マルウェア検体のデータ欠損がマルウェア同定に与える影響の調査

小久保 博崇[1,a)]　大山 恵弘[2]

概要：マルウェアを使った攻撃を受けた時，どのようなマルウェアが攻撃に使われたかを知ることは被害状況を知るためにも重要である．しかし，マルウェアの中には証拠隠滅や調査妨害等の目的で，自分自身を消去する物も存在する．消去されてしまったマルウェアはデジタルフォレンジック技術によってある程度復元することが可能な場合もあるが，マルウェアに限らず一度消去されたファイルはストレージの使用とともにデータの一部や全部が欠損していく．本稿ではこのようなファイル消去によりデータの一部が欠損したマルウェアに着目し，欠損した状態でどこまでマルウェアを同定できるのか，マルウェアのデータのうちどの部分が欠損すると同定に影響が出るのかについて調査を行った．NTFS におけるファイルの欠損の仕方をもとに人工的にマルウェアを欠損させ調査したところ，特定のアドレスからの欠損がマルウェアの同定や悪性判定に特に悪影響を及ぼすことがわかった．

キーワード：マルウェア，　データ欠損

Effect of Malware Data Loss on Malware Identification

Hirotaka Kokubo[1,a)]　Yoshihiro Oyama[2]

Abstract: It is important to know what kind of malware is used for an attack when it is attacked with malware in order to know the damage situation. However, some types of malware erase themselves for the purpose of destruction of evidence or obstruction of investigation. Erased malware can sometimes be recovered by digital forensic technology. Nevertheless some of the data in the deleted file are lost when using storage. In this paper, we focus on the malware which has lost parts of the data by such file deletion. We investigate the accuracy of identification of malware with missing data and which part of the data of the malware affects the identification. We artificially generate corrupted malware by referring to the file data loss in NTFS and analyze that corrupted malware. It was found from the result that the data loss in the specific address space affected the identification of the malware negatively.

Keywords: Malware, data loss

1. はじめに

サイバー攻撃対策において，攻撃に使われたマルウェアがどんな種類のマルウェアであったのかを同定する作業は，攻撃による被害状況を知る上で重要である．攻撃に使われたマルウェアを被害端末から入手し，マルウェアの種類を突き止めることで，情報流出や組織内感染の可能性の有無，感染経路の推定，改竄を受けうるファイルの特定，マルウェア除去の方法など，攻撃の事後対応に役立つ情報を手に入れることができる．

しかし，攻撃に使われたマルウェアが被害端末上に残っていないケースが存在する．役目を終えて不要になったマルウェアを，攻撃者や次段階のマルウェア，もしくはマルウェア自身が削除する場合である．マルウェアに限らず，

[1] 株式会社富士通研究所
FUJITSU LABORATORIES LTD.
[2] 筑波大学
University of Tsukuba
[a)] kokubo.hirotaka@fujitsu.com

削除されてしまったファイルはOS上から行う通常のファイル操作によって取り出すことはできないが，デジタルフォレンジック技術を用いることで復元できる場合がある [1]．Windows環境においてよく使用されるFAT（File Allocation Table）やNTFS（NT File System）などのファイルシステムでは，ファイルを削除するとファイルシステム上に存在するそのファイルのための管理データ領域の削除フラグが立つ．削除フラグが立つと，OS上からそのファイルを参照することができなくなる．しかし，この状態でもファイルのデータ自体はストレージ上に残存している．この残存したデータを集めて結合することで元のファイルを復元することができる．しかし，削除フラグが立ったファイルのデータ領域は空き領域として扱われるため，新しく作られた他のファイルのデータが上書きされることがある．通常，端末の使用にはファイル作成が伴うため，端末使用時間の経過とともに削除済みデータが新規ファイルデータにより上書きされてしまう可能性は高まる．そのため，削除されて間もないマルウェアは完全な状態で復元できる可能性が高いといえるが，削除されてから時間が経過したマルウェアのデータは一部または全部が欠損している可能性がある．マルウェアによる攻撃の被害状況を知るうえでは，データの一部が欠損してしまったマルウェア（以下，欠損マルウェアと呼ぶ）からでも，マルウェアの同定が行えることが望ましい．

そこで本研究では，このような欠損マルウェアに着目する．欠損の無いマルウェア検体を様々なアドレスから人工的にデータ欠損させることで欠損マルウェアを作り出し，欠損マルウェアのうち，どのような欠損の仕方をした検体が元のマルウェアと同定できなくなるのかを調査した．対象とするマルウェアはx86-64向けのWindows用PE形式のバイナリプログラムとする．

まず本論文では，現実の環境ではどのような形でデータ欠損が起こるかについて述べる．Windows 10で標準的に使われているNTFS上でのデータ欠損の起こり方について実験を行い確認し，現実に生成されうる欠損マルウェアの形式について言及する．また，上記から本論文で扱う欠損マルウェアについて定義する．

次に，実マルウェア検体を欠損させることで欠損マルウェアを作成し，欠損のさせ方によってアンチウイルスによる検知率やマルウェアの同定の成功率がどう変わるかを調査するとともに，なぜそのような結果になったかについて考察を行う．

本論文の貢献は以下の通りである．

- 検体のデータのうち，どのアドレス帯が欠損するとマルウェアの同定に特に悪影響を与えるかを明らかにした．
- 悪影響の大きかった欠損位置において，実際にどのような意味のデータが欠損していたかを明らかにした．
- 欠損が起きたマルウェアのうち，どの程度の割合のマルウェアが同定できなくなるかを明らかにした．

2. ファイルのデータ欠損

まず，消去されたファイルがどのように欠損するのかを確認するために，下記の実験を行った．実験環境のOSはWindows 10，ファイルシステムはNTFS，セクタサイズは512 bytes，クラスタサイズは4096 bytes（8セクタ）である．仮想マシンではなく物理マシン上で実験を行った．

(1) ハードディスクストレージ上にターゲットファイルを作成し，まだ欠損していない現時点でのファイルのデータ，ファイルのSHA256ハッシュ値，ストレージ上でのデータ位置・サイズを記録する．
(2) OS上の操作でターゲットファイルを消去する．
(3) ストレージ上に新規ファイルを作成する．ただし，新規ファイルのサイズはセクタサイズの非整数倍かつターゲットファイルのサイズ未満で，データの中身は非ゼロのバイト列とする．
(4) (1) で記録したストレージ上でのデータ位置とサイズを基にターゲットファイルを復元し，SHA256ハッシュ値を計算して元のSHA256ハッシュ値と比較する．SHA256ハッシュ値が同一であった場合は欠損が起きていないということであるため，(3) に戻る．
(5) 復元したターゲットファイルのデータと欠損前のデータを比較し，ターゲットファイルのどの部分がどれだけ欠損したかを確認する．

実験の結果，我々の環境では，データ欠損はクラスタサイズである4096 bytes単位で起こった．ターゲットファイルのデータ領域において，クラスタサイズの整数倍のアドレスから新規ファイルのデータによる上書きが始まり，新規ファイルの内容を書き込み尽くしたあとはクラスタサイズの整数倍に到達するまで0で埋められた．

この実験結果から，本研究ではマルウェア検体に対してクラスタサイズの整数倍アドレス，すなわち $4096 \times N$ 番地から，クラスタサイズである4096 bytes分のデータを値0で上書きすることにより欠損マルウェアを生成するものとする．

他にも欠損のさせ方としては，セクションやPEヘッダといった意味を持って区切られているデータ領域ごとに欠損させることも考えられるが，消去されたファイルを復元したときに起こる自然欠損においてはデータの意味を考慮した欠損は起きないため，本研究では取り扱わない．

3. 実験

まず，我々が独自に収集したx86-64向けのPE形式のマルウェア924検体に対して欠損処理を行い，自然な形で欠損したマルウェアを人工的に生成する．検体データのアドレス0x0000から4096 bytes分の領域を0で上書きしたグ

表 1 欠損マルウェアグループ

	欠損開始アドレス	欠損サイズ	検体数
Group 1	0x0000	4096	924
Group 2	0x1000	4096	924
Group 3	0x2000	4096	924
Group 4	0x3000	4096	924
Group 5	0x4000	4096	924
Group 6	0x5000	4096	924
Group 7	0x6000	4096	924
Group 8	0x7000	4096	924
Group 9	0x8000	4096	924
Original	—	—	924

表 2 検知率

	平均値	最小値	最大値	中央値
Group 1	0.010	0	0.439	0
Group 2	0.356	0	0.761	0.384
Group 3	0.385	0	0.786	0.464
Group 4	0.372	0	0.783	0.386
Group 5	0.397	0	0.789	0.479
Group 6	0.403	0	0.789	0.500
Group 7	0.400	0	0.786	0.496
Group 8	0.396	0	0.775	0.486
Group 9	0.403	0	0.786	0.500
Original	0.478	0	0.851	0.567

ループ，アドレス 0x1000 から 4096 bytes 分の領域を 0 で上書きしたグループのように，欠損開始アドレスを 0x0000 から 0x8000 まで 0x1000（4096 の 16 進数表現）ずつ増やしていき，そのアドレスから 4096 bytes 分を 0 で欠損させることで，9 つの欠損マルウェアグループ（各グループあたり 924 検体，合計 8,316 検体）を用意した（**表 1**）．欠損処理を行うためには元々の検体のデータサイズが開始アドレスと 4096 bytes の合計以上でなければならないが，ファイルサイズが小さく欠損処理を行えない検体はあらかじめ除外しているため，元々の 924 検体には含まれていない．

それぞれのグループに対して VirusTotal[*1]を使用し，各社のアンチウイルス製品（以下，AV と呼ぶ）でのスキャンを行い，結果を分析した．

4. 結果

4.1 欠損が検知率に与える影響

まずは，マルウェアの欠損が AV の検知率に与える影響について示す．

4.1.1 欠損マルウェアグループごとの検知率

表 2 に各欠損マルウェアグループと元データ（Original）の検知率を記載した．本項での検知率とは，検体を悪性と判定した AV の種類数をその検体の判定に使用した全 AV の種類数で割った値である．この検知率を各グループごとに集計し平均値，最小値，最大値，中央値を算出した．

*1 https://www.virustotal.com/

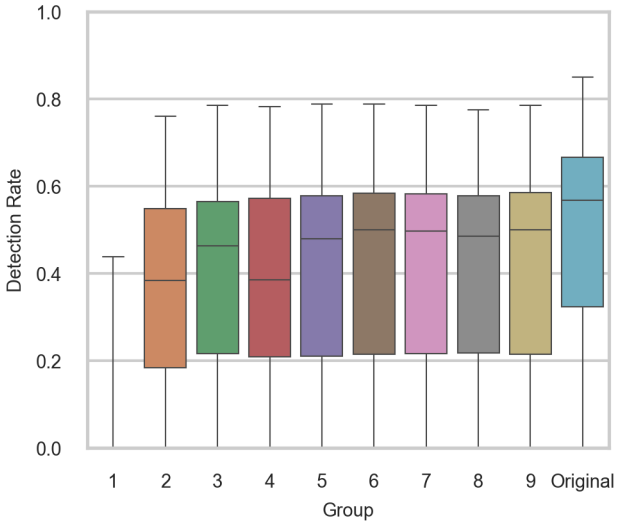

図 1 欠損マルウェア検知率の分布

また，**図 1** に各欠損マルウェアグループと元データの検知率（Detection Rate）の分布を箱ひげ図で表現した．ひげの下端が最小値，箱の下辺が 25 パーセンタイル，箱の内部の線が 50 パーセンタイル（中央値），箱の上辺が 75 パーセンタイル，ひげの上端が最大値を表している．これら区切れと区切れの間にはいずれも同量のデータが属しているため，区切れ間の長さが短いほどその区間にデータが集中して出現している．また，箱の中には全体の半数のデータが属している．

全ての Group において，検知率の最小値は 0 であったため，ひげの下端は X 軸と重なっている．Group 1 に関しては，ほとんどの検体の検知率が 0 であるため，箱部分が見えていない．Group 2 と Group 4 は似た分布となっており，中央値の位置が元データに比べて箱の中央付近まで低下している．Group 1，2，4 を除く Group も似た分布になっており，元データの検知率分布がそのまま下方向へシフトした形になっている．

Group 1，すなわちアドレス 0x0000 から 4096 bytes 分の欠損が，検体の検知率に特に悪い影響を与えている．中央値で比較すると，Group 2 及び Group 4 の検知率が Group 1 に次いで悪く，それ以外の Group はあまり差がない．このことから，マルウェアデータのアドレス 0x8000 までから始まるデータ欠損は，場所に依らず AV による悪性判定に悪い影響を与えており，特にマルウェア先頭の欠損が致命的であるといえる．また，アドレス 0x1000 及び 0x3000 付近から始まる欠損は，マルウェア先頭からの欠損ほどではないが他の欠損箇所よりも悪い影響を与える傾向にあるといえる．

図 2 に，各欠損マルウェアグループの検知率から元データの検知率を引いた値（Difference Value of Detection Rate）の分布を箱ひげ図で表現した．元データよりも検知率が低

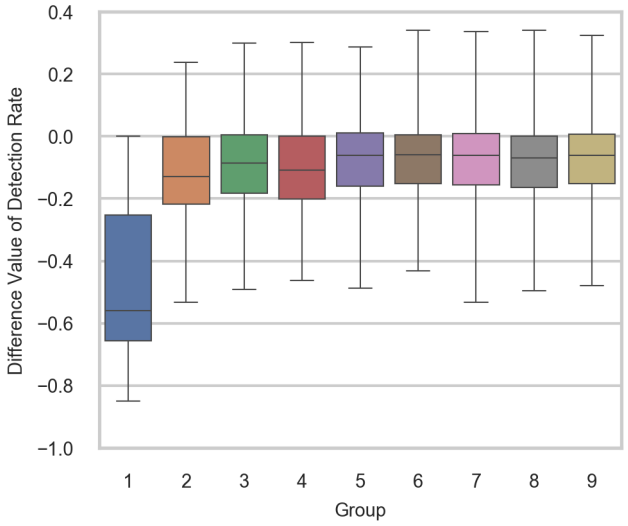

図 2 欠損マルウェア検知率の差分

下すると Y 軸の値である差分が 0.0 を下回り，検知率が上昇すると 0.0 を上回る．差分が 0 に等しい場合は検知率が変化しなかったことを表している．Group 1 は全ての検体の検知率が低下している．Group 1 の多くの検体は検知率が 0 まで低下しているため，図 1 の元データのグラフをマイナス方向へ反転させたようなグラフとなっており，グラフの形状にはあまり意味がないといえる．それ以外の Group は 75%程度の検体が検知率の低下を起こしているが，欠損しても検知率が変化しないもしくは，検知率が上昇する検体も合計して 25%ほど存在している．欠損前後で悪性判定に使用する AV の種類を統一した上で，全ての検体の検知率が低下した Group 1 を除いた各 Group において，欠損により検知率が上昇する検体の割合を調べたところ，各 Group で 17%から 18.3%の検体が該当した．

4.1.2 各 AV ごとの検知率

図 3 に，各 AV の欠損マルウェアに対する検知率（Detection Rate）を AV ごとに示した．4.1.1 項と異なり，本項での検知率とは，ある特定の AV が欠損マルウェアグループ全体をスキャンし，そのうち悪性と判定できた検体の割合を表す値である．スキャンに使われた AV のうち著名な物を 5 種類選出し，それぞれ AV01-05 と名付けた．

AV によらず，Group 1 に対する検知率はほぼ 0 である．これは今回選出しなかった他の AV でも同一であった．AV02-04 は，欠損前の検体に対しては AV01 や AV05 と比べて 0.1 から 0.34 程度良好な検知率を示すが，欠損により検知率が 0.4 以上低下する．逆に AV01 及び AV05 は，欠損前よりも欠損後のほうが検知率が向上している．

4.2 欠損がマルウェア同定に与える影響

次に，マルウェアの欠損がマルウェア同定に与える影響について示す．実験結果から，欠損後の検体から欠損前のマルウェア名を導き出せているかどうかを調査した．

図 4 に，各 AV の欠損マルウェアに対する同定成功率（Accuracy Rate）を示す．同定成功率とは，各 AV を用いて欠損マルウェアグループ中の検体に対してマルウェア名をつけたとき，欠損前と同じマルウェア名を命名できた割合を示す値である．欠損前の時点で悪性と判定できなかったために欠損前のマルウェア名をつけられなかった検体については，その AV の同定成功率の計算から除外している．つまり同定成功率は，欠損前後でマルウェア名が変わらなかった検体数を，欠損前にマルウェア名を付けられた検体数で割った値である．

AV に依らず，Group 1 に対する同定成功率は検知率と同様にほぼ 0 であった．AV01 は欠損後の検体に対しても比較的高い同定成功率となっており，Group 1 及び Group 2 以外で 0.6 以上を維持している．AV01 の Group 2 については Group 3 以降の半分の同定成功率となっており，アドレス 0x1000 からの欠損が同定に悪影響を与えていることがわかる．AV02 は全体的に同定成功率が低いが，Group5-7, 9 は他の Group の 2 倍以上である 0.3 近くまで高くなっている．AV03 及び AV04 は，欠損箇所に依らず同定成功率が 0.2 未満と低くなっている．AV05 は，Group 1 以外で 0.8 以上と高い同定成功率を示している．しかし，AV05 により付けられた名称を確認すると，欠損前後ともに特定のマルウェアを表す名称ではなく汎用的な名称が多く付けられており，それが影響して同定成功率が押し上げられていた．

これらから，マルウェアのデータ欠損は場所に依らずマルウェアの同定に悪影響を与え，特にデータ先頭からの欠損は検知率の場合と同様に著しい悪影響を与えるといえる．アドレス 0x1000 以降の欠損は，同定に使用する AV によって影響度合いが異なるが，概ね若いアドレスから始まる欠損がより悪い影響を与える．

5. 考察

4 章の結果から，若いアドレスからの欠損，特にデータ先頭からの欠損が検知率及びマルウェア同定に悪影響を与えることがわかった．そこで，各欠損マルウェアグループで具体的にどのような意味のデータが欠損しているかを，検体の PE ファイルとしての構造 [2] に着目して解析し調査した．

Group 1 の欠損箇所であるアドレス 0x0000 から 4096 bytes の範囲には，IMAGE_DOS_HEADER 構造体等から成る MS-DOS スタブ，IMAGE_NT_HEADERS 構造体に含まれるシグネチャや，IMAGE_FILE_HEADER 構造体から成る COFF ファイルヘッダ，データディレクトリを含む IMAGE_OPTIONAL_HEADER 構造体から成るオプショナルヘッダといった，PE ファイルにおいて非常に重要な情報や，IMAGE_SECTION_HEADER 構造体から構

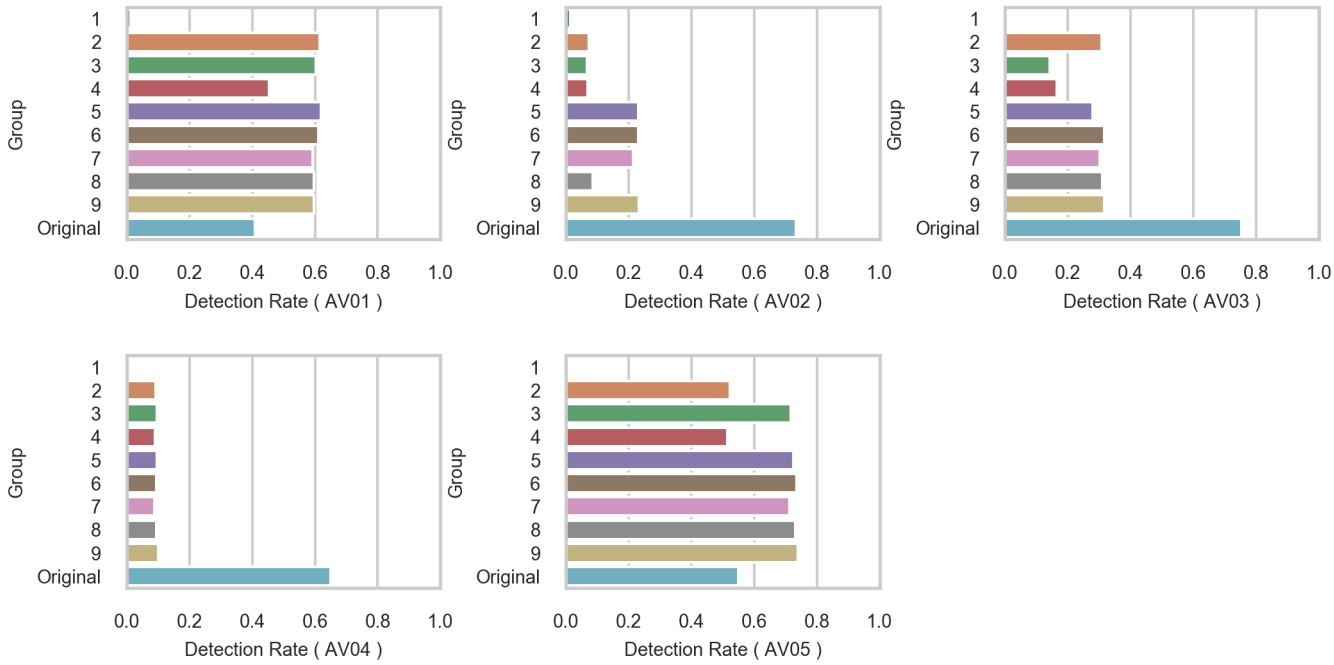

図 3　AV 別の欠損マルウェア検知率

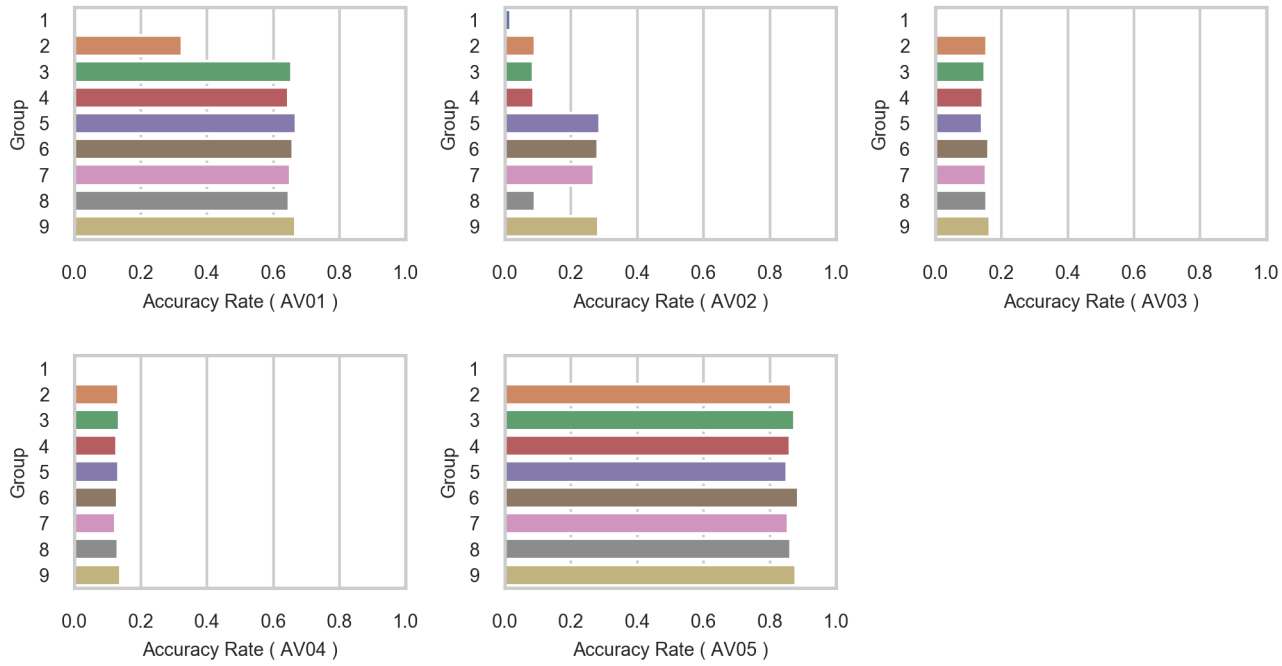

図 4　AV 別のマルウェア同定成功率

成されているセクションテーブルが格納されている．今回収集した検体群においては，これら情報は Group 2 以降の欠損箇所であるアドレス 0x1000 より後には含まれていなかった．また，IMAGE_DIRECTORY_ENTRY_IAT 用の IMAGE_DATA_DIRECTORY 構造体が指しているインポートアドレステーブルや IMAGE_DIRECTORY_ENTRY_DEBUG 用の同構造体が指しているデバッグディレクトリが格納されている領域が，アドレス 0x1000 以降に比べて多く含まれていた．IMAGE_DIRECTORY_ENTRY_IMPORT 用の IMAGE_DATA_DIRECTORY 構造体が指しているインポートテーブルが格納されている領域も含まれていたが，他の Group と比較して多く含まれているわけではなかった．このように，アドレス 0x0000 から 4096 bytes の範囲には，検体を PE ファイルと断定するためのデータや検体中の重要なデータへのポインタといったヘッダ情報の多く

が格納されているため，この領域の欠損は AV によるシグネチャマッチングやバイナリ解析に悪影響を与え，検知率や同定成功率を著しく下げていると考えられる．

セクションテーブルが指しているセクションデータ本体は，どのセクションも各欠損箇所にほぼ同量出現しており，特色は見られなかった．

Group 1 に次いで悪影響の大きかった Group 2 及び 4，すなわちアドレス 0x1000 と 0x3000 において欠損したデータは，Group 3 及び 5 以降で欠損したデータと比較しても，今回の調査においてはあまり差異が見られなかった．

6. 関連研究

Shafiq ら [3] は，機械学習アルゴリズムを使用して PE ファイルフォーマットの構造的特徴からファイルの良性・悪性判定をリアルタイムで判定する仕組みである PE-Miner を提案しており，テスト環境で 99%を上回る検知率と 0.5%を下回る誤検知率を出している．ファイル先頭付近の構造から得られる情報を有用な特徴量として採用しているため，Group 1 のようなファイル先頭からの欠損はこういった機械学習手法による悪性検知に対しても致命的な悪影響を及ぼすものと考えられる．

Hand ら [1] は，マルウェアが自身を削除した時など x86 用の実行可能ファイルが削除されたとき，それを復元する手法 Bin-Carver を提案している．EXT2 ファイルシステムと ELF 形式の実行可能ファイルにおいて，削除されたファイル群のうち 93.1%のファイルを復元することに成功している．ただし，ELF フッタやファイル先頭に存在する ELF ヘッダを復元のための重要な情報として扱っているため，ファイル先頭が欠損した検体の復元にどれほど適用できるかは未知数である．

大坪ら [4] は，コードの断片からコンパイラを推定する手法について提案している．16 命令分の小さなコード断片からコンパイラの種類を推定しており，コード断片の位置に依らず高い精度を出している．マルウェアの種類を直接的に推定するわけではないものの，欠損マルウェアのコンパイラ情報がわかれば，あるマルウェアと同一のコンパイラが使われているといった情報が得られるため，欠損位置に関わらずに使用できるマルウェア同定のための指標として使用できる可能性がある．

7. まとめと今後の課題

本論文では，ファイル消去されたことによりデータ欠損を起こしたマルウェアに着目し，欠損がマルウェア検知率やマルウェア名の同定に与える影響について明らかにした．ファイルデータの先頭から 4096 bytes の欠損が最も検知率への悪影響が大きくほぼ 0 まで検知率が低下し，ファイルデータのアドレス 0x1000 及び 0x3000 から 4096 bytes の欠損が次いで悪影響が大きかった．欠損位置によらずマルウェア名の同定成功率が 7 割前後減少するが，アンチウイルス製品によってはアドレス 0x2000 以降からの 4096 bytes の欠損であれば 6 割強の同定成功率を維持できた．

今回は x86-64 用のマルウェアのみを対象としたため x86 用のマルウェアも対象としデータの種類や数を増やして再度実験することや，欠損位置を増やすこと，アドレス 0x1000 及び 0x3000 からの欠損による悪影響が大きい理由を調査することが今後の課題である．

参考文献

[1] Hand, S., Lin, Z., Gu, G. and Thuraisingham, B.: Bin-Carver: Automatic Recovery of Binary Executable Files, *Proceedings of the 12th Annual Digital Forensics Research Conference (DFRWS'12)*, Washington DC (2012).

[2] Microsoft: Windows Dev Center - PE Format, `https://docs.microsoft.com/en-us/windows/win32/debug/pe-format`.

[3] Shafiq, M. Z., Tabish, S. M., Mirza, F. and Farooq, M.: PE-Miner: Mining Structural Information to Detect Malicious Executables in Realtime, *Recent Advances in Intrusion Detection* (Kirda, E., Jha, S. and Balzarotti, D., eds.), Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 121–141 (2009).

[4] 大坪雄平, 大塚玲, 三村守, 榊剛史, 受川弘, 岩田吉弘: コード断片からのコンパイラ推定手法, コンピュータセキュリティシンポジウム 2018 論文集, pp. 1259 – 1265 (2018).