(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5682089号
(P5682089)
(45)発行日　平成27年3月11日(2015.3.11)　(24)登録日　平成27年1月23日(2015.1.23)

(51)Int.Cl.　　F I
H04L 12/70 (2013.01)　　H04L 12/70 100Z
G06F 13/00 (2006.01)　　G06F 13/00 510A

請求項の数 6　(全 18 頁)

(21)出願番号　特願2011-33932(P2011-33932)
(22)出願日　平成23年2月18日(2011.2.18)
(65)公開番号　特開2012-175296(P2012-175296A)
(43)公開日　平成24年9月10日(2012.9.10)
審査請求日　平成25年7月18日(2013.7.18)

(73)特許権者　000004226
日本電信電話株式会社
東京都千代田区大手町一丁目５番１号
(73)特許権者　899000068
学校法人早稲田大学
東京都新宿区戸塚町１丁目１０４番地
(74)代理人　100107766
弁理士　伊東　忠重
(74)代理人　100070150
弁理士　伊東　忠彦
(74)代理人　100124844
弁理士　石原　隆治
(72)発明者　森　達哉
東京都千代田区大手町二丁目３番１号　日本電信電話株式会社内

最終頁に続く

(54)【発明の名称】通信分類装置及び方法

(57)【特許請求の範囲】
【請求項１】
通常の通信、悪意のある通信を弁別するための通信分類装置であって、
通信を発生させた端末の判定対象アドレスを取得し、該判定対象アドレスを構成するビット列の構造的な性質に基づいて、該判定対象アドレスに固有な特徴を特徴ベクトルとして抽出する特徴ベクトル抽出手段と、
随時または所定の周期で取得した悪意性の有無を示すラベルが付与されたアドレスのリストを格納した訓練データ記憶手段と、
前記訓練データ記憶手段の前記アドレスのリストに対し、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、該特徴ベクトルに対して教師付き機械学習を適用して訓練を実施し、訓練結果を出力するアドレス訓練手段と、
前記アドレス訓練手段の前記訓練結果と前記特徴ベクトル抽出手段で抽出された前記特徴ベクトルを用いて、通信が通常の通信か、または、悪意のある通信かを確率的に判定する判定手段と、
を有し、
前記特徴ベクトル抽出手段は、
前記判定対象アドレスを構成するビット列をサブビット列に分割し、各サブビット列を任意の関数で変換した値を前記特徴ベクトルの要素とする手段を含む
ことを特徴とする通信分類装置。
【請求項２】

通常の通信、悪意のある通信を弁別するための通信分類装置であって、
ある一定の性質を持つアドレスを収集したアドレスリストからなる特徴抽出用アドレス情報を格納するアドレスリスト記憶手段と、
通信を発生させた端末の判定対象アドレスを取得し、該判定対象アドレスを構成するビット列の構造的な性質に基づいて、該判定対象アドレスに固有な特徴を特徴ベクトルとして抽出する特徴ベクトル抽出手段と、
随時または所定の周期で取得した悪意性の有無を示すラベルが付与されたアドレスのリストを格納した訓練データ記憶手段と、
前記訓練データ記憶手段の前記アドレスのリストに対し、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、該特徴ベクトルに対して教師付き機械学習を適用して訓練を実施し、訓練結果を出力するアドレス訓練手段と、
前記アドレス訓練手段の前記訓練結果と前記特徴ベクトル抽出手段で抽出された前記特徴ベクトルを用いて、通信が通常の通信か、または、悪意のある通信かを確率的に判定する判定手段と、
を有し、
前記特徴ベクトル抽出手段は、
前記判定対象アドレスが前記アドレスリスト記憶手段に格納されている前記特徴抽出用アドレス情報に含まれるか否かにより値を定め、特徴ベクトルの構成要素の値とする手段を含む
ことを特徴とする通信分類装置。

【請求項３】
前記アドレス訓練手段及び前記判定手段は、
前記訓練結果として、悪意の有無を２値で表す、または、悪意性の距離もしくは確率を用いたスコアで表現する手段を含む
請求項１または２記載の通信分類装置。

【請求項４】
通常の通信、悪意のある通信を弁別するための通信分類方法であって、
特徴ベクトル抽出手段が、通信を発生させた端末の判定対象アドレスを取得し、該判定対象アドレスを構成するビット列の構造的な性質に基づいて、該判定対象アドレスに固有な特徴を特徴ベクトルとして抽出する特徴ベクトル抽出ステップと、
アドレス訓練手段が、随時または所定の周期で取得した悪意性の有無を示すラベルが付与されたアドレスのリストを格納した訓練データ記憶手段の前記アドレスのリストに対し、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、該特徴ベクトルに対して教師付き機械学習を適用して訓練を実施し、訓練結果を出力するアドレス訓練ステップと、
判定手段が、前記アドレス訓練ステップで出力された前記訓練結果と前記特徴ベクトル抽出ステップで抽出された前記特徴ベクトルを用いて、通信が通常の通信か、または、悪意のある通信かを確率的に判定する判定ステップと、
を行い、
前記特徴ベクトル抽出ステップにおいて、
前記判定対象アドレスを構成するビット列をサブビット列に分割し、各サブビット列を任意の関数で変換した値を前記特徴ベクトルの要素とする
ことを特徴とする通信分類方法。

【請求項５】
通常の通信、悪意のある通信を弁別するための通信分類方法であって、
特徴ベクトル抽出手段が、通信を発生させた端末の判定対象アドレスを取得し、該判定対象アドレスを構成するビット列の構造的な性質に基づいて、該判定対象アドレスに固有な特徴を特徴ベクトルとして抽出する特徴ベクトル抽出ステップと、
アドレス訓練手段が、随時または所定の周期で取得した悪意性の有無を示すラベルが付与されたアドレスのリストを格納した訓練データ記憶手段の前記アドレスのリストに対し

、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、該特徴ベクトルに対して教師付き機械学習を適用して訓練を実施し、訓練結果を出力するアドレス訓練ステップと、

判定手段が、前記アドレス訓練ステップで出力された前記訓練結果と前記特徴ベクトル抽出ステップで抽出された前記特徴ベクトルを用いて、通信が通常の通信か、または、悪意のある通信かを確率的に判定する判定ステップと、

を行い、

前記特徴ベクトル抽出ステップにおいて、

ある一定の性質を持つアドレスを収集したアドレスリストからなる特徴抽出用アドレス情報を格納するアドレスリスト記憶手段を参照し、前記判定対象アドレスが、該アドレスリスト記憶手段に格納されている該特徴抽出用アドレス情報に含まれるか否かにより値を定め、特徴ベクトルの構成要素の値とする

ことを特徴とする通信分類方法。

【請求項６】

前記アドレス訓練ステップ及び前記判定ステップにおいて、

前記訓練結果として、悪意の有無を２値で表す、または、悪意性の距離もしくは確率を用いたスコアで表現する

請求項４または５記載の通信分類方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、通信分類装置及び方法に係り、特に、悪意のあるソフトウェア等の通信を弁別するための通信分類装置及び方法に関する。

【背景技術】

【０００２】

ワームやボットネットと呼ばれるマルウェア（悪意のあるソフトウェア）による被害が拡大・深刻化している。マルウェアに感染したコンピュータはネットワークに接続された他のコンピュータに対して不正あるいは有害な動作を行うことが特徴であり、迷惑メールの大量送信や、サーバへの不正な大量アクセスによるサービス妨害攻撃といった悪質な行動を行うためのツールとして使われる。マルウェアの脅威は外部に対する攻撃のみならず、感染したコンピュータからクレジット番号やアドレス帳などの個人情報を抽出し、外部のコンピュータに送信する活動も存在する。このようなマルウェアによる被害を未然に防ぐためにはマルウェア本体を送受信している悪意のある通信を未然に検出する技術が必要となる。

【０００３】

一般に悪意のある通信を検出する方法として、DPI (Deep Packet Inspection) 技術が広く普及している（例えば、非特許文献１参照）。DPIは時々刻々と到来するパケット群をリアルタイムに分析することによって通信の中身を再構成し、その中身に対して復号化やパターンマッチングを適用することによって悪意のある通信に固有な既知のパターンを発見する手法である。この手法は復号に要する演算リソースが必要であることや、復号が困難な暗号化された通信に対して適用ができないこと、およびパターンが未知である場合には適用できないという問題があった。特に演算リソースに関しては、ネットワーク回線の超高速化が進むに連れ、顕著な問題となることが予想される。

【０００４】

上記問題の解決をサポートする一つの手段として、悪意のある通信を開始する端末に振られているアドレスの評判(レピュテーション)を用いる手法がある(例えば、非特許文献２参照)。ここでアドレスとはIPアドレスのようにネットワーク上の端末等に対して相互に到達性を得る為につけられたアドレスを指す。この方法は悪意のある通信を行う端末のアドレスがある一定のアドレス空間に集中しやすいという性質を利用したものであり、予め評判が悪いあるいは良いアドレスのリストを収集し、得られた評判リストと新たに観測

した通信のアドレスを照合することによって、該当する通信の悪意性を推定的に判定する。当該技術は、用いる情報が端末アドレスだけであるので、簡便で軽量な手法によって超高速回線における適用が可能である。

【先行技術文献】

【非特許文献】

【０００５】

【非特許文献１】John Pirc "Common Network Security Misconceptions: Firewalls Exposed" http://www.sans.edu/resources/securitylab/pirc_john_firewalls.php

【非特許文献２】Commtouch Unveils New IP Reputation Service Based on Global Real-Time Data http://www.commtouch.com/press-releases/commtouch-unveils-new-ip-reputation-service-based-global-real-time-data

【発明の概要】

【発明が解決しようとする課題】

【０００６】

しかしながら、上記非特許文献２のアドレスの評判リストを用いる方法は、いかにアドレスが収集されたかに大きく依存するため、一度アドレスが得られた後にはその精度を改善することが困難であったり、評判の収集時において観測されなかったアドレスに対しては判定ができない問題がある。例えば悪意のある通信に地域的な特徴が存在する場合、地域Aでの観測を元に構成した評判リストが地域Bではあまり役に立たないという問題がある。また、複数のアドレス評判リストや、アドレスの悪意性に関連するいくつかのヒントが得られているとき、各々の情報を用いて統合的な判断を下す一般的な手法はこれまでになかった。

【０００７】

本発明は、上記の点に鑑みなされたもので、過去に観測されなかった未知のアドレスにも対応可能なアドレスの評判判定を行い、結果として悪意のある通信の検出を実現することが可能な通信分類装置及び方法を提供することを目的とする。

【課題を解決するための手段】

【０００８】

上記の課題を解決するため、本発明は、通常の通信、悪意のある通信を弁別するための通信分類装置であって、

通信を発生させた端末の判定対象アドレスを取得し、該判定対象アドレスを構成するビット列の構造的な性質に基づいて、該判定対象アドレスに固有な特徴を特徴ベクトルとして抽出する特徴ベクトル抽出手段と、

随時または所定の周期で取得した悪意性の有無を示すラベルが付与されたアドレスのリストを格納した訓練データ記憶手段と、

前記訓練データ記憶手段の前記アドレスのリストに対し、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、該特徴ベクトルに対して教師付き機械学習を適用して訓練を実施し、訓練結果を出力するアドレス訓練手段と、

前記アドレス訓練手段の前記訓練結果と前記特徴ベクトル抽出手段で抽出された前記特徴ベクトルを用いて、通信が通常の通信か、または、悪意のある通信かを確率的に判定する判定手段と、

を有し、

前記特徴ベクトル抽出手段は、

前記判定対象アドレスを構成するビット列をサブビット列に分割し、各サブビット列を任意の関数で変換した値を前記特徴ベクトルの要素とする手段を含む。

【発明の効果】

【０００９】

上記のように、本発明は、通信を発生させた端末のアドレス構造から特徴ベクトルを抽出し、得られた特徴に対して教師付き機械学習の手法を適用することにより、通信が通常

かあるいは悪意があるかという、通信の悪意性の種別を推定的に判定することにより、過去に観測されなかった未知のアドレスにも対応可能なアドレスの評価判定が可能となり、結果として悪意のある通信を検出することができる。

【図面の簡単な説明】

【００１０】

【図１】本発明の一実施の形態における通信分類装置の構成図である。

【図２】本発明の一実施例の通信分類装置の動作のフローチャートである。

【図３】本発明の一実施例のシステムの適用例である。

【発明を実施するための形態】

【００１１】

以下図面と共に、本発明の実施の形態を説明する。

【００１２】

図１は、本発明の一実施の形態における通信分類装置の構成を示す。

【００１３】

同図に示す通信分類装置１００は、判定対象アドレス受信部１１０、特徴抽出用アドレス受信部１２０、訓練用データ受信部１３０、アドレス特徴抽出部１４０、アドレス訓練部１５０、アドレス判定部１６０、判定出力部１７０、判定結果キャッシュ部１８０、アドレス情報記憶部１０１、訓練元データ記憶部１０２、訓練データ記憶部１０３から構成される。

【００１４】

アドレス情報記憶部１０１は、特徴抽出用アドレス情報受信部１２０により外部から入力されたアドレス情報を格納する。

【００１５】

訓練元データ記憶部１０２は、訓練用データ受信部１３０において、随時または所定の周期で受信されたアドレス毎に悪意性の有無に関するラベルが付与された訓練元データを格納する。

【００１６】

訓練データ記憶部１０３は、訓練元データのアドレスに対して特徴ベクトルを抽出したデータ(訓練データ)を格納する。

【００１７】

判定対象アドレス受信部１１０は、外部から判定対象となるアドレスを受信する。

【００１８】

特徴抽出用アドレス情報受信部１２０は、外部から特徴を抽出するためのアドレスを受信し、アドレス情報記憶部１０１に格納する。

【００１９】

訓練用データ受信部１３０は、随時または所定の周期で外部から訓練用のデータを受信し、訓練元データ記憶部１０２に格納する。

【００２０】

アドレス特徴抽出部１４０は、端末のアドレスを構成するビット列の構造的な性質に基づいて、アドレスに固有な特徴を特徴ベクトルとして抽出する。詳細については後述する。

【００２１】

アドレス訓練部１５０は、訓練元データ記憶部１０２から取得した、予め悪意性のある有無でラベル付けされたアドレスのリストに対し、アドレス毎に、アドレスに固有な特徴を特定することにより特徴ベクトルを抽出し、当該特徴ベクトルに対して公知技術である教師付機械学習の方法を用いて教師データの訓練を実施する。

【００２２】

アドレス判定部１６０は、アドレス訓練部１５０における訓練結果と特徴ベクトルに基づいて、新たに観測したアドレスによる通信が通常であるか、あるいは悪意のある通信であるのかの悪意性を判断する。

【００２３】
判定結果出力部１７０は、アドレス判定部１６０による判定結果を外部及び判定キャッシュ部１８０に出力する。
【００２４】
判定キャッシュ部１８０は、過去の判定結果を保存し、当該判定結果は、判定対象アドレス受信部１１０が受信したアドレスと同じアドレスが判定キャッシュ部１８０の内容と照合する際に利用される。
【００２５】
最初に、アドレス特徴抽出部１４０において、アドレスから特徴ベクトルを抽出する方法について説明する。
【００２６】
アドレス特徴抽出部１４０において、特徴ベクトルを生成する方法として以下に示す４つの手法を示す。
【００２７】
（１）抽出する始点と始点からの抽出個数を指定して特徴ベクトルを構成する方法：
一般に、N 個のビット列 {b1, b2, ..., bN} によって構成されるアドレスに対して、上位ｊ番目からｋ個のビットを用いて構成したビット列 {bj, bj+1, ..., bj+k-1} を特徴ベクトルとして抽出する。
【００２８】
ここで、各ビット bj (j =1, ..., N) は"０"か"１"の値を取り、ｊ，ｋは
$$1 \leqq j < j + k - 1 \leqq N$$
を満たす任意に設定が可能な値である。
【００２９】
ここで、特徴ベクトルの抽出における、始点jと抽出個数kの値の定め方について説明する。
【００３０】
IPv4では経験則から定め方を例示可能だが、その他アドレス体系における具体的な定め方については本発明では範疇外とする。なお、IPv4アドレスの場合は上位ビット（第1～第3オクテット）がネットワークアドレスに近いため、より重要な意味を持ち、始点は j=1、抽出個数は k=24 （1オクテットは8ビットであり、3オクテットは24ビットに相当する）という定め方が一つの実施形態となる。詳細については実施例において後述する。一方、IPv6をはじめ、他のアドレス体系の場合IPv4の経験則が当てはまるとは限らない。
【００３１】
（２）サブビット列に分割し、各サブビット列を任意の関数で値に変換して特徴ベクトルの要素とする方法：
一般に、N 個のビット列 {b1, b2,…, bN} によって構成されるアドレスを、上位から順番に T 個のサブビット列 sj (j=1,…,T) に分割する。当該分割条件を以下に示す。
【００３２】
・T個に分割された各サブビット列 sj のサイズ(サブビット列を構成するビットの総数) Sj は、
$$S1 + S2 + \ldots + ST \leqq N \quad (式1)$$
を満たす、任意の値となるように設定される。
【００３３】
・1番目のサブビット列 s1 の開始位置は、上記(式1)が満たされている限り、任意の値をとることができる。
【００３４】
例えば N = 10個のビット列 {b1, b2, ..., b10} によって構成されるアドレスに対し、
$$s1 = \{b3, b4, b5\}$$
のようにサブビット列を構成することが可能である。

【００３５】
・各サブビット列 sj に対し、各sjを構成する Sj 個のビットに対して、任意に定義可能な関数f(sj)を適用した結果の値を構成要素とする特徴ベクトルを構成する。
【００３６】
上記の分割する個数T, サブビットs1の開始位置, 特徴ベクトルの構成要素を作成するための任意に定義可能な関数f(sj)とその関数の適用方法の定め方としては以下の２つの方法がある。
【００３７】
・任意のサイズ(分割個数T, s1の開始位置)で定義が可能なサブビットに分割する：
・分割されたサブビットから特徴ベクトルを構成できる関数f(sj)を適用する：
なお、特徴ベクトルを構成できる関数f(sj)は、既存の(任意の)関数を利用することとする。例えば、
・10進表記を利用した特徴ベクトル構成方法；
・10進表記した値に対してハッシュ関数を適用する方法；
などがある。
【００３８】
（３）アドレス情報記憶部１０１に格納された数個のアドレスリスト(アドレスリスト数A個)に対し、各アドレスリストが特定のアドレスを含むか否かで値を定め、特徴ベクトルの構成要素の値(要素数A個＝アドレスリストの個数)とする方法：
ここで、アドレスリストとは、ある一定の性質を持つアドレスを収集したものであり、単一のアドレスあるいは複数の連続するアドレスをネットワークプレフィックスによって表記したものから構成される。
【００３９】
例えば、アドレスリストとは、過去の実績に基づいて、悪意のある通信を発生する可能性が高いアドレスのみを収集したリスト、通常の通信を発生する可能性が高いアドレスのみを収集したリスト、ある特定の国に所属するアドレスリスト、ある特定のAS (Autonomous System) に所属するアドレスリスト、ある特定のサブネットワークに所属するアドレスリスト、あるサーバにて観測されたアドレスリスト、等である。以下に、IPv4アドレスの場合を示す。
【００４０】
単一のアドレスの例： 192.168.1.1
ネットワークプレフィックス表記の例：192.168.1.0/24
各アドレスリストは、同一の性質を有するアドレスを収集したものであり、これらのアドレスリストはリアルタイムに更新することが可能である。
【００４１】
（４）上記の（１）～（３）の方法を任意に組み合わせて特徴ベクトルを抽出する方法も可能である。
【００４２】
悪意性判定対象のアドレスが、アドレス情報記憶部１０１のアドレスリストLj (j=1,2,...,A , Aは用意したアドレスリストの個数)のリスト内に含まれるか否かの結果を特徴ベクトルの要素の値として、特徴ベクトル{l1, l2,..., lA} を構成する。このため、特徴ベクトルの要素数は、用意したアドレスリストLj (j=1,2,...,A) のリスト数A個(任意)に対応する。
【００４３】
アドレス特徴抽出部１４０は、悪意性判定対象のアドレスがリストに含まれるか否かを判断する際に、上記のIPv4アドレスの場合では、単一アドレスの場合は該アドレスが一致するか否かで判断し、ネットワークプレフィックスの場合は、該アドレスに対してネットワークプレフィックスによって一意に指定されるサブネットマスクを適用した場合に一致するか否かで判断する。ひとつでも一致した場合にリストに含まれるとする。一般に、悪意性判定対象のアドレスが、アドレスリストLj に含まれるとき、lj = 1 とし、そうでな

いときlj = 0 とする。なお、悪意性判定対象のアドレスの与え方は実施例にて後述する。

【００４４】

次に、アドレス訓練部１５０について説明する。

【００４５】

アドレス訓練部１５０は、予め定めた周期毎に、訓練元データ記憶部１０２に格納されているラベル付きアドレスリスト訓練データ、及び訓練データ記憶部１０３に格納されているラベルと特徴ベクトルからなる訓練データを用いて機械学習の訓練を実施し、訓練結果として、悪意の有無の２値、または、悪性の距離または確率を用いたスコアをアドレス判定部１６０に出力する。訓練元データ記憶部１０２のラベル付きアドレスリスト訓練データは、悪意の有無の可能性を示すラベルとアドレスから構成され、訓練用データ受信部１３０より新たなリストを得ることにより随時更新可能である。訓練データ記憶部１０３は、アドレス訓練部１５０によって生成されたラベル付きアドレスリスト訓練データのアドレスに対して抽出された特徴ベクトルとラベルの組を格納する。

【００４６】

アドレス判定部１６０は、アドレス訓練部１５０から取得した訓練結果と、アドレス特徴抽出部１４０から取得した特徴ベクトルを用いて、判定対象アドレス受信部１１０にて受信した受信アドレス情報に対して、通信が通常であるか否かを判定し、通信の悪意性を確率的に推定する。

【００４７】

判定出力部１７０は、アドレス判定部１６０の判定結果を外部に出力すると共に、判定結果キャッシュ部１８０にも出力する。

【実施例】

【００４８】

以下、本発明の実施例を図面と共に説明する。

【００４９】

本実施例では、アドレスとしてIPv4アドレスを用い、公知の教師付機械学習の手法として２クラスのサポートベクターマシン（以下SVM）を用いるが、本発明の適用範囲はこの限りではない。

【００５０】

図２は、本発明の一実施例の通信分類装置の動作のフローチャートである。

【００５１】

ステップ１０１） 通信分類装置１００の判定対象アドレス受信部１１０は、判定の対象となるアドレス情報を受信する。

【００５２】

ステップ１０２） 判定対象アドレス受信部１１０は、受信したアドレスに対する判定が過去になされていたかを判定結果キャッシュ部１８０に問い合わせる。判定結果が既にキャッシュされている場合はステップ１０３に移行し、キャッシュされていない場合はステップ１０４に移行する。

【００５３】

ステップ１０３） 判定結果キャッシュ部１８０に記録されている判定結果を基に、判定結果を抽出し、結果を判定出力部１７０に出力する。

【００５４】

ステップ１０４） 判定結果キャッシュ部１８０にキャッシュされていない場合は該アドレス情報をアドレス特徴抽出部１４０に出力し、特徴ベクトルを抽出する。

【００５５】

ステップ１０５） 一方、アドレス訓練部１５０では、訓練元データ記憶部１０２の予め準備した通常と悪意を区別するラベル付きの訓練データを用いて機械学習の訓練を実施しておく。この訓練データは、予め定めた周期が到来する毎に更新することができる。また、ラベル付きの訓練データは訓練用アドレスリスト受信部１３０より新たなリストを得

ることによって随時更新可能である。図３に示すように訓練データはネットワークの内部に設置した各種の侵入・攻撃・異常検知システム２３０などの出力を利用し、実際に悪意のある通信を行ったアドレスに対してラベルを"＋１"と付与することによって構成可能である。そのほか、外部で公開しているブラックリストやホワイトリストを訓練データとして用いることが可能である。訓練した結果はアドレス判定部１６０によって利用される。学習や訓練した結果の詳細については後述する。

【００５６】

アドレス特徴抽出部１４０で抽出された特徴ベクトルは、アドレス判定部１６０に出力され、アドレス判定部１６０ではアドレス訓練部１５０より得た訓練結果を用いて、アドレス受信部１１０にて受信した該受信したアドレス情報に対して、通信が通常であるか、悪意があるかの通信の悪意性を確率的に推定し、出力値ｙを出力する。

【００５７】

ステップ１０６）　アドレス判定部１６０の出力値ｙ＞０であれば、ステップ１０８に移行し、ｙ≦０であればステップ１０７に移行する。

【００５８】

ステップ１０７）　悪意はないと判定し、その結果を判定出力部１７０と判定結果キャッシュ部１８０に出力する。

【００５９】

ステップ１０８）　悪意があると判定し、その結果を判定出力部１７０と判定結果キャッシュ部１８０に出力する。

【００６０】

以下に、上記のステップ１０４の特項ベクトルの抽出方法について説明する。

【００６１】

以下ではIPv4アドレス"192.168.5.88 "を例として特徴ベクトルを抽出する方法を例示する。

【００６２】

一般にIPv4 アドレスは32ビット長で定義され、広く用いられている"192.168.5.88 "という表記は32ビットを4個の8ビットに区切り、各々の8ビットの値を10進表記したものを". "で結合したものである。

【００６３】

（１）抽出する始点と始点からの抽出個数を指定して特徴ベクトルを構成する方法：

"192.168.5.88 "を2進数表記すると

11000000101010000000010101011000

のように32個のビットから構成されるビット列となる。このうち上位 j 番目から k 個のビットを用いて構成したビット列を特徴ベクトルとする。

【００６４】

j=1, k=24 とすると、特徴ベクトルは24個の要素を持ち、

{1,1,0,0,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,1,0,1}

となる。

【００６５】

なお、IPv4アドレスの場合は、実験による経験則から、上位ビット（第1～第3オクテット）がネットワークアドレスに近いため、より重要な意味を持ち、始点は j=1、抽出個数は k=24 という定め方が一つの実施形態となりえる。

【００６６】

（２）サブビット列に分割し、各サブビット列を任意の関数で値に変換して特徴ベクトルの要素とする方法：

一般に、N 個のビット列 {b1, b2, ..., bN} によって構成されるアドレスを上位から順番に T 個のサブビット列 sj (j=1,...,T) に分割する。

ここでは前述と同じ"192.168.5.88 "を用いると、元のビット列は、

11000000101010000000010101011000

である。N = 32, T = 5 とし、サブビット列を
s1 = 11000000
s2 = 1010
s3 = 1000
s4 = 0000
s5 = 0101
と定義する。サブビット列 sj (j=1,2,..,4) のサイズは
S1 = 8, S2 = S3 = S4 = S5 = 4
である。

【００６７】
各々のサブビット列 sj に対して、適用する関数 f(sj) を「sj を10進表記する」と定義した場合、
{192, 10, 8, 0, 5}
のように特徴ベクトルを構成できる。

【００６８】
なお、各サブビット列に関数 f(sj)を適用した値は次のとおりとなる。

【００６９】
$f(s1) =(11000000)_2 = (192)_{10}$, $f(s2) = (1010)_2 = (10)_{10}$, $f(s3) = (1000)_2 = (8)_{10}$, $f(s4) = (0000)_2 = (0)_{10}$,
$f(s5) = (0101)_2 = (5)_{10}$
あるいは関数 f(sj) を「{s1,s2,...,sj} のビットを結合して得られたビットを10進表記する」と定義した場合、
{192, 3082, 49320, 789120, 12625925}
のように特徴ベクトルを構成できる。

【００７０】
なお、各サブビット列に関数 f(sj)を適用した値は次のとおりとなる。

【００７１】
$f(s1) =(11000000)_2 = (192)_{10}$,
$f(s2) = (110000001010)_2 = (3082)_{10}$,
$f(s3) = (1100000010101000)_2 = (49320)_{10}$,
$f(s4) = (11000000101010000000)_2 = (789120)_{10}$,
$f(s5) = (110000001010100000000101)_2 = (12625925)_{10}$
あるいは関数を以下のような特徴ベクトルを構成する関数として定義することもできる。
例えば、簡単のため、T=1 とし、s1 = 110000001010100000000101 とする。サブビットのサイズは S1 = 24 である。24ビット長の最大値は10進表記で 2^{24}=16,777,216 であるので、サブビットs1の10進表記 12,625,925 を用い、12,625,925 番目のビットのみが"1"で他はすべて"0"の粗な特徴ベクトルを構成することもできる。

【００７２】
上記のバリエーションとして、
s1 = 11000000
s2 = 10101000
s3 = 00000101
とした場合、それぞれのサブビットは8ビット長であるため、0から255の256通りの値をとることができる。

【００７３】
それらの値を利用して 256 x 3 = 768 通りの要素からなる特徴ベクトルを構成する。関数 f(sj)として、最初の256ビットのうち、s1 を10進表記した値である192番目のビットを1、次の256ビットのうち、s2を10進表記した値である168番目のビットを1、最後の256ビットのうち、s3を10進表記した値である5番目のビットを1とし、残りのすべてのビッ

トを0とする関数を定義することもできる。

【００７４】

上述のとおり、特徴ベクトルを構成するために、分割されたサブビット列に適用される関数f(sj)は、既存の任意の関数が利用され、例えば、「10進表記を利用した特徴ベクトル構成方法」、「10進表記した値に対してハッシュ関数を適用する方法」などがある。

【００７５】

（３）アドレス情報記憶部１０１に格納された数個のアドレスリスト(アドレスリスト数A個)に対し、各アドレスリストが特定のアドレスを含むか否かで値を定め、特徴ベクトルの構成要素の値(要素数A個＝アドレスリストの個数)とする方法：

別途アドレス情報記憶部１０１に用意したA種類のアドレスデータ L1,L2,...,LA を用いてA個の要素からなる特徴ベクトルを構成する方法を述べる。

【００７６】

表1にアドレスデータの例を示す。これは日本のIPv4アドレスの一部を抽出した例である。

【００７７】

【表１】

アドレスデータの例

1.21.0.0/16
1.33.0.0/16
1.66.0.0/15
1.72.0.0/13
1.112.0.0/14
14.0.8.0/22

上記の表１のアドレスデータは、ネットワークプレフィックス表記である 1.21.0.0/16 は、1.21.0.0 から 1.21.255.255 までの 65536 個のアドレス集合に対応する。

【００７８】

このようなアドレスデータは、特徴抽出用アドレス受信部１２０を通じて通信分類システムの外部から取り込み、アドレス情報記憶部１０１に格納される。

【００７９】

アドレス特徴抽出部１４０は、悪意性判定対象のアドレスが、アドレス情報記憶部１０１の Lj (j=1,2,...,A) のデータ内に含まれるか否かでベクトル{l1, l2,..., lA} を構成し、悪意性判定対象のアドレスの特徴として抽出する。例えばA =10であり、悪意性判定対象のアドレスが 1, 2, 10番目のデータにアドレス情報記憶部１０１の掲載のアドレスに合致するとき、

{1,1,0,0,0,0,0,0,0,1}

のように、特徴ベクトルを抽出する。

【００８０】

該アドレスがアドレス情報記憶部１０１のアドレスデータに合致するか否かの判断を上記の表1の例を用いて説明する。

【００８１】

悪意性判定対象のアドレスが"1.21.3.1"であったとすると、これは表１における"1.21.0.0/16"で表記するアドレス集合に包含されるので、合致すると判断する。従って、特徴

ベクトルは、{1,0,0,0,0,0,0,0,0,0}となる。

【００８２】

判定対象のアドレスが"4.1.2.3"であったとすると、表1で示されるアドレスデータに含まれないので合致しないと判断する。従って、特徴ベクトルは、{0,0,0,0,0,0,0,0,0,0}となる。

【００８３】

（４）上記の（１）～（３）の方法を任意に組み合わせて特徴ベクトルを抽出する方法：

例えば、上記の（１）の方法と（３）の方法を組み合わせ、"192.168.5.88"に対して、前述の（１）の例(j=1, k=24の特徴ベクトル)である、

{1,1,0,0,0,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,1}

と、同様に（３）の最初の例である、『A =10であり、悪意性判定対象のアドレスが 1, 2, 10番目のリストに掲載のアドレスに合致する』からなる特徴ベクトル

{1,1,0,0,0,0,0,0,0,0,1}

とを組み合わせ、

{1,1,0,0,0,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,1,1,1,1,0,0,0,0,0,0,0,0,1}

のように34次元の特徴ベクトルを構成することが可能である。本実施例ではアドレスデータとして国情報に相当する情報を用いたが、この他BGP経路情報あるいはWhoisといった外部公開されているIPアドレスに関する情報を利用することができる。

【００８４】

次に、アドレス訓練部１５０の処理について説明する。

【００８５】

以下では、２クラスのSVMを適用してアドレスを訓練する手順を例示する。

【００８６】

SVMは入力データを高次元に写像した上でデータを識別する超平面を構築することである。写像された高次元空間において線形分離を試みる。

【００８７】

訓練データを用いた学習では、超平面による識別境界と訓練データ間の距離、すなわちマージンを最大化するようパラメタを最適化する。

【００８８】

はじめに、訓練元データ記憶部１０２に格納されている訓練データの元となるデータ（訓練元データ）の例を表２に示す。ここでラベルは２値{－１,＋１}をとり、"－１"なら通常である可能性が高いアドレス、"＋１"なら悪意のある可能性が高いアドレスであると定義する。

【００８９】

【表２】

訓練元データ

ラベル	アドレス
+1	192.168.5.88
−1	10.23.10.7
−1	129.60.21.1
....	....

前記の訓練元データに記載の各々のアドレスに対して特徴ベクトルを抽出することによ

って表３のような訓練データを生成し、訓練データ記憶部１０３に格納する。ここでは特徴ベクトルは24次元であり、上位24ビットのビット列によって構成した。

【００９０】

【表３】

訓練データ

ラベル	特徴ベクトル
+1	{1,1,0,0,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,1,0,1}
−1	{0,0,0,0,1,0,1,0,0,0,0,1,0,1,1,1,0,0,0,0,1,0,1,0}
−1	{1,0,0,0,0,0,0,1,0,0,1,1,1,1,0,0,0,0,0,1,0,1,0,1}
....	....

訓練データにおけるi番目のサンプルの特徴ベクトルを

Xi = {xi1,xi2,...,xi24}

と表記する。同様にi番目のサンプルのラベルを ci と表記する。例えば、1番目の特徴ベクトルとラベルの値は

X1 = {1,1,0,0,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,1,0,1}

c1 = +1

である。

【００９１】

またSVMのパラメタをベクトル W = {w1,w2,...,w24} と定義する。ここではXi, W共に行ベクトルである。

【００９２】

訓練データからマージン最大化を実現するようなパラメタWを学習するには、すべてのi=1,2,...,M に対し、

【００９３】

【数１】

$$ci(W^T\varphi(X_i)+b)\geq 1$$

という条件の下で、目的関数‖W‖^{2}を最小とするような W を求める事（二次計画法）であり　例えば、文献「J.C Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization." Advances in Kernel Methods - Support Vector Learning, B. Scholkopf, c. Burges,and A. Smola, eds., pp. 185-208, MIT Press, (1999)」に記載のSequential minimal optimization(SMO) などの数値的解析の手法によって高速に算出する事ができる。ここで、ｂはバイアスパラメタである。W^TはWの転置行列φ(X_i)

はベクトル Xi に特徴空間の変換を施す非線形関数を適用したものである。‖W‖はベクトル W のユークリッドノルムである。M は訓練データにおけるサンプル数である。

【００９４】

特徴空間の変換を施す非線形関数φ(X_i)は既存技術を利用する。SVM においてはカーネルトリックと呼ばれる技術を利用することにより、φ(X_i)自体を明示的に定義することな

く、マージン最大化の計算を行うことが可能である。すなわち$\phi(X_i)$を直接定義せずに、一般的に下記のような$\phi(X_i)$の内積で定義されるカーネル関数 k(x, x')を用いて定義する。

【００９５】

$$k(X_i,X_i') = \phi(X_i)^T \phi(X_i')$$

ここで、$\phi(X_i)^T$は$\phi(X_i)$の転置行列である。

【００９６】

よく利用されるガウスカーネルは、次の式で定義される。

【００９７】

$$k(Xi,Xi') = \exp(-\|Xi-Xi'\|^2 / 2\sigma^2)$$

このようなカーネル関数を導入すると、マージン最大化はカーネル関数によって表現される二次計画法を解く問題に帰着する。

【００９８】

次に、アドレス判定部１６０の処理について説明する。

【００９９】

SVM を適用することにより、アドレスの通信の悪意性の有無を判定可能である。新たに観測したアドレスに対して、アドレス特徴抽出部１４０において特徴ベクトルを抽出し、X = {x1,x2,...,x24} を得たとする。アドレス判定部１６０では、

$$y = W^T \phi(X) + b$$

を計算し、y が正であれば悪意性がある、負の値であれば通常のアドレスである、と判定する。ここでW, b は前記のアドレス訓練部１５０で計算済みの値を用いる。なお、 y の計算はカーネル関数を用いて、

【０１００】

【数２】

$$y = \sum_i a_i c_i k(X, X_i) + b$$

と計算することができる。ここでa_iは二次計画法を解くために導入する制約式ごとに定義したラグランジュ乗数（正数）であり、訓練時に数値的に求めることができる。

【０１０１】

次に、図２に沿って新規のアドレスを受信してからそのアドレスを判定するまでの処理を説明する。なお、アドレス判定において、通信の悪意性の度合いを示す要素である「ラベル」の値の表現方法として、前述したように「悪意性の有/無」＝有る(+１)／無い(-1)の2値で示す方法の他、悪意性の距離もしくは確率というスコア(連続値)で表現する方法もとり得る。ここで、悪意性の距離の測り方、確率の測り方については文献「John C. Platt "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods", ADVANCES IN LARGE MARGIN CLASSIFIERS, pp. 61-74, 1999 http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.1639」に記載の方法を本発明に適用した例を示す。ある観測したアドレスに悪意性がある状態をC=1 と書くことにする。Cはクラスの意である。同様にあるアドレスに対してC=－1であればそのアドレスに悪意性が存在しない状態を示す。

【０１０２】

あるアドレスの特徴ベクトルに対してSVMのアドレス判定によって算出した出力値 y を得たとする(ステップ１０５)。yが所与である条件の下で、実際にそのアドレスに悪意性がある確率（条件付き確率）を次式に示すようにシグモイド関数でパラメトリックにモデル化し、シグモイド関数のパラメタA, B を訓練データからフィッティングする。

【０１０３】
P(C=1 | y) = 1 / (1 + exp(Ay+B))
フィッティングの方法は下記の通りである。はじめに訓練集合のCi (i=1,2,…)を用い、
ti = (yi+1)/2
という変数を定義する。A, B は次式で表現される訓練データの負の対数尤度を最小化する値を数値計算によって求める事ができる。
【０１０４】
【数３】

$$-\sum_i t_i \log(p_i) + (1-t_i)\log(1-p_i)$$

ただし、
$p_i = 1/(1 + \exp(Ay_i + B))$
であり、y_i は訓練データx_iに対してSVMの出力値 y を計算した結果である。
【０１０５】
以上が本願発明の各構成要素の例であるが、アドレス判定部１６０において、悪意性判定対象のアドレスを判定する方法の使用例として、次のようなケースがある。
【０１０６】
＜ケース１＞
ゲートウェイルータ等で、外部インターネットから内部組織への通信の送信元IPアドレスを得ることができるため、ルータが観測したすべてのアドレスを、本装置１００への入力値である悪意性判定対象アドレスとし、悪意性があると判断したアドレスをフィルターするケース。
【０１０７】
＜ケース１の例＞
本発明の通信分類装置１００はネットワーク上のゲートウェイルータ２１０と連携をする形で利用される。図３に構成例を示す。ネットワークの管理者は予め本システムのアドレス訓練部が保持する訓練元データを準備し、通信分類装置１００に投入する。この訓練元データはアドレス毎に悪意性の有無に関するラベルが付与されたものである。同様にネットワーク内部の侵入・異常・攻撃検知システム２３０の出力結果を訓練元データ(訓練元データ記憶部１０２)として利用することができる。本装置１００では前述のアドレス訓練を実施し、アドレス判定が可能な状態を保持する。
【０１０８】
ゲートウェイルータ２１０で判定の対象とするアドレスが観測されたら、そのアドレス情報を本発明の通信分類装置１００に送信する。本発明の通信分類装置１００はアドレス受信部１１０にてアドレス情報を受信し、アドレス特徴抽出部１４０へと転送する。アドレス特徴抽出部１４０が特徴ベクトルの抽出を行い、特徴ベクトルデータをアドレス判定部１６０へと転送する。アドレス判定部１６０では前記得られた特徴ベクトルに対してアドレス判定を行い、結果を判定結果送信部(図示せず)へ送信する。判定結果送信部は前記ルータ２１０に対して判定結果データを送信する。ルータ２１０は判定結果に基づいて該アドレスから送信されたパケットに対して適切な処理を行う。例えば悪意性があるパケットに対しては送信レートに制限をかけたり、パケットを廃棄したりする。前記の処理はルータの処理負荷によって変更しても構わない。
【０１０９】
＜ケース２＞
あるプログラムが、特定のアドレスからの通信を受信した際に、その通信を受け取るか

否かを判断するために、通信分類装置１００の出力である悪意性判定結果を利用するケース。

【０１１０】

＜ケース２の例＞

あるサーバ上で動作するプログラムがあるIPアドレスを持つクライアントからリクエストを受けたものとする。この際にそのリクエストをプログラムが実際に処理する前に本発明の通信分類装置１００に問い合わせをし、悪意性のあるアドレスであることが判明したらリクエストを廃棄するか、サーバの負荷が高い場合は他の通常のリクエストを優先し、悪意性のあるアドレスからのリクエストを最低優先とする。

【０１１１】

なお、図１に示す通信分類装置の構成要素の各動作をプログラムとして構築し、通信分類装置として利用されるコンピュータにインストールして実行させる、または、ネットワークを介して流通させることが可能である。

【０１１２】

本発明は、上記の実施の形態及び実施例に記載した方法に限定されることなく、特許請求の範囲内において、種々変更・応用が可能である。

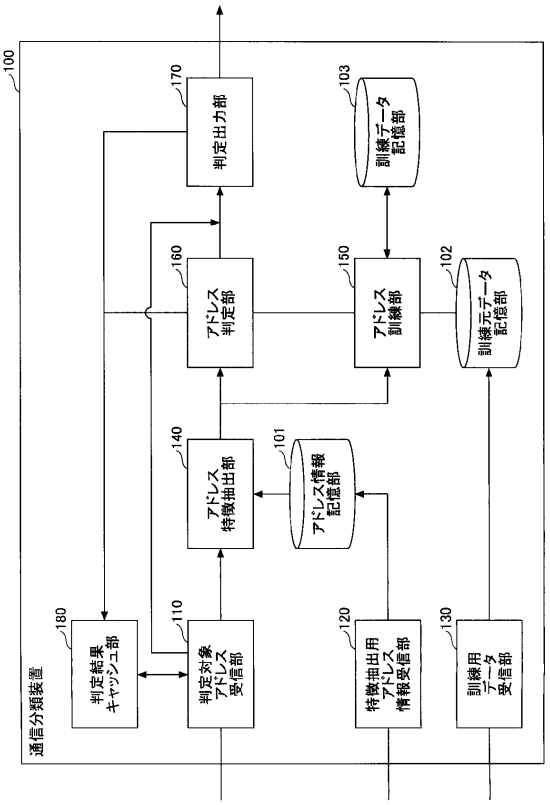
【符号の説明】

【０１１３】

１００　通信分類装置
１０１　アドレス情報記憶部
１０２　訓練元データ記憶部
１０３　訓練データ記憶部
１１０　判定対象アドレス受信部
１２０　特徴抽出用アドレス情報受信部
１３０　訓練用データ受信部
１４０　アドレス特徴抽出部
１５０　アドレス訓練部
１６０　アドレス判定部
１７０　判定出力部
１８０　判定結果キャッシュ部

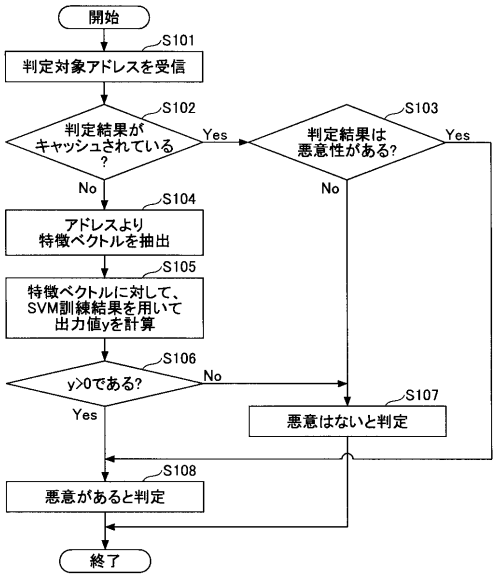
【図１】

本発明の一実施の形態における通信分類装置の構成図



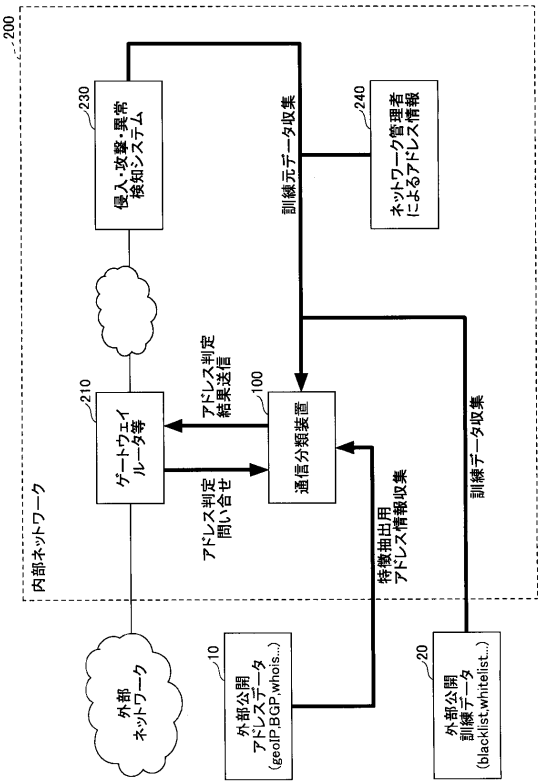
【図２】

本発明の一実施例の通信分類装置の動作のフローチャート



【図３】

本発明の一実施例のシステムの適用例



フロントページの続き

(72)発明者　千葉　大紀
東京都新宿区戸塚町１丁目１０４番地　学校法人早稲田大学内
(72)発明者　後藤　滋樹
東京都新宿区戸塚町１丁目１０４番地　学校法人早稲田大学内

審査官　廣川　浩

(56)参考文献　特開２０１１－０３４４１６（ＪＰ，Ａ）
澤谷　雪子，メッセージ本文受信前でのスパムメール検知方式の精度向上に関する一検討，電子情報通信学会技術研究報告，日本，社団法人電子情報通信学会，２００９年１１月　６日，第109巻 第285号，pp.19～24

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　１２／００－１２／９５５
Ｇ０６Ｆ　１３／００