

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年11月17日(17.11.2022)



(10) 国際公開番号

WO 2022/239147 A1

(51) 国際特許分類:
G06F 21/55 (2013.01)

(21) 国際出願番号: PCT/JP2021/018049

(22) 国際出願日: 2021年5月12日(12.05.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 越智 勇貴(OCHI, Yuki); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 久田 裕介(HISADA,

Yusuke); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).

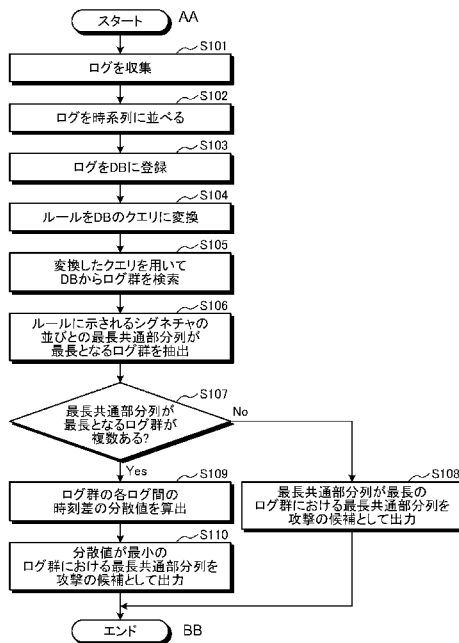
(74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎の門三井ビルディング Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): 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,

(54) Title: EXTRACTION DEVICE, EXTRACTION METHOD, AND EXTRACTION PROGRAM

(54) 発明の名称: 抽出装置、抽出方法、および、抽出プログラム

[図6]



S101 Collect logs
S102 Chronologically arrange logs
S103 Register logs in database
S104 Convert rule into database query
S105 Search database for log groups using converted query
S106 Extract log group for which maximum length common subsequence with respect to signature sequence indicated in rule is longest
S107 Are there plurality of log groups for which maximum length common subsequence is longest?
S108 Output, as attack candidate, maximum length common subsequence of log group if maximum length common subsequence is longest
S109 Calculate variance value of time difference between logs in each log group
S110 Output, as attack candidate, maximum length common subsequence of log group for which variance value is smallest
AA Start
BB End

(57) Abstract: An extraction device (10) collects log groups from a computer under investigation. The extraction device (10) then refers to a rule in which a plurality of signatures indicating attacks on computers are arranged in an order characteristic of the attacks, and extracts log groups matching any of the signatures from the collected data. The extraction device (10) then extracts, from among the extracted log groups, each log group for which the maximum length common subsequence between the chronological signature sequence matched by the logs of the log group and the signature sequence

[続葉有]

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) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (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).

添付公開書類：

一 国際調査報告(条約第21条(3))

indicated in the rule is the longest. The extraction device (10) then calculates, for each log group for which the maximum length common subsequence with respect to the signature sequence indicated in the rule is the longest, the variance value of the time difference between chronologically adjacent logs in the log group. The extraction device (10) then outputs, as an attack trace candidate, the maximum length common subsequence of the log group for which the variance value is the smallest among the extracted log groups.

(57) 要約：抽出装置(10)は、調査対象のコンピュータのログ群を収集する。そして、抽出装置(10)は、コンピュータへの攻撃を示す複数のシグネチャを攻撃に特徴的な順序に並べたルールを参照して、上記の収集したデータから、いずれかのシグネチャにマッチするログ群を抽出する。次に、抽出装置(10)は、抽出されたログ群の各ログがマッチするシグネチャの時系列での並びとルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する。その後、抽出装置(10)は、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群それぞれについて、当該ログ群における時系列で隣接する各ログ間の時刻差の分散値を算出する。そして、抽出装置(10)は、抽出したログ群のうち、上記の分散値が最も小さいログ群における最長共通部分列を攻撃の痕跡の候補として出力する。

明 細 書

発明の名称：抽出装置、抽出方法、および、抽出プログラム

技術分野

[0001] 本発明は、攻撃の痕跡を示すログを抽出するための、抽出装置、抽出方法、および、抽出プログラムに関する。

背景技術

[0002] 近年、コンピュータフォレンジック調査において、攻撃の特徴を示すシグネチャを用いて、PC (Personal Computer) のログから攻撃の痕跡を示すログを抽出することが行われている（非特許文献1 参照）。

先行技術文献

非特許文献

[0003] 非特許文献1：Sigma-Generic Signatures for SIEM Systems、[2021年4月22日検索]、インターネット<URL：https://www.slideshare.net/secret/gvgxeXoKblXRcA>

発明の概要

発明が解決しようとする課題

[0004] しかし、従来のシグネチャを用いて攻撃の痕跡を示すログを抽出する方法は、攻撃以外の通常の操作で出力されるようなログが抽出されることもある。そこで、本発明は、前記した問題を解決し、攻撃の痕跡を精度よく抽出することを課題とする。

課題を解決するための手段

[0005] 前記した課題を解決するため、本発明は、調査対象のコンピュータのログを収集するログ収集部と、コンピュータへの攻撃を示す複数のシグネチャを攻撃に特徴的な順序に並べたルールを参照して、前記収集したログから、前記ルールに示されるいずれかのシグネチャにマッチするログ群を抽出する第1の抽出部と、前記抽出されたログ群の各ログがマッチするシグネチャの時系列での並びと、前記ルールに示される複数のシグネチャの並びとの最長共

通部分列が最長となるログ群を抽出する第2の抽出部と、前記最長共通部分列が最長となるログ群それぞれについて、前記ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する算出部と、前記算出した分散値が最小となるログ群における前記最長共通部分列を攻撃の候補として出力する出力処理部とを備えることを特徴とする。

発明の効果

[0006] 本発明によれば、攻撃の痕跡を精度よく抽出することができる。

図面の簡単な説明

[0007] [図1]図1は、抽出装置の概要を説明するための図である。

[図2]図2は、抽出装置の概要を説明するための図である。

[図3]図3は、抽出装置の構成例を示す図である。

[図4]図4は、図3のルールに用いられるシグネチャの例を示す図である。

[図5]図5は、図3のルールの例を示す図である。

[図6]図6は、抽出装置の処理手順の例を示すフローチャートである。

[図7]図7は、抽出プログラムを実行するコンピュータの例を示す図である。

発明を実施するための形態

[0008] 以下、図面を参照しながら、本発明を実施するための形態（実施形態）について説明する。本発明は以下に説明する実施形態に限定されない。

[0009] [概要]

まず、図1および図2を用いて、本実施形態の抽出装置の概要を説明する。まず、抽出装置は、事前にコンピュータへの攻撃を示す複数のシグネチャを、攻撃に特徴的な順序に並べたルール（図1の符号101参照）を用意しておく。

[0010] 次に、抽出装置は、調査対象のコンピュータのログ群を時系列に並べ、ルールに示される各シグネチャにマッチするログ群（符号102参照）を抽出する。次に、抽出装置は、抽出したログ群のログがマッチするシグネチャの並びと、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する（符号103～符号105参照）。

[0011] ここで、図1の符号103～符号105に示すように、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群が複数である場合、抽出装置は、ログ群それぞれについて、当該ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する。

[0012] 例えば、抽出装置は、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群が、図2の符号103に示すログ群である場合、当該ログ群において時系列で隣接する各ログ間の時刻差は「2:00、1:00、3:00」である（符号201参照）。よって、抽出装置は、「2:00、1:00、3:00」の分散値を算出する。例えば、抽出装置は、「2:00、1:00、3:00」を秒に変換し、秒に変換した値（「120、60、180」）の分散値を算出する。

[0013] 抽出装置は、上記の分散値の算出を、ルールに示されるシグネチャの並びとの最長共通部分列が最長となる他のログ群についても実行し、分散値が最小のログ群における最長共通部分列を攻撃の候補として出力する。例えば、図1の符号103～105に示すログ群のうち、符号103に示すログ群の分散値が最小である場合、抽出装置は、当該ログ群における最長共通部分列（シグネチャAにマッチ→シグネチャBにマッチ→シグネチャCにマッチ→シグネチャDにマッチ）を攻撃の痕跡の候補として出力する。

[0014] このような抽出装置によれば、調査対象のコンピュータのログから、ルールに示される一連のシグネチャとの最長共通部分列が最長となる一連の攻撃を、攻撃の候補として抽出することができる。これにより、抽出装置は、調査対象のコンピュータのログから、攻撃の痕跡の候補を精度よく抽出することができる。

[0015] [構成例]

次に、図3を用いて抽出装置10の構成例を説明する。抽出装置10は、入出力部11と、記憶部12と、制御部13とを備える。

[0016] 入出力部11は、各種データの入出力を司るインタフェースである。例えば、入出力部11は、調査対象となるコンピュータのログの入力を受け付けたり、攻撃の痕跡の候補を出力したりする。

- [0017] 記憶部 12 は、ルールを記憶する。このルールは、コンピュータへの攻撃を示す複数のシグネチャを、攻撃に特徴的な順序に並べたものである（図 1 の符号 101 参照）。
- [0018] シグネチャは、コンピュータが攻撃を受けた際に、当該コンピュータにより記録されるログの特徴を示したものである。例えば、シグネチャは、コンピュータにおけるマルウェアの動作を示したものである（図 4 参照）。
- [0019] ルールは、例えば、図 5 に示すように、各シグネチャに対し、攻撃に特徴的な順序を示す値を付与することにより記述される。ここで各シグネチャに付与される順序を示す値は、同じ値のものがあってもよい。また、シグネチャの中に順序を示す値が付与されていないシグネチャがあってもよい。
- [0020] 図 3 の説明に戻る。また、記憶部 12 は DB を備える。この DB には、DB 登録部 133 により登録されたログ（調査対象のコンピュータのログを時系列で並べたもの）が記憶される。
- [0021] 制御部 13 は、抽出装置 10 全体の制御を司る。この制御部 13 は、ログ収集部 131 と、時系列化部 132 と、DB 登録部 133 と、ルール変換部 134 と、DB 検索部（第 1 の抽出部）135 と、抽出部（第 2 の抽出部）136 と、判定部 137 と、算出部 138 と、絞り込み部（出力処理部）139 とを備える。
- [0022] ログ収集部 131 は、調査対象のコンピュータからログを収集する。例えば、ログ収集部 131 は、調査対象のコンピュータからイベントログ、レジストリ、ファイル操作履歴等を収集する。例えば、ログ収集部 131 は、CDIR-C (Cyber Defense Institute Incident Response Collector) 等により、上記のログを収集する。
- [0023] 時系列化部 132 は、ログ収集部 131 により収集されたログを時系列に並べ替える。DB 登録部 133 は、時系列化部 132 により時系列に並べ替えられたログを DB に登録する。
- [0024] ルール変換部 134 は、ルールに記載されたシグネチャを、当該シグネチャにマッチするログを検索するための検索クエリに変換する。DB 検索部 1

35は、ルール変換部134により変換された検索クエリを用いて、DB内のログを検索する。つまり、DB検索部135は、ルールに示されるシグネチャにマッチするログ群をDBから抽出する。例えば、DB検索部135は、図1の符号101に示す各シグネチャにマッチするログ群を、DBから抽出する（図1の符号102参照）。

[0025] 図3の説明に戻る。抽出部136は、DB検索部135により検索されたログ群から、当該ログ群における各ログがマッチするシグネチャの時系列での並びに基づき、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する。例えば、抽出部136は、図1の符号102に示すログ群から、図1の符号101に示すシグネチャの並びとの最長部分文字列が最長となるログ群を抽出する（図1の符号103～符号105参照）。

[0026] 図3の説明に戻る。判定部137は、抽出部136により抽出された、最長共通部分列が最長となるログ群が複数あるか否かを判定する。判定部137により、最長共通部分列が最長となるログ群が複数あると判定された場合、算出部138は、最長共通部分列が最長となるログ群それぞれについて、当該ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する。

[0027] 例えば、算出部138は、図1の符号103に示す、最長共通部分列が最長となるログ群について、当該ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する（図2の符号201参照）。また、算出部138は、図1の符号104、105に示すログ群についても同様に、当該ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する。

[0028] 図3の説明に戻る。絞り込み部139は、攻撃の痕跡の候補の絞り込みを行う。例えば、絞り込み部139は、抽出部136により抽出された、最長共通部分列が最長となるログ群のうち、算出部138により算出された分散値が最も小さいログ群における最長共通部分列を、攻撃の痕跡の候補（攻撃の候補）として出力する。

[0029] なお、抽出部136により抽出された最長共通部分列が最長となるログ群が1つのみである場合、絞り込み部139は、抽出部136により抽出されたログ群における最長共通部分列を攻撃の候補として出力する。以上説明した抽出装置10によれば、攻撃の候補を精度よく抽出することができる。

[0030] [処理手順の例]

次に、図6を用いて、抽出装置10の処理手順の例を説明する。まず、ログ収集部131は、入出力部11経由で、調査対象のコンピュータから処理対象のログを収集する(S101)。次に、時系列化部132は、S101で収集されたログを時系列に並べる(S102)。そして、DB登録部133は、S102で並べ替えられたログをDBに登録する(S103)。

[0031] また、ルール変換部134は、記憶部12のルールを読み出し、当該ルールをDBのクエリに変換する(S104)。そして、DB検索部135は、S104で変換されたクエリでDBからログ群を抽出する(S105)。つまり、DB検索部135は、ルールに示されるシグネチャにマッチするログ群をDBから抽出する。

[0032] S105の後、抽出部136は、S105で抽出されたログ群から、当該ログ群を構成するログに対応するシグネチャの並びと、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する(S106)。

[0033] S106の後、判定部137は、S106で抽出された最長共通部分列が最長となるログ群が複数あるか否かを判定し(S107)、複数ない場合(S107でNo)、S106で抽出されたログ群における最長共通部分列を攻撃の候補として出力する(S108)。

[0034] 一方、S107で、判定部137が、S106で抽出された最長共通部分列が最長となるログ群が複数あると判定した場合(S107でYes)、算出部138は、S106で抽出されたログ群ごとに、当該ログ群の各ログ間の時刻差の分散値を算出する(S109)。そして、絞り込み部139は、S106で抽出されたログ群のうち、S109で算出した分散値が最小の口

グ群における最長共通部分列を、攻撃の候補として出力する（S 1 1 0）。このようにすることで抽出装置 1 0 は、攻撃の候補を精度よく抽出することができる。

[0035] [その他の実施形態]

なお、絞り込み部 1 3 9 は、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群における最長共通部分列を攻撃の候補として出力することとしたが、これに限定されない。例えば、絞り込み部 1 3 9 は、攻撃の候補として、上記の最長共通部分列のみならず、最長共通部分列が最長となるログ群（各ログの時刻情報を含む）も出力してもよい。このようにすることで、抽出装置 1 0 のユーザは、攻撃の候補の内容を、より詳細に分析することができる。

[0036] また、前記した実施形態において、抽出装置 1 0 は、コンピュータから取得したログを時系列に並べてから、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出することとしたがこれに限定されない。

[0037] 例えば、抽出装置 1 0 は、コンピュータから取得したログのうち、ルールに示されるいずれかのシグネチャにマッチするログ群を抽出した後、時系列に並べ替える。そして、抽出装置 1 0 は、時系列に並べ替えたログがマッチするシグネチャの並びと、ルールに示されるシグネチャの並びとの最長共通部分列が最長となるログ群を抽出してもよい。

[0038] [システム構成等]

また、図示した各部の各構成要素は機能概念的なものであり、必ずしも物理的に図示のように構成されていることを要しない。すなわち、各装置の分散・統合の具体的形態は図示のものに限られず、その全部又は一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的又は物理的に分散・統合して構成することができる。さらに、各装置にて行われる各処理機能は、その全部又は任意の一部が、CPU 及び当該 CPU にて実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実

現され得る。

[0039] また、前記した実施形態において説明した処理のうち、自動的に行われるものとして説明した処理の全部又は一部を手動的に行うこともでき、あるいは、手動的に行われるものとして説明した処理の全部又は一部を公知の方法で自動的に行うこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。

[0040] [プログラム]

前記した抽出装置 10 は、パッケージソフトウェアやオンラインソフトウェアとしてプログラムを所望のコンピュータにインストールさせることによって実装できる。例えば、上記のプログラムを情報処理装置に実行させることにより、情報処理装置を抽出装置 10 として機能させることができる。ここで言う情報処理装置には、デスクトップ型又はノート型のパーソナルコンピュータが含まれる。また、その他にも、情報処理装置にはスマートフォン、携帯電話機やPHS (Personal Handyphone System) 等の移動体通信端末、さらには、PDA (Personal Digital Assistant) 等の端末等がその範疇に含まれる。

[0041] また、抽出装置 10 は、ユーザが使用する端末装置をクライアントとし、当該クライアントに上記の処理に関するサービスを提供するサーバ装置として実装することもできる。この場合、サーバ装置は、Webサーバとして実装することとしてもよいし、アウトソーシングによって上記の処理に関するサービスを提供するクラウドとして実装することとしてもかまわない。

[0042] 図 7 は、抽出プログラムを実行するコンピュータの例を示す図である。コンピュータ 1000 は、例えば、メモリ 1010、CPU 1020 を有する。また、コンピュータ 1000 は、ハードディスクドライバインタフェース 1030、ディスクドライバインタフェース 1040、シリアルポートインタフェース 1050、ビデオアダプタ 1060、ネットワークインタフェース 1070 を有する。これらの各部は、バス 1080 によって接続される。

- [0043] メモリ 1 0 1 0 は、ROM (Read Only Memory) 1 0 1 1 及び RAM (Random Access Memory) 1 0 1 2 を含む。ROM 1 0 1 1 は、例えば、BIOS (Basic Input Output System) 等のブートプログラムを記憶する。ハードディスクドライブインタフェース 1 0 3 0 は、ハードディスクドライブ 1 0 9 0 に接続される。ディスクドライブインタフェース 1 0 4 0 は、ディスクドライブ 1 1 0 0 に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ 1 1 0 0 に挿入される。シリアルポートインタフェース 1 0 5 0 は、例えばマウス 1 1 1 0、キーボード 1 1 2 0 に接続される。ビデオアダプタ 1 0 6 0 は、例えばディスプレイ 1 1 3 0 に接続される。
- [0044] ハードディスクドライブ 1 0 9 0 は、例えば、OS 1 0 9 1、アプリケーションプログラム 1 0 9 2、プログラムモジュール 1 0 9 3、プログラムデータ 1 0 9 4 を記憶する。すなわち、上記の抽出装置 1 0 が実行する各処理を規定するプログラムは、コンピュータにより実行可能なコードが記述されたプログラムモジュール 1 0 9 3 として実装される。プログラムモジュール 1 0 9 3 は、例えばハードディスクドライブ 1 0 9 0 に記憶される。例えば、抽出装置 1 0 における機能構成と同様の処理を実行するためのプログラムモジュール 1 0 9 3 が、ハードディスクドライブ 1 0 9 0 に記憶される。なお、ハードディスクドライブ 1 0 9 0 は、SSD (Solid State Drive) により代替されてもよい。
- [0045] また、上述した実施形態の処理で用いられるデータは、プログラムデータ 1 0 9 4 として、例えばメモリ 1 0 1 0 やハードディスクドライブ 1 0 9 0 に記憶される。そして、CPU 1 0 2 0 が、メモリ 1 0 1 0 やハードディスクドライブ 1 0 9 0 に記憶されたプログラムモジュール 1 0 9 3 やプログラムデータ 1 0 9 4 を必要に応じて RAM 1 0 1 2 に読み出して実行する。
- [0046] なお、プログラムモジュール 1 0 9 3 やプログラムデータ 1 0 9 4 は、ハードディスクドライブ 1 0 9 0 に記憶される場合に限らず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ 1 1 0 0 等を介して CPU 1 0 2

0によって読み出されてもよい。あるいは、プログラムモジュール1093及びプログラムデータ1094は、ネットワーク（LAN（Local Area Network）、WAN（Wide Area Network）等）を介して接続される他のコンピュータに記憶されてもよい。そして、プログラムモジュール1093及びプログラムデータ1094は、他のコンピュータから、ネットワークインタフェース1070を介してCPU1020によって読み出されてもよい。

符号の説明

- [0047] 10 抽出装置
- 11 入出力部
 - 12 記憶部
 - 13 制御部
 - 131 ログ収集部
 - 132 時系列化部
 - 133 DB登録部
 - 134 ルール変換部
 - 135 DB検索部
 - 136 抽出部
 - 137 判定部
 - 138 算出部
 - 139 絞り込み部

請求の範囲

[請求項1]

調査対象のコンピュータのログを収集するログ収集部と、

コンピュータへの攻撃を示す複数のシグネチャを攻撃に特徴的な順序に並べたルールを参照して、前記収集したログから、前記ルールに示されるいずれかのシグネチャにマッチするログ群を抽出する第1の抽出部と、

前記抽出されたログ群の各ログがマッチするシグネチャの時系列での並びと、前記ルールに示される複数のシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する第2の抽出部と、

前記最長共通部分列が最長となるログ群それぞれについて、前記ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する算出部と、

前記算出した分散値が最小となるログ群における前記最長共通部分列を攻撃の痕跡の候補として出力する出力処理部とを備えることを特徴とする抽出装置。

[請求項2]

前記最長共通部分列の長さは、

前記ルールに示されるいずれかのシグネチャにマッチするログ群を時系列に並び替えたときのシグネチャの並びと、前記ルールに示される複数のシグネチャの並びとの最長共通部分列の長さ

であることを特徴とする請求項1に記載の抽出装置。

[請求項3]

前記最長共通部分列が最長となるログ群が複数あるか否かを判定する判定部をさらに備え、

前記判定部により、前記最長共通部分列が最長となるログ群が複数あると判定された場合、前記算出部は、前記最長共通部分列が最長となるログ群それぞれについて、前記ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する

ことを特徴とする請求項1に記載の抽出装置。

[請求項4]

前記判定部により、前記最長共通部分列が最長となるログ群が複数

ないと判定された場合、

前記出力処理部は、前記最長共通部分列が最長となるログ群における前記最長共通部分列を攻撃の痕跡の候補として出力することを特徴とする請求項3に記載の抽出装置。

[請求項5]

前記出力処理部は、さらに、
前記算出した分散値が最小となるログ群を出力することを特徴とする請求項1に記載の抽出装置。

[請求項6]

抽出装置により実行される抽出方法であって、
調査対象のコンピュータのログを収集する工程と、
コンピュータへの攻撃を示す複数のシグネチャを攻撃に特徴的な順序に並べたルールを参照して、前記収集したログから、前記ルールに示されるいずれかのシグネチャにマッチするログ群を抽出する工程と、
、

前記抽出されたログ群の各ログがマッチするシグネチャの時系列での並びと、前記ルールに示される複数のシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する工程と、

前記最長共通部分列が最長となるログ群それぞれについて、前記ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する工程と、

前記算出した分散値が最小となるログ群における前記最長共通部分列を攻撃の痕跡の候補として出力する工程と
を含むことを特徴とする抽出方法。

[請求項7]

調査対象のコンピュータのログを収集する工程と、
コンピュータへの攻撃を示す複数のシグネチャを攻撃に特徴的な順序に並べたルールを参照して、前記収集したログから、前記ルールに示されるいずれかのシグネチャにマッチするログ群を抽出する工程と、
、
前記抽出されたログ群の各ログがマッチするシグネチャの時系列で

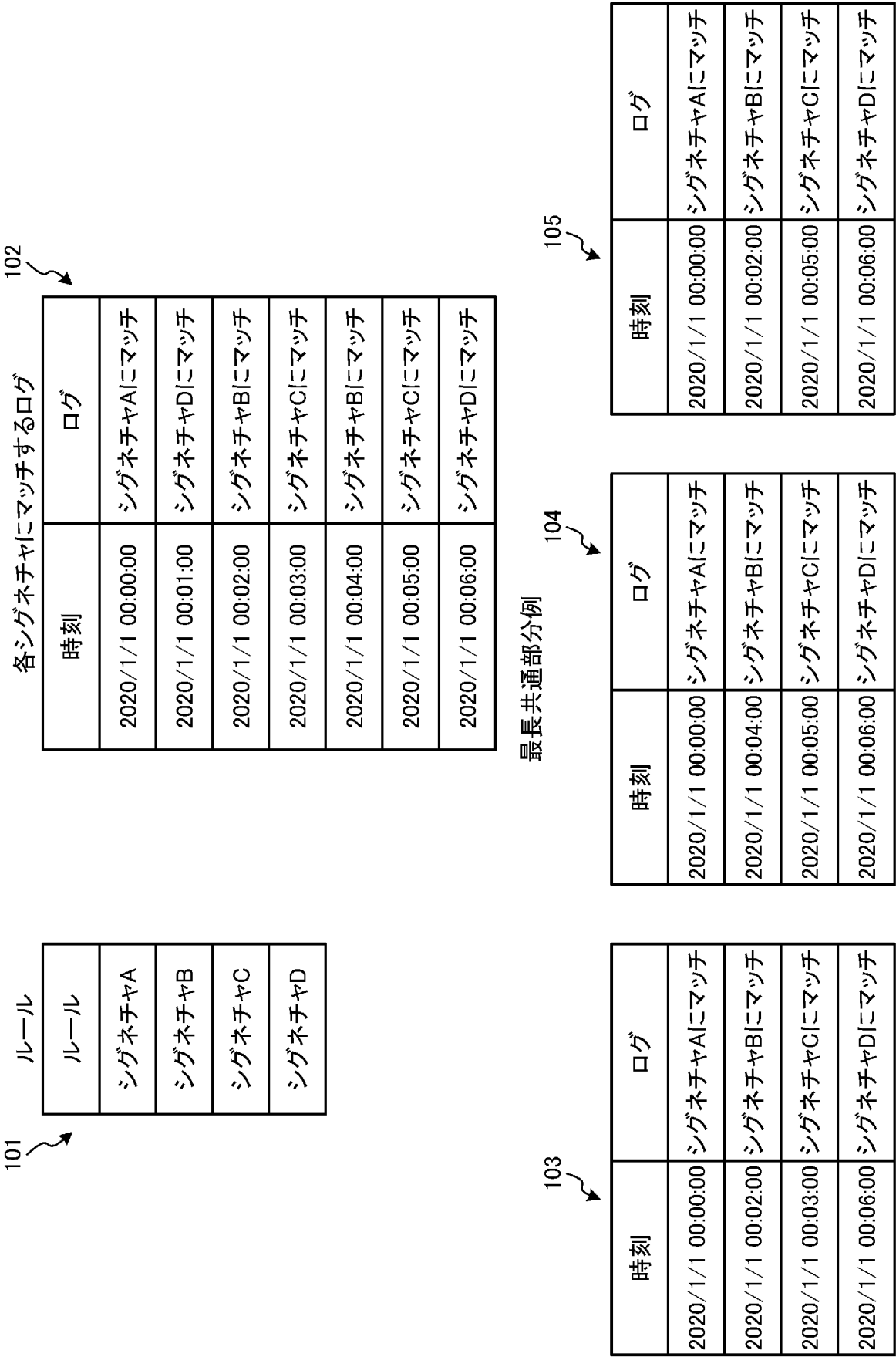
の並びと、前記ルールに示される複数のシグネチャの並びとの最長共通部分列が最長となるログ群を抽出する工程と、

前記最長共通部分列が最長となるログ群それぞれについて、前記ログ群において時系列で隣接する各ログ間の時刻差の分散値を算出する工程と、

前記算出した分散値が最小となるログ群における前記最長共通部分列を攻撃の痕跡の候補として出力する工程と

をコンピュータに実行させるための抽出プログラム。

[図1]



[図2]

最長共通部分例											
103	<table><tr><th>時刻</th><th>ログ</th></tr><tr><td>2020/1/1 00:00:00</td><td>シグネチャAにマッチ</td></tr><tr><td>2020/1/1 00:02:00</td><td>シグネチャBにマッチ</td></tr><tr><td>2020/1/1 00:03:00</td><td>シグネチャCにマッチ</td></tr><tr><td>2020/1/1 00:06:00</td><td>シグネチャDにマッチ</td></tr></table>	時刻	ログ	2020/1/1 00:00:00	シグネチャAにマッチ	2020/1/1 00:02:00	シグネチャBにマッチ	2020/1/1 00:03:00	シグネチャCにマッチ	2020/1/1 00:06:00	シグネチャDにマッチ
時刻	ログ										
2020/1/1 00:00:00	シグネチャAにマッチ										
2020/1/1 00:02:00	シグネチャBにマッチ										
2020/1/1 00:03:00	シグネチャCにマッチ										
2020/1/1 00:06:00	シグネチャDにマッチ										

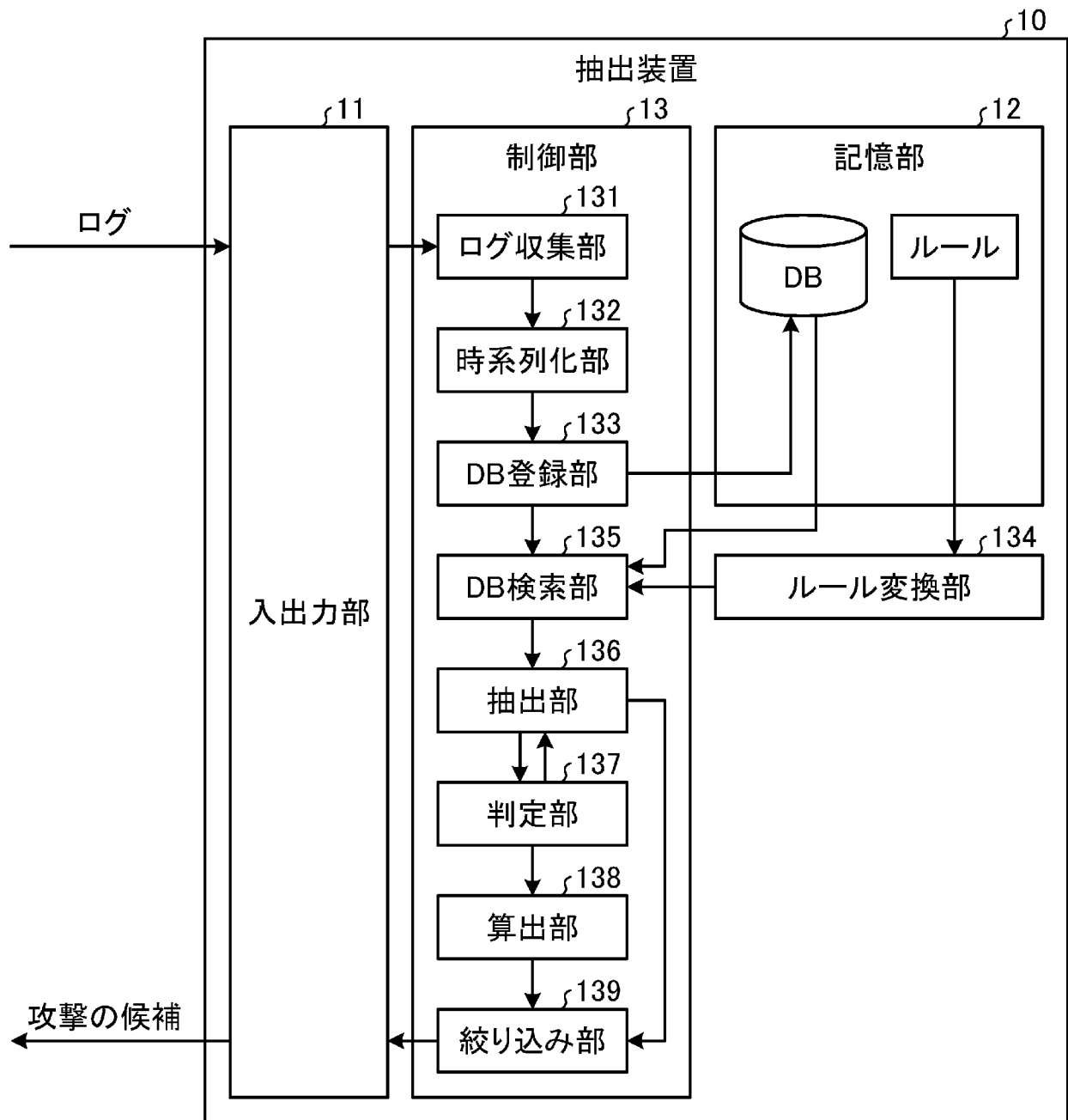
時刻差の分散値											
201	<table><tr><th>時刻</th><th>ログ</th></tr><tr><td>2020/1/1 00:00:00</td><td>シグネチャAにマッチ</td></tr><tr><td>2020/1/1 00:02:00</td><td>シグネチャBにマッチ</td></tr><tr><td>2020/1/1 00:03:00</td><td>シグネチャCにマッチ</td></tr><tr><td>2020/1/1 00:06:00</td><td>シグネチャDにマッチ</td></tr></table>	時刻	ログ	2020/1/1 00:00:00	シグネチャAにマッチ	2020/1/1 00:02:00	シグネチャBにマッチ	2020/1/1 00:03:00	シグネチャCにマッチ	2020/1/1 00:06:00	シグネチャDにマッチ
時刻	ログ										
2020/1/1 00:00:00	シグネチャAにマッチ										
2020/1/1 00:02:00	シグネチャBにマッチ										
2020/1/1 00:03:00	シグネチャCにマッチ										
2020/1/1 00:06:00	シグネチャDにマッチ										

2:00

1:00

3:00

[図3]



[図4]

シグネチャ

```

title: `wannacry service install`
detection:
  f1:
    system_evt_id: 7045
  f2:
    system_evt_service_name: `Microsoft Security Center (2.0) Service`
  f3:
    system_evt_service_file: `*mssecsvc.exe*`
condition: f1 and f2 and f3

```

[図5]

ルール

```

desc: Upjko malware activity

```

```

flow:

```

- _fid: 1
order: 1
phase: Login
file: winLogin.yml
- _fid: 2
order: 2
phase: Download malware
file: UpjkoCreated.yml
- _fid: 3
order: 2
phase: Download malware
file: clgyCreated.yml
- _fid: 4
order:
phase: Execution
file: cmdexe.yml
- _fid: 5
order: 4
phase: Execution
file: netexe.yml
- _fid: 6
order: 4
phase: Execution
file: systeminfo.yml

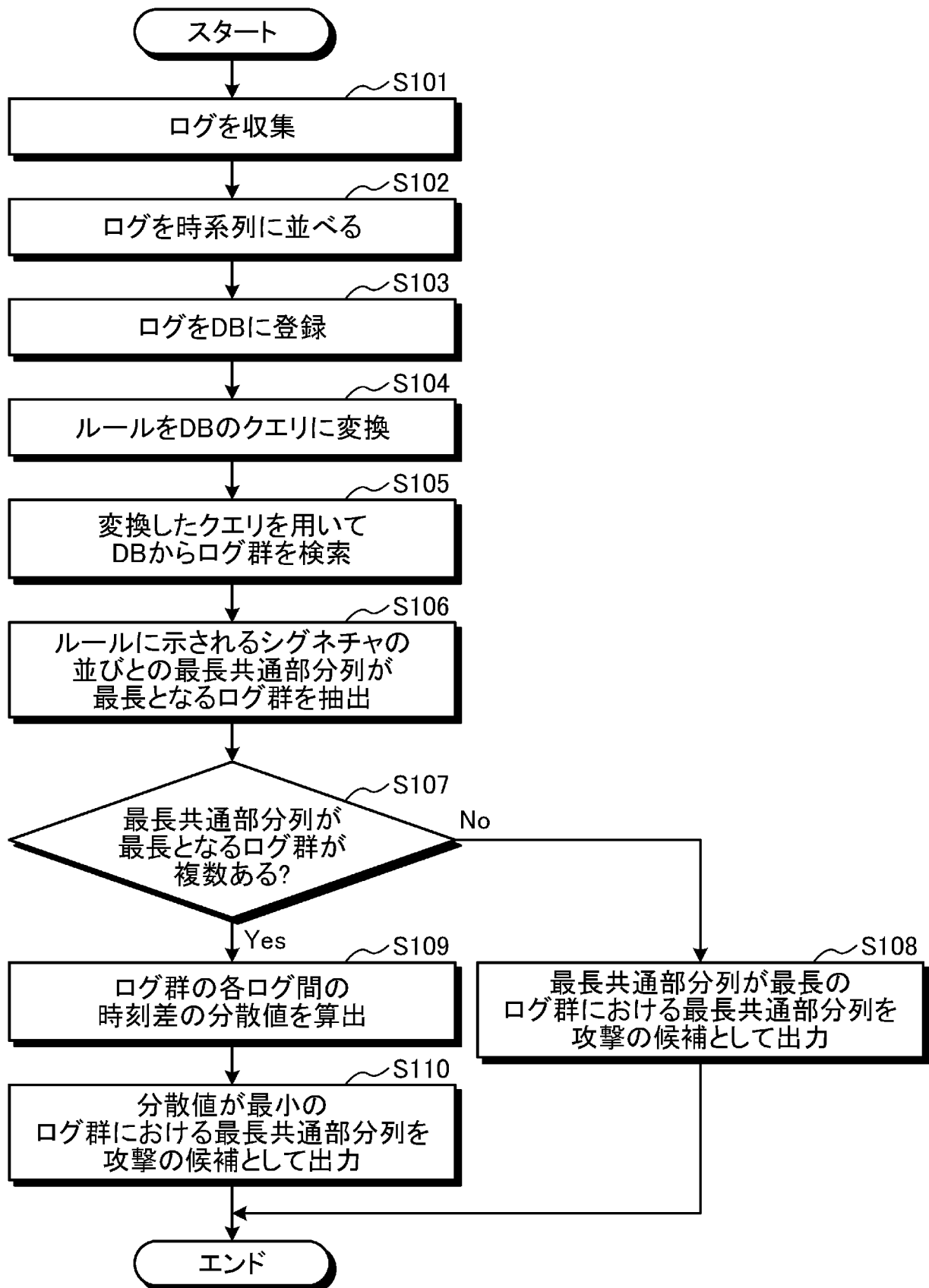
シグネチャ

```

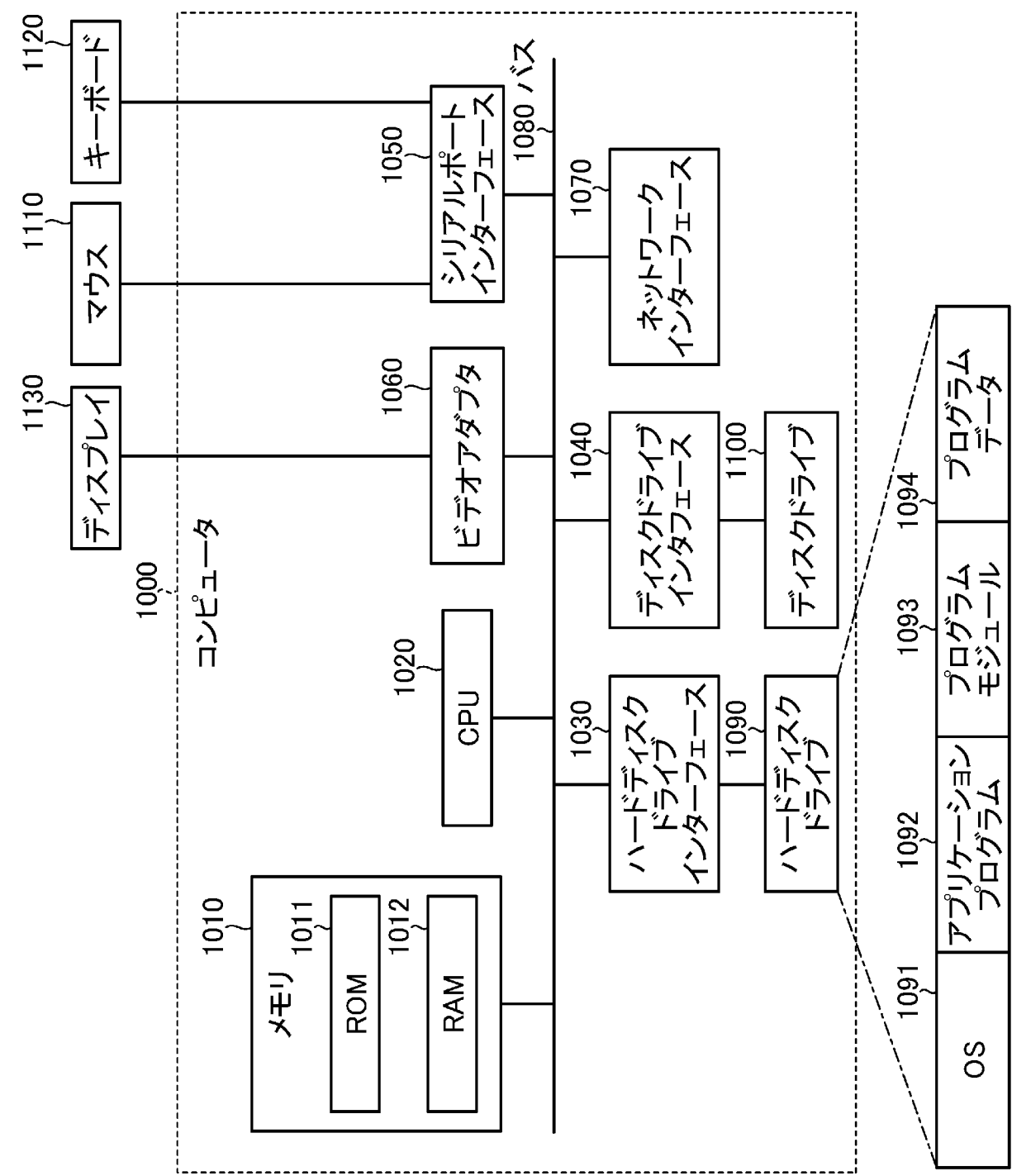
title: windows login
target: evt_system
detection:
  f1:
    target: evt_system
  f2:
    event_id: 6005
  f3:
    event_source: EventLog
condition: f1 and f2 and f3

```

[図6]



[図7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/018049

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06F21/55(2013.01)i

FI: G06F21/55

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)

Int.Cl. G06F21/55

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2016/147944 A1 (NIPPON TELEGR & TELEPH CORP.) 22 September 2016 (2016-09-22), paragraph [0032]	1 - 7
A	WO 2018/159362 A1 (NIPPON TELEGR & TELEPH CORP.) 07 September 2018 (2018-09-07), abstract	1 - 7

☐ 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

21 July 2021

Date of mailing of the international search report

03 August 2021

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2021/018049

WO 2016/147944 A1 22 September 2016 US 2018/0046800 A1
paragraph [0051]
EP 3258409 A1
CN 107408181 A

WO 2018/159362 A1 07 September 2018 US 2019/0387012 A1
abstract

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 21/55(2013.01)i FI: G06F21/55														
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） G06F21/55 最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2021年</td> </tr> </table> 国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2021年	日本国実用新案登録公報	1996 - 2021年	日本国登録実用新案公報	1994 - 2021年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2021年													
日本国実用新案登録公報	1996 - 2021年													
日本国登録実用新案公報	1994 - 2021年													
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	WO 2016/147944 A1（日本電信電話株式会社）22.09.2016（2016 - 09 - 22） 段落[0032]	1-7												
A	WO 2018/159362 A1（日本電信電話株式会社）07.09.2018（2018 - 09 - 07） [要約]	1-7												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 21.07.2021	国際調査報告の発送日 03.08.2021													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 平井 誠 5S 9071 電話番号 03-3581-1101 内線 3546													

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2021/018049

引用文献			公表日	パテントファミリー文献			公表日
WO	2016/147944	A1	22.09.2016	US	2018/0046800	A1	
				par.	[0051]		
				EP	3258409	A1	
				CN	107408181	A	
WO	2018/159362	A1	07.09.2018	US	2019/0387012	A1	
				ABSTRACT			