(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6623128号
(P6623128)

(45)発行日　令和1年12月18日(2019.12.18)　(24)登録日　令和1年11月29日(2019.11.29)

(51)Int.Cl.		F I	
G06F 16/28	(2019.01)	G06F	16/28
G06F 16/904	(2019.01)	G06F	16/904
G06Q 50/10	(2012.01)	G06Q	50/10

請求項の数 15　(全 18 頁)

(21)出願番号　特願2016-151528(P2016-151528)
(22)出願日　平成28年8月1日(2016.8.1)
(65)公開番号　特開2018-22248(P2018-22248A)
(43)公開日　平成30年2月8日(2018.2.8)
審査請求日　平成30年11月21日(2018.11.21)

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　110002365
特許業務法人サンネクスト国際特許事務所
(72)発明者　林　直樹
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　鬼頭　哲郎
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　仲小路　博史
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内

最終頁に続く

(54)【発明の名称】ログ分析システム、ログ分析方法及びログ分析装置

(57)【特許請求の範囲】
【請求項１】
ログ情報を分析するログ分析システムであって、
複数のイベントを含むログ情報を基に、複数の前記イベントの発生の因果関係があるか否かを判定し、
個々の１つの前記イベントを１つのノードとすること及び前記因果関係が存在した場合に当該イベントに対応する前記ノードの間にエッジを追加することで、グラフ構造を構築する関係性ツリー構築部と、
指定された任意の前記イベントについて、前記グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定し、
指定された当該イベントを発生させた木であるかの判定において、各木を構成する前記イベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する関係性ツリー不審度計算部と、
前記確からしさの度合いを加味し、前記因果関係を補った前記グラフ構造を結果として出力する結果出力部と
を備えるログ分析システム。
【請求項２】
実際に不審な動作を確認することで前記因果関係を補う不審性確認部
を備える請求項１に記載のログ分析システム。
【請求項３】

分析対象とする前記イベントを所定の条件で絞り込む周辺情報取得部
を備える請求項１に記載のログ分析システム。
【請求項４】
結果出力部が出力する前記結果は、
前記グラフ構造の前記確からしさの度合い及び分析結果の概要を表示する関係性ツリー一覧と、
前記関係性ツリー一覧から選択された前記グラフ構造の詳細が表示される分析結果描画領域と
を備える請求項１に記載のログ分析システム。
【請求項５】
前記イベントは、アクセス先のロケーション情報と遷移元のロケーション情報とを備え、
前記関係性ツリー不審度計算部は、アクセス先のロケーション情報が前記グラフ構造の途中で変化している場合に、当該グラフ構造の前記確からしさの度合いを不審だと判定する向きに変化させる
請求項１に記載のログ分析システム。
【請求項６】
前記イベントは、アクセス先のロケーション情報と遷移元のロケーション情報とを備え、
前記関係性ツリー不審度計算部は、グラフ構造中に２回以上のリダイレクト遷移が発生している場合に、当該グラフ構造の前記確からしさの度合いを不審だと判定する向きに変化させる
請求項１に記載のログ分析システム。
【請求項７】
前記イベントは、アクセスに使用したエージェント情報を備え、
前記関係性ツリー不審度計算部は、指定された前記イベントと同じエージェント情報であり指定された前記イベントと同じアクセス先であって不審な動作がない前記イベントがグラフ構造中にある場合に、当該グラフ構造の前記確からしさの度合いを不審ではないと判定する向きに変化させる
請求項１に記載のログ分析システム。
【請求項８】
ログ情報を分析するログ分析システムにおけるログ分析方法であって、
前記ログ分析システムは、関係性ツリー構築部と、関係性ツリー不審度計算部と、結果出力部とを有し、
前記関係性ツリー構築部が、複数のイベントを含むログ情報を基に、複数の前記イベントの発生の因果関係があるか否かを判定する第１のステップと、
前記関係性ツリー構築部が、個々の１つの前記イベントを１つのノードとすること及び前記因果関係が存在した場合に当該イベントに対応する前記ノードの間にエッジを追加することで、グラフ構造を構築する第２のステップと、
前記関係性ツリー不審度計算部が、指定された任意の前記イベントについて、前記グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定する第３のステップと、
前記関係性ツリー不審度計算部が、指定された当該イベントを発生させた木であるかの判定において、各木を構成する前記イベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する第４のステップと、
前記結果出力部が、前記確からしさの度合いを加味し、前記因果関係を補った前記グラフ構造を結果として出力する第５のステップと
を備えるログ分析方法。
【請求項９】
前記ログ分析システムは不審性確認部を有し、

不審性確認部が実際に不審な動作を確認することで前記因果関係を補う第６のステップ
を備える請求項８に記載のログ分析方法。
【請求項１０】
前記ログ分析システムは周辺情報取得部を有し、
前記周辺情報取得部が分析対象とする前記イベントを所定の条件で絞り込む第６のステップ
を備える請求項８に記載のログ分析方法。
【請求項１１】
前記結果出力部が出力する前記結果は関係性ツリー情報一覧と、分析結果描画領域とを有し、
前記結果出力部が、関係性ツリー一覧に前記グラフ構造の前記確からしさの度合い及び分析結果の概要を表示する第６のステップと、
前記結果出力部が、分析結果描画領域に前記関係性ツリー一覧から選択された前記グラフ構造の詳細を表示する第７のステップと
を備える請求項８に記載のログ分析方法。
【請求項１２】
前記イベントは、アクセス先のロケーション情報と遷移元のロケーション情報とを有し、
前記関係性ツリー不審度計算部がアクセス先のロケーション情報が前記グラフ構造の途中で変化している場合に、当該グラフ構造の前記確からしさの度合いを不審だと判定する向きに変化させる第６のステップを備える
請求項８に記載のログ分析方法。
【請求項１３】
前記イベントは、アクセス先のロケーション情報と遷移元のロケーション情報とを有し、
前記関係性ツリー不審度計算部は、グラフ構造中に２回以上のリダイレクト遷移が発生している場合に、当該グラフ構造の前記確からしさの度合いを不審だと判定する向きに変化させる第６のステップを備える
請求項８に記載のログ分析方法。
【請求項１４】
前記イベントは、アクセスに使用したエージェント情報を有し、
前記関係性ツリー不審度計算部は、指定された前記イベントと同じエージェント情報であり指定された前記イベントと同じアクセス先であって不審な動作がない前記イベントがグラフ構造中にある場合に、当該グラフ構造の前記確からしさの度合いを不審ではないと判定する向きに変化させる第６のステップを備える
請求項８に記載のログ分析方法。
【請求項１５】
ログ情報を分析するログ分析装置であって、
複数のイベントを含むログ情報を基に、複数の前記イベントの発生の因果関係があるか否かを判定し、
個々の１つの前記イベントを１つのノードとすること及び前記因果関係が存在した場合に当該イベントに対応する前記ノードの間にエッジを追加することで、グラフ構造を構築する関係性ツリー構築部と、
指定された任意の前記イベントについて、前記グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定し、
指定された当該イベントを発生させた木であるかの判定において、各木を構成する前記イベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する関係性ツリー不審度計算部と、
前記確からしさの度合いを加味し、前記因果関係を補った前記グラフ構造を結果として出力する結果出力部と

を備えるログ分析装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ログ分析システム、ログ分析方法及びログ分析装置に関し、例えば類似又は同一のアラート発生の再発を防止するログ分析システムに適用して好適なものである。
【背景技術】
【０００２】
組織のシステムを守るためには、外部からの攻撃等によってシステムに異常が発生した際、当該異常を速やかに発見して復旧を図るだけではなく、当該異常の原因を調査して対策を講じることで、同じような異常の再発を防止することが重要である。異常の発見及び異常の原因の調査のため、システムは記録しているログを分析する。
【０００３】
近年被害が増加している攻撃としては、例えば攻撃者が、標的が普段利用するようなＷｅｂサイトを改ざんし、閲覧した者の計算機に異常を発生させるものがある。この攻撃では多くの場合、改ざんされたＷｅｂサイトに標的がアクセスすると、標的が意図しない複数回のリダイレクトが発生し、最終的に攻撃者が悪用しているマルウェアの配布サイトまで誘導される。閲覧した者が、誘導されたマルウェアの配布サイトでマルウェアを取得し、閲覧した者の計算機にマルウェアをインストールすると、閲覧した者の計算機から組織のシステム内にマルウェアによる攻撃が行われてしまう。
【０００４】
このような攻撃では、攻撃者が複数のマルウェアの配布サイトを準備し、そのうちのいずれかにランダムに誘導することもある。従って、そのような攻撃による異常を検知した場合、個々のマルウェアの配布サイトへのアクセスをブロックするだけではなく、改ざんされたＷｅｂサイトを発見することで、当該Ｗｅｂサイトへのアクセスを一時的に遮断する又は当該Ｗｅｂサイトの管理者に改ざんの旨を通知することで復旧を促す、などの再発防止策を講じることが重要である。
【０００５】
すなわち、例えば、マルウェアをダウンロードした旨のアラートが発生した場合には、当該アクセスがどのような経緯で発生したのかについて、通信ログなどを基に調査し、入り口となったＷｅｂサイトを特定する必要がある。
【０００６】
一方で、近年の組織のシステムは多種多様及び多数の機器やアプリケーションで構成されているため、ログに記載された事象間の因果関係を辿り異常の発生原因を特定するには、大量のシステムログやセキュリティログを調査する必要があり、長時間を要するようになっている。
【０００７】
このような情報間の関係性の分析を効率化するログ分析システムの１つとして、特許文献１には、情報をグラフ構造として保持し、検索条件の基点となったノードに指定した初期値を設定して一定比率で減少させながら伝播させ、最終的に閾値以上となったノード集合を検索結果として出力することで、既存の検索においては容易にたどり着かない情報を効率的に取得する発明が開示されている。
【先行技術文献】
【特許文献】
【０００８】
【特許文献１】特開２０１０－１９１９０２号公報
【発明の概要】
【発明が解決しようとする課題】
【０００９】
特許文献１によれば、情報間の関係性をグラフ構造として保持してクラスタリングする

ことできるため、検索対象情報と関係がある情報を推移的に抽出することができる。従って当該技術を用いることで、ログに記載されたイベント間の関係性を基に構成できるグラフ構造を辿ることができ、発生原因を効率的に調べることができる。例えば、通常のＨＴＴＰ（Hyper Text Transfer Protocol）の遷移であれば、プロキシサーバ等の通信ログに、通信の遷移元情報であるＨＴＴＰヘッダのＲＥＦＥＲＥＲ（以下、リファラと呼ぶ）を記録させることが可能である。

【００１０】

しかしながら、ログに記載される情報には、イベント間の関係性が部分的に欠落することも多い。例えば、当該通信のプロトコルが途中で変化したり又はブラウザの脆弱性を突いた遷移であったりする場合には、通信ヘッダにリファラが含まれないため遷移元に関する情報をログとして残せず、従って、グラフ構造が途中で途切れてしまうことになる。

【課題を解決するための手段】

【００１１】

かかる課題を解決するため本発明においては、ログ分析システムは、ログ情報を分析するログ分析システムであって、複数のイベントを含むログ情報を基に、複数のイベントの発生の因果関係があるか否かを判定し、個々の１つのイベントを１つのノードとすること及び因果関係が存在した場合に当該イベントに対応するノード間にエッジを追加することで、グラフ構造を構築する関係性ツリー構築部と、指定された任意のイベントについて、グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定し、指定された当該イベントを発生させた木であるかの判定において、各木を構成するイベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する関係性ツリー不審度計算部と、確からしさの度合いを加味し、因果関係を補ったグラフ構造を結果として出力する結果出力部とを備えるようにした。

【００１２】

また、本発明においては、ログ分析方法は、ログ情報を分析するログ分析システムにおけるログ分析方法であって、ログ分析システムは、関係性ツリー構築部と、関係性ツリー不審度計算部と、結果出力部とを有し、関係性ツリー構築部が、複数のイベントを含むログ情報を基に、複数のイベントの発生の因果関係があるか否かを判定する第１のステップと、関係性ツリー構築部が個々の１つのイベントを１つのノードとすること及び因果関係が存在した場合に当該イベントに対応するノード間にエッジを追加することで、グラフ構造を構築する第２のステップと、関係性ツリー不審度計算部が、指定された任意のイベントについて、グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定する第３のステップと、指定された当該イベントを発生させた木であるかの判定において、各木を構成するイベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する第４のステップとを備えるようにした。

【００１３】

また、本発明においては、ログ分析装置は、ログ情報を分析するログ分析装置であって、複数のイベントを含むログ情報を基に、複数のイベントの発生の因果関係があるか否かを判定し、個々の１つのイベントを１つのノードとすること及び因果関係が存在した場合に当該イベントに対応するノード間にエッジを追加することで、グラフ構造を構築する関係性ツリー構築部と、指定された任意のイベントについて、グラフ構造で規定される木が指定された当該イベントを発生させた木であるかを判定し、指定された当該イベントを発生させた木であるかの判定において、各木を構成するイベントに関する判定条件と、木の構造的な特徴に関する判定条件とを基に確からしさの度合いを計算する関係性ツリー不審度計算部と、確からしさの度合いを加味し、因果関係を補ったグラフ構造を結果として出力する結果出力部とを備えるようにした。

【発明の効果】

【００１４】

本発明によれば、イベント間の関係性が部分的に欠落している場合に対応したログ分析システム、ログ分析方法及びログ分析装置を実現できる。

【図面の簡単な説明】
【００１５】
【図１】本発明の実施の形態によるシステム構成を示す図である。
【図２】本発明の実施の形態による分析対象であるログを示す図である。
【図３】本発明の実施の形態による分析処理の処理手順を示すフローチャートである。
【図４】本発明の実施の形態による周辺イベントの関係性ツリー構築処理の処理結果である。
【図５】本発明の実施の形態による不審度計算処理の処理手順を示すフローチャートである。
【図６】本発明の実施の形態による構造分析ルールである。
【図７】本発明の実施の形態による履歴照合ルールである。
【図８】本発明の実施の形態による不審度計算処理の処理結果である。
【図９】本発明の実施の形態による不審性確認処理の処理手順を示すフローチャートである。
【図１０】本発明の実施の形態による不審性確認処理の処理結果である。
【図１１】本発明の実施の形態による関係性ツリー保存処理の処理手順を示すフローチャートである。
【図１２】本発明の実施の形態による関係性ツリー保存処理の処理結果である。
【発明を実施するための形態】
【００１６】
以下図面について、本発明の一実施の形態を詳述する。
【００１７】
（１）本実施の形態によるログ分析システムの構成
図１において、１は全体として本実施の形態によるログ分析システムを示す。ログ分析システム１は、ログ分析装置１００にネットワーク１４１を介して分析対象であるログ情報２００を入力するプロキシサーバ等が接続されており、ログ分析装置１００には記憶媒体１７１を接続することができる。記憶媒体１７１は、ＩＣカードやＵＳＢメモリ等であり、可搬性を有する。
【００１８】
プロキシサーバ及びログ分析装置１００は、ＣＰＵ（Central Processing Unit）及びメモリ等の情報資源を備えたコンピュータ装置であり、例えば、オープン系のサーバや、メインフレームコンピュータなどから構成される。ログ分析装置１００は、メモリ１１０、外部記憶装置１２０、ＣＰＵ１３０、通信装置１４０、入力装置１５０、出力装置１６０及び読取装置１７０がインタフェース１８０を介して接続されている。
【００１９】
外部記憶装置１２０は、ＨＤＤ（Hard Disk Drive）などであり、ログ分析機能によってログ分析を行う際に、参照するテーブルである関係性ツリー履歴情報１２１、構造分析ルールテーブル１２２及び履歴照合ルールテーブル１２３を備える。なお、関係性ツリー履歴情報１２１、構造分析ルールテーブル１２２及び履歴照合ルールテーブル１２３はメモリ１１０に備えられていてもよい。
【００２０】
メモリ１１０は、分析要求入力部１１１、関係性ツリー構築部１１２、過去履歴管理部113、関係性ツリー不審度計算部１１４、周辺情報取得部１１５、確認用情報取得部１１６、不審性確認部１１７及び結果出力部１１８を備える。これらは、ログ分析機能を実現するために必要なプログラム等であって、メモリ１１０に直接ロードされてもよいし、外部記憶装置１２０や記憶媒体１７１に保存されていてもよく、ＣＰＵ１３０によって実行される。プログラムをＣＰＵ１３０に実行させるための媒体は外部記憶装置１２０や記憶媒体１７１に限らず、利用可能な媒体であればよく、ネットワーク１４１を伝搬する搬送波やディジタル信号でもよい。
【００２１】

通信装置１４０は、インターネットやＬＡＮ等のネットワーク１４１を介して、他の装置と通信を行う。入力装置１５０は、キーボードやマウス等であり、出力装置１６０は、モニタやプリンタ等である。読取装置１７０は、記憶媒体１７１を接続し、その内容を読み取る。

【００２２】

分析要求入力部１１１は本ログ分析装置１００を利用する利用者が入力装置１５０によって入力する分析要求を受け付ける。利用者は、分析対象とするイベント等を入力する。利用者はこの際にログ情報２００をログ分析装置１００へ送付してもよい。ログ情報２００は予め、定期的に送付されていてもよい。ログ情報２００は、ネットワーク１４１を介してログ分析装置１００に入力されてもよいし、記憶媒体１７１を介してログ分析装置１００に入力されてもよい。

【００２３】

分析要求入力部１１１が分析要求を受けると、ログ情報２００を分析するログ分析処理がログ分析装置１００によって開始される。なお不審な動作などによりセキュリティアラート（以下、アラートと呼ぶ）が発生した場合に、当該アラートの原因となったイベントをシステムが機械的に分析対象としてもよい。

【００２４】

ログ分析処理として、関係性ツリー構築部１１２は、ログ情報２００のイベント間の関係性を分析する。関係性とは、ログ情報２００に記載されたあるイベントが、同じくログ情報２００に記載された別のイベントを引き起こす契機になっているという関係を指す。また、関係性ツリー情報とは、個々のイベントをノード、関係性をエッジとして構築したグラフ構造（以下、ツリーと呼ぶ）のことである。ログ情報２００を基に関係性ツリー情報を構築すると、全てのイベントが相互に関係性を有すとは限らないため、互いに連結していない（関係性を有してない）複数の関係性ツリー情報が構築される。

【００２５】

過去履歴管理部１１３は、関係性ツリー構築部１１２で構築された関係性ツリー情報を関係性ツリー履歴情報１２１へ格納し、保管する。このことで、関係性ツリー構築部１１２で構築された関係性ツリー情報をのちに使うことができる。

【００２６】

関係性ツリー不審度計算部１１４は、関係性ツリー構築部１１２が構築した関係性ツリーの、分析対象であるイベントの原因である確からしさ（以下「不審度」と呼ぶ）を計算する。なお、関係性ツリー構築部１１２が構築した関係性ツリー情報の量が多い場合は、関係性ツリー不審度計算部１１４が計算対象とする関係性ツリー情報を絞り込んでもよい。

【００２７】

例えば、分析対象であるイベントの周辺イベントに関してのみ計算対象としてもよい。周辺イベントとは、分析対象とするイベントの周辺のイベントであり、例えばログ分析の対象とするイベントとユーザＩＤが同一であるイベントを指す。また、ログ分析の対象とするイベントと処理開始時間が近い（１秒以内など）イベントのログとしてもよいし、イベント発生前の１カ月又は１年などの期間のみのイベントのログを分析対象としてもよい。

【００２８】

周辺情報取得部１１５は、ログ情報２００のイベントの内容を分析するための周辺情報（以下、周辺イベントと呼ぶ）を特定するための付加的な情報（ユーザＩＤや処理開始時間）をログ情報２００から取得する。周辺情報取得部１１５は本実施の形態にとって必須ではないが、周辺情報取得部１１５を用いることで、関係のあるイベントのみを分析する（分析対象とするイベントの絞り込みを行う）ため、より詳細で迅速な分析が可能となる。特にログ情報２００の量が膨大な場合には、分析処理コスト低減のために、周辺情報取得部１１５での処理を行うことが望ましい。

【００２９】

また周辺情報取得部１１５は、インターネットなどから脅威情報やＤＮＳ（Domain Name System）登録情報などを取得し、関係性ツリー不審度計算部１１４が不審度を計算する際にそれらの情報を関係性ツリー不審度計算部１１４に提供してもよい。このためログ分析装置１００はインターネットに接続されていることが好ましい。

【００３０】

確認用情報取得部１１６は、不審性確認部１１７が行う不審性確認処理において用いるＷｅｂコンテンツの内容などの情報をインターネットなどから取得する。確認用情報取得部１１６は本実施の形態にとって必須ではないが、不審性確認処理を行う場合には必要である。このためログ分析装置１００はインターネットに接続されていることが好ましい。

【００３１】

不審性確認部１１７は、不審性確認処理を行う。不審性確認部１１７は、本実施の形態にとって必須の機能ではないが、不審性確認部１１７を用いて、ログ情報２００の不審な箇所の通りに動作環境をそろえて実動作を確認することで、利用者により正確な情報を提供することが可能となる。

【００３２】

結果出力部１１８は、ログ分析システム１の利用者に対して、分析要求入力部１１１を介して要求された情報の分析結果を、出力装置１６０を介して出力する。

【００３３】

関係性ツリー履歴情報１２１は、これまでに既に分析した関係性ツリー情報を履歴として保管しており、この関係性ツリー履歴情報１２１は不審性確認部１１７が行う不審性確認処理において使用される。なお、関係性ツリー不審度計算部１１４は、関係性ツリー履歴情報１２１を用いて、不審度を計算する。なお不審度は整数値で表し、数値が大きいほど不審であることを示すものとする。

【００３４】

構造分析ルールテーブル１２２は、関係性ツリー情報の不審度をどのように判定するかのルールを定めているテーブルであって、関係性ツリー履歴情報１２１を照合する。構造分析ルールテーブル１２２は、不審性確認部１１７が行う不審性確認処理で使用される。なお、関係性ツリー不審度計算部１１４は、構造分析ルールテーブル１２２を用いて、不審度を計算する。

【００３５】

履歴照合ルールテーブル１２３は、関係性ツリー情報の不審度をどのように判定するかのルールを定めているテーブルであって、関係性ツリー履歴情報１２１を照合する。履歴照合ルールテーブル１２３は、不審性確認部１１７が行う不審性確認処理において使用される。なお、関係性ツリー不審度計算部１１４でも、履歴照合ルールテーブル１２３を用いて、不審度を計算する。

【００３６】

（２）ログ分析機能

本実施の形態のログ分析機能は、入力されるログ情報２００を、ルールテーブルを参照して各種処理を行うことで分析し、分析結果を出力する。

【００３７】

（２－１）ログ

本実施の形態で分析されるログ情報２００は、図２に示すような、プロキシサーバが出力する典型的な構造となっており、各イベントのログが格納されている。各イベントのログは、ログ番号欄２０１、アクセス元ＩＰアドレス欄２０２、アクセス先ＵＲＬ欄２０３、ユーザＩＤ欄２０４、リファラ欄２０５、ユーザエージェント欄２０６及びレスポンスのバイト数欄２０７を含む。

【００３８】

ログ番号欄２０１には、各イベントのログを区別するために付けるログ番号が格納されており、「ｌｏｇ１」、「ｌｏｇ２」、……と、時系列順に数字が割り付けられている。アクセス元ＩＰアドレス欄２０２にはアクセス元のＩＰアドレスが格納され、例えば、「

１．１．１．１」といったＩＰｖ４プロトコルのＩＰアドレスを格納する。アクセス先ＵＲＬ２０３欄には、アクセス先ＵＲＬが格納されており、例えば、「ａ．ｃｏｍ／１．ｈｔｍｌ」といった形式で格納される。

【００３９】

ユーザＩＤ欄２０４には、ＨＴＴＰの認証ヘッダなどに含まれる認証情報のユーザＩＤが格納されており、例えば、「ａａａａａ」といった形式で格納される。リファラ欄２０５には、各イベントの遷移元を示す、ＨＴＴＰヘッダの１つであるリファラが格納されており、アクセス先ＵＲＬ欄２０３と同様の形式で格納されている。なお、リファラが「－」となっている場合は、ＨＴＴＰヘッダにリファラが含まれていないことを示す。

【００４０】

ユーザエージェント欄２０６には、各イベントがどのアプリケーションの、どのバージョンで実行されたかを示すユーザエージェントが格納しており、「ａｐｐ＿Ａ；ｖｅｒ＿１」といった形式で格納される。レスポンスのバイト数欄２０７には、各イベントによってサーバから何バイトのレスポンスが返ってきたかが格納されており、「３００」といった形式で格納されている。

【００４１】

関係性ツリー情報を構築する際に、アクセス先ＵＲＬ及びリファラを関係性として用いる。つまり、アクセス先ＵＲＬとリファラとの値が一致するイベントにおいては、前者のイベントが後者のイベントの発生原因（遷移元）であるとする。

【００４２】

例えば、ｌｏｇ５のリファラは「ａ．ｃｏｍ／０２／２．ｈｔｍｌ」となっており、ｌｏｇ３のアクセス先ＵＲＬである「ａ．ｃｏｍ／０２／２．ｈｔｍｌ」と一致する。このことから、ｌｏｇ５の遷移元はｌｏｇ３となる。

【００４３】

（２－２）分析処理

図３にログ情報２００を関係性に基づいて分析する分析処理の処理手順を示す。分析処理は、関係性が明確（アクセス先ＵＲＬとリファラとの値が一致するイベントなど）な関係性ツリー情報を構築し、関係性を有していないようにみえる複数の関係性ツリー情報間についても、目安となる基準（不審度）を設けて関係性（因果関係）を発見する。

【００４４】

実際上、分析要求入力部１１１は、分析対象とするイベントが入力されると、入力されたイベントを分析の対象イベントとし、調査対象イベントを決定する（ＳＰ３０１）。なお、セキュリティアラートの原因となったイベントを分析要求入力部１１１が機械的に対象イベントとしてもよい。以下の説明では、調査対象イベントを、セキュリティアラートなどのアラートの原因となったｌｏｇ６に記録されているイベントとする。

【００４５】

分析要求入力部１１１によって調査対象イベントが決定されると、周辺情報取得部１１５は、ログ情報２００の中で分析対象とするイベントの絞り込みを行い、周辺イベントを特定する（ＳＰ３０２）。以下の説明では、ユーザＩＤが「ａａａａａ」のイベントに調査対象を絞り込む。

【００４６】

周辺情報取得部１１５によって周辺イベントが特定されると、関係性ツリー構築部１１２は、絞り込んだイベント間での関係性ツリー構築をする（ＳＰ３０３）。具体的には、図４（Ｂ）の関係性ツリーｔ２においてｌｏｇ３からｌｏｇ５に矢印が引いてあるように遷移元から遷移先へエッジを追加する。

【００４７】

ユーザＩＤが「ａａａａａ」のイベントに調査対象を絞り込んで、作成した関係性ツリー情報を図４（Ａ）～（Ｄ）に示す。関係性ツリー情報には、ｔ１～ｔ４の識別名が付けられている。図４（Ａ）に示す関係性ツリーｔ１を例に説明をすると、ｌｏｇ１として記録された通信（イベント）が遷移元となって、ｌｏｇ２～ｌｏｇ４として記録された通信

が発生している。さらに、ｌｏｇ３として記録された通信が遷移元となってｌｏｇ５として記録された通信が発生している。また、図４（Ｂ）の関係性ツリーｔ２に示した通りアラート情報も関係性ツリー情報として表示される。

【００４８】

なお関係性ツリー構築部１１２が作成した関係性ツリーｔ１～ｔ４は適宜、過去履歴管理部１１３によって関係性ツリー履歴情報１２１に保存される。

【００４９】

関係性ツリー構築部１１２が関係性ツリーｔ１～ｔ４を作成すると、関係性ツリー不審度計算部１１４は関係性ツリーｔ１～ｔ４の不審度を計算する不審度計算処理を行う（ＳＰ３０４）。

【００５０】

図５に不審度計算処理の処理手順を示す。不審度計算処理は、ルールテーブルである構造分析ルールテーブル１２２及び履歴照合ルールテーブル１２３を参照して、不審な通信（イベント）を特定する。

【００５１】

実際上、関係性ツリー不審度計算部１１４は調査対象イベントがノードに含まれていない、未処理の関係性ツリー情報が有るか否かを判定する（ＳＰ５０１）。そして関係性ツリー不審度計算部１１４はこの判定で否定結果を得るとこの不審度計算処理を終了する。未処理か否かは、例えば関係性ツリー情報に処理完了フラグを設けることで判定される。

【００５２】

これに対して、関係性ツリー不審度計算部１１４は、ステップＳＰ５０１の判定で肯定結果を得ると、関係性ツリー不審度計算部１１４は、未処理の関係性ツリーのうちの１つを選択する（ＳＰ５０２）。この際、選択された関係性ツリー情報の不審度の値を初期値として０とする。

【００５３】

関係性ツリー不審度計算部１１４に選択された関係性ツリー情報に対して、関係性ツリー不審度計算部１１４は、関係性ツリー情報の構造を基に、ツリー構造分析処理を行う（ＳＰ５０３）。ツリー構造分析処理には構造分析ルールテーブル１２２を用いる。なおツリー構造分析処理は、不審な構造を発見して、不審度を判定するため、各ルールによって不審度が加算される。本実施の形態では構造から判定する場合、不審度は加算のみされているが、一般的で安全な構造を発見して減算してもよいものとする。

【００５４】

図６に構造分析ルールテーブル１２２の内容を示す。構造分析ルールテーブル１２２には、各構造分析ルールが格納されており、各ルールは、ルール番号欄６０１、判定条件欄６０２、処理欄６０３及び不審度判定欄６０４を含む。

【００５５】

ルール番号欄６０１には、各ルールを区別するために付けるルール番号が格納されており、「ｓ１」、「ｓ２」、……と、不審度判定欄６０４の不審度が高い順に数字が割り付けられている。判定条件欄６０２には、不審な構造を判定するための判定条件が格納されている。処理欄６０３には、判定を行うための処理内容が格納されている。本実施の形態では説明のために判定条件欄６０２及び処理欄６０３の内容は自然言語で記載している。不審度判定欄６０４には、処理を行い、条件を満たした場合、どの関連ツリー情報の不審度をどのように増減させるかを定義している。

【００５６】

具体的には、ルールｓ１は、分析対象の関係性ツリー情報に含まれるイベントのアクセス先が、脅威情報を公開しているサイトで不審と判定されるものがあるか否かを調査するルールである。本ルールに合致した関係性ツリー情報は、悪性な通信ログが含まれている可能性が高いため、本実施の形態では、当該関係性ツリー情報の不審度を大きく上げるよう定義している（不審度に＋５）。

【００５７】

また、ルールｓ２は、アクセス先のロケーション情報を指すドメインの国が関係性ツリー情報の途中で変わったか否かを判定するルールである。悪性サイトへのリダイレクト遷移では、遷移の途中から遷移先が国外サイトに変化する場合が多く、本実施の形態では、本ルールに合致した場合は当該関係性ツリー情報の不審度を中程度上げるよう定義している（不審度に＋２）。

【００５８】

また、ルールｓ３は、アクセス先ドメインのＤＮＳの登録日が１年未満のイベントが１つ以上含まれているか否かを判定するルールである。悪性サイトはＤＮＳに登録されてから日が浅いことが多く、本実施の形態では、本ルールに合致した場合は不審度を中程度上げるよう定義している（不審度に＋２）。

【００５９】

また、ルールｓ４は、２回以上のリダイレクト遷移が発生した通信が関係性ツリー情報に含まれているか否かを判定するルールである。リダイレクト遷移は、通常のＷｅｂページの閲覧におけるページ遷移とは異なり、リソースファイル等を取得しない。すなわち、ツリー構造の次数を調査することで、リダイレクトが発生した可能性が高いことを調べることが出来る。リダイレクトは悪性な通信だけではなく、正常な通信でも発生することがあるため、本実施の形態では、本ルールに合致した場合は不審度を小程度上げるよう定義している（不審度に＋１）。

【００６０】

例えば、本実施の形態では、ルールｓ２～ｓ４を調べることで、全てのルールが合致した場合は、たとえルールｓ１に合致していなくても、ルールｓ１に合致したのと同等に不審度が上がる（＋５）判定になる。このように複数のルールを用いることで未知の悪性サイトであっても、ツリーの構造や周辺情報を用いて不審度を詳細に分析することが可能である。

【００６１】

関係性ツリー不審度計算部１１４は、ツリー構造分析処理で不審度を計算すると、さらに履歴照合処理を行う（ＳＰ５０４）。履歴照合処理には履歴照合ルールテーブル１２３を用いる。なお履歴照合処理においては、過去に問題のなかった履歴を照合して、不審度を判定するため、各ルールによって不審度が減算される。本実施の形態では履歴から判定する場合、不審度は減算のみされているが、過去に問題のあった履歴を照合して、不審度を加算してもよいものとする。

【００６２】

図７に履歴照合ルールテーブル１２３の内容を示す。履歴照合ルールテーブル１２３には、各履歴照合ルールが格納されており、各ルールは、ルール番号欄７０１、判定条件欄７０２、処理欄７０３及び不審度判定欄７０４を含む。

【００６３】

ルール番号欄７０１には、各ルールを区別するために付けるルール番号が格納されており、「ｒ１」、「ｒ２」、……と、不審度判定欄７０４の不審度が高い順に数字が割り付けられている。判定条件欄７０２には、安全な履歴を判定するための判定条件が格納されている。処理欄７０３には、判定を行うための処理内容が格納されている。本実施の形態では説明のために判定条件欄７０２及び処理欄７０３の内容は自然言語で記載している。不審度判定欄７０４には、処理を行い、条件を満たした場合、どの関連ツリー情報の不審度をどのように増減させるかを定義している。

【００６４】

具体的には、ルールｒ１は、関係性ツリー履歴情報１２１の中の、分析対象の関連ツリー情報に含まれる通信と同じアクセス先にアクセスしているにも関わらず、アラートが発生していない事例があったか否かを判定するルールである。本ルールに合致した関係性ツリー情報は、悪性な通信とは無関係な可能性が高いため、本実施の形態では、不審度を下げるよう定義している（不審度に－１）。

【００６５】

また、ルールｒ２は、ルールｒ１に合致し、関係性ツリー履歴情報１２１の該当イベントの発生日が１日以内であるか否かを判定するルールである。攻撃者による改ざんはいつ行われるかわからないため、例えば、アラートが発生しなかった事例が１日未満に発生していれば、本実施の形態では、不審度をさらに下げるよう定義している（不審度に－１）。

【００６６】

また、ルールｒ３は、ルールｒ１に合致し、関係性ツリー履歴情報１２１の該当イベントのユーザエージェントの値が同じであった場合には、同じ条件（ブラウザ）を用いてアクセスを行っているにも関わらずアラートが発生していない事例があったことを表す。このため、本実施の形態では、不審度をさらに下げるよう定義している（不審度に－１）。これは、近年の悪性サイトは、特定のブラウザなどに狙いを定め、狙いを定めたブラウザ以外には攻撃を行わない（すなわちアラートが発生しない）場合があるためである。

【００６７】

なお、上述したルールは一例であり、これ以外のルールを組み合わせて用いることも可能である。

【００６８】

不審度計算処理の計算結果は図８（Ａ）～（Ｄ）に示すように、関係性ツリー情報の不審度と不審度の内訳として合致したルールのルール番号を関係性ツリー情報に表示してもよい。

【００６９】

具体的には、図８（Ａ）に示す関係性ツリーｔ１は、不審理由として示したように、構造分析ルールテーブル１２２のルールｓ２とルールｓ４に合致するため、不審度として＋３となる。また、図８（Ｂ）に示す関係性ツリーｔ２は、分析対象イベントを含むツリーであるため、不審度判定の対象外である。図８（Ｃ）に示す関係性ツリーｔ３及び図８（Ｄ）に示すｔ４は関係性ツリーｔ１と同様である。

【００７０】

以上の通り不審度計算処理から、関係性ツリーｔ１がもっとも不審な一連の通信（イベント）であると知ることが出来る。

【００７１】

関係性ツリー不審度計算部１１４が不審度計算処理をしたのち、不審性確認部１１７は不審性確認処理を行う（ＳＰ３０５）。

【００７２】

図９に不審性確認処理の処理手順を示す。不審性確認処理は、不審度の高い処理を実施することで実際にアラートが発生するか否かを確認する。

【００７３】

実際上、不審性確認部１１７は調査対象イベントがノードに含まれていない、未処理の関係性ツリー情報が有るか否かを判定する（ＳＰ９０１）。そして不審性確認部１１７はこの判定で否定結果を得るとこの不審性確認処理を終了する。未処理か否かは、例えば関係性ツリー情報に処理完了フラグを設けることで判定する。

【００７４】

これに対して、不審性確認部１１７は、ステップＳＰ９０１の判定で肯定結果を得ると、不審性確認部１１７は、不審度が高い未処理の関係性ツリー情報のうちの１つを選択する（ＳＰ９０２）。この際、不審度が高い関係性ツリー情報を優先的に選択すると実際にアラートを発生させるまでの時間を短縮できる。

【００７５】

不審性確認部１１７は、不審度が高い未処理の関係性ツリー情報を選択すると、不審性確認部１１７は、関係性ツリー情報のルートノード（基点となっているイベント）を取得する（ＳＰ９０３）。不審性確認部１１７は、取得したルートノードから当該イベントのアクセス先であるアクセス先ＵＲＬ含むリクエストヘッダを取得する。

【００７６】

当該イベントのアクセス先であるアクセス先ＵＲＬを含むリクエストヘッダを取得すると、不審性確認部１１７は、確認用情報取得部１１６を用いて、当該アクセス先ＵＲＬへリクエストヘッダを当該イベントのログと同じにしてアクセスし（ＳＰ９０４）、不審性確認部１１７はステップＳＰ９０１へ戻る。

【００７７】

不審性確認部１１７は、アラートを発生させたログと同じセキュリティ機構を用いてアクセスによって、再度アラートが発生するか否かを判定する（ＳＰ９０５）。そして不審性確認部１１７はこの判定で否定結果を得ると、不審性確認部１１７は当該イベントのログと実動作との関係性ツリー情報の構造やレスポンスのバイト数が異なっていることを確認し、当該関係性ツリー情報の不審度を１加算し（ＳＰ９０６）、不審性確認部１１７はステップＳＰ９０１へ戻る。

【００７８】

これに対して、不審性確認部１１７は、ステップＳＰ９０５の判定で肯定結果を得ると、当該関係性ツリー情報を、アラートを発生させたイベントを含む関係性ツリー情報と紐づけ（ＳＰ９０７）、不審性確認部１１７はステップＳＰ９０１へ戻る。このことで、アラートを発生させたログの原因となったイベントの一群のイベントである関係性ツリー情報を発見したこととする。

【００７９】

なお、アラートが発生した時点で、イベントの一群のイベントである関係性ツリー情報を発見し、因果関係を発見したこととなるので、アラートが発生した際に不審性確認部１１７は不審性確認処理を終了してもよい。なお、実際にアラートを発生させずに、不審度が高いものがアラートの発生原因だと推定するのみにとどまってもよい。

【００８０】

不審性確認部１１７による不審性確認処理が終了すると、結果出力部１１８が分析処理の結果を出力する（ＳＰ３０６）。

【００８１】

図１０に結果出力部１１８が出力する出力結果表示を示す。出力結果表示は分析対象のイベントのログのログ番号１００１、各関係性ツリー情報の不審度及び分析結果の概要を表示する関係性ツリー一覧１００２及び不審度計算処理結果にさらに不審性確認処理での確認結果を反映させた分析結果描画領域１００３を含む。

【００８２】

分析結果描画領域１００３は、利用者が関係性ツリー一覧１００２から選択し、当該選択した関係性情報ツリーの分析結果の詳細を閲覧するための描画領域である。本実施の形態では、欠落していた因果関係を補うことで、実際には因果関係のあった２つの関係性ツリーｔ１及びｔ２を連結させて描画する。このことで、当該分析の経過を知ることが出来る。本出力を閲覧することで、利用者は、分析対象のアラート等の異常イベントが他のどのイベントが原因となって発生したのか容易に知ることが出来る。

【００８３】

（２－３）関係性ツリー保存処理

不審度計算処理で使用する関係性ツリー情報については、不審度計算処理の前に適宜関係性ツリー履歴情報１２１に格納するが、これとは別に日ごとにバッチ処理として別のタイミングで実施してもよい。この場合、対象とするイベントの絞り込みは行わない。

【００８４】

図１１に関係性ツリー保存処理の処理手順を示す。関係性ツリー保存処理は、関係性ツリー情報を関係性ツリー履歴情報１２１に格納する。

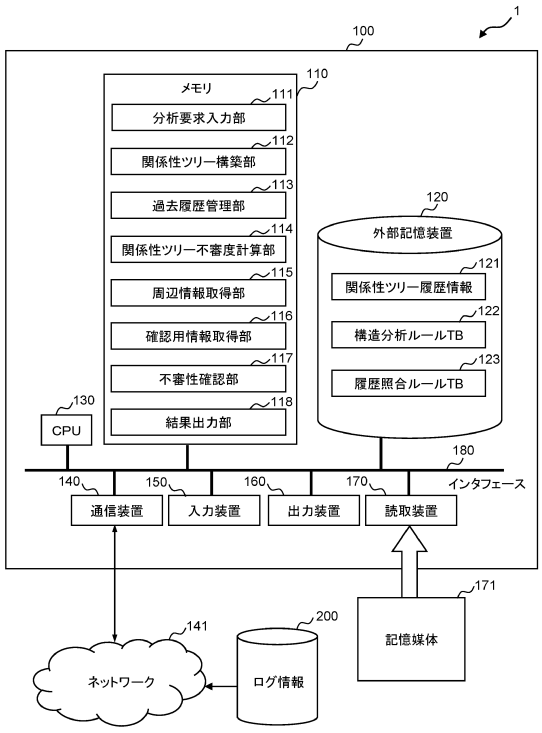
【００８５】

実際上、関係性ツリー構築部１１２は、イベント間での関係性ツリー構築をする（ＳＰ１１０１）。関係性ツリー構築部１１２が、イベント間での関係性ツリー構築をしたあと、過去履歴管理部１１３は、関係性ツリー情報を関係性ツリー履歴情報１２１に格納する（ＳＰ１１０２）。

【００８６】
関係性ツリー保存処理の結果を図１２に示す。図１２には、関係性ツリー履歴ｔｒ１が示されている。ｌｏｇ９及びｌｏｇ１０が含まれた関係性ツリー情報が表示されている。
【００８７】
（３）本実施の形態の効果
以上のように本実施の形態のログ分析システム１では、リファラから追えるひとまとまりのイベントごとの関係性ツリー情報間の関係を、ルールに基づいて不審度を設定することで推定し、実動作によりアラートを発生させることで、欠落していた因果関係を確認し、補うことができる。このことにより、分析に要する時間を短縮することができる。
【符号の説明】
【００８８】
１……ログ分析システム、１００……ログ分析装置、１１１……分析要求入力部、１１２……関係性ツリー構築部、１１３……過去履歴管理部、１１４……関係性ツリー不審度計算部、１１５……周辺情報取得部、１１６……確認用情報取得部、１１７……不審性確認部、１１８……結果出力部、１２１……関係性ツリー履歴情報、１２２……構造分析ルールテーブル、１２３……履歴照合ルールテーブル、１４１……ネットワーク、２００……ログ情報。

【図１】

図 1



【図２】

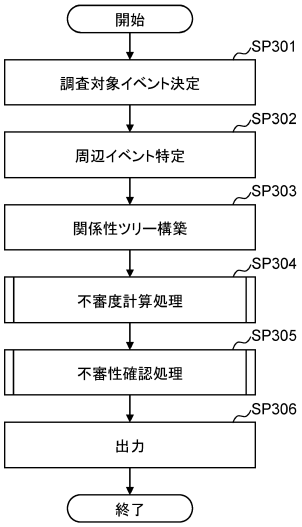
図 2

#	from	to	user id	referer	user agent	dst byte
log1	1.1.1.1	a.com/1.html	aaaaa	-	app_A; ver_1	300
log2	1.1.1.1	a.com/1.css	aaaaa	a.com/1.html	app_A; ver_1	10
log3	1.1.1.1	a.com/02/2.html	aaaaa	a.com/1.html	app_A; ver_1	200
log4	1.1.1.1	a.com/1.jpg	aaaaa	a.com/1.html	app_A; ver_1	3000
log5	1.1.1.1	a.ABCD/02/2.html	aaaaa	a.com/02/2.html	app_A; ver_1	400
log6	1.1.1.1	mal.abc/hoge.exe	aaaaa	-	app_A; ver_1	1500
log7	1.1.1.1	bbb.jp/1.html	aaaaa	-	app_A; ver_1	100
log8	1.1.1.1	bbb.jp/1.css	aaaaa	bbb.jp/1.html	app_A; ver_1	120
log9	2.2.2.2	bbb.jp/1.html	bbbbb	-	app_A; ver_1	100
log10	2.2.2.2	bbb.jp/1.html	bbbbb	bbb.jp/1.html	app_A; ver_1	100
log11	1.1.1.1	ccc.jp/1.html	aaaaa	-	app_A; ver_1	600
log12	1.1.1.1	ccc.jp/1.jpg	aaaaa	ccc.jp/1.html	app_A; ver_1	2000
lob13	1.1.1.1	ccc.jp/2.jpg	aaaaa	ccc.jp/1.html	app_A; ver_1	3000
…						

log6への遷移元（ログには明示的に情報無し）
アンチウィルスアラート

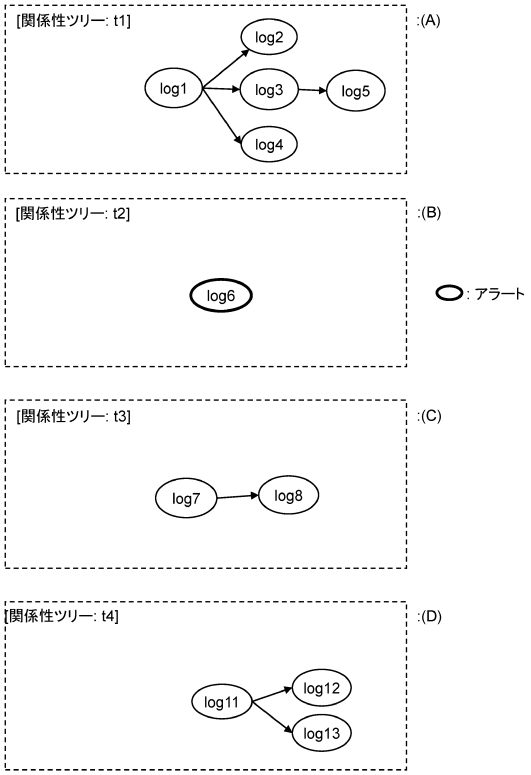
【図３】

図 3



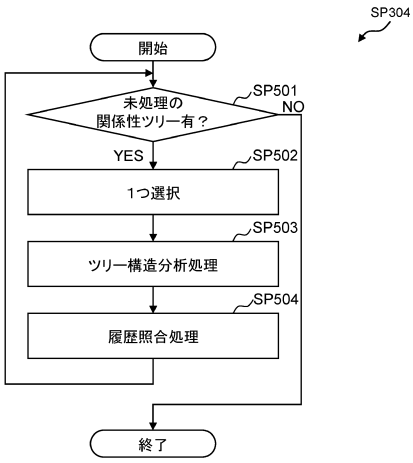
【図４】

図 4



【図５】

図 5



【図６】

図 6

	#	条件	処理	不審度判定
601	s1	脅威情報公開サイトで悪性と評価されるアクセス先が1つ以上含まれているか？	脅威情報公開サイトから評価情報を取得	当該ツリーの不審度を+5
	s2	アクセス先ドメインの国が途中から変化しているか？	ツリー内のアクセス先のドメイン文字列を比較	当該ツリーの不審度を+2
	s3	アクセス先ドメインのDNS登録日が1年未満のイベントがツリーに1つ以上含まれるか？	DNS登録日を取得し、イベントの日付と比較	当該ツリーの不審度を+2
	s4	2段以上のリダイレクト遷移をおこなった通信が含まれているか？	関係性ツリーの各ノードの出次数を計算し、1以下のノードが2つ以上連結しているか判定	当該ツリーの不審度を+1
	…			

602 603 604 122

【図７】

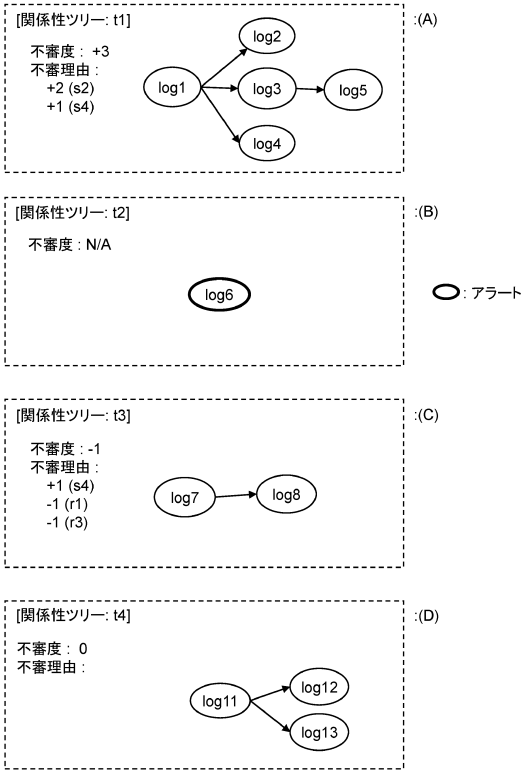
図 7

#	条件	処理	不審度判定
r1	調査対象ツリーのいずれかのノードと同じアクセス先を表すノードを保持するツリーが存在し、アラートが発生していない	ノードのアクセス先を比較	当該ツリーの不審度を-1
r2	r1を満たし、さらに、履歴ツリーの日時が1日以内である	履歴ツリー内のノードの日時と、調査対象ツリーの日時を比較	当該ツリーの不審度を-1
r3	r1を満たし、さらに、ログに記載されたユーザエージェントが同じである	履歴ツリー内のユーザエージェントと、調査対象ツリーのユーザエージェントを比較	当該ツリーの不審度を-1
…			

701 702 703 704 123

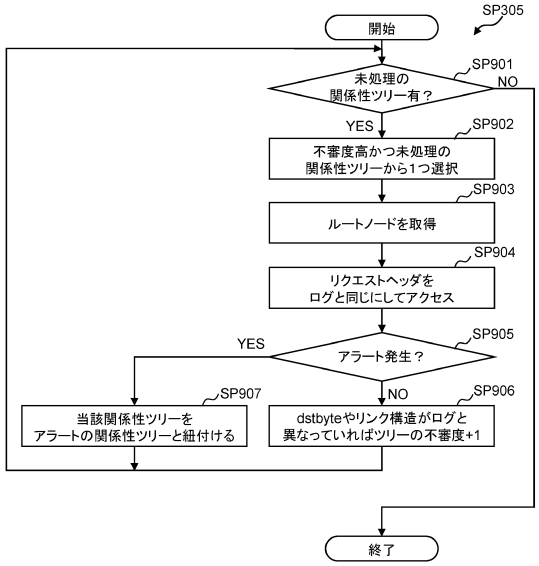
【図８】

図 8



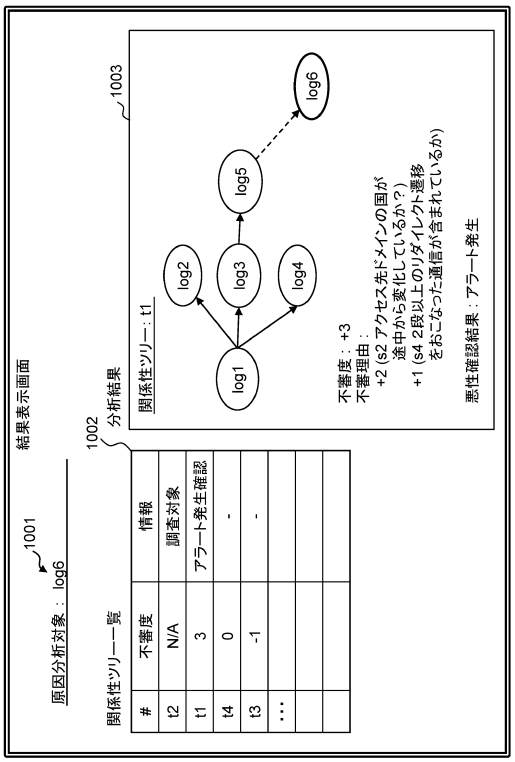
【図９】

図 9



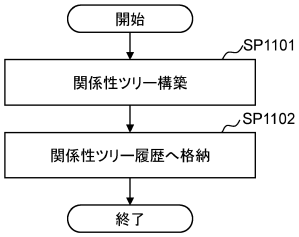
【図１０】

図 10



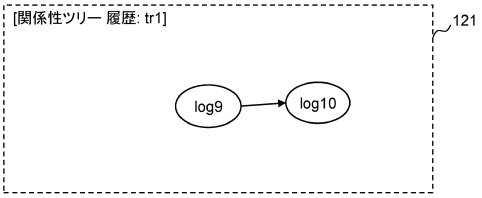
【図１１】

図 11



【図１２】

図 12



フロントページの続き

審査官　鹿野　博嗣

(56)参考文献　特開２０１６－０４５５５６（ＪＰ，Ａ）
特開２０１０－１２２８４７（ＪＰ，Ａ）
特開２０１０－０５０９３９（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１６／２８
Ｇ０６Ｆ　１６／９０４
Ｇ０６Ｑ　５０／１０