機密情報の拡散経路を可視化する機能の提案

福島 健太　　山内 利宏　　谷口 秀夫

岡山大学大学院自然科学研究科
700-8530 岡山市北区津島中 3 丁目 1 番 1 号
fukushima@swlab.cs.okayama-u.ac.jp　{yamauchi,tani}@cs.okayama-u.ac.jp

あらまし 機密情報が外部へ漏えいする事例が増加している．我々は，機密情報が拡散する契機となるシステムコール発行に着目し，機密情報が拡散する経路を追跡し，外部への漏えいを検知する機能を提案し，機密情報の拡散追跡機能として実現した．しかし，提案機能がテキスト形式のログで提供する情報は，利用者が機密情報の拡散状況を正確に把握するには不足している．また，ログは，機密情報の拡散が起こるたびに追加されるため，利用者が必要とする情報を探すには多くの手間がかかる．本稿では，拡散経路に関するログも出力する処理を実現し，有向グラフ形式で利用者の指定する情報を表示する可視化機能を提案する．また，提案機能がログを可視化した実例を示す．

Proposal of Function to Visualize Diffusion Path of Classified Information

Kenta Fukushima　　Toshihiro Yamauchi　　Hideo Taniguchi

Graduate School of Natural Science and Technology, Okayama University
3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530 JAPAN

Abstract The number of incidents leaking of classified information has increased. We proposed the function to trace the classified information diffusion and detect a leakage of classified information. However, it is insufficient to provide information about diffusion path of classified information for users because some information is not shown. Because information about diffusion path is shown by the text log, it is not easy for users to understand necessary information. In this paper, we propose a function to visualize diffusion path of classified information. In addition, we show the example of the log visualized by this function.

1　はじめに

機密情報の電子データ化により，機密情報が可搬可能な記憶媒体やネットワークを経由し，計算機外部へ漏えいする事例が増加している．情報漏えいの主な原因は，計算機の誤操作や管理ミスといった内部的要因が多く，情報漏えい事例の約 70 %を占めている [1]．

我々は，機密情報が拡散する契機となるシステムコール発行に着目し，機密情報が拡散する経路を追跡し，計算機外部への漏えいを検知する機能 [2]（以降，機密情報の拡散追跡機能と呼ぶ）を提案した．機密情報の拡散追跡機能は，機密情報の漏えいを計算機利用者へ通知するため，利用者は，計算機外部への機密情報の書き出しを制御できる．

しかし，機密情報の拡散追跡機能は，機密情報の拡散の検知や機密情報の管理に必要ない情報は取得せず，これに関するログも出力しない．また，利用者に提供される機密情報の拡散状況は，テキスト形式のログもしくは取得した情報から作成されるファイルとプロセスの一覧のみである．このため，既存のログと一覧では，機密情報の拡散状況を正確に把握できない．また，ログは，機密情報の拡散が起こるたびに追加されるため，利用者が必要とする情報を探すには多くの手間がかかる．したがって，機密情報の拡散状況を正確に把握できる情報を利用者に提供し，表示する必要がある．

本稿では，拡散経路に関するログも出力する処理を実現し，有向グラフ形式で利用者の指定する情報を表示する可視化機能を提案する．また，提案機能がログを可視化した実例を示す．

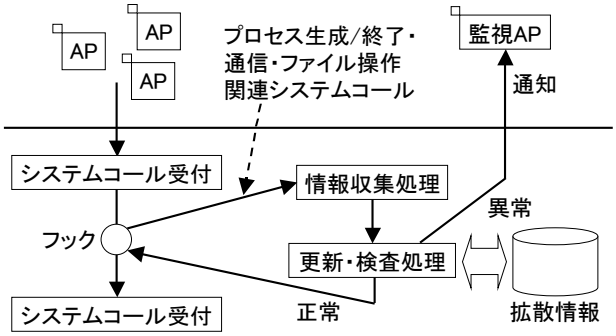

図 1 情報漏えい検知の基本機構

2 機密情報の拡散追跡機能

2.1 情報の拡散経路

機密情報はファイルの形式で存在し，どれが機密情報に相当するかは利用者が指定していると仮定する．以降では，機密情報を有する可能性があるため，機密情報の拡散追跡機能が拡散情報として管理しているファイルとプロセスを管理対象ファイルと管理対象プロセスと呼ぶ．

機密情報の拡散は，機密情報をプロセスが読み込み，さらに他のプロセスやファイルへその内容を伝えることで行われる．機密情報の拡散追跡機能が追跡する 3 つの経路を以下に説明する．

(1) ファイル操作

機密情報が拡散するファイル操作として，ファイルの読み込みとファイルへの書き出しがある．プロセスが管理対象ファイルを読み込んだ場合，プロセスを管理対象とする．また，管理対象プロセスがファイルに情報を書き出した場合，ファイルを管理対象とする．

(2) プロセス間通信

管理対象プロセスは，プロセス間通信により，他のプロセスに情報を伝達できる．このため，プロセス間通信を監視する．送信元プロセスが管理対象であるとき，送信を仲介する通信用資源と受信プロセスを管理対象とする．

(3) 子プロセスの生成

子プロセスの生成時に，子プロセスが親プロセスの資源を引き継ぐ機能がある場合（UNIX の fork 処理など），この機能によりプロセス間で情報が拡散する．したがって，親プロセスが管理対象の場合，子プロセスも管理対象とする．

2.2 基本機構

図 1 に機密情報の漏えいを検知する基本機構を示し，以下に説明する．

表 1 ログの種類

通番	ログの種類
(1)	管理対象でないファイルなどが管理対象になった際のログ
(2)	管理対象の消去などにより管理対象から外れた際のログ
(3)	管理対象ファイルが外部デバイスに書き出された際のログ
(4)	管理対象ファイルが外部計算機に送信された際のログ

表 2 ログの保持情報

保持情報	情報を含むログの種類
ログの動作が行われた時刻	(1) ～ (4)
ログの動作が行われた対象を特定するための情報	(1) ～ (4)
ログの動作を行った対象を特定するための情報	(1) ～ (4)
ファイルが書き出された外部デバイスのデバイス番号	(3)
送信先計算機の IP アドレス	(4)
送信先計算機のポート番号	(4)

(1) 機密情報の拡散契機となるシステムコールをフック

(2) 情報収集処理では，機密情報の拡散追跡に必要な情報を取得

(3) 更新・検査処理では，取得した情報をもとに拡散情報を更新し，漏えいの可能性を検査

(4) 検査によって漏えいの可能性（異常）が発見されると，監視アプリケーションプログラム（以降，監視 AP と呼ぶ）へその旨を通知

(5) 検査によって漏えいの可能性がない場合（正常）は，システムコールフック元へ復帰

上記の処理（4）が行われた場合，監視 AP は，漏えいの可能性に関する警告を利用者に表示する．このとき，利用者が漏えいを検知された機密情報の書き出しの可否を判断することで，書き出しを制御できる．

2.3 機密情報の拡散状況を確認する方法と課題

利用者が機密情報の拡散追跡機能から機密情報の拡散状況を確認する方法は以下の 2 つがある．

(1) 機密情報の拡散検知時に記録されるログから確認

(2) 機密情報の拡散追跡機能が提供するメンテナンスツールから確認

表 1 に出力されるログの種類を示し，表 2 にログが保持する情報の一覧を示す．ログからは，管理対象に関する情報と共に機密情報の拡散状況を確認できる．また，メンテナンスツールは，拡散情報から情報を取得し，管理対象プロセスと管理対象ファイルの一覧をそれぞれ表示する．しかし，既存の確認方法は，以下に述べる 3 つの課題があるため，利用者が機密情報の拡散状況を確認するのは困難である．

（課題 1） ログに表示される情報の不足
（課題 2） ログとメンテナンスツールに表示される情報の誤り
（課題 3） 機密情報の拡散状況の視認性が低い

（課題 1）については，機密情報の拡散追跡機能は，ファイルやプロセスへの機密情報の拡散を検知し，管理対象として管理する機能であり，拡散の検知や管理対象の管理に必要ない情報は取得せずログにも出力されない．つまり，拡散の範囲のみを把握するのに必要な情報のみをログに出力する．このため，既存のログだけでは，利用者が機密情報の拡散状況を正確に把握するには情報が不足している．また，機密情報の拡散追跡機能のログは，機密情報が拡散した過程の一部しかたどることができない．

（課題 2）については，管理対象ファイルのファイル名が変更されたとき，拡散情報に登録されているファイル名は更新されない．このため，メンテナンスツールは，誤った情報を表示する．また，ファイル名が変更されたことは，ログにも表示されないため，利用者は確認することができない．

（課題 3）については，機密情報の拡散状況を確認するには，ログを時系列順に読んでいく必要があり，ログが増えることで 1 つの管理対象の情報が数百行に渡って分散されて表示されることもある．このため，管理対象プロセスと管理対象ファイル間の機密情報の拡散関係を正確にかつ短時間で把握することが難しい．

3　機密情報の拡散追跡機能における可視化機能

3.1　可視化の目的

以下に可視化の目的を述べる．

（目的 1） 機密情報の書き出し検知時に適切な情報を表示し，利用者の判断ミスによる機密情報の漏えいを防止する．
（目的 2） 現在の機密情報の拡散状況を利用者が視覚的に確認できるようにし，計算機内で機密情報が意図していない拡散をしていないかを容易に確認できるようにする．
（目的 3） 機密情報が計算機外部へ漏えいした際，漏えい経路を表示し，迅速に情報漏えいの原因を調査できるようにする．

3.2　設計方針

3.1 節の目的を達成するには，2.3 節で述べた課題を解決する必要がある．本稿では，拡散経路に関するログも出力する処理を実現し，有向グラフ形式で利用者の指定する情報を表示する可視化機能を提案する．本機能に求められる 2 つの要件を以下に述べる．

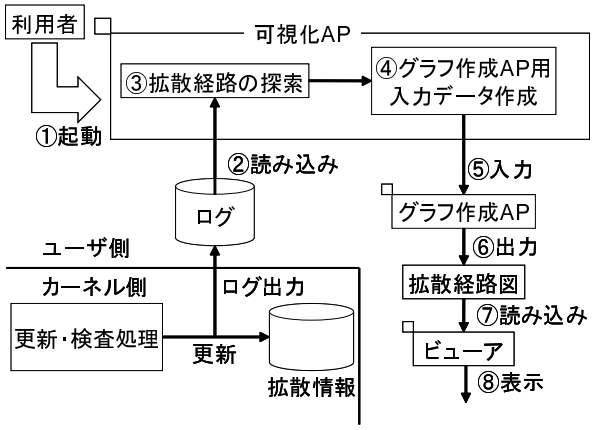

図 2　可視化機能の基本機構

（要件 1） 表示する情報に漏れと誤りがないこと
（要件 2） 表示する情報が簡潔であり，機密情報の拡散状況を容易に把握できること

（要件 1）を満たすには，必要な情報をログに追加し，拡散情報を正確に更新することで（課題 1）と（課題 2）を解決する必要がある．

（要件 2）を満たすには，利用者に表示する情報を必要最低限のものとし，機密情報の拡散状況を明確に表示することで（課題 3）を解決する必要がある．

3.3　基本機構

3.3.1　処理の流れ

図 2 に可視化機能の基本機構を示し，図中の番号に対応した処理の流れを以下に述べる．

(1) 利用者が可視化 AP を起動
(2) 可視化 AP がテキスト形式のログを読み込み
(3) 読み込んだログから機密情報の拡散経路を探索
(4) 探索した結果から，グラフ作成 AP への入力データを作成
(5) グラフ作成 AP へデータを入力
(6) グラフ作成 AP が拡散経路図を出力
(7) 利用者がビューアを起動し（6）で出力した図を読み込み
(8) 拡散経路図を利用者に表示

上記の処理のように，ログから機密情報の拡散経路を探索することにより，ログに漏れと誤りがない限り，（要件 1）を満たせる．なお，ログに漏れと誤りをなくすための対処は 4 章で述べる．また，拡散経路探索時に利用者が指定した管理対象に関する情報のみを探索して表示する．これにより，表示する情報が簡潔になるため，（要件 2）を満たせる．

3.3.2 グラフ作成 AP

グラフ作成 AP は，可視化 AP が出力したデータから機密情報の拡散経路を表す有向グラフを作成する．有向グラフは，表示する管理対象の数によって図の大きさや各ノードの配置を変える必要がある．このため，表示する管理対象の数に合わせて図のレイアウトを自動で調整できるフリーウェアのグラフ作成 AP である Graphviz [3] を使用する．Graphviz は，DOT 言語と呼ばれる言語で記述されたテキスト形式のファイルを読み込むことで有向グラフを作成できる．

3.4 機密情報の拡散状況の表示方法

3.4.1 表示方法

可視化機能では，機密情報の拡散経路を拡散経路図として表示する．拡散経路図は，利用者が機密情報の拡散状況を容易に把握できるようにするため，表示する情報を整理し，利用者に必要な情報のみを表示する必要がある．そこで，可視化機能は，表示する情報にフィルタをかける機能を提供する．フィルタの種類は，以下に示す 3 つである．

(1) 指定したファイルから拡散した機密情報の拡散経路のみを表示
(2) 指定したファイルに拡散した機密情報の拡散経路のみを表示
(3) 指定した時刻の間の機密情報の拡散経路のみを表示

また，これらのフィルタを組み合わせて機密情報の拡散経路を表示することもできる．

3.4.2 拡散経路図の構成

拡散経路図は，有向グラフで表示する．有向グラフは，ノード（接点）とエッジ（辺）からなる図形でエッジに向きを表す矢印がついている．可視化機能では，管理対象ファイルと管理対象プロセスをノードで表現し，機密情報が拡散した向きをエッジで表現する．これにより，ログでは分散されていた 1 つの管理対象に関する情報は，1 つのノードに集約して表示される．以下に，拡散経路図に表示する情報について述べ，図 3 に拡散経路図の例を示す．

(1) 管理対象ファイルノードに表示する情報

ノード内には，ファイルを識別する情報としてファイル名を表示する．ノードの形は四角形で表示し，現存するファイルは実線，消去されたファイルは破線で表示する．また，計算機外部に書き出された管理対象ファイルのノードの形は，2 重の四角形で表示し，書き出し先のデバイス番号または IP アドレスを表示する．

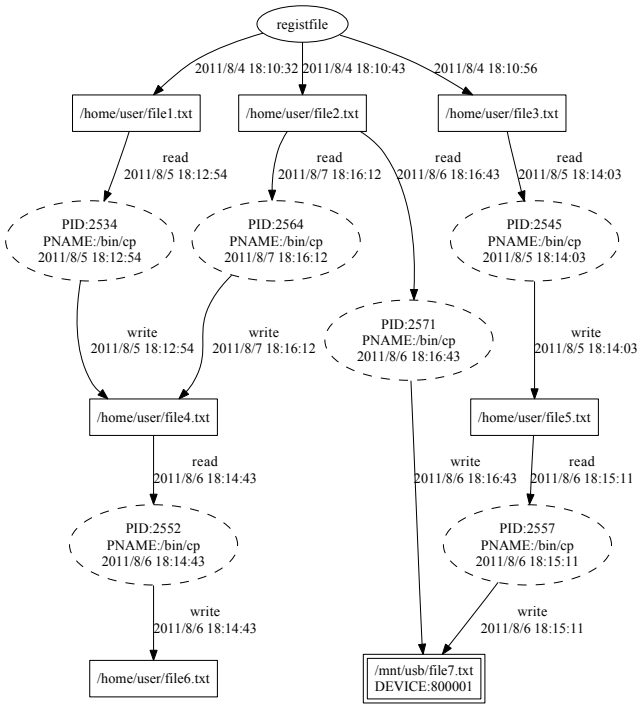

図 3 拡散経路図の例

(2) 管理対象プロセスノードに表示する情報

ノード内には，プロセス ID，プロセス名およびプロセス起動日時を併記して表示する．これらを併記することにより，同じプロセス ID とプロセス名を有したプロセスが複数存在しても利用者が識別できるようになる．ノードの形は楕円形で表示し，起動中のプロセスは実線，終了したプロセスは破線で表示する．また，ファイルを管理対象として登録するプログラム registfile は，それ自体が管理対象ではないものの，拡散経路の始点として経路図に表示する．registfile のノードは，拡散経路の始点として同一に扱うため，プロセス名だけ表記する．

(3) エッジに表示する情報

機密情報は，同じ管理対象に対して複数回拡散する可能性があるため，機密情報が拡散したすべての時刻を管理対象のノードに表示することはできない．よって，エッジの横に機密情報が拡散した時刻をすべて表示する．また，どのシステムコールによって機密情報が拡散したか区別するため，システムコール名も表示する．また，registfile のノードから伸びるエッジについては，registfile がファイルを管理対象として登録した時刻をエッジの横に表示する．

3.5 実現時の対処

大きく以下の 3 つの対処を行った．

(1) 機密情報の再拡散の流れも把握できるように機密情報の拡散追跡機能が管理対象への機密情報の再拡散を検知した際，機密情報の再拡散を伝えるログを出力する処理を追加した．

(2) 利用者が管理対象ファイルの正確なファイル名を確認できるようにするため，ファイル名変更の際に発行される rename システムコールをフックすることでファイル名の変更を検知し，ファイル名の変更を伝えるログを出力する処理を追加した．

(3) 利用者がプロセスを識別する情報としてプロセスの起動日時をプロセスが関連するログに追加した．同時には，同じ ID を持ったプロセスは複数存在しないため，プロセスを識別できる．

4 要件に対する評価

4.1 表示する情報の漏れと誤りについて

以降では,（要件 1）を 3.5 節で示した対処によって 2.3 節の（課題 1）と（課題 2）を解決することで満たせたかを図 3 を用いて確認する．図 3 は，以下の順序で操作した後に可視化した拡散経路図である．

(1) file0.txt を管理対象として登録
(2) file2.txt を管理対象として登録
(3) file3.txt を管理対象として登録
(4) file0.txt のファイル名を file1.txt に変更
(5) file1.txt を file4.txt にコピー
(6) file3.txt を file5.txt にコピー
(7) file4.txt を file6.txt にコピー
(8) file5.txt を USB メモリ内の file7.txt にコピー
(9) file2.txt を USB メモリ内の file7.txt にコピー
(10) file2.txt を file4.txt にコピー

図 3 からは，file4.txt や file7.txt に対して複数のファイルから機密情報が拡散したことを確認できる．また，プロセスのノードには，プロセスの起動日時が表示されている．さらに，上記の手順（4）で file0.txt のファイル名を file1.txt に変更している．図 3 では，file0.txt のファイル名は file1.txt に変更されており，正確に変更後のファイル名を表示できている．

4.2 簡潔な情報の表示による機密情報の拡散状況の把握の容易化について

以降では,（要件 2）を図 3 で示した拡散経路図を用いて，3.1 節で示したそれぞれの目的について（課題 3）を解決することで満たせているかを示す．

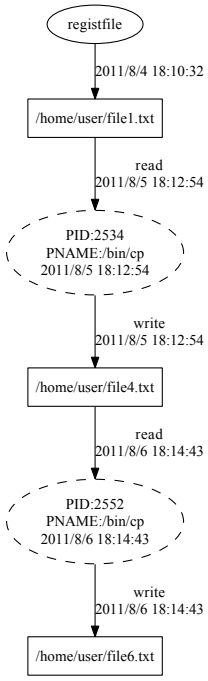

図 4 指定したファイルへの拡散経路図

（目的 1）については，利用者が書き出そうとしたファイルを管理対象だと認識していなかった場合，機密情報が拡散した原因を検証し，今後不必要な機密情報の拡散を起こさないように対策する必要がある．また，管理対象ファイルを計算機外部に書き出す必要があり，書き出そうとした場合，書き出すファイルに他の管理対象ファイルから機密情報が拡散していないかを確認する必要がある．

そこで，可視化機能では利用者が指定したファイルに拡散した機密情報の拡散経路のみを表示するフィルタを提供する．図 3 に示した拡散経路図において file6.txt に拡散した機密情報の拡散経路のみを可視化した図を図 4 に示す．図 3 では，file2.txt から file4.txt を介して file6.txt に機密情報が拡散したように表示されている．しかし，file2.txt から file4.txt へ機密情報が拡散したのは 8 月 7 日であり，file4.txt から file6.txt に機密情報が拡散したのは 8 月 6 日であるため，図 4 には表示されていない．このように，図 4 には，file6.txt に関係する情報しか表示されていないため，機密情報の拡散した原因や他の機密情報が拡散していないかを容易に確認できる．

（目的 2）については，利用者は，機密情報の漏えいを防止するため，機密情報の拡散状況を把握しておく必要がある．このため，可視化機能は，機密情報の全拡散経路を表示する機能を提供する．図 3 に示したように全体の拡散経路図が表示できている．

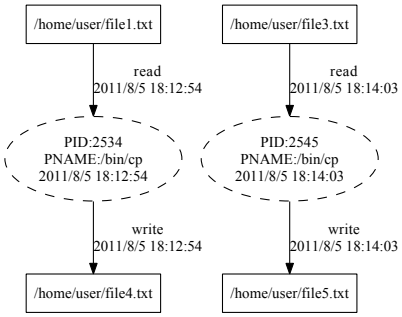

図 5 時刻指定をした拡散経路図

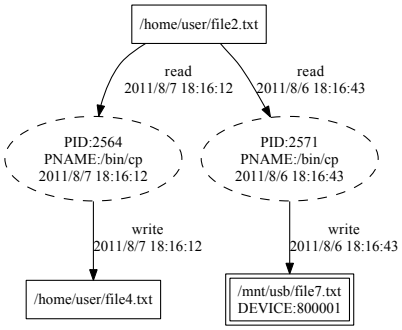

図 6 指定したファイルからの拡散経路図

また，ログは，記録する時間が長くなるほど増加していく．このため，ログを全て可視化した場合，表示される情報が増加し，利用者の視認性が低下する．

そこで，指定した時刻間に起きた機密情報の拡散経路のみを表示するフィルタを提供する．これにより，利用者の視認性を高める．図 3 に示した拡散経路において 2011 年 8 月 5 日の機密情報の拡散経路を表示した図を図 5 に示す．

（目的 3）については，漏えい原因を調査する際，漏えいした情報を保持しているファイルからの機密情報の拡散経路を確認する必要がある．

そこで，可視化機能では利用者が指定したファイルから拡散した機密情報の拡散経路のみを表示するフィルタを提供する．図 3 に示した拡散経路図において file2.txt の機密情報が漏えいしたと仮定し，file2.txt から拡散した機密情報の拡散経路のみを可視化した図を図 6 に示す．図 6 でも，図 4 の例で述べたように file2.txt の機密情報は，file6.txt に拡散していないため表示されていない．このように，図 6 には，file2.txt に関係する情報しか表示されていないため，漏えい原因を容易に確認できる．

5 関連研究

文献 [4] では，機密情報の拡散を追跡し，機密情報を有するファイルの親子関係をツリー状で可視化し，機密情報の拡散先を確認できる．しかし，表示される情報が少ないため，詳細な情報の確認にはログを見る必要がある．

また，文献 [5] では，機密情報の拡散を追跡したログから機密情報の拡散経路を 5 つの手法で可視化している．各手法は，目的ごとに表示情報を分け，表示の煩雑化を防いでいる．しかし，情報が分散されているため，各手法を切り替えながら確認する必要がある．

6 おわりに

機密情報の拡散状況を有向グラフ形式で利用者の指定する情報を表示する可視化機能を提案し，ログを可視化した実例を示した．

機密情報の拡散追跡機能がテキスト形式のログで提供する情報は，機密情報の拡散の検知や機密情報の管理に必要ない情報が出力されない．このため，利用者は正確な機密情報の拡散状況を把握できない．そこで，機密情報の拡散追跡機能が提供するログに拡散経路に関するログも出力する処理を実現した．

また，ログは機密情報の拡散が起こるたびに追加されるため，利用者が必要とする情報を探すには手間がかかる．そこで，可視化機能に利用者の指定する情報を表示する機能を実現した．

さらに，提案機能がログを可視化した実例を示し，正確な機密情報の拡散状況を把握できることと利用者が指定する情報を表示できていることを確認した．

残された課題として，機密情報の拡散追跡機能に追加した機能のオーバヘッドの評価と新たな表示方法の検討がある．

謝辞 本研究の一部は，科学研究費補助金若手研究 (B) (課題番号：21700034) による．

参考文献

[1] 日本ネットワークセキュリティ協会：2010 年度情報セキュリティインシデントに関する調査報告書 Ver.1.1, http://www.jnsa.org/result/incident/2010.html, 2011.

[2] 田端利宏，箱守　聰，大橋　慶，植村晋一郎，横山和俊，谷口秀夫：機密情報の拡散追跡機能による情報漏えいの防止機構，情報処理学会論文誌，Vol.50, No.9，pp.2088–2102，2009.

[3] AT&T, Graphviz, http://www.graphviz.org/

[4] Kida, K., Sakamoto, H., Shimazu, H. and Tarumi, H.: InfoCage: A Development and Evaluation of Confidential File Lifetime Monitoring Technology by Analyzing Events from File Systems and GUIs, Proceedings of the 2nd International Workshop on Security (IWSEC 2007), Springer, LNCS, Vol.4752, pp.246–261, 2007.

[5] 中山佑輝，小崎真寛，芝口誠仁，岡田謙一：機密情報追跡データの可視化による情報漏洩対策支援，情報処理学会論文誌，Vol.52，No.1，pp.24–32，2011.