Vol.2013-DPS-154 No.14
Vol.2013-CSEC-60 No.14
2013/3/14

時間経過に着目した HDD のデータ復元に関する実験と解析

林健[†1]　　佐々木良一[†1]

ファイルの削除と復元に関連して，自身がデータを削除したつもりでも復元技術を使ってデータが復元されてしまう場合と，間違えて消してしまったことに気づきデータを復元しようとする場合の 2 つの問題がある．自身がデータを削除したつもりのデータが消えていない場合についての研究はすでに行われており，様々な問題点が指摘されてはいる．しかし，間違えて消してしまったデータを復元する場合については，いつまで復元できるのか，パソコンによる作業によって復元の精度がどのように変化するかも明らかでない．そこで，間違って消してしまったファイルが何時まで残っているのか，復元を行うまでの間の作業内容による復元の可否を明らかにする為に，有償と無償のソフトを用いたデータ復元の特徴と時間経過に伴う復元精度ついての調査が必要であると考え,これらを明らかにする為に，時間経過に伴う HDD のデータ復元の確率について実験を行った．本論文ではこの実験の結果とそれに対する考察を報告するとともに考察を行う．

Experiments and considerations of data restoration of HDD focused on the passage of time

KEN HAYASHI[†1]
RYOICHI SASAKI[†1]

Concerning deletion and the reconstruction of data, there are two issues: (1) The reconstruction of the data is performed by others, though user of the data thought that the data was deleted.　(2) The user of the data tries to reconstruct the data deleted by mistake. The study to prevent the issue (1) has been already performed. However the study on the issue (2) is not enough. The report with regard to the reconstruction probability depending on the time after deletion cannot be found. Therefore, we made the experiment to find the reconstruction probability depending on the time after deletion by using the two type of recover software. This paper reports the result of the experiments with the consideration to the result.

1. はじめに

近年様々な企業で情報漏洩が問題になっており，個人情報の情報管理の重要性が増してきている[1]．NPO 日本ネットワークセキュリティ協会(JNSA)が公開した「2010 年情報セキュリティインシデントに関する調査報告書～個人情報漏えい編～」報告書によるとパソコンなどの誤操作の個人情報漏洩が 32%，データの紛失，置き忘れによる個人情報漏洩の件数は 13%となっており，ヒューマンエラーによる個人情報漏洩の件数は全体の 45%と半数に近い値となっている．特に近年では個人が所有するパソコン内にも重要なデータが保存されており，個人のパソコンからのデータの抜き取りや復元による個人情報の漏洩問題が問題となっている．そのため，様々な場面でデータの復元と抹消の技術[2]は重要となる．

しかし，個人によるファイルの復元と抹消を行う場合，自身がデータを削除したつもりでも復元技術を使ってデータが復元されてしまう場合と，間違えて消してしまったことに気づきデータを復元しようとする場合の 2 つの問題がある．

自身がデータを削除したつもりのデータが消えていない場合についての研究[3]はすでに行われており，様々な問題点が指摘されてはいるが，間違えて消してしまったデータを復元する場合については，いつまで復元できるのか，パソコンによる作業によって復元の精度がどのように変化するかも明らかでない．

そこで，間違って消してしまったファイルが何時まで残っているのか，復元を行うまでの間の作業内容による復元の可否を明らかにする為に，有償と無償のソフトを用いたデータ復元の特徴と時間経過に伴う復元精度ついての調査が必要であると考えた．

本研究では，これらを明らかにする為に，時間経過に伴う HDD のデータ復元の確率について実験し，考察を行った．

2. データの復元と抹消

2.1 ファイル構造

ファイルはインデックス部と実データ部の 2 つから構成されており，普段我々が扱っているのは実データの部分

†1 東京電機大学
Tokyo Denki University.

である．インデックス部にはファイル名やタイムスタンプ等のファイルの属性とファイルの位置情報にまとめられている．

2.2 データの削除とは

一般的なデータを削除する方法では，ゴミ箱を空にする，フォーマットを行うなどの作業は，実際は，ファイルのインデックス部のみが削除されたと印を付けてデータはその領域にないと認識されて他のデータで上書きできる状態になるだけで，パソコン上のデータ本体は削除されずに，ファイルシステムが指し示すデータの格納場所などが記載されたインデックス部のみが削除するだけなのでほとんどの情報が残っている事となる．そのためデータ削除を行う事で，見かけではファイルが存在はしなくなっても，HDD内にはデータの残骸である実データ部が残存しているため，データ復元ツールを使用する事によってデータを復元できることが多い．

2.3 なぜ復元が可能なのか

データの復元が可能になるかどうかは，ファイルシステムによって異なる．

一つは実データ部のみの復元で，これはドライブ全体を検索し，実データを全て復元していく方法である．この方法の特徴として，ドライブ全体を検索するので，時間がかかるが復元の可能性が高いという事がある．

二つ目にはインデックス部と実データ部の復元がある．これは，消されてしまったインデックス部を検索し，実データ部領域に残ったデータ領域インデックス部を解析して，ファイルのインデックス部を推測する．この方法の特徴として，短時間で復元ができる点が上げられる．しかし，インデックス部をたよりに類推するため，インデックス部と実データ部の両方がなければ復元ができないことがあげられる．

2.4 ファイルシステムとは

OS がファイルを管理し，データを読み書きできるようにする仕組みである．記憶装置にファイルやフォルダを作成したり，移動や削除を行ったりする方法や，データを記録する方式，管理領域の場所や利用方法などが定められている．ファイルシステムは OS の持つ機能の 1 つとして提供され，OS ごとに異なるファイルシステムを用いている．

2.5 FAT と NTFS について

Windows で使用されるハードディスク（HDD）には「NTFS」と「FAT32」の 2 種類のファイルシステムがある．

FAT32 とは Windows で採用されているファイルシステムである FAT（File Allocation Table）のひとつで、テーブル長が 32 ビット値のクラスタ識別子で管理される方式のことである．

NTFS とは Microsoft 社の OS である Windows NT/2000/XP で使われるファイルシステム．NTFS は複数ユーザがアクセスするサーバでの運用を念頭において設計されているため、MS-DOS や Windows 95 などの FAT/FAT32 ファイルシステムにはない，アカウントごとのアクセス権設定機能を持つ。また、Windows 2000 以降で使われている NTFS（NTFS 5 とか NTFS 2000 と呼ばれているもの）はジャーナリングファイルシステムとしての機能を持つようになり，突然の停電などに見舞われた時にデータが失われる可能性が低くなっている．また，Windows 2000 の NTFS ではファイルシステムレベルでの暗号化なども可能になっている．

現在では一般的に，内蔵 HDD に NTFS，外付け HDD に FAT32 が使用されていることが多い. 外付け HDD に FAT32 が使用されている理由としては，FAT32 が様々な OS に対応しているためで，どの種類の OS が入ったパソコンに接続しても使用できるためである．また，ファイルシステムは HDD をフォーマットして再設定する際に変換可能である．

本研究では[4][5]ファイルシステムの「NTFS」と「FAT32」の違いから NTFS 形式を使用して実験を行う．

2.6 復元ツール

復元を行うに際して復元ツールを使用して，削除したデータの復元を行う．復元ツールには様々な物があるが[6]ファイルシステムと時間経過に着目した HDD のデータ復元に関する研究より「DataRecovery」「簡単ファイル復活 2」「Recuva」3 つの無償の復元ツールを用いて実験を行った．その中でも一番復元率のよかった Recuva を今回は使用して実験を行った．

また，今回は無償の復元ツールの他に有償の復元ツールである AccessData 社の「Forensic tool kit」[6]を使用して無償と有料の復元ツールを用いたデータ復元の特徴について調査を行う．

Forensic tool kit とは無償の復元ソフトに比べて復元に際して様々さ設定を行う事ができ，色々なケースに合わせた復元をする事はできる．また，ファイル形式別の仕分けや E-mail やグラフィック等の調査目的別にタグが搭載されているためデータの素早いアクセス・回覧をすることができる．

Recuva とはスキャンスピードが速く，復元前にファイルをプレビュー表示する事ができる．検出されたファイルの状態を 4 段階で評価する機能がある．また，名前の一部やファイルの種類でフィルタリングできる．

Vol.2013-DPS-154 No.14
Vol.2013-CSEC-60 No.14
2013/3/14

3. 実験内容

3.1 時間経過に伴う復元実験

内蔵 HDD の C ドライブのデスクトップ上とマイドキュメント直下の 2 カ所で PPT，DOCX，JPG の三種類のファイル形式および 100KB，1MB，10MB の三種類のファイルサイズを用いて，それぞれのファイルを削除後にどのような作業を行ったのか，どの程度の時間が経過したのかによってどの程度復元できるか調査するデータ復元実験を行った．

今回、上記の拡張子とファルサイズを用いた理由としては自分たちがよう使用する拡張子であるということとそれらの拡張子のサイズとしてよく使用されているサイズだと考えられた為である．

3.1.1 実験環境

以下の OS 環境で実験を行った．
実験は研究室のメンバー3 人で行った．また実験時のそれぞれの環境を A，B，C として表記する．

今回復元を行った際に使用した HDD のファイルシステムは NFST を使用した．

表 1. 実験の OS 環境

Table1. OS environment of the experiment

	OS	Memory	CPU	HDD 容量
A	Windows7	16.00G	Intel Core i7	160G
B	Windows7	4.00G	Intel Core i7	500G
C	Windows7	2.00G	Intel Core i7	200G

実験に用いるファイルとして以下の物を用意した．

① Microsoft Office Word
ファイル名は「desktop.docx」「documents.docx」，サイズは 100KB，1MB，10MB である．

② Microsoft Power Point
ファイル名は「desktop.pptx」「documents.pptx」，サイズは 100KB，1MB，10MB である．

③ Microsoft Excel
ファイル名は「desktop.docx」「desktop.docx」，サイズは 100KB，1MB，10MB である．

また，復元ソフトとして①無料の復元ソフト Recuva，②有償の復元ソフト Forensic tool kit を用いた．

3.1.2 実験方法

実験の手順を以下に示す．画像の場合を例に挙げて説明する．

①実験環境を記載する
②指定の場所からファイルをダウンロードする
　このとき DL するファイル形式は PPT，DOCX，JPG のいずれかであり，ファイルサイズも 100KB，1MB，10MB のいずれかを選ぶ
③ファイルは C ドライブのマイドキュメント直下とデスクトップ上の 2 か所に保存する
④ファイルの中身を確認し，保存したファイルを削除する
⑤ごみ箱を空にし，消えたことを確認する
⑥パソコンを普段通りに使用し，そのときの使用方法を記録する
⑦復元ソフトを使用する
　このとき復元ソフトはファイル削除直後（11 時）から始め，一日に 2 回（11 時と 17 時に）復元を行う
⑧該当するファイルが存在するか確認する
　復元したファイルは実験に影響が出る可能性があるため，のドライブにまとめて保存を行う
⑨復元が不可能になるまで⑥～⑧を繰り返し行う．

3.1.3 実験結果

以下に主要なデータとして PPT，DOCX，JPG の 3 種類の拡張子を用いて実験した復元率と復元時間を表と図で示す. 今回の実験では拡張子ごとに 10 回の実験をそれぞれ行ない，そのときの復元率を表記した．また，今回はマイドキュメント直下とデスクトップ上の 2 カ所で実験を行ったが得られた結果が同じだったため一つにまとめて表記した．

表2. DOCXの場合の実験結果（復元率）

Table2. Experimental results in the case of DOCX (restoration rate)

	サイズ	削除直後	6時間後	24時間後	30時間後
Recuva	100KB	8/10	4/10	0/10	-
	1MB	8/10	5/10	0/10	
	10MB	8/10	4/10	0/10	
FTK	100KB	8/10	6/10	0/10	
	1MB	8/10	6/10	0/10	
	10MB	8/10	5/10	0/10	

Vol.2013-DPS-154 No.14
Vol.2013-CSEC-60 No.14
2013/3/14

表3.　PPTの場合の実験結果（復元率）

Table3.　Experimental results in the case of the PPT (restoration rate)

	サイズ	削除直後	6時間後	24時間後	30時間後
Recuva	100KB	7/10	5/10	0/10	-
	1MB	7/10	5/10	0/10	
	10MB	7/10	4/10	0/10	
FTK	100KB	9/10	6/10	0/10	
	1MB	9/10	5/10	0/10	
	10MB	8/10	4/10	0/10	

表4.　JPGの場合の実験結果（復元率）

Table4.　Experimental results in the case of the JPG (restoration rate)

	サイズ	削除直後	6時間後	24時間後	30時間後
Recuva	100KB	8/10	5/10	0/10	-
	1MB	8/10	5/10	0/10	
	10MB	7/10	4/10	0/10	
FTK	100KB	9/10	5/10	0/10	
	1MB	9/10	5/10	0/10	
	10MB	9/10	5/10	0/10	

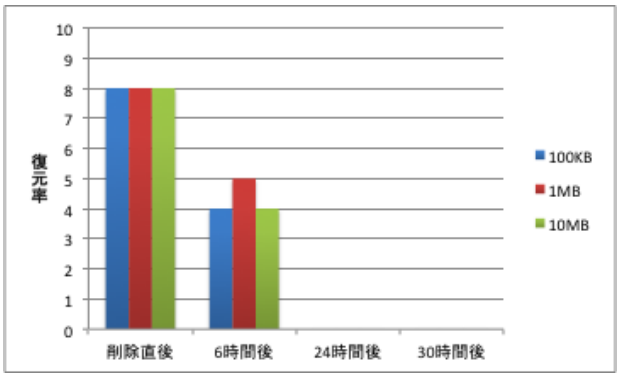

図1.　Recuvaを用いたDocxの実験結果（復元率）

Figre1.　Docx experimental results using Recuva (restoration rate)

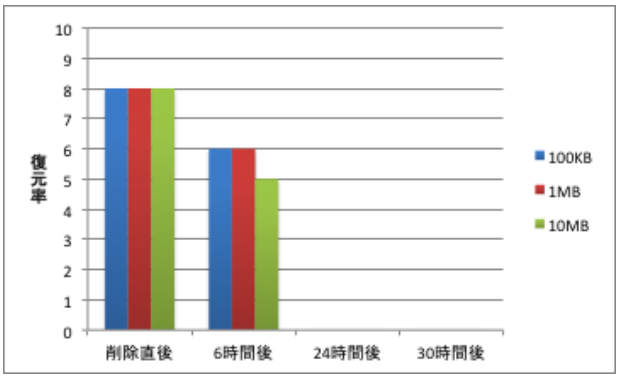

図2.　FTKを用いたDocxの実験結果（復元率）

Figre2.　Docx experimental results using FTK (restoration rate)

3.2　ドライブ別による復元実験

内蔵 HDD の C ドライブと D ドライブで DOCX ファイルおよび 100KB，1MB，10MB の三種類のファイルサイズを用いてドライブごとによる時間経過に伴う復元の差異を調べる復元実験を行った．

3.2.1 実験環境

以下の物を用意した．

①Microsoft Office Word

ファイル名は「test.docx」，文書や画像の入ったデータでページ数は 10 ページである．

②復元ソフト

Forensic tool kit

3.2.2 実験方法

実験の手順を以下に示す．

①実験環境を記載する

②指定の場所からファイルをダウンロードする

③ファイルは C ドライブ，D ドライブともにドライブ直下に保存する

④ファイルの中身を確認し，保存したファイルを削除する

⑤ごみ箱を空にし，消えたことを確認する

⑥パソコンを普段通りに使用し，その時の使用方法を記載する

⑦復元ソフトを使用する

このとき復元ソフトはファイル削除直後（11 時）から始め，一日に 2 回（11 時と 17 時に）復元を行う

⑧該当するファイルが存在するか確認する

復元したファイルは実験に影響が出る可能性があるため．別のドライブにまとめて保存を行う

⑨復元が不可能になるまで⑥～⑧を繰り返し行う

3.2.3 実験結果

以下に復元不可能になるまでに要した時間と結果を表にしたものを示す.
今回はC ドライブ，D ドライブでそれぞれのファイルサイズごとに 10 回の復元実験を行い, そのときの復元率を表記した.

表5.　DOCXの場合の実験結果（復元率）
Table5.　Experimental results in the case of DOCX (restoration rate)

	サイズ	削除直後	6時間後	24時間後	30時間後	48時間後	54時間後
Cドライブ	100KB	8/10	4/10	-	-	-	-
	1MB	8/10	5/10	-	-	-	-
	10MB	8/10	4/10	-	-	-	-
Dドライブ	100KB	10/10	10/10	10/10	10/10	10/10	10/10
	1MB	10/10	10/10	10/10	10/10	10/10	10/10
	10MB	10/10	10/10	10/10	10/10	10/10	10/10

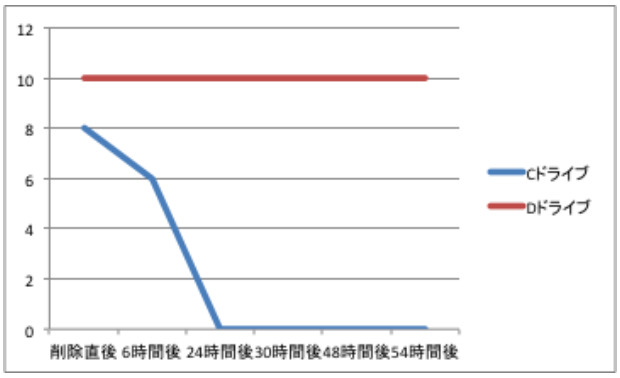

図3.　CドライブとDドライブを用いたDocxの実験結果（復元率）
Figre3.　Docx experimental results using the C drive and D drive (restoration rate)

3.3 考察

実験 1 の結果から C ドライブ上にあったファイルは削除してから 24 時間後にはファイルを復元する事ができなくなる事が分かった．また，表 2，3，4 より削除してから 6 時間後には実験を行った回数の約半分の回数しかしか復元できず、復元精度は約半分にまで落ちる事が分かった．
また，ファイルサイズ別の復元では大きな差はみられず，削除直後から 6 時間後には 10 回の実験のうち 5 回前後しか復元できずに復元率は半分となっていることがわかった．
復元ソフト別では，Recuva に比べて FTK のほうが全体的に 10 回の実験のうち 1，2 回の回数だけ復元率が高いことが分かった．今回の実験で A,B,C の実験環境で復元作業を行ったが, PC の違いによる復元率の変化は見られる事ができず，また，作業内容による復元率の変化もあまり見る事ができなかった．

実験 2 より D ドライブで削除されたファイルは完全に消えるまでに最低でも 3 日必要なことがわかった．ファイルサイズ別ではそれぞれの復元率に差は見られず全て同様の結果となった．

また，今回データを削除して直後に復元ツールを用いて復元を行ったが，復元が不可能となっていた時があった．データの復元が不可能にするためにはファイルの実データの上書きを行うなどの作業が必要であり，WindowsOS が何かしらの動作を行ってデータに影響を与えている可能性があったが, 削除直後では考えられない結果となっていた．考えられる要因としては，復元作業によってデータの復元は可能であったがデータを復元する際にファイル名が文字化けしてしまい見つける事ができなかった可能性などがある．

4.　今後の展望

今後の課題として，実験 1 での実験回数を増やしてもっと正確な復元率の傾向をつかむ事，拡張子を増やす，ファイルサイズをより大きいもと小さいもの使用して実験を行う，ファイルを保存する場所を変更したりする．
実験 2 では D ドライブ上でデータがどれくらいの期間復元を行う事ができるのか，C ドライブとの際をより詳しくする必要がある．D ドライブではデータの書き込みがほとんど行われず，削除されたデータは完全に復元が可能となっていた．最長で 3 日間の復元が可能となっていたが，4 日目以降はどれくらい復元が可能なのかが分からないため，さらなる調査が必要となる．

実験環境では HDD の残り容量による復元率の差異，HDD 内に存在するデータ数による復元率の差異，HDD の使用されている年月による復元率の差異を調査する．
データの復元に関しては，復元ソフトを増やし，拡張子ごとの復元率の差異，ファイルサイズ別の復元率の差異，データの復元を行った際に復元したファイルとオリジナルファイルのハッシュを比較し，完全にデータを復元できたのかを調査する，FTK の機能をより使いこなすなどがある．
また，データの削除直後に復元できない時がある事について何が原因となっているのかを調査するなどがあげられる．

Vol.2013-DPS-154 No.14
Vol.2013-CSEC-60 No.14
2013/3/14

5. まとめ

本稿では，データ復元に関する疑問を解決するために，間違えて消してしまったことに気づきデータを復元しようとする場合に関しての実験を行いデータ復元の特徴と時間経過に伴う復元精度ついての調査を行った．
今回の実験を通して分かったことはCドライブでの復元実験ではPPT，DOCX，JPGの三種類のファイル形式および100KB，1MB，10MB の三種類のファイルサイズで，データは削除してから24時間以内には上書きされてしまい，復元を行うことができない状態になっているとこが明らかになった．復元率ではファイルサイズ別に見るとあまりそれぞれのサイズでの差はあまり見られなかったが，拡張子で見ると時間経過による復元率としては DOCX が最も復元率が良かった．

また，復元ソフトとしてはFTKとRecuvaではFTKのほうが拡張子でもファイルサイズでも全体的に復元率は優れていた．

また，Cドライブと Dドライブを用いた復元実験からはDドライブのデータの復元可能期間がとても長い事が分かった．大事なデータはDドライブに保存しておく事で，データを誤って削除してしまったとしても，復元を行う事が可能となる確率が高くなると考えられる．

参考文献

1) 「2010 年情報セキュリティインシデントに関する調査報告書」:
http://www.jnsa.org/result/incident/data/2010incident_survey_PIL_v1.4.pdf
2)フリーソフトによるデータ抹消・復元大全:
http://www.cybernetic-survival.net/w_s.htm
3) パソコンの破棄・譲渡時におけるハードディスク上のデータ消去に関するガイドライン:
http://home.jeita.or.jp/page_file/20110511155520_8vAEy2Fi5d.pdf
4)ファイルシステム「NTFS」と「FAT32」の違い:
http://tennensui.sakura.ne.jp/hddrecover/category8
5) データ障害に強い FTFS と FAT の違い:
http://jisaku-pc.net/hddhukyuu/archives/741
6) AccessData:
http://www.accessdata.com/products/digital-forensics/ftk