情報処理学会論文誌　Vol. 52　No. 1　24–32 (Jan. 2011)

機密情報追跡データの可視化による情報漏洩対策支援

中山佑輝[†1]　小崎真寛[†1]
芝口誠仁[†1]　岡田謙一[†2,†3]

情報漏洩対策として，管理者が機密情報の流れを把握することが重要となっている．そして，管理者による情報の流れの理解を迅速化することにより，漏洩を事前防止する効果や漏洩後の早急な原因究明の実現が期待される．そこで，本論文では機密情報の追跡データを可視化する手法を提案する．本手法は，組織内での機密情報の流れをマクロ視点からミクロ視点まで計 5 つの手法で可視化し，それらを互いに関連付けて提供することで，管理者がより迅速かつ正確に機密情報の流れを把握することを可能とする．提案コンセプトをもとに構築したプロトタイプシステム CROWS Up Viewer を用いた評価実験により，従来の表形式でのログ解析に比べ，解析の正確さおよび解析速度を向上させることができたことを確認した．また，各手法を切り替える際に生じる解析精度の劣化と解析の遅延を測定し，このオーバヘッドが解析に影響を及ぼさない程度にまで抑えられたことを示した．

A Supportive Method for Preventing Information Leakage with Visualization of Tracing Pathways of Confidential Information

Yuki Nakayama,[†1] Masahiro Kozaki,[†2]
Seiji Shibaguchi[†1] and Ken-ichi Okada[†2,†3]

In this paper, we describe a visualization technique for tracing confidential data. Recently, the increasing damage caused by information leakage has become a growing public concern. Some current studies aim to counteract information leakage by developing techniques which enable administrators to determine which hosts have confidential documents and the means by which secret information is transmitted, received, and duplicated. Although the techniques create large amount of log data, there are no systematized analytic techniques. Therefore, we aim to develop a scalable and seamless visualization framework for analyzing these log data that can meet the needs of diversely sized organizations and that can track various means of propagation seamlessly. Lastly, we verify through an experiment that our visualization method is helpful for providing accurate analysis, rapid analysis and is easy to understand.

1. はじめに

情報漏洩の脅威が深刻化しており[1)]，機密情報の所在を把握することや従業員の PC を用いた業務内容の監視を行うことが重要となっている．それにともない機密情報を追跡する技術が発展してきた[2)–6)]．当技術を用いることにより，追跡結果をログデータとして得ることができる．そして，このログデータを解析することによって，管理者は漏洩の危険性を事前調査することや漏洩原因の事後調査を行うことが可能となる．その際の解析作業の迅速さは事前調査では漏洩を未然に食い止められるか否かを左右し，事後調査では組織の信頼をどの程度維持できるかに大きく影響する．ゆえに，解析作業をより迅速かつ正確にすることが重要となる．

一方で，近年デジタルフォレンジックについて頻繁に議論されている．デジタルフォレンジックは「インシデント・レスポンスや法的紛争・訴訟に対し，電磁的記録の証拠保全および調査・分析を行うとともに，電磁的記録の改ざん・毀損などについての分析・情報収集などを行う一連の科学的調査手法・技術」と定義することができる[7)]．情報漏洩対策におけるデジタルフォレンジックは，(1) まず文書の交換履歴や操作履歴のログの保全を行い，(2) 次に漏洩原因究明のためにログの解析を行い，(3) その結果を証拠として提示する，という 3 つのフェーズからなる．本論文が対象とするログの解析支援はこの流れにおける「(2) 漏洩原因究明のためのログ解析」のフェーズにあたり，デジタルフォレンジックの観点からもログ解析の迅速化・正確化は重要である．

以上をふまえ，本論文では管理者による機密情報追跡ログの解析作業を支援する可視化手法を提案する．漏洩調査においては非常に多くの情報を処理しなければならず，可視化の実現にあたっては表示の煩雑化が問題となる．そこで，複数の可視化を併用することで可視化対象となる事象を分散させる解決策が考えられるが，可視化を複数併用することには解析時

†1 慶應義塾大学大学院理工学研究科
Graduate School of Science and Technology, Keio University
†2 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University
†3 科学技術振興機構
Japan Science and Technology Agency

に各手法を切り替えなければならないことでオーバヘッド（ログ解析の遅延やログ解析精度の劣化）が発生してしまうという問題が生じる．提案手法は，複数の可視化を階層的に分類することで可視化の煩雑化を防ぎ，かつ解析手法の切替えにともなうオーバヘッドを解析に影響を及ぼさない程度にまで抑えることが可能な手法である．

以降，2 章で関連研究を紹介し，3 章で提案可視化手法の概要について述べる．4 章では提案手法のプロトタイプシステムの詳細を記し，5 章で可視化手法の有用性および手法切替えによるオーバヘッドを評価する．最後に 6 章を本論文の結びとする．

2. 関 連 研 究

セキュリティの分野において可視化技術が注目されている．そして，単一のログを可視化する技術がその大半を占めている．代表的なものとして，見えログ[8] はシステムのログを記録する際に汎用的に用いられる syslog に対し，頻度情報やログのアウトラインなどを用いて 1 画面上に比較的長期のログを表示させることを可能とし，概要と詳細を把握しつつ調査を行うことを可能としている．一方で近年では，複数ログを可視化する手法も提案されている．たとえば，時系列視覚化システム[9] は，侵入検知に関係する複数のログを同時に表示することでログ間の関連性調査を支援する手法である．また，IDS ログのインシデント分析システム[10] では，インシデント分析システムにおける可視化システムの要求条件をまとめている．しかし，これらの中に複数の可視化を階層に分け上位の階層から掘り下げる形で解析を行っていくというものはなく，また，複数の手法を切り替えながら解析を行う際の切替えにかかるオーバヘッドを測定したものもない．

一方で，漏洩対策としてユーザの操作を監視し機密情報の追跡を実現する製品が多くのベンダから発表されている．ログの解析にフォーカスすると，これらの特徴は以下のようである．まず，ログデータを分析・統計し解析者に提示するものがある[11),12)]．これにより，解析者は組織全体をマクロ視点でとらえたユーザの行動やログの記録内容を視覚的に判断することが可能となる．次に，ログを表形式で解析者に提示し，解析者が検索・絞り込みを行っていくことで機密情報の追跡を行うものがある[11),13),14)]．特に後者に関して，既存の製品のほとんどが機密情報を追跡するにあたってはログを表形式で提示することで解析を支援している．しかし，我々は情報を追跡するためには情報の流れ（以下，情報の伝搬経路）をより迅速かつ正確に理解できることが重要であり，表形式でのログ解析では十分にそれを実現できていないと考える．

3. 機密情報の伝搬経路可視化手法

3.1 解決すべき課題

漏洩を未然に防ぐための現状把握を可能とするためには，ログデータの解析をより迅速かつ正確とすることで機密文書の所在や機密文書の所有者を迅速に把握できることが重要となる．これによって，本来所有すべきでない従業員が機密文書を所持していたり，PC の貸し借りによって機密文書を偶然社外に持ち出してしまっていたりといった状況を知ることが可能となる．このような状況を把握すること自体は既存のログ解析からも可能であるが，漏洩を未然に防ぐためには頻繁にログ解析を行わなければならないうえ，把握後に具体的な対策，たとえば該当する従業員に電話をかけ注意を促すなどの対応を迅速に行う必要がある．そのため，従来までの表形式でのログ提示に比べ可視化ログを用いて解析作業を支援することが有効となる．

また，事後調査に際しては，漏洩に関与したクライアント（漏洩した文書を所持していた履歴）の特定や，漏洩文書の外部記憶媒体へのコピー履歴の調査をより迅速に行えることが重要となる．先と同様，従来からの表形式でのログ参照でもこれらの事象を把握することは可能であるが，漏洩インシデント発生後の原因調査の迅速化は，関連組織への報告をより迅速なものとし信頼低下を最小限にとどめることにもつながる．そのため，漏洩に関与したクライアントの特定や，漏洩文書の外部記憶媒体へのコピー履歴の調査を行うためのログ解析を迅速化するための解析支援は非常に重要となる．

以上のことから，可視化によって従来の表形式による提示よりも迅速かつ正確なログ解析を実現することを提案の目的とする．

一方で，従来から可視化に際しては，なるべく多くの情報を可視化すべきという要求と可視化結果はなるべくシンプルで見やすいものであるべきであるという要求の相反が問題となる．我々の提案する機密情報の伝搬経路可視化手法は，中小企業（従業員数 300 人以下）やそれと同規模の大企業の 1 部署・1 支社を想定し，表示を煩雑化させずシンプルな可視化を実現することを 2 つ目の提案目的とする．

3.2 実 現 方 法

3.1 節に記した課題を解決するため，本提案手法では以下のような可視化の構成と表示方式をとる．

可視化対象の分類と可視化の構成

以下，可視化の分類について**表 1** に沿って詳述する．実際の漏洩調査においては，解析

表 1　可視化手法と可視化対象の対応
Table 1　Visualization methods and watch lists.

監視対象	可視化手法の名称	可視化対象
ネットワーク	グループベース手法	グループ間での文書交換
	ネットワークベース手法	あるグループに属するクライアント間での文書交換
クライアント	クライアントベース手法	機密文書の出入り
	ディレクトリベース手法	ディレクトリ間の伝搬
	アプリケーションベース手法	アプリケーションを用いた編集・複製

者がはじめから 1 台 1 台のクライアントを精査していく必要があるという状況は限られており，多くの場合は追跡対象の文書を受け取ったクライアントの絞り込みを行い，次にそれらのクライアントを順に精査していくこととなる．そこで，可視化の分類の仕方として，監視対象をネットワーク経由での文書交換とするものと，各クライアント内での文書の移動や複製とするものの 2 つに大別する．つまり，追跡対象文書を所持し得たクライアントを絞り込むためのネットワーク経由での可視化と，各クライアントを精査するための可視化に役割を分担させる．

そのうちの前者，ネットワーク経由の文書交換を可視化するにあたっては，クライアント数が増大するに従って表示が煩雑化してしまうという問題点がある．そこで，複数のクライアントの集合を 1 つのグループと定義し，グループ間での文書交換をグループベース手法として可視化し，各グループ内のクライアント間での文書交換をネットワークベース手法として可視化する．これにより，より多くのクライアントを 1 画面内に，表示を煩雑化させることなく収容することが可能となる．また，このグループベース手法とネットワークベース手法の階層化は，実際の組織の構成とも酷似しており，解析者は組織の構成と文書の流れを直感的に関連付けることが可能となる（**図 1**）．

漏洩調査に際しては先のグループベース手法とネットワークベース手法を用いて調査対象の文書を所持している，あるいは所持していたクライアントを特定する．その後，当該クライアントが USB メモリやプリントアウトなどによってクライアント外部へ文書を持ち出した経緯を調査することとなる．そこで本可視化手法では，クライアントベース手法として文書がクライアントを出入りした履歴のみを最もシンプルに可視化する．これにより，解析者の最大の関心事である当該クライアントが機密文書を外部へ持ち出した経歴の有無のみを容易に知ることが可能となる．そして，解析者によってさらなる精査が必要と判断されたクライアントのみ，さらなる調査を行う．ここで，クライアントの精査といっても各クライアント内での文書に対するユーザ操作は多岐にわたり，それらを 1 画面で表現することは非常

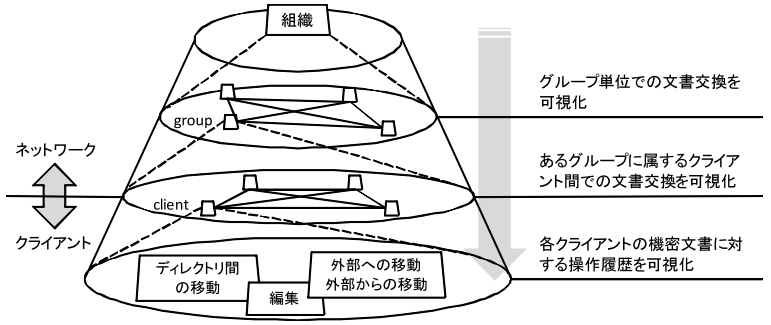

図 1　可視化の構成
Fig. 1　Structure of visualization.

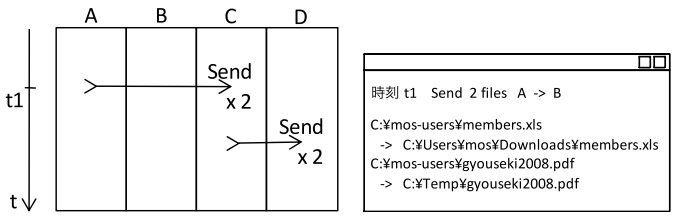

図 2　可視化の実現方法
Fig. 2　Basic visualization format.

に困難であり，表示の煩雑化を招くこととなる．そこで，クライアント内でのユーザによる文書操作が，ディレクトリ上での複製やディレクトリ間での移動と，アプリケーションを用いた編集・複製に大別できることから，本可視化手法では可視化対象をディレクトリ上での文書移動および複製とするものと，アプリケーションを用いた編集・複製とするものに分類し可視化を行う．つまり，各機密文書をソースとしたディレクトリ上での伝搬経路（移動・複製）をディレクトリベース手法によって可視化し，機密文書をアプリケーションで編集していた各時間帯に対する文書の編集・複製をアプリケーションベース手法で可視化する．これにより，ユーザ操作を直観的に把握することが可能となると同時に，可視化対象を分散させることで表示の煩雑化を防ぐことが可能となる．

表 示 方 式

提案手法では現在だけでなく過去の情報も可視化の対象とすることで，事前調査と事後調査の双方に貢献する可視化手法の実現を目指すため，**図 2** 左に示すような時系列に沿った

可視化を採用する．縦に時間軸をとり，横に情報をやりとりする主体を並べる．そして，その主体間における矢印で情報の交換を表現することで，操作が行われた時刻と情報の流れが明確となる．また，交換の対象となったデータや文書の詳細が表示からは欠落しているが，それらの情報は矢印と関連付けてダイアログ（図 2 右）などを用いて表示することで可視化の煩雑化を防ぎ，よりシンプルな可視化を実現する．

4. 伝搬経路可視化システム CROWS Up Viewer

本章では提案コンセプトをもとに構築した可視化システム CROWS Up Viewer（CROWS: Catch Reveal by Observing and WitneSsing）について詳述する．本アプリケーションは C++で構築され，Windows XP/Vista/7 上で動作を確認している．

4.1 システムアーキテクチャ

図 3 に CROWS Up Viewer のシステム構成を示す．本システムは可視化アプリケーションのインストールされた管理者 PC と監視データを格納するログ収集サーバ，グループ構成を格納したデータベース，および監視プログラムをインストールした監視対象となる従業員 PC から構成される．可視化プログラムはログ収集サーバに蓄積されたログを可視化し管理者に提供する．その際，可視化プログラムはデータベースを参照することにより，グループ構成の情報を得て，その情報を可視化に反映させる．なお，クライアントの操作監視ログと機密情報の伝搬追跡ログを収集するための研究は従来から多々行われているため[2)–6),11)–19)]，ここでは監視の実現方法およびログの収集に関しては議論の対象とはせず，既存の様々なベンダがリリースしている監視アプリケーションによって取得可能と判断できる情報からなるログを独自に仮定したものを使用した．なお，適応可能なログのデータ量に関しては，特に制約は設けておらず，ログが非常に膨大となると必要となる解析用 PC のメモリリソースが増大してしまうという問題がある．しかし，本論文において，提案手法の有用性を評価する目的では特に考慮する必要がないと判断したため，その問題に対して CROWS Up Viewer では特に対策は行っていない．実社会で導入する際には，ログデータを日付ごと，あるいは月ごとに分割し表示するなどの工夫で解決できる．

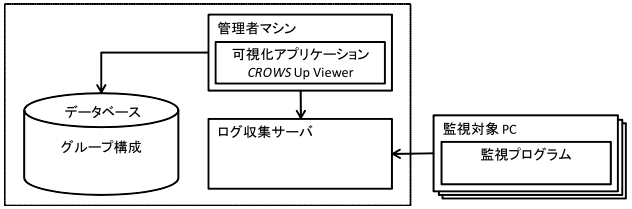

図 3　システム構成
Fig. 3　System architecture.

4.2 ウィンドウ構成

CROWS Up Viewer のアプリケーションウィンドウは，**図 4** のようにメインパネル，サブパネル，ツリービューで構成されている．以下ではこれら 3 つのパネルの表示について詳述する．

メインパネル

メインパネルには 5 つの可視化手法，すなわち時系列に沿った可視化を描画する．メインパネルに表示される 5 つの可視化手法の実現例を**図 5** に，CROWS Up Viewer で可視化の対象とした項目を**表 2** にそれぞれ示す．

サブパネル

サブパネルはメインパネルにおける解析を補佐するため，グループベース手法とネットワークベース手法の双方に対してシミュレーションモニタとトレーシングモニタを提供す

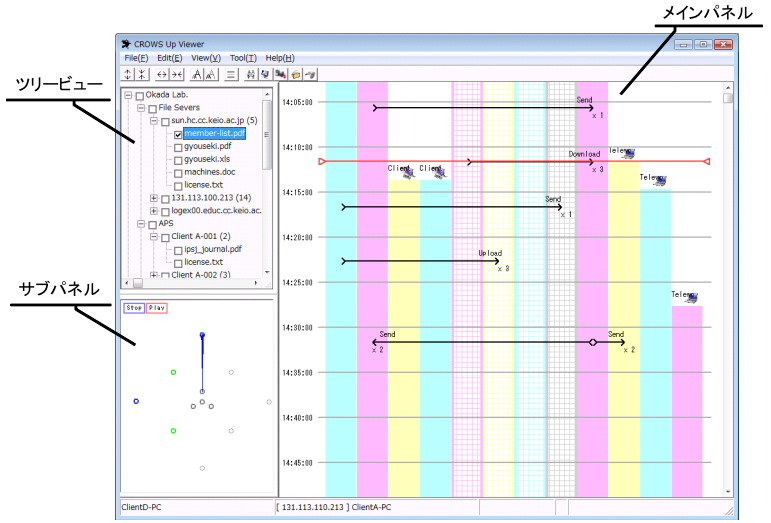

図 4　伝搬経路可視化システム CROWS Up Viewer
Fig. 4　Visualization system CROWS Up Viewer.

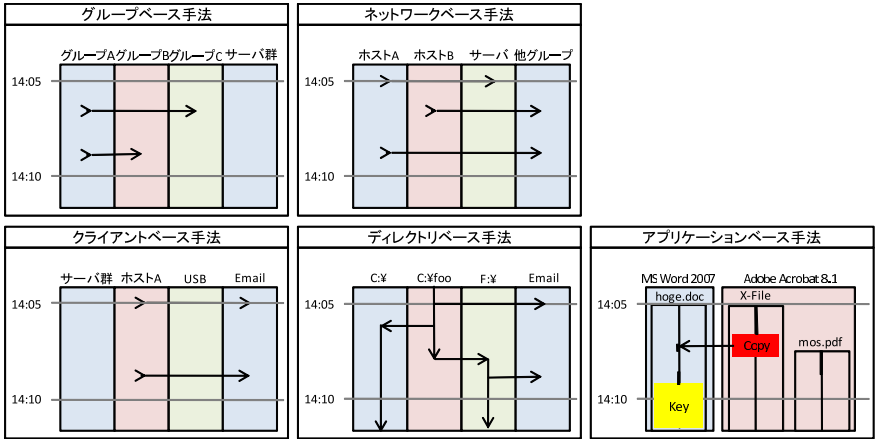

図 5　可視化手法ごとのメインパネルの表示
Fig. 5　Main panel display showing five visualization methods.

表 2　可視化対象
Table 2　Watch lists of five visualization methods.

可視化手法	可視化対象
グループベース手法	FTP を用いたダウンロード/アップロード，E メールを用いた送受信
ネットワークベース手法	FTP を用いたダウンロード/アップロード，E メールを用いた送受信
クライアントベース手法	FTP を用いたダウンロード/アップロード，E メールを用いた送受信，リムーバブルメディアへのコピー，磁気媒体への書き込み，プリントアウト
ディレクトリベース手法	機密文書のコピー，削除/復元，オープン/クローズ，Save As による複製，プリントアウト，アップロード，送信，圧縮/解凍，変換
アプリケーションベース手法	テキストデータやビットマップデータのコピー&ペースト，ウィンドウのポジション・サイズ・Z-index，キーボード入力，文書のオープン/クローズ，Save・Save As，プリントアウト

る．シミュレーションモニタでは**図 6** に示すように，情報交換の主体をノードで表し，ノード間を有向線で連結することで文書交換を表現する．そして，この表示を連続的に切り替えることで文書交換の様子をディスプレイ上でシミュレートする．シミュレーションモニタを用いることにより，文書交換の様子をより直感的に把握することが可能となる．一方のトレーシングモニタは 1 つの機密文書に着眼し，その文書のネットワーク経由での伝搬を示している（**図 7**）．線の太さで時間の遷移，色の変化で伝搬の分岐を表している．トレース

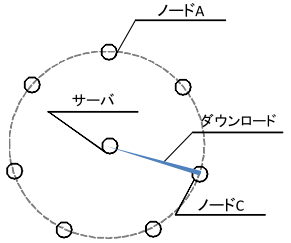

図 6　シミュレーションモニタ
Fig. 6　Simulation-based visualization method.

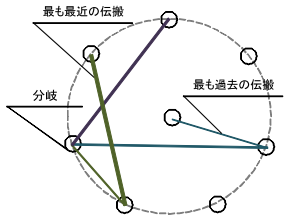

図 7　トレーシングモニタ
Fig. 7　Trace-based visualization method.

開始時刻と終了時刻を指定でき，特定期間中に文書交換に関与したグループおよびクライアントの直観的な特定を支援する．

ツリービュー

ツリービューでは各ファイルサーバ，およびクライアント内に存在する機密文書をリストアップする．ツリー上で解析対象となるグループやクライアント，あるいは文書の指定を行うことでメインパネルとサブパネルに表示される可視化手法の切替えが可能である．**図 8** に示すように，ツリーのトップを選択することでグループベース手法に表示を切り替えることができ，グループ名を指定することによりネットワークベース手法，クライアント名を指定することによりクライアントベース手法あるいはアプリケーションベース手法，文書名を指定することによりディレクトリベース手法に表示を切り替えることができる．さらに，チェックボックスを用いることにより，追跡を可視化する文書を絞り込むことが可能である．

4.3　適 応 環 境

提案可視化手法が，想定する組織規模である中小企業（従業員数 300 人以下）や，それ

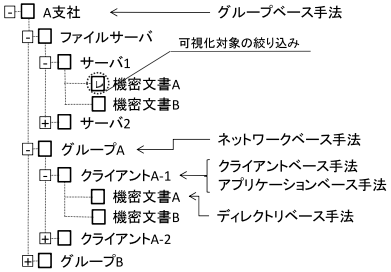

図 8　ツリービュー
Fig. 8　Tree view.

と同規模の大企業の 1 部署・1 支社に適応可能であるかを考察する．

収容可能なクライアント数（スケーラビリティ）に最も影響を及ぼす項目はディスプレイによる制約である．本可視化手法は，情報交換の主体を横並びに配置しているため，組織の規模拡大にともなって表示は横に拡大していくこととなる．一方で，縦方向には時間軸を設けているため可視化対象とする時間帯が長くなるにともなって表示は縦長となる．しかし，迅速なログ解析を実現するにあたって，縦・横の両スクロールを必要とすることで解析時の操作が煩わしくなってしまう．そこで，時間軸方向（縦）のスクロールを用いず可視化することは解析を行える時間帯を極端に制限してしまうため好ましくないことと，横スクロールを必要としてしまうことで文書交換を表現する矢印が 1 画面内に収まらずソース・ディスティネーションの関係が不明確となってしまうことから，横方向のスクロールを用いず可視化できることを可視化の設計目標とした．

ディスプレイ上に何台のクライアントおよびグループを横スクロールなしに表示可能であるか，実際に CROWS Up Viewer を用いて検証を行った．ディスプレイはサイズが 17 インチのもの，ログはネットワーク経由での文書交換を記録したもので，24 時間に文書交換が 1,000 回（時刻と送受信クライアントはランダムに決定）行われたものを使用した．このとき，グループベース手法にグループ 20 組，ネットワークベース手法にクライアント 20 台，計 400 台（20×20）のクライアントを横スクロールを要せず，また表示を煩雑化させることなく提示することができた．その際の使用メモリ容量は 7 MByte 強であり，表示の遅延による影響も特に感じることなく動作させることができた．このことから，本提案可視化手法は 400 台規模の組織規模，たとえば中小企業や大企業の 1 部署・1 支社まで適応可能といえる．

ところで，縦軸（時間軸）方向においても，表示対象のクライアント数やグループ数が多くなると矢印の表示が重なってしまうことも懸念される．しかし，上記の規模（20 クライアントまたは 20 グループ）で機密文書交換のタイミングが重なり，かつその交換が表示の重なる位置での交換であることの偶然さを考慮すると，縦方向（時間軸方向）のスケーラビリティは横方向に比べるとクリティカルに解析に影響を及ぼすものではないと考えられる．また，CROWS Up Viewer では縦方向のスケールを解析者が自由に変更できることから，時間軸方向に拡大することで送受信のタイミングが秒単位で異なっていれば表示をずらして表示させることが可能である．さらに，マウスオーバによって，その箇所の詳細を表示する機能も搭載しており，一部の表示が重なった場合でも，重なっていない部分をマウスオーバさせることで解析者は詳しい情報を知ることが可能となっている．なお，上記の横方向でのスケーラビリティをテストしたログを使用した場合でも矢印が重なることはなかった．

次に，クライアントベース手法，ディレクトリベース手法，アプリケーションベース手法においても考察する．これらの手法において縦軸方向は，1 人のユーザが同時に複数の操作を行えないと考えると，矢印が重なることはない．また可視化方式の統一性の面から，この 3 つの手法においても，縦軸方向ではなく横軸方向でスケーラビリティを考慮すべきである．まず，クライアントベース手法は，横に並ぶ情報交換の主体がリムーバブルメディアや E メールなどに限られるため，スケールを考慮する必要はない．ディレクトリベース手法では，1 つの機密文書から派生した文書を 20 以上のディレクトリに分散させて保存している状況は非常に珍しいケースであると考えると，十分に収容可能といえる．アプリケーションベース手法は機密文書を編集している時間帯ごとに可視化を提供しており，同時に 20 以上の機密文書をアプリケーション上で編集している状況を珍しいケースと考えると，十分に収容可能である．

最後に，ログの取得に関して，既存の製品 SKYSEA Client View [20] は，PC 1,400 台からのログ取得を実現しており，提案手法の想定する 400 台規模の組織においてログを収集することは十分に実現可能である．

5. 評価実験

5.1 実験内容

可視化手法の有用性および可視化手法の切替えによる効果を検証するための評価実験を行った．被験者は情報工学を専攻する大学生・大学院生 10 名である．1 つのログと複数のテキストによる記述（以下，問い）を提示し，被験者は問いの内容がログの内容と一致する

か/一致しないかを回答する．問いは「14:59 ClientA-3 から ClientA-1 に送信された文書は 15:03 Group B の ClientB-1 に送信された」「15:08 に Removable media からコピーされた mos.pdf は，その後，複製はされずにプリントアウトされている」といった複数ログにまたがる解析を必要とするものを提示した．

出題数は合計 64 題であり，その出題内容は用いるログの内容とログの提示方式によって以下のように分類される．まず，出題するログの内容は以下の 4 つに大別できる．

(A)　ネットワーク経由での文書交換に関する 11 行のログ
(B)　ネットワーク経由での文書交換および文書のクライアント出入りに関する 8 行のログ
(C)　1 つのクライアント内における文書の移動・複製に関する 10 行のログ
(D)　1 つのクライアント内における文書の移動・複製およびアプリケーションを用いた文書の編集に関する 13 行のログ

また，それぞれのログに対して提示方式は以下の 4 通りである．

(1)　表形式（テキストログ）で提示したもの
(2)　複数の可視化手法で可視化したもの（可視化ログを被験者が自由に切替え可能）
(3)　複数の可視化手法で可視化したもの（可視化手法の切替えを制限し理想的な切替えのみ可能）
(4)　複数の可視化手法を 1 つの手法（1 画面）に統合して表示したもの

つまり (A) のログデータは提案手法ではグループベース手法とネットワークベース手法によって提示されるログであり，それを (1) テキストログとして解析した場合と，(2) 提案手法の各手法を適宜被験者自身で切り替えて解析する場合，(3) 提案手法で各手法の切替えを理想切替えのみ実行可能とした場合，(4) 切替えを用いず 1 画面で表示させた場合，以上の 4 つの場合を比較・検討する．同様に，(B) のログはネットワークベース手法とクライアントベース手法で提示され，(C) ではクライアントベース手法とディレクトリベース手法，(D) ではディレクトリベース手法とアプリケーションベース手法で提示されるログである．これらの結果から各手法間での切替えにかかるオーバヘッド（遅延と誤認識）を測定する．

各ログの各提示方式に対して問いは 4 題出題した（被験者あたりの問い出題数は計 64 題）．そして，得られた正答率と所要時間を用いて，解析がテキストログと比較して正確かつ迅速なものとなったか，被験者が適切に可視化手法を切り替えて解析を実行することができたかを検証する．

なお，本実験では被験者に対して問いを提示する前に，先にログデータのみを提示し，被験者がこれから解析を行うログデータを十分に理解する時間を設けた．これは解析順によって後半の提示方法に対する解析に慣れてしまうことを抑えるためである．

また，本実験では独自に仮定した非常に短いログデータを用いている．独自の実験用ログを仮定して用いた理由は，一般的なログの形式というものがないためと可視化ログとの公平性を保つためである．文書交換やユーザ操作を監視しログを取得するアプリケーションは様々なベンダからリリースされているが，そのログ形式は各ベンダあるいは各アプリケーションによって異なっており統一されていない．また，テキストログから得られる情報量と可視化ログの情報量が等しい状況で実験を行うことで対等な条件下で解析精度と解析時間の測定が可能となる．しかし，既存のツールからそのようなログを得ることはできなかった．ゆえに，独自のログを仮定し使用することとした．短いログを用いた理由としては，可視化手法どうしの比較において，被験者による各手法の切替えが解析結果により大きく反映され，切替えの効果が測定しやすくなるという点があげられる．また，一般的に可視化技術を用いる利点として，テキストログの解析に比べ，膨大な量のログ情報をより直観的に理解できる点があげられているが，一般的には可視化技術が優位とはされていない非常に短いログを用いた解析においても可視化の有用性を示すことには価値があると考えたためである．

5.2　結果と考察

正答率の結果（平均値と標準偏差）を**表 3** に，所要時間の結果を**表 4** にそれぞれ示す．正答率において (A)，(B) のログに関しては提示方式による差はほとんど生じなかった．

表 3　ユーザ実験の結果–正答率（平均値と標準偏差）
Table 3　Accuracy of the experiment (average and standard deviation).

		提示方式			
		(1)	(2)	(3)	(4)
ログの種類	(A)	0.98 (0.08)	1 (0)	1 (0)	0.98 (0.08)
	(B)	1 (0)	1 (0)	1 (0)	1 (0)
	(C)	0.75 (0.22)	1 (0)	1 (0)	0.89 (0.13)
	(D)	0.83 (0.18)	1 (0)	1 (0)	0.78 (0.23)

表 4　ユーザ実験の結果–所要時間（平均値と標準偏差）
Table 4　Time of the experiment (average and standard deviation).

		提示方式			
		(1)	(2)	(3)	(4)
ログの種類	(A)	122 (20)	81 (13)	61 (11)	59 (11)
	(B)	101 (12)	87 (12)	72 (23)	64 (4)
	(C)	174 (15)	89 (11)	68 (18)	94 (8)
	(D)	177 (25)	65 (10)	49 (19)	79 (15)

しかし，(C)，(D) では (1) の表形式での提示に正答率の低下が確認されたほか，(D) においては (4) の 1 画面での提示にも正答率の低下がみられた．その理由として，(C)，(D) のログは (A)，(B) に比べて可視化対象となる事象が多いため，表形式のログでは直感的な理解が困難となり可視化ログとの差分が顕著に現れたことが考えられる．(D) ではアプリケーションの操作履歴が含まれているため特に表示対象が多い．そのため，(4) の可視化を用いた 1 画面での提示において表示の煩雑化が起こり，被験者の理解を妨げたものと考えられる．なお，本実験で用いたログは非常に小規模のものであるにもかかわらず，表示の煩雑化による解析精度の劣化がみられたことから，大規模なログではこの差分はより顕著に現れると予測される．

解析時間においては，いずれのログに対しても可視化した提示方式を用いることで解析時間が短縮できている．このことから，従来の表形式でのログ解析に比べ，可視化したログを用いて解析を行うことで解析を迅速化できることが分かる．

また，(A)，(B) のログにおいては，1 画面の可視化による提示方式が最も迅速に解析を行えている．そして，(2) 自由切替えの提示と (4) 1 画面の提示の間には有意水準 5%で有意差がみられたが，(3) 理想切替えの提示と (4) 1 画面の提示の間には有意水準 5%で有意差が確認できなかった．つまり，手法の切替えに不慣れな被験者では手法の切替えに際して解析の遅延が生じてしまう．しかし，理想切替えが可能な状況，つまり被験者が切替えを理想的に行えるまで習熟した状況では，切替えによる解析の遅延は生じないといえる．

一方で，(C)，(D) のログにおいては，(4) 1 画面による提示よりも (3) 理想切替えによる提示が早く解析を終えている（$p < .05$）．また，自由切替えと 1 画面を比較すると，(C)，(D) ともに自由切替えが早く解析を終えている（(C) においては有意差は確認できなかったが，(D) では有意差が確認できた（有意水準 5%））．1 画面提示の優位性が損なわれた理由としては，(C)，(D) は (A)，(B) に比べて可視化対象が多いため，1 画面で提示した際に表示が煩雑化してしまい被験者の理解を妨げてしまったと考えられる．そして，特に (D) では表示の煩雑化が起こり，その影響が結果に現れる結果となった．このことは，表示の煩雑化を防ぎシンプルな表示を行うことで解析速度の向上につながることも示唆している．

以上のことから，まず従来の表形式でのログ解析に比べ，可視化ログを用いた解析が迅速かつ正確な解析を実現できたといえる．また，提案可視化手法では複数の可視化を階層的に組み合わせて提供したが，手法を切り替えることにともなうオーバヘッドを抑え，切替えに十分に慣れた解析者にとっては 1 画面での提示と同等あるいはそれ以上の解析速度を実現した．

6. おわりに

情報漏洩対策として，情報の追跡が注目されている．情報の追跡によって得たログデータを解析することで，管理者は追跡結果を知ることが可能である．しかし，従来ログデータは表形式で参照されることがほとんどであり，解析を十分に支援できていなかった．そこで，本論文では複数の可視化を階層的に分類することで，機密情報の流れをマクロ視点からミクロ視点へ切り替えながら解析を行うことで，可視化の煩雑化を防ぐ可視化手法を提案した．プロトタイプシステム CROWS Up Viewer を用いた評価実験からは，従来の表形式でのログ解析に比べ解析の正確さおよび解析速度が向上できたことを確認した．また，各手法を切り替える際に生じる解析精度の劣化と解析の遅延の測定を行い，このオーバヘッドを解析に影響を及ぼさない程度にまで抑えられることを示した．

参考文献

1) 独立行政法人情報処理推進機構：情報セキュリティ白書 2008 第 II 部：10 大脅威 ますます進む『見えない化』(2008.5).
2) 箱守　聰，横山和俊，乃村能成，谷口秀夫：オープン処理に着目した情報拡散追跡法，情報処理学会研究報告 2005-OS-99，Vol.2005-OS, No.99, pp.127–134 (2005).
3) 大橋　慶，箱守　聰，田端利宏，横山和俊，谷口秀夫：機密情報の拡散追跡機能を利用した書き出し制御手法，マルチメディア・分散・協調とモバイル（DICOMO2007）シンポジウム論文集，pp.690–697 (2007).
4) 大橋　慶，田端利宏，谷口秀夫，横山利宏，箱守　聰：機密情報の拡散追跡機能における情報漏えいの検知精度の向上手法，情報処理学会研究報告 2008-DPS-134，2008-CSEC-40，pp.97–102 (2008).
5) 植村晋一郎，田端利宏，谷口秀夫，横山和俊，箱守　聰：機密情報の拡散追跡のソケット通信への適用手法，マルチメディア・分散・協調とモバイル（DICOMO2008）シンポジウム論文集，pp.768–775 (2008).
6) Sky 株式会社：ファイル監視装置およびファイル監視プログラム，公開特許公報（A），特開 2009-15659 号公報 (2009).
7) 辻井重夫，萩原栄幸：デジタルフォレンジック辞典，デジタルフォレンジック研究会 (2006).
8) 高田哲司，小池英樹：見えログ：情報視覚化とテキストマイニングを用いたログ情報ブラウザ，情報処理学会論文誌，Vol.41, No.12, pp.3265–3275 (2000).
9) 江端真行，小池英樹：不正侵入調査を目的とした複数ログの時系列視覚化システム，情報処理学会論文誌，Vol.47, No.4, pp.1099–1107 (2006).
10) 松本文子，堀　良彰，力武健次，馬場俊輔，鈴木和也，中尾康二：ネットワークインシ

デント分析システム構築運用におけるユーザインタフェースの検討，*SCIS2006* (2006).
11) FineArt Technology Co., Ltd. (online), available from ⟨http://www.fineart-tech.com/jp/⟩ (accessed 2010-3-22).
12) e-Tracker5 – e-System, corporation (online), available from ⟨http://www.e-system.co.jp/etr5/index.html/⟩ (accessed 2010-3-22).
13) MOTEX Inc. (online), available from ⟨http://www.motex.co.jp/index.html⟩ (accessed 2010-3-22).
14) InfoCage – NEC Corporation (online), available from ⟨http://www.nec.co.jp/cced/infocage/index.html⟩ (accessed 2010-3-22).
15) 株式会社日立製作所：アクセス状況監視システム，公開特許公報（A），特開 2008-310417 号公報 (2008).
16) 日立ソフトウェアエンジニアリング株式会社：ログファイルの送信システム及びその方法，公開特許公報（A），特開 2007-323320 号公報 (2007).
17) 富士ゼロックス株式会社：操作管理装置，操作内容判定方法，操作管理プログラム，操作管理システム，およびクライアント端末，公開特許公報（A），特開 2005-227866 号公報 (2005).
18) 三菱電機株式会社：情報管理システム及び端末装置及び情報管理方法及びプログラム，公開特許公報（A），特開 2008-225830 号公報 (2008).
19) Sky 株式会社：ファイル監視装置およびファイル監視プログラム，公開特許公報（A），特開 2009-15659 号公報 (2009).
20) SKYSEA Client View（オンライン），available from ⟨www.skyseaclientview.net/case/case015.html⟩ (参照 2010-7-25).

(平成 22 年 4 月 7 日受付)
(平成 22 年 10 月 4 日採録)

中山　佑輝（学生会員）
2009 年慶應義塾大学理工学部情報工学科卒業．現在，同大学大学院理工学研究科修士課程に在籍．ネットワークセキュリティ，ディジタルフォレンジックスに関する研究に従事．

小崎　真寛（学生会員）
2010 年慶應義塾大学理工学部情報工学科卒業．現在，同大学大学院理工学研究科修士課程に在籍．ネットワークセキュリティ，ディジタルフォレンジックスに関する研究に従事．

芝口　誠仁（学生会員）
2008 年慶應義塾大学理工学部情報工学科卒業．2010 年同大学大学院理工学研究科修士課程修了．ネットワークセキュリティ，ディジタルフォレンジックスに関する研究に従事．2010 年任天堂株式会社入社．

岡田　謙一（フェロー）
慶應義塾大学理工学部情報工学科教授，工学博士．専門は，CSCW，グループウェア，ヒューマン・コンピュータ・インタラクション．情報処理学会誌編集主査，論文誌編集主査，GW 研究会主査等を歴任．現在，情報処理学会 MBL 研究会運営委員，BCC 研究グループ主査，日本 VR 学会理事，CS 研究会委員長．情報処理学会論文賞（1996，2001，2008 年），情報処理学会 40 周年記念論文賞，日本 VR 学会サイバースペース研究賞，IEEE SAINT'04 最優秀論文賞を受賞．情報処理学会フェロー，IEEE，ACM，電子情報通信学会，人工知能学会各会員．