

明 細 書

発明の名称：

情報処理装置、情報処理方法、及び情報処理プログラム

技術分野

[0001] 本発明の実施形態は、情報処理装置、情報処理方法、及び情報処理プログラムに関する。

背景技術

[0002] Webサーバは、サーバを管理する管理者により作成、更新されたコンテンツを保存して、クライアント端末からのリクエストに応じてコンテンツを返す動作を実行する。

[0003] ところで、Webサーバには、OSを含む各種のソフトウェアがインストールされており、これらのソフトウェアにより、コンテンツの作成、更新の支援やクライアント端末との通信などが実行されている。これらのソフトウェアに脆弱性（セキュリティ上の弱点）が存在すると、悪意の第三者がこの脆弱性を利用してサーバへの攻撃（例えば、SQLインジェクション）を実行するおそれがある。

[0004] 悪意の第三者による攻撃が成功した場合、悪意の第三者はサーバ内へ侵入して、コンテンツの内容を意図的に改ざんすることで水飲み場型攻撃などを仕掛ける場合がある。水飲み場型攻撃とは、ユーザがインターネットを介して普段アクセスする、正規のWebサイトを悪意の第三者が改ざんして、正規のWebサイトにアクセスしたクライアント端末を、マルウェアなどをダウンロードする悪性のサイトへ誘導するサイバー攻撃の一種である。

[0005] マルウェアなどをダウンロードする悪性URLへのアクセスを防止するためには、悪性URLを予め多く取得することが重要となる。従来では、ユーザの通信ログを介して取得したURL等に基づいて取得したマルウェアファイルを仮想的に実行する実行装置を備えて、マルウェアを実行させた際に、このマルウェアファイルが通信を行う宛先となるURL等を取得して接続先

のURLをブラックリストとして収集する技術が開示されている。

先行技術文献

特許文献

[0006] 特許文献1：特開2014-179025号公報

発明の概要

発明が解決しようとする課題

[0007] しかしながら、マルウェアをダウンロードさせるサイト、あるいはフィッシングサイト（詐欺）等に誘導する悪性のURLを多く検出できたとしても、正規なWebサイトに対応するコンテンツに意図的な改ざんが実行されて、正規なWebサイトへアクセスした場合に悪性のサイトに遷移させる攻撃が仕掛けられた場合、改ざんされた正規なWebサイトのURLを即時に検知することは困難である。

[0008] このため、正規なサイトに対応するサーバを管理する管理者は、コンテンツの改ざんに早期に気付くことはできず、正規なサイトがユーザにとって有害なWebサイトのまま放置されてしまうおそれがある。

[0009] 本発明はこのような事情を考慮してなされたもので、Webサーバのコンテンツの改ざんを早期に検知できる情報処理装置、情報処理方法、及び情報処理プログラム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

[0010] 本発明の実施形態に係る情報処理装置は、クライアント端末がWebサーバに対してコンテンツをリクエストした際のアクセスログを収集するアクセスログ収集部と、悪性URLを予め登録するデータベースと、クライアント端末の接続先に対応する接続先URLをデータベースと照合して、接続先URLが悪性URLと一致する場合に、コンテンツの改ざんを検知する改ざん検知部と、を備えることを特徴とする。

発明の効果

[0011] 本発明の実施形態により、Webサーバのコンテンツの改ざんを早期に検

知できる情報処理装置、情報処理方法、及び情報処理プログラムが提供される。

図面の簡単な説明

[0012] [図1]第1実施形態に係る情報処理装置の構成の一例を示す構成図。

[図2] (A) クライアント端末からのリクエスト後に、URLリダイレクトにより、リクエスト先のURLとは別のURLに接続された場合の、アクセスの流れを説明する説明図、(B) 情報処理装置のログ抽出部において抽出されたアクセスログの一例を示す図。

[図3] (A) データベースで保存されるURLの一例を示す説明図、(B) 良性URLに関連付けられた管理者への通知先の一例を示す説明図。

[図4]第1実施形態に係る情報処理方法の一例を示すフローチャート。

[図5] (A) ドライブバイダウンロード攻撃の一例であり、A社のWebサーバで管理されるhtmlデータが悪意の第三者により改ざんされた場合に、特定の検索サイトのリファラを有した状態でリクエストがされたときのアクセスの流れを示す図、(B) 検索サイトを介さないでhtmlデータのリクエストがされたときのアクセスの流れを示す図。

[図6]第2実施形態に係る情報処理装置の構成の一例を示す構成図。

[図7]第2実施形態に係る情報処理方法の一例を示すフローチャート。

[図8]第3実施形態に係る情報処理装置の構成の一例を示す構成図。

[図9]第3実施形態に係る情報処理方法の一例を示すフローチャート。

発明を実施するための形態

[0013] (第1実施形態)

以下、本発明の実施形態を添付図面に基づいて説明する。

図1は、第1実施形態に係る情報処理装置10の構成の一例を示す図である。

[0014] まず、全体の構成について説明する。

クライアント端末11は、インターネットを介してWebサーバ12と接続されて、URLにより特定されるWebサーバ12のコンテンツをリクエ

ストする。そして、Webサーバ12から返されたコンテンツを、Webブラウザを介して表示させる。なお、コンテンツとしては、htmlデータ、画像データ、pdfファイル等の種々のデータが例示される。図1では、コンテンツとしてhtmlデータを例示している。

[0015] Webサーバ12は、クライアント端末11から要求されたコンテンツをクライアント端末11に返す。

[0016] 管理者端末13は、Webサーバ12に接続可能なコンピュータであり、管理者によりコンテンツの作成、更新が実行される。

[0017] 実施形態に係る情報処理装置10は、ユーザのクライアント端末11から収集したアクセスログを、悪性URLを予め登録したデータベース16と照合することで、ユーザのアクセス元のURLに対応するコンテンツの改ざんを検知する。

[0018] 情報処理装置10は、クライアント端末11のWebサーバ12への接続情報に関するアクセスログを収集する。情報処理装置10は、クライアント端末11から直接アクセスログを取得してもよいし、クライアント端末11が、インターネットへの接続を制御するためのプロキシサーバを介してWebサーバ12に接続される場合には、情報処理装置10はプロキシサーバからアクセスログを取得してもよい。また、外部のネットワークへの出入り口でクライアント端末11の通信制御を行うファイアウォールなどを介してアクセスログを取得してもよい。また、クライアント端末11からアクセスログを随時収集する別個のサーバを経由して情報処理装置10はアクセスログを取得してもよい。

[0019] なお、図1では、クライアント端末11を1つで記載しているが、情報処理装置10は、複数のクライアント端末11に接続されて、クライアント端末11のそれぞれからアクセスログを収集してもよい。

[0020] 第1実施形態に係る情報処理装置10の具体的な構成について説明する。

情報処理装置10は、アクセスログ収集部14と、ログ抽出部15と、データベース16と、改ざん検知部19と、改ざん通知部20と、を備えてい

る。

- [0021] なお、情報処理装置 10 を構成する各ユニットの機能は、所定のプログラムコードを、プロセッサを用いて実行することによって実現しても良く、このようなソフトウェア処理に限らず、例えば、ASIC 等を用いたハードウェア処理で実現してもよいし、ソフトウェア処理とハードウェア処理とを組み合わせることで実現してもよい。
- [0022] アクセスログ収集部 14 は、クライアント端末 11 が Web サーバ 12 に対してコンテンツをリクエストした際のアクセスログを収集する。アクセスログ収集部 14 は、クライアント端末 11 が Web サーバ 12 に接続した際の、クライアント端末 11 を識別する識別情報（例えば、IP アドレス）、アクセス元 URL、接続先 URL、コンテンツのアクセスに用いた Web ブラウザ名、アクセス時間などをアクセスログとして取得する。
- [0023] アクセス元 URL は、クライアント端末 11 からコンテンツをリクエストするための URL を意味する。接続先 URL は、コンテンツのリクエスト後に、実際にクライアント端末 11 が接続された URL を意味する。通常、クライアント端末 11 がコンテンツをリクエストするためにアクセスする URL と、そのリクエスト後に、実際にクライアント端末 11 が接続される URL とは一致する。つまりアクセス元 URL と接続先 URL とは一致する。
- [0024] 一方で、クライアント端末 11 からのリクエスト後に、URL リダイレクト（URL の参照先が自動的に変更されること）されて、リクエスト先の URL とは異なる URL にクライアント端末 11 が自動で接続された場合、クライアント端末 11 からコンテンツがリクエストされた URL と、リクエスト後に実際にクライアント端末 11 が接続された URL とは一致しない。つまりアクセス元 URL と接続先 URL とは一致しないものとなる。
- [0025] 図 2（A）は、クライアント端末 11 からのリクエスト後に、URL リダイレクトにより、リクエスト先の URL とは別の URL にクライアント端末 11 が接続された場合の、アクセスの流れを説明する説明図である。
- [0026] クライアント端末 11 が、アクセス元 URL となる「<http://www.aaa.com/>

aaa.html」にアクセスして、Webサーバ12にコンテンツ「aaa.html」をリクエストする。クライアント端末11には、httpレスポンスが返される。そして、クライアント端末11は、「http://www.ccc.com/ccc.html」にURLリダイレクトされて、Webサーバ50にコンテンツ「ccc.html」を要求する。クライアント端末11は、接続先URLからccc.htmlデータを取得する。このように、URLリダイレクトにより、リクエスト先のURLとは一致しないURLにクライアント端末11が接続された場合、アクセス元URLと接続先URLとは一致しない。

[0027] ログ抽出部15は、アクセスログ収集部14で収集されたアクセスログから、クライアント端末11の接続先に対応する接続先URLと、コンテンツのリクエスト先に対応するアクセス元URLとが異なる、つまり2つのURLが一致しないアクセスログを抽出する。

[0028] 図2(B)は、情報処理装置10のログ抽出部15において抽出されたアクセスログの一例を示す図である。

[0029] 図2(B)に示すように、ログ抽出部15は、ユーザのクライアント端末11を識別するユーザID(IPアドレス)ごとに、クライアント端末11が実際に接続した接続先URL、ユーザがコンテンツをリクエストするためにアクセスしたアクセス元URLなどを抽出する。

[0030] データベース16は、悪性URLが予め登録された悪性URL保存部17と、良性URLが予め登録された良性URL保存部18と、を有している。なお、以下の実施形態では、データベース16において、URLに対応させて良性または悪性のカテゴリに分けて登録しているが、IPアドレスに対応させて良性または悪性のカテゴリに分けて登録してもよい。

[0031] 悪性URLとは、クライアント端末11にウィルスなどのマルウェアをダウンロードさせるコンテンツ、フィッシング(詐欺)サイトなど、悪意のあるファイルのダウンロードやWebサイトへの接続にクライアント端末11を誘導する有害なURLを意味する。一方、良性URLとは、クライアント端末11にとって無害なURLを意味する。データベース16では、登録さ

れたURLのそれぞれに悪性、または、良性のカテゴリが付与されて保存されている。

[0032] データベース16は、良性URLのそれぞれに対応させて、良性URLに対応するコンテンツを管理する管理者端末13（または管理者）への通知先を保存している。通知先としては、例えばメールアドレスが例示される。この通知先は、コンテンツの改ざんを検知したときに、改ざんの発生を管理者側に通知する際に用いられる。

[0033] 図3（A）は、データベース16で保存されるURLの保存例を示す図である。図3（A）に示すように、データベース16で登録されるURLのそれぞれについて、良性または悪性のカテゴリ情報が関連付けて保存される。

[0034] 図3（B）は、良性URLに関連付けられた管理者への通知先の一例を示す説明図である。図3（B）に示すように、登録される良性URLのそれぞれについて、良性URLに対応するコンテンツを管理する管理者への通知先となるメールアドレスが保存される。

[0035] 図1に戻って説明を続ける。

改ざん検知部19は、ログ抽出部15において抽出された、クライアント端末11の接続先に対応する接続先URLと、コンテンツのリクエスト先に対応するアクセス元URLとが一致しないログをログ抽出部15から取得する。そして、クライアント端末11の実際の接続先となる接続先URLをデータベース16と照合して、接続先URLが悪性URLと一致する場合に、アクセス元URLに対応するコンテンツの改ざんを検知する。なお、データベース16において、IPアドレスに対応させて良性または悪性のカテゴリに分けて登録している場合には、接続先URLに対応するIPアドレスが悪性のIPアドレスと一致する場合に、コンテンツの改ざんを検知する。

[0036] 通常、ユーザが正規のWebサイトにアクセスする場合、アクセス元のURLとクライアント端末11による実際の接続先URLは一致し、いずれのURLも良性URLに属する。このため、接続先URLとアクセス元URLとが一致せず、接続先URLが悪性URLと一致する場合、正規のWebサ

イトへの接続が意図的に変更されているとして、正規のWebサイトの改ざん、すなわちアクセス元URLに対応するコンテンツの改ざんが発生していると判定できる。

[0037] なお、正規のWebサイトが、Webサイトにアクセスしてきたクライアント端末11を意図的にURLリダイレクトするときには、アクセス元のURLと実際の接続先URLは一致しない場合が想定される。この場合、接続先URLが悪性URLと一致することは無いため、コンテンツの改ざんが検知されることは無い。

[0038] 改ざん通知部20は、改ざんが検知された場合に、改ざんが検知されたコンテンツに対応するWebサーバ12の管理者に改ざんの発生を通知する。具体的には、良性URLに対応するコンテンツを管理する管理者端末13（または管理者）の通知先を用いて改ざんの発生を通知する。また、管理者への通知方法として、管理者端末13の通知先を予め保存しておく必要は無く、改ざんが検知されたコンテンツに対応するアクセス元URLに基づき管理者のアクセス先を検索して、改ざんの発生をWebサーバ12の管理者に通知してもよい。例えば、アクセス元URLからドメイン部分を抽出して、このドメイン部分と、問い合わせ先（コンタクト先）のURLにおいて一般的に用いられる「/content/」、「/info/」等を連結させることで通知先を探索し、この通知先に基づいて改ざんの発生をWebサーバ12の管理者に通知する。

[0039] 続いて、第1実施形態に係る情報処理装置10の動作について説明する。

図4は、第1実施形態に係る情報処理方法のフローチャートである（適宜、図1参照）。

[0040] アクセスログ収集部14は、クライアント端末11から、インターネットへのアクセスログを取得する（S10）。

[0041] ログ抽出部15は、クライアント端末11から取得したアクセスログの中から、実際にクライアント端末11に接続された接続先URLと、コンテンツのリクエスト先に対応するアクセス元URLとが一致しないログを抽出す

る（S 1 1）。

[0042] 改ざん検知部 1 9 は、接続先 URL をデータベース 1 6 で照合して、接続先 URL が悪性 URL に一致するか否かを判定する。接続先 URL が悪性 URL に一致する場合には、アクセス元 URL に対応するコンテンツの改ざんを検知する（S 1 2, S 1 3 : YES, S 1 4）。一方、改ざん検知部 1 9 は、一致しない場合には、コンテンツの改ざんは発生していない判定する（S 1 3 : NO, 終了）。

[0043] 改ざん通知部 2 0 は、改ざんが検知された場合に、改ざんが検知されたコンテンツに対応する Web サーバ 1 2 の管理者に改ざんの発生を通知する（S 1 5）。また、改ざんが検知されたコンテンツに対応するアクセス元 URL を、データベース 1 6 に悪性 URL として保存してもよい。すでに良性 URL にカテゴリされている URL については、悪性 URL にカテゴリを変更する。

[0044] このように、第 1 実施形態に係る情報処理装置 1 0 は、ユーザのクライアント端末 1 1 から収集したアクセスログを、悪性 URL を予め登録したデータベース 1 6 と照合することで、Web サーバ 1 2 の改ざんを早期に検知できる。そして、Web サーバ 1 2 の管理者にコンテンツの改ざんを通知することができる。これにより、管理者は、Web サーバ 1 2 の改ざんを早期に気付くことができ、水飲み場型攻撃などの悪意の第三者による攻撃を短期間で終息させることができる。

[0045] （第 2 実施形態）

まず、図 5 を用いて、水飲み場型攻撃を利用したドライブバイダウンロード攻撃の一例について説明する。ここでは、A 社の Web サーバ 1 2 で管理される html データ（aaa.html）が悪意の第三者により改ざんされて、特定の検索サイトのリファラを有した状態で aaa.html のリクエストがされたときに、悪意の第三者の Web サーバ 5 0 への URL リダイレクトが実行される特殊なリダイレクトスクリプトが aaa.html に埋め込まれているものとする。リファラとは、クライアント端末 1 1 に接続されている URL に対して、当該

URLに遷移する直前にアクセスしていたURLを意味する。

[0046] 図5（A）に示すように、ユーザが、検索サイトXにA社に関連するキーワードを入力すると、クライアント端末11にはキーワードに応じたA社のコンテンツのURLが表示される。そして、ユーザのクライアント端末11は、A社のWebサーバ12にaaa.htmlをリクエストする。このとき、検索サイトXのリファラを有した状態でアクセスを実行したクライアント端末11は、aaa.htmlに埋め込まれたリダイレクトスクリプトにより、悪意の第三者が管理するWebサーバ50に誘導されて、悪性のコンテンツであるccc.htmlを取得する。

[0047] 一方、図5（B）に示すように、クライアント端末11が、検索サイトXを介さないでaaa.htmlのリクエストがされたときに、埋め込まれたリダイレクトスクリプトは実行されず、クライアント端末11はA社のWebサーバ12からaaa.htmlを取得する。このように、ブックマークによるアクセスなど、URLに直接アクセスした際にはスクリプトは実行されず、検索サイトを介したアクセス時にのみクライアント端末11をURLリダイレクトさせる。このため、通常検索サイトを介してアクセスすることの無い管理者は、改ざんが発生している場合でも気づきにくく、改ざんの検知が遅れるおそれがある。

[0048] そこで、第2実施形態に係る情報処理装置10では、クライアント端末11の実際の接続先URLが悪性URLと一致する場合において、情報処理装置10からアクセス元URLに直接リクエスト、すなわちリファラ無しでコンテンツをリクエストすることで、接続先URLの検証を実行する。

[0049] 図6は、第2実施形態に係る情報処理装置10の構成の一例を示す図である。なお、図6において第1実施形態（図1）と共通の構成又は機能を有する部分は、同一符号で示し、重複する説明を省略する。以下では、アクセスログから抽出された接続先URLを「第1の接続先URL」として、アクセス検証部21においてアクセス元URLが直接リクエストされた際の接続先URLを「第2の接続先URL」として区別して記載する。

- [0050] ログ抽出部 15 は、アクセスログ収集部 14 で収集されたアクセスログから、クライアント端末 11 の接続先に対応する第 1 の接続先 URL と、コンテンツのリクエスト先に対応するアクセス元 URL とが一致しないアクセスログを抽出する。
- [0051] アクセス検証部 21 は、第 1 の接続先 URL をデータベース 16 で照合して、第 1 の接続先 URL が悪性 URL と一致する場合に、Web サーバ 12 に対して、アクセスログから抽出されたアクセス元 URL を直接リクエストする。そして、アクセス検証部 21 は、この直接リクエスト実行時において、第 2 の接続先 URL とアクセス元 URL とを比較する。
- [0052] 改ざん検知部 19 は、アクセス検証部 21 の比較によりアクセス元 URL と第 2 の接続先 URL とが一致する場合に、コンテンツの改ざんを検知する。
- [0053] 収集されたアクセスログにおいて、アクセス元 URL と第 1 の接続先 URL とが一致せず、第 1 の接続先 URL が悪性 URL に一致するにも関わらず、Web サーバ 12 に対してアクセス元 URL を直接リクエストした場合に、第 2 の接続先 URL とアクセス元 URL とが一致することは、正規の Web サイトへの接続がアクセス方法に応じて意図的に変更されているとして、正規の Web サイトの改ざん、すなわちアクセス元 URL に対応するコンテンツの改ざんが発生していると判定できる。
- [0054] 一方、アクセス検証部 21 の比較によりアクセス元 URL と第 2 の接続先 URL が一致せず、クライアント端末 11 が悪意の第 3 者の Web サーバ 50 に URL リダイレクトされることで第 2 の接続先 URL が悪性 URL に一致するときにはコンテンツの改ざんを検知する。なお、第 2 の接続先 URL が悪性 URL に一致しないときは、接続された第 2 の接続先 URL は無害となるものの、アクセス元 URL に対応するコンテンツの改ざんの有無が未定となるため、情報処理装置 10 の管理者よりアクセス元 URL が再検証されることで、コンテンツの改ざんの有無が確認される。
- [0055] 続いて、第 2 実施形態に係る情報処理装置 10 の動作について説明する。

図7は、第2実施形態に係る情報処理方法の一例を示すフローチャートである（適宜、図6参照）。

- [0056] アクセスログ収集部14は、クライアント端末11から、インターネットへのアクセスログを取得する（S20）。
- [0057] ログ抽出部15は、クライアント端末11から取得したアクセスログの中から、ユーザがリクエストするコンテンツに対応するアクセス元URLと、実際にクライアント端末11に接続された第1の接続先URLとが一致しないログを抽出する（S21）。
- [0058] アクセス検証部21は、第1の接続先URLが悪性URLに一致するかどうかを判定する（S22）。
- [0059] そして、アクセス検証部21は、第1の接続先URLが悪性URLと一致する場合に、リファラを無しの状態で、Webサーバ12に対してアクセス元URLに直接アクセスを実行して第2の接続先URLを取得する（S22；YES、S23）。第1の接続先URLが悪性URLと一致しない場合、情報処理装置10の管理者によりアクセス元URLが再検証されることで、コンテンツの改ざんの有無を確認する（S22：NO、S28）。
- [0060] 再検証の結果、改ざんが確認された場合には、情報処理装置10の管理者はコンテンツに対応するアクセス元URLを悪性URLとしてデータベース16に保存する。加えて、アクセス元URLに対応するコンテンツの管理者に改ざんの発生を通知する。改ざんが確認されなかった場合は、アクセス元URLを良性URLとしてデータベース16に保存する。
- [0061] 改ざん検知部19は、アクセス検証部21の比較によりアクセス元と第2の接続先URLが一致する場合には、アクセス元URLに対応するコンテンツの改ざん検知する（S24：YES、S26）。
- [0062] 一方、アクセス検証部21の比較によりアクセス元URLと第2の接続先URLが一致しない場合において、クライアント端末11が悪意の第3者のWebサーバ50にURLリダイレクトされることで第2の接続先URLが悪性URLに一致するときには、コンテンツの改ざんを検知する（S24：

NO、S 2 5 : YES、S 2 6)。

[0063] また、アクセス検証部 2 1 の比較によりアクセス元 URL と第 2 の接続先 URL が一致しない場合において、第 2 の接続先 URL が悪性 URL に一致しないときは、情報処理装置 1 0 の管理者によりアクセス元 URL が再検証されることで、コンテンツの改ざんの有無が確認される (S 2 4 : NO、S 2 5 : NO、S 2 8)。

[0064] 再検証の結果、改ざんが確認された場合には、情報処理装置 1 0 の管理者はコンテンツに対応するアクセス元 URL 及びアクセス検証時に悪性 URL と一致しなかった第 2 の接続先 URL を、悪性 URL としてデータベース 1 6 に保存する。加えて、アクセス元 URL に対応するコンテンツの管理者に改ざんの発生を通知する。一方、改ざんが確認されなかった場合は、アクセス元 URL を良性 URL としてデータベース 1 6 に保存する。

[0065] 最後に、改ざん通知部 2 0 は、改ざんが検知されたとき、アクセス元 URL に対応するコンテンツの管理者に改ざんの発生を通知する (S 2 7)。また、改ざんが検知されたコンテンツに対応するアクセス元 URL を、データベース 1 6 に悪性 URL として保存してもよい。すでに良性 URL にカテゴリされている URL については、悪性 URL にカテゴリを変更する。

[0066] このように、第 2 実施形態に係る情報処理装置 1 0 は、情報処理装置 1 0 からアクセス元 URL にリファラ無しでコンテンツをリクエストして、接続先の検証を実行することで、Web サーバ 1 2 の改ざんを早期かつ正確に検知できる。管理者は、悪意の第三者により、Web サイトへのアクセス方法に応じて意図的に接続先 URL が変更されるような巧妙な改ざん行為を早期に気付くことができ、上述のドライブバイダウンロード攻撃などの悪意の第三者による巧妙な攻撃を短期間で終息させることができる。

[0067] (第 3 実施形態)

図 8 は、第 3 実施形態に係る情報処理装置 1 0 の構成の一例を示す図である。なお、図 8 において第 1 実施形態 (図 1) と共通の構成又は機能を有する部分は、同一符号で示し、重複する説明を省略する。

- [0068] 第3実施形態に係る情報処理装置10では、データベース16に登録されていないURLがアクセスログにおいて抽出された場合に、この未分類のURLを解析してカテゴリ分類を実行し、カテゴリ分類の結果に基づいてコンテンツの改ざんを検知する。
- [0069] ログ抽出部15は、収集されたアクセスログから、データベース16に登録されたURLと一致しない未分類URLを検出する。
- [0070] URL解析部22は、未分類URLを悪性URLか否かを解析して、未分類URLを良性URL、または、悪性URLのいずれかに登録する。
- [0071] 未分類URLを解析する方法として、例えば未分類のURLにアクセスして、取得した実行ファイルを実行できる仮想環境において、当該ファイルを実行した結果などに基づいて未分類URLを悪性URLか否かを解析する。より具体的には、悪性ファイルに属するハッシュを予め保存しておき、このハッシュ値と比較することで取得したファイルが悪性か否かを判定して、未分類URLが悪性URLか否かを解析する。また、一般的に使用されるアンチウィルスソフトを使用して、未分類のURLにアクセスして、取得した実行ファイルがマルウェアか否かを判定することで、未分類URLを悪性URLか否かを解析してもよい。
- [0072] また、無害となる良性URLのドメインを予め取得しておき、そのドメインを含む未分類URLについては良性として、ドメインを含まないURLについては悪性URLと分類してもよい。なお、良性または悪性に登録された未分類URLについて、情報処理装置10の管理者によりカテゴリを変更可能な構成としてもよい。
- [0073] 改ざん検知部19は、アクセス元URLが良性URLに一致する一方、接続先URLが悪性URLに一致して、URLのカテゴリが良性から悪性に遷移する場合に改ざんを検知する。コンテンツの改ざんが発生していなければ、URLが良性URLから悪性URLへのカテゴリ遷移は起こり得ないため、良性から悪性へのカテゴリ遷移は、コンテンツの改ざんと判定することができる。

[0074] 続いて、第3実施形態に係る情報処理装置10の動作について説明する。

図9は、第3実施形態に係る情報処理方法の一例を示すフローチャートである（適宜、図8参照）。

[0075] アクセスログ収集部14は、クライアント端末11から、インターネットへのアクセスログを取得する（S30）。

[0076] ログ抽出部15は、クライアント端末11から取得したアクセスログの中から、ユーザがリクエストするコンテンツに対応するアクセス元URLと、実際にクライアント端末11に接続された接続先URLとが一致しないログを抽出する（S31）。

[0077] ログ抽出部15は、接続先URLをデータベース16に照合して、接続先URLがデータベース内で登録されていない未分類URLを抽出する（S32）。

[0078] URL解析部22は、抽出された未分類URLが悪性URLか否かを解析して、この未分類URLを良性または悪性のいずれかのカテゴリに分類する（S33）。

[0079] 改ざん検知部19は、アクセス元URL及び接続先URLがデータベース内でいずれのカテゴリに属するかを判定する。改ざん検知部19は、アクセス元URLが良性カテゴリに分類される一方、接続先URLが悪性カテゴリに分類されて、アクセス元URLから接続先URLのカテゴリが良性から悪性に遷移する場合には、アクセス元URLに対応するコンテンツの改ざん検知する（S34：YES，S35）。

[0080] 一方、アクセス元URLから接続先URLのカテゴリが良性のままで変化しない場合には、改ざんは発生していないと判定する（S34：NO，終了）。

[0081] 最後に、改ざん通知部20は、改ざんが検知されたとき、アクセス元URLに対応するコンテンツの管理者に改ざんの発生を通知する（S37）。また、改ざんが検知されたコンテンツに対応するアクセス元URLを、データベース16に悪性URLとして保存してもよい。すでに良性URLにカテゴ

リされているURLについては、悪性URLにカテゴリを変更する。

[0082] 第3実施形態に係る情報処理装置10は、未分類URLについて悪性URLか否かのカテゴリ分類を行って、この分類結果を用いて改ざんを検知する。悪性URLは常に増加して、その発見には時間を要し、網羅的に悪性サイトを即時に検知することは困難となるが、データベース16に登録されていない未分類URLについて悪性URLか否かのカテゴリ分類を実行することで、Webサーバ12の改ざんを早期に検知できる。

[0083] 以上述べた各実施形態の情報処理装置によれば、ユーザのクライアント端末11から収集したアクセスログを、悪性URLを予め登録したデータベースと照合することで、Webサーバの改ざんを早期に検知でき、Webサーバ12の管理者にコンテンツの改ざんを早期に通知することができる。これにより、管理者は、Webサーバ12の改ざんを早期に気付くことができ、水飲み場型攻撃などの悪意の第三者による攻撃を短期間で終息させることができる。

[0084] なお、情報処理装置10で実行されるプログラムは、ROM等の記憶回路に予め組み込んで提供される。もしくは、このプログラムは、インストール可能な形式または実行可能な形式のファイルでCD-ROM、CD-R、メモ리카ード、DVD、フレキシブルディスク等のコンピュータで読み取り可能な記憶媒体に記憶されて提供するようにしてもよい。また、情報処理装置10で実行されるプログラムは、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせて提供するようにしてもよい。

[0085] 本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。

符号の説明

[0086] 10…情報処理装置、11…クライアント端末、12…Webサーバ、13…管理者端末、14…アクセスログ収集部、15…ログ抽出部、16…データベース、17…悪性URL保存部、18…良性URL保存部、19…改ざん検知部、20…改ざん通知部、21…アクセス検証部、22…URL解析部、50…悪意の第三者のWebサーバ。

請求の範囲

- [請求項1] クライアント端末がW e bサーバに対してコンテンツをリクエストした際のアクセスログを収集するアクセスログ収集部と、
悪性U R Lを予め登録するデータベースと、
前記クライアント端末の接続先に対応する接続先U R Lを前記データベースと照合して、前記接続先U R Lが前記悪性U R Lと一致する場合に、前記コンテンツの改ざんを検知する改ざん検知部と、
を備えることを特徴とする情報処理装置。
- [請求項2] 前記アクセスログ収集部で収集された前記アクセスログから、前記コンテンツのリクエスト先に対応するアクセス元U R Lと前記接続先U R Lとが一致しないアクセスログを抽出するログ抽出部を備えることを特徴とする請求項1に記載の情報処理装置。
- [請求項3] 前記改ざん検知された場合に、改ざんが検知された前記コンテンツに対応する前記W e bサーバの管理者に前記改ざんの発生を通知する改ざん通知部を備えることを特徴とする請求項1に記載の情報処理装置。
- [請求項4] 前記接続先U R Lが前記悪性U R Lと一致する場合に、前記W e bサーバに対して、前記アクセスログから抽出されたアクセス元U R Lを直接リクエストして、前記アクセス元U R Lと前記直接リクエスト時における接続先U R Lとを比較するアクセス検証部を備えて、
前記改ざん検知部は、前記アクセス検証部の比較により前記アクセス元U R Lと前記直接リクエスト時における接続先U R Lとが一致するときに、前記コンテンツの改ざんを検知する、
ことを特徴とする請求項1に記載の情報処理装置。
- [請求項5] 収集された前記アクセスログから、前記データベースに登録されたU R Lと一致しない未分類U R Lを抽出し、
前記データベースは、前記悪性U R Lとともに良性U R Lを予め登録し、

前記未分類URLについて悪性URLか否かを解析して、前記未分類URLを前記良性URL、または、前記悪性URLのいずれかに登録するURL解析部を備える、

ことを特徴とする請求項1に記載の情報処理装置。

[請求項6]

前記データベースは、前記悪性URLとともに良性URLを予め登録し、

改ざん検知部は、アクセス元URLが前記良性URLに一致する一方、前記接続先URLが前記悪性URLに一致して、URLのカテゴリが良性から悪性に遷移する場合に前記コンテンツの改ざんを検知する、

ことを特徴とする請求項1に記載の情報処理装置。

[請求項7]

クライアント端末がWebサーバに対してコンテンツをリクエストした際のアクセスログを収集するアクセスログ収集部と、

悪性のIPアドレスを予め登録するデータベースと、

前記クライアント端末の接続先に対応する接続先URLを前記データベースと照合して、前記接続先URLに対応するIPアドレスが前記悪性のIPアドレスと一致する場合に、前記コンテンツの改ざんを検知する改ざん検知部と、

を備えることを特徴とする情報処理装置。

[請求項8]

クライアント端末がWebサーバに対してコンテンツをリクエストした際のアクセスログを収集するステップと、

悪性URLをデータベースに予め登録するステップと、

前記クライアント端末の接続先に対応する接続先URLを前記データベースと照合して、前記接続先URLが前記悪性URLと一致する場合に、前記コンテンツの改ざんを検知するステップと、

を含むことを特徴とする情報処理方法。

[請求項9]

コンピュータを、

クライアント端末がWebサーバに対してコンテンツをリクエスト

した際のアクセスログを収集する機能と、
悪性URLをデータベースに予め登録する機能と、
前記クライアント端末の接続先に対応する接続先URLを前記データベースと照合して、前記接続先URLが前記悪性URLと一致する場合に、前記コンテンツの改ざんを検知する機能と、
として機能させることを特徴とする情報処理プログラム。