized user of the unauthorized service account login. In some embodiments, when a service account login is deemed to be unauthorized by the unauthorized service account detection unit 140, access to a service account using a service account login may be restricted by the service account system or by an administrator of the service account system that is notified of the unauthorized service account login.

[0024] In some embodiments, the unauthorized service account detection unit 140 includes a service account login collection unit 205 and a feature extraction and matrix generation unit 210. In some embodiments, the feature extraction and matrix generation unit 210 includes a probability generation and replacement unit 220 and an authorization threshold comparison unit 230. In some embodiments, the service account login collection unit 205 is executable code configured to collect service account login history associated with users of source machines and destination machines that access, login, or attempt to login into service accounts via system 100 (e.g., a service account system). In some embodiments, the service account login history is collected by service account login collection unit 205 and transformed to a service account login history probability matrix by probability generation and replacement unit 220. In some embodiments, the service account login history probability matrix is a three-dimensional matrix (3D) of probability values that indicate the probabilities of a user or users of a service account making a connection from a source machine to a destination machine based on, for example, the service account login history of the user or users. In some embodiments, the probability values may range from, for example, 0 to 1. In some embodiments, the probability values in the service account login history probability matrix are identified using a service account identification (ID), a source machine ID, and a destination machine ID. In some embodiments, the service account ID is an ID that indicates the identity of a service account (and thus the identity of the user

associated with the service account). In some embodiments, the source machine identification is an ID that identifies a source machine that a user has utilized to access or attempt to access a service account. In some embodiments, the destination machine ID is an ID that identifies a destination machine that a user of a source machine has accessed or attempted to access via a service account. In some embodiments, the service account login history or service account login history matrix used to generate the service account login history probability matrix may be collected over a predetermined time period, such as, for example, sixty days, ninety days, or some other predetermined time period.

[0025] In some embodiments, the feature extraction and matrix generation unit 210 is executable code configured to utilize a service account login history probability matrix to determine whether a service account login by a user of system 100 is unauthorized. In some embodiments, as stated previously, the feature extraction and matrix generation unit 210 includes probability generation and replacement unit 220 and authorization threshold comparison unit 230. In some embodiments, the probability generation and replacement unit 220 is executable code configured to generate a service-account-login-probability that is utilized by the feature extraction and matrix generation unit 210 to determine whether a service account login by a user of system 100 is unauthorized. In some embodiments, the probability generation and replacement unit 220 is configured to utilize a source-machine service-account matrix and a service-account destination matrix to generate the service-account-login-probability. In some embodiments, as described further herein, probability generation and replacement unit 220 utilizes machine learning based matrix factorization to generate the source-machine service-account matrix and the service-account destination matrix that are used to generate the service-account-login-probability. In some embodiments, the service-account-login-probability is utilized by the feature extraction and

matrix generation unit 210 to determine whether a service account login by a user of system 100 is unauthorized, described further herein with reference to FIG. 2. In some embodiments, the authorization threshold comparison unit 230 is executable code configured to utilize the service-account-login-probability to determine whether a service account login is unauthorized.

[0026] In some embodiments, in operation, probability generation and replacement unit 220 of feature extraction and matrix generation unit 210 receives a service account login history matrix from service account login collection unit 205. In some embodiments, after receiving the service account login history matrix, probability generation and replacement unit 220 transforms the service account login history matrix into a service account login history probability matrix. In some embodiments, as stated previously, the service account login history probability matrix is a 3D matrix of probability values that indicate the probabilities of a user of a service account making a connection from a source machine to a destination machine via system 100. In some embodiments, due the nature of data collection for the service account login history matrix, the service account login history probability matrix may include missing entries or empty cells that do not have probability values in the service account login history probability matrix. In some embodiments, the missing entries in the service account login history probability matrix may be caused by, for example, data collection or preprocessing issues during the collection of service account data by service account login collection unit 205. As a result, in some embodiments, feature extraction and matrix generation unit 210 may utilize a machine learning algorithm that utilizes a matrix factorization model to fill the missing entries of the service account login history probability matrix. In some embodiments, matrix factorization is a factorization technique utilized in, for example, machine learning applications, to fill missing entries in a matrix. In some embodiments, matrix factorization is performed by the probability generation

and replacement unit 220 by decomposing an original matrix (e.g., a 2D matrix) into lower-dimensional representations and using the lower-dimensional representations to estimate the missing values in the original matrix, as described further herein.

[0027] In some embodiments, prior to performing matrix factorization operations, upon receipt of the service account login history probability matrix, the probability generation and replacement unit 220 of feature extraction and matrix generation unit 210 performs unfolding or separation operations to separate the service account login history probability matrix into two 2D matrices. For example, in some embodiments, feature extraction and matrix generation unit 210 separates the service account login history probability matrix into a source-machine service-account matrix and a service-account destination matrix. In some embodiments, the source-machine service-account matrix is a 2D matrix of the probability values from the service account login history probability matrix associated with a user accessing a service account from a source machine. In some embodiments, the service-account destination matrix is a 2D matrix of the probability values from the service account login history probability matrix associated with a user accessing a destination machine utilizing service account. Examples of a source-machine service-account matrix and a service-account destination matrix are described herein with reference to FIG. 3C and FIG. 3D, respectively.

[0028] FIG. 3A illustrates an example three-dimensional coordinate system with x-y-z axis utilized to describe an example service account login history probability matrix. In some embodiments, in FIG. 3A, the x-axis is configured to represent a destination machine axis, the y-axis is configured to represent a source machine axis, and the z-axis is configured to represent a service account axis. FIG. 3B illustrates example 2D plane representations of an example service account login history probability matrix with example probability values in accordance with

some embodiments. In some embodiments, in FIG. 3B, the x-axis represents a destination machine axis and the y-axis represents a source machine axis. In FIG. 3B, there are three service accounts illustrated, service account 1 (e.g., svc acc), service account 2 (e.g., svc acc2) and service account 3 (svc acc 3). FIG. 3C illustrates a source-machine service-account matrix of the service account login history probability matrix of FIG. 3B with example probability values in accordance with some embodiments. In some embodiments, the source-machine service-account matrix corresponds to service account 3 of FIG. 3B. FIG. 3D illustrates a service-account destination matrix of the service account login history probability matrix of FIG. 3B with example probability values in accordance with some embodiments. In some embodiments, the service-account destination matrix corresponds to service account 3 of FIG. 3B.

[0029] In some embodiments, with further reference to FIG. 3C and FIG. 3D, the source-machine service-account matrix (as well as the service-account destination matrix), may be represented as an M x N matrix, where M represents the rows in the matrix and N represents the columns in the matrix. In some embodiments, for example, for a source-machine service-account matrix, M represents the source machine (e.g., the source machine that the user logged in from) and N represents the service account ID (e.g., the service account that represents the user). In some embodiments, for example, for a service-account destination matrix, M represents the service account ID and N represents the destination machine (e.g., the destination machine that the user logged into utilizing the service account).

[0030] In some embodiments, with further reference to FIG. 2, after probability generation and replacement unit 220 separates the service account login history probability matrix into the source-machine service-account matrix and the service-account destination matrix, probability generation and replacement unit 220 commences the process of utilizing machine learning based

matrix factorization operations on both the source-machine service-account matrix and the service-account destination matrix to estimate the missing probability values in each matrix. In some embodiments, as stated previously, probability generation and replacement unit 220 utilizes a machine learning algorithm that performs matrix factorization by decomposing the two 2D matrices (e.g., source-machine service-account matrix or service-account destination matrix) into lower-dimensional representations and using the representations to estimate the missing probability values in the source-machine service-account matrix and the service-account destination matrix. In some embodiments, as a result of the matrix factorization, the 2D matrix (represented as an M x N matrix) is represented as two matrices, an M x F matrix and F x N matrix, where F are features in each matrix. In some embodiments, because there is a source-machine service-account matrix and a service-account destination matrix, the feature extraction and feature extraction and matrix generation unit 210 performs two matrix factorization operations, a source-service account matrix factorization and a service-account destination matrix factorization.

[0031] In some embodiments, for each source-machine service-account matrix and service-account destination matrix, probability generation and replacement unit 220 trains a matrix factorization model to fill the missing entries in the M x F and F x N matrices to attain the desired output, M x N (e.g., source-machine service-account matrix or service-account destination matrix). In some embodiments, when training a matrix factorization model with M x F and F x N lower-dimensional matrices, the feature extraction and matrix generation unit 210 initializes the M x F and F x N matrices, which includes the feature factors in each matrix. In some embodiments, when initializing the M x F matrix and F x N matrix (which may denoted as, for example, U and V, respectively) in matrix factorization, the probability values in the missing

entries are initialized with random probability values or with predefined probability values. In some embodiments, the initialization method may vary depending on the specific implementation or algorithm being used during the matrix factorization operations. In some embodiments, for random initialization, the elements of U and V are randomly initialized and the probability values may be drawn from a uniform distribution or a Gaussian distribution. In some embodiments, for predefined initialization, the elements of U and V can be initialized with predefined values based on prior knowledge or domain expertise. In some embodiments, irrespective of the initialization method, the training process may update the elements in U and V based on the observed or actual values in the input matrix (e.g., source-machine service-account matrix or service-account destination matrix), iteratively optimizing the matrix factorization model to capture the underlying patterns and relationships in the source-machine service-account matrix and service-account destination matrix.

[0032] In some embodiments, after the M x F and F x N matrices are initialized by the feature extraction and matrix generation unit 210, probability generation and replacement unit 220 utilizes a loss function to measure the performance of the matrix factorization model. In some embodiments, the loss function is configured to quantify the discrepancy between the estimated probability values and the actual probability values in the M x N matrix (e.g., from the factorized matrices M x F and F x N). In some embodiments, the loss function may be, for example, a mean squared error (MSE), mean absolute error (MAE), or other loss function. In some embodiments, the actual execution of the loss function occurs during the training process when the matrix factorization model iteratively updates the factorized matrices M x F and F x N to minimize the loss function.

[0033] In some embodiments, during the training process, feature extraction and matrix generation unit 210 trains the matrix factorization model utilizing a machine learning optimization algorithm. In some embodiments, the optimization algorithm may be, for example, a stochastic gradient descent (SGD), alternating least squares (ALS), or any other suitable optimization algorithm configured to be utilized to iteratively update the elements in M x F and F x N.

[0034] In some embodiments, during the training process, probability generation and replacement unit 220 utilizes the factorized matrices M x F (e.g., matrix U) and F x N (e.g., matrix V) to estimate the missing values (e.g., missing entries) in the M x N matrix (e.g., matrix R). In some embodiments, the use of the terms R, U, and V to represent M x F, F x N, and M x N, respectively, are utilized for explanation purposes. In some embodiments, to predict or estimate the missing entries in matrix M x N (e.g., matrix R), probability generation and replacement unit 220 computes the dot product of the corresponding row in U (e.g., M x F) and column in V (e.g., F x N), which provides an estimation of the missing entry (e.g., missing probability value) in R. In some embodiments, the weights, denoted as W, are used to scale the contribution of each feature factor to the prediction. In some embodiments, the prediction for an entry r_ij in the matrix R may be computed by the probability generation and replacement unit 220 as follows:

[0035] $r_{ij} \approx \Sigma(W_k * U_{ik} * V_{kj})$

[0036] where, W_k represents the weight associated with the k-th feature factor, U_{ik} denotes the i-th user's strength of association with the k-th feature factor, and V_{kj} represents the j-th item's strength of association with the k-th feature factor.

[0037] In some embodiments, during the training process, the weights W_k are learned by optimizing the loss function using the aforementioned optimization algorithms or techniques, such as stochastic gradient descent. In some embodiments, by updating the weights iteratively, the model learns to assign appropriate importance to different feature factors, capturing the underlying patterns and relationships in the data. In some embodiments, the optimization process involves iterating through the actual probability values and adjusting the corresponding elements in M x F and F x N to improve the fit between the estimated and actual probability values. In summary, matrix factorization decomposes the input matrix R into matrices U and V and the weights W are used to scale the contribution of feature factors in U and V in predicting the missing entries (e.g., missing probability values) in R.

[0038] In some embodiments, after the probability values have been estimated, feature extraction and matrix generation unit 210 evaluates the performance of the matrix factorization model by comparing the estimated values with the actual values in the matrix. In some embodiments, feature extraction and matrix generation unit 210 utilizes standard evaluation metrics to assess the accuracy and quality of the filled-in missing probability values. In some embodiments, feature extraction and matrix generation unit 210 refines the matrix factorization model by adjusting hyperparameters, such as the number of feature factors, and, if necessary, and repeats the training process to further improve the results. In some embodiments, by factorizing the input matrix into U and V matrices and optimizing the U and V matrices based on the observed or actual values, the matrix factorization model may estimate the missing probability values by computing the values through the dot product of the corresponding rows in U and V.

[0039] In some embodiments, after the matrix factorization model has been trained and the probability values for M x F and F x N for each source-machine service-account matrix or

service-account destination matrix have been estimated, feature extraction and matrix generation unit 210 utilizes the source-machine service-account matrix or service-account destination matrix to determine whether a service account login (e.g., a new service account login) by a user of the system 100 is unauthorized.

[0040] In some embodiments, feature extraction and matrix generation unit 210 determines whether a service account login by a user of the system 100 is unauthorized by generating a service-account-login-probability based on the source-machine service-account matrix and the service-account destination matrix and comparing the service-account-login-probability to an authorization threshold value. In some embodiments, probability generation and replacement unit 220 calculates the service-account-login-probability by utilizing the product of the source-machine service-account matrix and the service-account destination matrix (e.g., multiplying the source-machine service-account matrix by the service-account destination matrix to compute service-account-login-probability). In some embodiments, each probability in the resulting matrix corresponds to the service-account-login-probability of the corresponding service account login (e.g., new service account login).

[0041] In some embodiments, after the service-account-login-probability has been calculated by the probability generation and replacement unit 220, authorization threshold comparison unit 230 compares the service-account-login-probability value to the authorization threshold value. In some embodiments, the authorization threshold value may be, for example, a value indicative of the service account login being authorized, such as, 0.7, 0.8, or some other a value that is indicative of the service account login being authorized. In some embodiments, the authorization threshold comparison unit 230 compares the service-account-login-probability value to the authorization threshold value by determining whether the service-account-login-probability value

is less than, equal to, or greater than the authorization threshold value. In some embodiments, when authorization threshold comparison unit 230 determines that the service-account-login-probability value is equal to or greater than the authorization threshold value, the service account login (e.g., new service account login) is considered authorized by the feature extraction and matrix generation unit 210. In some embodiments, when the service-account-login-probability value is less than the authorization threshold value, the service account login is considered not authorized by the feature extraction and matrix generation unit 210. In some embodiments, when the feature extraction and matrix generation unit 210 determines that the service account login is unauthorized access to system 100, the service account login may be restricted by the service account system or by an administrator of the service account system that is notified of the unauthorized service account login.

[0042] FIG. 4 is a flow diagram illustrating an unauthorized service account detection method 400 in accordance with some embodiments. In some embodiments, the unauthorized service account detection method 400 is a method utilized by unauthorized service account detection unit 140 of FIG. 2 to detect an unauthorized service account login by a user of system 100. The method, process steps, or stages illustrated in the figures may be implemented as an independent routine or process, or as part of a larger routine or process. Note that each process step or stage depicted may be implemented as an apparatus that includes a processor executing a set of instructions, a method, or a system, among other embodiments. In some embodiments, the unauthorized service account detection unit 140 is configured to perform the unauthorized service account detection method 400 in accordance with some embodiments. In some embodiments, the unauthorized service account detection method 400 is described with reference to the figures described herein.

[0043] In some embodiments, at operation 420, probability generation and replacement unit 220 of feature extraction and matrix generation unit 210 receives a service account login history matrix from service account login collection unit 205. In some embodiments, after receiving the service account login history matrix, probability generation and replacement unit 220 transforms the service account login history matrix into a service account login history probability matrix. In some embodiments, after transforming the service account login history information into the service account login history probability matrix, operation 420 proceeds to operation 430.

[0044] In some embodiments, at operation 430, probability generation and replacement unit 220 generates a source-machine service-account matrix from the service account login history probability matrix. In some embodiments, at operation 440, probability generation and replacement unit 220 generates a service-account destination matrix from the service account login history probability matrix. In some embodiments, as stated previously, the probability generation and replacement unit 220 generates the source-machine service-account matrix and the service-account destination matrix by utilizing unfolding operations configured to separate the service account login history probability into the source-machine service-account matrix and the service-account destination matrix (illustrated by way of example in FIG. 3C and FIG. 3D, respectively). In some embodiments, after generating the source-machine service-account matrix and the service-account destination matrix, operation 440 proceeds to operation 450.

[0045] In some embodiments, at operation 450, probability generation and replacement unit 220 generates estimated probability values to replace the values in the missing entries of the source-machine service-account matrix and the service-account destination matrix. In some embodiments, as stated previously, probability generation and replacement unit 220 utilizes machine learning based matrix factorization (e.g., source-machine service-account matrix

factorization and the service-account destination matrix factorization) to generate and replace the values in the missing entries of the source-machine service-account matrix and the service-account destination matrix. In some embodiments, after replacing the missing entries of the source-machine service-account matrix and the service-account destination matrix with the estimated probability values, operation 450 proceeds to operation 460.

[0046] In some embodiments, at operation 460, probability generation and replacement unit 220 utilizes the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability. In some embodiment, as stated previously, the service-account-login-probability is generated by performing product operations between the source-machine service-account matrix and the service-account destination matrix after performing the machine learning based matrix factorization operations (e.g., source-machine service-account matrix factorization and the service-account destination matrix factorization). In some embodiments, as stated previously, the probability generation and replacement unit 220 performs the matrix operations on both the source-machine service-account matrix and the service-account destination matrix in order to update the missing entries of the source-machine service-account matrix and the service-account destination matrix with estimated probability values. In some embodiments, after generating the service-account-login-probability, operation 460 proceeds to operation 470.

[0047] In some embodiments, at operation 470, authorization threshold comparison unit 230 of feature extraction and matrix generation unit 210 utilizes the service-account-login-probability to indicate whether a service account login is unauthorized. In some embodiments, as stated previously, in order to determine whether a service-account login is authorized, the authorization threshold comparison unit 230 compares the service-account-login-probability with an

authorization threshold value. In some embodiments, when the authorization threshold comparison unit 230 determines that the service-account-login-probability value is equal to or greater than the authorization threshold value, the service account login associated service-account-login-probability value is considered authorized by the feature extraction and matrix generation unit 210.

[0048] FIG. 5 illustrates an example implementation of the unauthorized service account detection method of FIG. 4 in accordance with some embodiments. For example, as illustrated in FIG. 5, in some embodiments, for a user 510 attempting to login to a service account via a source machine 520 with a source-machine service-account matrix probability of 0.8 and a destination machine 530 with a service-account destination matrix probability of 0.8, the service-account-login-probability calculated by the probability generation and replacement unit 220 is 0.64. In some embodiments, with an authorization threshold value of 0.6, authorization threshold comparison unit 230 determines that the service-account-login-probability is greater than the authorization threshold value, and thus the service account login attempt is deemed authorized by the feature extraction and matrix generation unit 210.

[0049] In some embodiments, with further reference to FIG. 2 – FIG. 5, an unsupervised learning approach may be utilized by the feature extraction and matrix generation unit 210 to group users of service machines into distinct groups based on user data related to service account access. In some embodiments, the feature extraction and matrix generation unit 210 may be configured to utilize a grouping algorithm that utilizes the user data (representative of, for example, the users of service machines) to group the users into distinct groups (e.g., grouped users or clusters). In some embodiments, the distinct groups may be utilized by the probability generation and replacement unit 220 to replace the missing probability values with the estimated

probability values. In some embodiments, for example, users of service machines may be grouped into a data scientist group or other distinct group. In some embodiments, the service accounts associated with the data scientist group may only be accessible to service account users belonging to the data scientist group. In some embodiments, as a result, a service account user may not be authorized to access service accounts affiliated with the data scientist group unless the service account user is a member of the data scientist group and thus, the service account may not be recommended or accessible to, for example, a user interface (UI) developer or a user experience (UX) developer.

[0050] In some embodiments, the unsupervised grouping algorithm that may be utilized by the feature extraction and matrix generation unit 210 to generate the group of users (e.g., K groups or clusters) is based on a feature matrix (e.g., a service account feature matrix, etc.), and may be, for example, a k-means clustering algorithm. In some embodiments, in the context of the matrix factorization described herein, a k-means clustering algorithm or k-means grouping algorithm is a clustering algorithm used to partition data into k distinct clusters based on similarity or distance metrics. K-means is an iterative algorithm that aims to minimize the within-cluster sum of squares by iteratively updating cluster assignments. In some embodiments, the service accounts are clustered into K groups based on a variance (e.g., a minimum or closest variance) between each member of each group. In some embodiments, using the clustered groups, whenever a service account login (e.g., new service account login) is attempted by a user of a service machine, the service account login may be matched initially with the clustered groups and, then the weights are assigned by the probability generation and replacement unit 220 in a group that matches the service account (e.g., the same group of the service account).

[0051] In some embodiments, a computer-implemented method includes receiving a service account login history associated with users of a service account, the service account login history being transformed to a service account login history probability matrix; generating a source-machine service-account matrix from the service account login history probability matrix; generating a service-account destination matrix from the service account login history probability matrix; and utilizing the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

[0052] In some embodiments of the computer-implemented method, the service account login history probability matrix is a three-dimensional (3D) probability matrix.

[0053] In some embodiments, the computer-implemented method further includes separating the service account login history probability matrix into the source-machine service-account matrix and the service-account destination matrix.

[0054] In some embodiments, the computer-implemented method further includes performing a source-machine service account matrix factorization of the source-machine service-account matrix and a service-account destination matrix factorization of the service-account destination matrix.

[0055] In some embodiments, the computer-implemented method further includes utilizing the source-machine service account matrix factorization to estimate source-machine service account matrix probability values in the source-machine service-account matrix.

[0056] In some embodiments, the computer-implemented method further includes utilizing the service-account destination matrix factorization to estimate service-account destination matrix probability values in the service-account destination matrix.

[0057] In some embodiments, the computer-implemented method further includes multiplying the source-machine service-account matrix that has been updated with estimated source-machine service account matrix probability values and the service-account destination matrix that has been updated with estimated service-account destination matrix probability values to ascertain the service-account-login-probability.

[0058] In some embodiments of the computer-implemented method, the service-account-login-probability is utilized to determine whether the service account login is authorized by comparing the service-account-login-probability to an authorization threshold value.

[0059] In some embodiments of the computer-implemented method, the service-account-login-probability is generated based upon a k-means grouping algorithm.

[0060] In some embodiments, a system includes a processor; and a non-transitory computer readable medium coupled to the processor, the non-transitory computer readable medium including code that: receives a service account login history matrix associated with users of a service account, the service account login history matrix being transformed to a service account login history probability matrix; generates a source-machine service-account matrix from the service account login history probability matrix; generates a service-account destination matrix from the service account login history probability matrix; and utilizes the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

[0061] In some embodiments of the system, the code separates the service account login history probability matrix into the source-machine service-account matrix and the service-account destination matrix.

[0062] In some embodiments of the system, the code performs a source-machine service account matrix factorization of the source-machine service-account matrix and a service-account destination matrix factorization of the service-account destination matrix.

[0063] In some embodiments of the system, the code utilizes the source-machine service account matrix factorization to estimate source-machine service account matrix probability values in the source-machine service-account matrix.

[0064] In some embodiments of the system, the code utilizes the service-account destination matrix factorization to estimate service-account destination matrix probability values in the service-account destination matrix.

[0065] In some embodiments of the system, the code multiplies the source-machine service-account matrix that has been updated with estimated source-machine service account matrix probability values and the service-account destination matrix that has been updated with estimated service-account destination matrix probability values to ascertain the service-account-login-probability.

[0066] In some embodiments of the system, the code determines whether the service account login is authorized by comparing the service-account-login-probability to an authorization threshold value.

[0067] In some embodiments of the system, the service-account-login-probability is generated based upon a k-means grouping algorithm.

[0068] In some embodiments, a computer-implemented method includes ascertaining a source-machine service-account matrix probability; ascertaining a service-account destination matrix probability; multiplying the source-machine service-account matrix probability by the service-account destination matrix probability to ascertain a service-account-login-probability; and

utilizing the service-account-login-probability to determine whether to report an unauthorized service account login.

[0069] In some embodiments of the computer-implemented method, the source-machine service-account matrix probability is ascertained utilizing a source-machine service account matrix factorization of a source-machine service-account matrix.

[0070] In some embodiments of the computer-implemented method, the service-account destination matrix probability is ascertained utilizing a service-account destination matrix factorization of a service-account destination matrix.

[0071] In one aspect, the embodiments described herein improve upon other computer systems by receiving a service account login history associated with users of a service account, the service account login history being transformed to a service account login history probability matrix; generating a source-machine service-account matrix from the service account login history probability matrix; generating a service-account destination matrix from the service account login history probability matrix; and utilizing the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized. In some aspects, the practical application of embodiments described herein aim to address challenges presented using other computer systems (e.g., lack of speed, cost-effectiveness, etc.) by introducing embodiments described herein to indicate whether a service account is unauthorized. By leveraging the unique characteristics and functionalities of embodiments described herein, the practical application of the embodiments optimizes existing processes, resulting in superior outcomes with increased speed, precision, and/or cost-effectiveness.

WHAT IS CLAIMED IS:

1. A computer-implemented method, comprising:

receiving a service account login history associated with users of a service account, the service account login history being transformed to a service account login history probability matrix;

generating a source-machine service-account matrix from the service account login history probability matrix;

generating a service-account destination matrix from the service account login history probability matrix; and

utilizing the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

2. The computer-implemented method of claim 1, wherein:

the service account login history probability matrix is a three-dimensional (3D) probability matrix.

3. The computer-implemented method of claim 2, further comprising:

separating the service account login history probability matrix into the source-machine service-account matrix and the service-account destination matrix.

4. The computer-implemented method of claim 3, further comprising:

performing a source-machine service account matrix factorization of the source-machine service-account matrix and a service-account destination matrix factorization of the service-account destination matrix.

5. The computer-implemented method of claim 4, further comprising:

utilizing the source-machine service account matrix factorization to estimate source-machine service account matrix probability values in the source-machine service-account matrix.

6. The computer-implemented method of claim 5, further comprising:

utilizing the service-account destination matrix factorization to estimate service-account destination matrix probability values in the service-account destination matrix.

7. The computer-implemented method of claim 6, further comprising:

multiplying the source-machine service-account matrix that has been updated with estimated source-machine service account matrix probability values and the service-account destination matrix that has been updated with estimated service-account destination matrix probability values to ascertain the service-account-login-probability.

8. The computer-implemented method of claim 7, wherein:

the service-account-login-probability is utilized to determine whether the service account login is authorized by comparing the service-account-login-probability to an authorization threshold value.

9. The computer-implemented method of claim 8, wherein:

the service-account-login-probability is generated based upon a k-means grouping algorithm.

10. A system, comprising:

a processor; and

a non-transitory computer readable medium coupled to the processor, the non-transitory computer readable medium including code that:

receives a service account login history matrix associated with users of a service account, the service account login history matrix being transformed to a service account login history probability matrix;

generates a source-machine service-account matrix from the service account login history probability matrix;

generates a service-account destination matrix from the service account login history probability matrix; and

utilizes the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability, the service-account-login-probability being utilized to indicate whether a service account login is unauthorized.

11. The system of claim 10, wherein:

the code separates the service account login history probability matrix into the source-machine service-account matrix and the service-account destination matrix.

12. The system of claim 11, wherein:

the code performs a source-machine service account matrix factorization of the source-machine service-account matrix and a service-account destination matrix factorization of the service-account destination matrix.

13. The system of claim 12, wherein:

the code utilizes the source-machine service account matrix factorization to estimate source-machine service account matrix probability values in the source-machine service-account matrix.

14. The system of claim 13, wherein:

the code utilizes the service-account destination matrix factorization to estimate service-account destination matrix probability values in the service-account destination matrix.

15. The system of claim 14, wherein:

the code multiplies the source-machine service-account matrix that has been updated with estimated source-machine service account matrix probability values and the service-account destination matrix that has been updated with estimated service-account destination matrix probability values to ascertain the service-account-login-probability.

16. The system of claim 15, wherein:

the code determines whether the service account login is authorized by comparing the service-account-login-probability to an authorization threshold value.

17. The system of claim 16, wherein:

the service-account-login-probability is generated based upon a k-means grouping algorithm.

18. A computer-implemented method, comprising:

ascertaining a source-machine service-account matrix probability;

ascertaining a service-account destination matrix probability;

multiplying the source-machine service-account matrix probability by the service-account destination matrix probability to ascertain a service-account-login-probability; and

utilizing the service-account-login-probability to determine whether to report an unauthorized service account login.

19. The computer-implemented method of claim 18, wherein:

the source-machine service-account matrix probability is ascertained utilizing a source-machine service account matrix factorization of a source-machine service-account matrix.

20. The computer-implemented method of claim 19, wherein:

the service-account destination matrix probability is ascertained utilizing a service-account destination matrix factorization of a service-account destination matrix.

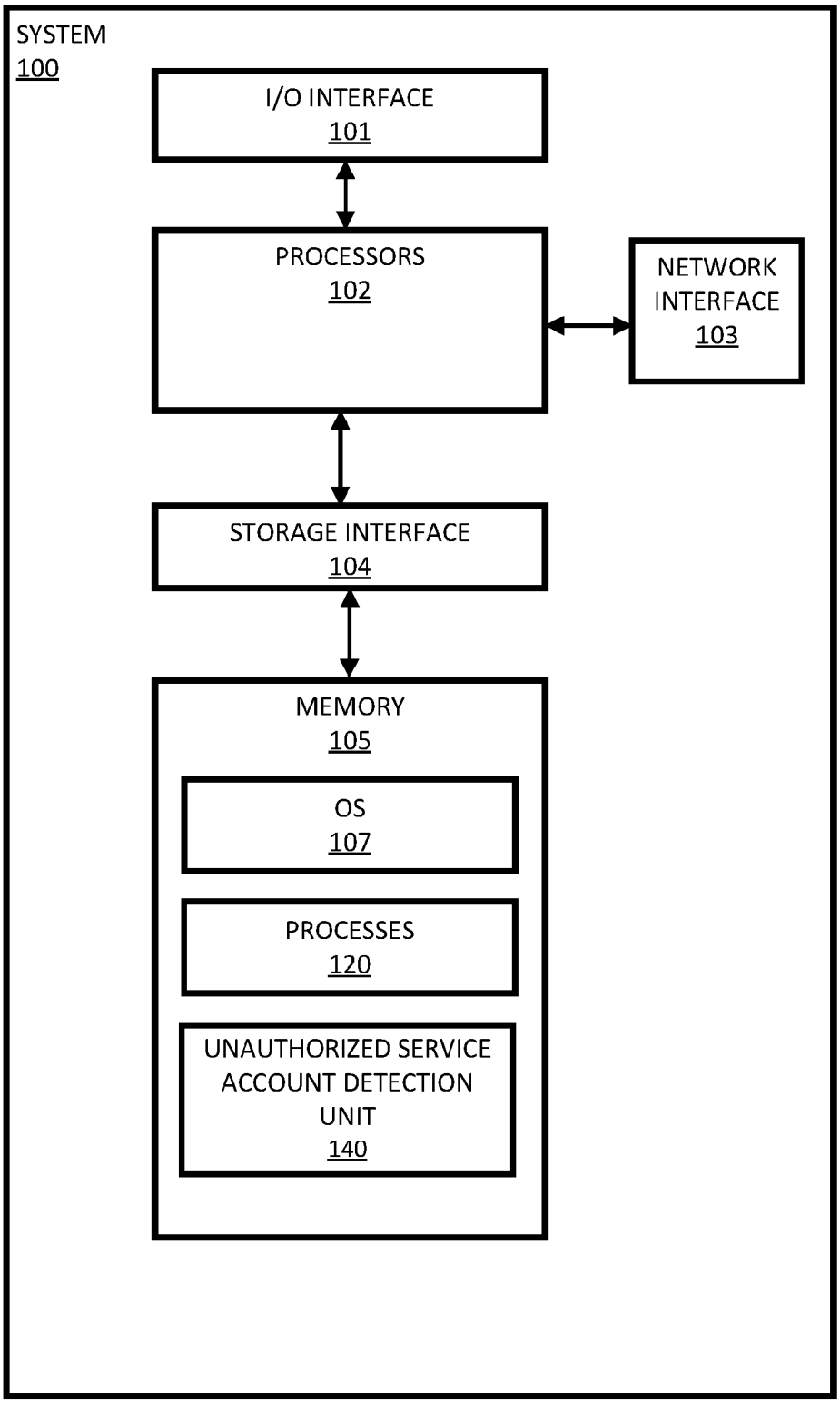
SYSTEM
100
I/O INTERFACE
101
PROCESSORS
102
NETWORK
INTERFACE
103
STORAGE INTERFACE
104
MEMORY
105
OS
107
PROCESSES
120
UNAUTHORIZED SERVICE
ACCOUNT DETECTION
UNIT
140

FIG. 1

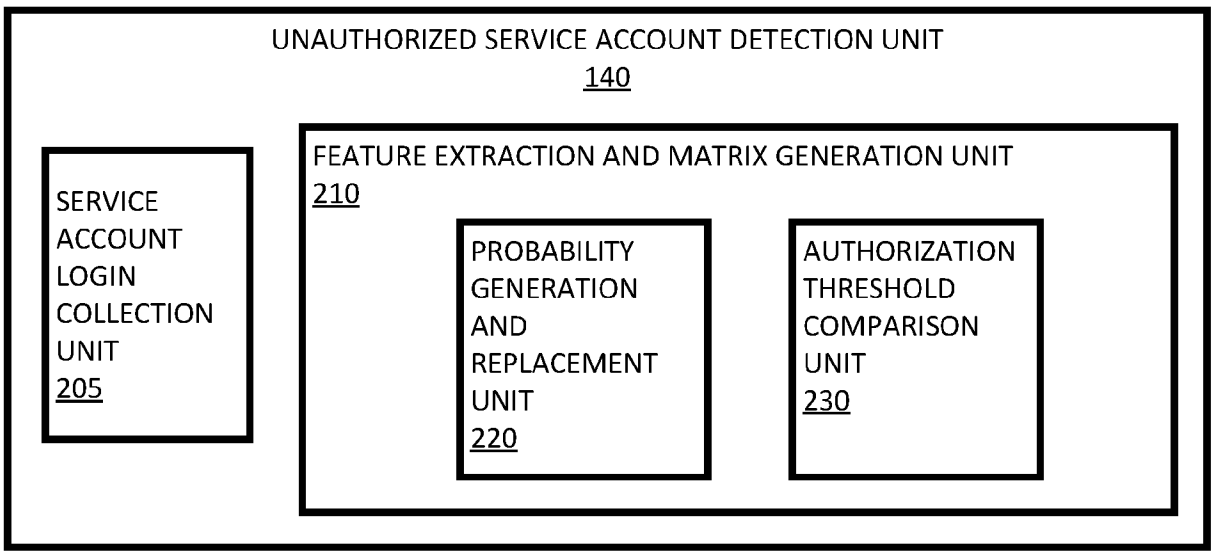
UNAUTHORIZED SERVICE ACCOUNT DETECTION UNIT
140
SERVICE ACCOUNT LOGIN COLLECTION UNIT
205
FEATURE EXTRACTION AND MATRIX GENERATION UNIT
210
PROBABILITY GENERATION AND REPLACEMENT UNIT
220
AUTHORIZATION THRESHOLD COMPARISON UNIT
230

FIG. 2

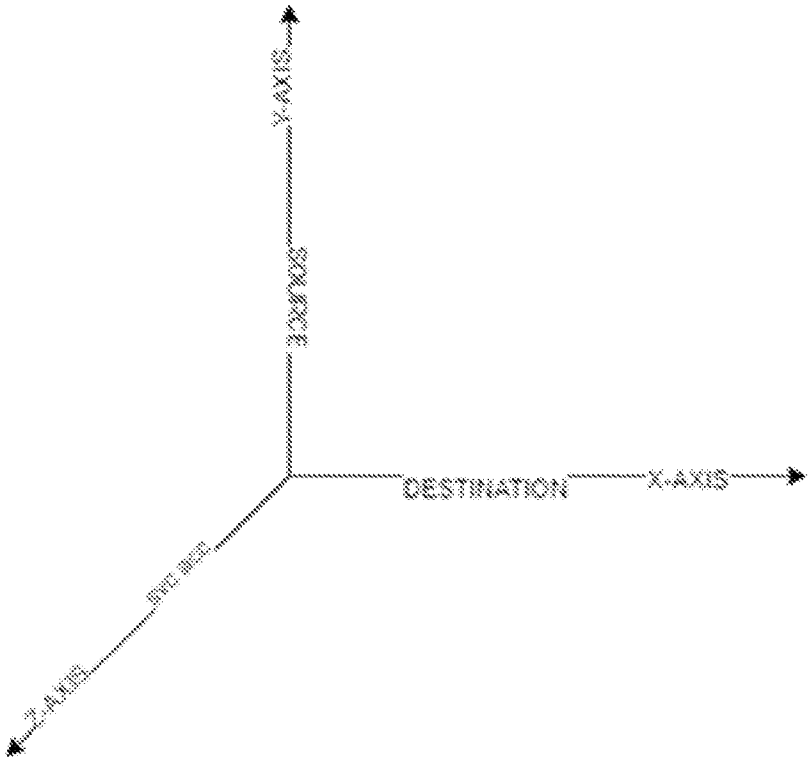

FIG. 3A

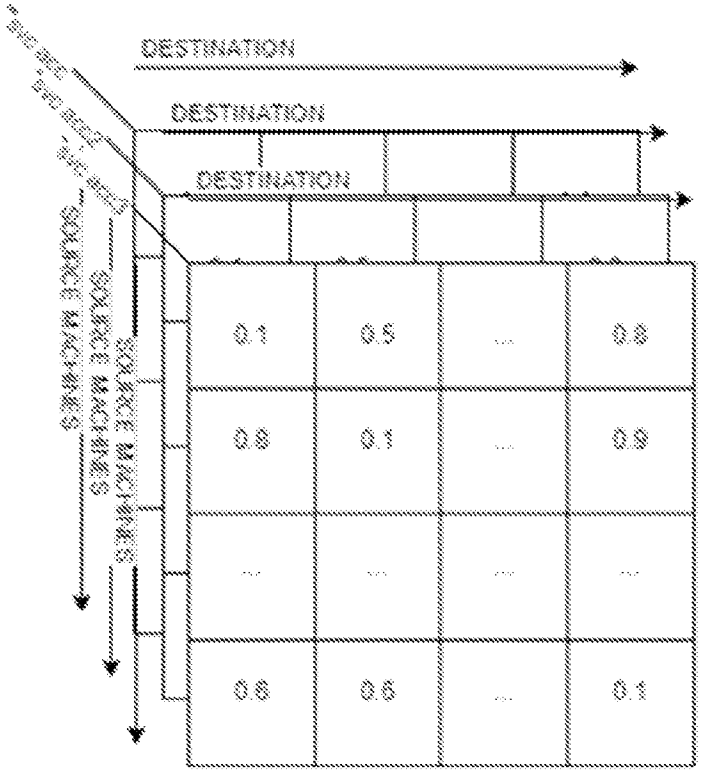
DESTINATION
DESTINATION
DESTINATION
SOURCE WORKERS
SOURCE WORKERS
SOURCE WORKERS
0.1
0.5
...
0.8
0.8
0.1
...
0.9
...
...
...
...
0.6
0.6
...
0.1
x
y

FIG. 3B

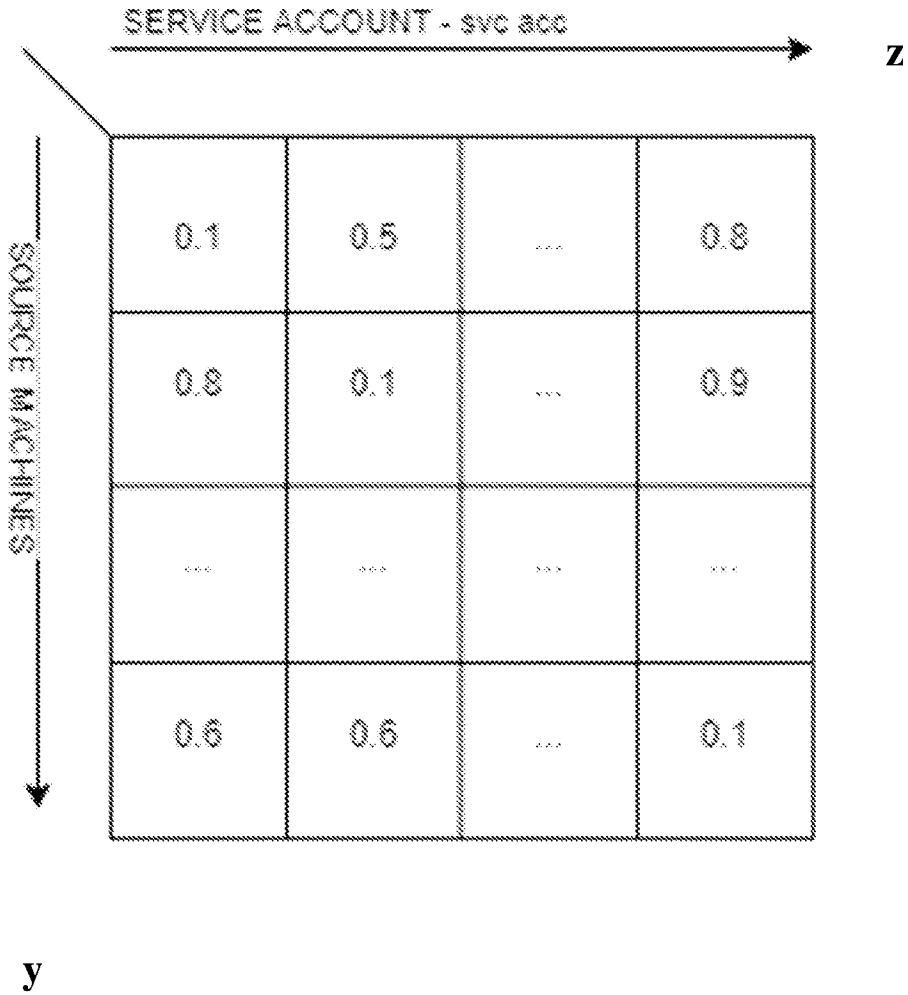
SERVICE ACCOUNT - svc acc
z
SOURCE MACHINES
y
0.1
0.5
...
0.8
0.8
0.1
...
0.9
...
...
...
...
0.6
0.6
...
0.1

FIG. 3C

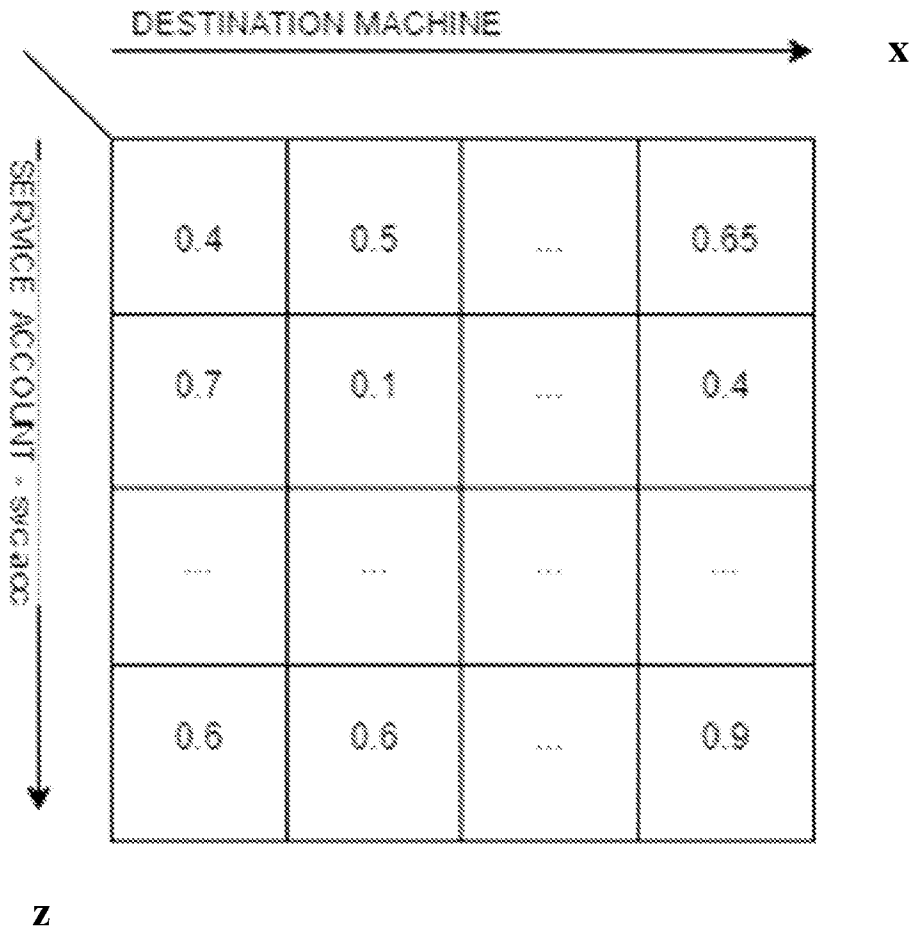
DESTINATION MACHINE
x
SERVICE ACCOUNT - SVCacc
0.4
0.5
...
0.65
0.7
0.1
...
0.4
...
...
...
...
0.6
0.6
...
0.9
z

FIG. 3D

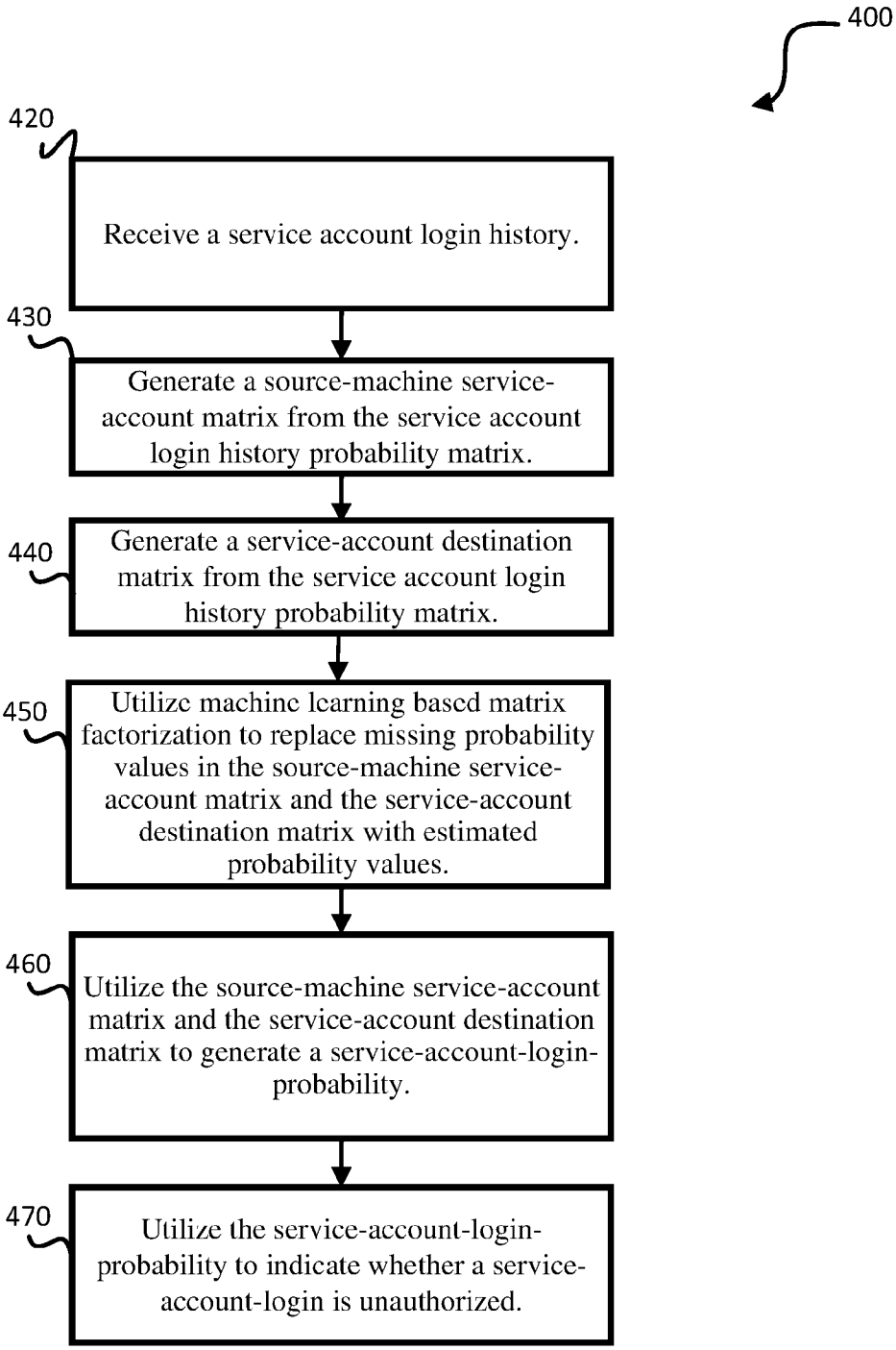
400
420
Receive a service account login history.
430
Generate a source-machine service-account matrix from the service account login history probability matrix.
440
Generate a service-account destination matrix from the service account login history probability matrix.
450
Utilize machine learning based matrix factorization to replace missing probability values in the source-machine service-account matrix and the service-account destination matrix with estimated probability values.
460
Utilize the source-machine service-account matrix and the service-account destination matrix to generate a service-account-login-probability.
470
Utilize the service-account-login-probability to indicate whether a service-account-login is unauthorized.

FIG. 4

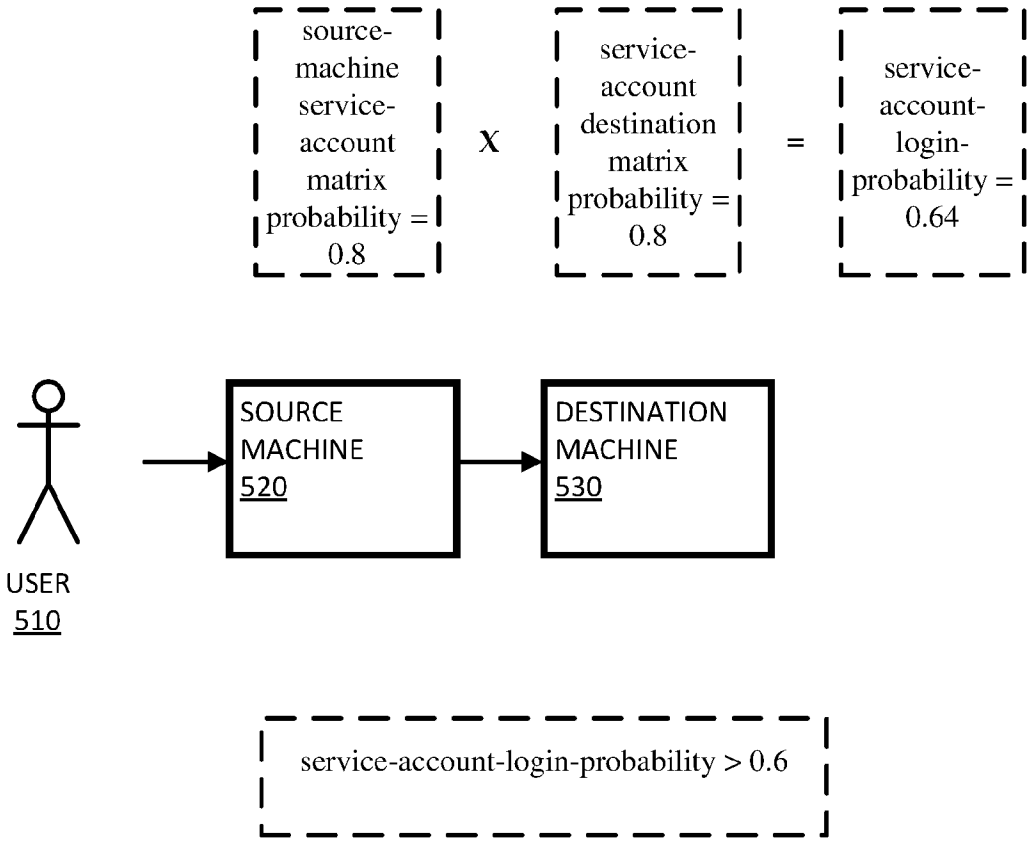
source-machine service-account matrix probability = 0.8
X
service-account destination matrix probability = 0.8
=
service-account-login-probability = 0.64
USER
510
SOURCE MACHINE
520
DESTINATION MACHINE
530
service-account-login-probability > 0.6

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/026162

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F21/55
ADD.

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)
G06F

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)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2018/124082 A1 (SIADATI SEYEDHOSSEIN [US] ET AL) 3 May 2018 (2018-05-03) paragraph [0027] - paragraph [0028]	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

"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
15 January 2024	24/01/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Vinck, Bart

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2023/026162

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018124082 A1	03-05-2018	NONE	