(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
2018年9月7日(07.09.2018)

WIPO | PCT

(10) 国際公開番号
WO 2018/159337 A1

(51) 国際特許分類:
G06F 21/55 (2013.01)

(21) 国際出願番号： PCT/JP2018/005601

(22) 国際出願日： 2018年2月16日(16.02.2018)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2017-041124 2017年3月3日(03.03.2017) JP

(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: 折原 慎吾 (ORIHARA, Shingo); 〒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). 佐藤 徹 (SATO, Tohru); 〒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). 嶋田 陽介 (SHIMADA, Yohsuke); 〒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). 鐘揚 (ZHONG, Yang); 〒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). 岩城 悠太 (IWAKI, Yuta); 〒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).

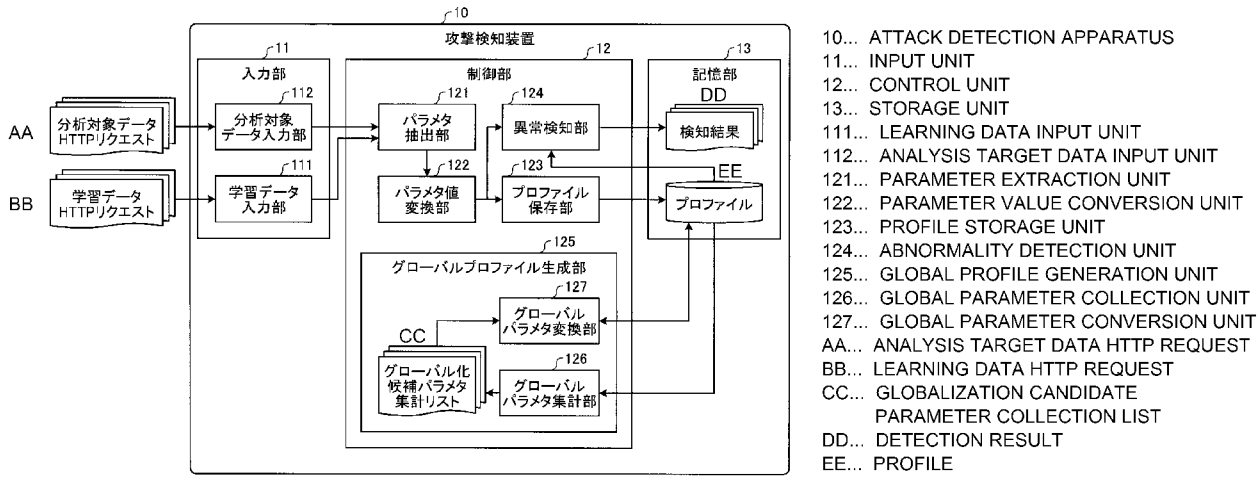
(74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関３丁目８番１号 虎の門三井ビルディング Tokyo (JP).

(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,

(54) **Title:** PROFILE GENERATION DEVICE, ATTACK DETECTION APPARATUS, PROFILE GENERATION METHOD, AND PROFILE GENERATION PROGRAM

(54) 発明の名称: プロファイル生成装置、攻撃検知装置、プロファイル生成方法、および、プロファイル生成プログラム

[図1]



(57) **Abstract:** A global profile generation unit (125) acquires profiles that each include, as an entry, information about a parameter value for a combination of a parameter name and a path included in a normal HTTP request to a web server. Then, the global profile generation unit (125) generates a global profile that aggregates entries of parameter names in the acquired profiles when entries which have different paths but have identical parameter names are present in the acquired profile.

WO 2018/159337 A1

[続葉有]

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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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:

— 国際調査報告(条約第21条(3))

(57) 要約: グローバルプロファイル生成部(１２５)は、Webサーバへの正常なHTTPリクエストに含まれるパス部とパラメタ名との組み合わせに対するパラメタ値の情報をエントリとして含むプロファイルを取得する。その後、グローバルプロファイル生成部(１２５)は、取得したプロファイルに、パス部は異なるがパラメタ名が同じエントリが存在する場合、取得したプロファイルにおける当該パラメタ名のエントリを集約したグローバルプロファイルを生成する。

明　細　書

発明の名称 ：

プロファイル生成装置、攻撃検知装置、プロファイル生成方法、および、プロファイル生成プログラム

技術分野

[0001] 本発明は、ネットワークセキュリティに関する技術であり、特に、Webサーバおよびwebアプリケーションに対して攻撃を行うアクセスに対して、アクセスを分析し、検知する技術に関する。

背景技術

[0002] インターネットの普及に伴い、Webサーバに対する攻撃が急増している。攻撃の対策としてIDS（Intrusion Detection System）、IPS（Intrusion Prevention System）、WAF（Web Application Firewall）等によるネットワークへの不正侵入検知・防御システムが知られている。これらのシステムは、ブラックリストやシグネチャファイルを用いたパターンで攻撃の検知を行うため、既知の攻撃の検知や防御しかできないという問題がある。

[0003] 一方、未知の攻撃の検知技術として、アノマリ型検知技術がある。このアノマリ型検知技術は、正常なWebサーバへのアクセス要求のパラメタ値の特徴（プロファイル）を学習しておき、検知対象のアクセス要求の特徴と比較することにより攻撃（正常なアクセス要求ではない）か否かを判定する技術である。

[0004] 例えば、このアノマリ型検知技術として、Webサーバへの正常なHTTP（Hypertext Transfer Protocol）リクエストにおけるパス部とパラメタの組み合わせごとに、パラメタ値の文字列構造を抽象化した情報（文字クラス列）をプロファイルとして学習しておき、識別対象のHTTPリクエストのパラメタ値の文字クラスとの類似度を算出することで、攻撃か否かを判定する技術がある（特許文献１参照）。このような技術によれば、プロファイルに含まれないパラメタ値や、文字列の自由度の高いパラメタ値を含む正常なHTTPリクエ

ストを攻撃として検知してしまう（誤検知してしまう）可能性を低減することができる。

先行技術文献

特許文献

[0005] 特許文献1：国際公開第２０１５／１８６６６２号

発明の概要

発明が解決しようとする課題

[0006] しかし、上記のアノマリ型検知技術は、Webサーバ（サイト）へのHTTPリクエストにおけるパス部が動的に生成されるWebアプリケーションの場合は、適切なプロファイルを生成することができない。

[0007] 例えば、動的に生成されるパス部ごとに、当該パス部のパラメタのパラメタ値を学習し、プロファイルを生成すると、当該プロファイルにおける学習対象のパラメタに対するパラメタ値の種類数が不足してしまう。そして、種類数が不足しているプロファイルで攻撃検知を行うと、正常なHTTPリクエストに対する誤検知が多発してしまう。

[0008] また、動的に生成されるパス部は、Webアプリケーションの仕様によっては、一度しか生成されない場合もある。そのため、ユーザがアクセスする度にHTTPリクエストにおけるパス部が変わることがある。このような場合、当該パス部のパラメタのパラメタ値を学習し、プロファイルを作成しても、検知の際に対応するパス部がプロファイル中に存在しない状態となり、攻撃を見逃すおそれがある。

[0009] さらに、動的に生成されるパス部に対するパラメタを、パス部ごとに学習する場合、生成されたパス部の種類数の増加に比例して、プロファイルのサイズも増大してしまう。その結果、検知対象のHTTPリクエストとプロファイルとの照合に要する時間の増大やシステムの性能低下が生じるおそれがある。

[0010] そこで、本発明は、前記した問題を解決し、リクエストにおけるパス部が

動的に生成されるWebアプリケーションに対しても、精度よく、かつ、効率よく攻撃を検知することを課題とする。

課題を解決するための手段

[0011] 前記した課題を解決するため、本発明は、検知対象となるWebサーバへのリクエストの特徴を示すプロファイルを生成するプロファイル生成装置であって、学習用データであるリクエストに含まれるパス部とパラメタとの組み合わせを含むプロファイル情報を取得する取得部と、取得したプロファイル情報群に、パス部は異なるがパラメタの名称の類似度が所定値以上のプロファイル情報が所定数以上存在する場合、前記プロファイル情報群を集約したプロファイルを生成するプロファイル生成部とを備えることを特徴とする。

発明の効果

[0012] 本発明によれば、リクエストにおけるパス部が動的に生成されるWebアプリケーションに対しても、精度よく、かつ、効率よく攻撃を検知することができる。

図面の簡単な説明

[0013] [図1]図１は、本実施形態のプロファイル生成装置を含む攻撃検知装置の構成を示す図である。

[図2]図２は、学習データからプロファイル（グローバルパラメタ変換処理後）を生成する例を示す図である。

[図3]図３は、学習データからプロファイル（グローバルパラメタ変換処理前）を生成する例を示す図である。

[図4]図４は、図１の攻撃検知装置におけるパラメタ値の文字クラス列への変換およびプロファイルへの保存を説明するための図である。

[図5]図５は、グローバルパラメタ変換処理前のプロファイルの例を示す図である。

[図6]図６は、図１の攻撃検知装置における攻撃（異常）の検知を説明するための図である。

[図7]図７は、図１のグローバルパラメタ集計部による、グローバル化候補パ

ラメタ集計リストの生成を説明するための図である。

[図8]図８は、図１のグローバルパラメタ変換部による、パラメタのグローバル化の判定を説明するための図である。

[図9]図９は、図１のグローバルパラメタ変換部による、プロファイルの更新を説明するための図である。

[図10]図１０は、図１の攻撃検知装置がグローバルプロファイルを生成する際の処理手順の例を示すフローチャートである。

[図11]図１１は、第２の実施形態の攻撃検知装置における文字クラス列の集約を説明するための図である。

[図12]図１２は、第３の実施形態の攻撃検知装置における文字クラス列の集約を説明するための図である。

[図13]図１３は、攻撃検知プログラムを実行するコンピュータを示す図である。

発明を実施するための形態

[0014] 以下、図面を参照しながら、本発明の実施形態を、第１の実施形態から第３の実施形態に分けて説明する。本発明は、以下に示す各実施形態に限定されない。

[0015] （第１の実施形態）

（概要）

図１～図３を用いて、第１の実施形態のプロファイル生成装置（グローバルプロファイル生成部１２５）を含む攻撃検知装置１０の概要を説明する。

[0016] なお、攻撃検知装置１０による、学習データおよび分析対象データ（攻撃の検知対象となるデータ）は、Webサーバ（サイト）へのHTTPリクエストである場合を例に説明する。

[0017] また、攻撃検知装置１０は、いわゆるアノマリ型検知を行う場合を例に説明する。また、攻撃検知装置１０は、検知に用いるプロファイル（プロファイル情報）として、正常なHTTPリクエストの含まれるパス部とパラメタの名称（パラメタ名）の組み合わせごとに、当該パラメタのパラメタ値を示す情

報（パラメタ値情報）を示した情報を用いる。なお、攻撃検知装置１０は、パラメタ値情報として、パラメタ値の文字列の構造を抽象化した文字クラス列（後記）を用いる場合を例に説明する。

[0018] 図１に示す攻撃検知装置１０は、まず、学習データのHTTPリクエストに基づき、正常なHTTPリクエストの特徴を示すプロファイルを生成しておく。そして、攻撃検知装置１０は、分析対象データであるHTTPリクエストの入力を受け付けると、当該HTTPリクエストとプロファイルとを比較することにより、当該HTTPリクエストが正常なHTTPリクエストか否かを判定する。

[0019] ここで、攻撃検知装置１０は、学習データである正常なHTTPリクエストに基づき、プロファイルとして、例えば、正常なHTTPリクエストに含まれるパス部とパラメタの名称の組み合わせごとに、当該パラメタのパラメタ値の文字クラス列を示す情報を生成する。

[0020] この攻撃検知装置１０は、プロファイルの生成にあたり、学習データのHTTPリクエストに、パス部は異なるが名称が同じパラメタ（グローバルパラメタ）があれば、このグローバルパラメタの文字クラス列を示した情報を生成する（図２参照）。

[0021] この文字クラス列は、パラメタ値を構成する文字それぞれを、AL：アルファベットを表す文字クラス（alpha）、NU：数字を表す文字クラス（numeric）、SY：記号を表す文字クラス（symbol）等のクラスに分類し、その分類結果を文字列の並びに従い並べたものである。

[0022] 例えば、攻撃検知装置１０は、まず、図３に示す学習データのHTTPリクエストから、HTTPリクエストの含まれるパス部とパラメタ名との組み合わせごとに、パラメタ値の文字クラス列を示したプロファイル（図３に示すグローバルパラメタ変換処理前のプロファイル）を生成する。その後、攻撃検知装置１０は、このプロファイルにおける、グローバルパラメタ（例えば、図２におけるパラメタ「file」）のエントリを集約したプロファイル（グローバルパラメタ変換処理後のプロファイル）を生成する（図２参照）。

[0023] このように、攻撃検知装置１０は、パス部は異なるが名称が同じパラメタ

のエントリを集約したプロファイルを生成する。これにより、攻撃検知装置１０は、パス部が動的に変わるHTTPリクエストを学習データに用いた場合でも、プロファイル中のパス部とパラメタ名との組み合わせに対するパラメタ値の種類数の不足を解消できる。その結果、攻撃検知装置１０は、当該プロファイルを用いた攻撃の検知における誤検知を低減できる。また、攻撃検知装置１０は、一度しか生成されないパス部に適用可能なプロファイルを生成できるので、当該プロファイルを用いることで攻撃の見逃しを低減することができる。さらに、攻撃検知装置１０は、プロファイルのサイズを低減できるので、分析対象のHTTPリクエストとプロファイルとの照合に要する時間を低減できる。

[0024] （構成）

図１を用いて、攻撃検知装置１０の構成を説明する。攻撃検知装置１０は、入力部１１と、制御部１２と、記憶部１３とを備える。

[0025] 入力部１１は、学習データ入力部（取得部）１１１と、分析対象データ入力部１１２とを有する。学習データ入力部１１１は、学習データとして、Webサーバ（図示省略）への正常なHTTPリクエストの入力を受け付ける。分析対象データ入力部１１２には、分析対象データであるHTTPリクエストの入力を受け付ける。

[0026] 制御部１２は、パラメタ抽出部１２１と、パラメタ値変換部１２２と、プロファイル保存部１２３と、異常検知部１２４と、グローバルプロファイル生成部１２５とを備える。

[0027] パラメタ抽出部１２１は、学習データのHTTPリクエストから、パス部、パラメタ名およびパラメタ値を抽出する。また、パラメタ抽出部１２１は、分析対象データのHTTPリクエストから、パス部、パラメタ名およびパラメタ値を抽出する。

[0028] パラメタ値変換部１２２は、パラメタ抽出部１２１により抽出された学習データのHTTPリクエストのパラメタ値を文字クラス列に変換する。また、パラメタ値変換部１２２は、パラメタ抽出部１２１により抽出された分析対象

データのHTTPリクエストのパラメタ値を文字クラス列に変換する。

[0029] 例えば、パラメタ値変換部１２２は、予め用意した文字クラス列の正規表現を参照して、パラメタ値に対し最長一致した部分を１つのクラスと判定し、文字クラス列への変換を行う。

[0030] 一例を挙げる。ここでは、パラメタ抽出部１２１により学習データのHTTPリクエストから抽出されたパラメタ値が「img.jpg」である場合を考える。この場合、パラメタ値変換部１２２は、図４の（１）に示すように、「img.jpg」を（img,.,jpg）というように３つの文字列に区切り、（img,.,jpg）を、（alpha,symbol,alpha）というクラスに変換する。なお、alphaはアルファベットを示し、symbolは記号を示す。

[0031] 図１のプロファイル保存部１２３は、パラメタ値変換部１２２から学習データのHTTPリクエストのパス部と、パラメタ名と、パラメタ値を示す文字クラス列とを対応付けた情報を受け取ると、この情報を記憶部１３にプロファイルとして保存する。例えば、プロファイルは、図５に示すように、パス部およびパラメタ（パラメタ名）の組み合わせごとに、当該パラメタのパラメタ値の示す文字クラス列と、出現回数とを対応付けた情報である。

[0032] 例えば、図５に示すプロファイルにおける１行目のエントリは、学習データに、パス部「/index.php」、パラメタ名「id」、パラメタ値の文字クラス列「 {AL, NU} 」のHTTPリクエストが含まれており、その数（出現回数）は「200（回）」であることを示す。

[0033] なお、図１のプロファイル保存部１２３は、同じパス部およびパラメタ名の組み合わせに対し、複数の文字クラス列が存在する場合、これらのうち出現回数または出現頻度が最も高い文字クラス列を、当該パス部およびパラメタ名の組み合わせとして対応付けた情報をプロファイルとして保存してもよい。

[0034] 一例を挙げる。例えば、プロファイル保存部１２３は、同じパス部およびパラメタ名に対し、図４の（２）に示す３つの文字クラス列が存在した場合、これらの文字クラス列のうち、出現頻度が最大（「１００」）の（alpha,s

ymbol,alpha）を当該パス部およびパラメタ名に対応付けた情報をプロファイルとして保存する。

[0035] 図１の異常検知部１２４は、グローバルプロファイル生成部１２５により生成されたプロファイル（グローバルプロファイル）を参照し、分析対象データのHTTPリクエストが正常なHTTPリクエストか否かを判定する。

[0036] 例えば、異常検知部１２４は、パラメタ値変換部１２２から分析対象データのHTTPリクエストのパス部、パラメタ名および文字クラス列を対応付けた情報を受け取る。その後、異常検知部１２４は、分析対象データのHTTPリクエストのパス部およびパラメタ名の組み合わせに対する文字クラス列と、プロファイルにおける当該パス部およびパラメタ名の組み合わせに対する文字クラス列との類似度を算出する。

[0037] その後、異常検知部１２４は、算出した類似度が所定の閾値以上であれば、分析対象データのHTTPリクエストは正常なHTTPリクエストと判定し、算出した類似度が所定の閾値未満であれば、分析対象データのHTTPリクエストは攻撃を示すHTTPリクエストと判定する。

[0038] 一例を挙げる。まず、パラメタ値変換部１２２は、パラメタ抽出部１２１により抽出された分析対象データのHTTPリクエストのパラメタ値を文字クラス列に変換する（図６の（１））。

[0039] その後、異常検知部１２４は、分析対象データのHTTPリクエストの文字クラス列と、プロファイルの文字クラス列との類似度を式（１）により算出する（図６の（２））。すなわち、異常検知部１２４は、分析対象データのHTTPリクエストのパス部およびパラメタ名の組み合わせに対する文字クラス列と、プロファイルにおける当該パス部およびパラメタ名の組み合わせに対する文字クラス列との類似度を式（１）により算出する。

[0040] 例えば、異常検知部１２４は、図６の（２）に示すように、プロファイルの文字クラス列の長さをＸ、分析対象データの文字クラス列の長さをＹ、プロファイルの文字クラス列と分析対象データの文字クラス列の最長共通部分列（LCS：Longest Common Subsequence）の長さをＺとし、式（１）により

類似度Ｓを算出する。

[0041] 類似度Ｓ＝Ｚ／（Ｘ＋Ｙ－Ｚ）…式（１）

[0042] 例えば、異常検知部１２４が、図６の（２）に示すように、プロファイルの文字クラス列（alpha,numeric,symbol,alpha）と、分析対象データの文字クラス列（alpha,symbol,numeric,symbol,alpha）との類似度を算出する場合、Ｘ＝４、Ｙ＝５、Ｚ＝４であるので、これらの値を式（１）に代入すると、類似度Ｓは「０．８」となる。

[0043] 次に、異常検知部１２４は、算出した文字クラス列の類似度Ｓが閾値Ｓｔよりも小さいか否かを判定し、類似度Ｓが閾値Ｓｔよりも小さい場合、分析対象データのHTTPリクエストを異常と判定する（図６の（３））。つまり、異常検知部１２４は、分析対象データのHTTPリクエストを攻撃と判定する。一方、類似度Ｓが閾値Ｓｔ以上である場合、分析対象データのHTTPリクエストを正常と判定する。つまり、当該HTTPリクエストを攻撃ではないと判定する。

[0044] 図１のグローバルプロファイル生成部１２５は、学習データのHTTPリクエスト群のうち、パス部は異なるがパラメタ名が同じHTTPリクエストについて、当該パラメタ名が同じHTTPリクエストを集約したプロファイル（グローバルプロファイル）を生成する。なお、以下では、グローバルプロファイル生成部１２５は、プロファイル保存部１２３により保存されたプロファイルを用いてグローバルプロファイルを生成する場合を例に説明するが、これに限定されない。例えば、グローバルプロファイル生成部１２５は、学習データのHTTPリクエスト群から、直接グローバルプロファイルを生成してもよい。

[0045] グローバルプロファイル生成部１２５は、グローバルパラメタ集計部１２６と、グローバルパラメタ変換部１２７とを備える。

[0046] グローバルパラメタ集計部１２６は、プロファイルに存在するパラメタ名ごとに、当該パラメタ名が出現したパス部の種類を集計する。例えば、グローバルパラメタ集計部１２６は、図７に示すグローバルパラメタ変換処理前のプロファイルについて、当該プロファイルに存在するパラメタ（パラメタ

名）ごとに、当該パラメタ（パラメタ名）が出現したパス部と、当該パス部の種類数とを示すグローバル化候補パラメタ集計リストを生成する。

[0047] 図１のグローバルパラメタ変換部１２７は、グローバルパラメタ集計部１２６により生成されたグローバル化候補パラメタ集計リスト（図７参照）を参照して、パラメタ名に対し、パス部の種類数が所定値以上のパラメタ名をグローバル化が必要なパラメタ名と判定する。その後、グローバルパラメタ変換部１２７は、グローバル化が必要と判定したパラメタ名について、プロファイルにおける当該パラメタ名のエントリをマージ（集約）する。

[0048] 例えば、グローバルパラメタ変換部１２７は、図８に示すグローバル化候補パラメタ集計リストにおいて、符号８０１に示すパラメタ名「file」は、パス部種類数≧パス部種類数の閾値（例えば、「２」）のため、グローバル化が必要と判定する。なお、パラメタ名「id」、「cc」は、パス部種類数＜パス部種類数の閾値（例えば、「２」）のため、グローバルパラメタ変換部１２７は、グローバル化は不要と判定する。

[0049] その後、グローバルパラメタ変換部１２７は、図９に示すように、グローバルパラメタ変換処理前のプロファイルにおけるグローバル化「要」と判定されたパラメタ名（「file」）のエントリにおける、パス部を「＊（ワイルドカード）」に変換する。また、グローバルパラメタ変換部１２７は、図９における、グローバル化「要」と判定されたパラメタ名（「file」）に対する文字クラス列（符号９０１～符号９０３に示す文字クラス列）を、符号９０１および符号９０４に示す文字クラス列に集約（マージ）する。

[0050] 例えば、グローバルパラメタ変換部１２７は、図９の符号９０１に示す文字クラス列はそのままとする（当該文字クラス列の出現回数も継承する）が、符号９０２と符号９０３に示す文字クラス列は、同じ文字クラス列なので、符号９０４に示す文字クラス列にマージする。なお、グローバルパラメタ変換部１２７は、符号９０２と符号９０３に示す文字クラス列のマージにあたり、それぞれの文字クラス列の出現回数を合算してもよい。

[0051] 上記のようにしてグローバルパラメタ変換部１２７がグローバルプロファ

イルを生成すると、生成したグローバルプロファイルを記憶部１３に保存する。その後、異常検知部１２４は、記憶部１３に保存されたグローバルファイル（プロファイル）を用いて、分析対象データのHTTPリクエストに対する異常検知を行う。

[0052] 記憶部１３は、制御部１２により生成されたプロファイルを記憶する。なお、グローバルプロファイル生成部１２５による、プロファイルのグローバルパラメタ変換処理の実行後は、グローバルプロファイル（図２参照）が記憶される。また、この記憶部１３は、さらに、分析対象データのHTTPリクエストの検知結果を記憶してもよい。

[0053] （処理手順）

次に、図１０を用いて、攻撃検知装置１０がグローバルプロファイルを生成する際の処理手順の例を説明する。なお、攻撃検知装置１０は、学習データのHTTPリクエストの入力後、プロファイル保存部１２３により、プロファイルが生成または更新されたことを契機として、グローバルプロファイルの生成を開始するものとする。また、プロファイルの生成および保存は、例えば、特許文献１に記載の処理と同様の処理手順により行われるものとする。

[0054] まず、グローバルパラメタ集計部１２６は、記憶部１３からプロファイル（図２参照）を取得し（Ｓ１）、グローバル化候補パラメタ集計リスト（図８参照）を初期化する（Ｓ２）。Ｓ２の後、グローバルパラメタ集計部１２６は、取得したプロファイルの先頭行のパス部、パラメタ名、文字クラス列、および、出現回数を抽出する（Ｓ３：パス部・パラメタ情報を抽出）。Ｓ３の後、グローバルパラメタ集計部１２６は、Ｓ３で取得したプロファイルのパラメタ名がグローバルパラメタへの変換の候補か否かを判定する（Ｓ４）。

[0055] 例えば、Ｓ４において、グローバルパラメタ集計部１２６が、プロファイルから取り出した行のパス部が「＊」である場合、当該行は既にグローバルパラメタに変換されているのでグローバルパラメタへの変換の候補ではないと判定し（Ｓ４でＮｏ）、Ｓ３に戻る。そして、グローバルパラメタ集計部

１２６は、Ｓ１で取得したプロファイルの次の行の処理を行う。一方、グローバルパラメタ集計部１２６が、プロファイルから取り出した行のパス部が「＊」以外である場合（Ｓ４でＹｅｓ）、当該行をグローバルパラメタへの変換の候補であると判断し、Ｓ５へ進む。

[0056] Ｓ４でＹｅｓの場合、グローバルパラメタ集計部１２６は、グローバル化候補パラメタ集計リスト（図８参照）の更新を行う（Ｓ５）。例えば、グローバルパラメタ集計部１２６は、Ｓ４で、グローバルパラメタへの変換の候補と判断した行に含まれるパラメタ名およびパス部の組み合わせが、グローバル化候補パラメタ集計リストに既に存在するかを確認し、確認結果に応じて、グローバル化候補パラメタ集計リストに対し、以下の（ａ）、（ｂ）に示すいずれかの更新処理を行う。

[0057] （ａ）グローバル化候補パラメタ集計リストに、当該パラメタ名も当該パス部も存在しない場合、グローバルパラメタ集計部１２６は、グローバル化候補パラメタ集計リストに当該パラメタ名および当該パス部の組み合わせのエントリを新規作成する。そして、グローバルパラメタ集計部１２６は、当該エントリにおけるパス部種類数を「１」にする。

[0058] （ｂ）グローバル化候補パラメタ集計リストに、当該パラメタ名は存在するが、当該パス部は存在しない場合、グローバルパラメタ集計部１２６は、当該パス部を、グローバル化候補パラメタ集計リストの当該パラメタ名のエントリに追加する。そして、グローバルパラメタ集計部１２６は、当該エントリにおけるパス部種類数の値を加算する。

[0059] なお、グローバル化候補パラメタ集計リストに、既に、当該パラメタ名および当該パス部名の組み合わせが存在する場合、グローバルパラメタ集計部１２６は、当該パラメタ名部も当該パス部も、グローバル化候補パラメタ集計リストに追加しない。

[0060] Ｓ５の後、グローバルパラメタ集計部１２６は、Ｓ１で取得したプロファイルに未処理の行があれば（Ｓ６でＮｏ）、Ｓ３へ戻り、Ｓ１で取得したプロファイルに未処理の行がなければ（Ｓ６でＹｅｓ）、Ｓ７へ進み、グロー

バル化候補パラメタ集計リストを出力する（Ｓ７）。

[0061] Ｓ７の後、グローバルパラメタ変換部１２７は、グローバル化候補パラメタ集計リストを取得する（Ｓ８）。その後、グローバルパラメタ変換部１２７は、グローバル化候補パラメタ集計リストのパラメタ名、パス部およびパス部種類数を抽出する（Ｓ９：パラメタ・パス部情報を抽出）。そして、グローバルパラメタ変換部１２７は、抽出したパラメタ名に対するパス部種類の数が所定の閾値以上であれば、当該パラメタ名をグローバル化が必要なパラメタ名と判定し（Ｓ１０でＹｅｓ）、Ｓ１１へ進む。つまり、グローバル化候補パラメタ名集計リストに、当該パラメタ名に対応する所定の閾値以上の種類のパス部が存在する場合、グローバルパラメタ名変換部１２７は、当該パラメタ名をグローバル化が必要なパラメタ名と判定する。

[0062] 一方、グローバルパラメタ変換部１２７は、抽出したパラメタ名に対するパス部種類の数が所定の閾値未満であれば、当該パラメタ名をグローバル化が必要なパラメタ名と判定せず（Ｓ１０でＮｏ）、グローバル化候補パラメタ集計リストの次の行に、Ｓ９の処理を実行する。

[0063] 例えば、パス部種類の閾値＝２の場合、グローバルパラメタ変換部１２７は、図８に示すグローバル化候補パラメタ集計リストのうち、パラメタ名「id」、「cc」については、グローバル化は不要と判定し、パラメタ名「file」については、グローバル化が必要と判定する。

[0064] 図１０のＳ１０でＹｅｓの場合、グローバルパラメタ変換部１２７は、プロファイルを更新する（Ｓ１１）。具体的には、グローバルパラメタ変換部１２７は、プロファイルにおいてグローバル化が必要なパラメタ名について、当該パラメタ名のエントリをマージする。

[0065] 例えば、グローバルパラメタ変換部１２７は、当該パラメタ名のエントリにおけるパス部を「＊」に変換する。また、グローバルパラメタ変換部１２７は、当該パラメタ名のエントリにおける文字クラス列を集約する。

[0066] 例えば、グローバルパラメタ変換部１２７は、図９に示すように、グローバルパラメタ変換処理前のプロファイルにおけるパラメタ名「file」に対す

るパス部を「＊」に変換し、パラメタ名「file」に対する符号９０１～９０３に示す文字クラス列を、符号９０１に示す文字クラス列と符号９０４に示す文字クラス列に集約する。

[0067] なお、グローバルパラメタ変換部１２７は、プロファイルにおけるグローバルパラメタへの変換対象となるエントリにおける文字クラス列の出現回数については、以下のように処理する。すなわち、グローバルパラメタ変換部１２７は、グローバルパラメタへの変換対象となるエントリ間で、同じ文字クラス列があるものについては、当該文字クラス列の出現回数を合算し、グローバルパラメタへの変換対象となるエントリ間で、同じ文字クラス列がないものについては、当該文字クラス列の出現回数を継承する。

[0068] 例えば、グローバルパラメタ変換部１２７は、図９に示すように、グローバルパラメタ変換処理前のプロファイルにおける符号９０１に示す文字クラス列の出現回数は、グローバルパラメタ変換処理後のプロファイルでもそのまま継承する。一方、グローバルパラメタ変換処理前のプロファイルにおける符号９０２，９０３に示す文字クラス列の出現回数は合算し、グローバルパラメタ変換処理後のプロファイルにおいて符号９０４に示す出現回数とする。

[0069] 図１０のＳ１１の後、グローバルパラメタ変換部１２７は、Ｓ８で取得したグローバル化候補パラメタ集計リストに未処理の行があれば（Ｓ１２でＮｏ）、Ｓ９へ戻り、Ｓ８で取得したプロファイルに未処理の行がなければ（Ｓ１２でＹｅｓ）、Ｓ１３へ進み、グローバル化候補パラメタ集計リストを削除する（Ｓ１３）。

[0070] 以上の処理により、攻撃検知装置１０は、プロファイルにパス部は異なるがパラメタ名が同じエントリを集約したプロファイル（グローバルプロファイル）を生成する。よって、例えば、攻撃検知装置１０は、学習データとしてパス部が動的に変わるHTTPリクエストを用いてプロファイルを生成した場合でも、プロファイル中のパラメタ名に対するパラメタ値の種類数の不足を抑制できる。その結果、攻撃検知装置１０は、当該プロファイルを用いて、

攻撃の検知を行う場合における、誤検知を低減できる。

[0071] また、攻撃検知装置１０は、例えば、一度しか生成されないパス部に適用可能なプロファイルを生成できるので、当該プロファイルを用いることで攻撃の見逃しを低減することができる。さらに、攻撃検知装置１０は、プロファイルのサイズを低減できるので、当該プロファイルを用いて、攻撃の検知を行う場合における、分析対象のHTTPリクエストとプロファイルとの照合に要する時間を低減できる。

[0072]（第２の実施形態）

なお、図１０のＳ１１においてグローバルパラメタ変換部１２７は、同じパラメタ名に対応付けられた文字クラス列を集約する際、同じ文字クラス列のみならず、類似度が所定の閾値以上の文字クラス列を対象として集約してもよい。ここでの類似度の算出は、例えば、文字クラス列の編集距離を用いる。また、文字クラス列の集約する方法としては、最長共通部分文字列を用いる。

[0073] 例えば、グローバルパラメタ変換部１２７が、類似度の算出に、文字クラス列の編集距離を用い、編集距離が「３」未満の文字クラス列群を集約の対象とする場合を考える。この場合、図１１に示すように、グローバルパラメタ変換処理前のプロファイルにおけるパラメタ名「file」に対する符号１１０１に示す文字クラス列と符号１１０２に示す文字クラス列との編集距離は「２」である。

[0074] したがって、グローバルパラメタ変換部１２７は、符号１１０１に示す文字クラス列と符号１１０２に示す文字クラス列とは類似していると判断し、これらの文字クラス列を集約する。なお、文字クラス列の集約にあたり、最長共通部分文字列を用いる場合、グローバルパラメタ変換部１２７は、符号１１０１に示す文字クラス列および符号１１０２に示す文字クラス列について、これらの文字クラス列の最長共通部分列である｛AL,NU,SY,AL｝に集約する。また、グローバルパラメタ変換部１２７は、符号１１０１に示す文字クラス列と符号１１０２に示す文字クラス列の出現回数を合算し、符号１１０

３に示す出現回数（「３」）とする。

[0075] このようにすることで、攻撃検知装置１０は、プロファイル（グローバルプロファイル）のサイズをさらに低減できるので、当該プロファイルを用いて、攻撃の検知を行う場合における、分析対象のHTTPリクエストとプロファイルとの照合に要する時間を低減できる。

[0076] （第３の実施形態）

また、グローバルパラメタ変換部１２７は、グローバルプロファイルの生成にあたり、出現回数や出現率の低い文字クラス列を集約の対象外としてもよい。例えば、グローバルパラメタ変換部１２７は、第１の実施形態または第２の実施形態において、集約の候補となる文字クラス列それぞれについて、集約の候補となる文字クラス列群における相対的な出現率を求める。そして、グローバルパラメタ変換部１２７は、出現率が所定の閾値よりも低い文字クラス列を、集約の対象外とする。

[0077] 例えば、グローバルパラメタ変換部１２７は、集約の候補となる文字クラス列群での相対的な出現率が、５％未満の文字クラス列を集約の対象外とする場合を考える。この場合、図１２に示すように、グローバルパラメタ変換部１２７は、グローバルパラメタ変換処理前のプロファイルにおけるパラメタ名「file」に対する符号１２０１～符号１２０３に示す文字クラス列のうち、符号１２０３に示す文字クラス列は集約の候補となる文字クラス列群における出現率が５％未満なので、集約の対象外とする。一方、符号１２０１に示す文字クラス列と符号１２０２に示す文字クラス列は、集約の候補となる文字クラス列群における出現率が５％以上なので、集約の対象とする。

[0078] このようにグローバルパラメタ変換部１２７は、グローバルプロファイルの生成にあたり、出現回数の低い文字クラス列を集約の対象外とする。これにより、攻撃検知装置１０は、誤入力されたパラメタ値を持つHTTPリクエストや攻撃を示すHTTPリクエストの学習の結果がプロファイルに含まれている場合でも、当該学習の結果をグローバルプロファイルに反映させないようにすることができる。その結果、攻撃検知装置１０は、当該グローバルプロフ

ァイルを用いた攻撃の検知の精度を向上させることができる。

[0079] なお、グローバルパラメタ変換部１２７は、第１の実施形態または第２の実施形態と同様のグローバルプロファイルを生成した後で、上記のように、出現回数や出現率の低い文字クラス列の情報を削除してもよい。

[0080] （その他の実施形態）

なお、グローバルパラメタ変換部１２７は、グローバルプロファイルの生成にあたり、パス部は異なるがパラメタ名の類似度が所定値以上のエントリが存在する場合、当該エントリも集約の対象としてもよい。このようにすることで、攻撃検知装置１０は、プロファイル（グローバルプロファイル）のサイズをさらに低減できる。

[0081] なお、攻撃検知装置１０は、学習データおよび分析対象データとして、Webサーバ（サイト）へのHTTPリクエストを用いる場合を例に説明したが、HTTPリクエスト以外のアクセス要求を用いてももちろんよい。

[0082] さらに、攻撃検知装置１０は、上記のようなパス部は異なるがパラメタ名が同じエントリの集約を、いわゆるシグネチャ型検知に用いられるプロファイルの生成に適用してもよい。例えば、攻撃検知装置１０は、学習データとして悪性の通信に用いられるHTTPリクエストを取得し、このHTTPリクエストからパス部は異なるがパラメタ名が同じHTTPリクエストにおけるパラメタ値（あるいはパラメタ値を示す文字クラス列）を集約し、プロファイルを生成する。そして、攻撃検知装置１０は、分析対象データのHTTPリクエストに対し、当該プロファイルを用いてシグネチャ型検知を行う。

[0083] （プログラム）

また、各実施形態で述べた攻撃検知装置１０の機能を実現するプログラムを所望の情報処理装置（コンピュータ）にインストールすることによって実装できる。例えば、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提供される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、情報処理装置を攻撃検知装置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、デスクトップ型また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 Assistants）等がその範疇に含まれる。また、攻撃検知装置１０を、クラウドサーバに実装してもよい。

[0084] 以下に、上記のプログラム（攻撃検知プログラム）を実行するコンピュータの一例を説明する。図１３は、攻撃検知プログラムを実行するコンピュータを示す図である。図１３に示すように、コンピュータ１０００は、例えば、メモリ１０１０と、ＣＰＵ（Central Processing Unit）１０２０と、ハードディスクドライブインタフェース１０３０と、ディスクドライブインタフェース１０４０と、シリアルポートインタフェース１０５０と、ビデオアダプタ１０６０と、ネットワークインタフェース１０７０とを有する。これらの各部は、バス１０８０によって接続される。

[0085] メモリ１０１０は、ＲＯＭ（Read Only Memory）１０１１およびＲＡＭ（Random Access Memory）１０１２を含む。ＲＯＭ１０１１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１０３０は、ハードディスクドライブ１０９０に接続される。ディスクドライブインタフェース１０４０は、ディスクドライブ１１００に接続される。ディスクドライブ１１００には、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挿入される。シリアルポートインタフェース１０５０には、例えば、マウス１１１０およびキーボード１１２０が接続される。ビデオアダプタ１０６０には、例えば、ディスプレイ１１３０が接続される。

[0086] ここで、図１３に示すように、ハードディスクドライブ１０９０は、例えば、ＯＳ１０９１、アプリケーションプログラム１０９２、プログラムモジュール１０９３およびプログラムデータ１０９４を記憶する。前記した実施形態で説明した各種データや情報は、例えばハードディスクドライブ１０９０やメモリ１０１０に記憶される。

[0087] そして、ＣＰＵ１０２０が、ハードディスクドライブ１０９０に記憶され

たプログラムモジュール１０９３やプログラムデータ１０９４を必要に応じてＲＡＭ１０１２に読み出して、上述した各手順を実行する。

[0088] なお、上記の攻撃検知プログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ハードディスクドライブ１０９０に記憶される場合に限られず、例えば、着脱可能な記憶媒体に記憶されて、ディスクドライブ１１００等を介してＣＰＵ１０２０によって読み出されてもよい。あるいは、上記のプログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ＬＡＮ（Local Area Network）やＷＡＮ（Wide Area Network）等のネットワークを介して接続された他のコンピュータに記憶され、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

[0089] １０ 攻撃検知装置

１１ 入力部

１２ 制御部

１３ 記憶部

１１１ 学習データ入力部

１１２ 分析対象データ入力部

１２１ パラメタ抽出部

１２２ パラメタ値変換部

１２３ プロファイル保存部

１２４ 異常検知部

１２５ グローバルプロファイル生成部

１２６ グローバルパラメタ集計部

１２７ グローバルパラメタ変換部

請求の範囲

[請求項1] 検知対象となるWebサーバへのリクエストの特徴を示すプロファイルを生成するプロファイル生成装置であって、

学習用データであるリクエストに含まれるパス部とパラメタとの組み合わせを含むプロファイル情報を取得する取得部と、

取得したプロファイル情報群に、パス部は異なるがパラメタの名称の類似度が所定値以上のプロファイル情報が所定数以上存在する場合、前記プロファイル情報群を集約したプロファイルを生成するプロファイル生成部とを備えることを特徴とするプロファイル生成装置。

[請求項2] 前記プロファイル生成部は、

取得したプロファイル情報群に、パス部は異なるがパラメタの名称が同じプロファイル情報群が所定数以上存在する場合、前記パラメタの名称が同じプロファイル情報群を集約の対象とすることを特徴とする請求項１に記載のプロファイル生成装置。

[請求項3] 前記プロファイル生成部は、

前記パラメタの名称の類似度が所定値以上のプロファイル情報群を集約する際、前記プロファイル情報群に含まれるパラメタ値情報同士の類似度が所定値以上のパラメタ値情報群を集約することを特徴とする請求項１に記載のプロファイル生成装置。

[請求項4] 前記プロファイル生成部は、

前記プロファイル情報群に含まれるパラメタ値情報同士の類似度として、パラメタ値の文字列構造を抽象化した文字クラス列同士の類似度を用いることを特徴とする請求項３に記載のプロファイル生成装置。

[請求項5] 前記プロファイル生成部は、

前記パラメタの名称の類似度が所定値以上のプロファイル情報群を集約する際、前記プロファイルにおける、集約されるプロファイル情報群におけるパス部として、ワイルドカードを設定することを特徴と

する請求項１に記載のプロファイル生成装置。

[請求項6] 前記プロファイル生成部は、

前記パラメタの名称の類似度が所定値以上のプロファイル情報群を集約する際、前記プロファイル情報群におけるパラメタ値情報の出現回数または出現率が所定値未満のパラメタ値情報を集約の対象外とすることを特徴とする請求項３に記載のプロファイル生成装置。

[請求項7] 前記プロファイルは、

リクエストに含まれるパス部と、パラメタの名称と、前記パラメタのパラメタ値情報とを対応付けた情報であることを特徴とする請求項１に記載のプロファイル生成装置。

[請求項8] 検知対象となるWebサーバへのリクエストの特徴を示すプロファイルを用いて攻撃の検知を行う攻撃検知装置であって、

学習用データであるリクエストに含まれるパス部とパラメタとの組み合わせを含むプロファイル情報を取得する第１の取得部と、

取得したプロファイル情報群に、パス部は異なるがパラメタの名称の類似度が所定値以上のプロファイル情報群が存在する場合、前記プロファイル情報群を集約したプロファイルを生成するプロファイル生成部と、

攻撃の検知対象となるリクエストを取得する第２の取得部と、

前記攻撃の検知対象となるリクエストにおけるパス部およびパラメタと、前記プロファイルに示されるパス部とパラメタとの組み合わせとを比較することにより、前記リクエストが攻撃か否かを検知する検知部とを備えることを特徴とする攻撃検知装置。

[請求項9] 検知対象となるWebサーバへのリクエストの特徴を示すプロファイルを生成するプロファイル生成装置を用いたプロファイル生成方法であって、

学習用データであるリクエストに含まれるパス部とパラメタとの組み合わせを含むプロファイル情報を取得するステップと、

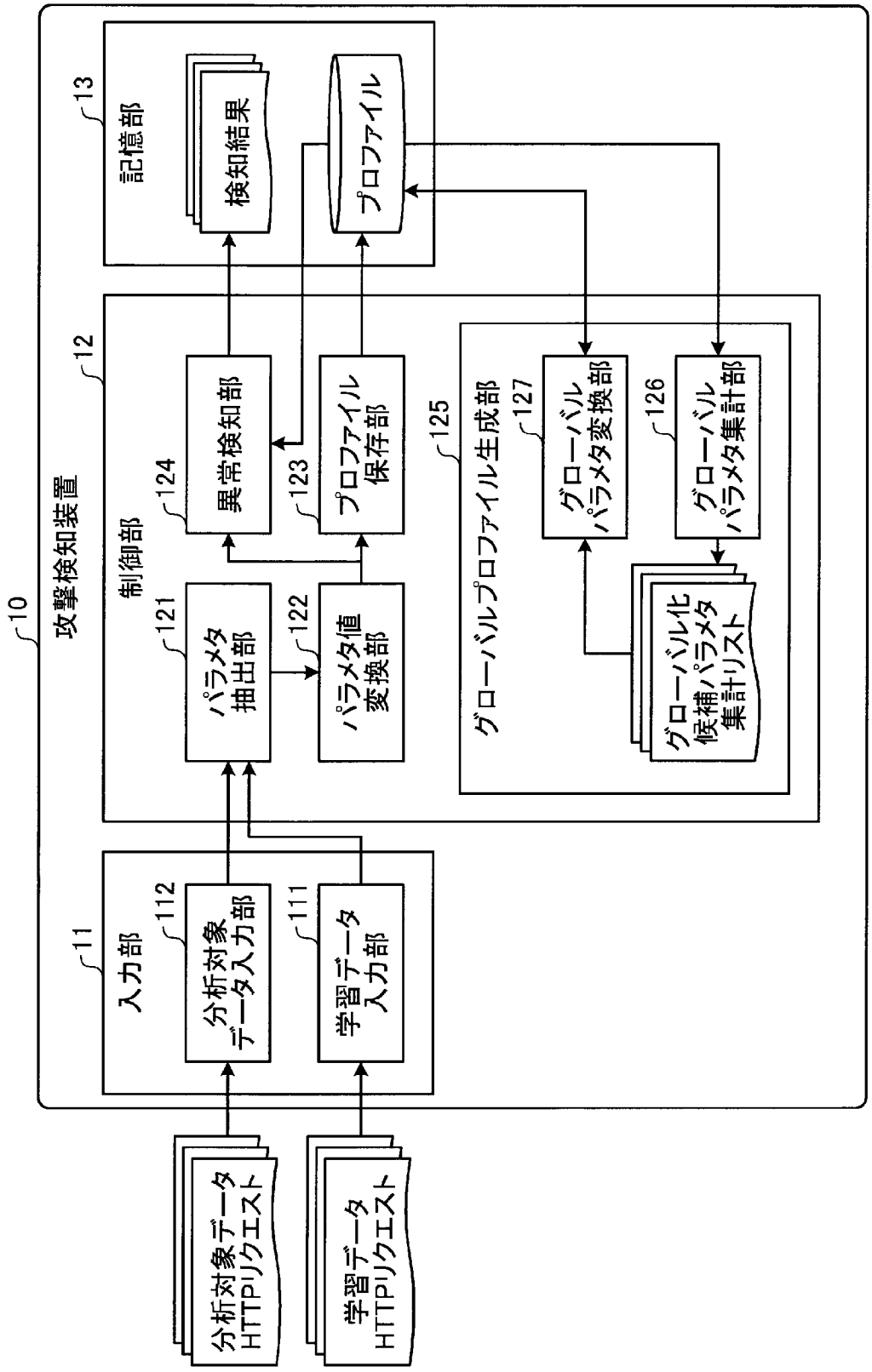
取得したプロファイル情報群に、パス部は異なるがパラメタの名称の類似度が所定値以上のプロファイル情報が所定数以上存在する場合、前記プロファイル情報群を集約したプロファイルを生成するステップとを含んだことを特徴とするプロファイル生成方法。

[請求項10] 検知対象となるWebサーバへの正常なリクエストの特徴を示すプロファイルを生成するためのプロファイル生成プログラムであって、

学習用データであるリクエストに含まれるパス部とパラメタとの組み合わせを含むプロファイル情報を取得するステップと、

取得したプロファイル情報群に、パス部は異なるがパラメタの名称の類似度が所定値以上のプロファイル情報が所定数以上存在する場合、前記プロファイル情報群を集約したプロファイルを生成するステップとをコンピュータに実行させることを特徴とするプロファイル生成プログラム。

[図1]


10
攻撃検知装置
11
入力部
112
分析対象データ入力部
111
学習データ入力部
12
制御部
121
パラメタ抽出部
122
パラメタ値変換部
124
異常検知部
123
プロファイル保存部
125
グローバルプロファイル生成部
127
グローバルパラメタ変換部
126
グローバルパラメタ集計部
グローバル化候補パラメタ集計リスト
13
記憶部
検知結果
プロファイル
分析対象データHTTPリクエスト
学習データHTTPリクエスト

[図2]

学習データのHTTPリクエスト

```
12/1 0:01:30  192.168.0.1 GET /index.php?id=user001
12/1 0:01:34  192.168.0.1 GET /index.php?id=user001&cc=jp
12/1 0:01:30  192.168.0.1 GET /dynamic001.php?file=img001.jpg
12/1 0:01:34  192.168.0.1 GET /dynamic002.php?file=img002-a.jpg
```

プロファイル(グローバルパラメタ変換処理後)

パス部	パラメタ	文字クラス列	出現回数
/index.php	id	{AL, NU}	2
/index.php	cc	{AL}	1
*	file	{AL, NU, SY, AL}	1
		{AL, NU, SY, AL, SY, AL}	1

[図3]

学習データのHTTPリクエスト

```
12/1 0:01:30  192.168.0.1 GET /index.php?id=user001
12/1 0:01:34  192.168.0.1 GET /index.php?id=user001&cc=jp
12/1 0:01:30  192.168.0.1 GET /dynamic001.php?file=img001.jpg
12/1 0:01:34  192.168.0.1 GET /dynamic002.php?file=img002-a.jpg
```

プロファイル(グローバルパラメタ変換処理前)

パス部	パラメタ	文字クラス列	出現回数
/index.php	id	{AL, NU}	2
/index.php	cc	{AL}	1
/dynamic001.php	file	{AL, NU, SY, AL}	1
/dynamic002.php	file	{AL, NU, SY, AL, SY, AL}	1

[図4]

(1)

文字クラス列の正規表現例

```
"numeric" : "[0-9]+", "hex" : "[0-9a-fA-F]+", "alpha" : "[a-zA-Z]+",
"space" : "[ ]+", "quote" : "[¥"¥']+", "pair" : "[¥(¥)¥{¥}¥[¥]<>]+",
"symbol" : "[¥!¥#¥$¥%¥&¥*¥+¥,¥-¥.¥/¥:¥;¥=¥?¥@¥^¥_¥|¥~]+",
"control" : "[¥x00-¥x1f¥x7f]+",
"noascii" : "[¥x80-¥xff]+", "multi_byte" : u"[¥u0100-¥uffff]+",
“url” : “http://[0-9a-z¥.]+”, “ip” : , “path” : , “mac” : , ...
```

img.jpg
top001.png
⇒ (img,.jpg)
(top,001,.,png)
⇒ (alpha, symbol, alpha)
(alpha, numeric, symbol, alpha)

(2)

出現頻度

(alpha, symbol, alpha) : 100
(alpha, numeric, symbol, alpha) : 50
(alpha) : 1

→ プロファイル
(alpha, symbol, alpha)

[図5]

プロファイルの例

パス部	パラメタ	文字クラス列	出現回数
/index.php	id	{AL, NU}	200
		{AL}	100
/index.php	cc	{AL}	100
/dynamic001.php	file	{AL, NU, SY, AL}	1
/dynamic002.php	file	{AL, NU, SY, AL}	1
⋮	⋮	⋮	⋮

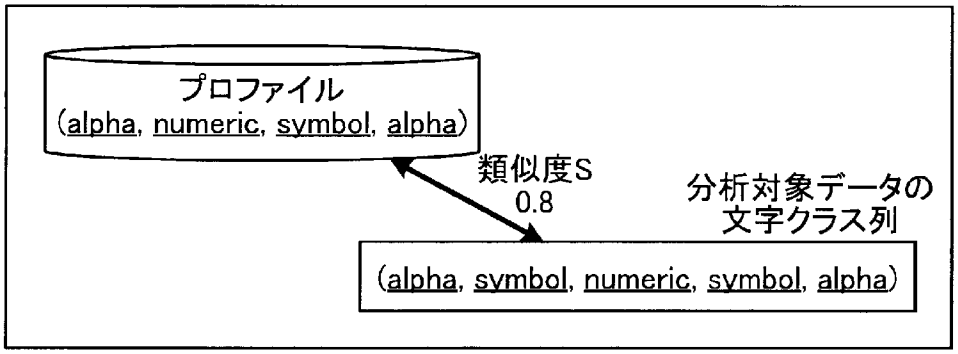
AL: アルファベットを表す文字クラス(alpha)
NU: 数字を表す文字クラス(numeric)
SY: 記号を表す文字クラス(symbol)

[図6]

(1) 分析対象データのHTTPリクエストのパラメタ値を文字クラス列に変換する
(2) プロファイルの文字クラス列との類似度を式(1)により算出する

X: プロファイルの文字クラス列の長さ
Y: 分析対象データの文字クラス列の長さ
Z: プロファイルの文字クラス列と分析対象データの
文字クラス列の最長共通部分列(LCS)の長さ

$$\text{類似度}S=\frac{Z}{X+Y-Z} \quad \cdots\text{式(1)}$$



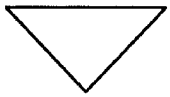
プロファイルとの類似度の計算

(3) 類似度Sが閾値Stより小さい場合、異常と判定

[図7]

プロファイル(グローバルパラメタ変換処理前)

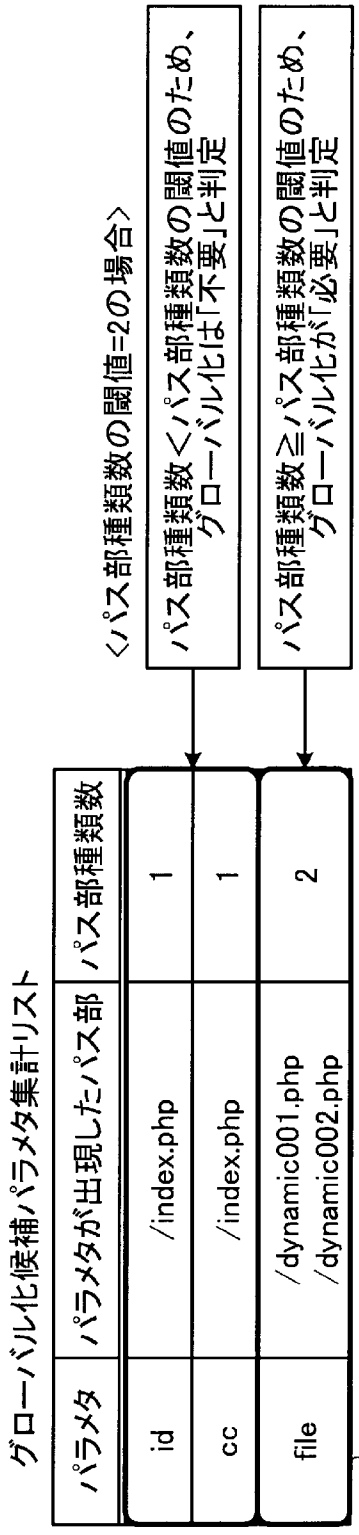
パス部	パラメタ	文字クラス列	出現回数
/index.php	id	{AL, NU}	2
/index.php	cc	{AL}	1
/dynamic001.php	file	{AL, NU, SY, AL}	1
/dynamic002.php	file	{AL, NU, SY, AL, SY, AL}	1



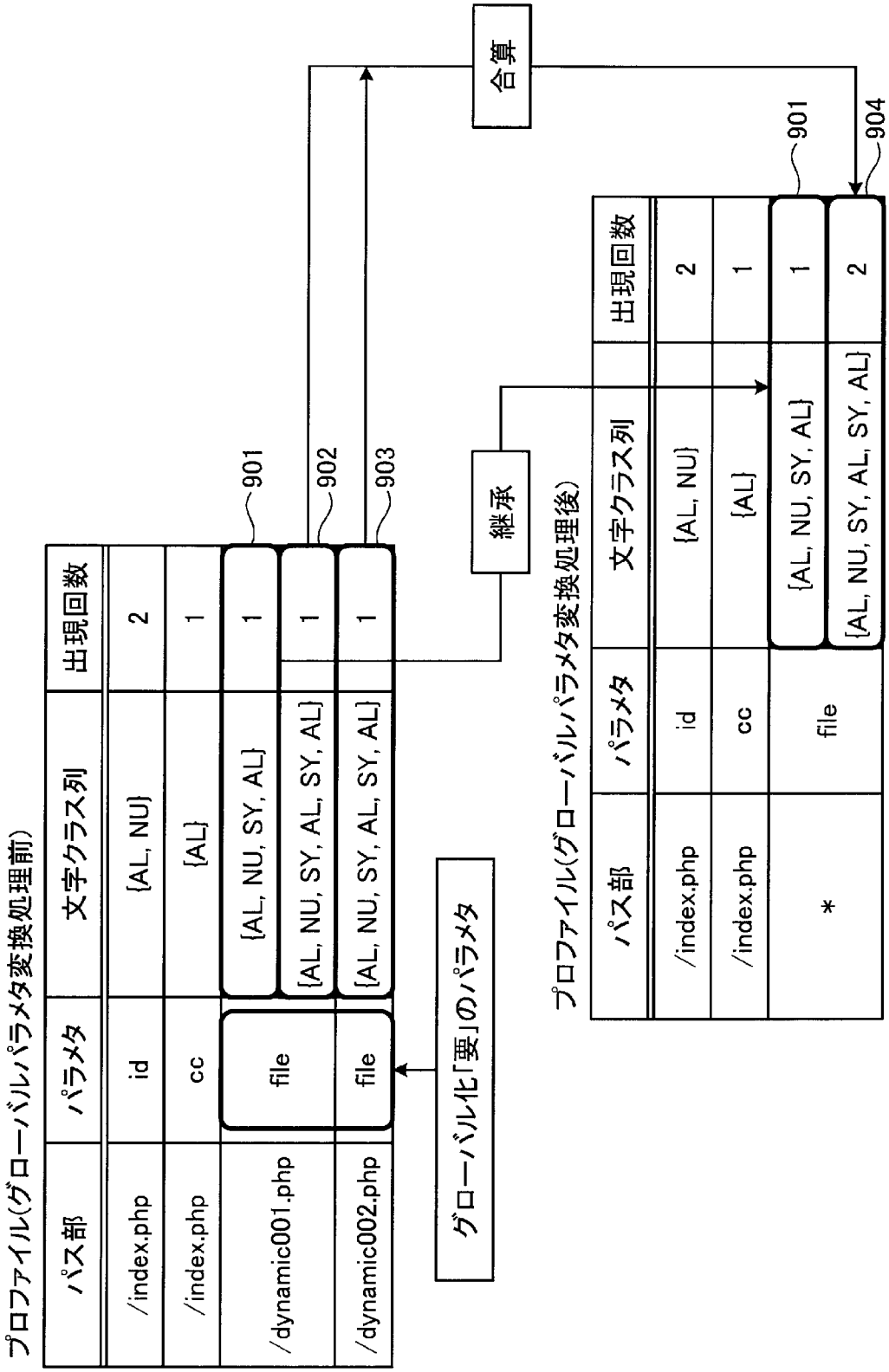
グローバル化候補パラメタ集計リスト

パラメタ	パラメタが出現したパス部	パス部種類数
id	/index.php	1
cc	/index.php	1
file	/dynamic001.php /dynamic002.php	2

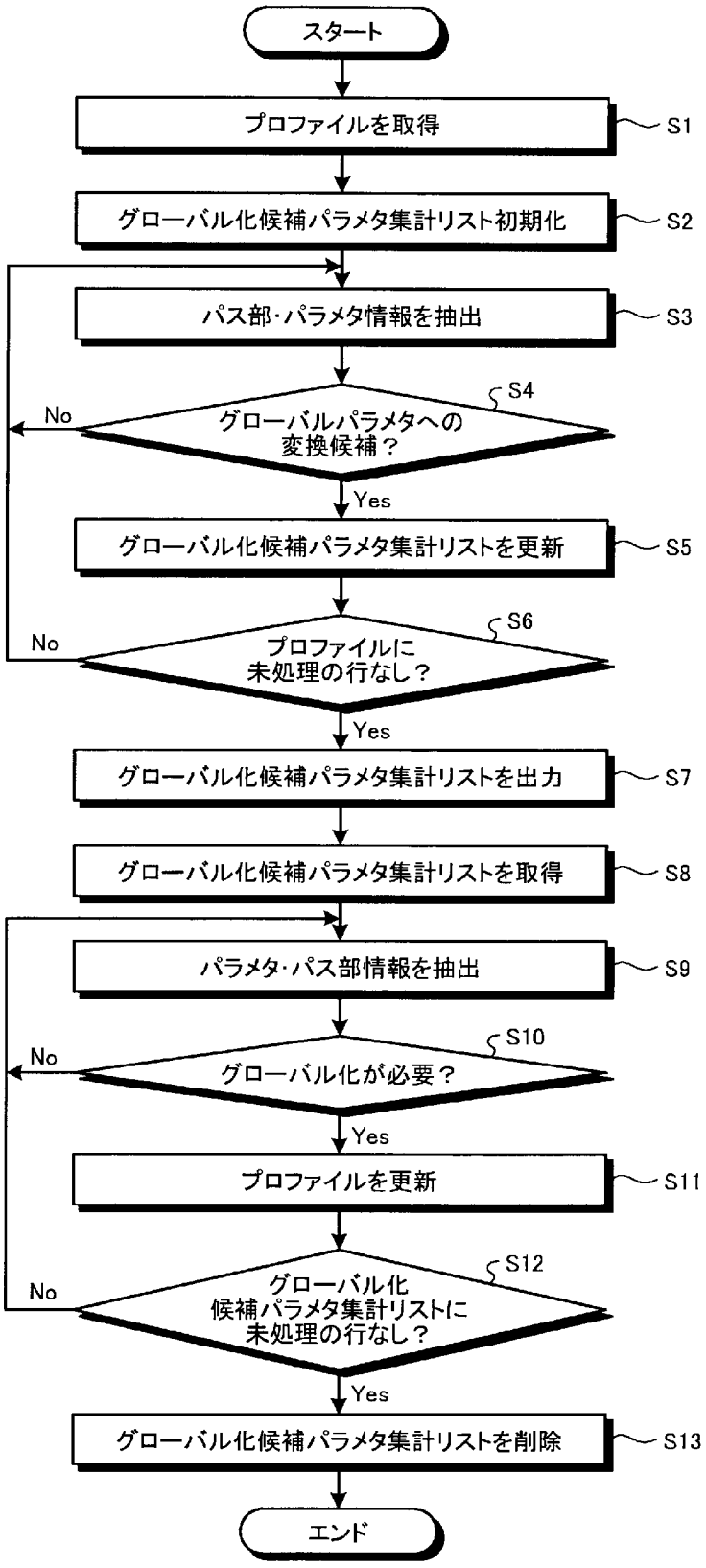
[図8]


グローバル化候補パラメタ集計リスト
パラメタ
パラメタが出現したパス部
パス部種類数
id
/index.php
1
cc
/index.php
1
file
/dynamic001.php
/dynamic002.php
2
801
〈パス部種類数の閾値=2の場合〉
パス部種類数＜パス部種類数の閾値のため、
グローバル化は「不要」と判定
パス部種類数≧パス部種類数の閾値のため、
グローバル化が「必要」と判定

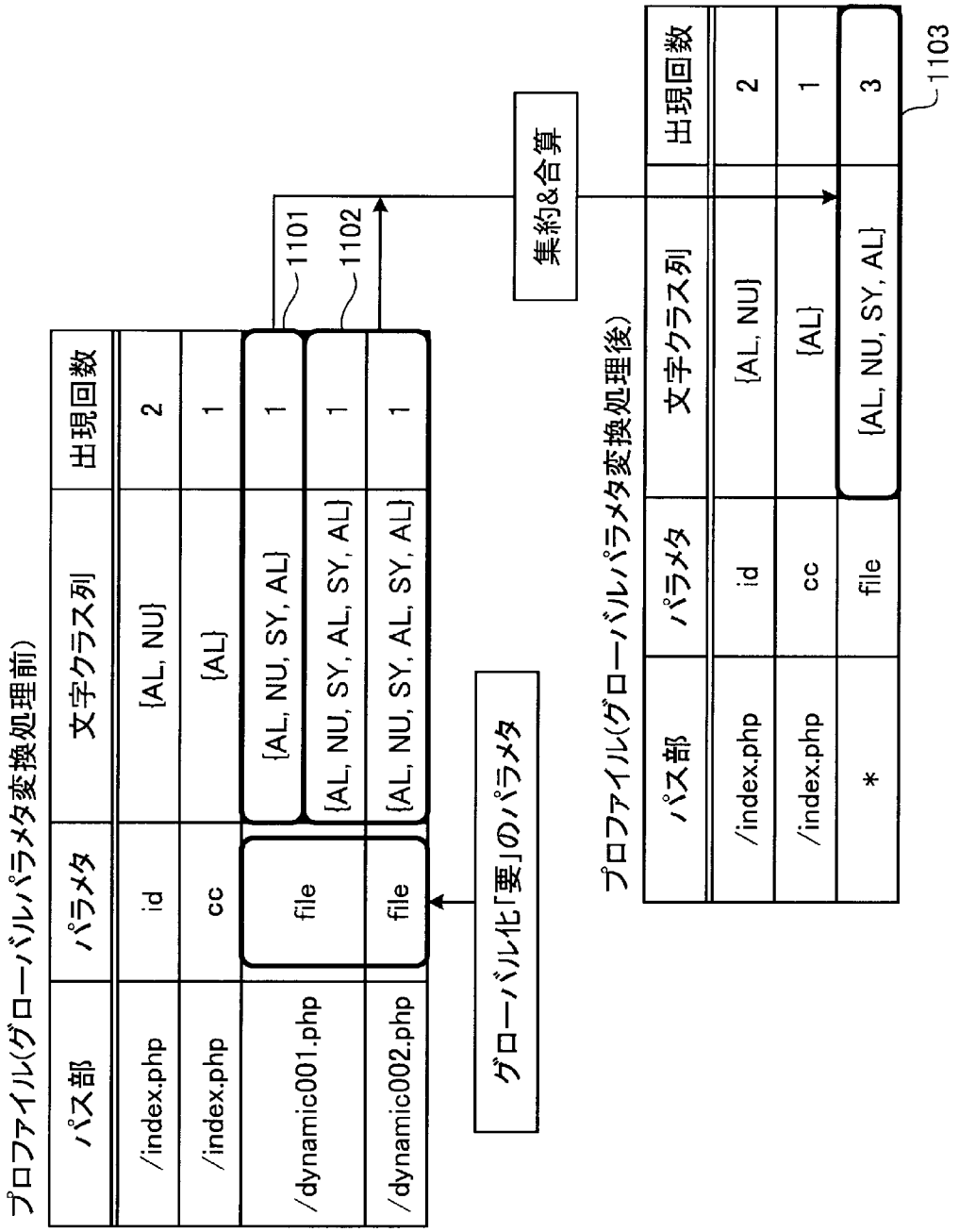
[図9]


プロファイル(グローバルパラメタ変換処理前)
パス部
パラメタ
文字クラス列
出現回数
/index.php
id
{AL, NU}
2
/index.php
cc
{AL}
1
/dynamic001.php
file
{AL, NU, SY, AL}
1
/dynamic002.php
file
{AL, NU, SY, AL, SY, AL}
1
901
902
903
合算
継承
グローバル化「要」のパラメタ
プロファイル(グローバルパラメタ変換処理後)
パス部
パラメタ
文字クラス列
出現回数
/index.php
id
{AL, NU}
2
/index.php
cc
{AL}
1
*
file
{AL, NU, SY, AL}
1
{AL, NU, SY, AL, SY, AL}
2
901
904

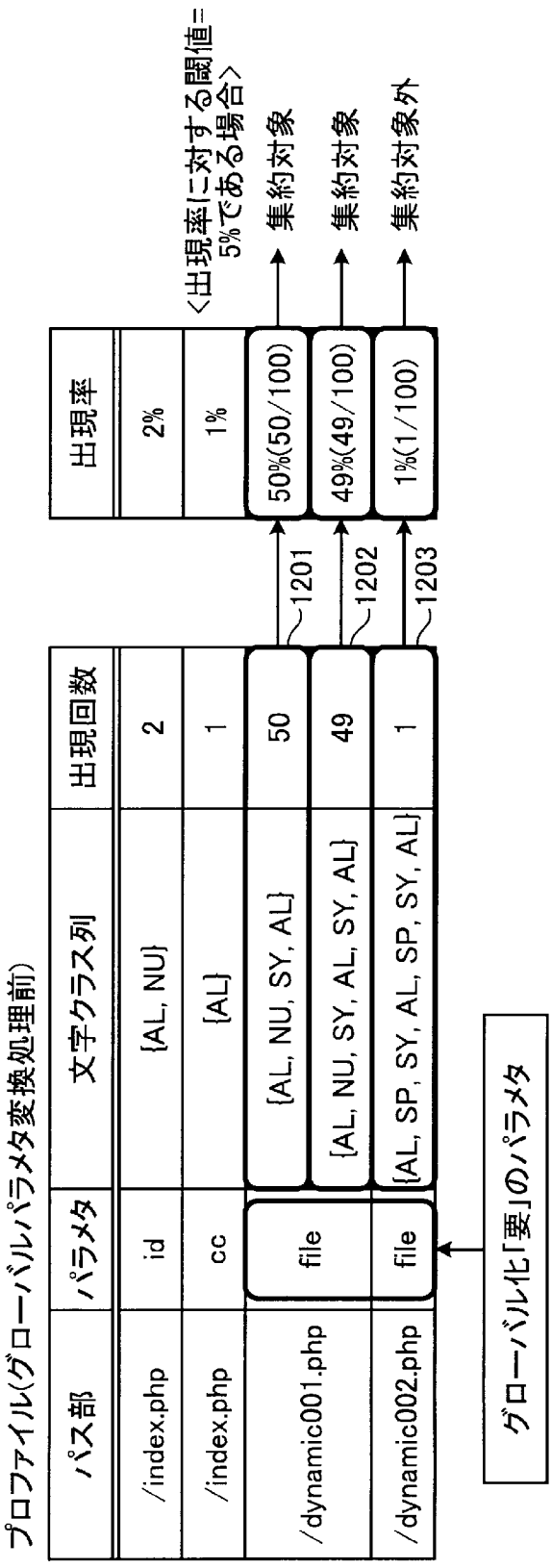
[図10]


スタート
プロファイルを取得
S1
グローバル化候補パラメタ集計リスト初期化
S2
パス部・パラメタ情報を抽出
S3
S4
No
グローバルパラメタへの
変換候補？
Yes
グローバル化候補パラメタ集計リストを更新
S5
S6
No
プロファイルに
未処理の行なし？
Yes
グローバル化候補パラメタ集計リストを出力
S7
グローバル化候補パラメタ集計リストを取得
S8
パラメタ・パス部情報を抽出
S9
S10
No
グローバル化が必要？
Yes
プロファイルを更新
S11
S12
No
グローバル化
候補パラメタ集計リストに
未処理の行なし？
Yes
グローバル化候補パラメタ集計リストを削除
S13
エンド

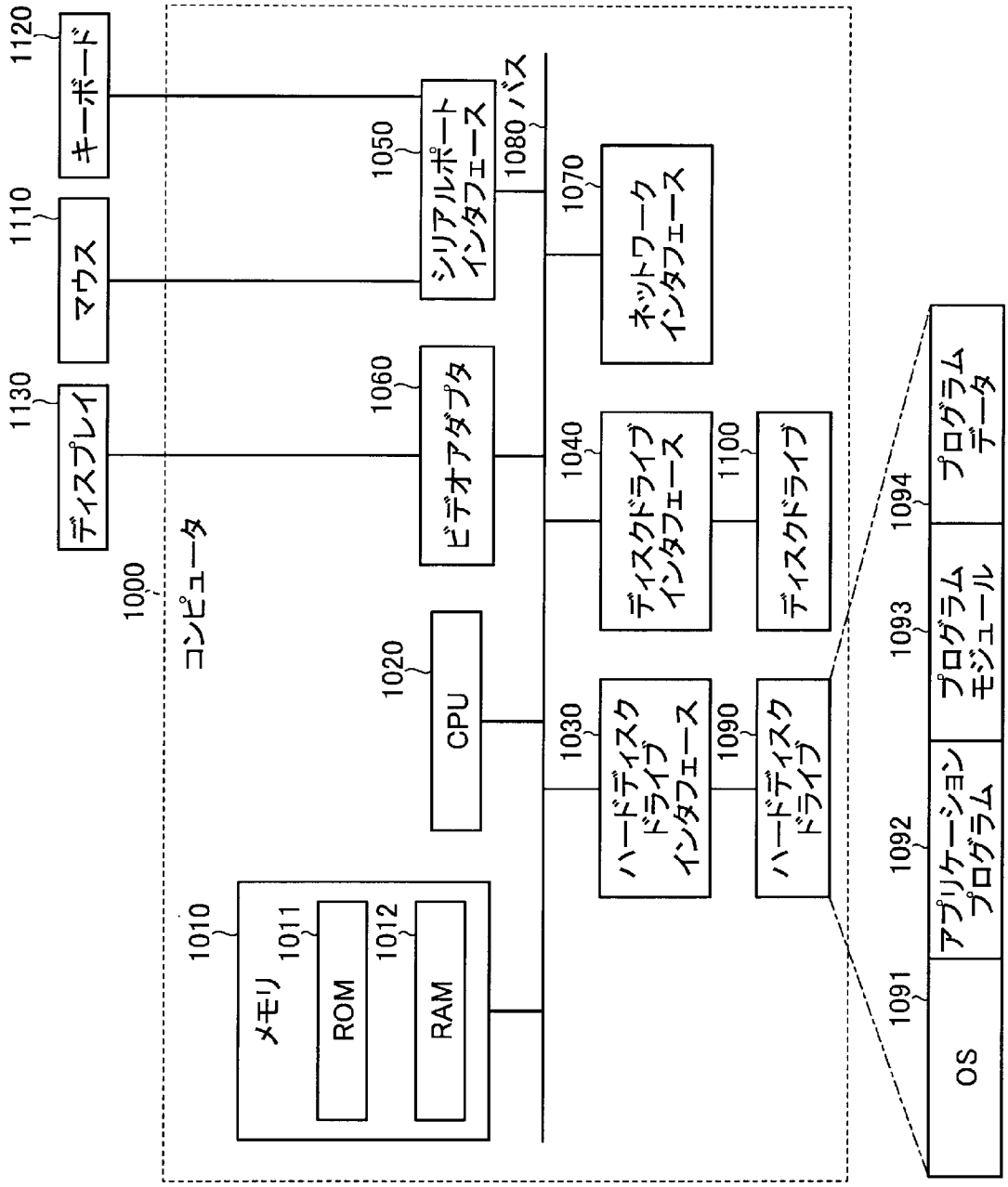
[図11]


プロファイル(グローバルパラメタ変換処理前)
パス部
パラメタ
文字クラス列
出現回数
/index.php
id
{AL, NU}
2
/index.php
cc
{AL}
1
/dynamic001.php
file
{AL, NU, SY, AL}
1
1101
{AL, NU, SY, AL, SY, AL}
1
1102
/dynamic002.php
file
{AL, NU, SY, AL, SY, AL}
1
グローバル化「要」のパラメタ
集約&合算
プロファイル(グローバルパラメタ変換処理後)
パス部
パラメタ
文字クラス列
出現回数
/index.php
id
{AL, NU}
2
/index.php
cc
{AL}
1
*
file
{AL, NU, SY, AL}
3
1103

[図12]


プロファイル(グローバルパラメタ変換処理前)
パス部
パラメタ
文字クラス列
出現回数
出現率
/index.php
id
{AL, NU}
2
2%
/index.php
cc
{AL}
1
1%
/dynamic001.php
file
{AL, NU, SY, AL}
50
50%(50/100)
{AL, NU, SY, AL, SY, AL}
49
49%(49/100)
/dynamic002.php
file
{AL, SP, SY, AL, SP, SY, AL}
1
1%(1/100)
1201
1202
1203
〈出現率に対する閾値=5%である場合〉
集約対象
集約対象
集約対象外
グローバル化「要」のパラメタ

[図13]


1000
コンピュータ
1010
メモリ
1011
ROM
1012
RAM
1020
CPU
1030
ハードディスクドライブインタフェース
1040
ディスクドライブインタフェース
1050
シリアルポートインタフェース
1060
ビデオアダプタ
1070
ネットワークインタフェース
1080
バス
1090
ハードディスクドライブ
1100
ディスクドライブ
1110
マウス
1120
キーボード
1130
ディスプレイ
1091
OS
1092
アプリケーションプログラム
1093
プログラムモジュール
1094
プログラムデータ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2018/005601

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G06F21/55(2013.01)i

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	1971-2018
Registered utility model specifications of Japan	1996-2018
Published registered utility model applications of Japan	1994-2018

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 2015/186662 A1 (NIPPON TELEGRAPH AND TELEPHONE CORP.) 10 December 2015, abstract & US 2017/0126724 A1, abstract & EP 3136249 A1 & CN 106415507 A	1-10
A	WO 2015/114804 A1 (HITACHI, LTD.) 06 August 2015, abstract (Family: none)	1-10

☐ 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
04 April 2018 (04.04.2018)	17 April 2018 (17.04.2018)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１８／００５６０１

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F21/55(2013.01)i

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））

Int.Cl.　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

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１８年
日本国実用新案登録公報	１９９６－２０１８年
日本国登録実用新案公報	１９９４－２０１８年

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）

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 2015/186662 A1（日本電信電話株式会社）2015.12.10, ［要約］ & US 2017/0126724 A1, ABSTRACT & EP 3136249 A1 & CN 106415507 A	1-10
A	WO 2015/114804 A1（株式会社日立製作所）2015.08.06, ［要約］（ファミリーなし）	1-10

☐　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。

＊　引用文献のカテゴリー
「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
「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
「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）
「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
「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矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「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
「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
「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
０４．０４．２０１８	１７．０４．２０１８

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｓ　９０７１
平井　誠
電話番号　０３－３５８１－１１０１　内線　３５４６