

高速簡易証拠保全ツールの提案と Autopsy の日本語化

長井 健^{1,a)} 上原 哲太郎^{2,b)}

概要：デジタル機器が関連する犯罪や不正の調査においては、二次記憶装置内のデータの証拠保全が不可欠である。しかしパソコンの二次記憶装置の容量は近年肥大化し、証拠保全にかかる時間が大きな問題となっている。本研究では、Windows の NTFS を対象とした証拠保全において、一部の重要なデータに絞って保全を行うことで保全にかかる時間を短縮する手法の提案と、保全すべきファイルを現場で同定するためのツールとしての Linux 版 Autopsy の日本語化について報告する。

1. はじめに

情報化社会の進展につれて、パソコンやスマートフォンなどの電子機器の普及が著しく、人々の生活においてその行動の多くがこれらの機器内の情報として記録されるようになってきた。このため、犯罪や不正の調査・捜査の現場においても、これらの機器内に残された記録（以下、電磁的記録と呼ぶ）を証拠とする機会が増えてきており、必然的にデジタルフォレンジックへの関心が高まってきている。犯罪に関わる電磁的記録が証拠となり得るとき、これを電磁的証拠と呼ぶ。電磁的証拠は改ざんや隠滅、調査時の操作誤り等による消失等の可能性がその他の証拠より高いため、より慎重かつ専門的な取扱いが求められている。そこで、警察をはじめとする司法機関においては、電磁的証拠が記録された機器のデータをまずそのまま複製して保全し、その時点でのデータの真正性をハッシュ値などで保存する「証拠保全」を解析前に行っている。

現在、警察等の専門員や鑑識員は、犯罪に関連しているすべてのハードディスク等の二次記憶装置に関して、その記憶領域を未使用領域も含めて全て複製するとともに、そのデータ内容のハッシュ値を計算して記録し、後の改ざん等が行われていないことを保証できるようにしている。これは、未使用領域内に削除されたファイルの痕跡等が残っている場合があり、重要な電磁的証拠となりかねないことと、一部のみの複製では裁判の場において被告弁護人から

複製の手続きに恣意性があり複製されていない領域に被告の有利となる電磁的証拠が残されていた可能性について指摘される可能性があるためである。しかし、近年の二次記憶装置の容量増大により、この証拠保全にかかる時間は増大しており、捜査の上では大きな支障となっているとされる。また、現在デジタル・フォレンジックにおける証拠保全や証拠分析を行える専門員の人数は不足しており、犯罪捜査において特に重要となる家宅搜索等の初期調査において全ての捜査現場に正しい証拠保全手続きを行うための人員を配置することは不可能である。よって、実際には必ずしも専門的知識を持たない捜査員が現場において証拠保全を行っている。そこで、証拠保全に際して出来るだけ平易な操作で確実に証拠保全ができるツールの開発が現場では求められている。

そこで本研究では、物理コピーではなく必要とされる電磁的証拠のみが短時間で抽出可能であり、デジタルフォレンジックの知識を十分に持ち合わせていなくても使用できる、証拠保全ツールを提案する。また、警察の専門員が保全すべき領域を現場で同定するためのツールを利用しやすくするため、Autopsy の日本語化について報告する。

2. 研究背景

2.1 デジタルフォレンジック

デジタルフォレンジックと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、電磁的記録の改竄・毀損等についての分析・情報収集等を行う一連の科学的調査手法・技術である。[1][2]

デジタルフォレンジックの応用分野は多岐にわたるが、ここでは応用分野を以下の 3 つに分けて整理する。[3]

- 企業活動

¹ 立命館大学 大学院情報理工学研究科
Graduate school of information science and engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

² 立命館大学 情報理工学部
College of information science and engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

a) cyse_nagai@cysec.cs.ritsumeikan.ac.jp

b) uehara@cs.ritsumeikan.ac.jp

企業の活動におけるデジタルフォレンジックとは、業務上の不正行為や業務不履行、守秘義務違反による契約違反、内規違反などによる民事訴訟や、不正競争防止法、会社法、金融商品取引法、個人情報保護法などが関連する民事訴訟または刑事訴訟への対応を行う技術である。

- インシデントレスポンス

インシデントレスポンスにおけるデジタルフォレンジックとは、インシデントが起きた場合の原因を探り、被害拡大の防止とシステムの速やかな回復、再発防止を図る技術である。さらに、裁判が起こる可能性に備えてその被害の状況や、解析によって判明した侵入経路、侵入者等の追跡の手がかりとなる証拠を保全する必要がある。刑事事件に発展することが想定される場合には、訴訟に備えて刑事事件に対応した証拠保全を行っておくことになる。

- 刑事事件

刑事事件におけるデジタルフォレンジックとは、犯罪の捜査で被疑者のパソコンやスマートフォンにあるメールの履歴やアプリケーションファイルなどから、犯罪の証拠となりうるものを取り出す技術である。刑事裁判においては証拠の取扱いは慎重であり、その評価が客観的であることが求められるため、特に証拠保全の手続きは重要である。しかし、実際の犯罪捜査現場においてはデジタルフォレンジックに関し高い技術を持つ専門家を十分に確保できるとは限らないため、専門の知識がなくても使用することができるフォレンジックツールが必要とされている。

これらの応用分野すべてにおいて、訴訟が生じた場合、証拠保全が必要となる可能性が高い。しかし、記憶領域を未使用領域も含めて全て複製することは、多大な時間を要する。そのため、物理コピーではなく必要とされる電磁的証拠のみ抽出することは必要である。ところが問題となるのが、物理コピーを行わず、必要とされる電磁的証拠のみ抽出した場合、抽出した証拠以外の箇所に被告に対して有利な証拠が存在するという疑念が生まれる。これに対して、NTFSにおける、MFT(Master File Table)の箇所を取ることににより、すべてのファイル名やアクセス時刻等を把握することができる。そのため、抽出した証拠以外の箇所を指摘された場合もアクセス時刻などから、被告に有利な証拠であるということが考えにくいと立証できる。しかし、本来、重要証拠である証拠が不足することを防ぐため、証拠保全を行う現場で判断し、ある程度の同定をする必要がある。そのためのソフトウェアも必要である。

また、証拠保全では、フォレンジック調査の対象物が、年々増加している。要因の一つとしてサイバー攻撃を行う際、攻撃者が組織内のネットワークへの侵入を行う際に周囲の不特定多数のパソコンが段階的に狙われ、アクセスさ

れる場合がある。攻撃者が捕まり、捜査に至った場合、周囲の不特定多数のパソコンについても捜査の対象になる可能性が高い。そのため、フォレンジック調査を行わなければならない対象物が増加し、どの証拠が必要かというものが区別がつかない状態となる。

また、犯人を逮捕してから、刑事訴訟法より23日間の間に起訴されない場合は、釈放されることになる。そのため、電子機器が関連する犯罪の場合、23日以内にデジタルフォレンジックと事件自体の捜査を行い、証拠を集めなければならない。しかし、現在の証拠保全は、時間の比重が非常に大きいため、十分な証拠が集まらないのを理由に不起訴になる可能性がある。これらの課題より、一部の記憶領域の保全によってこの時間を短縮する手法の提案をする。

2.2 Autopsy

Autopsyとは、ハードディスクなどを調査する際に使用する、フリーのフォレンジックツールである。このツールは、Webブラウザのインターフェースを提供して、調査を行いやすくするためのものである。実際のディスク調査を行うのは、The Sleuth Kitというフリーソフトウェアが行っている。[4] Autopsyでは、ファイルやディレクトリの内容を閲覧することに加え、削除されたファイル名やNTFSイメージの代替ストリームも表示が可能になる。また、メタデータ構造体の調査によるイメージの解析や修正時刻/アクセス時刻/変更時刻等の表示も可能である。このAutopsyを使用することで、2.1で述べた、証拠保全を行う現場で判断し、ある程度の同定を行うことにより、重要証拠が欠けることを防ぐことができる。しかし、Autopsyは英語ベースのツールであるため、ツールを使用する際に、抵抗を感じる場合がある。

3. 研究方法

3.1 提案手法

本研究では、デジタルフォレンジックの知識を十分に持ち合わせていない状態で証拠保全を行うことはできないという問題意識をもとにデジタルフォレンジックの知識を持ち合わせていない人でも証拠保全を行うことが可能にする、自作の証拠保全ツールを作成する。

自作の証拠保全ツールでは、部分的に必要な電磁的証拠のみをコピーする証拠保全の方法と部分的に証拠保全した内容の原本同一性を保証する方法を組み合わせることで、短時間で容易に証拠保全が行えるようにする証拠保全ツールを作成することが目的である。

自作の証拠保全ツールを作成するにあたって必要は、部分的に必要な電磁的証拠をコピーすることと、電磁的証拠の原本同一性を保証することであることから、主に以下の点に焦点をあてて作成を行った。

- 証拠保全箇所の限定

証拠保全箇所の限定については、WindowsOS のファイルシステムの多くのファイルから証拠保全に必要なファイルのみを選択した。今までの方法として、すべてのファイル領域を物理コピーする手法があるが、物理コピーの時間が多くかかるという問題がある。現代の記憶媒体容量の増加に対応できない節がある。

- 電磁的証拠の原本同一性の確立

電磁的証拠の原本同一性の確立について保証するにはタイムスタンプ署名（電磁的証拠がある日時に確かに存在していたという「存在証明」と、その日時以降改竄されていないという「非改竄証明」）を選択した。タイムスタンプ署名を行う箇所として、部分的に証拠保全したデータを対象にしている。

- Autopsy の日本語化

現場で警察が証拠を同定する際に英語ベースのソフトウェアの場合、抵抗を感じる人がいるため、Autopsy を日本語化することで警察でより扱いやすいツールとして対応した。

3.2 実装環境

近年、サイバー犯罪で不特定多数のパソコンがフォレンジックの対象になることも多い。そのため、何時でも何処でも簡単に証拠保全ができる体系を作成し、すぐに証拠保全ができる体制を整えることが重要である。

現在のパソコンには USB が大抵搭載されている。また、搭載されていない場合でも、PCI スロットや Card Bus を搭載していれば、USB スロットの増設が可能である。そこで、証拠保全がすぐに行えるように USB メモリーに OS をインストールを行った。他にも、CD-ROM や DVD-ROM、SD カードなどの記憶媒体が存在するが、いずれも、USB メモリーに比べて、搭載されている可能性は低くなる。また、USB メモリーは、ノートパソコンで見られる、取り外しができない HDD や SSD に対しても有効である点もある。USB メモリーに OS をインストールするにあたって、選択した OS は、Linux 系 OS である。WindowsOS において起動スクリプトの書き換えは権限の関係で困難な場合が多いことがある。また、WindowsOS は著作権や有償の問題があるため Linux 系の OS を採用した。

3.3 証拠保全箇所

表 1 は一部の記憶領域の保全によってこの時間を短縮するために必要だと考えられるものを示している。次節で、表 1 に保全箇所を限定した主な理由を示す。

3.3.1 \$MFT(Master File Table)

MFT(Master File Table) は NTFS のボリ्यूーム構造の中心であり、1kB ファイルレコードの配列として実装され、1 個のファイルレコードが 1 つのファイルを表す。各ファイルレコードには 0 から順番に番号がついており、これが

表 1 部分的にコピーする証拠保全箇所

部分的証拠保全箇所	内容
\$MFT	メタデータの格納域とデータ本体の配置情報
\$Bitmap	ファイル削除時の両方の空き領域の管理
hiberfil.sys	Windows が休止状態に移る前に必要なデータを一時保存しておくためのファイル
pagefile.sys	物理メモリの空きがない際にデータの一時退避に使用
Temp	編集ファイル(.word 等)などの一時ファイル
Prefetch	頻繁に起動するプログラムの記録
Registry 系	アプリケーションの設定内容等
User ディレクトリ	.word や.txt など個人のもの

小さなファイルの場合

基本情報	ファイル名	データ	空き
------	-------	-----	----

大きなファイルの場合

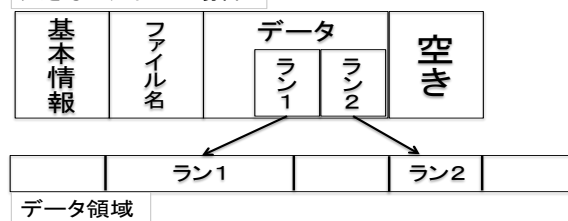


図 1 各 MFT エントリの構造

NTFS ファイル内の各ファイルに対する情報は持っている。各ファイルに関する情報はいずれかの 1 つのファイルレコードから得られる。この MFT を基準にすべてのデータが格納されている。ファイルの属性数が多い場合やファイルの断片化が進行している場合は、1 つのファイルに複数のファイルレコードが必要になる。MFT は属性情報により、ファイル名、ファイル属性、ファイル位置情報が加わる。MFT の属性情報は図 1 より、データが十分に小さなファイルの場合は常駐属性という。データが大きなファイルの場合は、非常駐属性という。非常駐属性は、値の部分がランと呼ばれるクラスタ領域を単位に表される。各ランは先頭のクラスタ番号とクラスタ数で表現できる。ファイルがすべての連続したランに収められた時は、1 つのランの情報が収められるが、メディアのいくつかの部分に分散して収められたときは、そのランの数だけ情報が収められる。メディア内でファイルがあまりにも細分化され、多数のランに分割されたときは、1 つのファイルレコードには収まりきらなくなるため、\$ATTRIBUTE_LIST と呼ばれる属性を用いて、複数のファイルレコードを連結して使用する。[5][6] すなわち、このときは例外的に 1 つのファイルが複数のファイルレコードに対応することになる。ファ

図 2 MFT の NFTS メタデータファイルのファイルレコード

パソコンのメモリの中身が保存されているということは、実行プロセスの状態、実行しているプログラムデータ、ネットワーク接続状況、カーネルモジュール（プログラムに機能を追加しようとする場合、必要に応じて動的に機能を追加することができる機能）、などの情報が含まれている．hiberfil.sys にはパソコンを休止状態にした時点でのメモリ内容のコピーが保存されるため、休止状態のパソコンから hiberfil.sys を抽出してフォレンジック調査することで、その時点でどのようなプログラムやプロセスが動作していたということや各プロセスがどのようなデータ情報を持っているのを知ることができる．また、アプリケーションや文書ファイルが暗号化されていた場合でも、プログラムがそのファイルにアクセスしている場合、例えば、Word ファイルを暗号化しているがそれを編集していると、通常の場合は word ファイルは復号化して格納されるため、その状態で休止状態にすると、hiberfil.sys には word ファイルの内容が平文で保存される．つまり、ファイルが暗号化

Regidry 系は，調査対象のパソコンにおける外部接続機器 (USB メモリーや SD カード，外付け HDD 等) の接続履歴を調査し，それらの機器にファイルがコピーされ持ち出された痕跡がないか調査を行うために使用する．[9] ファイルへのアクセス日時や外部接続機器の接続履歴等を時系列順に整理し，ファイルがコピーされた決定的な証拠を見つけ出すことなどに使用するためでもある．近年では，音楽プレーヤーや携帯電話にも外付けのデータストレージとして，ファイルをコピーできるものが存在しており，場合に

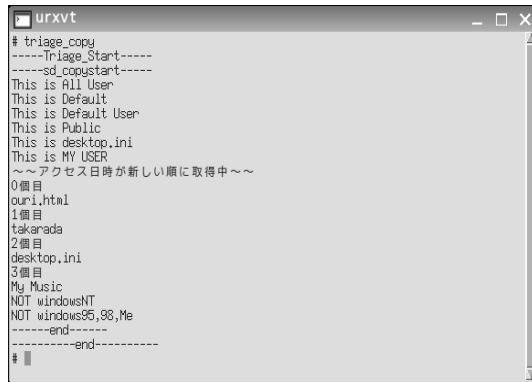


図 3 スクリプトの実行例

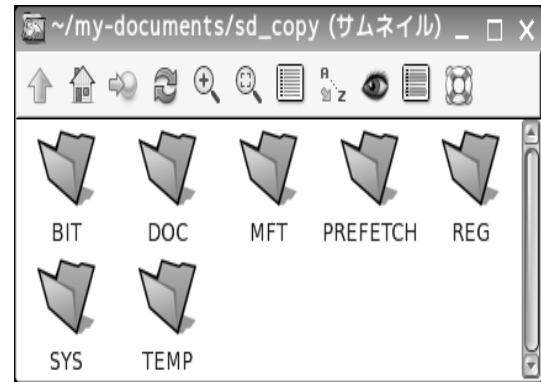


図 4 スクリプトの実行後のファイル内容

よってはこれらの電子機器も調査対象になる可能性がある。

3.3.8 User ディレクトリ

User ディレクトリ自体は、犯罪などに使われているファイル等は削除されている可能性が高いが、しかし、個人のファイルの中身から、更新日時が早いもの、すなわち、利用頻度が高いもしくは犯罪を起こす直前に使用した可能性が高いファイルがある場合があるため、個人ファイルの中身自体は多くは必要ではないが、ある一定数は必要だと考える。

4. 実装結果

一部の記憶領域の保全を行うにあたって、2つのスクリプトをOSの中に実装した。1つは原本同一性を保証する、起動時のスクリプトを実装した。2つ目は一部の記憶領域のみを保全する、部分的証拠保全スクリプトを実装した。また、保全すべき領域を現場で同定するためのツールとしてのLinux版Autopsyの日本語化について報告する。

4.1 起動時スクリプト

起動時スクリプトとは、証拠保全でコピーをした際、コピーの内容に変更を加えられていないことを保証するため、読み取り専用(Read Only)で起動しなければならない。

Linuxでは、起動時のマウントの変更を行うには、rootに存在するStartupファイルの中に新しくスクリプトを記述する必要がある。今回は「automount」というファイルを作成し、その中に起動時スクリプトが記述されている。記述内容はmnt(マウント)のファイルの箇所に[mkdir]コマンドでファイルを/devの中に存在するすべてのデバイスファイルのマウント先を作成する。その後、[mount]コマンドとオプション[-r]ですべてのデバイスファイルを読み取り専用にして/mntのファイルの箇所にマウントするというスクリプトである。これにより、起動した際にすべてのHDDやUSBなどの電磁記録媒体が、読み取り専用でマウントされるようになる。



図 5 Autopsy 日本語化の例

4.2 部分的証拠保全スクリプト

部分的証拠保全スクリプトとは、3.3で述べた箇所を証拠保全するものである。linuxでは、rootの直下に「trriage_copy」というファイル名で証拠保全実行スクリプトを作成した。このスクリプトを実行すると、図3のようになる。記述内容は、コピー先のフォルダを作成している。コピー先は「sd_copy」というフォルダを作成している。また、図4のように、種類別に分類している。分類種別は、7種類のフォルダに分類している。BITには、\$Bitmapが保存されている。MFTには\$MFTが保存されている。REGにはレジストリファイル系とNTUSER.DATを保存されている。SYSにはhiberfilとpagefileを保存されている。PREFETCHには、Prefetchファイルが保存されている。TEMPにはTempが保存されている。DOCには個人ファイルが保存されている。個人ファイルはすべて証拠保全すると莫大なファイル量になる可能性が高いため、設定では更新日時が早い順に並べ替えて、上から順にコピーをとる設定になっている。このスクリプトを実行するには、コマンドラインで「trriage_copy」と入力する。そこで、自動的にUSB上にファイルが作成され、必要部分のみ証拠保全される。

4.3 Autopsyの日本語化

Autopsyの日本語化は、必ずしも英語が堪能ではない警察の専門員が保全すべき領域を現場で同定するためのツールを利用しやすくするために行っている。図5は日本語化

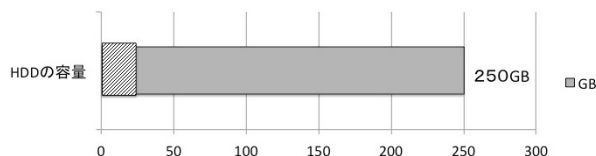


図 6 250GB の HDD のデータ領域の使用量

している例となる．Autopsy は，Perl のプログラムで構成されており，その Perl のプログラムの中に HTML タグを埋め込むことで Web ブラウザのインターフェースを提供している．そのため，Perl 内の英語文章を日本語文章に変更している．英語の部分は基本的に日本語に修正を行っているが，画像でアイコン表示しているタイプは，英語のままのため，画像の差し替えを行い，アイコンも日本語にする予定である．

4.4 タイムスタンプ署名の実装

電磁的証拠は，紙文書などの実際に存在する，存在していたものと比較して，改ざんが容易，改ざん跡が残らず改ざんを事後に検出することが困難，記録媒体の経年劣化による内容消失の可能性など，電磁的証拠特有の脆弱性を持っている．そこで，電磁的証拠を証拠として「完全性」の機能をもたせるためにタイムスタンプ署名を用いる．タイムスタンプ署名は，タイムスタンプに刻印されている時刻以前にその電子文書が存在していたこと（存在証明）と，その時刻以降、当該文書が改ざんされていないこと（非改ざん証明）を証明するものである．[10] そのため，必要とされる電磁的証拠のみ抽出したファイルを対象に署名を行う．しかし，今回タイムスタンプ機能については実装できていないため，今後実装する予定である．

5. ツールの使用時のデータ

ツール使用時のデータとして，今回作成した部分的証拠保全ツールと「これ do 台 Hi-Speed (KD25/35HS)」という完全コピーを行う製品で証拠保全のスピードを比較したグラフが図 7 である．今回は，250GB の Seagate 製 HDD を使用して実験を行った．250GB の HDD のデータ領域の使用量は図 6 で表しており，データ領域の使用量は 24GB となっている．

図 7 の上段のグラフである「ツール使用時」は，今回作成した，証拠保全ツールを使用した時の結果である．証拠保全ツールを使用して，部分的な証拠保全を行った場合，11 分 21 秒で証拠保全が終了した．一方，「高速完全コピーツール使用時」は，1 時間 39 分 11 秒，証拠保全に時間が必要としている．証拠保全の実行時間の差として今回作成した証拠保全ツールが 8.8 倍早く動くことが確認できた．そのため，証拠保全の時間短縮は大いに達成でき，有用性は高い．しかし，データ領域の使用量が 24GB であるため，

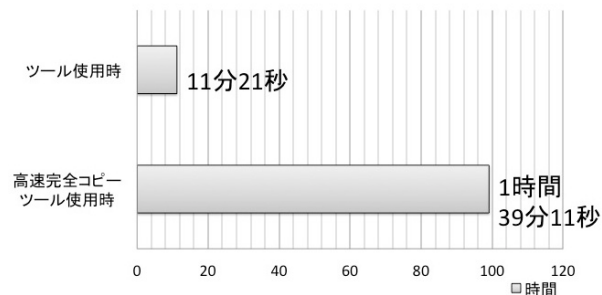


図 7 証拠保全時間の比較グラフ

11 分 21 秒で終了しているが，これが 100GB 等大きくになった場合，時間が延びることが予想されるため十分に高速化しているとは言えない．そのため，ツールの複製速度に関して改善の余地がある．

6. 今後の課題

今後の自作ツールの課題としてあげられるものは大きく分けて二つある．一つ目に原本同一性の保証を行うためのタイムスタンプ電子署名の実装である．コピー元の原本と同一であることを証明するため，電子データにタイムスタンプを付与することで，今ここに確実に存在していたという「存在時刻」とタイムスタンプ付与時の状態である，「完全性」が証明することができる．そのため，証拠保全ツールでコピーを行うと同時にタイムスタンプ電子署名という形で原本同一性の保証を行うことが今後の課題の一つである．二つ目に削除済みデータの復元である．ファイルを復活するには様々な方法があるが，一つ目に実装したいのが，MFT を復活させて MFT の内容から該当するデータ領域を復活させるという作業である．これは，\$Bitmap の操作が主になると考えられる．二つ目に削除ファイルのメタデータが壊れていた場合の復活も視野に入れていきたい．アイデアとして考えられるのは，セクター同士の連続性を判定して復活させる考え方と削除されたファイルのヘッダとフッタを検出して最尤法で連続していない部分を見つけ，その断片化しているところをつなげ合わせるといふ考え方の二つが思い浮かぶ．この二つの原本同一性の保証と削除済みデータの復元は確実にやりたい．他にも，他の OS への対応，ユーザインターフェースの向上などの各種ブラッシュアップも課題としてあげられる．

参考文献

- [1] 佐々木良一：改訂版 デジタルフォレンジック事典，日科技連 (2014)．
- [2] 特定非営利活動法人デジタル・フォレンジック研究会： デジタルフォレンジックとは ，入手先 <<https://digitalforensic.jp/home/what-df/>>
- [3] 上原哲太郎： デジタルフォレンジック 電磁的証拠の収集と分析の

- 技術 ,
情報処理, 48 巻, 8 号, pp.890-893, 2007.
入手先 <<http://ci.nii.ac.jp/naid/110006368639>>
- [4] Autopsy , 入手先 <<http://www.sleuthkit.org/index.php>>
- [5] Mark E Russinovich,David A. Solomon,Alex Ionescu : インサイド Windows 第 6 版 上 , 日経 BP 社 (2012).
- [6] Mark E Russinovich,David A. Solomon,Alex Ionescu : インサイド Windows 第 6 版 下 , 日経 BP 社 (2012).
- [7] 特定非営利活動法人デジタル・フォレンジック研究会 : 「証拠保全ガイドライン 第 3 版」 , 入手先 <<http://www.digitalforensic.jp/eximgs/20130930/gijutsu.pdf>>
- [8] Prefetch Files Revisited ,
入手先 <<http://42llc.net/index.php>>
- [9] レジストリとは ,
入手先 <<http://www.lifehacker.jp/2010/03/100318/registrywtf.html>>
- [10] IPA(情報処理推進機構) : タイムスタンプとは , 入手先 <<https://www.ipa.go.jp/files/000013707.pdf>> (2014)
- [11] 情報セキュリティ白書 2014 , 情報処理推進機構 (2014).