(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
2016 年 9 月 9 日(09.09.2016)

WIPO | PCT

(10) 国際公開番号
WO 2016/140038 A1

(51) 国際特許分類:
G06F 21/55 (2013.01)

(21) 国際出願番号: PCT/JP2016/054102

(22) 国際出願日: 2016 年 2 月 12 日(12.02.2016)

(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 2015-043940 2015 年 3 月 5 日(05.03.2015) 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: 千葉 大紀(CHIBA, Daiki); 〒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). 八木 毅(YAGI, Takeshi); 〒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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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.

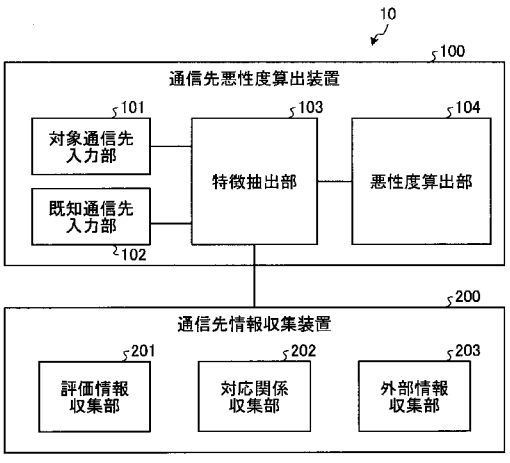
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[続葉有]

(54) **Title**: DEVICE FOR CALCULATING MALICIOUSNESS OF COMMUNICATION DESTINATION, METHOD FOR CALCULATING MALICIOUSNESS OF COMMUNICATION DESTINATION, AND PROGRAM FOR CALCULATING MALICIOUSNESS OF COMMUNICATION DESTINATION

(54) 発明の名称 : 通信先悪性度算出装置、通信先悪性度算出方法及び通信先悪性度算出プログラム

[図1]



100... DEVICE FOR CALCULATING MALICIOUSNESS OF COMMUNICATION DESTINATION
101... TARGET COMMUNICATION DESTINATION INPUT UNIT
102... ALREADY-KNOWN COMMUNICATION DESTINATION INPUT UNIT
103... CHARACTERISTIC EXTRACTION UNIT
104... MALICIOUSNESS CALCULATION UNIT
200... DEVICE FOR COLLECTING COMMUNICATION DESTINATION INFORMATION
201... EVALUATION INFORMATION COLLECTION UNIT
202... CORRESPONDENCE RELATION COLLECTION UNIT
203... EXTERNAL INFORMATION COLLECTION UNIT

(57) **Abstract**: A communication destination that is already known to be malicious or benign is inputted to an already-known communication destination input unit (102). A target communication destination, the maliciousness of which is to be calculated, is inputted to a target communication destination input unit (101). A change over time in whether the already-known communication destination and the target communication destination were entered into a list of malicious communication destinations or a list of benign communication destinations at a predetermined point of time in the past is extracted by a characteristic extraction unit (103) as characteristic information regarding the already-known communication destination and the target communication destination. On the basis of the characteristic information regarding the already-known communication destination and the target communication destination, a maliciousness calculation unit (104) calculates the maliciousness of the target communication destination.

(57) 要約:

[続葉有]

WO 2016/140038 A1

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)）

悪性または良性であることが既知の通信先を既知通信先入力部（１０２）に入力し、悪性度を算出する対象である対象通信先を対象通信先入力部（１０１）に入力し、既知通信先及び対象通信先の、悪性通信先リスト及び良性通信先リストへの過去の所定の時点における掲載の有無の時間経過に伴う変化を既知通信先及び対象通信先の特徴情報として特徴抽出部（１０３）により抽出し、既知通信先及び対象通信先の特徴情報に基づいて悪性度算出部（１０４）により対象通信先の悪性度を算出する。

明　細　書

発明の名称 ：

通信先悪性度算出装置、通信先悪性度算出方法及び通信先悪性度算出プログラム

技術分野

[0001] 本発明は、通信先悪性度算出装置、通信先悪性度算出方法及び通信先悪性度算出プログラムに関する。

背景技術

[0002] インターネットの普及に伴い、ＤＤｏＳ攻撃やスパムメール送信等のサイバー攻撃が急増している。これらの攻撃のほとんどは、マルウェアと呼ばれる悪意あるソフトウェアに起因している。攻撃者は一般ユーザの端末やサーバ等をマルウェアに感染させ、マルウェアを操作することで端末やサーバを不正に制御し、情報収集や新たな攻撃を実施している。これらの攻撃は近年社会問題化している。このため、マルウェア感染を中心としたサイバー攻撃への対策が急務となっている。

[0003] サイバー攻撃対策としては、端末上で実施する対策とネットワーク上で実施する対策とが利用されている。端末上で実施する対策としては、アンチウィルスソフトを用いる手法や、ホスト型ＩＤＳ（Intrusion Detection System）やホスト型ＩＰＳ（Intrusion Prevention System）を用いる手法が利用されている。これらの手法では、端末にソフトウェアをインストールして対策を実施する。

[0004] 一方、ネットワーク上で実施する対策としては、ネットワーク型のＩＤＳやＩＰＳ、ＦＷ（Firewall）やＷＡＦ（Web Application Firewall）等を用いる手法が利用されている。これらの手法では、ネットワークの通信経路上に検査装置を配置する。例えば、ネットワークの通信経路のうちＤＮＳサーバへの通信を監視できる箇所で、ＤＮＳクエリやＤＮＳレスポンスの通信の検査を行う手法が提案されている（例えば、非特許文献１または２を参照

）。また、近年では、端末や装置のログを分析して攻撃の痕跡を発見するＳＩＥＭ（Security Information and Event Management）サービス等も実施されている。

[0005] これらの手法においては、ハニーポットと呼ばれるおとりのシステム上でマルウェア感染攻撃やその他のサイバー攻撃の通信相手や通信内容を収集する。また、サンドボックスと呼ばれるマルウェア解析システムでマルウェアを実際に動作させてマルウェアの通信先や通信内容を収集したり、スパムメール対策システムやＤＤｏＳ対策システムで攻撃と判定された通信の通信先や通信内容を収集したりすることで、攻撃に関わる通信の情報を収集する。

[0006] そして、収集した攻撃について、例えば通信先のＩＰアドレス等をブラックリスト化し、当該ＩＰアドレスを相手とした通信を攻撃と判定する。なお、ブラックリスト化する情報は、統一資源位置指定子（ＵＲＬ：Uniform Resource Locator）やドメイン名とする場合もあるが、この際は、ＵＲＬやドメイン名を正規表現でブラックリスト化することもある。

[0007] なお、通常異なる装置やソフトウェアからトラヒックログやアラートを収集して通信相手や通信内容の情報を抽出する際、装置やソフトウェアに応じて各項目の表記方法が異なる場合があるが、近年ではＳＩＥＭ製品として異なる表記で示されたログ情報を統一的な表記方法に変換して集計する技術も普及している。

先行技術文献

非特許文献

[0008] 非特許文献1：M. Antonakakis, et al., “Building a Dynamic Reputation System for DNS,” Proc. USENIX conference on Security, 2010.

非特許文献2：L. Bilge, et al., “EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis,” Proc. NDSS, 2011.

発明の概要

発明が解決しようとする課題

[0009] しかし、上記の方法では、ハニーポットやサンドボックス等を利用してサイバー攻撃に関わる通信の情報を収集する際に、利用されるすべての悪性通信先を抽出することはできない。例えば、ハニーポットで収集可能な悪性サイトの通信先は無数に存在する上に、時間経過とともに当該悪性サイトの無効化や、別の通信先への移行が発生する。

[0010] また、サンドボックスでマルウェアを解析する場合に、マルウェアは解析の妨害やインターネットへの接続確認を目的として、良性通信先へのアクセスや、時間経過とともに変化する悪性通信先へのアクセスを発生させる。このようにサイバー攻撃に関わる通信の情報を収集するだけでは、悪性通信先を網羅的に、かつ、正確に特定し、ブラックリスト化することは困難である。

[0011] 例えば、非特許文献１または２においては、ある時点までに収集した情報を利用して、当該時点で特定されていない悪性通信先をブラックリスト化する手法が提案されているが、攻撃者によって一時的に利用される悪性通信先や攻撃者が攻撃の準備のために確保している悪性通信先を特定することができないという課題がある。

[0012] このため、現在では、サイバー攻撃を発見するために、ある時点で最新のブラックリストを参照しても、攻撃に利用される悪性通信先を特定できない場合があり、このような場合においては、解析者が手動で内容を分析する必要が生じる。この結果、攻撃に利用された通信先の情報を抽出するための時間的なコストと人的なコストが必要となり、攻撃が多種多様化している近年では、これらのコストがセキュリティベンダやサービスプロバイダにおいて大きなボトルネックとなっている。

[0013] 本発明の目的は、実通信を発生させることなく自動的に通信先の悪性度を算出し、最新のブラックリストを参照するだけでは判別できない悪性通信先を精度よく特定することにある。

課題を解決するための手段

[0014] 本発明の通信先悪性度算出装置は、悪性度を算出する対象である対象通信先を入力する対象通信先入力部と、悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力部と、前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出部と、前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出部と、を有することを特徴とする。

[0015] 本発明の通信先悪性度算出方法は、悪性度を算出する対象である対象通信先を入力する対象通信先入力工程と、悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力工程と、前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出工程と、前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出工程と、を含んだことを特徴とする。

[0016] 本発明の通信先悪性度算出プログラムは、コンピュータに、悪性度を算出する対象である対象通信先を入力する対象通信先入力ステップと、悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力ステップと、前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出ステップと、前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出ステップと、を実行させることを特徴とする。

発明の効果

[0017] 本発明によれば、実通信を発生させることなく自動的に通信先の悪性度を算出し、最新のブラックリストを参照するだけでは判別できない悪性通信先を精度よく特定することができる。

図面の簡単な説明

[0018] [図1]図１は、実施形態１に係る通信先悪性度算出システムの構成の一例を示す図である。

[図2]図２は、実施形態１に係る通信先悪性度算出装置における対象通信先の一例を示す図である。

[図3]図３は、実施形態１に係る通信先悪性度算出装置における既知通信先の一例を示す図である。

[図4]図４は、実施形態１に係る通信先悪性度算出装置における通信先の評価情報の一例を示す図である。

[図5]図５は、実施形態１に係る通信先悪性度算出装置における通信先の評価情報の一例を示す図である。

[図6]図６は、実施形態１に係る通信先悪性度算出装置における通信先の評価情報の一例を示す図である。

[図7]図７は、実施形態１に係る通信先悪性度算出装置における通信先の対応関係の一例を示す図である。

[図8]図８は、実施形態１に係る通信先悪性度算出装置におけるＩＰアドレスの外部情報の一例を示す図である。

[図9]図９は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群の抽出方法の一例を示す図である。

[図10]図１０は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群のリストの一例を示す図である。

[図11]図１１は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群から抽出した特徴情報の一例を示す図である。

[図12]図１２は、実施形態１に係る通信先悪性度算出装置におけるドメイン名の外部情報の一例を示す図である。

[図13]図１３は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群の抽出方法の一例を示す図である。

[図14]図１４は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群のリストの一例を示す図である。

[図15]図１５は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群から抽出した特徴情報の一例を示す図である。

[図16]図１６は、実施形態１に係る通信先悪性度算出装置が算出する悪性度の一例を示す図である。

[図17]図１７は、実施形態１に係る通信先悪性度算出装置における統合された特徴情報の一例を示す図である。

[図18]図１８は、実施形態１に係る通信先悪性度算出装置が算出する悪性度の一例を示す図である。

[図19]図１９は、実施形態１に係る通信先悪性度算出装置の処理の一例を示す図である。

[図20]図２０は、実施形態１に係る通信先悪性度算出装置の処理の一例を示す図である。

[図21]図２１は、通信先悪性度算出装置として機能するコンピュータの一例を示す図である。

発明を実施するための形態

[0019] 以下に添付図面を参照して、この発明に係る通信先悪性度算出装置、通信先悪性度算出方法及び通信先悪性度算出プログラムの実施形態を詳細に説明する。なお、この実施形態によりこの発明が限定されるものではない。

[0020] ［実施形態１に係る通信先悪性度算出装置の構成］

まず、図１を用いて、通信先悪性度算出システムの構成について説明する。図１は、実施形態１に係る通信先悪性度算出システムの構成の一例を示す図である。図１に示すように、通信先悪性度算出システム１０は、通信先悪性度算出装置１００及び通信先情報収集装置２００を有する。

[0021] 図１に示すように、通信先悪性度算出装置１００は、対象通信先入力部１

０１と、既知通信先入力部１０２と、特徴抽出部１０３と、悪性度算出部１０４と、を有する。

[0022] 対象通信先入力部１０１には、悪性度を算出する対象である対象通信先を入力する。対象通信先入力部１０１に入力されるデータの例について、図２を用いて説明する。図２は、実施形態１に係る通信先悪性度算出装置における対象通信先の一例を示す図である。図２に示すように、通信先の種別としては、ドメイン名、ＵＲＬ、ＩＰアドレス等が挙げられる。例えば、図２の通番１の行は、ドメイン名が「www.example.com」である通信先を示している。なお、通信先の種別は図示のものに限られず、ＦＱＤＮ（Fully Qualified Domain Name）等であっても良い。また、対象通信先入力部１０１に入力される通信先の情報として、評価情報や外部情報はこの時点で含まれている必要はない。

[0023] 既知通信先入力部１０２には、悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する。既知通信先入力部１０２に入力されるデータの例について、図３を用いて説明する。図３は、実施形態１に係る通信先悪性度算出装置における既知通信先の一例を示す図である。既知通信先のデータとしては、まず、図２に示す対象通信先と同様に、通信先の種別及び通信先が含まれる。さらに、図３のラベル列に示すような、悪性または良性を示す情報も必要となる。例えば、図３の通番１の行は、ドメイン名が「foo.example.com」である通信先のラベルが「良性」であることを示している。図３の例においては、ラベル列に示される情報は悪性または良性の２値となっているが、図３に示すものに限られず、悪性度を示す値等であっても良い。また、図２と同様に、通信先の種別は図示のものに限られない。

[0024] 特徴抽出部１０３は、既知通信先及び対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、掲載の有無の時間経過に伴う変化を既知通信先及び対象通信先の特徴情報として抽出する。また、特徴抽出部１０３は、既知通信先及び対象通信先の外部情報及び関連する通

信先との対応関係をさらに取得し、対応関係から抽出される関連する通信先群の外部情報の統計量を特徴情報としてさらに抽出する。特徴抽出部１０３における具体的な処理については通信先情報収集装置２００で収集される情報についての説明と共に後述する。

[0025] なお、特徴抽出部１０３は、通信先評価のためのリストから所定の周期で所定の期間に収集された掲載の有無を取得しても良い。また、既知通信先及び対象通信先をドメイン名とし、関連する通信先は、既知通信先及び対象通信先、及び既知通信先及び対象通信先のトップレベルドメイン、及び既知通信先及び対象通信先をトップレベルドメインとして持つドメイン名と対応付けられたＩＰアドレス、または、既知通信先及び対象通信先と同じＡＳ番号に所属するＩＰアドレスに対応付けられた履歴を持つドメイン名としても良い。

[0026] 悪性度算出部１０４は、既知通信先及び対象通信先の特徴情報に基づいて対象通信先の悪性度を算出する。ここで、悪性度算出部１０４は、既知通信先の特徴情報を入力データとし、既知通信先が悪性であるか良性であるかを出力データとする教師あり機械学習によって悪性度算出のためのモデルを生成し、モデルを用いて対象通信先の悪性度を算出するようにしても良い。

[0027] 図１に示すように、通信先情報収集装置２００は、評価情報収集部２０１と、対応関係収集部２０２と、外部情報収集部２０３と、を有する。通信先情報収集装置２００の各部で収集された情報は、通信先悪性度算出装置１００の特徴抽出部１０３に転送される。

[0028] 評価情報収集部２０１は、通信先の評価情報を収集する。評価情報収集部２０１は、通信先の評価情報として、予め定義された悪性通信先リストや良性通信先リスト等を収集する。また、予め設定された所定の期間及び周期等に従って収集を行うようにしても良い。収集方法としては、例えば、公知のＷｅｂクロール手法を用いて収集対象のリストの配信先へアクセスする。なお、収集するリストとしては、前述の悪性通信先リストや良性通信先リストのように悪性または良性を示すようなものに限られない。例えば、アクセス

数が多い通信先のリストといった、何らかの評価が行われており、定期的に掲載の開始や終了が行われているものであれば良い。

[0029] 図４を用いて、評価情報収集部２０１で収集される評価情報について説明する。図４は、実施形態１に係る通信先悪性度算出装置における通信先の評価情報の一例を示す図である。図４に示すように、評価情報収集部２０１は、複数の悪性通信先リストや良性通信先リストを収集し、それぞれのリストの所定の期間における掲載状況を評価情報としている。図４の表は、例えばt、t-1、t-2がそれぞれ今月、先月、先々月を示すものとし、各期間におけるリストへの掲載の有無を表している。

[0030] 例えば、図４の通番１の行は、通信先「www.example.com」がt-2の時点からtの時点まで「良性通信先リスト１」に掲載されていたことを示している。また、通番２の行は、通信先「www.example.com」がt-2の時点では、「良性通信先リスト２」に掲載が無く、t-1及びtの時点においては「良性通信先リスト２」に掲載されていたことを示している。

[0031] なお、評価情報収集部２０１は、悪性通信先リストとして、例えば公開ブラックリストの全部や一部を利用することが考えられる。また、評価情報収集部２０１は、良性通信先リストとして、例えばＷｅｂ上で公開されているような頻繁に閲覧される人気のドメインリストの全部や一部を利用することや、任意のネットワーク内で収集できる頻繁に閲覧されているようなドメインリストを利用することが考えられる。

[0032] また、評価情報収集部２０１は、各通信先がリストに記載された通信先と完全に一致している場合だけでなく、部分的に一致している場合や、別途定める通信先の類似性が一定の基準を満たしている場合にも通信先がリストへ掲載されているとみなすようにしても良い。

[0033] 特徴抽出部１０３は、例えば、図４の表を基に、図５または図６に示すような特徴情報を抽出する。図５及び図６は、実施形態１に係る通信先悪性度算出装置における通信先の評価情報の一例を示す図である。

[0034] 図５は、通信先の各リストへの掲載有無の時間的な変化を特徴情報として

抽出したものである。例えば、図５の通番１に示す通信先「www.example.com」は、t-2からtという範囲の各時点において、良性通信先リスト１に安定的に掲載されていることから、特徴抽出部１０３は、「安定掲載」という特徴情報を抽出する。同様に、特徴抽出部１０３は、途中から掲載が開始された場合は「途中掲載開始」、途中で掲載が終了した場合は「途中掲載終了」、どの時点でも掲載されていない場合「掲載無」などの特徴情報を抽出することができる。なお、特徴情報の抽出方法は、図５に示すものに限られず、例えば特徴情報を数値とし、掲載有の場合は数値に1だけ加えるといったルールに基づいて抽出されるようにしても良い。

[0035] 図６は、通信先の複数のリストへの掲載有無の時間的変化を組み合わせたものである。例えば、図６の通番１に示す通信先「www.example.com」は、t-2からtという範囲の各時点において、良性通信先リスト１に安定的に掲載されていると同時に、良性通信先リスト２に途中から掲載が開始されたという「安定掲載＆途中掲載開始」という特徴情報を抽出する。なお、特徴情報の抽出方法は、図６に示すものに限られず、例えば掲載有無の時間的変化を数値化した値の和や積によって表すようにしても良い。

[0036] 対応関係収集部２０２は、種別の異なる通信先の対応関係及びその履歴を収集する。対応関係収集部２０２は、例えば、ＤＮＳサーバでＤＮＳクエリを収集するＰａｓｓｉｖｅ　ＤＮＳと呼ばれる手法を利用して収集を行う。図７を用いて、対応関係収集部２０２が収集する対応関係及びその履歴について説明する。図７は、実施形態１に係る通信先悪性度算出装置における通信先の対応関係の一例を示す図である。

[0037] 図７は、ドメイン名とＩＰアドレスとの対応関係及びその履歴の例を示している。ドメイン名とＩＰアドレスとの対応関係は、例えばＤＮＳに代表されるプロトコルを利用して得ることができる。ドメイン名とＩＰアドレスとの対応関係は、運用やその形態に応じて時間経過とともに変化する場合がある。そのため、対応関係収集部２０２は、ドメイン名とＩＰアドレスの対応関係にタイムスタンプを付与し、履歴として収集する。

[0038] 例えば、図７の通番１の行は、ドメイン名「www.example.com」が2015年1月1日00:00:00にＩＰアドレス「192.0.2.1」が対応していたことを示している。通番２の行は、通番１の行の１時間後の2015年1月1日01:00:00にドメイン名「www.example.com」に対し、ＩＰアドレス「192.0.2.2」が対応していたことを示している。また、通番２０１の行は、2015年1月1日00:00:00にドメイン名「example.com」に対し、公知の負荷分散技術であるＤＮＳラウンドロビン等が利用され、複数のＩＰアドレス「192.0.2.201,192.0.2.202」が対応していたことを示している。

[0039] 対応関係収集部２０２は、ドメイン名とＩＰアドレスとの対応関係の履歴を収集する際には、例えばトップレベルドメインやセカンドレベルドメインを管理する権威ＤＮＳサーバや、任意の組織内ネットワークに配置されたキャッシュＤＮＳサーバでＤＮＳ通信を観測する手法を利用することができる。

[0040] 外部情報収集部２０３は、通信先の運用状況や利用状況等を示す外部情報を収集する。図８を用いて、外部情報収集部２０３が収集するＩＰアドレスの外部情報の例について説明する。図８は、実施形態１に係る通信先悪性度算出装置におけるＩＰアドレスの外部情報の一例を示す図である。

[0041] 図８に示すように、ＩＰアドレスの外部情報としては、ＩＰアドレスの属するアドレスプレフィックスやＡＳ番号、組織名、国、地域インターネットレジストリ（ＲＩＲ：Regional Internet Registry）、当該アドレスがＲＩＲから割り当てられたアドレス割当日等が挙げられる。なお、ＩＰアドレスに対する外部情報は図８に示すものに限られない。また、外部情報収集部２０３は、ＷＨＯＩＳプロトコルを利用して独自に収集した情報、各ＲＩＲが公開している情報や、MaxMind社のGeoIP（登録商標）等の公開されたサービスにより得られる情報を利用してＩＰアドレスに対する外部情報を収集することができる。

[0042] 例えば、図８の通番１の行では、ＩＰアドレス「192.0.2.1」の外部情報であるアドレスプレフィックスが「192.0.2.0/24」、ＡＳ番号が「64501」、組

織名が「ＴＥＳＴ－ＮＥＴ－１」、国が「ＵＳ」、ＲＩＲが「ＡＲＩＮ」、アドレス割当日が「2001年1月1日」であることを示している。

[0043] ここで、図９、１０及び１１を用いて、特徴抽出部１０３が、図７に示すドメイン名とＩＰアドレスとの対応関係及びその履歴と、図８に示すＩＰアドレスの外部情報とから特徴情報を抽出する場合の例を説明する。まず、特徴抽出部１０３は、図９に示す方法でドメイン名に関連するＩＰアドレス群を抽出する。図９は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群の抽出方法の一例を示す図である。

[0044] 図９を用いて、ドメイン名３００の関連ＩＰアドレス群を抽出する方法を説明する。図７に示すように、ドメイン名３００は、t-1即ち2015年1月1日00:00:00にＩＰアドレス「192.0.2.1」と対応しており、t即ち2015年1月1日01:00:00にはＩＰアドレス「192.0.2.2」と対応していた。さらに、上位ドメインであるドメイン名３５０は、t-1即ち2015年1月1日00:00:00にＩＰアドレス「192.0.2.201」及び「192.0.2.202」と対応しており、t即ち2015年1月1日01:00:00にもＩＰアドレス「192.0.2.201」及び「192.0.2.202」と対応している。これより、ドメイン名３００に関連するＩＰアドレス群として、ＩＰアドレス群３０１及び３５１に含まれる「192.0.2.1」「192.0.2.2」「192.0.2.201」「192.0.2.202」の４つが抽出される。

[0045] このようにして、図１０に示すように、各ドメイン名に関連するＩＰアドレス群のリストを抽出することができる。図１０は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群のリストの一例を示す図である。例えば、通番１で示されるドメイン名「www.example.com」に関連するＩＰアドレス群は、「192.0.2.1」「192.0.2.2」「192.0.2.201」「192.0.2.202」の４つであることが分かる。

[0046] 次に、特徴抽出部１０３は、各ドメイン名に関連するＩＰアドレス群に含まれるＩＰアドレスの外部情報から図１１に示すような統計量を算出したものを特徴量として抽出する。図１１は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するＩＰアドレス群から抽出した特徴情報の

一例を示す図である。

[0047] 例えば、図１１の通番１の行に記載されているドメイン名「www.example.com」に関連するＩＰアドレス群は、「192.0.2.1」「192.0.2.2」「192.0.2.201」「192.0.2.202」であり、特徴抽出部１０３は、図８を参照してこれらのＩＰアドレスについての算出可能な統計量であるアドレスプレフィックス数、ＡＳ番号数、組織名数、国数、ＲＩＲ数、アドレス割当日数等を算出し、特徴情報とする。例えば図１１の通番１の行は、通信先「www.example.com」から算出した関連するＩＰアドレスが4、アドレスプレフィックス数が1、ＡＳ番号数が1、組織名数が1、国数が1、ＲＩＲ数が1、アドレス割当日数が1であることを示している。なお、統計量の項目は、図１１に示すものに限られない。

[0048] ここまで、特徴抽出部１０３が関連するＩＰアドレス群の外部情報に基づいて特徴情報を抽出する場合の例について説明した。特徴抽出部１０３が特徴情報を抽出する他の方法として、関連するドメイン名群の外部情報に基づいて特徴情報を抽出する方法がある。なお、特徴抽出部１０３は、ＩＰアドレス群の外部情報に基づいて特徴情報を抽出する方法と、関連するドメイン名群の外部情報に基づいて特徴情報を抽出する方法のいずれかを採用しても良いし、両方を採用しても良い。

[0049] 特徴抽出部１０３が関連するドメイン群の外部情報に基づいて特徴情報を抽出する場合は、外部情報収集部２０３は、ドメイン名の外部情報を収集する。図１２を用いて、外部情報収集部２０３が収集するドメイン名の外部情報の例について説明する。図１２は、実施形態１に係る通信先悪性度算出装置におけるドメイン名の外部情報の一例を示す図である。

[0050] 図１２に示すように、ドメイン名の外部情報としては、ドメイン名の属するＴＬＤ（トップレベルドメイン）やＷＨＯＩＳサーバ名、ＮＳサーバ、ドメイン名登録日、ドメイン名更新日、ドメイン名失効日等が挙げられる。なお、ドメイン名に対する外部情報は図１２に示すものに限られない。また、外部情報収集部２０３は、ＷＨＯＩＳプロトコルを利用して独自に収集した

情報や、第３者が公開しているサービスにより得られる情報を利用してドメイン名の外部情報を収集することができる。

[0051] 例えば、図１２の通番１の行では、ドメイン名「www.example.com」の外部情報であるＴＬＤが「.com」、ＷＨＯＩＳサーバ名が「whois.example.com」、ＮＳサーバが「ns1.example.com」、ドメイン名登録日が「2001.1.1」、ドメイン名更新日が「2014.1.1」、ドメイン名失効日が「2015.1.1」であることを示している。

[0052] 図１３、１４及び１５を用いて、特徴抽出部１０３が、図７に示すドメイン名とＩＰアドレスとの対応関係及びその履歴と、図１２に示すドメイン名の外部情報とから特徴情報を抽出する場合の例を説明する。まず、特徴抽出部１０３は、図１３に示す方法でドメイン名に関連するドメイン名群を抽出する。図１３は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群の抽出方法の一例を示す図である。

[0053] 図１３を用いて、ドメイン名４００の関連ドメイン名群を抽出する方法を説明する。図７に示すように、ドメイン名４００は、t-1即ち2015年1月1日00:00:00にＩＰアドレス「192.0.2.1」と対応しており、t即ち2015年1月1日01:00:00にはＩＰアドレス「192.0.2.2」と対応していた。そして、「192.0.2.1」及び「192.0.2.2」と同じＡＳ番号「64501」を持つＩＰアドレス「192.0.2.101」や「192.0.2.201」が含まれるＩＰアドレス群４５０と対応したことがあるドメイン名４１０及び４２０が関連するドメイン名群として抽出される。

[0054] このようにして、図１４に示すように、各ドメイン名について関連するドメイン名群のリストを抽出することができる。図１４は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群のリストの一例を示す図である。例えば、通番１で示されるドメイン名「www.example.com」に関連するドメイン名群は、「www.example.com」「foo.example.com」「example.com」の３つであることが分かる。なお、統計量を計算する際の便宜上、あるドメイン名に関連するドメイン名群にはそのドメイン名自身も

含まれるものとする。

[0055] 次に、特徴抽出部１０３は、各ドメイン名に関連するドメイン群に含まれるドメインの外部情報から図１５に示すような統計量を算出したものを特徴量として抽出する。図１５は、実施形態１に係る通信先悪性度算出装置におけるドメイン名に関連するドメイン名群から抽出した特徴情報の一例を示す図である。

[0056] 例えば、図１５の通番１の行に記載されているドメイン名「www.example.com」に関連するドメイン名群は、「www.example.com」「foo.example.com」「example.com」であり、特徴抽出部１０３は、図１２を参照してこれらのドメイン名についての算出可能な統計量である関連するドメイン名数、ＴＬＤ数、ＷＨＯＩＳサーバ数、ＮＳサーバ数等を算出する。また、特徴抽出部１０３は、関連するドメイン名の文字列の文字数に関する統計量として、ドメイン名平均、ドメイン名中央値、ドメイン名標準偏差等を算出し、特徴情報としても良い。例えば図１５の通番１の行は、通信先「www.example.com」から算出した関連するドメイン名数が3、ドメイン名平均が13.7、ドメイン名中央値が15、ドメイン名標準偏差が2.31、ＴＬＤ数が1、ＷＨＯＩＳサーバ数が1、ＮＳサーバ数が3であることを示している。なお、統計量の項目は、図１５に示すものに限られない。

[0057] 悪性度算出部１０４は、図１６に示すような通信先毎の悪性度を算出する。図１６は、実施形態１に係る通信先悪性度算出装置が算出する悪性度の一例を示す図である。図１６の通番１の行は、例えば、図２に示す悪性度算出の対象となる通信先の１つである「www.example.com」について、悪性度算出部１０４が0.3という悪性度を算出したことを示している。

[0058] 悪性度の算出方法としては、公知の外れ値発見等の統計分析手法や、サポートベクターマシンやランダムフォレストやロジスティック回帰に代表される機械学習手法が利用できる。なお、悪性度は連続値で算出する場合だけでなく、離散値で算出する場合、連続値や離散値で算出した結果に応じて悪性度を任意の値やラベルに変換して出力する場合があり、図１６に示すものに

限られない。

[0059] 具体的に、まず、悪性度算出部１０４は、図３に示すような既知通信先の特徴情報に所定のアルゴリズムを適用して、悪性度算出のためのモデルである訓練モデルを生成する。特徴情報としては、これまで説明してきたように、図５及び６に示すような評価情報に基づくもの、また、図１１及び図１５に示すような通信先の外部情報や対応関係に基づくものが含まれる。そして、悪性度算出部１０４は、生成された訓練モデルを利用して、訓練モデルを生成した際のアルゴリズムを対象通信先に適用することで、悪性度を算出する。なお、訓練モデルは、図３に示すような良性ラベルが付された既知通信先は悪性度が小さくなり、悪性ラベルが付された既知通信先は悪性度が大きくなるように生成される。例えば、所定のアルゴリズムをロジスティック回帰とすると、訓練モデルとして、図３の通信先「foo.example.com」の悪性度を小さくし、通信先「bar.example.com」の悪性度を大きくするような回帰式が得られ、得られた回帰式から対象通信先の悪性度を得ることができる。

[0060] さらに、図１７に示すような未知の通信先についても、各種情報を収集し、統合された特徴情報を抽出することで、図１８に示すような悪性度を算出することが可能である。なお、図１７は、実施形態１に係る通信先悪性度算出装置における統合された特徴情報の一例を示す図である。また、図１８は、実施形態１に係る通信先悪性度算出装置が算出する悪性度の一例を示す図である。

[0061] ［実施形態１に係る通信先悪性度算出装置の処理］

図１９及び図２０を用いて、通信先悪性度算出装置１００の処理について説明する。図１９及び図２０は、実施形態１に係る通信先悪性度算出装置の処理の一例を示す図である。詳しくは、図１９は、前述の訓練モデル生成までの処理を示しており、図２０は生成された訓練モデルを利用して悪性度を算出する処理を示している。

[0062] 図１９を用いて、訓練モデル生成までの処理を説明する。図１９に示すように、まず、既知通信先入力部１０２に、既知悪性通信先と既知良性通信先

が入力される（ステップＳ１０１）。次に、特徴抽出部１０３は、入力された既知悪性通信先と既知良性通信先の、通信先リストへの掲載有無を参照する（ステップＳ１０２）。そして、特徴抽出部１０３は、通信先リストへの掲載有無の時間変化（ステップＳ１０３）及び通信先リストへの掲載有無の組合せの時間変化（ステップＳ１０４）を特徴情報として抽出する。

[0063] 特徴抽出部１０３は、ドメイン名とＩＰアドレスの対応関係の履歴情報の構築を行う（ステップＳ１０５）。そして、ＩＰアドレスの利用状況を示す外部情報の収集が行われ（ステップＳ１０６）、特徴抽出部１０３は、通信先に関連するＩＰアドレス群の関係を構築し（ステップＳ１０７）、通信先に関連するＩＰアドレス群の統計量を特徴情報として抽出する（ステップＳ１０８）。

[0064] また、ドメイン名の利用状況を示す外部情報の収集が行われ（ステップＳ１０９）、特徴抽出部１０３は、通信先に関連するドメイン名群の関係を構築し（ステップＳ１１０）、通信先に関連するドメイン名群の統計量を特徴情報として抽出する（ステップＳ１１１）。

[0065] そして、悪性度算出部１０４は、抽出した特徴情報を統合し（ステップＳ１１２）、悪性度算出のアルゴリズムを適用し（ステップＳ１１３）、訓練モデルを出力する（ステップＳ１１４）。

[0066] 図２０を用いて、訓練モデルを利用して悪性度を算出する処理を説明する。図２０に示すように、まず、対象通信先入力部１０１に、悪性度算出対象の通信先と訓練モデルが入力される（ステップＳ２０１）。そして、特徴抽出部１０３は、通信先リストへの掲載有無の時間変化（ステップＳ２０２）及び通信先リストへの掲載有無の組合せの時間変化（ステップＳ２０３）を特徴情報として抽出する。

[0067] 特徴抽出部１０３は、通信先に関連するＩＰアドレス群の統計量を特徴情報として抽出し（ステップＳ２０４）、さらに、通信先に関連するドメイン名群の統計量を特徴情報として抽出する（ステップＳ２０５）。そして、悪性度算出部１０４は、抽出した特徴情報を統合し（ステップＳ２０６）、訓

練モデルを利用して悪性度算出のアルゴリズムを適用し（ステップＳ２０７）、通信先に対する悪性度を出力する（ステップＳ２０８）。

[0068] ［実施形態１の効果］

通信先悪性度算出装置１００は、悪性度を算出する対象である対象通信先を対象通信先入力部１０１へ入力し、悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として既知通信先入力部１０２へ入力する。そして、特徴抽出部１０３は、既知通信先及び対象通信先の、悪性通信先リスト及び良性通信先リストへの所定の時点における掲載の有無の時間経過に伴う変化を既知通信先及び対象通信先の特徴情報として抽出する。悪性度算出部１０４は、既知通信先及び対象通信先の特徴情報に基づいて対象通信先の悪性度を算出する。このため、実通信を発生させることなく自動的に通信先の悪性度を算出し、最新のブラックリストを参照するだけでは判別できない悪性通信先を精度よく特定することが可能である。

[0069] また、悪性度算出部１０４は、既知通信先の特徴情報を入力データとし、既知通信先が悪性であるか良性であるかを出力データとする教師あり機械学習によって悪性度算出のためのモデルを生成し、生成したモデルを用いて対象通信先の悪性度を算出する。このため、例えば、既知通信先の時間経過に伴う変化等を加味したモデルに、対象通信先の特徴情報を入力するだけで、精度よく対象通信先の悪性度を自動的に算出することが可能である。

[0070] また、従来では、悪性度が未知である通信先には、攻撃者が一時的に利用する通信先や攻撃者が今後利用する可能性の高い通信先が含まれ、これらはブラックリストを参照するだけでは悪性であるか否かを判別できなかった。これに対して、実施形態１によれば、既知の悪性通信先リストと良性通信先リストを入手し、各通信先の時間経過に伴う変化（例えば、掲載開始や掲載終了）を解析して特徴情報として抽出し、解析対象の通信先リストの特徴情報と比較分析して、各解析対象の通信先の悪性度を算出しているので、悪性度が未知である通信先について、実通信を発生させることなく悪性度を算出することができる。さらに、特徴抽出部１０３は、通信先評価のためのリス

トから所定の周期で所定の期間に収集された掲載の有無を取得することで、複数のリストの情報を効率良く比較することができる。

[0071] また、特徴抽出部１０３は、既知通信先及び対象通信先の外部情報及び関連する通信先との対応関係の履歴情報をさらに取得し、履歴情報から抽出される関連する通信先群の外部情報の統計量を特徴情報としてさらに抽出する。なお、関連する通信先は、通信先のトップレベルドメインや、通信先をトップレベルドメインとして持つドメイン名等と対応付けられたＩＰアドレスや、同じＡＳ番号に所属するＩＰアドレスを持つドメイン名等である。

[0072] これにより、対象通信先や既知通信先だけでなく、それらに関連する通信先も含めた幅広い範囲の通信先の悪性度を算出することができ、また、より多くの特徴情報を得ることができるため、算出の精度を向上させることができる。

[0073] なお、実施形態の説明においては、モデルを生成する処理と生成したモデルによって対象通信先の悪性度を算出する処理を別々に行う場合について説明したが、本発明はこれに限定されるものではない。例えば、モデルを生成することなく、既知通信先と対象通信先とに関する情報を同時に入力し、既知通信先の特徴情報と対象通信先の特徴情報とを比較分析して、悪性度を算出するようにしても良い。

[0074] ［システム構成等］

また、図示した各装置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如く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、各装置にて行なわれる各処理機能は、その全部又は任意の一部が、ＣＰＵ（Central Processing Unit）及び当該ＣＰＵにて解析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。

[0075] また、本実施形態において説明した各処理のうち、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。

[0076] ［プログラム］

また、上記実施形態において説明した通信先悪性度算出装置が実行する処理について、コンピュータが実行可能な言語で記述したプログラムを作成することもできる。この場合、コンピュータがプログラムを実行することにより、上記実施形態と同様の効果を得ることができる。さらに、かかるプログラムをコンピュータが読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータに読み込ませて実行することにより上記実施形態と同様の処理を実現してもよい。以下に、図１に示した通信先悪性度算出装置と同様の機能を実現するプログラムを実行するコンピュータの一例を説明する。

[0077] 図２１は、通信先悪性度算出装置として機能するコンピュータの一例を示す図である。図２１に例示するように、コンピュータ１０００は、例えば、メモリ１０１０と、ＣＰＵ１０２０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とを有し、これらの各部はバス１０８０によって接続される。

[0078] メモリ１０１０は、図２１に例示するように、ＲＯＭ（Read Only Memory）１０１１及び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、図２１に例示するように、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、図２１に例示するように

、ディスクドライブ１１００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１１００に挿入される。シリアルポートインタフェース１０５０は、図２１に例示するように、例えばマウス１１１０、キーボード１１２０に接続される。ビデオアダプタ１０６０は、図２１に例示するように、例えばディスプレイ１１３０に接続される。

[0079] ここで、図２１に例示するように、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、プログラムデータ１０９４を記憶する。すなわち、上記のプログラムは、コンピュータ１０００によって実行される指令が記述されたプログラムモジュールとして、例えばハードディスクドライブ１０９０に記憶される。

[0080] また、上記実施形態で説明した各種データは、プログラムデータとして、例えばメモリ１０１０やハードディスクドライブ１０９０に記憶される。そして、ＣＰＵ１０２０が、メモリ１０１０や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、実行する。

[0081] なお、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、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、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、ネットワーク（ＬＡＮ（Local Area Network）、ＷＡＮ（Wide Area Network）等）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

[0082] １０ 通信先悪性度算出システム

１００ 通信先悪性度算出装置

１０１　対象通信先入力部

１０２　既知通信先入力部

１０３　特徴抽出部

１０４　悪性度算出部

２００　通信先情報収集装置

２０１　評価情報収集部

２０２　対応関係収集部

２０３　外部情報収集部

請求の範囲

[請求項1] 悪性度を算出する対象である対象通信先を入力する対象通信先入力部と、

悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力部と、

前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出部と、

前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出部と、

を有することを特徴とする通信先悪性度算出装置。

[請求項2] 前記悪性度算出部は、前記既知通信先の特徴情報を入力データとし、前記既知通信先が悪性であるか良性であるかを出力データとする教師あり機械学習によって悪性度算出のためのモデルを生成し、前記モデルを用いて前記対象通信先の悪性度を算出することを特徴とする請求項１に記載の通信先悪性度算出装置。

[請求項3] 前記特徴抽出部は、前記通信先評価のためのリストから所定の周期で所定の期間に収集された掲載の有無を取得することを特徴とする請求項１に記載の通信先悪性度算出装置。

[請求項4] 前記特徴抽出部は、前記既知通信先及び前記対象通信先の外部情報及び関連する通信先との対応関係の履歴情報をさらに取得し、前記履歴情報から抽出される関連する通信先群の前記外部情報の統計量を前記特徴情報としてさらに抽出することを特徴とする請求項１に記載の通信先悪性度算出装置。

[請求項5] 前記既知通信先及び前記対象通信先はドメイン名であり、

前記関連する通信先は、前記既知通信先及び前記対象通信先、及び前記既知通信先及び前記対象通信先のトップレベルドメイン、及び前

記既知通信先及び前記対象通信先をトップレベルドメインとして持つドメイン名と対応付けられたＩＰアドレスであることを特徴とする請求項４に記載の通信先悪性度算出装置。

[請求項6] 前記既知通信先及び前記対象通信先はドメイン名であり、

前記関連する通信先は、前記既知通信先及び前記対象通信先と同じＡＳ番号に所属するＩＰアドレスに対応付けられた履歴を持つドメイン名であることを特徴とする請求項４に記載の通信先悪性度算出装置。

[請求項7] 悪性度を算出する対象である対象通信先を入力する対象通信先入力工程と、

悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力工程と、

前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出工程と、

前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出工程と、

を含んだことを特徴とする通信先悪性度算出方法。

[請求項8] コンピュータに、

悪性度を算出する対象である対象通信先を入力する対象通信先入力ステップと、

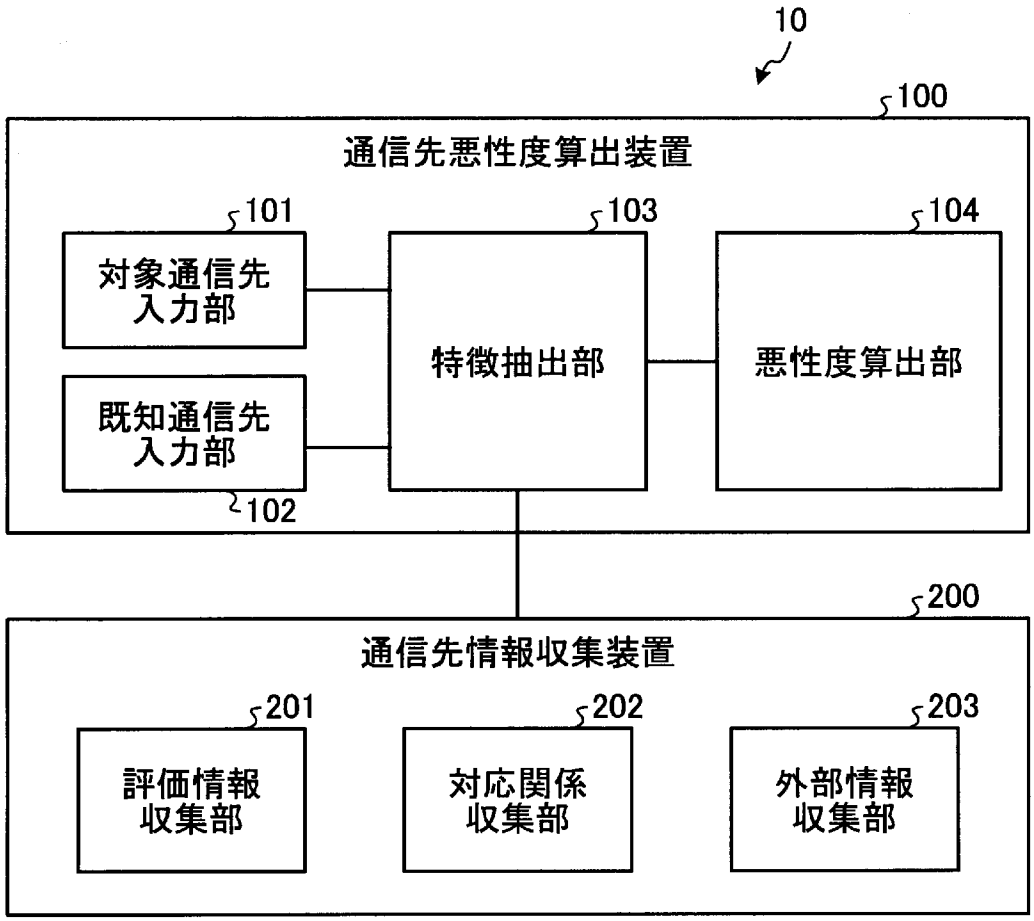
悪性であることが既知の通信先と、良性であることが既知の通信先と、を既知通信先として入力する既知通信先入力ステップと、

前記既知通信先及び前記対象通信先の、通信先評価のためのリストへの所定の時点における掲載の有無を取得し、前記掲載の有無の時間経過に伴う変化を前記既知通信先及び前記対象通信先の特徴情報として抽出する特徴抽出ステップと、

前記既知通信先及び前記対象通信先の前記特徴情報に基づいて前記対象通信先の悪性度を算出する悪性度算出ステップと、

を実行させることを特徴とする通信先悪性度算出プログラム。

[図1]


10
100
通信先悪性度算出装置
101
対象通信先
入力部
既知通信先
入力部
102
103
特徴抽出部
104
悪性度算出部
200
通信先情報収集装置
201
評価情報
収集部
202
対応関係
収集部
203
外部情報
収集部

[図2]

通番	種別	通信先
1	ドメイン名	www.example.com
2	ドメイン名	www.example.net
…	…	…
101	URL	http://www.example.com/abcdef/index.php?test=123
102	URL	http://www.example.net/index.php?num=2
…	…	…
201	IPアドレス	192.0.2.1
202	IPアドレス	203.0.113.2
…	…	…

[図3]

通番	ラベル	種別	通信先
1	良性	ドメイン名	foo.example.com
2	悪性	ドメイン名	bar.example.net
3	悪性	ドメイン名	hoge.example.net
…		…	…
101	良性	URL	http://foo.example.com/abcdef/index.php?test=123
102	悪性	URL	http://bar.example.net/index.php?num=2
103	悪性	URL	http://hoge.example.net/test/index.php
…		…	…
201	悪性	IPアドレス	192.0.2.101
202	良性	IPアドレス	203.0.113.202
203	悪性	IPアドレス	203.0.113.51
…		…	…

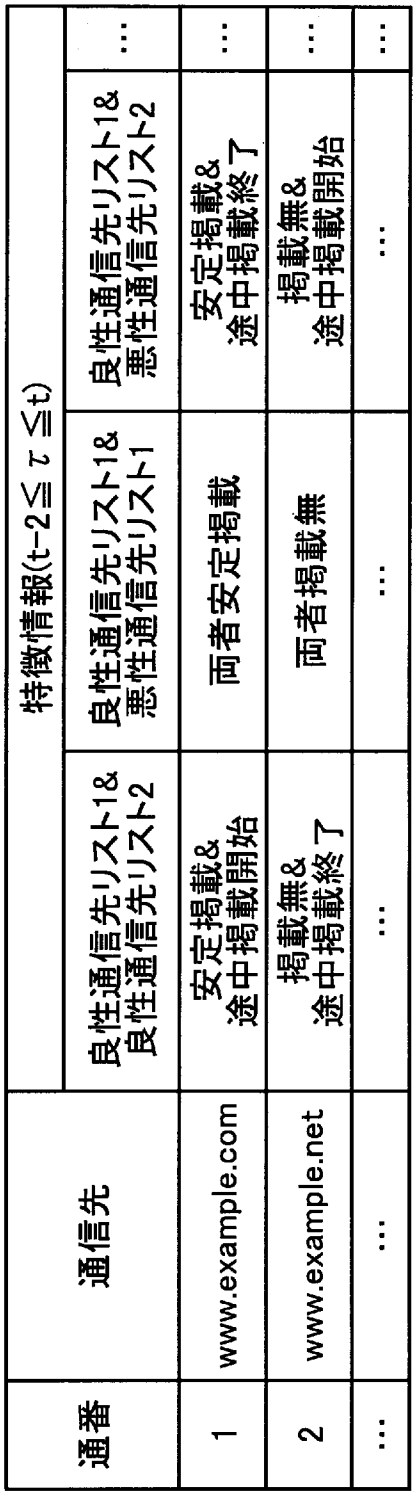
[図4]

通番	通信先	通信先リスト	掲載有無			
			…	t-2	t-1	t
1	www.example.com	良性通信先リスト1	…	掲載有	掲載有	掲載有
2	www.example.com	良性通信先リスト2	…	掲載無	掲載有	掲載有
…	…	…	…	…	…	…
11	www.example.com	悪性通信先リスト1	…	掲載有	掲載有	掲載有
12	www.example.com	悪性通信先リスト2	…	掲載有	掲載無	掲載無
…	…	…	…	…	…	…
21	www.example.net	良性通信先リスト1	…	掲載無	掲載無	掲載無
22	www.example.net	良性通信先リスト2	…	掲載有	掲載有	掲載無
…	…	…	…	…	…	…
31	www.example.net	悪性通信先リスト1	…	掲載無	掲載無	掲載無
32	www.example.net	悪性通信先リスト2	…	掲載無	掲載有	掲載有
…	…	…	…	…	…	…

[図5]

通番	通信先	特徴情報(t-2≦ τ ≦t)					
		良性通信先リスト1	良性通信先リスト2	…	悪性通信先リスト1	悪性通信先リスト2	…
1	www.example.com	安定掲載	途中掲載開始	…	安定掲載	途中掲載終了	…
2	www.example.net	掲載無	途中掲載終了	…	掲載無	途中掲載開始	…
…	…	…	…	…	…	…	…

[図6]



通番	通信先	特徴情報(t-2≦τ≦t)			
		良性通信先リスト1&良性通信先リスト2	良性通信先リスト1&悪性通信先リスト1	良性通信先リスト1&悪性通信先リスト2	…
1	www.example.com	安定掲載&途中掲載開始	両者安定掲載	安定掲載&途中掲載終了	…
2	www.example.net	掲載無&途中掲載終了	両者掲載無	掲載無&途中掲載開始	…
…	…	…	…	…	…

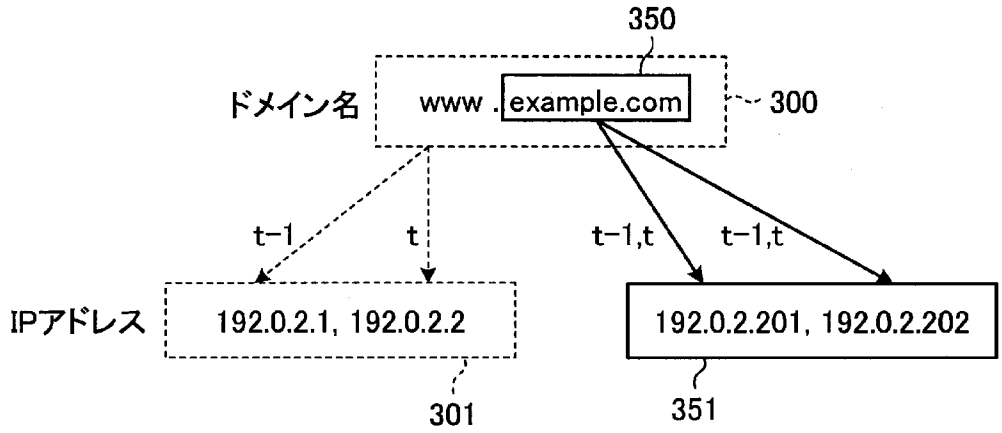
[図7]

通番	ドメイン名	タイムスタンプ	IPアドレス
1	www.example.com	2015年1月1日 00:00:00	192.0.2.1
2	www.example.com	2015年1月1日 01:00:00	192.0.2.2
…	…	…	…
101	foo.example.com	2015年1月1日 00:00:00	192.0.2.101
102	foo.example.com	2015年1月1日 01:00:00	192.0.2.102
…	…	…	…
201	example.com	2015年1月1日 00:00:00	192.0.2.201, 192.0.2.202
202	example.com	2015年1月1日 01:00:00	192.0.2.201, 192.0.2.202
…	…	…	…
301	www.example.net	2015年1月1日 00:00:00	203.0.113.1
302	www.example.net	2015年2月1日 01:00:00	203.0.113.2
…	…	…	…
401	bar.example.net	2015年2月2日 01:00:00	203.0.113.202
…	…	…	…
501	hoge.example.net	2015年2月3日 01:00:00	203.0.113.51
…	…	…	…

[図8]

通番	IPアドレス	アドレス プレフィックス	AS 番号	組織名	国	RIR	アドレス割当日	…
1	192.0.2.1	192.0.2.0/24	64501	TEST-NET-1	US	ARIN	2001年1月1日	…
2	192.0.2.101	192.0.2.0/24	64501	TEST-NET-1	US	ARIN	2001年1月1日	…
3	192.0.2.201	192.0.2.0/24	64501	TEST-NET-1	US	ARIN	2001年1月1日	…
…	…	…	…	…	…	…	…	…
101	203.0.113.2	203.0.113.0/24	64502	TEST-NET-2	US	ARIN	2001年2月1日	…
102	203.0.113.202	203.0.113.0/24	64502	TEST-NET-2	US	ARIN	2001年3月1日	…
103	203.0.113.51	203.0.113.0/24	64502	TEST-NET-2	US	ARIN	2001年4月1日	…
…	…	…	…	…	…	…	…	…

[図9]



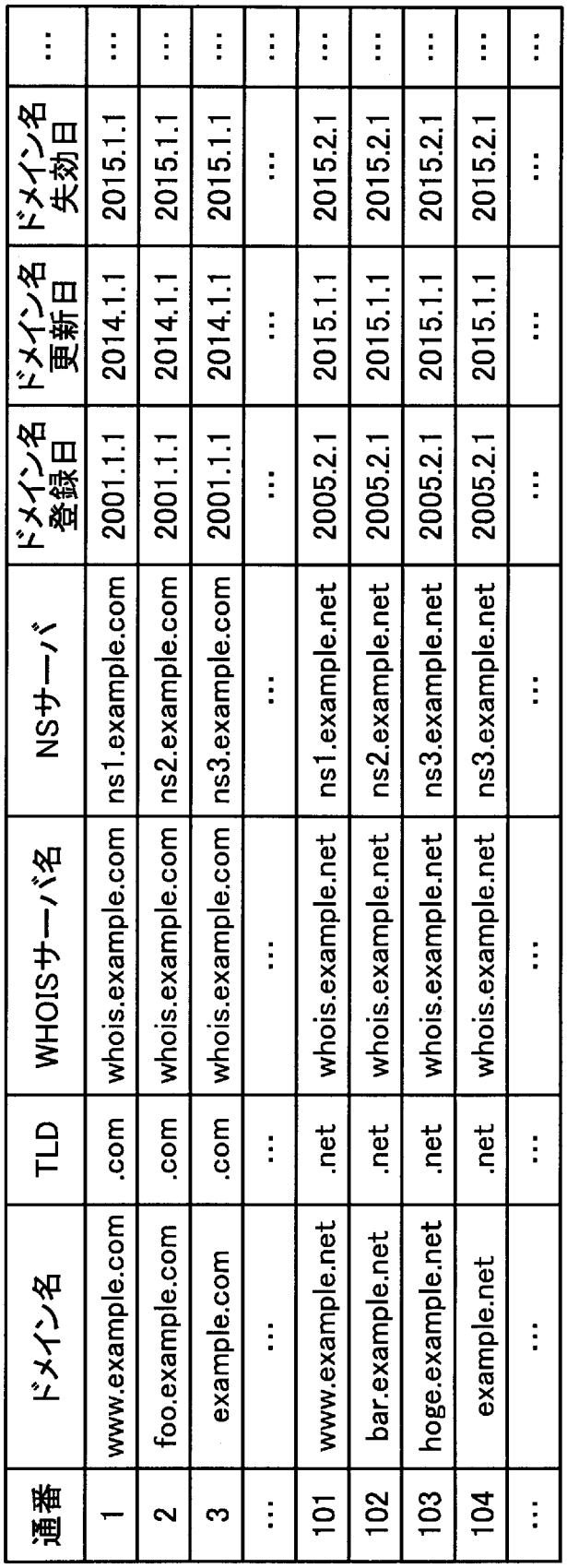
[図10]

通番	通信先	通信先に関連するIPアドレス群
1	www.example.com	192.0.2.1, 192.0.2.2, 192.0.2.201, 192.0.2.202
2	foo.example.com	192.0.2.101, 192.0.2.102, 192.0.2.201, 192.0.2.202
3	example.com	192.0.2.201, 192.0.2.202
…	…	…
101	www.example.net	203.0.113.1, 203.0.113.2, 203.0.113.101
102	bar.example.net	203.0.113.201, 203.0.113.202, 203.0.113.101
103	hoge.example.net	203.0.113.51, 203.0.113.101
104	example.net	203.0.113.101
…	…	…

[図11]

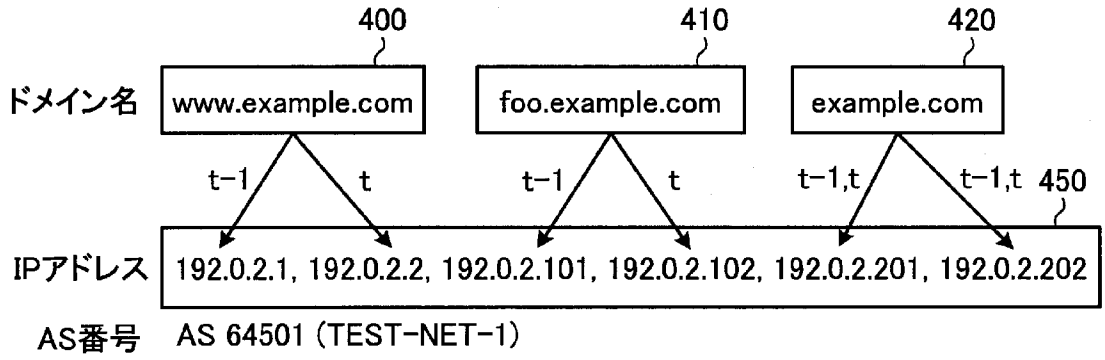
通番	通信先	特徴情報							
		関連するIPアドレス	アドレスプレフィックス数	AS番号数	組織名数	国数	RIR数	アドレス割当日数	…
1	www.example.com	4	1	1	1	1	1	1	…
2	foo.example.com	4	1	1	2	2	1	1	…
3	example.com	2	1	1	1	1	1	1	…
…	…	…	…	…	…	…	…	…	…
101	www.example.net	3	1	1	3	1	1	2	…
102	bar.example.net	3	1	1	3	1	1	2	…
103	hoge.example.net	2	1	1	1	1	1	1	…
104	example.net	1	1	1	1	1	1	1	…
…	…	…	…	…	…	…	…	…	…

[図12]



通番	ドメイン名	TLD	WHOISサーバ名	NSサーバ	ドメイン名登録日	ドメイン名更新日	ドメイン名失効日	…
1	www.example.com	.com	whois.example.com	ns1.example.com	2001.1.1	2014.1.1	2015.1.1	…
2	foo.example.com	.com	whois.example.com	ns2.example.com	2001.1.1	2014.1.1	2015.1.1	…
3	example.com	.com	whois.example.com	ns3.example.com	2001.1.1	2014.1.1	2015.1.1	…
…	…	…	…	…	…	…	…	…
101	www.example.net	.net	whois.example.net	ns1.example.net	2005.2.1	2015.1.1	2015.2.1	…
102	bar.example.net	.net	whois.example.net	ns2.example.net	2005.2.1	2015.1.1	2015.2.1	…
103	hoge.example.net	.net	whois.example.net	ns3.example.net	2005.2.1	2015.1.1	2015.2.1	…
104	example.net	.net	whois.example.net	ns3.example.net	2005.2.1	2015.1.1	2015.2.1	…
…	…	…	…	…	…	…	…	…

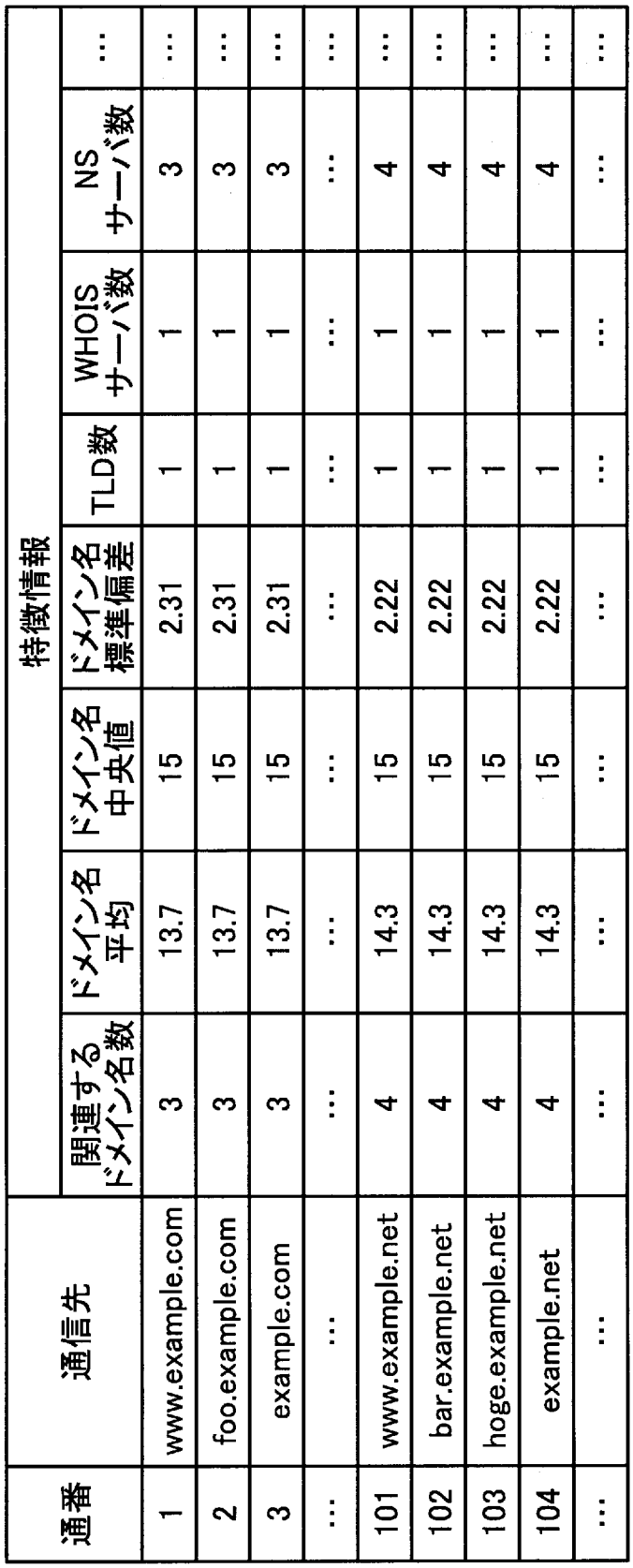
[図13]



[図14]

通番	通信先	通信先に関連するドメイン名群
1	www.example.com	www.example.com, foo.example.com, example.com
2	foo.example.com	www.example.com, foo.example.com, example.com
3	example.com	www.example.com, foo.example.com, example.com
…	…	…
101	www.example.net	www.example.net, bar.example.net, hoge.example.net, example.net
102	bar.example.net	www.example.net, bar.example.net, hoge.example.net, example.net
103	hoge.example.net	www.example.net, bar.example.net, hoge.example.net, example.net
104	example.net	www.example.net, bar.example.net, hoge.example.net, example.net
…	…	…

[図15]



通番	通信先	特徴情報							
		関連するドメイン名数	ドメイン名平均	ドメイン名中央値	ドメイン名標準偏差	TLD数	WHOISサーバ数	NSサーバ数	…
1	www.example.com	3	13.7	15	2.31	1	1	3	…
2	foo.example.com	3	13.7	15	2.31	1	1	3	…
3	example.com	3	13.7	15	2.31	1	1	3	…
…	…	…	…	…	…	…	…	…	…
101	www.example.net	4	14.3	15	2.22	1	1	4	…
102	bar.example.net	4	14.3	15	2.22	1	1	4	…
103	hoge.example.net	4	14.3	15	2.22	1	1	4	…
104	example.net	4	14.3	15	2.22	1	1	4	…
…	…	…	…	…	…	…	…	…	…

[図16]

通番	通信先	悪性度
1	www.example.com	0.3
2	foo.example.com	0.4
3	example.com	0.2
…	…	…
101	www.example.net	0.6
102	bar.example.net	0.7
103	hoge.example.net	0.9
104	example.net	0.7
…	…	…

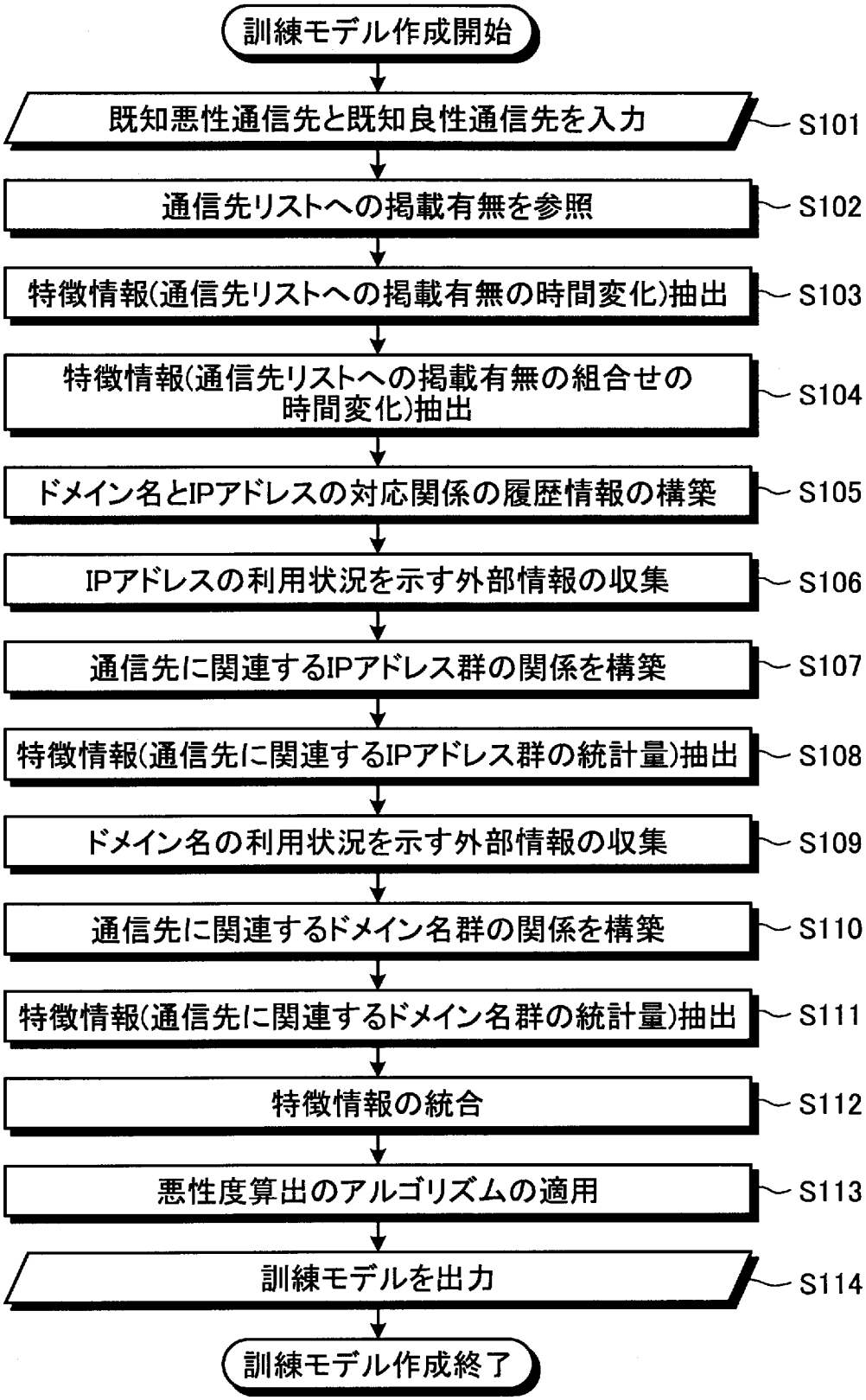
[図17]

通番	通信先	特徴情報					
		良性通信先リスト1	悪性通信先リスト2	…	関連するドメイン名数	ドメイン名平均	…
1	www.example.co.jp	安定掲載	途中掲載終了	…	5	17	…
2	hoge.example.co.jp	安定掲載	安定掲載	…	5	17	…
…	…	…	…	…	…	…	…

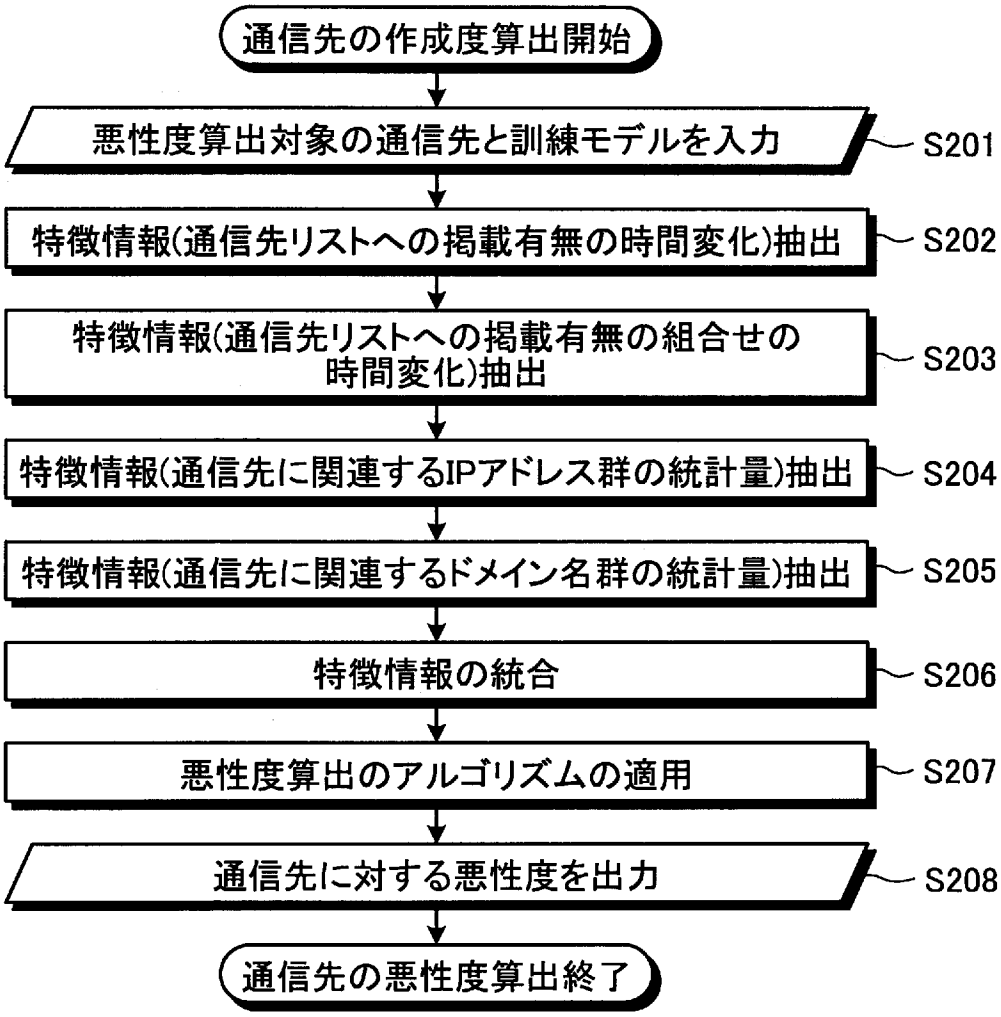
[図18]

通番	通信先	悪性度
1	www.example.co.jp	0.2
2	hoge.example.co.jp	0.5

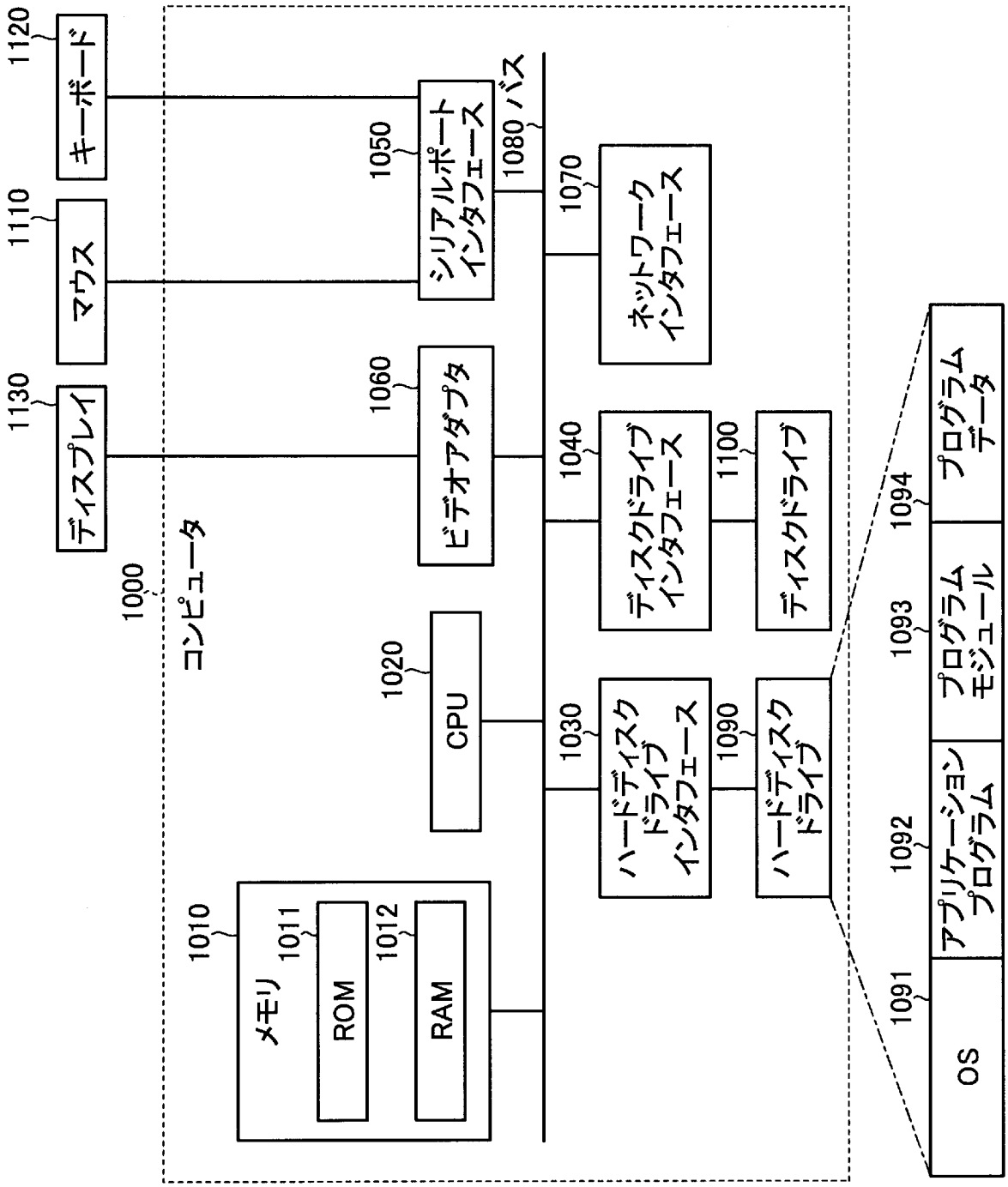
[図19]


訓練モデル作成開始
既知悪性通信先と既知良性通信先を入力
S101
通信先リストへの掲載有無を参照
S102
特徴情報(通信先リストへの掲載有無の時間変化)抽出
S103
特徴情報(通信先リストへの掲載有無の組合せの時間変化)抽出
S104
ドメイン名とIPアドレスの対応関係の履歴情報の構築
S105
IPアドレスの利用状況を示す外部情報の収集
S106
通信先に関連するIPアドレス群の関係を構築
S107
特徴情報(通信先に関連するIPアドレス群の統計量)抽出
S108
ドメイン名の利用状況を示す外部情報の収集
S109
通信先に関連するドメイン名群の関係を構築
S110
特徴情報(通信先に関連するドメイン名群の統計量)抽出
S111
特徴情報の統合
S112
悪性度算出のアルゴリズムの適用
S113
訓練モデルを出力
S114
訓練モデル作成終了

[図20]


通信先の作成度算出開始
悪性度算出対象の通信先と訓練モデルを入力
S201
特徴情報(通信先リストへの掲載有無の時間変化)抽出
S202
特徴情報(通信先リストへの掲載有無の組合せの時間変化)抽出
S203
特徴情報(通信先に関連するIPアドレス群の統計量)抽出
S204
特徴情報(通信先に関連するドメイン名群の統計量)抽出
S205
特徴情報の統合
S206
悪性度算出のアルゴリズムの適用
S207
通信先に対する悪性度を出力
S208
通信先の悪性度算出終了

[図21]


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
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
1110
マウス
1120
キーボード
1130
ディスプレイ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2016/054102

A. CLASSIFICATION OF SUBJECT MATTER
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)
G06F21/55

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2016
Kokai Jitsuyo Shinan Koho 1971-2016 Toroku Jitsuyo Shinan Koho 1994-2016

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	US 2014/0298460 A1 (MICROSOFT CORP.), 02 October 2014 (02.10.2014), claims 1, 7; paragraph [0055] (Family: none)	1-8
A	JP 2012-175296 A (Nippon Telegraph and Telephone Corp. et al.), 10 September 2012 (10.09.2012), claim 1 (Family: none)	1-8

☐ 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
27 April 2016 (27.04.16)	17 May 2016 (17.05.16)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3,Kasumigaseki,Chiyoda-ku, Tokyo 100-8915,Japan	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１６／０５４１０２

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	１９７１－２０１６年
日本国実用新案登録公報	１９９６－２０１６年
日本国登録実用新案公報	１９９４－２０１６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	US 2014/0298460 A1 (MICROSOFT CORPORATION) 2014.10.02, claims 1,7,[0055]（ファミリーなし）	1-8
A	JP 2012-175296 A（日本電信電話株式会社他）2012.09.10, 請求項１（ファミリーなし）	1-8

☐　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
２７．０４．２０１６	１７．０５．２０１６

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　９５５５
宮司　卓佳
電話番号　０３－３５８１－１１０１　内線　３５４６