情報処理学会論文誌 Vol.53 No.9 2171–2181 (Sep. 2012)

可視化とフィルタリング機能により機密情報の拡散追跡を支援する機構の実現

福島 健太[1] 山内 利宏[1,a)] 谷口 秀夫[1]

受付日 2011年12月5日, 採録日 2012年6月1日

概要：機密情報が外部へ漏えいする事例が増加している．情報の漏えいを防止するには，計算機の利用者が計算機内部の機密情報の利用状況を把握することが重要である．機密情報の利用状況を把握する手法としては，計算機内の機密情報を有するファイルへの操作を監視し，ログとして出力する機能が実現されている．しかし，機密情報の利用状況を把握するにはログを解析する必要があるため，未然に漏えいを防止することが難しい．我々は，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た．しかし，計算機利用者が提案機能から機密情報の利用状況を確認するには，テキスト形式のログを解析する必要があり，情報の漏えいが起こった際の原因特定には時間がかかる．そこで，本論文では，計算機での機密情報の利用状況を特定の機密情報ファイルや特定の期間などに着目し，視覚的に把握可能な機密情報利用状況の可視化機能を提案する．また，既存の機密情報の拡散追跡機能を拡張して，提案機能を実現する方式について述べ，評価結果を報告する．

キーワード：可視化，機密情報，拡散経路，情報漏えい対策

Implementation of Mechanism to Support Tracing Diffusion of Classified Information by Visualization and Filtering Function

Kenta Fukushima[1] Toshihiro Yamauchi[1,a)] Hideo Taniguchi[1]

Received: December 5, 2011, Accepted: June 1, 2012

Abstract: The number of incidents leaking of classified information has increased. To prevent leakage of information, it is important for users to understand the usage of classified information. To understand the usage of classified information, an method has implemented that monitors operations on the classified information and logs those operations. However, because an analysis of logs is necessary for understanding the usage of classified information, it is difficult to prevent leakage of classified information. We proposed the function to trace the classified information diffusion and detect a leakage of classified information. However, to understand the usage of classified information from the function by users, it is necessary to analyze the log in text format. Therefore, it takes long time to investigate the cause of the leakage of information. This paper proposes a function to visualize diffusion path of classified information. The function enables us to visualize the diffusion path of designated files that contain classified information. In addition, the function can visualize the diffusion paths focusing on the designated period of file operations. This paper also describes the implementation of the proposed function by extending the existing tracing function of classified information.

Keywords: Visualization, classified information, diffusion path, information leakage prevention

1. はじめに

近年，計算機で機密性の高い情報を扱う機会の増加とともに，機密情報が可搬記憶媒体やネットワークを経由し，

[1] 岡山大学大学院自然科学研究科
Graduate School of Natural Science and Technology, Okayama University, Okayama 700–8530, Japan
[a)] yamauchi@cs.okayama-u.ac.jp

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]．このような情報の漏えいを防止するには，計算機の利用者が計算機内部の機密情報の利用状況を把握することが重要である．また，利用者にとって，現在どのプロセスが機密情報をどのように利用しているのかを把握することは，不正なプログラムの兆候を把握するうえでも重要である．

しかし，既存 OS では，機密情報に着目して，情報の流れを追跡することは行われていないため，計算機内部の機密情報の利用状況を把握するのは難しい．また，機密情報の利用状況を把握する手法として計算機内の機密情報を有するファイル（以降，機密情報ファイルと略す）への操作を監視し，ログとして出力する機能 [2] がある．しかし，機密ファイルが伝播する情報を取得できたとしても，機密情報の利用状況を把握するにはログを解析する必要がある．ログは，発生したイベントが 1 行ごとに記録してあるため，ログの解析に時間を要する．このため，機密情報の流れを即座に把握することは難しく，情報の漏えいが起こった場合の原因の特定に時間を要してしまう．

我々は，機密情報の漏えいを未然に防ぐ手法として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 [3]（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から利用者が確認できる機密情報の利用状況は，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機密情報を有する可能性のあるファイルとプロセスの一覧のみである．このため，機密情報の利用状況を把握するには，文献 [2] と同様にログを解析する必要があり，情報漏えいの原因を特定する時間を短縮できない．また，機密情報の漏えいを検知した際も書き出し制御は，計算機利用者が書き出しの可否を判断する．このため，書き出しの可否を判断する際にその操作で漏えいする可能性がある機密情報ファイルを特定するには，ログを解析し，機密情報の利用状況を確認することが必要であり，利用者への負担が大きい．このとき，確認にミスがあった場合は，情報の漏えいを許す可能性もある．

そこで，本論文では，計算機での機密情報の利用状況について特定の機密情報ファイルや特定の期間などに着目し，視覚的に把握可能な機密情報利用状況の可視化機能を提案する．また，既存の機密情報の拡散追跡機能を拡張して，提案機能を実現する方式について述べる．さらに，可視化機能を実現した機密情報の拡散追跡機能で取得したログを基にした提案方式の評価を行い，評価結果を報告する．評価と関連研究との比較から，提案方式の有用性と長所を明らかにする．

2. 機密情報の拡散追跡機能

本論文で提案する可視化機能を実現する機密情報の拡散追跡機能とその問題点について述べる．

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事前に指定している．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．

(1) 機密情報の拡散契機となるシステムコールをフック

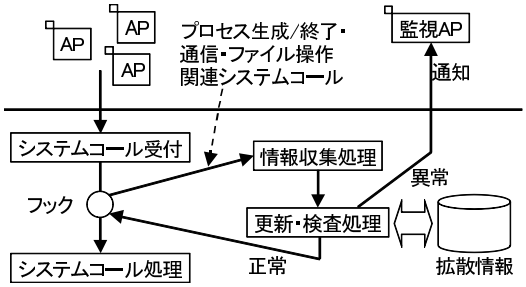

図 1 機密情報の拡散追跡機能の基本機構

Fig. 1 Basic mechanism of the diffusion tracing function of classified information.

(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 可視化が必要な状況

以下に機密情報の拡散追跡機能利用時において機密情報の拡散経路の可視化が必要になると想定される状況について述べる．また，各状況において計算機利用者に求められる作業について述べる．

（状況 1） 機密情報の拡散追跡機能が機密情報の漏えいの可能性を検知した際に，可否判断をする場合
（状況 2） 機密情報が計算機外部へと漏えいした際，漏えいの原因を特定する場合
（状況 3） 機密情報が意図していない拡散をしていないか確認する場合

（状況 1）については，機密情報の拡散追跡機能が計算機外部へのファイル書き出し時に情報の漏えいの可能性があるとして検知したファイルを，利用者自身が管理対象ファイルだと認識していなかった場合がある．この場合，どの機密情報ファイルから書き出し対象ファイルに，機密情報が拡散した可能性があるのかを利用者が検証し，計算機外部へのファイル書き出しを許可するか否か判断する必要がある．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るか否かを確認する必要がある．これにより，利用者の判断ミスによる情報漏えいを防止できる．

（状況 2）については，機密情報の漏えいが起こった場合，漏えいの原因を検証することで今後機密情報の漏えいを起こさないように対策する必要がある．これにより，再度の情報漏えいを防止できる．

（状況 3）については，定期的に機密情報の利用状況を確認することで利用者が意図していない機密情報の拡散を確認できる．これにより，マルウェアなどの不正なプロセスの存在や情報の漏えいを未然に防止できる．

以上の 3 つの状況は，文献 [2] や文献 [3] などの既存手法では，機密情報ファイルからの情報の伝搬に必要な情報がログに出力されていれば，それぞれ機密情報の利用状況を記したログを解析することで対応できる．しかし，広く利用されている既存 OS は，そのようなログを取得していない．また，文献 [3] では，新たに機密情報が拡散する契機でしかログを取得しておらず，1 度機密情報が拡散した対象に対して，それ以降の操作のログを取得していないという問題がある．

また，ログの解析には，以下の 2 つの問題点がある．

（問題点 1） ログは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わたって分散されて表示される可能性がある．
（問題点 2） ログの解析には時間がかかり，迅速に対応することが困難である．

（問題点 1）により，解析の際に人為的なミスが発生し誤った解析結果が出る可能性がある．また，（問題点 2）により，計算機利用者の作業効率が低下する．そこで，機密情報ファイルからの伝搬を追跡するのに必要な情報を取得し，ログを自動で解析することで，計算機での機密情報の利用状況の把握を支援でき，過去の利用状況も視覚的かつ詳細に把握可能な機密情報の拡散経路の可視化機能を提案する．

3. 機密情報利用状況の可視化機能

3.1 可視化の目的

以下に可視化の目的を述べる．

（目的 1） 機密情報の拡散経路を既存のログよりも正確に把握可能にする．
（目的 2） 機密情報の拡散経路を既存のログよりも迅速に把握可能にする．

2.3 節で述べた状況において，利用者の判断ミスによる漏えいを防止し，機密情報漏えい時の漏えい原因を特定するには，機密情報の拡散経路を正確に把握することが重要となる．

そこで，機密情報の拡散経路を既存のログよりも正確に把握可能にすることを 1 つ目の目的とする．また，2.3 節で述べた状況において，利用者の機密情報の拡散経路の認識速度を向上させ作業効率を高めるには，機密情報の拡散経路を迅速に把握することが重要となる．そこで，機密情報の拡散経路を既存のログよりも迅速に把握可能にすることを 2 つ目の目的とする．

3.2 考え方

文献 [3] で提案した機密情報の拡散追跡機能は，新たに機密情報が拡散したときの操作内容だけをログに出力する．このため，すべての機密情報に関する操作をログとして出力するように機能を追加する必要がある（3.4 節で後述する）．

また，機密情報の拡散に関わった操作に関するログをすべて出力したとしても，計算機利用者が機密情報の利用状

況を確認する際には，いつも機密情報の利用状況に関するすべての情報が必要になるわけではない．たとえば，ある特定の管理対象ファイルに関連する機密情報の利用状況のみが必要となる場合がある．また，ある特定の時刻における機密情報の利用状況のみが必要になる場合がある．このように，計算機利用者が確認したい情報が限られている際，すべての情報を表示した場合，計算機利用者は必要な情報を自身で探す必要がある．このため，問題点で述べたように解析の誤りと作業効率の低下が起こる可能性があり，計算機利用者が機密情報の拡散経路を正確かつ迅速に把握するのを阻害している．

そこで，可視化機能では，表示する情報をフィルタリングする機能を提供する．フィルタリング機能は，以下の 3 つに分けられる．

(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) 指定した期間の機密情報の拡散経路のみを表示

また，3 つのフィルタリング機能は，組み合わせて使うことも可能であり，指定したファイルから指定したファイル間の機密情報の拡散経路を表示したり，指定した時刻における指定したファイルからの機密情報の拡散経路を表示したりできる．

フィルタリングした情報をどのような形で表示するかについては次節で述べる．

3.3 可視化方式

（問題点 1）で述べたようにログが増えることで 1 つの管理対象の情報が数百行にわたって分散されて表示される可能性がある．そこで，分散されていた情報を集約するため機密情報の拡散経路を有向グラフ形式の拡散経路図として表示する．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 2** 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．

(2) 管理対象プロセスノードに表示する情報

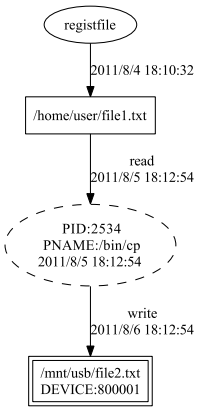

図 2 拡散経路図の例
Fig. 2 Example of diffusion path diagram.

表 1 ファイル操作時に取得する情報
Table 1 Information obtained when file operations.

通番	取得する情報の種類
(1)	ファイル操作が行われた時刻
(2)	ファイル操作を行ったプロセスを特定する情報（プロセス名，プロセス ID，起動日時）
(3)	ファイル操作が行われたファイルを特定する情報（inode 番号，デバイス番号）
(4)	ファイル操作に使用されたシステムコール名

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．このため，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．

3.4 可視化に必要な情報

機密情報の利用状況を可視化する際に必要な情報について 2.1 節で述べた機密情報が拡散する契機ごとに示す．ファイル操作時に取得する情報を**表 1**，プロセス間通信時に取得する情報を**表 2**，および子プロセス生成時に取得す

表 2 プロセス間通信時に取得する情報
Table 2 Information obtained when inter-process communication.

通番	取得する情報の種類
(1)	プロセス間通信が行われた時刻
(2)	送信元プロセスを特定する情報（プロセス名，プロセス ID，起動日時）
(3)	送信を仲介した通信用資源を特定する情報（ソケットアドレス，パイプの inode 番号，FIFO の inode 番号，メッセージキュー ID，共有メモリ ID）
(4)	受信先プロセスを特定する情報（プロセス名，プロセス ID，起動日時）
(5)	プロセス間通信に使用されたシステムコール名
(6)	受信先の計算機を特定する情報（IP アドレス）

表 3 子プロセス生成時に取得する情報
Table 3 Information obtained when child process creation.

通番	取得する情報
(1)	子プロセスの生成が行われた時刻
(2)	親プロセスを特定する情報（プロセス名，プロセス ID，起動日時）
(3)	子プロセスを特定する情報（プロセス名，プロセス ID，起動日時）
(4)	子プロセスの生成に使用されたシステムコール名

表 4 ファイル名変更時に取得する情報
Table 4 Information obtained when a file name is renamed.

通番	取得する情報の種類
(1)	ファイル名変更が行われた時刻
(2)	ファイル名が変更されたファイルを特定する情報（inode 番号，デバイス番号）
(3)	変更後のファイル名

る情報を**表 3** に示す．

既存の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ことで，計算機外部への漏えいを防止することが目的であり，新たに機密情報が拡散した場合のみ操作の情報を取得していた．そこで，すべての機密情報の拡散経路を可視化するためにすべての拡散契機で表 1，表 2，および表 3 に示した情報を取得するようにした．また，既存のログは正確に機密情報の拡散経路を表示するには情報が不足していたため，プロセスを特定するための情報にプロセスの起動日時を追加した．さらに，機密情報の拡散の契機とはならないが，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，ファイル名変更時に**表 4** に示す情報を取得する処理を追加した．取得した情報は，図 1 で示した更新・検査処理の際，テキスト形式のログとして出力する．

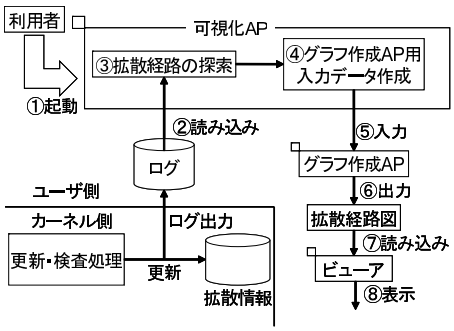

図 3 可視化機能の基本機構
Fig. 3 Basic mechanism of function to visualize diffusion path.

4. 機密情報の拡散追跡機能における設計と実現方式

4.1 設計方針

3.1 節の目的を達成するために可視化機能に求められる 2 つの要件を以下に述べる．

（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ること

（要件 1）は，機密情報の拡散経路を正確に把握するうえで必要になる．

（要件 2）は，機密情報の拡散経路を正確かつ迅速に把握するうえで必要になる．

4.2 基本機構

4.2.1 処理の流れ

図 3 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

処理 (1) の際，利用者は可視化する対象のファイルや期間を指定できる．可視化機能は，処理 (3) の際，利用者が指定した対象に関する機密情報の拡散経路のみを探索する．また，各機密情報の拡散経路について，機密情報が拡散した時刻の順序関係を用いて，指定された期間に機密情報が拡散した経路か否かを判別する．これにより，表示する拡散経路をフィルタリングできる．

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．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．

4.2.2 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 [4] 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．

DOT 言語で記述されたファイルは，可視化 AP が出力する．具体的には，可視化 AP が機密情報の拡散追跡機能が出力するテキスト形式のログを読み込む．次に，リスト形式の構造体である管理対象ファイルに関する情報を格納する管理対象ファイルリストと管理対象プロセスに関する情報を格納する管理対象プロセスリストを作成する．各リストには，管理対象に機密情報を拡散したファイルまたはプロセスの情報が格納されている．その後，各リストを相互にたどることで機密情報の拡散経路を探索し，DOT 言語で記述されたグラフ作成用のファイルを出力する．

5. 評価

5.1 目的と観点

3.1 節で述べた 2 つの可視化の目的を満たしているか否かを明らかにするため，以下の 3 つの評価を行った．

（**評価 1**） 各状況でのフィルタと拡散経路図の有用性の評価
提案機能が 2.3 節で示した各状況において有効か否か評価する．

（**評価 2**） フィルタリングによる拡散経路図の複雑化防止の評価
ログを蓄積し続けたときに，蓄積したログからすべての情報を可視化すると，表示が複雑化する．そこで，フィルタリングにより，表示の複雑さを防止できるか否かを評価する．

（**評価 3**） 提示方式の違いによる拡散経路把握の正確性と迅速性の評価
機密情報の利用状況の 3 つの提示方式を比較し，各方式における拡散経路把握の正確性と迅速性を評価する．

5.2 各状況でのフィルタと拡散経路図の有用性の評価

可視化機能を実装した機密情報の拡散追跡機能を有効にし，以下の手順で機密情報を拡散させた後，可視化機能によって出力した拡散経路図である**図 4** を用いて，可視化機能が 2.3 節で示した各状況において有効な表示方法ができているかを示す．

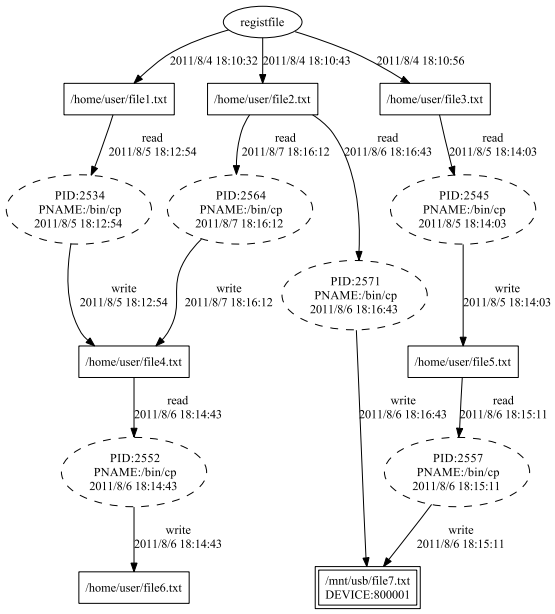

図 4 全体の拡散経路図
Fig. 4 All of diffusion path diagram.

(1) file1.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) file1.txt を file4.txt にコピー
(5) file3.txt を file5.txt にコピー
(6) file4.txt を file6.txt にコピー
(7) file5.txt を USB メモリ内の file7.txt にコピー
(8) file2.txt を USB メモリ内の file7.txt にコピー
(9) file2.txt を file4.txt にコピー

（状況 1）においては，外部へ書き出そうとしているファイルへの機密情報の拡散経路を確認することで，利用者の意図しない機密情報の拡散の原因や利用者の判断ミスによる漏えいを防止できる．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 4 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 5** に示す．図 4 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，file6.txt には file2.txt の機密情報は拡散しておらず，図 5 には表示されていない．このように，図 5 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たか否かを容易に確認できる．

（状況 2）においては，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ことで，機密情報の漏えいの原因を特定できる．そこで，可視化機能

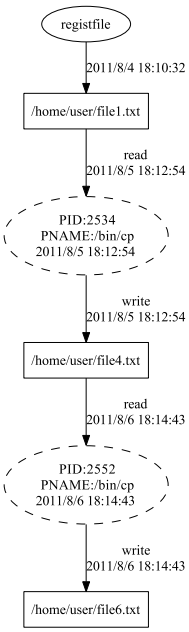

図 5 指定したファイルへの拡散経路図
Fig. 5 Diffusion path diagram to the specified file.

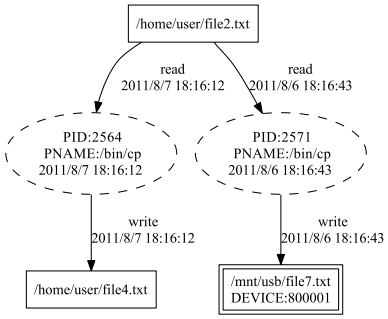

図 6 指定したファイルからの拡散経路図
Fig. 6 Diffusion path diagram from the specified file.

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 4 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 5 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．

（状況 3）においては，利用者は，定期的に機密情報の拡散経路を把握することで，利用者の意図しない機密情報の拡散を防止し，情報の漏えいを未然に防止できる．拡散経路図は，可視化対象のログの容量が増加するに比例して複雑になる．そこで，指定した期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，利用者は，前回確認した以降の機密情報の拡散のみを確認できるようになる．図 4 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 7** に示す．図 7 には，指定した期間に発生した機密情報の拡散経路しか表示されていない．このように，期間を指定することで表示される情報が増えることを防ぐことができる．

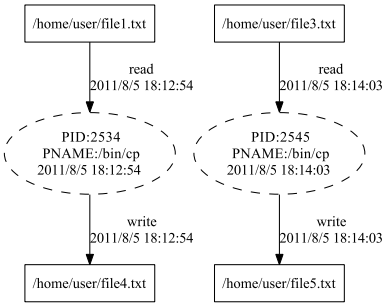

図 7 時刻指定をした拡散経路図
Fig. 7 Diffusion path diagram at the specified time.

この評価では，3 つのフィルタを単体で適用した場合の例を示した．実際には，これらの 3 つのフィルタを任意に組み合わせて利用することができ，利用者の要求に合わせた機密情報の伝搬の可視化が可能である．

5.3 フィルタリングによる拡散経路図の複雑化防止の評価

5.3.1 目的と評価内容

提案機能は，可視化するログの量が増えるほど表示するノードの量が増加し，ノード間の依存関係が複雑になる．このような表示の複雑化を防止するため，提案機能は，表示する拡散経路のフィルタリング機能を提供する．そこで，取得したログを提案機能によって全経路を可視化した拡散経路図と，提案機能のフィルタリング機能により特定のファイルに関する経路のみを可視化した拡散経路図に表示されるノード数を比較する．これにより，フィルタリング機能により，表示の複雑化をどの程度防止できるかを評価する．また，評価には約 250 行のログを用いた．このログは，機密情報の拡散追跡機能を実装した計算機を使用し，7 日間にわたって計算機内で機密情報を拡散させることで取得した．

5.3.2 評価結果と考察

全経路を可視化した拡散経路図を**図 8** に，フィルタリング機能により特定のファイルへの拡散経路のみを可視化した拡散経路図を**図 9** に示す．図 8 と図 9 から，ノード数は特定のファイルのみへの拡散経路にフィルタリングすることにより，129 ノードから 25 ノードに約 80%減少しており，表示の複雑化を防止できていることが分かる．

長期間にわたってログを収集した場合でも，提案機能は可視化する期間を指定して，特定のファイルに着目してフィルタリングした結果を可視化できる．このため，可視化する期間を指定するなどフィルタリングの粒度を細かくすることにより表示の複雑化は防止できると考える．

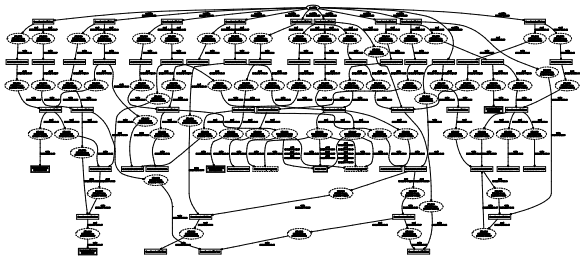

図 8 全経路を可視化した拡散経路図
Fig. 8 Diffusion path diagram includes all paths.

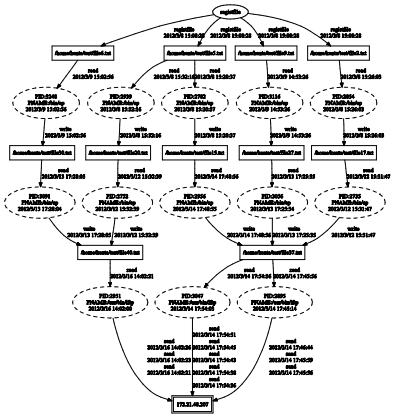

図 9 特定経路のみの拡散経路図
Fig. 9 Diffusion path diagram includes specified paths.

5.4 正確性と迅速性の評価

5.4.1 評価内容

機密情報の利用状況を有向グラフ形式の拡散経路図で表示することの有用性について検証する．評価では，ログと拡散経路図を被験者に提示し，被験者はログと拡散経路図に関する質問に回答する．

回答結果は，以下に示す 2 つの観点で評価する．

（観点 1） 質問への回答の正答率

（観点 2） 質問への回答までの時間

（観点 1）を比較することで拡散経路図がログと比べて正確に機密情報の利用状況を確認できることを示す．また，（観点 2）を比較することで拡散経路図がログと比べて迅速に機密情報の利用状況を確認できることを示す．

評価では，機密情報の拡散追跡機能が出力する 30 行程度のログ 9 種類を使用する．**図 10** は，使用したログの 1 つである．また，以下の 3 つの提示方式を評価した．

（提示方式 1） 可視化機能が出力する拡散経路図と情報量を同じにしたログ

機密情報の拡散追跡機能が出力するログには，情報の拡散を追跡するには不必要なデータが多数入っている．そのデータを含んだログを評価に使用しては，可視化機能で出力した拡散経路図との公平性が保てないと考え，不必要なデータを削除したものを提示する．また，（提示方式 2）についても同様の編集を行っている．図 10 のログを（提示方式 1）に合わせて編集したログを**図 11** に示す．

1 Fri Dec 16 17:39:09 2011 892569240 PID: 2817 File Marked. FILE:/home/kenta/list/address/secret1.txt INODE:3286933 PID:2817 PNM:/home/kenta/rd-dev/tool/registfile MODE:1 NO:1 DEV:300003
2 Fri Dec 16 17:39:40 2011 141970632 PID: 2824 File Marked. FILE:/home/kenta/list/user/secret2.txt INODE:3286939 PID:2824 PNM:/home/kenta/rd-dev/tool/registfile MODE:1 NO:2 DEV:300003
3 Fri Dec 16 17:40:00 2011 934809640 PID: 2829 File Marked. FILE:/home/kenta/secret/secret3.txt INODE:3009092 PID:2829 PNM:/home/kenta/rd-dev/tool/registfile MODE:1 NO:3 DEV:300003
・・・
35 Fri Dec 16 17:50:41 2011 140483584 PID: 3002 Send to remote address! IP address:172.21.48.207 Port:5376 PID:3002 PNM:./writefile
36 Fri Dec 16 17:50:47 2011 880458952 PID: 2832 Process Unmarked PID:3002 PNM:./writefile

図 10 使用したログの例
Fig. 10 Example of logs used on the evaluation.

1 2011/12/16 17:39:09 File Marked. FILE:/home/kenta/list/address/secret1.txt PID:2817 PNM:./registfile
2 2011/12/16 17:39:40 File Marked. FILE:/home/kenta/list/user/secret2.txt PID:2824 PNM:./registfile
3 2011/12/16 17:40:00 File Marked. FILE:/home/kenta/secret/secret3.txt PID:2829 PNM:./registfile
・・・

図 11 提示方式 1 で利用したログ
Fig. 11 Logs used on the method 1.

（提示方式 2） ログを可視化機能と同様のフィルタリングを使用し，出力する拡散経路を指定したログ

フィルタリングは，質問内容に合わせて最適なものを使用する．ここでは，例として（提示方式 1）のログを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表示したログを**図 12** に示す．

（提示方式 3） フィルタリングを使用し，出力する拡散経路を指定した拡散経路図

（提示方式 2）と同様に，フィルタリングは，質問内容に合わせて最適なものを使用する．**図 13** に（提示方式 2）に示したログを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表示した拡散経路図を示す．

ログは，3 種類ごとに 3 つの提示方式に合わせて編集する．その後，被験者は，（提示方式 1）に合わせた 3 種類のログ，（提示方式 2）に合わせた 3 種類のログ，および（提示方式 3）に合わせた 3 種類の拡散経路図についての質問に回答する．

また，各提示方式についての質問は，2.3 節に示した各状況を想定したものであり以下の 3 種類に分けられる．

（質問 1） 対象ファイルにどこから機密情報が拡散したか

1 2011/12/16 17:43:05 Process Marked PID:2885 PNM:./writefile FILE:/home/kenta/list/address/secret1.txt (read)
2 2011/12/16 17:43:05 File Marked. FILE:/home/kenta/MyDocument/memo.txt PID:2885 PNM:./writefile
3 2011/12/16 17:43:05 Process Unmarked PID:2885 PNM:./writefile
5 2011/12/16 17:43:43 File Marked. FILE:/home/kenta/MyDocument/data/data2.txt PID:2895 PNM:/bin/cp
6 2011/12/16 17:43:43 Process Unmarked PID:2895 PNM:/bin/cp
7 2011/12/16 17:48:47 Process Marked PID:2986 PNM:/bin/cp FILE:/home/kenta/MyDocument/data/data2.txt (read)
8 2011/12/16 17:48:47 File Re-marked FILE:/home/kenta/sendinfo.txt PID:2986 PNM:/bin/cp
9 2011/12/16 17:48:47 Process Unmarked PID:2986 PNM:/bin/cp
10 2011/12/16 17:50:41 Process Marked PID:3002 PNM:./writefile FILE:/home/kenta/sendinfo.txt (read)
11 2011/12/16 17:50:41 Send to remote address! IP address:172.21.48.207 PID:3002 PNM:./writefile
12 2011/12/16 17:50:47 Process Unmarked PID:3002 PNM:./writefile

図 12 提示方式 2 で利用したログ
Fig. 12 Logs used on the method 2.

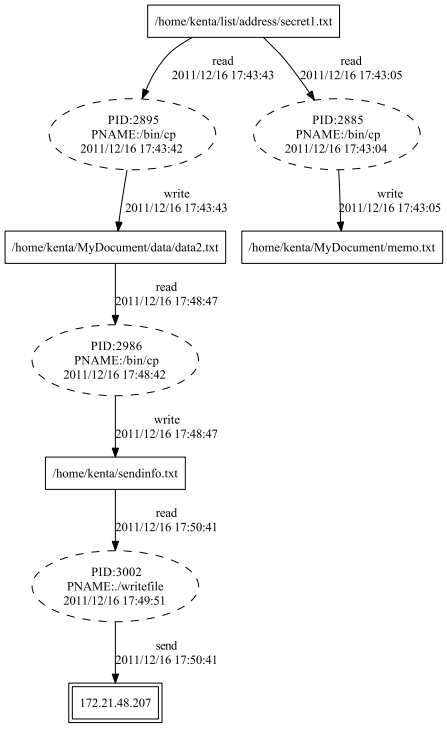

図 13 提示方式 3 で利用した拡散経路図
Fig. 13 Diffusion path diagram used on the method 3.

を問う質問
これは，（状況 1）を想定した質問である．この質問により，可視化機能が利用者の書き出しの可否判断時に有効に機能するかを示す．

（**質問 2**） 機密情報の漏えい先を問う質問
これは，（状況 2）を想定した質問である．この質問により，可視化機能が機密情報の漏えい原因の検証に有効に機能するかを示す．

（**質問 3**） 特定の期間内に機密情報を拡散させたプロセスを問う質問
これは，（状況 3）を想定した質問である．この質問により，可視化機能が利用者が意図していない機密情報の拡散の確認に有効に機能するかを示す．

（提示方式 1）と（提示方式 2）の差を調べることによって，可視化機能についている出力する拡散経路をフィルタリングする機能の性能を評価する．また，（提示方式 2）と（提示方式 3）の差を調べることによってログを拡散経路図として出力することの優位性を評価する．

5.4.2 ログに関する前提条件

機密情報の拡散と計算機外部への書き出しには，計算機利用者が意図して起こしたものとそうでないものがある．評価で使用するログは，被験者自身が計算機を利用して機密情報を拡散させた結果をログにしたものではなく，事前に用意したログである．このため，被験者はファイル名，プロセス名，および操作を行った日時などから意図的に起こった拡散かどうかを把握できない．そこで，利用者が意図して起こした拡散は，プロセス名を./writefile にすることによって表現する．このプロセスによって拡散し計算機外部へ書き出された機密情報は，利用者が機密情報と知っていて意図的に書き出したものと見なし情報の漏えいとはしない．他のプロセスによる拡散は，利用者が意図していなかったものと見なし，情報の漏えいとする．

5.4.3 評価結果と考察

5.4.1 項で示した観点に基づいて研究室内の大学生と大学院生 11 名を被験者として評価を行った．（質問 1），（質問 2），および（質問 3）ごとに，各提示方式における回答時間と正答率の平均値を求めたものを**表 5** に示す．また，（質問 1〜3）について，提示方式ごとに，回答時間と正答率の平均値を求めたものを**表 6** に示す．

回答時間については，表 6 より，（提示方式 1）と（提示方式 2）を比較すると（提示方式 2）の回答時間が 48%程度減少している．また，（提示方式 2）と（提示方式 3）を比較すると（提示方式 3）の回答時間が 69%程度減少している．これらの結果より，フィルタリング機能と可視化は，機密情報の利用状況を迅速に把握するのに有用であるといえる．

正答率については，表 6 より，ログのフィルタリングと可視化により正答率が向上している．正答率と回答時間の関係はトレードオフであり，回等時間が長くなるほど正答率は向上する．この評価では，時間の制限がないため，被験者は自分で正答を見つけたと判断するまで考えることができる．このため，回答時間に比べ，各提示方式間の正答

表 5 各質問における回答時間と正答率
Table 5 Response time and a percentage of correct answers of eash question.

	質問 1		質問 2		質問 3	
	回答時間	正答率	回答時間	正答率	回答時間	正答率
(提示方式 1)	4 分 37 秒	81.8%	4 分 16 秒	72.7%	2 分 33 秒	90.9%
(提示方式 2)	2 分 14 秒	81.8%	1 分 19 秒	100%	2 分 17 秒	100%
(提示方式 3)	46 秒	90.9%	30 秒	100%	35 秒	100%

表 6 提示方式ごとの回答時間の平均値と正答率
Table 6 Average of response time and a percentage of correct answers of each method.

	回答時間の平均値	正答率
(提示方式 1)	3 分 49 秒	81.8%
(提示方式 2)	1 分 58 秒	93.9%
(提示方式 3)	37 秒	97%

率の差は小さいと推察できる．

上記より，拡散経路の有向グラフ形式の可視化は，各質問において有効に機能しており，2.3 節に示した各状況において，ログよりも正確かつ迅速に機密情報の利用状況を把握できている．また，評価に使用したログは 30 行程度と短いものであるが回答時間と正答率に差が出ており，この差はログが増えるほど増大すると推察できる．

6. 関連研究

文献 [5] では，Jif（Java information flow）言語とセキュア OS の提供する強制アクセス制御を統合することで，OS レベルでのアクセス制御に加えて，AP 内における情報フローの追跡と制御を実現している．強制アクセス制御は，情報フローに基づいてアクセス制御できるものの，ポリシの設定ミスがあった場合，想定しないアクセスを許可する可能性がある．しかし，文献 [5] の手法においては，情報フローを視覚的に確認することはできないため，設定の不備などが原因となる情報漏えいの発見が遅れる可能性がある．一方，提案機能は，機密情報の拡散追跡機能が出力するログを用いて情報フローを拡散経路図として視覚的に示すことができるため，ポリシの設定ミスなどを発見することを支援することできる．機密情報の拡散追跡機能は，任意アクセス制御や強制アクセス制御と共存して動作することができるため，提案機能を用いて，強制アクセス制御などのアクセス制御の結果を可視化して，人間が確認することにも利用することができる．

文献 [6] は，ユーザレベルでシステムコールを捕捉し，ファイルの一部が共有メモリを通して複写された際の機密データの拡散を監視する方式を提案している．しかし，本方式においても，監視された機密データの拡散を視覚的に確認する手段は提供されていない．

機密情報の漏えいを防止するには，計算機利用者が機密情報が計算機内のどこにあるのかを認識することが重要であり，機密情報の拡散を追跡するだけでは対応できないこともある．

文献 [7] は，1 つの情報伝搬に関する操作の一連のログを，集約することにより，ログを削減することを実現している．また，蓄積したログから，機密情報の元となるファイルと，伝搬先のファイルなどの関係を木構造の親子関係として可視化し，機密情報の拡散先を確認できる．しかし，可視化された図では，機密情報を伝搬させたプロセスの情報や，プロセスのどのような処理によって，いつ伝搬したのかなどの詳細な情報を確認できない．一方，提案手法では，機密情報を伝搬させるすべての操作について，ログを取得することを実現している．このため，提案手法の可視化では，利用者に指定された任意の期間において，拡散経路上に機密情報を拡散，漏えいさせたプロセスが表示され，漏えい原因の特定が可能である．

文献 [8] 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，監視対象や可視化対象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，一連の機密情報の拡散についての情報が，各表示方法に分散されるため，1 つの表示手法では，一連の拡散に関する情報を把握できない．このため，各手法を切り替えながら，一連の機密情報の拡散について確認する必要があり，オーバヘッドが大きくなる．たとえば，ディレクトリ間の伝搬を表示する手法では，ディレクトリ間でのファイル複写などの流れしか分からない．どの AP によって，ファイル複写されたのかなどを確認する場合には，AP を用いた編集・複製を可視化する表示に切り替え，タイムラインで対応する箇所を探し，AP を特定する必要がある．このように，情報の伝搬経路を追跡するために，表示手法の切替えと経路特定の工数が増大する問題がある．また，可視化対象のファイルを指定して絞り込むことはできるものの，時間軸で絞り込むことはできない．文献 [8] の評価においても，状況によっては，情報を分散して表示するよりも 1 画面で表示した方が有効であるという結果が示されている．

一方，提案手法では，漏えいに関連するすべての情報を 1 画面で表示できるため，表示切替えのオーバヘッドがない．また，表示の煩雑化については，フィルタリング機能によって期間や伝搬経路を表示するファイルを絞り込むことで，煩雑化を防止できる．さらに，文献 [8] では，各方

式で可視化するために，どのようなタイミングで，どの情報を取得しなければならないかということは議論されていない．一方，本論文では，提案手法で可視化のために必要な情報やその取得手法について明らかにした．文献 [8] では，独自に仮定したログを用いているものの，提案方式では，拡張した機密情報の拡散追跡機能を用いて取得したログを用いて評価している．

7. おわりに

計算機での機密情報の利用状況を特定の機密情報ファイルや特定の期間などに着目し，視覚的に把握可能な機密情報利用状況の可視化機能を提案した．また，既存の機密情報の拡散追跡機能にすべての拡散経路に関するログを出力する処理を実現した．提案機能では，機密情報の拡散追跡機能が出力するテキスト形式のログを解析し，機密情報の拡散経路図を有向グラフ形式で利用者に表示する．また，表示する情報の煩雑化を防ぐため，特定の機密ファイルや特定の期間などに着目した拡散経路のみを表示するフィルタリングを行う．

提案機能が可視化の目的を満たしているか否かを明らかにするため，3 つの評価を行った．1 つ目の評価では，提案機能は，利用状況に合わせて，3 つのフィルタを組み合わせて，機密状況の伝搬の可視化ができることを示した．2 つ目の評価では，フィルタリングにより，可視化表示の複雑化を防ぐことができることを示した．3 つ目の評価では，被験者の回答時間と正答率を，フィルタリングと可視化により，向上させることができ，ログを用いるよりも，正確かつ迅速に機密情報の利用状況を把握できることを示した．

謝辞 本研究の一部は，科学研究費補助金若手研究（B）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，入手先 ⟨http://www.jnsa.org/result/incident/2010.html⟩ (2011).
[2] Goel, A., Po, K., Farhadi, K., Li, Z. and Lara, D.E.: The Taser Intrusion Recovery System, *Proc. 20th ACM Symposium on Operating Systems Principles* (*SOSP 2005*), pp.163–176 (2005).
[3] 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).
[4] AT&T, Graphviz, available from ⟨http://www.graphviz.org/⟩.
[5] Hicks, B., Rueda, S., Jaeger, T. and McDaniel, P.: From Trusted to Secure: Building and Executing Applications that Enforce System Security, *Proc. USENIX Annual Technical Conference*, pp.205–218 (2007).
[6] 鷲尾知暁，除補由紀子，大嶋嘉人，金井 敦：属性の伝播を利用した電子文書の柔軟な利用制御方式の提案，情報処理学会研究報告 2004-CSEC-28，Vol.2005, No.33, pp.375–380 (2005).
[7] 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, *Proc. 2nd International Workshop on Security* (*IWSEC 2007*), LNCS, Vol.4752, pp.246–261, Springer (2007).
[8] 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).

福島 健太

2009 年岡山大学工学部情報工学科卒業．2012 年同大学大学院自然科学研究科博士前期課程修了．コンピュータセキュリティに興味を持つ．

山内 利宏 （正会員）

1998 年九州大学工学部情報工学科卒業．2000 年同大学大学院システム情報科学研究科修士課程修了．2002 年同大学院システム情報科学府博士後期課程修了．2001 年日本学術振興会特別研究員（DC2）．2002 年九州大学大学院システム情報科学研究院助手．2005 年岡山大学大学院自然科学研究科助教授．現在，同准教授．博士（工学）．オペレーティングシステム，コンピュータセキュリティに興味を持つ．電子情報通信学会，ACM，USENIX 各会員．

谷口 秀夫 （正会員）

1978 年九州大学工学部電子工学科卒業．1980 年同大学大学院修士課程修了．同年日本電信電話公社電気通信研究所入所．1987 年同所主任研究員．1988 年 NTT データ通信株式会社開発本部移籍．1992 年同本部主幹技師．1993 年九州大学工学部助教授．2003 年岡山大学工学部教授．博士（工学）．オペレーティングシステム，実時間処理，分散処理に興味を持つ．著書『並列分散処理』（コロナ社）等．電子情報通信学会，ACM 各会員．