

(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2007-102657

(P2007-102657A)

(43) 公開日 平成19年4月19日(2007.4.19)

(51) Int. Cl.

G06Q 10/00 (2006.01)

F I

G06F 17/60 174

テーマコード (参考)

審査請求 未請求 請求項の数 14 O L (全 28 頁)

(21) 出願番号 特願2005-294344 (P2005-294344)
 (22) 出願日 平成17年10月7日 (2005.10.7)

(71) 出願人 000005496
 富士ゼロックス株式会社
 東京都港区赤坂九丁目7番3号
 (74) 代理人 100086531
 弁理士 澤田 俊夫
 (74) 代理人 100093241
 弁理士 宮田 正昭
 (74) 代理人 100101801
 弁理士 山田 英治
 (72) 発明者 園田 隆志
 神奈川県足柄上郡中井町境430 グリー
 ンテクなかい 富士ゼロックス株式会社内
 (72) 発明者 倉林 則之
 神奈川県足柄上郡中井町境430 グリー
 ンテクなかい 富士ゼロックス株式会社内
 最終頁に続く

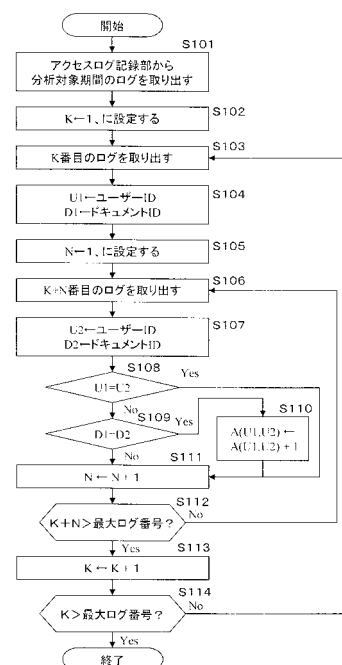
(54) 【発明の名称】 コミュニティ分析装置、およびコミュニティ分析方法、並びにコンピュータ・プログラム

(57) 【要約】

【課題】 電子文書記憶部に対するアクセスログ解析により、コミュニティ分析を行なう装置および方法を提供する。

【解決手段】 電子文書記憶部に対するアクセス情報としてユーザIDと、文書IDとをアクセス日時情報に対応付けて記録し、このアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザをコミュニティ関係を有するユーザとして選択抽出する。または、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する。例えば、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【選択図】 図5



【特許請求の範囲】**【請求項 1】**

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録するアクセスログ記憶部と、

前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析部とを有し、

前記アクセスログ分析部は、

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成であることを特徴とするコミュニティ分析装置

【請求項 2】

前記アクセスログ分析部は、

前記コミュニティ関係を有するユーザの各々をノードとし、コミュニティ関係を有するユーザ間をリンクで結び付けたネットワーク図を生成する処理を実行する構成であることを特徴とする請求項 1 に記載のコミュニティ分析装置。

【請求項 3】

前記アクセスログ分析部は、

予め定められた時間間隔としての閾値インターバル内において、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成であることを特徴とする請求項 1 に記載のコミュニティ分析装置。

【請求項 4】

前記アクセスログ分析部は、

前記閾値インターバルを、アクセス密度の高い時間間隔として設定する処理を実行する構成であることを特徴とする請求項 3 に記載のコミュニティ分析装置。

【請求項 5】

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録するアクセスログ記憶部と、

前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析部とを有し、

前記アクセスログ分析部は、

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成であることを特徴とするコミュニティ分析装置。

【請求項 6】

前記アクセスログ分析部は、

予め定められた時間間隔としての閾値インターバル内において、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成であることを特徴とする請求項 5 に記載のコミュニティ分析装置。

【請求項 7】

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、

アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを有し、

前記アクセスログ分析ステップは、

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユー

10

20

30

40

50

ザとして選択抽出する処理を実行することを特徴とするコミュニティ分析方法。

【請求項 8】

前記コミュニティ分析方法は、さらに、

前記コミュニティ関係を有するユーザの各々をノードとし、コミュニティ関係を有するユーザ間をリンクで結び付けたネットワーク図を生成する処理を実行するステップを有することを特徴とする請求項 7 に記載のコミュニティ分析方法。

【請求項 9】

前記アクセスログ分析ステップは、

予め定められた時間間隔としての閾値インターバル内において、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行することを特徴とする請求項 7 に記載のコミュニティ分析方法。

10

【請求項 10】

前記アクセスログ分析ステップにおいて、

前記閾値インターバルを、アクセス密度の高い時間間隔として設定する処理を実行することを特徴とする請求項 9 に記載のコミュニティ分析方法。

【請求項 11】

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザ ID と、アクセス対象文書識別子である文書 ID とをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、

アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを有し、

20

前記アクセスログ分析ステップは、

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行することを特徴とするコミュニティ分析方法。

【請求項 12】

前記アクセスログ分析ステップは、

予め定められた時間間隔としての閾値インターバル内において、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成であることを特徴とする請求項 11 に記載のコミュニティ分析方法。

30

【請求項 13】

コミュニティ分析処理をコンピュータにおいて実行させるコンピュータ・プログラムであり、

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザ ID と、アクセス対象文書識別子である文書 ID とをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、

アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを、

実行させ、

前記アクセスログ分析ステップにおいては、

40

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行させることを特徴とするコンピュータ・プログラム。

【請求項 14】

コミュニティ分析処理をコンピュータにおいて実行させるコンピュータ・プログラムであり、

電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザ ID と、アクセス対象文書識別子である文書 ID とをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、

アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分

50

析処理を実行するアクセスログ分析ステップとを、
実行させ、

前記アクセスログ分析ステップにおいては、

前記アクセスログ記憶部に記録されたアクセスログに基づいて、

同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理
を実行させることを特徴とするコンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、例えばデータベースに格納された電子文書に対するアクセスログの解析処理
により、人と人との関連性などのコミュニティ分析を行なうコミュニティ分析装置、およ
びコミュニティ分析方法、並びにコンピュータ・プログラムに関する。

10

【背景技術】

【0002】

従来、組織活動の特徴を調査するために、組織のメンバー間の社内メールを使っ
ての連絡や口頭での報告など、さまざまな形態での交信を記録し分析する手法が利用されてきた。
例えば、非特許文献1に見られるソーシャル・ネットワーク分析と呼ばれる手法では、
組織のメンバー間のネットワークを、ネットワーク図で表現する。これにより、人事上の
組織図では表現されないメンバーの役割やメンバー間の依存関係を知ることができる。し
かしながら、ネットワーク分析は、分析者が組織のメンバーにインタビューやアンケート
を行なうことによって進められるため、時間がかかるのが課題である。さらに、分析のた
めには分析者の経験が必要となる。

20

【0003】

近年では、インターネットの発達により多くの交信が電子メールで行なわれるようにな
ってきている。電子メールは、メールサーバと呼ばれる電子メールの送受信を管理するコ
ンピュータを通して行なわれる。メールサーバには通信記録（電子メールログ）が残され
ており、これらを利用することでネットワーク分析も容易に行うことができる。

【0004】

非特許文献2には、この電子メールログを使ったネットワーク分析の手法が提案されて
いる。ネットワーク分析は、誰と誰がコミュニケーションを取ったことがあるかをグラフ
で表し、そのグラフの構造を調べることによって組織のコミュニケーションを分析しよう
とするものである。

30

【0005】

図16を参照して、グラフの構造を調べることによる組織コミュニケーションの分析手
法について説明する。図16において、コミュニケーションの参加者はノード11で、そ
のノード11間のコミュニケーションはリンク12で表される。そして、コミュニケーション
の回数はリンク12の太さで表される。

【0006】

アクセスログから作成されるネットワークはさまざまなネットワークの重ねあわせであ
る。すなわち、図16に示す(A)全ネットワークには、さまざまなネットワークの重ね
あわせである。ネットワーク分析手法を利用しようとする、これらを分離する必要があ
る。例えば図16に示すように、(B1)組織図に対応したネットワーク、(B2)庶務
ネットワーク、(B3)同期入社ネットワークなど、各ノードの情報に基づいて分離した
ネットワークが取得される。

40

【0007】

ネットワーク分析で重要な指標は、コミュニケーションの参加者であるノードに対する
入次数と出次数、2点間の距離、中心媒介性などである。ノードに向かうリンクの数は入
次数、ノードから出るリンクを出次数と呼ぶ。また2点間の距離は、あるノードから別の
ノードまで、直接電子メールをやりとりしたノードをたどり、いくつのノードでたどりつ
けるかのその人数（ノード数）である。さらに、媒介中心性とは、あるノードを取り除い

50

た時に情報が伝わる度合いを表している。

【0008】

また特許文献1では、利用者のメッセージのやりとりの履歴情報を保存しておいて、電子メールのメッセージを返信した相手数、返信を受けた相手数、投稿したメッセージの総量（長さ）、投稿したスレッドの数からグループ活動上の役割に対する利用者の適合度を求める手法を開示している。

【特許文献1】特開2003-216785号公報

【非特許文献1】ローレンス・プルサック、ロブ・クロス、西尚久訳『ソーシャル・ネットワーク：組織活力の源泉』DIAMONDハーバード・ビジネス・レビュー、2002年10月号、p.96-107

10

【非特許文献2】安田雪著、「ネットワーク分析」、新曜社、1997年発行

【発明の開示】

【発明が解決しようとする課題】

【0009】

このように、電子データの利用により、大量のデータであっても分析が可能となる。しかしながら、電子データには、電子メールのように直接コミュニケーションを表すようなデータでないものが存在する。

【0010】

例えば、電子文書を格納した電子ファイリングシステムでは、電子文書の格納や閲覧の情報が記録される。これらのデータは、組織における人の活動を分析する情報であるが、人と人の関係が直接記述されていないが、同じ文書にアクセスした人の間に関係を付けることでネットワークグラフを作成できる。このグラフを使ってネットワーク分析の手法で分析が可能である。しかしながら、アクセスの関係で記述されるネットワークはさまざまな種類のネットワークが含まれている可能性があり、これらを分離する必要がある。この関係は、インタビューや、アクセスした文書の本文を分析することで分離可能であるが、処理の簡便性は失われてしまう。さらに、アクセスした文書の内容を調査することは、組織メンバーのプライバシーを侵害する可能性があり、この点でも実施が困難である。

20

【0011】

本発明は上記事情に鑑みてなされたものであり、調査対象者となるメンバーのプライバシーを保ちつつ、直接人と人の関係が記述されていない電子文書アクセス記録からのコミュニティ分析を実現するコミュニティ分析装置、およびコミュニティ分析方法、並びにコンピュータ・プログラムを提供することを目的とする

30

【課題を解決するための手段】

【0012】

本発明の第1の側面は、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録するアクセスログ記憶部と、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析部とを有し、前記アクセスログ分析部は、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成であることを特徴とするコミュニティ分析装置にある。

40

【0013】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一文書に対する興味を持つユーザを効率的に選択抽出することが可能となり、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【0014】

さらに、本発明のコミュニティ分析装置の一実施態様において、前記アクセスログ分析部は、前記コミュニティ関係を有するユーザの各々をノードとし、コミュニティ関係を有するユーザ間をリンクで結び付けたネットワーク図を生成する処理を実行する構成である

50

ことを特徴とする。

【0015】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて抽出された同一文書に対する興味を持つユーザの集合情報を適用することで、効率的なネットワーク図の作成、コミュニティ解析を行なうことができる。

【0016】

さらに、本発明のコミュニティ分析装置の一実施態様において、前記アクセスログ分析部は、予め定められた時間間隔としての閾値インターバル内において、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成であることを特徴とする。

10

【0017】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において実行された同一文書に対するアクセスユーザを抽出する構成としたので、例えばあるプロジェクトの特定の処理フェーズなど、特定期間内に同一文書に対する興味を示したユーザの集合を選択することが可能であり、より深いコミュニティ関係のあるユーザの抽出処理、およびネットワーク図の作成、コミュニティ解析を行なうことができる。

【0018】

さらに、本発明のコミュニティ分析装置の一実施態様において、前記アクセスログ分析部は、前記閾値インターバルを、アクセス密度の高い時間間隔として設定する処理を実行する構成であることを特徴とする。

20

【0019】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において実行された同一文書に対するアクセスユーザを抽出する際、規定時間をアクセス密度の高い時間間隔として設定する構成としたので、アクセス処理状況に応じた柔軟な解析処理を行なうことができる。

【0020】

さらに、本発明の第2の側面は、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録するアクセスログ記憶部と、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析部とを有し、前記アクセスログ分析部は、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成であることを特徴とするコミュニティ分析装置にある。

30

【0021】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として効率的に抽出することが可能となり、これらの文書関連情報に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【0022】

さらに、本発明のコミュニティ分析装置の一実施態様において、前記アクセスログ分析部は、予め定められた時間間隔としての閾値インターバル内において、同一ユーザによってアクセスの実行された異なる文書を関連文書として抽出する処理を実行する構成であることを特徴とする。

40

【0023】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において同一ユーザによってアクセスの実行された異なる文書を関連文書として抽出する構成としたので、例えばあるプロジェクトの特定の処理フェーズなど、特定期間内にアクセスされた関連文書を選択することが可能となる。

【0024】

さらに、本発明の第3の側面は、電子文書記憶部に対するアクセス情報として、アクセ

50

スユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを有し、前記アクセスログ分析ステップは、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行することを特徴とするコミュニティ分析方法にある。

【0025】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一文書に対する興味を持つユーザを効率的に選択抽出することが可能となり、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

10

【0026】

さらに、本発明のコミュニティ分析方法の一実施態様において、前記コミュニティ分析方法は、さらに、前記コミュニティ関係を有するユーザの各々をノードとし、コミュニティ関係を有するユーザ間をリンクで結び付けたネットワーク図を生成する処理を実行するステップを有することを特徴とする。

【0027】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて抽出された同一文書に対する興味を持つユーザの集合情報を適用することで、効率的なネットワーク図の作成、コミュニティ解析を行なうことができる。

20

【0028】

さらに、本発明のコミュニティ分析方法の一実施態様において、前記アクセスログ分析ステップは、予め定められた時間間隔としての閾値インターバル内において、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行することを特徴とする。

【0029】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において実行された同一文書に対するアクセスユーザを抽出する構成としたので、例えばあるプロジェクトの特定の処理フェーズなど、特定期間内に同一文書に対する興味を示したユーザの集合を選択することが可能であり、より深いコミュニティ関係のあるユーザの抽出処理、およびネットワーク図の作成、コミュニティ解析を行なうことができる。

30

【0030】

さらに、本発明のコミュニティ分析方法の一実施態様において、前記アクセスログ分析ステップにおいて、前記閾値インターバルを、アクセス密度の高い時間間隔として設定する処理を実行することを特徴とする。

【0031】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において実行された同一文書に対するアクセスユーザを抽出する際、規定時間をアクセス密度の高い時間間隔として設定する構成としたので、アクセス処理状況に応じた柔軟な解析処理を行なうことができる。

40

【0032】

さらに、本発明の第4の側面は、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを有し、前記アクセスログ分析ステップは、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行することを特徴とするコミュニティ分析方法にある。

50

【0033】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として効率的に抽出することが可能となり、これらの文書関連情報に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【0034】

さらに、本発明のコミュニティ分析方法の一実施態様において、前記アクセスログ分析ステップは、予め定められた時間間隔としての閾値インターバル内において、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成であることを特徴とする。

10

【0035】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、規定時間内において同一ユーザによってアクセスの実行された異なる文書を関連文書として抽出する構成としたので、例えばあるプロジェクトの特定の処理フェーズなど、特定期間内にアクセスされた関連文書を選択することが可能となる。

【0036】

さらに、本発明の第5の側面は、コミュニティ分析処理をコンピュータにおいて実行させるコンピュータ・プログラムであり、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを実行させ、前記アクセスログ分析ステップにおいては、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行させることを特徴とするコンピュータ・プログラムにある。

20

【0037】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一文書に対する興味を持つユーザを効率的に選択抽出することが可能となり、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析をコンピュータによって行なうことができる。

30

【0038】

さらに、本発明の第6の側面は、コミュニティ分析処理をコンピュータにおいて実行させるコンピュータ・プログラムであり、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記憶部に記録するアクセスログ記憶ステップと、アクセスログ分析部において、前記アクセスログ記憶部に記録されたアクセスログの分析処理を実行するアクセスログ分析ステップとを実行させ、前記アクセスログ分析ステップにおいては、前記アクセスログ記憶部に記録されたアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行させることを特徴とするコンピュータ・プログラムにある。

40

【0039】

本構成によれば、電子文書記憶部に対するアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として効率的に抽出することが可能となり、これらの文書関連情報に基づいて効率的にネットワーク図の作成、コミュニティ解析をコンピュータによって行なうことができる。

【0040】

なお、本発明のコンピュータ・プログラムは、例えば、様々なプログラム・コードを実行可能なコンピュータシステムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体、例えば、CDやFD、MOなどの記録媒体、あるいは、ネットワークなどの通

50

信媒体によって提供可能なコンピュータ・プログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、コンピュータシステム上でプログラムに応じた処理が実現される。

【0041】

本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

【発明の効果】

【0042】

本発明の一実施例の構成によれば、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録し、このアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成としたので、同一文書に対する興味を持つユーザを効率的に選択抽出することが可能となり、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

10

【0043】

さらに、本発明の一実施例の構成によれば、電子文書記憶部に対するアクセスログに基づいて、アクセスログ記憶部に記録されたアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成としたので、関連文書の効率的な抽出処理が可能となり、これらの文書関連情報に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

20

【発明を実施するための最良の形態】

【0044】

以下、図面を参照しながら本発明のコミュニティ分析装置、およびコミュニティ分析方法、並びにコンピュータ・プログラムの詳細について説明する。

【0045】

まず、本発明のコミュニティ分析の概要について説明する。本発明のコミュニティ分析では、例えばネットワークを介してアクセス可能なデータベースに格納されたデータファイルに対するアクセスログの解析を実行する。

30

【0046】

例えば、同じファイルにアクセスを行なっているアクセスログを解析して、そのアクセス時間の間隔によって、ユーザ間、すなわち人と人の関係を分類する。例えば、同じ文書にアクセスするユーザを解析してネットワークを作成しネットワーク分析の手法を用いてコミュニティ分析を行う。

【0047】

一般的なネットワーク分析の教科書などで示されるネットワーク分析手法は、多くの場合、理想的なネットワークを対象とした分析手法を示している。たとえば、組織には複数のコミュニティが存在するが、コミュニティの重なりがないネットワークを対象とした分析を示している。しかし、実際の組織は、組織図上の組織のほかに、新人社員のコミュニティや、組織図上の各グループに属する庶務担当のコミュニティなどが重なり合って存在する。このようなネットワークは、従来のネットワーク分析手法を直接適用することは困難であることが多い。

40

【0048】

本発明においては、例えば新人社員のコミュニティなどの同じコミュニティに属する人たちは、データベースに格納された同じ文書に、ある期間に集中してアクセスを行なうという仮定に基づいて解析処理を行なう。たとえば、同期入社コミュニティでは、入社後、一定年数後に行われる研修の資料をアクセスすることがある。この仮定に基づき、同一

50

文書に対して所定期間内にアクセスするユーザをクラスター化し分類することでネットワークを作成し、ネットワーク分析を可能にする。

【0049】

本発明のコミュニティ分析の実行環境としてのネットワークシステムの構成例について、図1を参照して説明する。ネットワークシステムは、例えば、ある企業内の複数の事業所にまたがって構築された広域ネットワーク(WAN)であって、図1に示すように複数のユーザ端末(クライアント)51~53がネットワーク60によって接続されている。ネットワーク60には、電子文書ファイリング装置70が接続され、ユーザ端末51~53は、電子文書ファイリング装置70に格納された電子文書を取得、あるいは閲覧を行なうことができる。

10

【0050】

電子文書ファイリング装置70にはコミュニティ分析装置100が接続されている。コミュニティ分析装置100は、電子文書ファイリング装置70に対するユーザ端末(クライアント)51~53からのアクセス状況を解析する。

【0051】

図2にコミュニティ分析装置100の構成を示す。図2に示すようにコミュニティ分析装置100は、アクセスログ記憶部101と、アクセスログ分析部102を有する。アクセスログ記憶部101には、電子文書ファイリング装置70に対するユーザ端末(クライアント)51~53からのアクセスログが記録され、アクセスログ分析部102は、アクセスログ記憶部101に記録されたアクセスログの解析を実行する。

20

【0052】

なお、コミュニティ分析装置100は、例えばPCなどの情報処理装置によって構成され、アクセスログ記憶部101はハードディスク等の記憶手段によって構成される。また、アクセスログ分析部102は、ROM等の記憶部に格納されたコンピュータ・プログラムを実行するCPU等を有するデータ処理部によって構成される。アクセスログ分析部102において実行する処理の具体例については、以下、詳細に説明する。

【0053】

図3にアクセスログ記憶部101の記録データの例を示す。アクセスログ記憶部101には、図3に示すように、電子文書ファイリング装置70に対するユーザ端末(クライアント)51~53からのアクセスログとして、アクセス時刻、アクセスを実行したユーザの識別子としてのユーザIDと、アクセス対象となった文書の識別子としての文書IDの対応データを記録する。

30

【0054】

アクセスログ分析部102は、このアクセスログ記憶部101に記録されたアクセスログの解析を実行する。以下、アクセスログ分析部102の実行するアクセスログ解析処理の異なる解析処理例について、順次説明する。以下の4つの処理例について、順次、説明する。

(処理例1) 共通文書に対するアクセスユーザの解析によるコミュニティ分析

(処理例2) 規定時間内の共通文書に対するアクセスユーザの解析によるコミュニティ分析

40

(処理例3) 共通のユーザによる異なる文書に対するアクセス解析によるコミュニティ分析

(処理例4) 規定時間内の同一ユーザによる異なる文書に対するアクセスの解析によるコミュニティ分析

【0055】

(処理例1) 共通文書に対するアクセスユーザの解析によるコミュニティ分析

まず、コミュニティ分析処理例1として、共通文書に対してアクセスを行なうユーザの解析によるコミュニティ分析処理について説明する。アクセスログ分析部102の実行する本処理例における解析処理の一例について、図4を参照して説明する。アクセスログ分析部102は、アクセスログ記憶部101に記録されたアクセスログから、同一文書に対

50

してアクセスを実行した複数のユーザを抽出し、これら同一文書にアクセスしたユーザをコミュニティ関係にあるユーザであると判定する。

【0056】

図4に示すように、図1に示すユーザ端末(クライアント)51~53を利用するユーザAとユーザBが、電子文書ファイリング装置70に格納された同一文書(文書Da)に対してアクセスを行ったことを示すアクセスログがアクセスログ記憶部101に存在する場合、アクセスログ分析部102は、これらのログ情報に基づいてユーザAとユーザBが同一コミュニティに属するユーザである可能性が高いと判定する。

【0057】

具体的には、異なるユーザが同一文書に対して共通にアクセスを実行している回数をカウントし、カウント値に基づいて、ユーザのコミュニティ分析を行なう。例えば、アクセスログ記憶部101に記録されたアクセスログから、ユーザ[U1]とユーザ[U2]との共通文書に対するアクセス回数をA(U1, U2)として、アクセス回数をカウントする。共通文書に対するアクセス回数A(U1, U2)が大きな値を示すほど、ユーザ[U1], [U2]は、同一コミュニティに属する可能性が高いと判定する。

【0058】

このように、アクセスログ分析部102は、同一文書に対するアクセス回数が高いユーザの集合を同一コミュニティに属するユーザであると判定する。この解析処理シーケンスについて、図5に示すフローチャートを参照して説明する。

【0059】

まず、アクセスログ分析部102は、ステップS101において、アクセスログ記憶部101に記録されたアクセスログ(図3参照)から、分析対象期間のログを取り出す。なお、分析期間は、オペレータによって任意の期間を分析期間として予め設定する。

【0060】

次に、ステップS102において、アクセスログのログ番号を示す変数[K]を初期値K=1に設定する。ステップS103において、K番目のログを選択し、ステップS104において、取得したK番目のログに記録されたユーザIDを[U1]とし、ドキュメントIDを[D1]とする。

【0061】

次に、ステップS105において、比較ログのログ番号[K+N]を設定するための変数[N]の初期設定としてN=1に設定し、ステップS106において、K+N番目のログを取得し、ステップS107において、取得したK+N番目のログに記録されたユーザIDを[U2]とし、ドキュメントIDを[D2]とする。

【0062】

ステップS108では、K番目とK+N番目のログのユーザIDを比較し、ステップS109では、K番目とK+N番目のログのドキュメントIDを比較する。ステップS108において、K番目とK+N番目のログのユーザIDが同一であると判定された場合は、同一ユーザによるアクセスログ記録であるので、異なるユーザ間のコミュニティを解析する対象としてのデータに選定されず、ステップS111に進む。

【0063】

ステップS108において、K番目とK+N番目のログのユーザIDが同一でない場合は、ステップS109に進み、K番目とK+N番目のログのドキュメントIDが同一であるか否かを判定する。K番目とK+N番目のログのドキュメントIDが同一である場合は、K番目とK+N番目のログは、異なるユーザによる同一ドキュメントに対するアクセス情報を示すログであると判定される。すなわち、図4を参照して説明した関係を提示していることになる。

【0064】

この場合は、ステップS110に進み、ユーザ[U1]とユーザ[U2]との同一文書に対するアクセス回数を示すデータ[A(U1, U2)]を1つ増加させる処理を行なう。すなわち、

10

20

30

40

50

$A(U_1, U_2) = A(U_1, U_2) + 1$
とする。

【0065】

$[A(U_1, U_2)]$ は、ユーザ U_1, U_2 の同一文書に対するアクセス数を示すデータであり、この数値が高いほど、ユーザ U_1, U_2 が同一コミュニティに属する確率が高いと判定される。

【0066】

ステップS111では、比較ログのログ番号 $[K+N]$ を設定するための変数 $[N]$ の値を1つ増加させる変数更新処理を行なう。すなわち、 K 番目のログとの比較ログを次のログデータに設定する処理である。ステップS112では、比較ログのログ番号 $[K+N]$ が最大ログ番号を超えたか否か、すなわち解析対象として取得したログに存在するか否かを判定する。比較ログのログ番号 $[K+N]$ が最大ログ番号を超えていない場合は、ステップS106に戻り、 K 番目のログと、 $K+N$ 番目のログとの比較処理を繰り返す。

10

【0067】

すなわち、ユーザIDとドキュメントIDの比較を実行し、異なるユーザによる同一ドキュメントに対するアクセスログの関係にあれば、ユーザ $[U_1]$ とユーザ $[U_2]$ との同一文書に対するアクセス回数を示すデータ $[A(U_1, U_2)]$ を1つ増加させる処理を行なう。

【0068】

ステップS112において、比較ログのログ番号 $[K+N]$ が最大ログ番号を超えていると判定され、最終ログまでの比較が終了したと判定すると、ステップS113に進み、比較元のログ番号 $[K]$ を1つ増加させる処理を行い、ステップS114において、比較元のログのログ番号 $[K]$ が最大ログ番号を超えたか否かを判定する。すなわち、最終ログのログ番号を超えていないかをチェックする。ログ番号 $[K]$ が最大ログ番号を超えていない場合は、ステップS103に戻り、 K 番目のログと、 $K+N$ 番目のログとの比較処理を繰り返す。

20

【0069】

ステップS114において、比較元のログのログ番号 $[K]$ が最大ログ番号を超え、取得ログに存在しないと判定されると、すべての取得ログ間の比較処理が終了したことになり、アクセスログ分析部102の処理を終了する。

30

【0070】

このようなログ分析によって、アクセスログ分析部102は、同一文書に対するアクセス回数が高いユーザの組み合わせを解析し、このようなユーザの集合を取得することができる。これらのユーザ集合は、同一コミュニティに属するユーザであると推定することができる。図5を参照して説明した処理フローによって、各ユーザの同一文書に対するアクセス回数 $A(U_n, U_m)$ が算出され、この各ユーザのアクセス回数 $A(U_n, U_m)$ に基づいて、ネットワーク図を作成する。アクセスログ分析部102は、コミュニティ関係を有するユーザの各々をノードとし、コミュニティ関係を有するユーザ間をリンクで結び付けたネットワーク図を生成する処理を実行する。

【0071】

各ユーザの共通文書に対するアクセス回数 $A(U_n, U_m)$ が大きいほどユーザ間の結びつきが高いと判定して、各ユーザ間のアクセス回数 $A(U_n, U_m)$ をユーザ間の結びつきを示す指標、すなわち同一コミュニティに属する判定指標として適用してネットワーク図を作成する。本手法によって、例えば、図6(a)に示すネットワーク図が作成される。図6(a)に示すネットワーク図は、ノード201がユーザを示し、リンク202がユーザ(ノード)間の結びつきを示す。

40

【0072】

リンクの設定は、上述のアクセス回数 $A(U_n, U_m)$ に基づいて設定される。例えば予め設定された閾値以上のアクセス回数を持つユーザ間にリンクを設定する。なお、ネットワーク構成手法の具体例については、例えば、ローレンス・プルサック、ロブ・クロス

50

『ソーシャル・ネットワーク：組織活力の源泉』DIAMONDハーバード・ビジネス・レビュー、2002年10月号96ページに記載されている。この手法を適用し、上述したログ分析によって算出される各ユーザ間のアクセス回数 $A(U_n, U_m)$ をユーザ間の結びつきを示す指標、すなわち同一コミュニティに属する判定指標として適用してネットワーク図を生成する。このネットワーク図によって、人事上の組織図では表現されない、構成人員の役割や、構成人員間の依存関係を知ることができる。

【0073】

このような、ネットワーク図を利用し、ネットワーク図の解析手法において利用される最短経路媒介性を計算することで、コミュニティの発見を行なうことができる。例えば図6(a)に示すように、ユーザノードAと、ユーザノードBとを結びつける最短経路A~C~Bを最短経路媒介性の計算により検出し、図6(b)に示すように、ユーザAの属するコミュニティ1と、ユーザBの属するコミュニティ2を発見することができる。これらの各コミュニティは、例えば、同期入社のコミュニティや、類似製品の開発エンジニアのコミュニティなど、ある共通のカテゴリを持つコミュニティである。

10

【0074】

(処理例2)規定時間内の共通文書に対するアクセスユーザの解析によるコミュニティ分析

次に、コミュニティ分析処理例2として、予め定めた規定時間内に共通文書に対してアクセスを行なうユーザの解析によるコミュニティ分析処理について説明する。アクセスログ分析部102の実行する本処理例における解析処理の一例について、図7を参照して説明する。アクセスログ分析部102は、アクセスログ記憶部101に記録されたアクセスログから、予め定めた規定時間内に同一文書に対してアクセスを実行した複数のユーザを抽出し、これら規定時間内に同一文書にアクセスしたユーザをコミュニティ関係にあるユーザであると判定する。

20

【0075】

すなわち、先の処理例では、時間を考慮することなく、同一文書に対してアクセスを実行したユーザであるか否かのみを判定指標として、同一コミュニティに属するユーザ関係にあるか否かを決定していた。本処理例では、さらに時間要素を加え、規定時間内の同一文書アクセスの有無に基づいて、同一コミュニティに属するユーザ関係にあるか否かを決定する。

30

【0076】

図7に示すように、ユーザU1~U3が、電子文書ファイリング装置70に格納された同一文書(文書D1)に対してアクセスを行ったものとする。これらのアクセス情報を示すアクセスログがアクセスログ記憶部101に存在するものとする。

【0077】

アクセスログ分析部102は、これらのログ情報に基づく解析を実行する。ユーザU1とユーザU2の文書D1に対するアクセス時間の差 t_1 は、予め定めた時間間隔としての閾値インターバル $[I_a]$ 未満であり、ユーザU1とユーザU3の文書D1に対するアクセス時間の差 t_2 は、予め定めた時間間隔としての閾値インターバル $[I_a]$ 以上となっている。これらの時間情報は、アクセスログ記憶部101に記録されたログ情報(図3参照)に基づいて取得される。

40

【0078】

この場合、アクセスログ分析部102は、ユーザU1とユーザU2の関係は、同一コミュニティに属する可能性が高く、ユーザU1とユーザU3の関係は、同一コミュニティに属する可能性が低い関係であると判定する。

【0079】

本処理例においても、先の処理例と同様、異なるユーザが、規定時間(閾値インターバル I_a)未満の時間間隔で、同一文書に対して共通にアクセスを実行している回数をカウントし、カウント値に基づいて、ユーザのコミュニティ分析を行なう。例えば、アクセスログ記憶部101に記録されたアクセスログから、ユーザ $[U_1]$ とユーザ $[U_2]$ との

50

規定時間（閾値インターバル I_a ）未満の時間間隔で実行された共通文書に対するアクセス回数を $A(U_1, U_2)$ として、アクセス回数をカウントする。規定時間未満に実行された共通文書に対するアクセス回数 $A(U_1, U_2)$ が大きな値を示すほど、ユーザ $[U_1]$ 、 $[U_2]$ は、同一コミュニティに属する可能性が高いと判定する。

【0080】

このように、アクセスログ分析部 102 は、規定時間内の同一文書に対するアクセス回数が高いユーザの集合を同一コミュニティに属するユーザであると判定する。この解析処理シーケンスについて、図 8 に示すフローチャートを参照して説明する。

【0081】

まず、アクセスログ分析部 102 は、ステップ S 201 において、アクセスログ記憶部 101 に記録されたアクセスログ（図 3 参照）から、分析対象期間のログを取り出す。なお、分析期間は、オペレータによって任意の期間を分析期間として予め設定する。

【0082】

次に、ステップ S 202 において、アクセスログのログ番号を示す変数 $[K]$ を初期値 $K = 1$ に設定する。ステップ S 203 において、 K 番目のログを選択し、ステップ S 204 において、取得した K 番目のログに記録されたアクセス時刻を $[T_1]$ 、ユーザ ID を $[U_1]$ とし、ドキュメント ID を $[D_1]$ とする。

【0083】

次に、ステップ S 205 において、比較ログのログ番号 $[K + N]$ を設定するための変数 $[N]$ の初期設定として $N = 1$ に設定し、ステップ S 206 において、 $K + N$ 番目のログを取得し、ステップ S 207 において、取得した $K + N$ 番目のログに記録されたアクセス時刻を $[T_2]$ 、ユーザ ID を $[U_2]$ とし、ドキュメント ID を $[D_2]$ とする。

【0084】

ステップ S 208 では、 K 番目と $K + N$ 番目のログのユーザ ID を比較し、ステップ S 209 では、 K 番目と $K + N$ 番目のログのドキュメント ID を比較し、ステップ S 210 では、 K 番目と $K + N$ 番目のログのアクセス時刻の差 $[T_2 - T_1]$ と予め定めた閾値インターバル $[I_a]$ とを比較する。すなわち、2 つのログのアクセス時間の差が閾値インターバル $[I_a]$ 未満であるか否かを判定する。

【0085】

ステップ S 208 において、 K 番目と $K + N$ 番目のログのユーザ ID が同一であると判定された場合は、同一ユーザによるアクセスログ記録であるので、異なるユーザ間のコミュニティを解析する対象としてのデータに選定されず、ステップ S 212 に進む。

【0086】

ステップ S 208 において、 K 番目と $K + N$ 番目のログのユーザ ID が同一でない場合は、ステップ S 209 に進み、 K 番目と $K + N$ 番目のログのドキュメント ID が同一であるか否かを判定する。 K 番目と $K + N$ 番目のログのドキュメント ID が同一である場合は、 K 番目と $K + N$ 番目のログは、異なるユーザによる同一ドキュメントに対するアクセス情報を示すログであると判定される。

【0087】

この場合は、さらに、ステップ S 210 に進み、 K 番目と $K + N$ 番目のログのアクセス時刻の差 $[T_2 - T_1]$ と予め定めた閾値インターバル $[I_a]$ とを比較する。すなわち

$$T_2 - T_1 < I_a$$

上記式が成立するか否かを判定する。これは、2 つのログのアクセス時間の差が閾値インターバル $[I_a]$ 未満であるか否かを判定する処理である。

【0088】

上記式が成立する場合は、ステップ S 211 に進み、ユーザ $[U_1]$ とユーザ $[U_2]$ との規定時間内の同一文書に対するアクセス回数を示すデータ $[A(U_1, U_2)]$ を 1 つ増加させる処理を行なう。すなわち、

$$A(U_1, U_2) = A(U_1, U_2) + 1$$

とする。このケースは、図 7 に示すユーザ U 1 とユーザ U 2 との関係に相当する。

【 0 0 8 9 】

ステップ S 2 1 0 において、K 番目と K + N 番目のログのアクセス時刻の差 [T 2 - T 1] が、

$$T 2 - T 1 < I a$$

を満たさない場合、すなわち、2つのログのアクセス時間の差が閾値インターバル [I a] 以上である場合は、2つのログのユーザは、同一文書に対するアクセスを行ってはいいるが、予め定めた閾値インターバル以上の差を持ったアクセスであり、このログに示されるユーザ間の共通性は低いと判定し、ステップ S 2 1 1 のカウントアップ処理を実行することなく、ステップ S 2 1 2 に進む。このケースは、図 7 に示すユーザ U 1 とユーザ U 3 との関係に相当する。

10

【 0 0 9 0 】

本処理例では、[A (U 1 , U 2)] は、ユーザ U 1 , U 2 の規定時間内の同一文書に対するアクセス数を示すデータであり、この数値が高いほど、ユーザ U 1 , U 2 が同一コミュニティに属する確率が高いと判定される。

【 0 0 9 1 】

ステップ S 2 1 2 では、比較ログのログ番号 [K + N] を設定するための変数 [N] の値を 1 つ増加させる変数更新処理を行なう。すなわち、K 番目のログとの比較ログを次のログデータに設定する処理である。ステップ S 2 1 3 では、比較ログのログ番号 [K + N] が最大ログ番号を超えたか否か、すなわち解析対象として取得したログに存在するか否かを判定する。比較ログのログ番号 [K + N] が最大ログ番号を超えていない場合は、ステップ S 2 0 6 に戻り、K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

20

【 0 0 9 2 】

すなわち、ログ間のユーザ ID とドキュメント ID の比較と、ログ間の時間差 [T 2 - T 1] と閾値インターバル [I a] との比較を実行し、異なるユーザによる閾値時間 T a 内の同一ドキュメントに対するアクセスログの関係にあれば、ユーザ [U 1] とユーザ [U 2] との規定時間内の同一文書に対するアクセス回数を示すデータ [A (U 1 , U 2)] を 1 つ増加させる処理を行なう。

【 0 0 9 3 】

ステップ S 2 1 3 において、比較ログのログ番号 [K + N] が最大ログ番号を超えていると判定され、最終ログまでの比較が終了したと判定すると、ステップ S 2 1 4 に進み、比較元のログ番号 [K] を 1 つ増加させる処理を行い、ステップ S 2 1 5 において、比較元のログのログ番号 [K] が最大ログ番号を超えたか否かを判定する。すなわち、最終ログのログ番号を超えていないかをチェックする。ログ番号 [K] が最大ログ番号を超えていない場合は、ステップ S 2 0 3 に戻り、更新したログ番号 [K] によって指定される K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

30

【 0 0 9 4 】

ステップ S 2 1 5 において、比較元のログのログ番号 [K] が最大ログ番号を超え、取得ログに存在しないと判定されると、すべての取得ログ間の比較処理が終了したことになり、アクセスログ分析部 1 0 2 の処理を終了する。

40

【 0 0 9 5 】

このようなログ分析によって、アクセスログ分析部 1 0 2 は、同一文書に対する規定時間内のアクセス回数が高いユーザの組み合わせを解析し、このようなユーザの集合を取得することができる。これらのユーザ集合は、同一コミュニティに属するユーザであると推定することができる。このように、図 8 を参照して説明した処理フローによって、各ユーザの規定時間内の同一文書に対するアクセス回数 A (U n , U m) が算出され、この規定時間内の異なるユーザのアクセス回数 A (U n , U m) に基づいて、ネットワーク図を作成する。

【 0 0 9 6 】

各ユーザの閾値インターバル [I a] 内の共通文書に対するアクセス回数 A (U n , U

50

m) が大きいほどユーザ間の結びつきが高いと判定して、各ユーザ間のアクセス回数 $A(U_n, U_m)$ をユーザ間の結びつきを示す指標、すなわち同一コミュニティに属する判定指標として適用してネットワーク図を作成する。本手法によっても、先に図6を参照して説明したと同様のネットワーク図が作成される。

【0097】

このような、閾値インターバル $[I_a]$ 内の共通文書に対するアクセス回数 $A(U_n, U_m)$ に基づくネットワーク分析によって、例えば、図9に示すような、異なるコミュニティ、例えば課長コミュニティや、同期入社コミュニティの存在を把握することが可能となる。例えば、年次教育の資料に対するアクセスログの解析によって、このような、組織図からは直接読み取ることが困難なコミュニティの存在を解析することができる。

10

【0098】

なお、上述した処理例では、図7を参照して説明したように、アクセスログ分析部102が実行するログ情報に基づく解析において、異なるユーザ間の同一文書 D_1 に対するアクセス時間の差と、予め定めた時間間隔としての閾値インターバル $[I_a]$ とを比較して、ユーザ間の関係を判断する際、適用する閾値インターバル $[I_a]$ を予め定めた固定時間とした例を説明したが、この閾値インターバルを文書に対するアクセス状況に応じて決定する構成としてもよい。

【0099】

この処理例について、図10を参照して説明する。図10(a)は、上述した処理例と同様、予め定めた時間間隔としての閾値インターバル $[I_a]$ を適用した場合の処理である。(b)は、閾値インターバルを文書に対するアクセス状況に応じて決定する構成例を示している。図において、左から右に時間経過を示しており、マーク301はユーザによるアクセスタイミングを示している。これらの時間情報は、アクセスログ記憶部101に記録されたログ情報(図3参照)に基づいて取得される。

20

【0100】

このアクセス履歴から、時間 $T_1 \sim T_2$ の間に、アクセスが集中していることが解析される。このアクセス密度の高い期間をユーザ間の関係を判断する際に適用する閾値インターバル $[I_b]$ とする。具体的には、予め定めたアクセス密度の値以上のアクセス密度を持つ期間を閾値インターバル $[I_b]$ として設定する。このようにアクセス密度に応じて閾値インターバル $[I_b]$ を設定することで、より柔軟な解析が可能となる。

30

【0101】

(処理例3) 共通のユーザによる異なる文書に対するアクセス解析によるコミュニティ分析

次に、コミュニティ分析処理例3として、共通のユーザによる異なる文書に対するアクセス解析によるコミュニティ分析処理について説明する。アクセスログ分析部102の実行する本処理例における解析処理の一例について、図11を参照して説明する。アクセスログ分析部102は、アクセスログ記憶部101に記録されたアクセスログから、共通のユーザによる異なる文書に対するアクセスログを抽出し、これら共通のユーザによってアクセスされた文書を関連文書であると判定する。

【0102】

図11に示すように、同一ユーザが、電子文書ファイリング装置70に格納された異なる文書、すなわち文書 D_a と文書 D_b に対してアクセスを行ったことを示すアクセスログがアクセスログ記憶部101に存在する場合、アクセスログ分析部102は、これらのログ情報に基づいて文書 D_a と文書 D_b とが関連性の高い文書であると判定する。

40

【0103】

具体的には、同一ユーザが異なる文書に対してアクセスを実行している回数をカウントし、カウント値に基づいてコミュニティ分析を行なう。例えば、アクセスログ記憶部101に記録されたアクセスログから、同一のユーザ $[U_1]$ による異なる文書 $[D_1]$, $[D_2]$ に対するアクセス回数を $A(D_1, D_2)$ として、アクセス回数をカウントする。共通ユーザによる異なる文書に対するアクセス回数 $A(D_1, D_2)$ が大きな値を示すほ

50

ど、文書 [D 1] , [D 2] は、関連性の高い文書である可能性が高いと判定する。この解析処理シーケンスについて、図 1 2 に示すフローチャートを参照して説明する。

【 0 1 0 4 】

まず、アクセスログ分析部 1 0 2 は、ステップ S 3 0 1 において、アクセスログ記憶部 1 0 1 に記録されたアクセスログ (図 3 参照) から、分析対象期間のログを取り出す。なお、分析期間は、オペレータによって任意の期間を分析期間として予め設定する。

【 0 1 0 5 】

次に、ステップ S 3 0 2 において、アクセスログのログ番号を示す変数 [K] を初期値 $K = 1$ に設定する。ステップ S 3 0 3 において、K 番目のログを選択し、ステップ S 3 0 4 において、取得した K 番目のログに記録されたユーザ ID を [U 1] とし、ドキュメント ID を [D 1] とする。

10

【 0 1 0 6 】

次に、ステップ S 3 0 5 において、比較ログのログ番号 [K + N] を設定するための変数 [N] の初期設定として $N = 1$ に設定し、ステップ S 3 0 6 において、K + N 番目のログを取得し、ステップ S 3 0 7 において、取得した K + N 番目のログに記録されたユーザ ID を [U 2] とし、ドキュメント ID を [D 2] とする。

【 0 1 0 7 】

ステップ S 3 0 8 では、K 番目と K + N 番目のログのドキュメント ID を比較し、ステップ S 3 0 9 では、K 番目と K + N 番目のログのユーザ ID を比較する。ステップ S 3 0 8 において、K 番目と K + N 番目のログのドキュメント ID が同一であると判定された場合は、同一ドキュメントに対するアクセスログ記録であるので、関連文書判定処理対象としてのデータに選定されず、ステップ S 3 1 1 に進む。

20

【 0 1 0 8 】

ステップ S 3 0 8 において、K 番目と K + N 番目のログのドキュメント ID が同一でない場合は、ステップ S 3 0 9 に進み、K 番目と K + N 番目のログのユーザ ID が同一であるか否かを判定する。K 番目と K + N 番目のログのユーザ ID が同一である場合は、K 番目と K + N 番目のログは、同一ユーザによる異なるドキュメントに対するアクセス情報を示すログであると判定される。すなわち、図 1 1 を参照して説明した関係を提示していることになる。

【 0 1 0 9 】

この場合は、ステップ S 3 1 0 に進み、同一ユーザの異なる文書に対するアクセス回数

30

を示すデータ [A (D 1 , D 2)] を 1 つ増加させる処理を行なう。すなわち、
 $A (D 1 , D 2) \quad A (D 1 , D 2) + 1$
 とする。

【 0 1 1 0 】

[A (D 1 , D 2)] は、文書 D 1 , D 2 に対する共通ユーザによるアクセス数を示すデータであり、この数値が高いほど、文書 D 1 , D 2 の関連性が高いと判定される。

【 0 1 1 1 】

ステップ S 3 1 1 では、比較ログのログ番号 [K + N] を設定するための変数 [N] の値を 1 つ増加させる変数更新処理を行なう。すなわち、K 番目のログとの比較ログを次のログデータに設定する処理である。ステップ S 3 1 2 では、比較ログのログ番号 [K + N] が最大ログ番号を超えたか否か、すなわち解析対象として取得したログに存在するか否かを判定する。比較ログのログ番号 [K + N] が最大ログ番号を超えていない場合は、ステップ S 3 0 6 に戻り、K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

40

【 0 1 1 2 】

すなわち、ドキュメント ID とユーザ ID の比較を実行し、同一ユーザによる異なるドキュメントに対するアクセスログの関係にあれば、同一ユーザによるドキュメント [D 1] とドキュメント [D 2] のアクセス回数

を示すデータ [A (D 1 , D 2)] を 1 つ増加させる処理を行なう。

【 0 1 1 3 】

50

ステップ S 3 1 2 において、比較ログのログ番号 [K + N] が最大ログ番号を超えていると判定され、最終ログまでの比較が終了したと判定すると、ステップ S 3 1 3 に進み、比較元のログ番号 [K] を 1 つ増加させる処理を行い、ステップ S 3 1 4 において、比較元のログのログ番号 [K] が最大ログ番号を超えたか否かを判定する。すなわち、最終ログのログ番号を超えていないかをチェックする。ログ番号 [K] が最大ログ番号を超えていない場合は、ステップ S 3 0 3 に戻り、K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

【 0 1 1 4 】

ステップ S 3 1 4 において、比較元のログのログ番号 [K] が最大ログ番号を超え、取得ログに存在しないと判定されると、すべての取得ログ間の比較処理が終了したことになり、アクセスログ分析部 1 0 2 の処理を終了する。

10

【 0 1 1 5 】

このようなログ分析によって、アクセスログ分析部 1 0 2 は、同一ユーザによる異なる文書に対するアクセス回数が高い文書の組み合わせを解析し、このような文書の集合を取得することができる。これらの文書集合は、同一コミュニティに属するユーザによってアクセスされる可能性の高い文書集合であると推定することができる。図 1 2 を参照して説明した処理フローによって、同一ユーザの異なる文書に対するアクセス回数 $A(D_n, D_m)$ が算出され、このアクセス回数 $A(D_n, D_m)$ に基づいて、ネットワーク図を作成する。

【 0 1 1 6 】

同一ユーザの異なる文書に対するアクセス回数 $A(D_n, D_m)$ が大きいほど文書間の関連性が高いと判定して、これらの文書集合は、同一コミュニティに属するユーザによってアクセスされる可能性の高い文書集合であると推定し、このような共通の文書集合との結びつきを持つ異なるユーザ間の結びつきについても高い結びつきがあると推定し、この推定をユーザ間の結びつきを示す指標、すなわち同一コミュニティに属する判定指標として適用してネットワーク図を作成する。本手法によっても、例えば、図 6 を参照して説明したと同様のネットワーク図が設定され、ネットワーク解析を行なうことができる。

20

【 0 1 1 7 】

(処理例 4) 規定時間内の同一ユーザによる異なる文書に対するアクセスの解析によるコミュニティ分析

30

次に、コミュニティ分析処理例 4 として、予め定めた規定時間内に同一ユーザによって異なる文書に対して行なわれたアクセスの解析によって行なわれるコミュニティ分析処理について説明する。アクセスログ分析部 1 0 2 の実行する本処理例における解析処理の一例について、図 1 3 を参照して説明する。アクセスログ分析部 1 0 2 は、アクセスログ記憶部 1 0 1 に記録されたアクセスログから、同一のユーザが予め定めた規定時間内に異なる文書に対してアクセスを行なったことを示すログを抽出し、これら規定時間内に同一ユーザがアクセスした文書を関連性の高い文書であると判定する。

【 0 1 1 8 】

すなわち、先の処理例 3 では、時間を考慮することなく、同一ユーザによるアクセス文書を関連性のある文書として判定する関連性判定を実行したが、本処理例では、さらに時間要素を加え、同一ユーザによって、規定時間内に文書アクセスが発生した場合に関連性のある文書であると判定する。

40

【 0 1 1 9 】

図 1 3 に示すように、ユーザ U 1 が、電子文書ファイリング装置 7 0 に格納された異なる文書、すなわち、文書 D 1 ~ D 3 に対してアクセスを行ったものとする。これらのアクセス情報を示すアクセスログがアクセスログ記憶部 1 0 1 に存在するものとする。

【 0 1 2 0 】

アクセスログ分析部 1 0 2 は、これらのログ情報に基づく解析を実行する。ユーザ U 1 による文書 D 1 と文書 D 2 に対するアクセス時間の差 t_1 は、予め定めた時間間隔としての閾値インターバル [I b] 未満であり、文書 D 1 と文書 D 3 に対するアクセス時間の差

50

t 2 は、予め定めた時間間隔としての閾値インターバル [I b] 以上となっている。これらの時間情報は、アクセスログ記憶部 1 0 1 に記録されたログ情報 (図 3 参照) に基づいて取得される。

【 0 1 2 1 】

この場合、アクセスログ分析部 1 0 2 は、文書 D 1 と文書 D 2 は関連性が高は、文書 D 1 と文書 D 3 の関連性は低いと判定する。

【 0 1 2 2 】

本処理例においては、各文書の組み合わせ (D n , D m) について、規定時間 (閾値インターバル I a) 未満の時間間隔で同一ユーザによってアクセスされた回数をカウントし、カウント値に基づいて、文書の関連性を判断する。例えば、アクセスログ記憶部 1 0 1 に記録されたアクセスログから、文書 [D 1] と文書 [D 2] とに対して、規定時間 (閾値時間 T a) 未満の時間間隔で実行された同一ユーザからのアクセス回数を A (D 1 , D 2) として、アクセス回数をカウントする。規定時間未満に同一ユーザによって実行された異なる文書に対するアクセス回数 A (D 1 , D 2) が大きな値を示すほど、文書 [D 1] , [D 2] は関連性が高いと判定する。この解析処理シーケンスについて、図 1 4 に示すフローチャートを参照して説明する。

【 0 1 2 3 】

まず、アクセスログ分析部 1 0 2 は、ステップ S 4 0 1 において、アクセスログ記憶部 1 0 1 に記録されたアクセスログ (図 3 参照) から、分析対象期間のログを取り出す。なお、分析期間は、オペレータによって任意の期間を分析期間として予め設定する。

【 0 1 2 4 】

次に、ステップ S 4 0 2 において、アクセスログのログ番号を示す変数 [K] を初期値 K = 1 に設定する。ステップ S 4 0 3 において、K 番目のログを選択し、ステップ S 4 0 4 において、取得した K 番目のログに記録されたアクセス時刻を [T 1] 、ユーザ ID を [U 1] とし、ドキュメント ID を [D 1] とする。

【 0 1 2 5 】

次に、ステップ S 4 0 5 において、比較ログのログ番号 [K + N] を設定するための変数 [N] の初期設定として N = 1 に設定し、ステップ S 4 0 6 において、K + N 番目のログを取得し、ステップ S 4 0 7 において、取得した K + N 番目のログに記録されたアクセス時刻を [T 2] 、ユーザ ID を [U 2] とし、ドキュメント ID を [D 2] とする。

【 0 1 2 6 】

ステップ S 4 0 8 では、K 番目と K + N 番目のログのドキュメント ID を比較し、ステップ S 4 0 9 では、K 番目と K + N 番目のログのユーザ ID を比較し、ステップ S 4 1 0 では、K 番目と K + N 番目のログのアクセス時刻の差 [T 2 - T 1] と予め定めた閾値インターバル [I b] とを比較する。すなわち、2 つのログのアクセス時間の差が閾値インターバル [I b] 未満であるか否かを判定する。

【 0 1 2 7 】

ステップ S 4 0 8 において、K 番目と K + N 番目のログのドキュメント ID が同一であると判定された場合は、同一文書に対するアクセスログ記録であるので、異なる文書間の関連性を判定する対象としてのデータに選定されず、ステップ S 4 1 2 に進む。

【 0 1 2 8 】

ステップ S 4 0 8 において、K 番目と K + N 番目のログのドキュメント ID が同一でない場合は、ステップ S 4 0 9 に進み、K 番目と K + N 番目のログのユーザ ID が同一であるか否かを判定する。K 番目と K + N 番目のログのユーザ ID が同一である場合は、K 番目と K + N 番目のログは、同一ユーザによる異なるドキュメントに対するアクセス情報を示すログであると判定される。

【 0 1 2 9 】

この場合は、さらに、ステップ S 4 1 0 に進み、K 番目と K + N 番目のログのアクセス時刻の差 [T 2 - T 1] と予め定めた閾値インターバル [I b] とを比較する。すなわち

、

10

20

30

40

50

$$T_2 - T_1 < I_b$$

上記式が成立するか否かを判定する。これは、2つのログのアクセス時間の差が閾値インターバル $[I_b]$ 未満であるか否かを判定する処理である。

【0130】

上記式が成立する場合は、ステップ S 4 1 1 に進み、同一ユーザの異なる文書に対するアクセス回数を示すデータ $[A(D_1, D_2)]$ を1つ増加させる処理を行なう。すなわち、

$$A(D_1, D_2) = A(D_1, D_2) + 1$$

とする。このケースは、図 1 3 に示す文書 D 1 と文書 D 2 との関係に相当する。

【0131】

$[A(D_1, D_2)]$ は、文書 D 1, D 2 に対する共通ユーザによるアクセス数を示すデータであり、この数値が高いほど、文書 D 1, D 2 の関連性が高いと判定される。

【0132】

ステップ S 4 1 0 において、K 番目と K + N 番目のログのアクセス時刻の差 $[T_2 - T_1]$ が、

$$T_2 - T_1 < I_b$$

を満たさない場合、すなわち、2つのログのアクセス時間の差が閾値インターバル $[I_b]$ 以上である場合は、2つのログの文書は、同一ユーザによるアクセスがなされているが、予め定めた閾値時間以上の差を持ったアクセスであり、このログに示される文書間の関連性は低いと判定し、ステップ S 4 1 1 のカウントアップ処理を実行することなく、ステップ S 4 1 2 に進む。このケースは、図 1 3 に示す文書 D 1 と文書 D 3 との関係に相当する。

【0133】

本処理例では、 $[A(D_1, D_2)]$ は、文書 D 1, D 2 の規定時間内の同一ユーザによるアクセス数を示すデータであり、この数値が高いほど、文書 D 1, D 2 の関連性が高いと判定される。

【0134】

ステップ S 4 1 2 では、比較ログのログ番号 $[K + N]$ を設定するための変数 $[N]$ の値を1つ増加させる変数更新処理を行なう。すなわち、K 番目のログとの比較ログを次のログデータに設定する処理である。ステップ S 4 1 3 では、比較ログのログ番号 $[K + N]$ が最大ログ番号を超えたか否か、すなわち解析対象として取得したログに存在するか否かを判定する。比較ログのログ番号 $[K + N]$ が最大ログ番号を超えていない場合は、ステップ S 4 0 6 に戻り、K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

【0135】

すなわち、ログ間のドキュメント ID とユーザ ID の比較と、ログ間の時間差 $[T_2 - T_1]$ と閾値インターバル $[I_b]$ との比較を実行し、同一ユーザによる閾値インターバル I_b 内の異なるドキュメントに対応するアクセスログの関係にあれば、文書 D 1, D 2 の規定時間内の同一ユーザによるアクセス数を示すデータ $[A(D_1, D_2)]$ を1つ増加させる処理を行なう。

【0136】

ステップ S 4 1 3 において、比較ログのログ番号 $[K + N]$ が最大ログ番号を超えていると判定され、最終ログまでの比較が終了したと判定すると、ステップ S 4 1 4 に進み、比較元のログ番号 $[K]$ を1つ増加させる処理を行い、ステップ S 4 1 5 において、比較元のログのログ番号 $[K]$ が最大ログ番号を超えたか否かを判定する。すなわち、最終ログのログ番号を超えていないかをチェックする。ログ番号 $[K]$ が最大ログ番号を超えていない場合は、ステップ S 4 0 3 に戻り、更新したログ番号 $[K]$ によって指定される K 番目のログと、K + N 番目のログとの比較処理を繰り返す。

【0137】

ステップ S 4 1 5 において、比較元のログのログ番号 $[K]$ が最大ログ番号を超え、取得ログに存在しないと判定されると、すべての取得ログ間の比較処理が終了したことになる

10

20

30

40

50

り、アクセスログ分析部 102 の処理を終了する。

【0138】

このようなログ分析によって、アクセスログ分析部 102 は、同一ユーザによる規定時間内のアクセス回数が高い異なる文書の組み合わせを解析し、このような文書の集合を取得することができる。これらの文書集合は関連性が高いと判定して、これらの文書集合は、同一コミュニティに属するユーザによってアクセスされる可能性の高い文書集合であると推定し、このような共通の文書集合との結びつきを持つ異なるユーザ間の結びつきについても高い結びつきがあると推定し、この推定をユーザ間の結びつきを示す指標、すなわち同一コミュニティに属する判定指標として適用してネットワーク図を作成する。本手法によっても、例えば、図 6 を参照して説明したと同様のネットワーク図が設定され、ネットワーク解析を行なうことができる。

10

【0139】

以上、複数のコミュニティ分析処理例として、ユーザによる文書アクセスログの解析によるコミュニティ分析処理例として、以下の 4 つの処理例、

(処理例 1) 共通文書に対するアクセスユーザの解析によるコミュニティ分析

(処理例 2) 規定時間内の共通文書に対するアクセスユーザの解析によるコミュニティ分析

(処理例 3) 共通のユーザによる異なる文書に対するアクセス解析によるコミュニティ分析

(処理例 4) 規定時間内の同一ユーザによる異なる文書に対するアクセスの解析によるコミュニティ分析

20

これらの処理例を説明した。

【0140】

例えば、処理例 2 や、処理例 4 では、規定の時間ごとに区切りを設定して、ユーザのコミュニティ関係を解析する構成としているが、例えば、このような時間区切りのコミュニティ解析を実行することで、特定のプロジェクトのメンバーの関わりなどを解析することが容易となる。

【0141】

図 15 を参照してコミュニティ解析処理例について説明する。例えば、ある会社においてある製品の開発プロジェクトがあった場合、その製品に関するプロセスの流れとしては、企画、設計、試作、評価、量産等の各ステップが時間の進行に従って行われる。これらのプロジェクトにかかわるメンバーは、各処理ステップのいずれかにかかわることになる。企画、設計、試作、評価、量産の全ステップのいずれかに参加したメンバーのネットワーク図として、図 15 (a) が得られるようなコミュニティがある場合、上述した処理例 2 や、処理例 4 を適用し、規定の時間ごとに区切りを設定して、ユーザのコミュニティ関係を解析することで、図 15 (b) ~ (e) のようなプロジェクトの段階ごとの時期に対応した個別のコミュニティネットワークを取得し、解析を行なうことが可能となる。

30

【0142】

また、(処理例 1) 共通文書に対するアクセスユーザの解析によるコミュニティ分析を適用することで、ある専門分野、例えば建築関係の文書にアクセスしているユーザのネットワークを解析することが加納であり、また、(処理例 3) 共通のユーザによる異なる文書に対するアクセス解析によるコミュニティ分析を実行することで、異なる文書の関連性の解析が可能となり、このような関連性の高い文書集合との結びつきを持つユーザ間の結びつきをネットワーク解析によって行なうことが可能となる。

40

【0143】

以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

50

【0144】

なお、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。

【0145】

例えば、プログラムは記録媒体としてのハードディスクやROM (Read Only Memory) に予め記録しておくことができる。あるいは、プログラムはフレキシブルディスク、CD-ROM (Compact Disc Read Only Memory)、MO (Magneto optical) ディスク、DVD (Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体に、一時的あるいは永続的に格納 (記録) しておくことができる。このようなりムーバブル記録媒体は、いわゆるパッケージソフトウェアとして提供することができる。

【0146】

なお、プログラムは、上述したようなりムーバブル記録媒体からコンピュータにインストールする他、ダウンロードサイトから、コンピュータに無線転送したり、LAN (Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

【0147】

なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

【産業上の利用可能性】

【0148】

以上、説明したように、本発明の一実施例の構成によれば、電子文書記憶部に対するアクセス情報として、アクセスユーザ識別子であるユーザIDと、アクセス対象文書識別子である文書IDとをアクセス日時情報に対応付けて記録し、このアクセスログに基づいて、同一文書に対するアクセスを実行した異なるユーザを、コミュニティ関係を有するユーザとして選択抽出する処理を実行する構成としたので、同一文書に対する興味を持つユーザを効率的に選択抽出することが可能となり、同一文書に対する興味を持つユーザの集合をコミュニティ関係のあるユーザとして判定し、これらのユーザ集合に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【0149】

さらに、本発明の一実施例の構成によれば、電子文書記憶部に対するアクセスログに基づいて、アクセスログ記憶部に記録されたアクセスログに基づいて、同一ユーザによってアクセスの実行された異なる文書を、関連文書として抽出する処理を実行する構成としたので、関連文書の効率的な抽出処理が可能となり、これらの文書関連情報に基づいて効率的にネットワーク図の作成、コミュニティ解析を行なうことができる。

【図面の簡単な説明】

【0150】

【図1】本発明にかかるコミュニティ分析が適用されるネットワークシステムの構成を示す図である。

【図2】本発明にかかるコミュニティ分析装置の構成を示す図である。

【図3】アクセスログ記憶部の記録データの例を示す図である。

【図4】アクセスログ分析部の実行する解析処理の一例について説明する図である。

【図5】アクセスログ分析部の実行する解析処理シーケンスについて説明するフローチャートを示す図である。

10

20

30

40

50

【図6】本発明にかかるコミュニティ分析装置において実行される解析情報を適用して生成されるネットワーク図の例を示す図である。

【図7】アクセスログ分析部の実行する解析処理の一例について説明する図である。

【図8】アクセスログ分析部の実行する解析処理シーケンスについて説明するフローチャートを示す図である。

【図9】本発明にかかるコミュニティ分析装置において実行される解析処理例について説明する図である。

【図10】アクセスログ分析部の実行する解析処理において、適用する閾値インターバルの一例について説明する図である。

【図11】アクセスログ分析部の実行する解析処理の一例について説明する図である。

10

【図12】アクセスログ分析部の実行する解析処理シーケンスについて説明するフローチャートを示す図である。

【図13】アクセスログ分析部の実行する解析処理の一例について説明する図である。

【図14】アクセスログ分析部の実行する解析処理シーケンスについて説明するフローチャートを示す図である。

【図15】本発明にかかるコミュニティ分析装置において実行される解析処理例について説明する図である。

【図16】グラフの構造を調べることによって行なわれる組織コミュニケーションの分析手法について説明する図である。

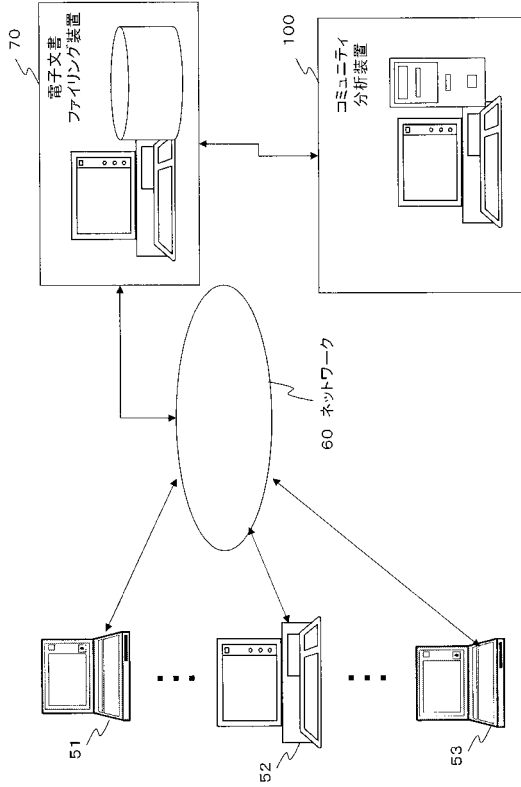
【符号の説明】

20

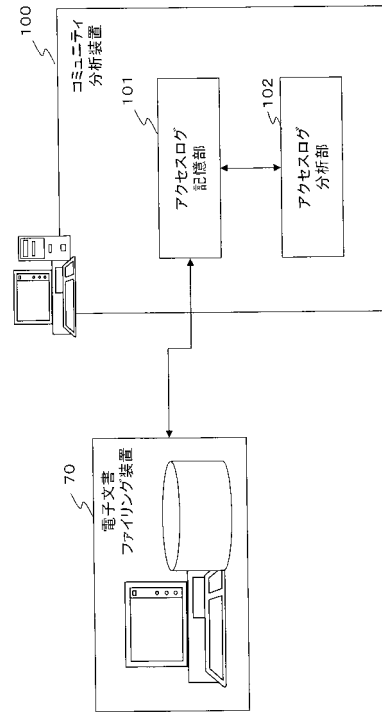
【0151】

- 11 ノード
- 12 リンク
- 51 ~ 53 ユーザ端末（クライアント）
- 60 ネットワーク
- 70 電子文書ファイリング装置
- 100 コミュニティ分析装置
- 101 アクセスログ記憶部
- 102 アクセスログ分析部

【 図 1 】



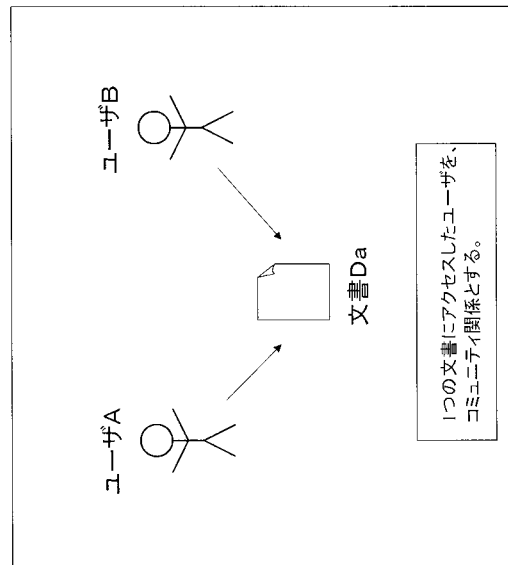
【 図 2 】



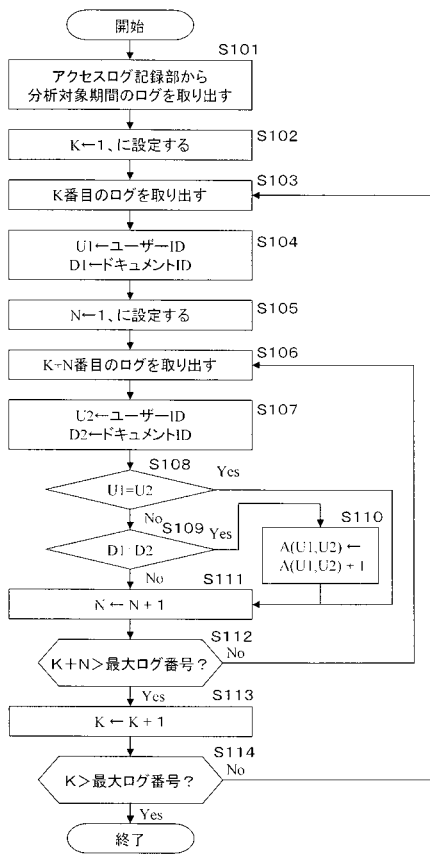
【 図 3 】

	ユーザーID	文書ID (ドキュメントID)
	U0001	D0002
2003/01/01/10:00	U0003	D0011
2003/01/01/10:20	U0011	D0002
2003/01/01/11:10	U0121	D0002
2003/01/01/12:00	⋮	⋮
⋮	⋮	⋮
2003/03/01/11:00	U0001	D0002
2003/03/04/12:07	U0001	D0021

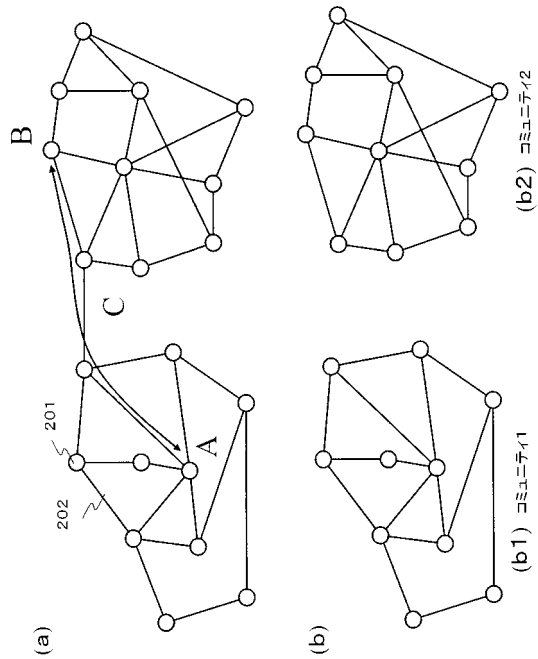
【 図 4 】



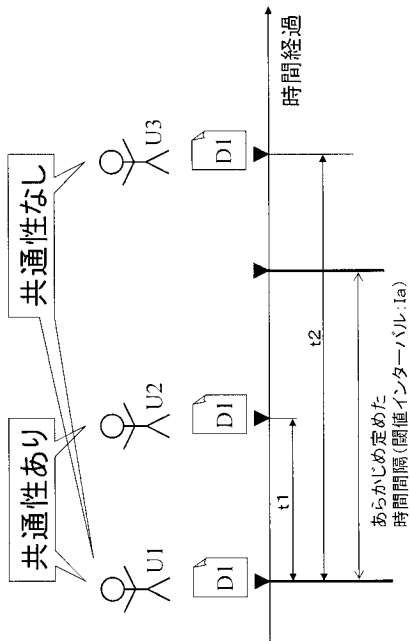
【 図 5 】



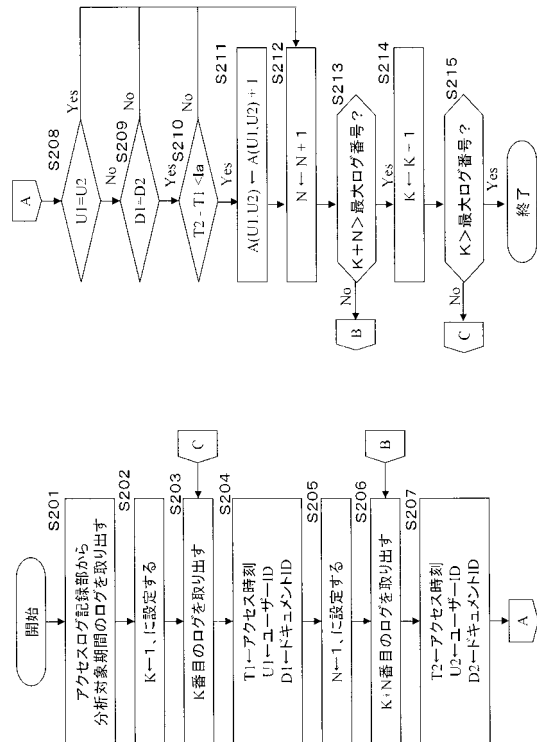
【 図 6 】



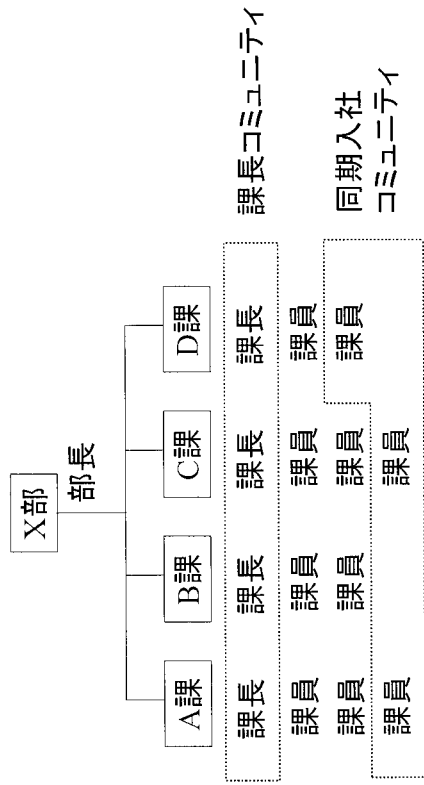
【 図 7 】



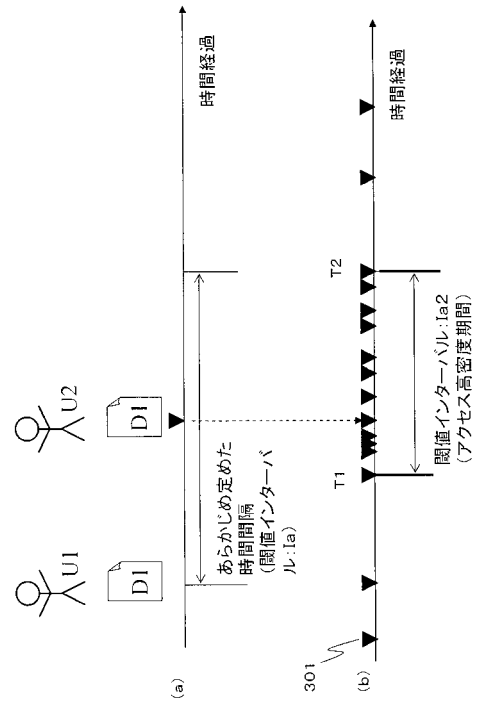
【 図 8 】



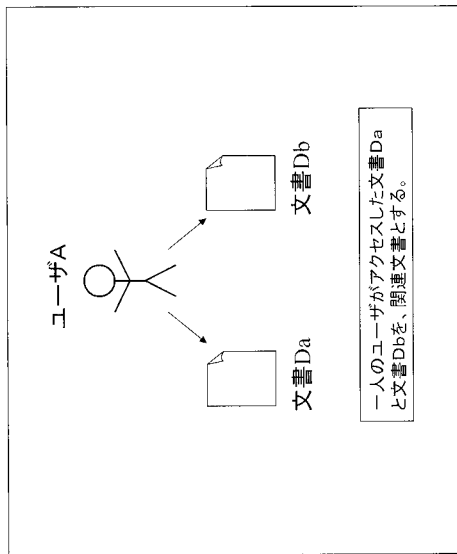
【 図 9 】



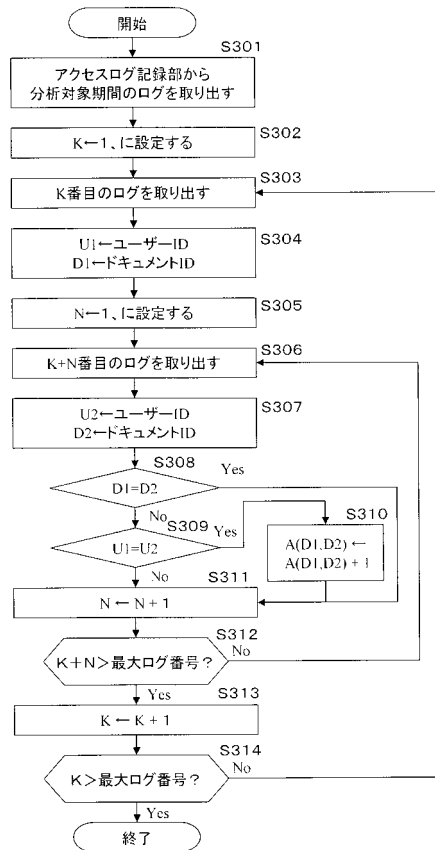
【 図 1 0 】



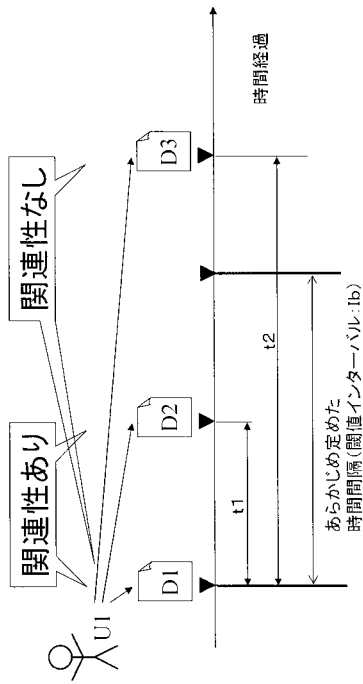
【 図 1 1 】



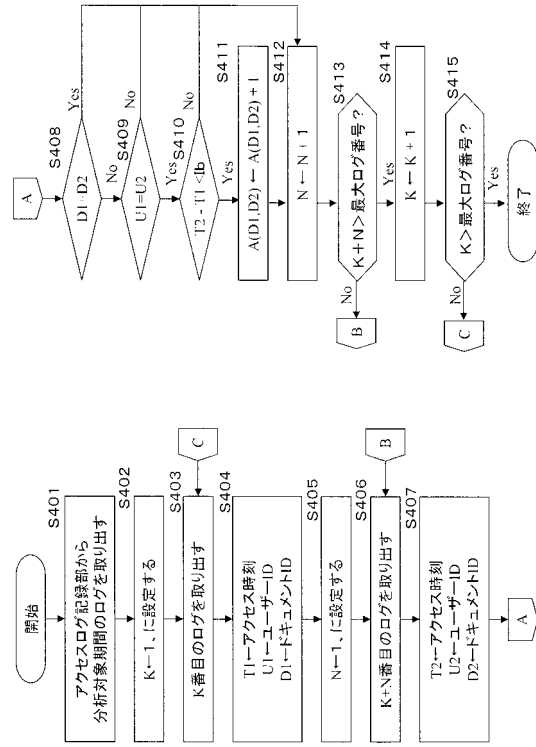
【 図 1 2 】



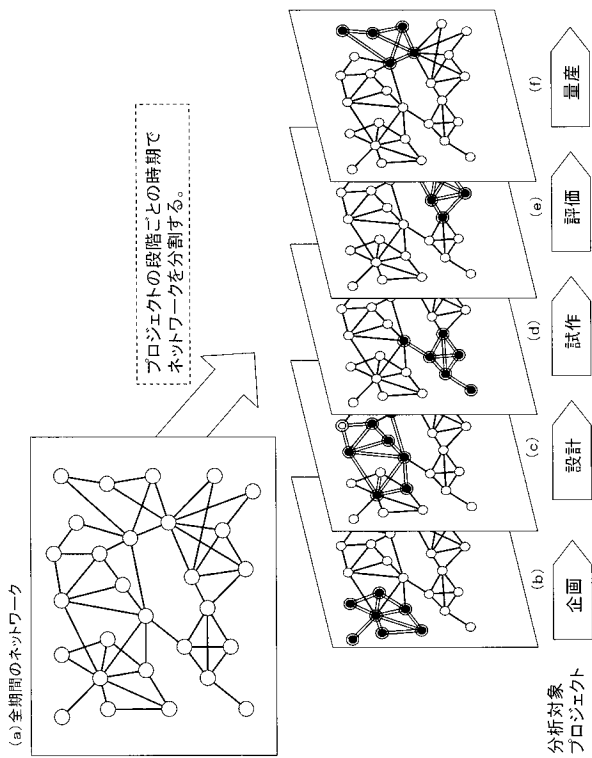
【 図 1 3 】



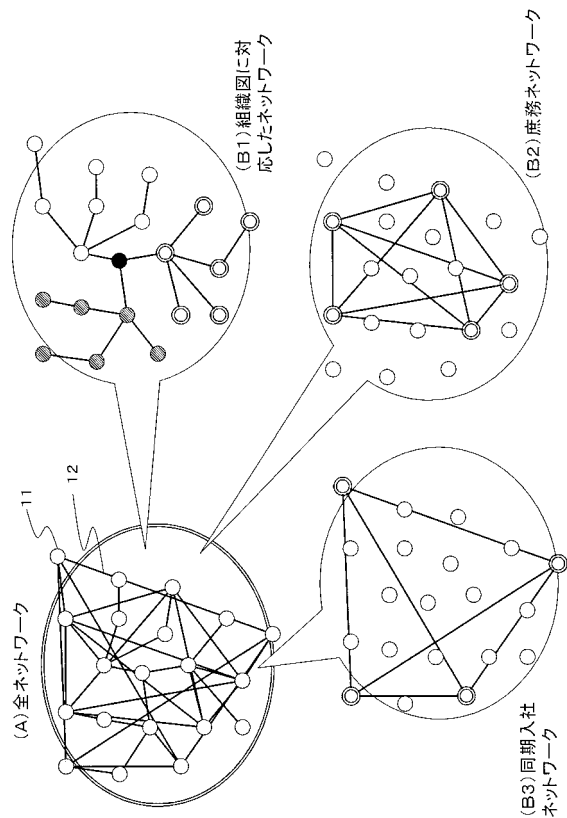
【 図 1 4 】



【 図 1 5 】



【 図 1 6 】



フロントページの続き

- (72)発明者 藤本 正和
神奈川県足柄上郡中井町境4 3 0 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 山崎 伸宏
神奈川県足柄上郡中井町境4 3 0 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 上野 裕一
神奈川県足柄上郡中井町境4 3 0 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 根本 啓一
神奈川県足柄上郡中井町境4 3 0 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 伊東 敦
神奈川県足柄上郡中井町境4 3 0 グリーンテクなかい 富士ゼロックス株式会社内