(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2006-260515
(P2006-260515A)
(43)公開日　平成18年9月28日(2006.9.28)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 13/00	(2006.01)	G06F	13/00	610Q	5K030
H04L 12/58	(2006.01)	H04L	12/58	100F	

審査請求 未請求　請求項の数 21　ＯＬ　（全 40 頁）

(21)出願番号　特願2005-150811 (P2005-150811)
(22)出願日　平成17年5月24日 (2005.5.24)
(31)優先権主張番号　特願2005-39351 (P2005-39351)
(32)優先日　平成17年2月16日 (2005.2.16)
(33)優先権主張国　日本国(JP)

特許法第３０条第１項適用申請有り　２００４年８月２０日　社団法人電子情報通信学会発行の「ＦＩＴ２００４　第３回　情報科学技術フォーラム　プログラム、講演論文集（ＣＤ－ＲＯＭ）、一般講演論文集　第４分冊」に発表

(71)出願人　304027349
国立大学法人豊橋技術科学大学
愛知県豊橋市天伯町雲雀ヶ丘１－１
(74)代理人　100103045
弁理士　兼子　直久
(74)代理人　100127605
弁理士　伊藤　愛
(74)代理人　100129447
弁理士　橋本　努
(72)発明者　山崎　仁
愛知県豊橋市天伯町雲雀ヶ丘１－１
国立大学法人豊橋技術科学大学内

最終頁に続く

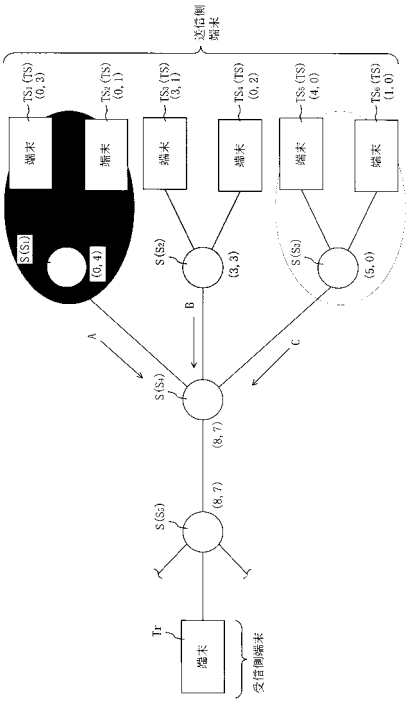
(54)【発明の名称】電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステム

(57)【要約】

【課題】　電子メール中に含まれる情報に依存することなく、配信経路上における中継装置が過去に中継した迷惑メールの頻度に基づいて、迷惑メールをフィルタリングすることが可能な電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムを提供すること。

【解決手段】　本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、配信経路上において特定された中継装置の迷惑メール頻度情報に基づいて迷惑メールであるか否かの判定を行うので、配信経路上を通って配信される電子メールが迷惑メールである場合にそれを確実に検出できる。

【選択図】　図８



【特許請求の範囲】
【請求項１】
送信元から送信された不特定多数の電子メールに含まれる迷惑メールのフィルタリングを実行させるための電子メールフィルタリングプログラムにおいて、
送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得ステップと、
その中継装置アドレス取得ステップにおいて取得された中継装置に対し、その中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得ステップと、
その迷惑メール中継確率取得ステップにおいて得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得ステップと、
その迷惑メール受信確率取得ステップにおいて得られた確率に応じて前記電子メールを所定区分に分類するメール判定ステップとを備えていることを特徴とする電子メールフィルタリングプログラム。
【請求項２】
前記メール判定ステップは、
前記迷惑メール受信確率取得ステップにおいて得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定ステップと、
前記迷惑メール受信確率取得ステップにおいて得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定ステップとを含み、
前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて判定された結果に応じて、前記中継装置アドレス取得ステップにおいて取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新ステップを備えていることを特徴とする請求項１記載の電子メールフィルタリングプログラム。
【請求項３】
前記メール判定ステップは、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識ステップをさらに含むことを特徴とする請求項２記載の電子メールフィルタリングプログラム。
【請求項４】
前記メール判定ステップは、前記認識ステップにおいて不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定ステップをさらに含むことを特徴とする請求項３記載の電子メールフィルタリングプログラム。
【請求項５】
前記中継装置アドレス取得ステップにより取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認ステップを備えていると共に、
そのアドレス確認ステップにより前記アドレスが不正なアドレスであると確認された場合には、前記メール判定ステップにおいて、前記送信元から送信された電子メールが迷惑メールであると判定することを特徴とする請求項２から４のいずれかに記載の電子メールフィルタリングプログラム。
【請求項６】

前記中継装置アドレス取得ステップにより取得された中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、前記迷惑メール中継確率取得ステップにおいて、その新出の中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項１から５のいずれかに記載の電子メールフィルタリングプログラム。

【請求項７】

前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得ステップと、

その中継ルータアドレス取得ステップにより得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得ステップとを備えていることを特徴とする請求項１から６のいずれかに記載の電子メールフィルタリングプログラム。

【請求項８】

送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることができる電子メールフィルタリング方法において、

送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得手段と、

その中継装置アドレス取得手段により取得された中継装置に対し、その中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、

その迷惑メール中継確率取得手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、

その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えていることを特徴とする電子メールフィルタリング方法。

【請求項９】

前記メール判定手段は、

前記迷惑メール受信確率取得手段により得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、

前記迷惑メール受信確率取得手段により得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、

前記迷惑メール判定手段又は前記正当メール判定手段により判定された結果に応じて、前記中継装置アドレス取得手段により取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えていることを特徴とする請求項８記載の電子メールフィルタリング方法。

【請求項１０】

前記メール判定手段は、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段により前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えていることを特徴とする請求項９記載の電子メールフィルタリング方法。

【請求項１１】

前記メール判定手段は、前記不確定メール認識手段により不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えていることを特徴とする請求項１０記載の電子メールフィルタリ

ング方法。
【請求項１２】
前記中継装置アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、
そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定することを特徴とする請求項９から１１のいずれかに記載の電子メールフィルタリング方法。
【請求項１３】
前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段により取得された中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項８から１２のいずれかに記載の電子メールフィルタリング方法。
【請求項１４】
前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、
その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段によりアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段とを備えていることを特徴とする請求項８から１３のいずれかに記載の電子メールフィルタリング方法。
【請求項１５】
電子メールを伝送可能な経路上において、送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることが可能な電子メールフィルタリングシステムにおいて、
送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得手段と、
１の中継装置に対し、その中継装置によって過去に中継された迷惑メール及び正当なメールの各頻度を示すメール情報を記憶する情報記憶手段と、
前記中継装置アドレス取得手段により取得された中継装置に対し、情報記憶手段に記憶されている前記メール情報に基づいて、その中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、
その迷惑メール中継確率演算手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、
その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えていることを特徴とする電子メールフィルタリングシステム。
【請求項１６】
前記メール判定手段は、
前記迷惑メール受信確率取得手段により得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、
前記迷惑メール受信確率取得手段により得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、
前記迷惑メール判定手段又は前記正当メール判定手段による判定結果に応じて、前記中継装置アドレス取得手段により取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えていることを特徴とする請求項１５記載の電子メールフィルタリングシステム。
【請求項１７】

前記メール判定手段は、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段による判定が前記迷惑メール又は前記正当なメールのいずれにも非該当である電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えていることを特徴とする請求項１６記載の電子メールフィルタリングシステム。

【請求項１８】

前記メール判定手段は、前記不確定メール認識手段において不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えていることを特徴とする請求項１７記載の電子メールフィルタリングシステム。

【請求項１９】

前記中継装置アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、

そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定するものであることを特徴とする請求項１６から１８のいずれかに記載の電子メールフィルタリングシステム。

【請求項２０】

前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段により取得された中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項１５から１９のいずれかに記載の電子メールフィルタリングシステム。

【請求項２１】

前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、

その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段においてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段とを備えていることを特徴とする請求項１５から２０のいずれかに記載の電子メールフィルタリングシステム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、不特定多数の送信元から送信された多数の電子メールの中から迷惑メールをフィルタリングすることが可能な電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムに関する。

【背景技術】

【０００２】

近年において、ネットワークの発展により、誰しもが気軽に簡単に電子メール（以下、必要に応じて単に「メール」と称する）を送受信できるようになったことに伴い、所謂スパムメール(spam mail)の数も増大している。ここで、「スパムメール」とは、受信者の意図を無視して事前の要請や同意なしに、無差別かつ大量発信されるメールを意味するものである。なお、このスパムメールの同義語として、「迷惑メール」、「ジャンクメール」、「ＵＣＥ(Unsolicited Commercial Email)」、「ＵＢＥ（Unsolicited Bulk Email）」などがある。

【０００３】

このようなスパムメールは、添付ファイルなどによるウイルス感染や、不要なメールの

増加による受信者の業務生産性及び効率の低下や、トラフィックの増加によるサーバ及びネットワークへの負荷増大や、詐欺サイトへの誘導などによるプライバシーや機密情報の漏洩などの点において、個人及び団体を問わずに脅威となり得るものである。

【０００４】

上記のようなスパムメールによる問題は既に社会問題の域にまで達している。メールアドレスが安価に入手可能であることや、定額料金の高速通信が安価で提供されていることなどを鑑みると、今後、スパムメールは減少することなくますます増加していくと考えられ、スパムメールに対する有効な対策が早急に要求されている。

【０００５】

現在、使用又は提案されているスパムメール対策としては、送信者を特定する技術や、受信メールをフィルタにかけて選別する方法などがある。

【０００６】

送信者を特定する技術をスパムメール対策として用いた場合、送信元を特定することによって、差出人を偽るスプーフィングやフィッシング（Phishing）などのメールを受信前に見分けることが可能となる。しかし、その反面で、ドメインを偽装しないスパム業者のメールは排除することができないという問題がある。

【０００７】

また、受信メールの選別に用いられるフィルタとして代表的なフィルタとしては、アドレスフィルタやテキストフィルタなどが挙げられる。ここで、アドレスフィルタは、メールに記されているメールアドレス（例えば、example@xxx.ac.jp）やＩＰアドレス（例えば、133.aa.bbb.cc）に基づいてメールを選別するフィルタである。

【０００８】

アドレスフィルタとしては、ブラックリスト(受信拒否リスト)に掲載された送信者からのメールを排除するブラックリストフィルタや、ホワイトリスト(受信許可リスト)に予め登録されることによって明示的に承認された送信者からのメールだけを受信するホワイトリストフィルタなどがある。

【０００９】

例えば、特開２００１－１５６８３４号公報（特許文献１）には、電子メールが到着した際に、その電子メールの差出人を、ユーザによって登録されたホワイトリスト又はブラックリストと対比させ、許可された電子メールのみをファクシミリ装置へ送出可能とするＦＡＸサーバシステムが開示されている。

【００１０】

一方で、テキストフィルタは、指定されたヘッダフィールドもしくはヘッダ全体、又は本文に含まれる文字列や文法規則に基づいてメールを選別するフィルタである。このテキストフィルタによれば、例えば、ヘッダフィールドの「Subject:」(題名)に「未承諾広告」などの特定の文字列を検出した場合にスパムメールとして検出するように設定することができる。また、ＲＦＣ（Request for Comments）の規定に則したヘッダであるかなどの文法規則をチェックした場合に、不適切なものが検出された場合にスパムメールとして検出するように設定することができる。

【００１１】

また、最近では、２００２年にPaul Grahamによって提案された、ベイズ理論を用いるベイジアンフィルタが有名である。ベイジアンフィルタとは、ベイズ単語分布フィルタとも呼ばれ、スパムメールに出現する単語とハムメールに出現する単語の出現確率の違いを利用したフィルタリング手法である。ベイジアンフィルタは、過去の情報を利用する学習型のフィルタであるので、学習するほど判定精度が向上するフィルタである。

【特許文献１】特開２００１－１５６８３４号公報

【発明の開示】

【発明が解決しようとする課題】

【００１２】

しかしながら、アドレスフィルタは管理（メンテナンス）の作業が煩雑である上に、管

理するアドレスの件数の多少に応じて、メールが過剰に受信拒否されたり、逆にスパムメールを容易に通過させたりする事例が生じ得るという問題点があった。

【００１３】

例えば、ブラックリストフィルタは、ブラックリストへの登録や削除などのメンテナンス作業が煩雑である。また、ブラックリストフィルタは、ブラックリストに登録されたメールアドレスのみが受信拒否されるので、メールアドレスが偽装されると、容易にブラックリストフィルタを通過してしまうことになる。

【００１４】

また、このブラックリストフィルタが、特定のＩＰアドレスの範囲にある送信元からの送信であった場合や特定の国の送信元からの送信であった場合に受信を拒否するものであれば、受信すべきハムメールまでも過剰に拒否されかねない。

【００１５】

一方で、ホワイトリストフィルタもまた、ホワイトリストへの登録や削除などのメンテナンス作業が煩雑である。また、スパムメールのアドレスがホワイトリストに登録されたメールアドレスと一致するように偽装された場合には、スパムメールがホワイトリストフィルタを通過し、ユーザの元に届いてしまうことになる。さらに、ホワイトリストフィルタが、送信元が特定のＩＰアドレスの範囲にあるメールの受信を許可するものである場合には、スパムメールを完全に受信拒否することができない。

【００１６】

また、テキストフィルタは、フィルタリングのルールを一つ一つ追加（学習）させる作業が煩雑である上に、例えば、ルールとして「未承諾広告」という文字列を追加したとしても、その文字列が「未承諾＿広告」のように偽装された場合にはフィルタリングされないなど、判定精度が低いという問題点があった。

【００１７】

上述したベイジアンフィルタにもいくつかの問題点がある。例えば、スパム単語データベースの更新に時間がかかることや、学習し続けることでそのデータベースのディスク消費量が他のフィルタよりも大きい。また、販売や広告などのメールを望んで受け取っている場合(オプトイン)であっても、そのメールがスパムメールと認識される確率が非常に高い。さらに、誤認識され易い公告付きメールなどをハムメールとして学習させると、今度はスパムメールがハムメールと認識されてしまうなど、学習サンプルによって判定精度に差が生じる。また、構文解析の難しい日本語で書かれたメールでの判定精度は、英語で書かれたメールに比べて低く、誤判定される傾向にある。

【００１８】

ここで、ベイジアンフィルタの最も重大な問題点は、正当な送信者に対しても、メールで使用可能な語句を制限してしまうことである。つまり、ベイジアンフィルタは、正当であり重要なメールであっても、スパムメールに度々使用される語句が使用されているメールをスパムメールと認識してしまう。一方で、ハムメールに頻繁に使用される語句を用いてスパムメールを構成した場合には、ベイジアンフィルタはそのメールをスパムメールとして認識しない。

【００１９】

上記のようなベイジアンフィルタの問題点を解消する方法としては、アドレスフィルタであるホワイトリストに受信を許可するユーザを登録して、両フィルタを併用する方法があるが、上記したようなホワイトリストの問題点によって十分な効果を上げることができない。

【００２０】

本発明は、上述した問題点を解決するためになされたものであり、電子メール中に含まれる情報（アドレスやテキストなど）に依存することなく、配信経路上における中継装置が過去に中継した迷惑メール（スパムメール）及び正当なメール（ハムメール）の各頻度に基づいて、不特定多数の送信元から送信された多数の電子メールの中から迷惑メールをフィルタリングすることが可能な電子メールフィルタリングプログラム、電子メールフィ

ルタリング方法、電子メールフィルタリングシステムを提供することを目的としている。
【課題を解決するための手段】
【００２１】
この目的を達成するために、請求項１記載の電子メールフィルタリングプログラムは、送信元から送信された不特定多数の電子メールに含まれる迷惑メールのフィルタリングを実行させるためのプログラムであり、このプログラムは、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得ステップと、その中継装置アドレス取得ステップにおいて取得された中継装置に対し、その中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得ステップと、その迷惑メール中継確率取得ステップにおいて得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得ステップと、その迷惑メール受信確率取得ステップにおいて得られた確率に応じて前記電子メールを所定区分に分類するメール判定ステップとを備えている。
【００２２】
請求項１記載の電子メールフィルタリングプログラムによれば、まず、中継装置アドレス取得ステップにより、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスが取得される。次に、迷惑メール中継確率取得ステップにより、アドレスの取得された中継装置に対し、情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率が得られる。ここで、この迷惑メール中継確率取得ステップにより得られる確率は、ベイズ確率モデルを用いて得られる確率である。
【００２３】
次に、迷惑メール受信確率取得ステップにより、迷惑メール中継確率取得ステップによって得られた確率に基づいて、送信元から送信された電子メールが迷惑メールである確率が得られる。そして、迷惑メール受信確率取得ステップによって得られた送信元から送信された電子メールが迷惑メールである確率に応じて、その電子メールが、メール判定ステップによって判定されて所定区分に分類される。
【００２４】
即ち、請求項１記載の電子メールフィルタリングプログラムによれば、電子メールの配信経路上の中継装置をアドレスによって特定した上で、メール情報に基づき、送信元から送信された電子メールを、例えば、迷惑メールであるとして分類することができる。
【００２５】
なお、特許請求の範囲における用語「迷惑メール」とは、受信者の意図を無視して事前の要請や同意なしに、無差別かつ大量発信されるメール（所謂「スパムメール」）を意味する。また、特許請求の範囲における用語「正当なメール」とは、「迷惑メール」の対義語であり、迷惑メールでないメール、即ち、送信者と受信者との間で互いにそのメールの受け渡しを行う必然性のあるメール（所謂「ハムメール」）を意味する。また、特許請求の範囲における用語「メール情報」とは、中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報を意味する。また、特許請求の範囲における用語「中継装置」とは、配信経路上においてメールが経由する装置及びメールが経由する可能性のある（経由すると推定される）装置を意味する。よって、特許請求の範囲における「中継装置」には、メールを中継する中継サーバだけでなく、メールの送信元となる端末や、配信経路追跡を行った場合にメールを中継したと推定される装置も含まれる。
【００２６】
請求項２記載の電子メールフィルタリングプログラムは、請求項１記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記迷惑メール受信確率取得ステップにより得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合

に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定ステップと、前記迷惑メール受信確率取得ステップにより得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定ステップとを含んでおり、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて判定された結果に応じて、前記中継装置アドレス取得ステップにおいて取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新ステップを備えている。

【００２７】

請求項２記載の電子メールフィルタリングプログラムによれば、請求項１記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、迷惑メール判定ステップと正当メール判定ステップとを含んでいる。即ち、迷惑メール受信確率取得ステップによって得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合には、迷惑メール判定ステップにより、送信元から送信された電子メールが迷惑メールであると判定される。一方で、迷惑メール受信確率取得ステップによって得られた確率が、第２閾値未満又は第２閾値以下の場合には、正当メール判定ステップにより、送信元から送信された電子メールが正当なメールであると判定される。

【００２８】

そして次に、情報更新ステップによって、上記の迷惑メール判定ステップ又は正当メール判定ステップの実行によって判定された結果に応じて、情報記憶手段に記憶されているメール情報が更新される。即ち、中継装置アドレス取得ステップによってアドレスが取得された中継装置に対応する情報記憶手段のメール情報が、上記の迷惑メール判定ステップ又は正当メール判定ステップによって判定された結果に応じて更新される。

【００２９】

請求項３記載の電子メールフィルタリングプログラムは、請求項２記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識ステップをさらに含む。

【００３０】

請求項３記載の電子メールフィルタリングプログラムによれば、請求項２記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、不確定メール認識ステップを含んでいる。即ち、第１閾値と第２閾値とが異なる場合に、迷惑メール判定ステップ又は正当メール判定ステップによって電子メールが迷惑メールでも正当なメールでもないと判定された場合には、不確定メール認識ステップにより、迷惑メールでも正当なメールでもないと判定された電子メールが不確定メールとして認識される。

【００３１】

請求項４記載の電子メールフィルタリングプログラムは、請求項３記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記不確定メール認識ステップにおいて不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定ステップをさらに含む。

【００３２】

請求項４記載の電子メールフィルタリングプログラムによれば、請求項３記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、不確定メール再判定ステップを含んでいる。即ち、不確定メールに認識ステップによって電子メールが不確定メールと認識された場合には、不確定メール再判定ステップにより、その不確定メールに含まれるテキスト情報に基づいて、その不確定メールが迷惑メールであるか又は正当なメールであるかが判定される。

【００３３】

なお、特許請求の範囲における「不確定メールに含まれるテキスト情報」との記載は、

本文テキストやメールヘッダにおけるメールアドレスなど、不確定メールの中でテキストによって記述された情報を全て包含している。

【００３４】

請求項５記載の電子メールフィルタリングプログラムは、請求項２から４のいずれかに記載の電子メールフィルタリングプログラムにおいて、前記中継装置アドレス取得ステップにより取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認ステップを備えていると共に、そのアドレス確認ステップにより前記アドレスが不正なアドレスであると確認された場合には、前記メール判定ステップにおいて、前記送信元から送信された電子メールが迷惑メールであると判定する。

【００３５】

請求項５記載の電子メールフィルタリングプログラムによれば、請求項２から４のいずれかに記載の電子メールフィルタリングプログラムと同様に作用する上、アドレス確認ステップにより、中継装置アドレス取得ステップによって取得された中継装置のアドレスが正当なアドレスであるかが確認される。その結果、アドレスが不正なアドレスであると確認された場合には、メール判定ステップにより、送信元から送信された電子メールが迷惑メールであると判定される。

【００３６】

請求項６記載の電子メールフィルタリングプログラムは、請求項１から５のいずれかに記載の電子メールフィルタリングプログラムにおいて、前記中継装置アドレス取得ステップにより取得された中継装置に対応するメール情報が前記情報記憶手段に不在の新出の中継装置である場合には、前記迷惑メール中継確率取得ステップにおいて、その新出の中継装置に対し、前記迷惑メールである確率として所定値を付与する。

【００３７】

請求項６記載の電子メールフィルタリングプログラムによれば、請求項１から５のいずれかに記載の電子メールフィルタリングプログラムと同様に作用する上、中継装置アドレス取得ステップによって取得された中継装置に対応するメール情報が情報記憶手段に記憶されていない新出の中継装置である場合には、迷惑メール中継確率取得ステップにより、その新出の中継装置に対し、迷惑メールである確率として所定値が付与される。

【００３８】

請求項７記載の電子メールフィルタリングプログラムは、請求項１から６のいずれかに記載の電子メールフィルタリングプログラムにおいて、前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得ステップと、その中継ルータアドレス取得ステップにより得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得ステップとを備えている。

【００３９】

請求項７記載の電子メールフィルタリングプログラムによれば、請求項１から６のいずれかに記載の電子メールフィルタリングプログラムと同様に作用する上、中継ルータアドレス取得ステップにより、配信経路上における中継ルータのアドレスが得られる。次に、補完アドレス取得ステップにより、中継ルータアドレス取得ステップによって得られた中継ルータのアドレスに基づいて、中継装置アドレス取得ステップによってアドレスが取得された中継装置以外の中継装置であって、配信経路を補完する中継装置のアドレスが得られる。

【００４０】

即ち、請求項７記載の電子メールフィルタリングプログラムによれば、中継ルータのアドレスを取得し、その中継ルータのアドレスに基づいて、ヘッダ情報を参照することによってアドレスが取得される中継装置以外の中継装置を、電子メールの配信経路上に補完することができる。

【００４１】

請求項８記載の電子メールフィルタリング方法は、送信元から送信された不特定多数の

電子メールに含まれる迷惑メールをフィルタリングすることができる方法であって、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得手段と、その中継装置アドレス取得手段により取得された中継装置に対し、その中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、その迷惑メール中継確率取得手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えている。

【００４２】

請求項９記載の電子メールフィルタリング方法は、請求項８記載の電子メールフィルタリング方法において、前記メール判定手段は、前記迷惑メール受信確率取得手段により得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、前記迷惑メール受信確率取得手段により得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、前記迷惑メール判定手段又は前記正当メール判定手段により判定された結果に応じて、前記中継装置アドレス取得手段により取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えている。

【００４３】

請求項１０記載の電子メールフィルタリング方法は、請求項９記載の電子メールフィルタリング方法において、前記メール判定手段は、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段により前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えている。

【００４４】

請求項１１記載の電子メールフィルタリング方法は、請求項１０記載の電子メールフィルタリング方法において、前記メール判定手段は、前記不確定メール認識手段により不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えている。

【００４５】

請求項１２記載の電子メールフィルタリング方法は、請求項９から１１のいずれかに記載の電子メールフィルタリング方法において、前記中継装置アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定する。

【００４６】

請求項１３記載の電子メールフィルタリング方法は、請求項８から１２のいずれかに記載の電子メールフィルタリング方法において、前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段により取得された中継装置に対応するメール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与する。

【００４７】

請求項１４記載の電子メールフィルタリング方法は、請求項８から１３のいずれかに記載の電子メールフィルタリング方法において、前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、その中継ルータアドレス取得手段により

得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段によりアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段とを備えている。

【００４８】

請求項１５記載の電子メールフィルタリングシステムは、電子メールを伝送可能な経路上において、送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることが可能なシステムであって、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも１の中継装置のアドレスを取得する中継装置アドレス取得手段と、１の中継装置に対し、その中継装置によって過去に中継された迷惑メール及び正当なメールの各頻度を示すメール情報を記憶する情報記憶手段と、前記中継装置アドレス取得手段により取得された中継装置に対し、情報記憶手段に記憶されているメール情報に基づいて、その中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、その迷惑メール中継確率演算手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えている。

【００４９】

請求項１６記載の電子メールフィルタリングシステムは、請求項１５記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記迷惑メール受信確率取得手段により得られた確率が、第１閾値を越えたか又は第１閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、前記迷惑メール受信確率取得手段により得られた確率が、第２閾値未満又は第２閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、前記迷惑メール判定手段又は前記正当メール判定手段による判定結果に応じて、前記中継装置アドレス取得手段により取得された中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えている。

【００５０】

請求項１７記載の電子メールフィルタリングシステムは、請求項１６記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記第１閾値と前記第２閾値とが異なる場合に、前記迷惑メール判定手段又は前記通常メール判定手段による判定が前記迷惑メール又は前記正当なメールのいずれにも非該当である電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えている。

【００５１】

請求項１８記載の電子メールフィルタリングシステムは、請求項１７記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記不確定メール認識手段において不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えている。

【００５２】

請求項１９記載の電子メールフィルタリングシステムは、請求項１６から１８のいずれかに記載の電子メールフィルタリングシステムにおいて、前記中継装置アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定する。

【００５３】

請求項２０記載の電子メールフィルタリングシステムは、請求項１５から１９のいずれかに記載の電子メールフィルタリングシステムにおいて、前記迷惑メール中継確率取得手

段は、前記中継装置アドレス取得手段により取得された中継装置に対応するメール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与する。

【００５４】

請求項２１記載の電子メールフィルタリングシステムは、請求項１５から２０のいずれかに記載の電子メールフィルタリングシステムにおいて、前記配信経路上における中継ルータのアドレスを得る中継ルータアドレス取得手段と、その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段においてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを得る補完アドレス取得手段とを備えている。

【発明の効果】

【００５５】

本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、電子メールのヘッダ情報における記述に基づいて配信経路上の中継装置をアドレスによって特定した上で、メール情報に基づき、送信元から送信された電子メールを、例えば、１区分として迷惑メールであると区分して分類することができる。

【００５６】

ここで、配信経路上における中継装置が過去に迷惑メールを中継した頻度の多さは、その中継装置が迷惑メールの配信経路上にある可能性の高さに対応する。即ち、ある中継装置に対応するメール情報が、過去における迷惑メールの中継頻度が高いことを示すものであれば、その中継装置は、悪質な送信者から送信された迷惑メールの配信経路上の中継装置である可能性が高い。

【００５７】

よって、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、配信経路の中継装置をアドレスによって特定した上で、その特定された中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用して、受信したメールを所定区分、例えば、迷惑メールとして分類することができる。また、その一方で、正当なメールが過剰に拒否されることを抑制できるという効果がある。このように、迷惑メールが確実に検出されると、その結果として、迷惑メールと判定されたメールを削除する、又は、受信を拒否するなどの処理を施すことが可能となり、ユーザが迷惑メールにより受ける実害を低減できる。

【００５８】

また、迷惑メールにおいて、配信経路における中継装置のアドレス（ＩＰアドレス、メールアドレス）が偽装される場合、同じアドレスが複数回使われることが少なく、結果として、偽装されたアドレスは１回のみの出現となる。そのため、単語の出現回数から判定する従来のベイズフィルタでは、そのような１回きりの偽装を迷惑メールとして判定することが困難である。

【００５９】

一方で、本発明では、中継装置が過去に中継したメール情報、即ち、迷惑メールの頻度と正当なメールの頻度とを組合わせてベイズ理論を適用することによって、迷惑メールであるか否かの判定力を向上させることができるという効果がある。

【００６０】

また、中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用するので、従来のテキストフィルタのような膨大なデータの蓄積を必要とせず、データベースによる記憶装置（メモリやディスクなど）の消費量を抑制することができるという効果がある。また、送信者による偽装が容易である電子メールに含まれるテキストを利用することなくメールを区分できるので、例えば、所定の閾値を境界としてそれ以上又はそれを越えた場合にそのメールを迷惑メールであると区分すること

により、迷惑メールを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。

【００６１】

さらに、送信者による偽装が容易である送信者のアドレス（メールアドレス）を利用することなくメールを区分できるので、例えば、所定の閾値を境界としてそれ以上又はそれを越えた場合にそのメールを迷惑メールであると区分することにより、迷惑メールを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。さらに、従来のアドレスフィルタのように、悪質な送信者との鼬ごっこのような偽装アドレスの登録及び削除を繰り返す必要がなくなり、管理の負担が軽減されるという効果もある。

【００６２】

加えて、配信経路上の中継装置のＩＰアドレスを利用してメールの判定を行うので、その判定結果は言語情報に依存しない。よって、構文解析が難しく従来のテキストフィルタでは誤判定されやすかった日本語のメールであっても、迷惑メールであるか正当なメールであるかを確実に検出することができるという効果がある。

【００６３】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、迷惑メール又は正当なメールであると判定されると、その結果に基づいて、該メールの配信経路上の中継装置に対応する情報記憶手段に記憶されているメール情報が更新される。即ち、判定結果を学習するので、学習するほどその判定精度が向上するという効果がある。

【００６４】

さらに、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、迷惑メールとも正当なメールとも判定がつかず不確定要素の高いメールが不確定メール（グレイメール）として認識される。このように曖昧な分類区分を設けることにより、正当なメールが迷惑メールとして、又は、迷惑メールが正当なメールとして誤判定されることを防止できるという効果がある。その結果として、そのような誤判定に基づいてユーザが被り得る問題を回避することができるという効果がある。

【００６５】

加えて、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、不確定メールであると認識されたメールに対し、そのメールに含まれるテキスト情報に基づいて迷惑メールであるか正当なメールであるかの判定が行われるので、配信経路上を通って配信される電子メールが迷惑メールである場合にそれを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。

【００６６】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、電子メールのヘッダ情報に基づいて特定された中継装置が正当に登録されているアドレスであるかを確認し、その際、不正なアドレスであれば、そのメールが迷惑メールであると判定される。よって、ヘッダ情報が明らかに不正であることが確認されれば迷惑メールとして処理されるので、処理を効率化できると共に、迷惑メールであるか又は正当なメールであるかの判定精度を向上させることができるという効果がある。

【００６７】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、情報記憶手段にメール情報の記憶されていない新出の中継装置に対しては、そのメールが迷惑メールである確率として所定値を用いるので、幅広い配信経路に対して適用可能であると共に、新たな配信経路が生じやすい初見メールに対しても対応可能であるという効果がある。

【００６８】
さらに、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、中継ルータのアドレスを取得し、その中継ルータのアドレスに基づいて、ヘッダ情報を参照することによってアドレスが取得されていた中継装置以外の中継装置が、電子メールの配信経路上に補完される。よって、迷惑メールであるかの判定のために利用する中継装置の数が増加するので、同一の経路を経由する重なりを増やすことができる。その結果として、迷惑メールが配信される傾向にある配信経路と正当なメールが配信される傾向にある配信経路との区別をより明確にすることができるので、迷惑メールであるか又は正当なメールであるかの判定精度を向上させることができるという効果がある。

【発明を実施するための最良の形態】
【００６９】
以下、本発明の好ましい実施例について、添付図面を参照しつつ説明する。まず、本発明の電子メールフィルタリングシステムの第１実施例について説明する。図１は、本発明の第１実施例における電子メールフィルタリングシステムが実装される電子メールの配信経路の一例を示す模式図である。図１に示すように、６つの端末Ｔｓ（Ts_1～Ts_6）を送信側の端末とした場合には、それらの端末から受信側の端末Ｔ（Ｔｒ）へ向けて送信されたメールは、それらの間に介在された５個のメールサーバＳ（S_1～S_5）により中継されて、端末Ｔｒへ配信される。この場合、メールの配信経路は、図１に示す通り、送信側端末Ｔｓの数、即ち、６通り存在することになる。

【００７０】
具体的には後で詳述するが、本実施例の電子メールフィルタリングシステムでは、配信経路を経由したメールの履歴に基づいて、配信経路における「スパムメールが中継される傾向」であるか「ハムメールが中継される傾向」であるかを推測し、そのような推測に基づいて、送信側端末Ｔｓ（図１では端末Ts_1～Ts_6）から送信されたメールがスパムメールであるか否かを判定する。

【００７１】
つまり、図１に示す例でいえば、６つの配信経路上においてメールが受信側端末Ｔｒに到達するまでに経由する装置（以下、「中継装置」と称する）、即ち、送信側端末Ｔｓ（Ts_1～Ts_6）及びメールサーバＳ（S_1～S_5）のそれぞれについて、過去にスパムメールが通過した頻度の情報及びハムメールが通過した頻度の情報を記憶しておく。次いで、その情報に基づいて、ある配信経路を通るメールがスパムメールである確率を、ベイズ確率モデルを用いて算出する。そして、そのように算出された確率に基づいて、その配信経路を通るメールがスパムメールであるか否かを判定する。

【００７２】
例えば、送信側端末Ts_1が受信側端末Ｔｒに対しスパムメールのみを送信していた場合には、メールサーバS_1，S_4，S_5を含む経路Ａを共有する送信側端末Ts_2が受信側Ｔｒへ最初のメール（以下、ある端末へ初めて送信したメールを「初見メール」と称する）を送信した場合に、そのメールを「スパムメールらしい」と推測することができる。同様に、例えば、送信側端末Ts_5が受信側端末Ｔｒに対しハムメールのみを送信していた場合には、メールサーバS_3，S_4，S_5を含む経路Ｃを共有する送信側端末Ts_6から受信側端末Ｔｒへ向けて送信されたメールは、初見メールであっても「ハムメールらしい」と推測することができる。

【００７３】
送信側端末Ts_1～Ts_6から受信側端末Ｔｒへメールが配信される際に経由するメールサーバＳは、そのメールにおけるヘッダ情報を参照することにより特定することができる。ヘッダ情報の１つである「Received:」フィールドは、メールが、送信側の端末Ｔｓ（図１では端末Ts_1～Ts_6）から送信されてから、受信側の端末Ｔｒに到達するまでに経由する（中継される）メールサーバＳ（図１ではメールサーバS_1～S_5）を示すフィールド（情報）である。

【００７４】
ここで、図２を参照して、「Received:」フィールドに記録されている情報について説明する。図２は、メールに付加された「Received:」フィールドを示す模式図である。ここで、メールサーバＳは、上記のように「Received:」フィールドへ情報を追加することは許可されているが、既に存在する「Received:」フィールドの情報を消去や変更することは禁止されている。
【００７５】
よって、送信側端末Ｔｓから受信側端末Ｔｒへ到達するまでに経由する（中継される）メールサーバＳの数が多いほど、「Received:」フィールドの数は増えることになる。図２に示す例では、情報Ｒ１及び情報Ｒ２の２つの「Received:」フィールドが存在している。即ち、図２に示す「Received:」フィールドが記録されたメールは、２つのメールサーバＳによって中継されたことを示す。
【００７６】
「Received:」フィールドは、図２に示すように、このフィールドの開始文字列である「Received:」が必須で記録され、この文字列「Received:」以下に、文字列「from」で始まる送信ホストの情報や、文字列「by」で始まる受信ホストの情報などが任意に記憶される。ここで、文字列「from」以下に記録される送信ホストの情報は、多くの場合、「送信ホスト名　（ＦＱＤＮ名　［ＩＰアドレス］）」の書式で記載されている。なお、「ＦＱＤＮ」とは、「Fully Qualified Domain Name」の略である。一方、文字列「by」以下に記録される受信ホストの情報は、多くの場合、「ＦＱＤＮ名　（付加情報）」の書式で記載されている。
【００７７】
また、「Received:」フィールドが追加される場合には、下から上に向かって追加されるので、下に行くほど送信側端末Ｔｓに近い情報となる。よって、この「Received:」フィールドを遡ることによって、メールの配信経路を受信者側端末Ｔｒから送信者側端末Ｔｓまで遡ることができる。
【００７８】
よって、図２に示す「Received:」フィールドの記録は、送信者側端末から送信されたメールが、『（１）送信者側端末→（２）ＦＱＤＮ名が「ceres.xxx.ne.jp」であり、ＩＰアドレスが「211.6.xxx.78」であるメールサーバ→（３）ＦＱＤＮ名が「mx1.xxx.or.jp」であるメールサーバ→（４）受信側端末』の配信経路を通って、受信側端末に到達していることを示している。
【００７９】
ここで、上記したように、「Received:」フィールドはメールを受信したメールサーバＳによって追加される。よって、送信者により送信側端末Ｔｓのメールアドレスやヘッダ情報が偽装されていたとしても、その偽装されたメールを受信したメールサーバＳは、偽装されたメールアドレスやヘッダ情報とは無関係に、そのメールをどのホスト（ＩＰアドレス）からいつ受信したかを「Received:」フィールドに記録する。このように、「Received:」フィールドに記録される情報は、そのメールサーバＳが正規に管理されているものである場合には、送信者による偽装が困難な情報である。
【００８０】
本実施例の電子メールフィルタリングシステムは、受信側端末Ｔｒに電子メールフィルタリングプログラムを実装することにより、受信側端末Ｔｒにおいて、受信したメールがスパムメールであるか否かを判定するものである。
【００８１】
ここで、図３を参照して、本実施例の電子メールフィルタリングシステムを機能させる受信側端末Ｔｒの構成について説明する。図３は、受信側端末Ｔｒの構成を示すブロック図である。図３に示すように、受信側端末Ｔｒは、受信側端末Ｔｒ全体の動作を制御するＣＰＵ１０と、そのＣＰＵ１０により実行される制御プログラム１２ａや固定値データを記憶するＲＯＭ１２と、ＣＰＵ１０により実行される各種処理に必要なデータやプログラ

ム等を一時的に記憶するためのメモリであるＲＡＭ１４と、記憶部１６と、公衆通信網などの通信回線を介してメールサーバＳと接続するためのインターフェイス１８（Ｉ／Ｆ１８）とを主に備えており、これらの構成がバスライン２０によって互いに接続されている。

【００８２】

ここで、ＲＯＭ１２に格納されている制御プログラム１２ａには、本実施例の電子メールフィルタリングシステムを機能させる電子メールフィルタリングプログラムが含まれている。この制御プログラム１２ａに含まれる電子メールフィルタリングプログラムによって実行される処理については後述する。

【００８３】

ＲＡＭ１４は、受信メールメモリ１４ａと、経路情報メモリ１４ｂと、配信木情報メモリ１４ｃとを備えている。なお、これらのメモリ１４ａ～１４ｃは、いずれも受信側端末Ｔｒの電源投入時に初期化される。

【００８４】

受信メールメモリ１４ａは、受信したメールを一時的に記憶するメモリであり、この受信メールメモリ１４ａに記憶されたメールに対し、後述する電子メールフィルタリングプログラムに従う処理が実行される。

【００８５】

経路情報メモリ１４ｂは、送信側端末Ｔｓから送信されたメールが受信側装置Ｔｒに受信されるまでの配信経路上の中継装置（送信側端末Ｔｓ及びメールサーバＳ）のＩＰアドレスを記憶するメモリである。なお、この経路情報メモリ１４ｂに記憶されるメールサーバＳのＩＰアドレスは、受信したメール（受信メールメモリ１４ａに記憶されたメール）の「Received:」に記録されているＩＰアドレスだけでなく、後述する中継ルータの探索に基づいて取得されたメールサーバのＩＰアドレスも含まれる。

【００８６】

配信木情報メモリ１４ｃは、経路情報メモリ１４ｂに記憶されたＩＰアドレスに対応する中継装置（送信側端末Ｔｓ及びメールサーバＳ）について、その中継装置が過去に中継したスパムメール及びハムメールの数を後述する中継装置メモリ１６ａから読み出して一時的に記憶するためのメモリである。

【００８７】

記憶部１６は、書き換え可能な大容量の記憶装置であり、電源断後もデータを保持する不揮発性のメモリであるハードディスクなどの書き換え可能な不揮発性メモリである。この記憶部１６は、中継装置メモリ１６ａと、スパム受信カウンタ１６ｂと、ハム受信カウンタ１６ｃとを備えている。

【００８８】

中継装置メモリ１６ａは、中継装置、即ち、送信者側端末Ｔｓ（図１ではＴｓ$_1$～Ｔｓ$_6$）及びメールサーバＳ（図１ではＳ$_1$～Ｓ$_5$）が過去に中継したスパムメール及びハムメールの数を記憶するメモリである。この中継装置メモリ１６ａは、第１中継装置メモリ１６ａ$_1$から第ｎ中継装置メモリ１６ａ$_n$までのｎ個のメモリから構成されており、その数（ｎ個）は、過去に受信側端末Ｔｒが受信したメールの配信経路上の中継装置（送信側端末Ｔｓ及びメールサーバＳ）として検出された全ての中継装置の数に対応する。

【００８９】

第１～第ｎ中継装置メモリ１６ａ$_1$～１６ａ$_n$は、それぞれ、スパム中継カウンタ１６ａ$_{11}$～１６ａ$_{n1}$とハム中継カウンタ１６ａ$_{12}$～１６ａ$_{n2}$とを備えている。ここで、スパム中継カウンタ１６ａ$_{11}$～１６ａ$_{n1}$は、それぞれ、対応する中継装置がスパムメールを中継（経由）した数を計数するカウンタであり、一方で、ハム中継カウンタ１６ａ$_{12}$～１６ａ$_{n2}$は、対応する中継装置がハムメールを中継した数を計数するカウンタである。

【００９０】

なお、これらのスパム中継カウンタ１６ａ$_{11}$～１６ａ$_{n1}$及びハム中継カウンタ１６

a_{12}～１６a_{n2}の値は、後述する学習処理（図９参照）において、新規の中継装置メモリ１６ａが作成された場合に、そのメールに対するメール判定処理（図７参照）による判定結果に基づいて、初期値として「０」又は「１」が設定される。

【００９１】

スパム受信カウンタ１６ｂは、受信側端末Ｔｒが受信したスパムメールの数を計数するカウンタであり、ハム受信カウンタ１６ｃは、受信側端末Ｔｒが受信したハムメールの数を計数するカウンタである。

【００９２】

なお、１の中継装置メモリ１６a_x（ｘ＝１～ｎ）に対応するスパム中継カウンタ１６a_{x1}及びハム中継カウンタ１６a_{x2}（ｘ＝１～ｎ）の値と、スパム受信カウンタ１６ｂ及びハム受信カウンタ１６ｃの値とから得られる値が、その中継装置メモリ１６a_xに対応する中継装置（メールサーバＳ及び送信側端末Ｔｓ）の「メール情報」である。

【００９３】

次に、上記のように構成された受信側端末Ｔｒに実装された電子メールフィルタリングプログラムによって実行される各処理について説明する。図４は、電子メールフィルタリングプログラムによって実行されるメール受信処理を示すフローチャートである。図４に示すメール受信処理は、受信側端末Ｔｒにおいて、ユーザがメール受信の指示を行った場合に起動する処理である。

【００９４】

図４に示すように、メール受信処理は、まず、その受信側端末Ｔｒに接続されるメールサーバＳ（図１に示すサーバS_5）にメールがあるかを確認し（Ｓ１）、メールがあれば（Ｓ１：Ｙｅｓ）、そのメールサーバS_5から１のメールを受信し、受信メールメモリ１４ａにその受信メールのデータを記憶する（Ｓ２）。

【００９５】

Ｓ２の処理後、その受信メールにおけるメールヘッダ（「Received:」フィールド）を参照することによって、メールの配信経路の経路情報を得る経路情報取得処理を実行し（Ｓ３）、その経路情報取得処理（Ｓ３）の処理の結果として取得された経路情報に基づいて、１の配信経路を表す配信木と呼ばれる木構造を構築する配信木構築処理を実行する（Ｓ４）。

【００９６】

配信木構築処理（Ｓ４）の実行後、構築された配信木によって表される１の配信経路を通るメールがスパムメールであるか否かを判定するメール判定処理（Ｓ５）を実行する。そして、メール判定処理（Ｓ５）の実行後、その判定結果に基づいて、記憶部１６に記憶される「メール情報」の更新を行う学習処理（Ｓ６）を実行する。なお、これらの経路情報取得処理（Ｓ３）、配信木構築処理（Ｓ４）、メール判定処理（Ｓ５）、学習処理（Ｓ６）における具体的な処理については、それぞれ、図５～図７，図９のフローチャートを参照しつつ後述する。

【００９７】

そして、学習処理（Ｓ６）の実行後、メールサーバS_5に受信すべきメールがなくなるまで、Ｓ１～Ｓ６の処理を繰り返す。そして、Ｓ１の処理によって確認した結果、メールサーバS_5に受信すべきメールがなくなると（Ｓ１：Ｎｏ）、このメール受信処理を終了する。本実施例の電子メールフィルタリングプログラムに従う上記のメール受信処理の実行によって、不特定多数の送信側端末Ｔｓから送信された電子メールがスパムメールであるか否かを判定することができる。

【００９８】

次に、図５のフローチャートを参照して、上記した経路情報取得処理（Ｓ３）について説明する。図５は、経路情報取得処理（Ｓ３）を示すフローチャートである。図５に示すように、経路情報取得処理（Ｓ３）は、まず、受信メールメモリ１４ａに記憶されているメールの「Received:」フィールドを参照し、配信経路上の中継装置（メールサーバＳ及び送信側端末Ｔｓ）のＩＰアドレスを取得し（Ｓ３０１）、取得したＩＰアドレスを経路

情報メモリ１４ｂに記憶し（Ｓ３０２）、この経路情報取得処理（Ｓ３）を終了する。上記した経路情報取得処理（Ｓ３）によって、配信経路上の送信側端末Ｔｓ及びメールサーバＳを特定することができる。

【００９９】

次に、図６のフローチャートを参照して、上記した配信木構築処理（Ｓ４）について説明する。図６は、配信木構築処理（Ｓ４）を示すフローチャートである。なお、「配信木」とは、配信経路を木構造で表したものである。この配信木は、「根」である受信側端末Ｔｒと、「ノード（所謂「節」とみなされる）」であるメールサーバＳと、「葉」である送信側端末Ｔｓとから構成される。

【０１００】

図６に示すように、配信木構築処理（Ｓ４）は、まず、経路情報メモリ１４ｂに記憶されているＩＰアドレスの中から最も受信側端末に近い中継装置であるメールサーバＳ（図１では$Ｓ_5$）のＩＰアドレスを読み込み（Ｓ４０１）、読み込んだＩＰアドレスが中継装置メモリ１６ａ（第１中継装置メモリ$１６ａ_1$～第ｎ中継装置メモリ$１６ａ_n$）に既存する中継装置のＩＰアドレスであるかを確認する（Ｓ４０２）。

【０１０１】

Ｓ４０２の処理により確認した結果、読み込んだＩＰアドレスが中継装置メモリ１６ａに既存する中継装置のものであれば（Ｓ４０２：Ｙｅｓ）、対応するスパム中継カウンタ$１６_{x\ 1}$及びハム中継カウンタ$１６_{x\ 2}$（ｘは１～ｎのうち対応する値）に記憶されている値を、その中継装置の配信木情報として配信木情報メモリ１４ｃに記憶する（Ｓ４０３）。

【０１０２】

一方で、Ｓ４０２の処理により確認した結果、既存する中継装置のＩＰアドレスでなければ（Ｓ４０２：Ｎｏ）、新規の配信木情報として配信木情報メモリ１４ｃに記憶する（Ｓ４０５）。

【０１０３】

Ｓ４０３又はＳ４０５の処理後、経路情報メモリ１４ｂから読み込んだＩＰアドレスが送信元のＩＰアドレス、即ち、送信側端末$Ｔ_s$のＩＰアドレスであるか否かを確認し（Ｓ４０４）、そうでなければ（Ｓ４０４：Ｎｏ）、経路情報メモリ１４ｂに記憶されているＩＰアドレスの中で、直前に読み込んだＩＰアドレスに対する１段上位（送信側端末Ｔｓの側）のＩＰアドレスを読み込み（Ｓ４０６）、Ｓ４０２の処理へ移行する。そして、Ｓ４０４において、読み込んだＩＰアドレスが送信元のＩＰアドレスであることが確認されるまで、Ｓ４０２～Ｓ４０６を繰り返す。

【０１０４】

Ｓ４０４の処理により確認した結果、読み込んだＩＰアドレスが送信元のＩＰアドレスであれば（Ｓ４０４：Ｙｅｓ）、この配信木構築処理（Ｓ４）を終了する。この配信木構築処理（Ｓ４）により、１の配信経路を表す木構造である配信木が構築される。

【０１０５】

次に、図７のフローチャートを参照して、上記したメール判定処理（Ｓ５）について説明する。図７は、メール判定処理（Ｓ５）を示すフローチャートである。図７に示すように、メール判定処理（Ｓ５）では、まず、配信木情報メモリ１４ｃに記憶されている １の配信木情報を読み出し（Ｓ５０１）、その配信木情報が新規の中継装置に対するものであるかを確認する（Ｓ５０２）。

【０１０６】

Ｓ５０２の処理により確認した結果、配信木情報が新規の中継装置に対するものでなければ（Ｓ５０２：Ｎｏ）、その中継装置に対する配信木情報（その中継装置が過去に中継したスパムメールの数及びハムメールの数）とスパム受信カウンタ１６ｂ及びハム受信カウンタ１６ｃの値とを用いて、即ち、その中継装置に対するメール情報を用いて、下記式（１）に従ってベイズ確率モデルに基づくベイズ確率「$ｐｇ_n$」を求める（Ｓ５０３）。

【０１０７】

【数１】

$$pg_n = \frac{\frac{b}{nbad}}{\frac{g}{ngood} + \frac{b}{nbad}} \qquad (1)$$

式（１）において、「ｂ」は、その中継装置が過去に中継したスパムメールの数、即ち、その中継装置に対応するスパム中継カウンタ１６ａ$_{x1}$（ｘ＝１～ｎ）の値であり、「ｇ」は、その中継装置が過去に中継したハムメールの数、即ち、その中継装置に対応するハム中継カウンタ１６ａ$_{x2}$（ｘ＝１～ｎ）の値である。また、「ｎｂａｄ」は、スパム受信カウンタ１６ｂの値であり、「ｎｇｏｏｄ」は、ハム受信カウンタ１６ｃの値である。なお、このベイズ確率ｐｇ$_n$の上限は「０．９９」であり、下限は「０．０１」であるとする。

【０１０８】

一方で、Ｓ５０２の処理により確認した結果、配信木情報が新規の中継装置に対するものであれば（Ｓ５０２：Ｙｅｓ）、ベイズ確率ｐｇ$_n$の値を「０．５」とする（Ｓ５１２）。このように、配信経路上に新規の中継装置が確認された場合には、ベイズ確率「ｐｇ$_n$」の値を所定の定数として処理するので、新たな配信経路が生じ易い初見メールに対しても対応可能となる。また、幅広い配信経路に対しても適用可能となる。

【０１０９】

Ｓ５１２又はＳ５０３の処理後、配信木情報メモリ１４ｃに記憶されている全ての配信木情報を読み出したかを確認し（Ｓ５０４）、まだ読み出していない配信木情報があれば（Ｓ５０４：Ｎｏ）、Ｓ５０１の処理へ戻り、全ての配信木情報が読み出されるまで、Ｓ５０１～Ｓ５０４，Ｓ５１２の処理を繰り返す。そして、Ｓ５０４において、全ての配信木情報が読み出されたと確認されたら（Ｓ５０４：Ｙｅｓ）、Ｓ５０５の処理へ移行する。

【０１１０】

Ｓ５０５では、Ｓ５０３又はＳ５１２の処理によって、配信経路上の各中継装置に対して得られたベイズ確率ｐｇ$_n$の値から、配信経路全体のベイズ確率ｐｇを下記式（２）から求める（Ｓ５０５）。

【０１１１】

【数２】

$$pg = \frac{pg_1 \times pg_2 \times \cdots \times pg_n}{pg_1 \times pg_2 \times \cdots \times pg_n + (1-pg_1) \times (1-pg_2) \times \cdots \times (1-pg_n)} \qquad (2)$$

Ｓ５０５の処理後、配信経路全体のベイズ確率ｐｇの値が０．９を越えるかを確認し（Ｓ５０６）、ベイズ確率ｐｇの値が０．９を越えていれば（Ｓ５０６：Ｙｅｓ）、受信メールメモリ１４ａに記憶され現在判定中のメールをスパムメールであると判定する（Ｓ５１３）。

【０１１２】

一方で、Ｓ５０６の処理により確認した結果、ベイズ確率ｐｇの値が０．９以下であれば（Ｓ５０６：Ｎｏ）、ベイズ確率ｐｇの値が０．１より小さいかを確認し（Ｓ５０７）、ベイズ確率ｐｇの値が０．１より小さければ（Ｓ５０７：Ｙｅｓ）、受信メールメモリ１４ａに記憶され現在判定中のメールをハムメールであると判定する（Ｓ５１４）。

【０１１３】
また、Ｓ５０７の処理により確認した結果、ベイズ確率ｐｇの値が０．１以上であれば、即ち、ベイズ確率ｐｇが０．１以上かつ０．９以下である場合には（Ｓ５０７：Ｎｏ）、受信メールメモリ１４ａに記憶され現在判定中のメールを、スパムメールともハムメールとも判定がつかず不確定要素の高いメールであるグレイメール（特許請求の範囲における「不確定メール」に該当する）であると判定する（Ｓ５０８）。
【０１１４】
Ｓ５０８の処理後、スパムメールともハムメールとも判定できないとされたグレイメールに対し、別のフィルタをかけてスパムメールであるかハムメールであるかを分類するグレイメール再判定処理を実行する（Ｓ５０９）。このグレイメール再判定処理（Ｓ５０９）では、例えば、グレイメールの本文中のテキスト情報に対して従来のテキストフィルタを適用することによって、グレイメールをスパムメール又はハムメールのいずれかに分類することができる。
【０１１５】
Ｓ５０９、Ｓ５１３又はＳ５１４の処理後、ベイズ確率ｐｇに基づいてスパムメールと判定されたメール、及びＳ５０９におけるグレイメールの再判定によってスパムメールと判定されたメールに対して行う処理であるスパムメール対応処理を実行し（Ｓ５１０）、配信木情報メモリ１４ｃ及びメール受信メモリ１４ａをクリアし（Ｓ５１１）、このメール判定処理（Ｓ５）を終了する。
【０１１６】
なお、上記のスパムメール対応処理（Ｓ５１０）として実行できる処理としては、例えば、スパムメールと判定されたメールを受信メールメモリ１４ａから削除したり、そのスパムメールを受信拒否としたり、そのスパムメールを専用のフォルダに格納したり、スパムメールが受信されたことをユーザに報知したりなどの各種処理が挙げられる。このように、スパムメールと判定されたメールに対して削除や受信拒否するなどのスパムメール対応処理（Ｓ５１０）を実行することによって、ユーザがスパムメールにより受ける実害を低減することができる。なお、スパムメール対応処理（Ｓ５１０）においてスパムメールに対して所定の処理を行う一方で、ハムメールと判定されたメールについては、受信メールメモリ１４ａから記憶部１６に設けられたメール格納部（非図示）に不揮発的に記憶するように構成すればよい。
【０１１７】
また、このメール判定処理（Ｓ５）におけるＳ５０８、Ｓ５１３又はＳ５１４の処理の結果として、受信メールメモリ１４ａに一時的に記憶されているメールが、スパムメール、ハムメール、グレイメールの３種類に分類される。ここで、図８を参照して、グレイメールについて説明する。
【０１１８】
図８は、図１に示した配信経路におけるスパムメール及びハムメールの偏りを説明するための模式図である。なお、図８において、丸括弧内に記載された数値は、左側が「その中継サーバを経由したハムメールの数（ハムメールの経由数）」を示す数値であり、右側が「その中継サーバを経由したスパムメールの数（スパムメールの経由数）」を示す数値である。即ち、左側の数値が、各中継サーバに対するハム中継カウンタ16_{x2}（ｘ＝１～ｎ）の値であり、右側の数値が、各中継サーバに対するスパム中継カウンタ16_{x1}（ｘ＝１～ｎ）の値である。
【０１１９】
図８に示すように、送信側端末Ts_1及びTs_2からスパムメールのみが送られた場合、これらの２つの配信経路における重なり部分であるメールサーバS_1，S_4，S_5を含む経路Ａを通るメールは、スパムメールである可能性が高い。その結果として、メール判定処理（Ｓ５）では、経路Ａを通るメールをスパムメールとして判定する。
【０１２０】
一方で、送信側端末Ts_5及びTs_6からハムメールのみが送られた場合、これらの２

つの配信経路における重なり部分であるメールサーバS_3，S_4，S_5を含む経路Ｃを通るメールは、初見メールであってもハムメールである可能性が高い。その結果として、メール判定処理（Ｓ５）では、経路Ｃを経由するメールをハムメールとして判定する。

【０１２１】

しかし、メールサーバS_2，S_4，S_5を含む経路Ｂように、その経路Ｂを経由するメールが、スパムメール又はハムメールのいずれでもあり得る場合には、その経路Ｂを経由するメールがスパムメールかハムメールかを区別し難い。ここで、そのように区別し難いメールに対し、スパムメールであるかハムメールであるかを厳密に区別を付けるような判定を行った場合には、それが誤判定となる可能性が高い。

【０１２２】

上記したメール判定処理（Ｓ５）では、スパムメールであるかハムメールであるかを区別し難いメール、即ち、ベイズ確率ｐｇが０．１以上かつ０．９以下であったメールはグレイメールとして判定されるので、そのような誤判定を防止することができる。

【０１２３】

さらに、グレイメール再判定処理（Ｓ５０９）において、グレイメールと判定されたメールに対して別のフィルタをかけることによって、スパムメールであるかハムメールであるかの判定精度を向上させることができる。

【０１２４】

次に、図９のフローチャートを参照して、上記した学習処理（Ｓ６）について説明する。図９は、学習処理（Ｓ６）を示すフローチャートである。図９に示すように、学習処理（Ｓ６）は、まず、経路情報メモリ１４ｂに記憶されているＩＰアドレスの中から最も受信側端末に近い中継装置であるメールサーバＳ（図１ではS_5）のＩＰアドレスを読み込み（Ｓ６０１）、読み込んだＩＰアドレスが中継装置メモリ１６ａ（第１中継装置メモリ$16a_1$～第ｎ中継装置メモリ$16a_n$）に既存する中継装置のＩＰアドレスであるかを確認する（Ｓ６０２）。

【０１２５】

Ｓ６０２の処理により確認した結果、読み込んだＩＰアドレスが中継装置メモリ１６ａに既存する中継装置のものであれば（Ｓ６０２：Ｙｅｓ）、対応するスパム中継カウンタ16_{x1}及びハム中継カウンタ16_{x2}（ｘは１～ｎのうち対応する値）に記憶されている値を、Ｓ５１３又はＳ５１４で判定されたメールの判定結果に基づいて更新する（Ｓ６０３）。

【０１２６】

一方で、Ｓ６０２の処理により確認した結果、読み込んだＩＰアドレスが中継装置メモリ１６ａに既存する中継装置のＩＰアドレスでなければ（Ｓ６０２：Ｎｏ）、そのＩＰアドレスに対応する中継装置を新規の中継装置として中継装置メモリ１６ａに記憶する（Ｓ６０６）。即ち、その新たな中継装置に対応する新たな第ｎ中継装置メモリ$16a_n$を中継装置メモリ１６ａに作成し、スパム中継カウンタ$16a_{n1}$及びハム中継カウンタ$16a_{n2}$の値を、Ｓ５１３又はＳ５１４で判定されたメールの判定結果に基づいて「０」又は「１」に設定する。

【０１２７】

Ｓ６０６又はＳ６０３の処理後、スパム受信カウンタ１６ｂ又はハム受信カウンタ１６ｃを、Ｓ５１３又はＳ５１４で判定されたメールの判定結果に基づいて更新する（Ｓ６０４）。次いで、経路情報メモリ１４ｂから読み込んだＩＰアドレスが送信元のＩＰアドレス、即ち、送信側端末T_sのＩＰアドレスであるか否かを確認し（Ｓ６０５）、そうでなければ（Ｓ６０５：Ｎｏ）、経路情報メモリ１４ｂに記憶されているＩＰアドレスの中で、直前に読み込んだＩＰアドレスに対する１段上位（送信側端末Ｔｓの側）のＩＰアドレスを読み込み（Ｓ６０７）、Ｓ６０２の処理へ移行する。そして、Ｓ６０５において、読み込んだＩＰアドレスが送信元のＩＰアドレスであることが確認されるまで、Ｓ６０２～Ｓ６０７を繰り返す。Ｓ６０５の処理により確認した結果、読み込んだＩＰアドレスが送信元のＩＰアドレスであれば（Ｓ６０５：Ｙｅｓ）、経路情報メモリ１４ｂをクリアし（

Ｓ６０８）、この学習処理（Ｓ６）を終了する。

【０１２８】

この学習処理（Ｓ６）によって、メール判定処理（Ｓ５）によるメールに対する判定結果に応じて、記憶部１６（中継装置メモリ１６ａ、スパム受信カウンタ１６ｂ、ハム受信カウンタ１６ｃ）に記憶されるメール情報が更新される。即ち、メール情報がメールの判定結果に応じて学習されていくので、学習すればするほどその判定精度を向上させることができる。

【０１２９】

上記のように機能する第１実施例の電子メールフィルタリングシステムによるメールの判定能力を検証した。なお、第１実施例の電子メールフィルタリングシステムを機能させるための電子メールフィルタリングプログラムは、Ｃ言語を用いてＵＮＩＸ（登録商標）上に作成した。

【０１３０】

この検証において３種類のサンプルセットを用いた。この３種類のサンプルセットとは、表１に示す「サンプルセット１」、表２に示す「サンプルセット２」、表３に示す「サンプルセット３」である。なお、検証に使用されるすべての電子メールは、複数のメールサーバにより実際に受信されたものである。

【０１３１】

【表１】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	198	1	1500	0
英語	2	199	0	1500

【０１３２】

【表２】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	200	200	300	300
英語	0	0	0	0

【０１３３】

【表３】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	200	100	800	400
英語	0	100	0	400

表１～表３に示すように、この検証においてサンプルとするスパムメールは、サンプルセット１では英語で記述されているものとし、サンプルセット２では日本語で記述されているものとし、サンプルセット３では半分が日本語で記述されたものであり、残りの半分が英語で記述されたものとした。

【０１３４】

なお、表１～表３に示す各サンプルセットにおいて、「学習メール」とは、本実施例の電子メールフィルタリングシステムにおいて判定されたメールの判定結果を学習させるメールであり、「テストメール」とは、学習による判定精度向上の効果を比較するために、メールの判定結果を学習させないメールである。

【０１３５】

上記のサンプルセット１～３に対し、第１実施例の電子メールフィルタリングシステムを適用することによってメールの判定を行った。即ち、メールの「Received:」フィールドから取得したＩＰアドレスのみを使用したメール判定を行った。その結果を表４～表６に示す。なお、以下において、「ｈａｍ」及び「ｓｐａｍ」の表記は、それぞれ、ハムメール及びスパムメールを示す。

【０１３６】

【表４】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	0	0
40	77	0	0	0
50	93	0	0	0
60	97	0	0	0
70	104	0	0	0
80	104	0	0	0
90	104	0	0	0
100	112	0	0	0
200	125	2	0	0
300	135	1	0	0
400	141	1	0	0
500	142	0	0	0
1000	165	1	0	0
1500	170	0	1	0

【０１３７】

【表５】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	2	0
40	77	0	2	0
50	93	0	2	0
60	97	0	4	0
70	104	0	4	0
80	104	0	4	0
90	104	1	4	0
100	112	1	4	0
200	125	1	8	0
300	136	1	14	0

【０１３８】

【表６】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	2	0
40	77	0	2	0
50	93	0	2	0
60	97	0	2	0
70	104	0	2	0
80	104	0	2	0
90	104	0	2	0
100	112	0	2	0
200	125	0	2	0
300	136	1	2	0
400	141	2	8	0
500	142	1	8	0
800	163	1	8	0

表４～表６に示すように、サンプルセット１～３のいずれも、学習メールが多ければ多いほどハムメールを正しく判定した。即ち、メールの言語に依存することなく、ハムメールを正しく判定した。また、スパムメールをハムメールとして誤判定されることはなかった。

【０１３９】

本実施例の電子メールフィルタリングシステムでは、学習用のハムメールのＩＰアドレスと重複しない限り、スパムメールがハムメールとして誤判定されない。仮に、「Received:」フィールドに複数のＩＰアドレスを偽装して付加したとしても、新規の中継装置のＩＰアドレスはベイズ確率pg_nが０．５として処理されるので、ハムメールと誤判定されることがないのである。

【０１４０】

次に、図１０及び図１１を参照して、本発明の電子メールフィルタリングシステムの第２実施例について説明する。なお、この第２実施例において、上記した第１実施例と同一の部分には同一の符号を付して、その説明を省略する。

【０１４１】

上記した第１実施例の電子メールフィルタリングシステムでは、メールの「Received:」フィールドから取得したＩＰアドレスのみから特定されたメールサーバＳに基づいてメールの判定を行った。しかし、不特定多数の送信側端末Ｔｓから１の受信側端末Ｔｒまでの各配信経路において、配信経路に重複部分が少ない場合には、初見メールが「スパムメールが中継される傾向にある経路」を通過するか、「ハムメールが中継される傾向にある経路」を通過するかを判定するための情報が特に少なくなる。その場合、初見メールの判定精度に影響が生じることになる。

【０１４２】

そこで、この第２実施例の電子メールフィルタリングシステムでは、第１実施例における経路情報取得処理（Ｓ３）におけるＳ３０２の処理後、中継装置として推定されるメールサーバＳを配信経路上に補完するための経路追跡処理（Ｓ３０３）を実行する。なお、以下の説明では、経路追跡処理（Ｓ３０３）によって推定されて補完されたメールサーバＳを便宜的にメールサーバＳ’として表すことがある。

【０１４３】

ここで、図１０は、第２実施例の電子メールフィルタリングシステムにおける経路情報取得処理（Ｓ３）のフローチャートである。図１０に示すように、経路追跡処理（Ｓ３０３）は、以下に説明するＳ３０３ａ～Ｓ３０３ｄの各ステップから構成される。

【０１４４】

まず、Ｓ３０３ａの処理として、中継ルータの探索を行い、そのＩＰアドレスを取得する。このＳ３０３ａの処理において行う中継ルータの探索は、Traceroute（例えば、W. R

ichard Stevens.「詳解TCP/IP Vol.1 プロトコル」Pearson Education Japan. 2004年６月20日新装版第５刷を参照のこと）などの技術を使用することができる。Traceroute の実行により、目的のホストまでに経由した中継ルータのＩＰアドレスを調べることができる。なお、本実施例では、このTracerouteを実行する上で、「Received:」フィールドに記録されているＩＰアドレス群のうち、最も受信者に近いＩＰアドレスを使用する。

【０１４５】

Ｓ３０３ａの処理後、取得した中継ルータのＩＰアドレスが所属するネームサーバの情報（ネームサーバ情報）を取得する（Ｓ３０３ｂ）。このＳ３０３ｂの処理において行うネームサーバ情報の取得は、例えば、Whoisプロトコルを使用するWhoisサーバへの問い合わせを利用することができる。なお、Whoisプロトコルとは、ドメイン名の登録やＩＰアドレスの割当に関する情報をオンラインで提供する仕組みである。このWhoisサーバへの問い合わせにより取得されるネームサーバ情報は、ネームサーバのドメイン名である。

【０１４６】

Ｓ３０３ｂの処理後、取得したネームサーバ情報によって表されるネームサーバにＤＮＳ（Domain Name System）の問い合わせを行い、メールの配信経路上において補完されるメールサーバＳ’のＩＰアドレスを取得し（Ｓ３０３ｃ）。取得したメールサーバＳ’のＩＰアドレスを経路情報メモリ１４ｂに記憶し（Ｓ３０３ｄ）、第２実施例の経路情報取得処理（Ｓ３）を終了する。

【０１４７】

ＤＮＳサーバには、Ａレコード（名前→ＩＰアドレスの定義）やＰＴＲレコード（ＩＰアドレス→名前の定義）、ＮＳレコード(ネームサーバの定義）、ＳＯＡレコード（ドメインのオーソリティ情報の定義）、ＭＸレコードなど、さまざまなレコード(情報）が登録されている。ここで、ＭＸレコードは、Mail Exchangerの略であり、そのドメインにおけるメールサーバに関する情報が登録されている。

【０１４８】

あるドメインから別のドメインに対してメール送信しようとするとき、送信元のメールサーバは、送信先ドメインのＤＮＳサーバに対してＭＸレコードの情報を問い合わせる。ＭＸレコードには、当該ドメインにおけるメールサーバとなっているコンピュータの名前（ＦＱＤＮ名）とＩＰアドレスなどの情報が含まれており、これを基にしてメールの送信先（通常はＳＭＴＰのサーバ）を知ることができるのである。

【０１４９】

Ｓ３０３ｃの処理では、ＤＮＳに対してＭＸレコードへの参照を要求する。このアクセスは、リゾルバ(resolver)を経由して行われる。例えば、「xxx.ac.jp」のＭＸレコードの情報を問い合わせた結果が以下の通りであれば、補完されるメールサーバＳ’のＩＰアドレスは、「133.15.xxx.1」である。

【０１５０】

```
xxx.xx.jp MX preference = 0, mail exchanger = server1.xxx.xx.jp
xxx.xx.jp MX preference = 100, mail exchanger = server.xxx.xx.jp
xxx.xx.jp nameserver = server.xxx.xx.jp
xxx.xx.jp nameserver = nameserv.gw.xxx-u.xx.jp
server1.xxx.xx.jp      internet address = 133.15.xxx.1
server.xxx.xx.jp      internet address = 133.15.xxx.1
nameserv.gw.xxx-u.xx.jp  internet address = 192.50.xx.9
```

【０１５１】

ここで、図１１を参照して、上記の経路情報取得処理（Ｓ３）における経路追跡処理（Ｓ３０３）による結果を模式的に説明する。図１１（ａ）は、各配信経路に重複部分が存在しない場合を模式的に示す図であり、図１１（ｂ）は、経路追跡処理（Ｓ３０３）によってメールサーバＳ’（中継装置）が補完された状態を模式的に示す図である。

【０１５２】

上記したＳ３０３ａ～Ｓ３０３ｄから構成される経路追跡処理（Ｓ３０３）が実行され

た結果、メールの配信経路は図１１（ｂ）に示す通りとなる。即ち、図１１（ａ）に示した配信経路における、送信側端末Ｔｓに接続されるメールサーバＳと受信側端末に接続されるメールサーバＳとの間にメールサーバＳ’が補完される。配信経路上にメールサーバＳ’を補完することによって、各配信経路上において重複する中継装置を増やすことができる。その結果として、スパムメールが中継され易い経路と、ハムメールが中継され易い経路とを確率的に区別し易くすることができるのである。

【０１５３】

Ｓ３０３の処理後、取得したメールサーバＳ’のＩＰアドレスを経路情報メモリ１４ｂに格納し（Ｓ３０３ｄ）、この経路情報取得処理（Ｓ３）を終了する。なお、この第２実施例では、配信木における「ノード」には、メールサーバＳとして補完されたメールサーバＳ’も含むものとする。

【０１５４】

サンプルセット１（表１参照）に対し、中継ルータの探索（経路追跡）を行いメールサーバＳ’の補完を行った場合と、行わない場合とについて、配信経路上におけるメールサーバＳ（補完されたメールサーバＳ’を含む）の数の変化を表７に示す。

【０１５５】

【表７】

深さ	1	2	3	4	5	6	7	8	9	10	平均
ham(経路追跡前)	89	72	11	7	0	20	1	0	0	0	2.1
spam(経路追跡前)	191	8	1	0	0	0	0	0	0	0	1.1
ham(経路追跡後)	11	27	34	26	42	22	28	9	1	0	4.5
spam(経路追跡後)	0	0	1	90	56	31	16	4	1	1	5.0

なお、表７において、「経路追跡前」とは、メールサーバＳ’の補完を行う前、即ち、メールの「Received:」フィールドから取得したＩＰアドレスのみからメールサーバＳを特定した場合を示す。一方、「経路追跡後」とは、上記した経路追跡処理（Ｓ３０３）によってメールサーバＳ’が補完された場合を示す。

【０１５６】

表７に示すように、ハムメール及びスパムメールのいずれも、経路追跡前の状態では、根（受信側端末Ｔｒ）から葉（送信側端末Ｔｓ）までの深さの大半が１または２であった。この状態では、配信経路が重複することは少ない。

【０１５７】

一方で、経路追跡後、即ち、経路追跡処理（Ｓ３０３）の実行後は、両メールとも、同一の経路を経由する重なりが増え、深さの平均が４を超えた。よって、経路追跡処理（Ｓ３０３）の実行によって、よりハムメール的な経路と、よりスパムメール的な経路と、グレイメール的な経路との区別をより明確にすることができ、その結果、スパムメールであるか又はハムメールであるかの判定精度をより向上させることができるのである。

【０１５８】

次に、第２実施例の電子メールフィルタリングシステム及び従来のホワイトリストフィルタについて、ハムメールの判定精度を比較した。図１２は、第２実施例の電子メールフィルタリングシステムによるハムメールに対する判定結果と従来のホワイトリストフィルタによるハムメールに対する判定結果とを記したグラフである。なお、ホワイトリストフィルタに用いるホワイトリストは、サンプルセット１～３のそれぞれにおける「学習メール」としたメールのメールアドレスを全て登録することによって作成した。

【０１５９】

図１２（ａ）～（ｃ）は、それぞれサンプルセット１～３に対する結果を示すグラフである。ここで、グラフ１２０ａ，１２１ａ，１２２ａは、それぞれ、サンプルセット１～３について、本実施例の電子メールフィルタリングシステムによってハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１２０ｂ，１２１ｂ，１２２ｂは、

それぞれ、サンプルセット１～３について、従来のホワイトリストフィルタによってハムメールを正しく判定した検出率を示すグラフである。また、グラフ１２０ｃ，１２１ｃ，１２２ｃは、それぞれ、サンプルセット１～３について、本実施例の電子メールフィルタリングシステムによりハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１２０ｄ，１２１ｄ，１２２ｄは、それぞれ、サンプルセット１～３について、従来のホワイトリストフィルタによりハムメールをスパムメールとして誤判定した検出率を示すグラフである。

【０１６０】

図１２（ａ）～図１２（ｃ）に示すように、サンプルセット１～３のいずれの場合も、本実施例の電子メールフィルタリングシステムを用いた方が、従来のホワイトリストフィルタを用いた場合に比べ、ハムメールの判定精度が全体に渡って良好であった。ここで、図１２（ａ）に示すように、サンプルセット１に対しては、学習メールが１００通及び１０００通の場合に、ハムメールの判定精度がそれぞれ６２．０％及び１６．５％向上した。また、図１２（ｂ）に示すように、サンプルセット２に対しては、学習メールが１００通及び３００通の場合に、ハムメールの判定精度がそれぞれ６２．０％及び２８．５％向上した。また、図１２（ｃ）に示すように、サンプルセット３に対しては、学習メールが１００通及び５００通の場合に、ハムメールの判定精度がそれぞれ６０．５％及び２５．５％向上した。この結果は、従来のホワイトフィルタでは初見メールに対する判定が困難であるのに対し、本実施例の電子メールフィルタリングシステムでは、配信経路上の中継装置におけるメール情報に基づいてメールを判定するので、初見メールであっても精度よくハムメールを判定できることを示している。また、日本語、英語といった言語情報の差違にかかわらず、精度よくハムメールを判定できることを示している。

【０１６１】

次に、第２実施例の電子メールフィルタリングシステム及び従来のブラックリストフィルタについて、スパムメールの判定精度を比較した。図１３は、第２実施例の電子メールフィルタリングシステムによるスパムメールに対する判定結果と従来のブラックリストフィルタによるスパムメールに対する判定結果とを記したグラフである。なお、ブラックリストへの登録内容は、学習メールにおけるスパムメールのメールアドレスを全て登録した。同時に、ＤＮＳブラックリストに登録されている２０万件のＩＰアドレスを使用した。

【０１６２】

図１３（ａ）～（ｃ）は、それぞれサンプルセット１～３に対する結果を示すグラフである。ここで、グラフ１３０ａ，１３１ａ，１３２ａは、それぞれ、サンプルセット１～３について、本実施例の電子メールフィルタリングシステムによってスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１３０ｂ，１３１ｂ，１３２ｂは、それぞれ、サンプルセット１～３について、従来のブラックリストフィルタによってスパムメールを正しく判定した検出率を示すグラフである。また、グラフ１３０ｃ，１３１ｃ，１３２ｃは、それぞれ、サンプルセット１～３について、本実施例の電子メールフィルタリングシステムによりスパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１３０ｄ，１３１ｄ，１３２ｄは、それぞれ、サンプルセット１～３について、従来のブラックリストフィルタによりスパムメールをハムメールとして誤判定した検出率を示すグラフである。

【０１６３】

図１３（ａ）～図１３（ｃ）に示すように、サンプルセット１～３のいずれの場合も、本実施例の電子メールフィルタリングシステムを用いた方が、従来のブラックリストフィルタを用いた場合に比べ、スパムメールの判定精度が全体に渡って良好であった。ここで、図１３（ａ）に示すように、サンプルセット１に対しては、学習メールが１００通及び１０００通の場合にいずれも、スパムメールの判定精度が８９．０％向上した。また、図１３（ｂ）に示すように、サンプルセット２に対しては、学習メールが１００通及び３００通の場合に、スパムメールの判定精度がそれぞれ１６．０％及び２５．５％向上した。また、図１３（ｃ）に示すように、サンプルセット３に対しては、学習メールが１００通

及び５００通の場合に、スパムメールの判定精度がそれぞれ６０．５％及び２５．５％向上した。

【０１６４】

よって、本実施例の電子メールフィルタリングシステムによれば、サンプルセット１のような送信元の地域を日本と海外とである程度の分類が可能である場合に、ブラックリストフィルタよりスパムメールの判定精度が格段に向上することを示す。また、サンプルセット２のような、発信元が主に日本である故にハムメール及びスパムメールの発信元地域を分類し難い場合であっても、学習が行われることによって判定精度が向上することを示している。また、日本語、英語といった言語情報の差違にかかわらず、精度よくスパムメールを判定できることを示している。

【０１６５】

次に、図１４を参照して、第２実施例の電子メールフィルタリングシステムにおけるグレイメール再判定処理（Ｓ５０９：図７参照）の実行による効果について検証する。図１４は、サンプルセット３について、グレイメール再判定処理（Ｓ５０９）を実行した場合と実行しなかった場合におけるメールの判定精度を比較するグラフである。図１４（ａ）は、ハムメールに対する判定精度を示すグラフであり、図１４（ｂ）は、スパムメールに対する判定精度を示すグラフである。なお、テキスト型ベイジアンフィルタとして、「bsfilter」（http://bsfilter.org/）を用い、Paul Graham方式で実行した。

【０１６６】

図１４（ａ）において、グラフ１４０ａは、グレイメール再判定処理（Ｓ５０９）を実行した場合におけるハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１４０ｂは、グレイメール再判定処理（Ｓ５０９）を実行しなかった場合におけるハムメールを正しく判定した検出率を示すグラフである。また、グラフ１４０ｃは、グレイメール再判定処理（Ｓ５０９）を実行した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１４０ｄは、グレイメール再判定処理（Ｓ５０９）を実行しなかった場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。

【０１６７】

図１４（ｂ）において、グラフ１４１ａは、グレイメール再判定処理（Ｓ５０９）を実行した場合におけるスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１４１ｂは、グレイメール再判定処理（Ｓ５０９）を実行しなかった場合におけるスパムメールを正しく判定した検出率を示すグラフである。また、グラフ１４１ｃは、グレイメール再判定処理（Ｓ５０９）を実行した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１４１ｄは、グレイメール再判定処理（Ｓ５０９）を実行しなかった場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。

【０１６８】

図１４（ａ）に示すように、ハムメールの判定精度は、学習メールが１００通及び５００通の場合にそれぞれ１３．５％及び５．５％向上した。一方で、スパムメールをハムメールとして誤判定した検出率が悪化した（学習メール１００通の場合に２９．０％、学習メール５００通の場合に１５．５％）が、これは、グレイメール再判定処理（Ｓ５０９）を実行した結果として生じたハムメールの誤判定率の悪化（グラフ１４０ｄにおける各値に対するグラフ１４０ｃにおける各値の上昇）は、ハムメールでもスパムメールでもない第３の区分であるグレイメールを考慮した結果として、グレイメールに対する再判定を実行しなかった場合における誤判定率の低下分がグレイメールに振り分けられたことに起因すると考えられる。

【０１６９】

また、図１４（ｂ）に示すように、学習メールが１００通及び５００通の場合に、スパムメールの判定精度がそれぞれ８．５％及び１７．５％向上した。一方で、ハムメールをスパムメールとして誤判定した検出率には変化がなかった。この結果は、グレイメールに

対しテキスト型ベイジアンフィルタを適用して再判定を行うことによって、ハムメール及びスパムメールの判定精度が向上することを示している。

【０１７０】

次に、グレイメール再判定処理（Ｓ５０９：図７参照）としてテキスト型ベイジアンフィルタを併用してメール判定を行った場合と、第２実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタのみを用いてメール判定を行った場合について比較した。

【０１７１】

図１５は、サンプルセット３について、本実施例の電子メールフィルタリングシステムにベイジアンフィルタを併用した場合の判定結果と、テキスト型ベイジアンフィルタのみを使用した場合の判定結果とを比較するグラフである。図１５（ａ）は、ハムメールに対する判定精度を示すグラフであり、図１５（ｂ）は、スパムメールに対する判定精度を示すグラフである。なお、テキスト型ベイジアンフィルタは、図１４に示した結果を得るために用いたものと同じく「bsfilter」を用い、Paul Graham方式で実行した。

【０１７２】

図１５（ａ）において、グラフ１５０ａは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合におけるハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１５０ｂは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合におけるハムメールを正しく判定した検出率を示すグラフである。また、グラフ１５０ｃは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１５０ｄは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。

【０１７３】

図１５（ａ）に示すように、学習メールが１００通及び５００通の場合に、ハムメールの誤判定精度をそれぞれ１４．５％及び５．５％改善した。この結果は、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用することによって、ハムメールの誤判定の程度を改善することを示している。即ち、従来のテキスト型ベイジアンフィルタでは、ハムメールに重みを置かれていることに基づく誤判定のし易さや、グレイメールを設けることなくハムメールかスパムメールかを二者択一的に判定することに基づく誤判定のし易さや、構文解析の難しい日本語に対する判定程度の低さなどが問題であったが、本実施例の電子メールフィルタリングシステムのようにテキスト型ベイジアンフィルタを併用することによってこれらの問題を解決し得ることを示している。

【０１７４】

ここで、この図１５（ａ）におけるグラフ１５０ｃは、図１４（ａ）におけるグラフ１４０ｃと同一のグラフである。よって、図１４（ａ）と図１５（ａ）とから、スパムメールをハムメールとして誤判定した検出率に関し、（１）グレイメールに対する再判定処理（Ｓ５０９：図７参照）を実行しない場合（グラフ１４０ｄに対応）と、（２）グレイメールに対し、従来のテキスト型ベイジアンフィルタを併用した場合（グラフ１４０ｃ及びグラフ１５０ｃに対応）と、（３）従来のテキスト型ベイジアンフィルタを使用した場合（グラフ１５０ｄに対応する）とを比較することができる。

【０１７５】

よって、図１４（ａ）に示すように、グレイメール再判定処理（Ｓ５０９）を実行した場合のハムメールの誤判定率（グラフ１４０ｃ）が、グレイメール再判定処理（Ｓ５０９）を実行しなかった場合（グラフ１４０ｄ）に比べて悪化したとしても、それは従来技術（従来のテキスト型ベイジアンフィルタ）に対する悪化を示すものではない。即ち、図１５（ａ）によれば、グレイメール再判定処理（Ｓ５０９）を実行した結果として、ハムメールの誤判定率（グラフ１５０ｃ（グラフ１４０ｃに対応））は、従来テキスト型ベイジ

アンフィルタにおけるハムメールの誤判定率（グラフ１５０ｄ）に比べて改善されていることが明白である。

【０１７６】

一方、図１５（ｂ）において、グラフ１５１ａは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合におけるスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ１５１ｂは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合におけるスパムメールを正しく判定した検出率を示すグラフである。また、グラフ１５１ｃは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ１５１ｄは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。

【０１７７】

図１５（ｂ）に示すように、学習メールが１００通及び５００通の場合に、スパムメールの判定精度がそれぞれ１４．５％及び５．５％向上した。また、誤判定精度には変化がなかった。この結果は、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用することによって、スパムメールの判定精度が、テキスト型ベイジアンフィルタのみを使用する場合に比べて向上することを示している。

【０１７８】

次に、本発明の電子メールフィルタリングシステムの第３実施例について説明する。なお、この第３実施例において、上記した第１及び第２実施例と同一の部分には同一の符号を付して、その説明を省略する。

【０１７９】

この第３実施例の電子メールフィルタリングシステムは、「Received:」フィールドに記録されている情報が偽装された情報であることが確認された場合には、そのメールをスパムメールであると判定するものであり、「Received:」フィールドに記録されている情報が偽装されたものであるか否かを確認するために、ＩＰアドレス確認処理（Ｓ７）を行う。

【０１８０】

このＩＰアドレス確認処理（Ｓ７）は、まず、メールの「Received:」フィールドに記録されているＩＰアドレスを、受信側に近い方から順にＤＮＳサーバに問い合わせることによってＦＱＤＮ名を逆引きする処理（Ｓ７ａ）と、その逆引きによって得られたＦＱＤＮ名と「Received:」フィールドに記録されている送信ホスト名とが一致するかを確認する処理（Ｓ７ｂ）とから構成される処理である。

【０１８１】

このＩＰアドレス確認処理（Ｓ７）は、その判定結果を学習させるために学習処理（Ｓ６）の前に実行することが好ましい。ここで、ＩＰアドレス確認処理（Ｓ７）が、経路情報取得処理（Ｓ３）より前に実行される場合には、Ｓ７ｂの処理により確認した結果、逆引きによって得られたＦＱＤＮ名と「Received:」フィールドに記録されている送信ホスト名とが不一致である場合には（Ｓ７ｂ：Ｎｏ）、スパムメールであると判定し（Ｓ８）、その一方で、一致する場合には（Ｓ７ｂ：Ｙｅｓ）、未判定メールと認定する（Ｓ９）。そして、未判定メールであった場合には、経路情報取得処理（Ｓ３）、配信木構築処理（Ｓ４）、メール判定処理（Ｓ５）を実行して、そのメールがスパムメールであるか等を判定するように構成すればよい。なお、Ｓ８においてスパムメールと判定された場合には、そのメールがスパムメールであることを前提として経路情報取得処理（Ｓ３）、配信木構築処理（Ｓ４）、メール判定処理（Ｓ５）、学習処理（Ｓ６）を実行するように構成すればよい。

【０１８２】

あるいは、メール判定処理（Ｓ５）によりグレイメールと判定されたメールに対し、Ｉ

Ｐアドレス確認処理（Ｓ７）を実行するように構成してもよい。この場合には、メール判定処理（Ｓ５）におけるＳ５０８の処理以降、例えば、Ｓ５０８とＳ５０９との間などに実行することができる。即ち、Ｓ５０８の処理によりグレイメールと判定されたメールに対し、ＤＮＳ逆引きによって得られたＦＱＤＮ名と「Received:」フィールドに記録されている送信ホスト名とが不一致である場合には（Ｓ７ｂ：Ｎｏ）、スパムメールであると判定し（Ｓ８）、その一方で、一致する場合には（Ｓ７ｂ：Ｙｅｓ）、ハムメールと判定する（Ｓ１０）ように構成すればよい。

【０１８３】

なお、このようにＤＮＳ逆引きによって得られたＦＱＤＮ名と「Received:」フィールドに記録されている送信ホスト名とを比較する手法は、従来のメールサーバにセキュリティ機能として実装されている。しかし、従来のセキュリティ機能ではこの手法はホストの判定のみに使用されており、本実施例の電子メールフィルタリングシステムでは、一つのメールを中継する配信経路上のメールサーバＳを一組とした情報として取り扱う。即ち、この手法をメールの配信経路の全体に対して適用し、メールが中継される経路からスパムメールかどうかを判定する。よって、中継装置が動的ＩＰアドレス割り当てによって変化しても、柔軟にかつより詳細に対応できるのである。また、ヘッダが明らかに偽造されているものをスパムメールとして処理できるので、処理を効率化できると共に、スパムメールであるか又はハムメールであるかの判定精度を向上させることができる。

【０１８４】

以上説明したように、本発明の電子メールフィルタリングシステムによれば、配信経路上の中継装置（送信側端末Ｔｓ及びメールサーバＳ）をＩＰアドレスによって特定した上で、その特定された中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用することにより、配信経路上を通って配信されるメールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。

【０１８５】

また、配信経路上の中継装置（送信側端末Ｔｓ及びメールサーバＳ）が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用するので、従来のテキストフィルタのような膨大なデータの蓄積を必要しない。そのため、データベースによる記憶装置（メモリやディスクなど）の消費量を抑制することができる。さらに、スパムメールであるか否かの判定を行う場合に、送信者による偽装が容易である電子メールに含まれるテキストを利用しないので、配信経路上を通って配信される電子メールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。

【０１８６】

さらに、上記のように、配信経路上の中継装置のＩＰアドレスを利用してメールの判定を行うので、言語情報に依存しない。よって、構文解析が難しくテキストフィルタでは誤判定されやすい日本語のメールに対しても、スパムメールであるかハムメールであるかを確実に検出することができる。

【０１８７】

加えて、迷惑メールであるか否かの判定を行う場合に、送信者による偽装が容易である送信者のアドレス（メールアドレス）を利用しないので、配信経路上を通って配信されるメールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。さらに、従来のアドレスフィルタのように、悪質な送信者との鼬ごっこのような偽装アドレスの登録及び削除を繰り返す必要がなくなり、管理が容易になる。

【０１８８】

なお、請求項１記載の中継アドレス取得ステップ及び請求項８，１５記載の中継アドレス取得手段としては、Ｓ３０１の処理が該当する。また、請求項１記載の迷惑メール中継確率取得ステップ及び請求項８，１５記載の迷惑メール中継確率取得手段としては、Ｓ５

０３及びＳ５１２の処理が該当する。また、請求項１記載の迷惑メール受信確率取得ステップ及び請求項８，１５記載の迷惑メール受信確率取得手段としては、Ｓ５０５の処理が該当する。また、請求項１記載のメール判定ステップ及び請求項８，１５記載のメール判定手段としては、Ｓ５０６～Ｓ５０８，Ｓ５１３，Ｓ５１４の処理及びＳ８の処理が該当する。

【０１８９】

また、請求項２記載の迷惑メール判定ステップ及び請求項９，１６記載の迷惑メール判定手段としては、Ｓ５０６におけるＹｅｓの分岐処理及びＳ５１３の処理が該当する。また、請求項２記載の正当メール判定ステップ及び請求項９，１６記載の正当メール判定手段としては、Ｓ５０７におけるＹｅｓの分岐処理及びＳ５１４の処理が該当する。また、請求項２記載の情報更新ステップ及び請求項９，１６記載の情報更新手段としては、学習処理（Ｓ６）が該当する。また、請求項３記載のグレイメール認識ステップ及び請求項１０，１７記載のグレイメール認識手段としては、Ｓ５０７におけるＮｏの分岐処理及びＳ５０８の処理が該当する。また、請求項４記載の不確定メール再判定ステップ及び請求項１１，１８記載の不確定メール再判定手段としては、グレイメール再判定処理（Ｓ５０９）が該当する。

【０１９０】

また、請求項５記載のアドレス確認ステップ及び請求項１２，１９記載のアドレス確認手段としては、ＩＰアドレス確認処理（Ｓ７）が該当する。また、請求項７記載の中継ルータアドレス取得ステップ及び請求項１４，２１記載の中継ルータアドレス取得手段としては、Ｓ３０３ａの処理が該当する。また、請求項７記載の補完アドレス取得ステップ及び請求項１４，２１記載の補完アドレス取得手段としては、Ｓ３０３ｃの処理が該当する。

【０１９１】

以上、実施例に基づき本発明を説明したが、本発明は上述した実施例に何ら限定されるものではなく、本発明の趣旨を逸脱しない範囲内で種々の改良変更が可能であることは容易に推察できるものである。

【０１９２】

例えば、上記実施例の電子メールフィルタリングシステムは、受信側端末Ｔｒに実装されるシステムであるとして説明したが、これに限定されず、メールサーバＳや中継ルータに実装するシステムとして構成してもよい。また、一部の処理を受信側端末Ｔｒで行い、残りの処理をメールサーバＳで行うなど、複数の装置で分割された処理が実行されて全体として電子メールフィルタリングシステムとして機能するように構成してもよい。

【０１９３】

また、上記実施例の電子メールフィルタリングシステムを機能させるための電子メールフィルタリングプログラムは、所謂パーソナルコンピュータである受信側端末ＴｒのＲＯＭ１２に格納されている制御プログラム１２ａの一部であるとしたが、これに限定されるものではない。例えば、ファイアーウォールやアプライアンスやエンタープライズなどの製品や装置に実装されて、これらの製品や装置において、上記実施例によって説明したような電子メールフィルタリングシステムを機能させるものであってもよい。また、格納場所はＲＯＭに限定されず、ハードディスクなどの書き換え可能なメモリや、各種記憶媒体（ＣＤ－ＲＯＭなど）などであってもよい。

【０１９４】

また、上記実施例の電子メールフィルタリングシステムでは、ベイズ確率ｐｇの値が、０．１以上かつ０．９以下である場合に、そのメールをグレイメールであると判定するように構成したが、グレイメールとする範囲を設けることなく、スパムメールであるかハムメールであるかの二者択一の判定を行うように構成してもよい。例えば、ベイズ確率ｐｇの値が、０．９以上であればスパムメールであると判定し、０．９より小さければハムメールであると判定するように構成してもよい。なお、スパムメールであるか、ハムメールであるか、グレイメールであるかを区分するための閾値は、必要に応じて適宜設定可能で

あることは容易に推察可能な事項である。また、本実施例では、Ｓ５０６において、ベイズ確率ｐｇ＞０．９であるか否かを確認するように構成したが、ベイズ確率ｐｇ≧０．９であるか否かを確認するように構成することも容易に推察可能な事項である。なお、Ｓ５０７の場合についても同様である。

【０１９５】

また、メールの判定区分をスパムメールであるかハムメールであるかの二者択一とする場合には、式１に換えて下記式（３）を用いるように構成してもよい。

【０１９６】

【数３】

$$pg_n = \frac{\frac{b}{nbad}}{\frac{2\times g}{ngood} + \frac{b}{nbad}} \qquad (3)$$

また、この場合、本実施例においてベイズ確率「pg_n」に対して設定した上限「０．９８」に換えて「０．９９」とすることが好ましい。なお、ベイズ確率「pg_n」の値の上限及び下限についても必要に応じて適宜設定できることは容易に推察可能である。

【０１９７】

また、メールの判定区分をスパムメールであるかハムメールであるかの二者択一とする場合には、配信木情報が新規の中継装置に対するものであった場合に、ベイズ確率「pg_n」の値を「０．４」とすることが好ましいが、この定数の値もまた、適宜設定可能であることは容易に推察可能である。

【０１９８】

また、上記実施例では、グレイメール再判定処理（Ｓ５０９）において、グレイメールの本文中のテキスト情報に対して従来のテキストフィルタをかけるように説明したが、これに限定されるものではなく、グレイメールのメールヘッダにおけるメールアドレスなどの他の部分にテキストフィルタを適用するように構成してもよい。

【０１９９】

また、上記実施例では、Ｓ２において、１通ずつメールサーバＳから受信するように構成したが、メールサーバＳにあるメールを全て受信してから処理するように構成してもよい。また、上記実施例では、受信したメールをＲＡＭ１４（受信メールメモリ１４ａ）に一時的に記憶させるように構成したが、記憶部１６に記憶させるように構成してもよい。

【０２００】

また、上記実施例では、学習処理（Ｓ６）と配信木構築処理（Ｓ４）とをそれぞれ独立した処理として説明したが、学習処理（Ｓ６）におけるＳ６０３，Ｓ６０４，Ｓ６０６の処理を、配信木構築処理（Ｓ４）の中で実行されるように構成してもよい。

【０２０１】

また、上記実施例における配信木構築処理（Ｓ４）では、受信側端末Ｔｒに近い側の中継装置から順に、そのＩＰアドレスを経路情報メモリ１４ｂから読み出すように構成したが、経路情報メモリ１４ｂからＩＰアドレスを読み出す順序はこれに限定されるものではない。

【０２０２】

また、上記実施例では、受信したメールがメール判定手段（Ｓ５）によって判定された後に、その判定結果に基づいて学習処理（Ｓ６）が実行されるように構成されているが、学習処理（Ｓ６）は、ユーザが受信したメールに対してスパムメールであるか否かを判定した結果に基づいて実行するように構成してもよい。この場合は、図９に示した学習処理

（Ｓ６）において、Ｓ６０１の処理後に、経路情報取得処理（Ｓ３）を実行して、ユーザが判定したメールの経路情報を取得するように構成すればよい。

【図面の簡単な説明】

【０２０３】

【図１】本発明の第１実施例における電子メールフィルタリングシステムが実装される電子メールの配信経路の一例を示す模式図である。

【図２】メールの「Received:」フィールドを示す模式図である。

【図３】本実施例の電子メールフィルタリングシステムを機能させる受信側端末の構成を示すブロック図である。

【図４】受信側端末に実装された電子メールフィルタリングプログラムによって実行されるメール受信処理を示すフローチャートである。

【図５】経路情報取得処理を示すフローチャートである。

【図６】配信木構築処理を示すフローチャートである。

【図７】メール判定処理を示すフローチャートである。

【図８】図１の配信経路におけるメールの偏りを説明するための模式図である。

【図９】学習処理を示すフローチャートである。

【図１０】第２実施例における経路情報取得処理を示すフローチャートである。

【図１１】第２実施例の経路追跡処理による結果を説明するための模式図である。

【図１２】第２実施例の電子メールフィルタリングシステムと従来のホワイトリストフィルタとを比較するグラフである。

【図１３】第２実施例の電子メールフィルタリングシステムと従来のブラックリストフィルタとを比較するグラフである。

【図１４】グレイメール再判定処理を実行した場合と実行しなかった場合におけるメールの判定精度を比較するグラフである。

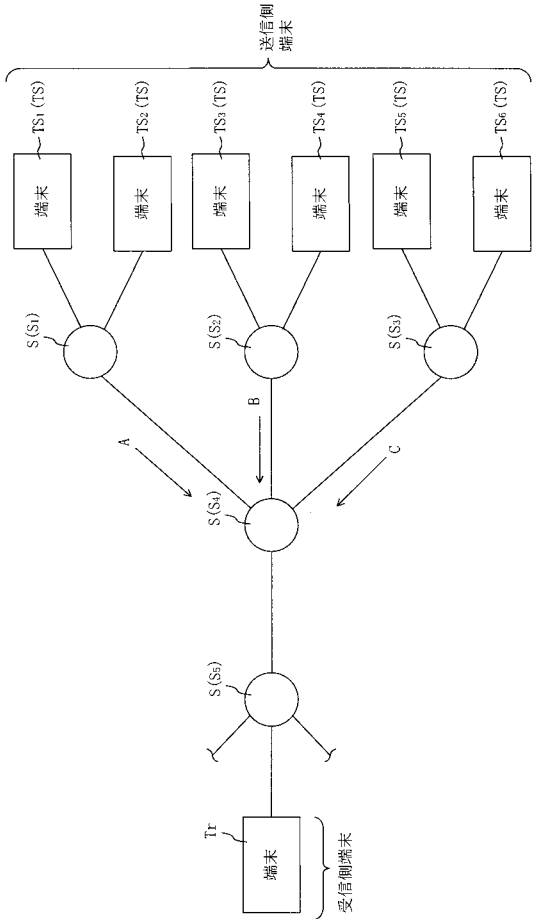
【図１５】テキスト型ベイジアンフィルタを併用した場合と、テキスト型ベイジアンフィルタのみを使用した場合とを比較するグラフである。

【符号の説明】

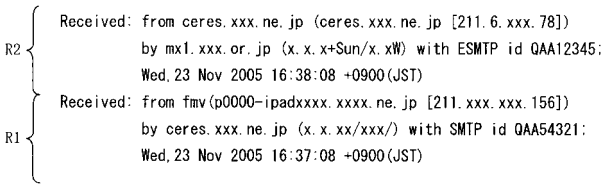
【０２０４】

１６ａ（$16a_1$～$16a_n$）	中継装置メモリ（情報記憶手段）
$16a_{11}$～$16a_{n1}$	スパム中継カウンタ（情報記憶手段）
$16a_{12}$～$16a_{n2}$	ハム中継カウンタ（情報記憶手段）
１６ｂ	スパム受信カウンタ（情報記憶手段）
１６ｃ	ハム受信カウンタ（情報記憶手段）
Ｓ	メールサーバ（中継装置）
Ｔｓ	送信側端末（中継装置）
Ｔｒ	受信側端末（コンピュータ）
Ｒ１，Ｒ２	「Received:」フィールド（ヘッダ情報）

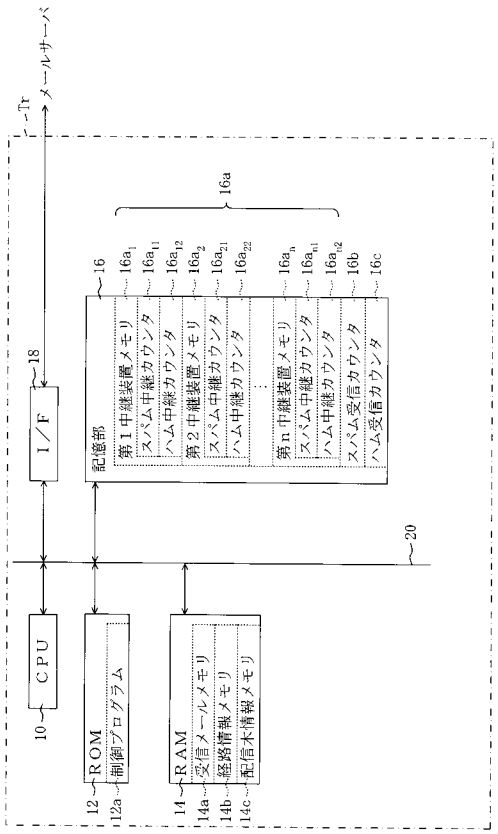
【図１】


送信側端末
TS1(TS)
TS2(TS)
TS3(TS)
TS4(TS)
TS5(TS)
TS6(TS)
端末
S(S1)
S(S2)
S(S3)
S(S4)
S(S5)
A
B
C
Tr
受信側端末

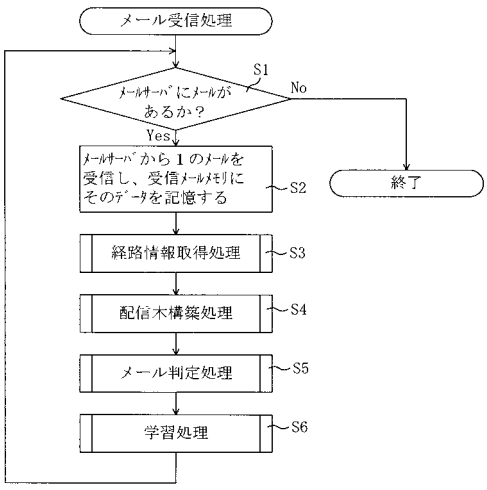
【図２】


R2
Received: from ceres.xxx.ne.jp (ceres.xxx.ne.jp [211.6.xxx.78])
by mx1.xxx.or.jp (x.x.x+Sun/x.xW) with ESMTP id QAA12345;
Wed,23 Nov 2005 16:38:08 +0900(JST)
R1
Received: from fmv(p0000-ipadxxxx.xxxx.ne.jp [211.xxx.xxx.156])
by ceres.xxx.ne.jp (x.x.xx/xxx/) with SMTP id QAA54321;
Wed,23 Nov 2005 16:37:08 +0900(JST)

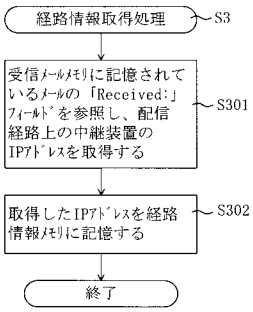
【図３】


メールサーバ
Tr
16a
16
記憶部
第1中継装置メモリ
スパム中継カウンタ
ハム中継カウンタ
第2中継装置メモリ
スパム中継カウンタ
ハム中継カウンタ
第n中継装置メモリ
スパム中継カウンタ
ハム中継カウンタ
スパム受信カウンタ
ハム受信カウンタ
16b
16c
18
I/F
20
CPU
10
ROM
制御プログラム
12
12a
RAM
受信メールメモリ
経路情報メモリ
配信木情報メモリ
14
14a
14b
14c

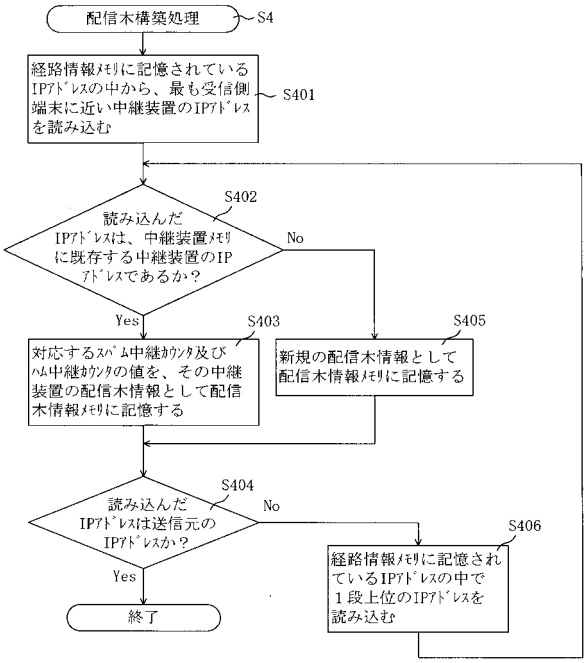
【図４】


メール受信処理
S1
メールサーバにメールがあるか？
No
Yes
終了
メールサーバから1のメールを受信し、受信メールメモリにそのデータを記憶する
S2
経路情報取得処理
S3
配信木構築処理
S4
メール判定処理
S5
学習処理
S6

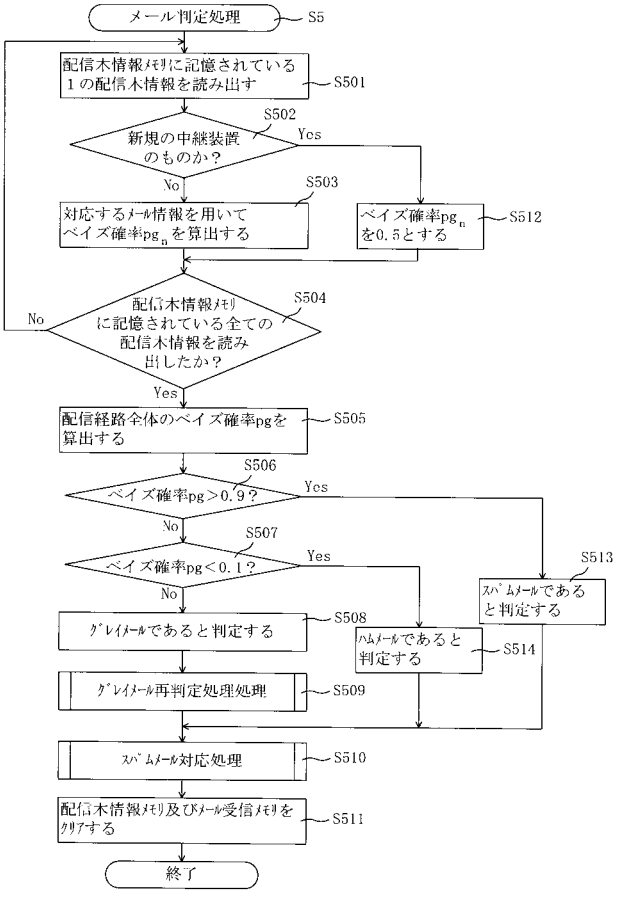
【図５】


経路情報取得処理
S3
受信ﾒｰﾙﾒﾓﾘに記憶されているﾒｰﾙの「Received:」ﾌｨｰﾙﾄﾞを参照し、配信経路上の中継装置のIPｱﾄﾞﾚｽを取得する
S301
取得したIPｱﾄﾞﾚｽを経路情報ﾒﾓﾘに記憶する
S302
終了

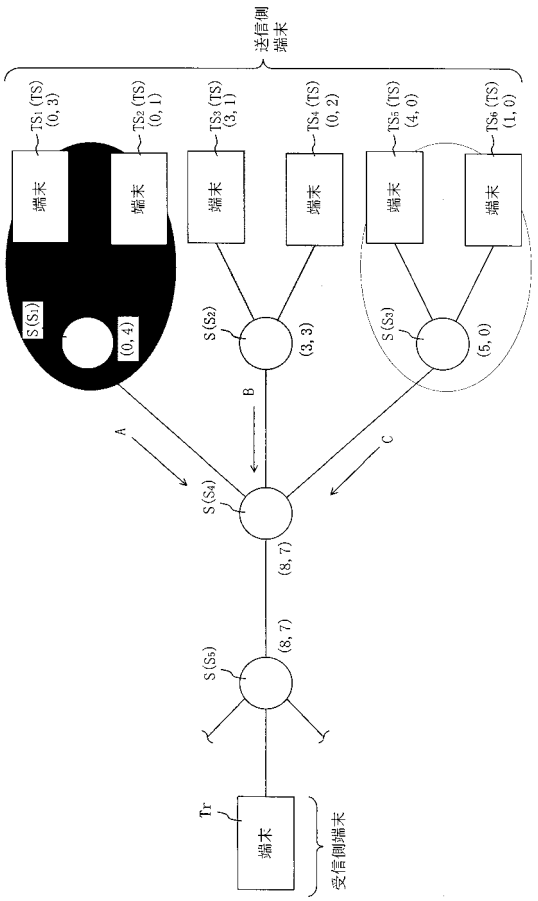
【図６】


配信木構築処理
S4
経路情報ﾒﾓﾘに記憶されているIPｱﾄﾞﾚｽの中から、最も受信側端末に近い中継装置のIPｱﾄﾞﾚｽを読み込む
S401
読み込んだIPｱﾄﾞﾚｽは、中継装置ﾒﾓﾘに既存する中継装置のIPｱﾄﾞﾚｽであるか？
S402
No
Yes
対応するｽﾊﾟﾑ中継ｶｳﾝﾀ及びﾊﾑ中継ｶｳﾝﾀの値を、その中継装置の配信木情報として配信木情報ﾒﾓﾘに記憶する
S403
新規の配信木情報として配信木情報ﾒﾓﾘに記憶する
S405
読み込んだIPｱﾄﾞﾚｽは送信元のIPｱﾄﾞﾚｽか？
S404
No
Yes
経路情報ﾒﾓﾘに記憶されているIPｱﾄﾞﾚｽの中で１段上位のIPｱﾄﾞﾚｽを読み込む
S406
終了

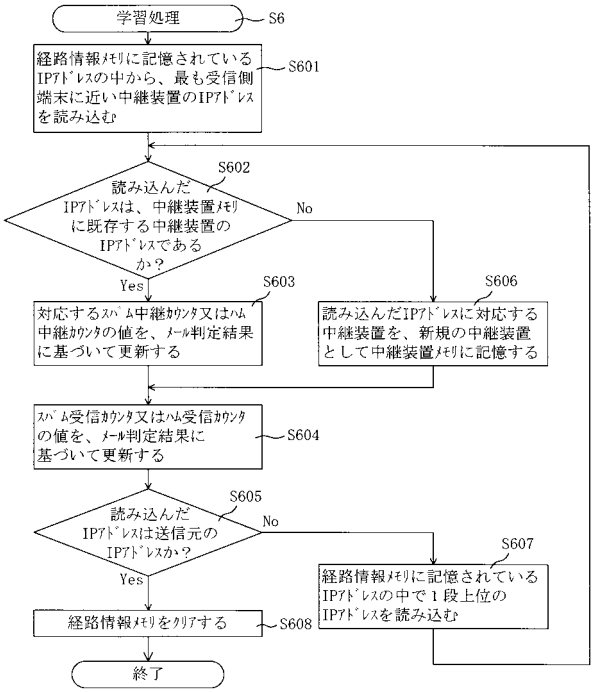
【図７】


メール判定処理
S5
配信木情報ﾒﾓﾘに記憶されている１の配信木情報を読み出す
S501
新規の中継装置のものか？
S502
Yes
No
対応するﾒｰﾙ情報を用いてベイズ確率pg_n を算出する
S503
ベイズ確率pg_n を0.5とする
S512
配信木情報ﾒﾓﾘに記憶されている全ての配信木情報を読み出したか？
S504
No
Yes
配信経路全体のベイズ確率pgを算出する
S505
ベイズ確率pg＞0.9？
S506
Yes
No
ベイズ確率pg＜0.1？
S507
Yes
No
ｽﾊﾟﾑﾒｰﾙであると判定する
S513
ｸﾞﾚｲﾒｰﾙであると判定する
S508
ﾊﾑﾒｰﾙであると判定する
S514
ｸﾞﾚｲﾒｰﾙ再判定処理処理
S509
ｽﾊﾟﾑﾒｰﾙ対応処理
S510
配信木情報ﾒﾓﾘ及びﾒｰﾙ受信ﾒﾓﾘをｸﾘｱする
S511
終了

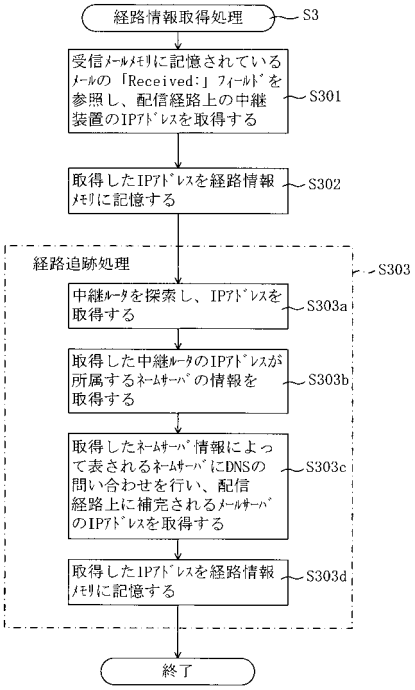
【図８】


送信側端末
TS1 (TS) (0, 3)
TS2 (TS) (0, 1)
TS3 (TS) (3, 1)
TS4 (TS) (0, 2)
TS5 (TS) (4, 0)
TS6 (TS) (1, 0)
端末
S(S1)
(0, 4)
S(S2)
(3, 3)
S(S3)
(5, 0)
A
B
C
S(S4)
(8, 7)
(8, 7)
S(S5)
Tr
端末
受信側端末

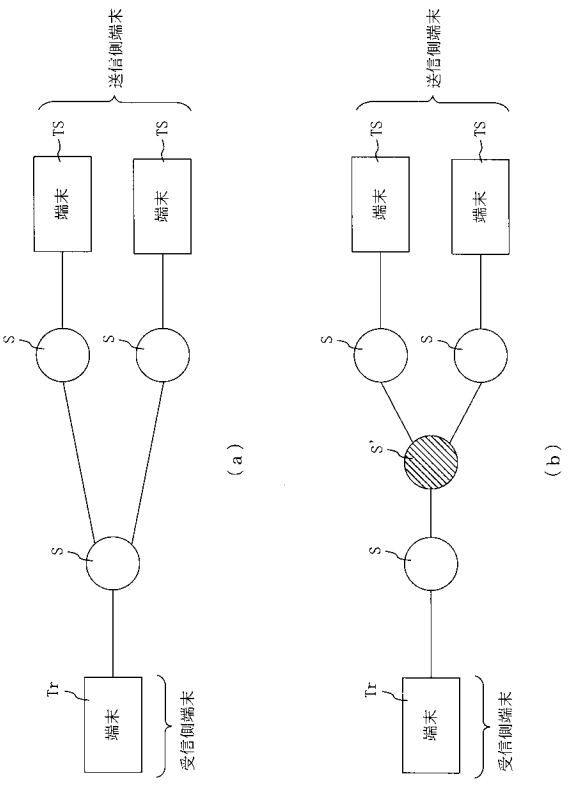
【図９】



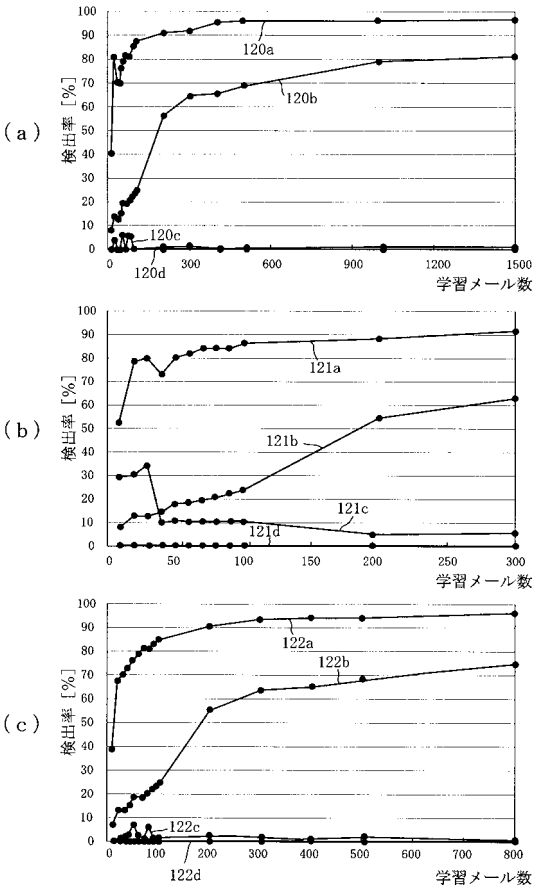
【図１０】



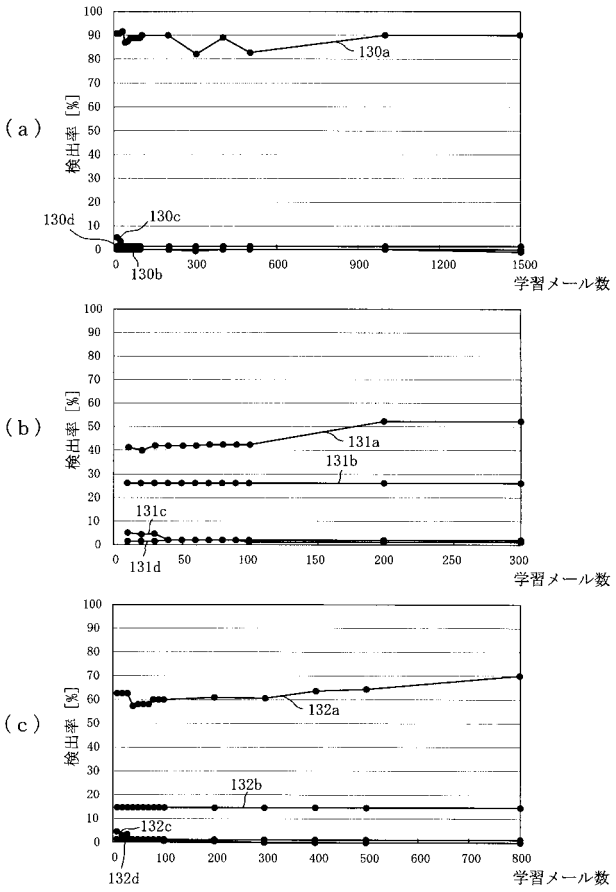
【図１１】



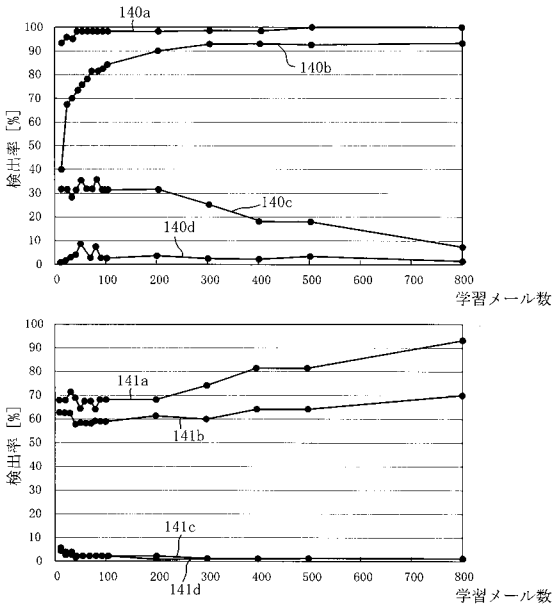
【図１２】



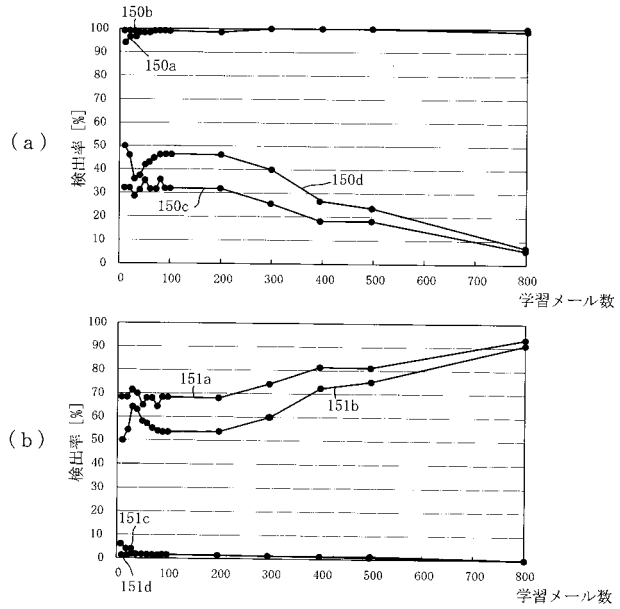
【図１３】


(a)
検出率［%］
学習メール数
130a
130b
130c
130d
(b)
検出率［%］
学習メール数
131a
131b
131c
131d
(c)
検出率［%］
学習メール数
132a
132b
132c
132d

【図１４】


(a)
検出率［%］
学習メール数
140a
140b
140c
140d
(b)
検出率［%］
学習メール数
141a
141b
141c
141d

【図１５】


(a)
検出率［%］
学習メール数
150a
150b
150c
150d
(b)
検出率［%］
学習メール数
151a
151b
151c
151d

フロントページの続き

(72)発明者　白川　正知
愛知県豊橋市天伯町雲雀ヶ丘１－１　国立大学法人豊橋技術科学大学内

Ｆターム(参考) 5K030 GA14 HA05 HA06 JA10 JA11 KA02 MB16