データ消去済み SMR 方式 HDD からのデジタル証拠復元を主眼としたデジタルフォレンジック調査の有用性

下垣内 太 [1,a)]

概要：企業や組織における営業秘密の不正持出しや背任などの事件において，証拠隠滅目的でデータが消された PC やサーバを対象にデジタルフォレンジック調査を行うことがある．しかし，デジタル証拠が復元される率は近年低くなりつつある．その要因として削除データの復元が困難な SSD の普及が挙げられる．こうした特性を持つデバイスが増えるほど，消されたデータの復元は困難になる．それでも被害企業は，不正や犯罪の証拠を集めるしかない．それが無理ならば泣き寝入りすることになってしまう．そこで被害者救済の観点から，データの消去処理済みであっても，デジタル証拠の収集が期待できる媒体として SMR 方式の HDD に着目した．SMR 方式の HDD ならば論理アドレスのないデータ記録領域があるはずで，被害や損害の立証に有用なデジタル証拠を復元できる可能性が残されていると考えられたからである．本稿では，そうしたデータ記録領域の容量を推定すべく，SMR 方式 HDD ついて物理セクタ数を調査した結果をまとめ，さらに日本及び米国で購入したデータ消去済みの SMR 方式 HDD から，機密ファイルが復元された実証実験についても報告する．

キーワード：デジタル証拠，フォレンジック，データ消去，SMR，HDD

Usefulness of digital forensic investigation focusing on digital evidence recovery from wiped SMR HDDs

Dai Shimogaito[1,a)]

Abstract: Digital forensic investigations are sometimes conducted on PCs that have been wiped to destroy evidence in cases of data theft or other crimes. However, the recovery rate of digital evidence has declined recently due to the widespread use of SSDs, which make it difficult to recover deleted data. In contrast, SMR HDDs could be a medium of interest for forensic investigations because they may contain data areas without LBAs, potentially allowing the recovery of digital evidence even after erasure. This paper investigates the number of physical sectors on SMR HDDs to estimate the capacity of these data areas. It also reports on a demonstration experiment where confidential files were recovered from wiped drives purchased in Japan and the U.S., highlighting the potential of SMR HDDs for supporting victims in digital forensic cases.

Keywords: Digital Evidence, Forensics, Data Erasure, SMR, HDD

1. はじめに

犯罪や不正におけるデジタルフォレンジック調査において，パソコンやサーバを解析したところ，証拠隠滅を目的としたデータ消去が行われていたことが原因で，デジタル証拠を十分に収集できないことがある．

もちろんデジタルデータには，消す方法や媒体特性によって消え残り方に違いがあり，その残存状況によっては復元できる場合もある．しかし全論理セクタ領域のバイナリ値がゼロの場合はどうだろう．これはデータ消去規格である NIST SP800-88 Rev.1[1]及び IEEE 2883-2022[2]が Clear と規定するデータ消去後の状態と一致するものであり，同領域に対して従来法のデータの証拠保全がいかに適切に行われ，かつ入念なデジタルフォレンジック解析が実施されたとしても，確保できるのはゼロ値のみであるため，消されたデータの復元及び検出は不可能である．

証拠隠滅により消えたデジタルデータが復元できない場合，犯罪や不正事件の被害者にとっては「デジタル証拠が無い.」という状況に置かれることになる．それにより加害者と被害者を公平かつ客観的に判断する材料が揃っていないにもかかわらず「証拠が無いから，被害は無い.」と結論付けられてしまえば，被害者は泣き寝入りするしかない．デジタル証拠を必要とする被害者にとって，加害者による証拠隠滅は，明らかに不利に作用するが，その一方で，NAND 型フラッシュメモリにおけるブロック単位のデータ消去[3]のように，消されたデータの復元ができない特性をもつ媒体及び装置の普及は進んでいる．

ここでデータ消去済み媒体からのデジタル証拠の復元の可能性について改めて考えてみる．まず前述のとおり物理的に消滅したデータの復元は不可能である．ならば復元の余地はどこにあるのか．それは，媒体のどこかで消えずに残っているデータを探し，かつ消えてしまう前までに読み

1 アイフォレンセ日本データ復旧研究所 株式会社
Aiforense Japan Data Recovery, inc.
a) dai@daillo.com

出すことである．

SSD ならば全論理セクタ領域のバイナリ値がゼロであったとしても，オーバープロビジョニング領域[4]にデータが残ることは十分に見込まれる．つまりデジタル証拠の残存に期待が持てる領域である．しかしながら事件の発生からある程度の時間を経てデジタルフォレンジック調査に着手する頃には NAND 型フラッシュメモリにおけるブロック単位のデータ消去[3]が実行されていると見込まれることから，デジタル証拠の残存は期待しにくい．

そこで本研究では SMR 方式[5]の HDD の特性に着目した． SMR 方式[5]の HDD には，論理アドレスが割り当てられた経緯のある物理セクタ領域が，トラックの重ね書きを実現するためにドライブの公称容量以上あると推定され，それならば全論理セクタ領域のバイナリ値がゼロであっても消え残るデータの存在に期待が持てると考えたからだ．さらに HDD の特性として SSD のようにデータを消去して書き込みに備える必要がない点も都合が良い．つまり SMR 方式[5]の HDD には，全論理セクタがデータ消去済みであってもデータが残存しうる領域があり，かつ SSD ほど積極的に余剰領域のデータを消す必要性がないことから，事件からデジタルフォレンジック調査までの経過時間の如何によらず，残存データが存在し続ける可能性はあると期待がもてる．

本稿では，まず調査対象として購入した使用済み中古 HDD16 点の仕様を調べ，かつデータ消去検証を行うことで，一般的なデジタルフォレンジック調査ではユーザデータの復元ができない状態にあることを確認した過程をまとめた．加えて HDD には SSD への搭載で知られる TRIM コマンド[6]に対応したものがあることと，物理セクタ容量と公称容量の差分が CMR 方式の HDD よりも SSD に近いことも示している．そして実証実験の結果として，社外極秘の文書ファイルや運転免許証のスキャン画像が復元されたほか，NTFS ファイルレコード[7]が検出された事例についても説明する．

2. HDD の調達と仕様確認

2.1 HDD の調達

デジタルフォレンジック調査における有用性を確認するために，検証用のデータやドライブを自前で用意することはせず，実際の使用済み中古品を購入し，調査の対象とした．2024 年 5 月から 8 月までの期間に，日本と米国で中古品を調達した．なお，購入にあたっては商品説明に，使用済み中古品であり，かつデータ消去処理済みであることが記載されているものを選定条件とした．また HDD が SMR 方式[5]であることは各メーカの WEB サイト等を参考にして型番情報などを集めた．

2.2 調査対象 SMR 方式 HDD16 点の概要

メーカ名（Make），フォームファクタ（Form Factor），ラベル表記容量（Capacity），製造日（Date），および公称容量（Nominal Capacity）を測定した 16 点分の結果を表 1 に示す．なお本稿では物理セクタサイズが 4K 仕様[8]の HDD であったとしても，セクタサイズは一律 512 バイトとして記載する．また，次項以降の表にある Drive #（もしくは，「D.#」と略記．）は，全て表 1 の Drive #と一致する．

表 1　調査対象 SMR 方式 HDD の概要

Drive #	Make	Form Factor	Capacity	Date	Nominal Capacity (Sectors)
1	A	2.5	500GB	2020/10/07	976,773,168
2	A	2.5	500GB	2020/10/08	976,773,168
3	A	2.5	1TB	2018/07/23	1,953,525,168
4	A	2.5	1TB	2019/05/15	1,953,525,168
5	A	2.5	1TB	2018/11/25	1,953,525,168
6	A	2.5	1TB	2019/09/04	1,953,525,168
7	A	2.5	1TB	2020/08/17	1,953,525,168
8	A	2.5	1TB	2018/11/19	1,953,525,168
9	B	2.5	1TB	2018/08/21	1,953,525,168
10	B	2.5	1TB	2019/09/08	1,953,525,168
11	C	2.5	500GB	2019/03/19	976,773,168
12	C	2.5	500GB	2020/03/15	976,773,168
13	C	2.5	500GB	2019/09/08	976,773,168
14	C	2.5	500GB	2019/07/13	976,773,168
15	C	2.5	1TB	2017/05/17	1,953,525,168
16	C	2.5	1TB	2017/04/13	1,953,525,168

2.3 データ消去検証

データ消去済みとして売られていた HDD ではあるが，適正な消去処理[9]が実施されていない可能性を想定し，データ消去検証を行った．

検証には，データ消去規格である NIST SP800-88 Rev.1[1]及び IEEE 2883-2022[2]の Clear 規定を基準とした．これを満たしている場合には，一般的なデジタルフォレンジック解析ではアーティファクト[10]の検出は不可能である．調べた結果，16 点のうち 12 点は Clear を満たしていることが判明した．

残る 4 点のうち 3 点は，おそらく Clear を満たす消去が実施されたあとに NTFS 初期化が行われたものと見込まれる．例えば#11 の HDD は，販売開始日時が 2024 年 4 月 24 日 10 時 34 分であったのに対し，初期化日時が同じ日の 8 時 37 分であることが分かった．この日時情報を信頼するならば，Clear を満たしているともいえよう．いずれにしても残る 4 点についても概ね同等のデータ消去処理が実施済みであることが確認された．

表 2 にはデータ消去検証の結果に加え，各 HDD の前所

有者が使用していたことを示す情報として，S.M.A.R.T.の Power on Hours 値も記載した．なお，#11 のみ記載がないのは筆者による記録忘れによるものである．

表 2　HDD のデータ消去検証結果と使用済時間

D.#	Clear	論理セクタ領域のデータ状態	Power On Hours
1	YES	全論理セクタ x00	12,288
2	YES	全論理セクタ x00	25,808
3	YES	全論理セクタ x00	732
4	-	約 1.8MB の連続領域を除き，全て x00.	3,012
5	YES	全論理セクタ x00	2,881
6	YES	全論理セクタ x00	175
7	YES	全論理セクタ x00	4,372
8	YES	全論理セクタ x00	1,659
9	YES	全論理セクタ x00	4,804
10	-	NTFS 初期化済み．空き領域は全て x00.	1,816
11	-	NTFS 初期化済み．空き領域は全て xFA.	n/a
12	YES	全論理セクタ x00	6,377
13	YES	全論理セクタ x30	1,024
14	YES	全論理セクタ xD4	4
15	YES	全論理セクタ x00	10,715
16	-	NTFS 初期化済み．空き領域は全て x00.	14,942

2.4 TRIM 関連コマンドの対応

TRIM 関連コマンド[6]の対応状況を，Ubuntu 24.04 LTS の hdparm で調べた結果が，次の表 3 のとおりである．

とくに DRAT（Deterministic Read After Trim）[6]及び RZAT（Read Zeroes After Trim）[6]については，ドライブ内部の物理的な記録が消えたか否かにかかわらず読み出し要求に対して値を返す際の機能である．つまり通常の読み出しコマンドで得られる情報は，媒体が保有している情報であるとは限らないことを示しており，前項で検証したような手段では見つからないデータが潜んでいる可能性を示しているともいえよう．

表 3　TRIM 関連コマンドの対応

Drive #	TRIM	DRAT	RZAT
1	YES	YES	NO
2	YES	YES	NO
3	YES	YES	NO
4	YES	YES	NO
5	YES	YES	NO
6	YES	YES	NO
7	YES	YES	NO
8	YES	YES	NO
9	NO	NO	NO
10	NO	NO	NO
11	NO	NO	NO
12	NO	NO	NO
13	NO	NO	NO
14	NO	NO	NO
15	NO	NO	NO
16	NO	NO	NO

2.5 物理セクタ容量

IDENTIFY DEVICE コマンド[11]では物理セクタ容量を知ることができないため，ファームウェアレベルでのアクセスが可能な装置αを介して各 HDD から磁気ディスクにおけるトラックの配置情報（Zone Allocation Table）であるを入手し，表計算ソフトを用いて算出した．なお，ファームウエア用に割り当てられた物理セクタはここでは対象外とし，ユーザデータが記録されうる領域を算入条件とした．そのためバッファ及びキャッシュとして確保された磁気ディスク上の領域も含まれている．

公称容量（Nominal Capacity），物理セクタ容量（Physical Capacity），公称容量と物理セクタ容量の差分（Diff），およびその差分のパーセンテージ表記を 16 点の HDD についてまとめたのが表 4 である．

表 4　公称容量と物理セクタ容量との差分

D. #	Nominal Capacity (Bytes)	Physical Capacity (Bytes)	Diff (Bytes)	Diff (%)
1	500,107,862,016	780,496,310,272	280,388,448,256	56.07%
2	500,107,862,016	781,435,535,360	281,327,673,344	56.25%
3	1,000,204,886,016	1,052,764,663,808	52,559,777,792	5.25%
4	1,000,204,886,016	1,052,772,532,224	52,567,646,208	5.26%
5	1,000,204,886,016	1,052,783,681,536	52,578,795,520	5.26%
6	1,000,204,886,016	1,052,784,099,328	52,579,213,312	5.26%
7	1,000,204,886,016	1,052,792,348,672	52,587,462,656	5.26%
8	1,000,204,886,016	1,052,819,316,736	52,614,430,720	5.26%
9	1,000,204,886,016	1,056,040,349,696	55,835,463,680	5.58%
10	1,000,204,886,016	1,056,306,704,384	56,101,818,368	5.61%
11	500,107,862,016	558,017,822,720	57,909,960,704	11.58%
12	500,107,862,016	566,660,747,264	66,552,885,248	13.31%
13	500,107,862,016	570,464,108,544	70,356,246,528	14.07%
14	500,107,862,016	570,488,713,216	70,380,851,200	14.07%
15	1,000,204,886,016	1,052,769,746,944	52,564,860,928	5.26%
16	1,000,204,886,016	1,062,163,931,136	61,959,045,120	6.19%

3. データ復元の実証実験

調査対象の HDD については，論理アドレスの割り当て領域にユーザデータが無いことが確認された．さらに，公称容量の 5%以上は論理アドレスをもたない物理セクタとして存在することも認められた．なかには 50%を超えるものすらあることも分かった．そこで残存するユーザデータを復元すべく装置αを用いてその領域にアクセスする方法で調査を行い．その結果を表 5 にまとめた．まず 4 点のドライブからは何も見つからなかった．

次に他の 12 点については何かしらの残存データは見つかった．暗号化されているデータの場合は復号用のキーがなければデータを復元できたとまでは言えないがデジタルフォレンジック調査の場合にはファイルコンテンツの可読性だけが問題ではないことから，ユーザデータの復元には至らなくとも有効な情報ないしデジタル証拠を確保できている可能性はある．実際に#3 の HDD は BitLocker が使用されていたとみられるが，搭載されていた PC のメーカ名及び型番の特定には至っており，その内蔵ドライブであったこと等は判明している．

さて 8 点の HDD においては暗号化されていないユーザデータが検出された．NTFS ファイルレコードしか見つからなかったドライブもあるが，文書，表計算ファイル，そして画像ファイルがみつかったものもあった．

本稿は復元できるデータの量に主眼をおくものではないので復元したファイルの数やデータの容量については掘り下げてはいないが，参考までに例を挙げると#2 の HDD からは PDF ファイルだけで 2,799 件が復元され，#13 の HDD からは 142,603 件の NTFS ファイルレコードが復元された．データ復元結果は表 5 にまとめた．

表 5　データの復元結果

D.#	Non-Encrypted Data	Encrypted Data	No Data	Description
1	YES	-	-	業務ファイル検出. 社名判明.
2	YES	-	-	業務ファイル検出. 社名判明.
3	YES	YES	-	BitLocker 使用と推定．PC 機種判明.
4	YES	-	-	業務ファイル検出. 社名判明.
5	-	YES	-	BitLocker 使用と推定.
6	YES	-	-	業務ファイル検出．使用者判明.
7	YES	-	-	業務ファイル検出. 社名判明.
8	-	-	YES	Enhanced Security Erase によるデータ消去と推定.
9	-	-	YES	-
10	-	-	YES	-
11	YES	-	-	NTFS ファイルレコード多数検出.
12	YES	-	-	過去の NTFS パーティション構成判明．ファイルレコード検出あり.
13	YES	-	-	NTFS ファイルレコード多数検出.
14	-	YES	-	暗号化機能を使用と推定.
15	-	-	YES	Enhanced Security Erase によるデータ消去と推定.
16	-	YES	-	暗号化機能を使用と推定.

4. 考察

4.1 結論

まずはデータ復元の実証実験の結果から，データ消去済み HDD であっても SMR 方式[5]ならば十分にデータを復元できる可能性があることが分かった．つまりデジタルフォレンジック調査において従来ならば得ることのできなかったデジタル証拠を確保できる可能性があるということである．

4.2 論理アドレスのない物理セクタ

本稿がテーマにするデータ復元の可能性は論理アドレスを持たない物理セクタにある．CMR 方式の HDD ならばその余剰容量は約 1%と見込まれている．しかも不良セクタの交替処理がなければユーザデータが記録されることはない．そのため通常はユーザデータの検出を期待することができない．つまり CMR 方式の HDD の場合は，今回の調査対象 HDD のようにデータ消去が実施済みの場合には，もうデータを復元できる見込みは極めて低いことになる．ただし一部の製品にはキャッシュ機能として使うものがあり，今回の実験においてもその領域からしか復元できていない可能性がある点は補足しておく．

さて SSD には論理アドレスのない物理セクタとしてオーバープロビジョニング領域[4]があることが知られている．そしてコンシューマ用 SSD のオーバープロビジョニング領域[4]は公称容量の 6～7%であり，表 4 の値が近い数値であることが分かった．SMR 方式[5]の HDD には SSD らしさが取り込まれていることがうかがえる調査結果である．

4.3 データ復元方法の改善の余地

本実証実験では全論理セクタが全面的に消去されたものが対象であったが，その消去範囲がより限定的な場合には，もしかするとより多くのデータを復元できるかもしれない．この点については今後さらに研究の余地がある．

また，本実験では論理アドレスをもたない物理セクタに装置αを用いてアクセスをしたが，そのアプローチを改善することによって，より多くのデータを復元できるかもしれない．この点についても今後さらに研究の余地がある．

4.4 本研究の法的及び倫理的懸念

本実証実験では,使用済み中古品として購入したHDDからデータ復元を行ったことによって，前所有者が保存していた社外極秘と明記された文書や個人情報が含まれたファイルも見つかった．しかもデータ消去規格である NIST SP800-88 Rev.1[1]及び IEEE 2883-2022[2]の両方においてClear規定を満たすデータ消去が実施されたHDDからの検出である．代替セクタ領域等にわずかに残る情報ならば復元される[12]こともあるだろうが，調べなくても事件調査にはまずもって支障がないことから，デジタルフォレンジック調査の対象から除外されているのが実情である．つまり1ファイル復元できるだけでも非常にまれなことであるのに対し，本実証実験ではその 1000 倍以上のファイルが復元できうることが判明した．繰り返しとなるが本研究の目的は，悪意のあるデータ消去行為によってデジタル証拠を復元できず，損害を証明できない被害者をデジタルフォレンジックスの観点から支援することである．その意味では非常に期待を持つことができる技術であることが本実証実験によって確認できた．ただ一方で，悪用されるようなことがあれば大きな問題になりかねない懸念も浮上した．

4.4.1 本稿では復元手順は非開示

本稿では実験結果の悪用を防ぐために，データの復元方法について基本的なアプローチは説明しているが，具体的な手順は説明をしていない．例えばランサムウェア攻撃[13]がそうであるように，企業の機密ファイルが第三者の手に渡った場合にはそのデータの返却と引換えに身代金が要求されるような事件を招きかねない．

なおデジタルフォレンジック調査を行う法執行機関にはすでに技術的な手順の開示を本稿執筆時点（2024 年 8 月）で始めている．

4.5 SMR 方式 HDD からの情報漏えい防止策

本実証実験でも該当するものがあったように，暗号化機能の活用により消去後の HDD からデータが復元されてしまうリスクは無くなる．

もしかするとデータ消去のことはよくわからないからHDDは壊して捨てよう,という意思決定がされているのかもしれないが 2022 年に制定された新たなデータ消去規格である IEEE2883-2022[2]においては破砕は規定から除外[14]されている点には注意が必要だ．媒体におけるデータ記録密度の向上により破砕では不十分と判断されたためである．SDGs 推進の社会的な観点からも，できるかぎり適正なデータ消去の実施とリユースの検討が望ましい．

4.6 本研究継続の意義

本稿ではデータ消去処理済みの HDD からであっても裁判用のデジタル証拠を確保できるか否かを確認するのが目的であった．そしてその結論は YES である．ただし調査対象となったHDDは全て2.5インチであり,製造日が最も新しいものでも 2020 年製である．そのため網羅性を向上させる余地があるものと考える．

4.7 データ消去方法の研究

本実証実験では#8 と#15 の HDD からは全くデータが検知されなかったことから，Enhanced Security Erase[15][16]が実行されたものと推定されている．この機能によるデータ消去は，データ消去規格である NIST SP800-88 Rev.1[1]及びIEEE 2883-2022[2]の両方においてPurgeと規定されている．論理アドレスのない物理セクタ領域のデータを消去する機能を有しているはずだが SMR 方式[5]の HDD においてどの程度の消去性能を有しているかが定かではない．とくにSMR 方式[5]の場合には SSD らしさが融合されている点でも注意が必要であろう．これについては今後新たなテーマとして研究する予定である．

5. おわりに

これまでの事件対応におけるデジタルフォレンジック調査では全論理セクタを調べてもデジタル証拠が検出されなければ，被害者は泣き寝入りするしかないような状況であったが，本実証実験により打つ手はまだあることが確認された．被害者救済のために活用されることが期待できる．

謝辞　本研究を進めるにあたり相談に乗ってくださった捜査機関を含む情報セキュリティの専門家の方々からは貴重なご助言をいただきました．また，アイフォレンセ日本データ復旧研究所のメンバーも，実験のサポートも含め，いつもながらありがとうございます．本稿の内容を鑑みてお名前の記載は控えさえていただきますが，ここに感謝の意を表します.

参考文献

[1] "Guidelines for Media Sanitization". https://nvlpubs.nist.gov/nistpuubs/specialpublications/nist.sp.800-88r1.pdf, (参照 2024-08-23).

[2] Cybersecurity and Privacy Standards Committee of the IEEE Computer Society. IEEE Standard for Sanitizing Storage. IEEE. 2022.

[3] "Read-Modify-Write-Erase". https://www.slideshare.net/slideshow/dbts-osaka-2014-b11-hardware-hironobuasano/36183117#21, (参照 2024-08-23).

[4] “SSD オーバープロビジョニング (OP) を理解する”. https://www.kingston.com/jp/blog/pc-performance/overprovisioning, （参照 2024-08-23）

[5] "SMR 技術とは". https://toshiba.semicon-storage.com/jp/storage/product/articles/what-is-smr-technology.html, (参照 2024-08-23).

[6] American National Standards Institute, Inc.. ATA Command Set - 5 (ACS-5) . 2021, 689p.

[7] “NTFS Master File Table (MFT)”. https://ntfs.com/ntfs-mft.htm, (参照 2024-08-23)

[8] “WD ドライブが 4K かどうかを確認する方法”. https://support-jp.wd.com/app/answers/detailweb/a_id/28713/kw/WD_BLACK%20SN770M%20SSD/related/, (参照 2024-08-23).

[9] “ADEC データ適正消去実行証明協議会 データ消去技術 ガイドブック 第 2.4 版”. https://adec-cert.jp/guidebook/pdf/DATA

WIPEGUIDEBOOK.pdf, （参照 2024-08-23）
[10] 安藤潔，上原哲太郎．基礎から学ぶデジタル・フォレンジック．日科技連，2019，39p.
[11] American National Standards Institute, Inc.. ATA Command Set - 5 (ACS-5) . 2021, 157p.
[12] 上原哲太郎. 神奈川県ハードディスク流出事件 ー HDD 廃棄時にデータ消去はどうあるべきかー. 情報処理, 2020, vol. 61, no. 3, 232p.
[13] “情報セキュリティ 10 大脅威 2024”. https://www.ipa.go.jp/security/10threats/10threats2024.html, （参照 2024-08-23）
[14] “Evolving Data Security: A Comparative Analysis of IEEE 2883-2022 and NIST SP 800-88r1 Standards”. https://circulardrives.org/evolving-data-security/, （参照 2024-08-23）
[15] American National Standards Institute, Inc.. ATA Command Set - 5 (ACS-5) . 2021, 268p.
[16] “[CB16] EXOTIC DATA RECOVERY & PARADAIS by Dai Shimogaito”. https://www.youtube.com/watch?v=lIJkFJTFKLY, （参照 2024-08-23）