Responsible disclosureの具体的な事例紹介

横浜国立大学
大学院環境情報研究院/先端科学高等研究院

吉岡克成

SCIS2017@沖縄

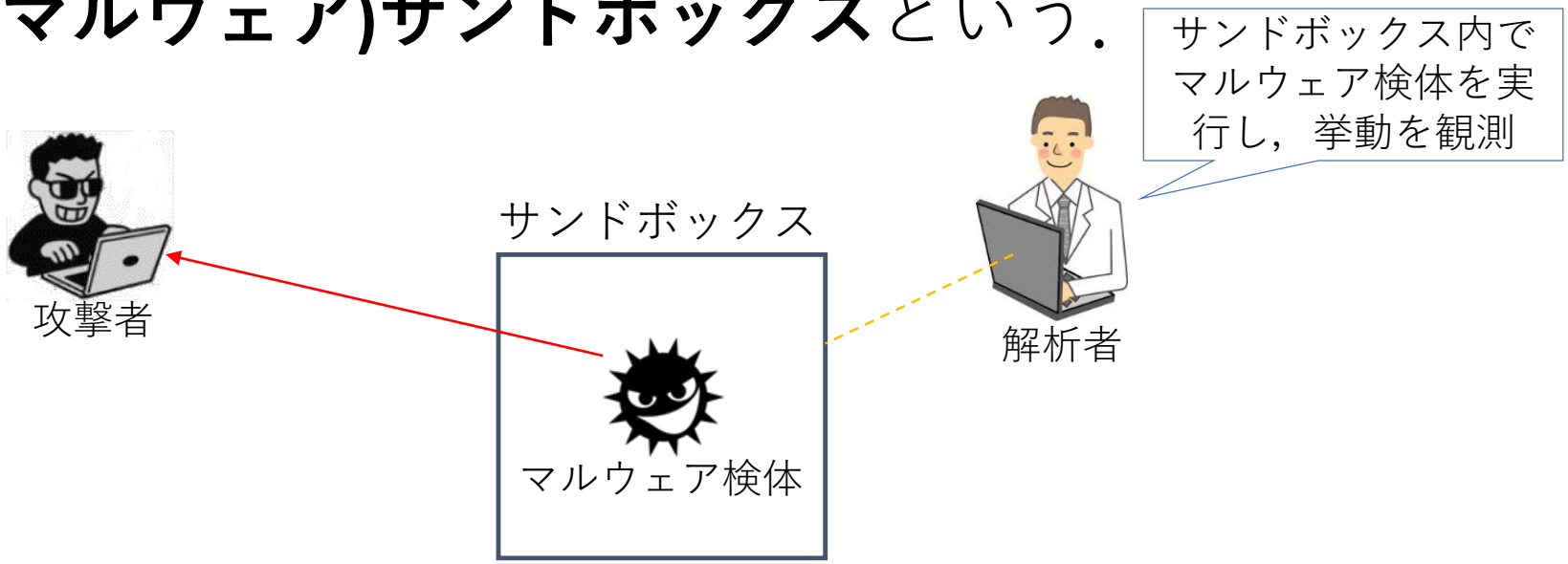
SandPrint: Fingerprinting Malware Sandboxes to Provide Intelligence for Sandbox Evasion

A. Yokoyama, K. Ishii, R. Tanabe, Y. M. P. Pa,
T. Kasama, K. Yoshioka, T. Matsumoto, D. Inoue,
C. Rossow, and M. Backes,

The 19th International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2016), 2016.

(マルウェア)サンドボックスとは？

- **マルウェア動的解析**とは，解析対象のマルウェア検体を解析環境内で実行し，その挙動を観測する解析手法.
- マルウェア動的解析に用いられる解析環境を**(マルウェア)サンドボックス**という.



こんなところで使われています(1)
セキュリティアプライアンス

- ネットワークトラヒックやメール添付ファイルを解析してマルウェアを検知する製品
- **サンドボックス**が内蔵されており、この中で検査対象を実行して悪性なファイルを検知

製造社名	アプライアンス名 / サービス名	種類	製造社名	アプライアンス名 / サービス名	種類
Bluecoat	Malware Analysis System[2]	オンプレミス型	Lastline	Lastline on-Premise[12]	オンプレミス型
Check Point	Threat Emulation[3]	オンプレミス型 / クラウド型	McAfee	Advanced Threat Defence[13]	オンプレミス型
Cisco	Advanced Malware Protection[4]	クラウド型	Paloalto	WildFire[14]	クラウド型
Dell	SonicWALL Capture[5]	クラウド型	Proofpoint	Targeted Attack Protection[15]	クラウド型
FFRI	FFR Yarai Analyzer[6]	オンプレミス型	Secure Brain	Zero-Hour Response[16]	オンプレミス型
FireEye	Malware Analysis[7]	オンプレミス型	Sophos	Sandstorm[17]	クラウド型
Fortinet	FortiCloud[8]	クラウド型	Symantec	Advanced Threat Protection[18]	クラウド型
Fortinet	FortiSandbox[9]	オンプレミス型	TrendMicro	Cloud App Security[19]	クラウド型
Hitachi	MAAS[10]	オンプレミス型 / クラウド型	TrendMicro	Deep Discovery Analyzer[19]	オンプレミス型
IIJ	SecureMX[11]	クラウド型	WatchGuard	APT Blocker[20]	クラウド型
Lastline	Lastline Cloud[12]	クラウド型	Websense	Sandbox Modules[21]	オンプレミス型

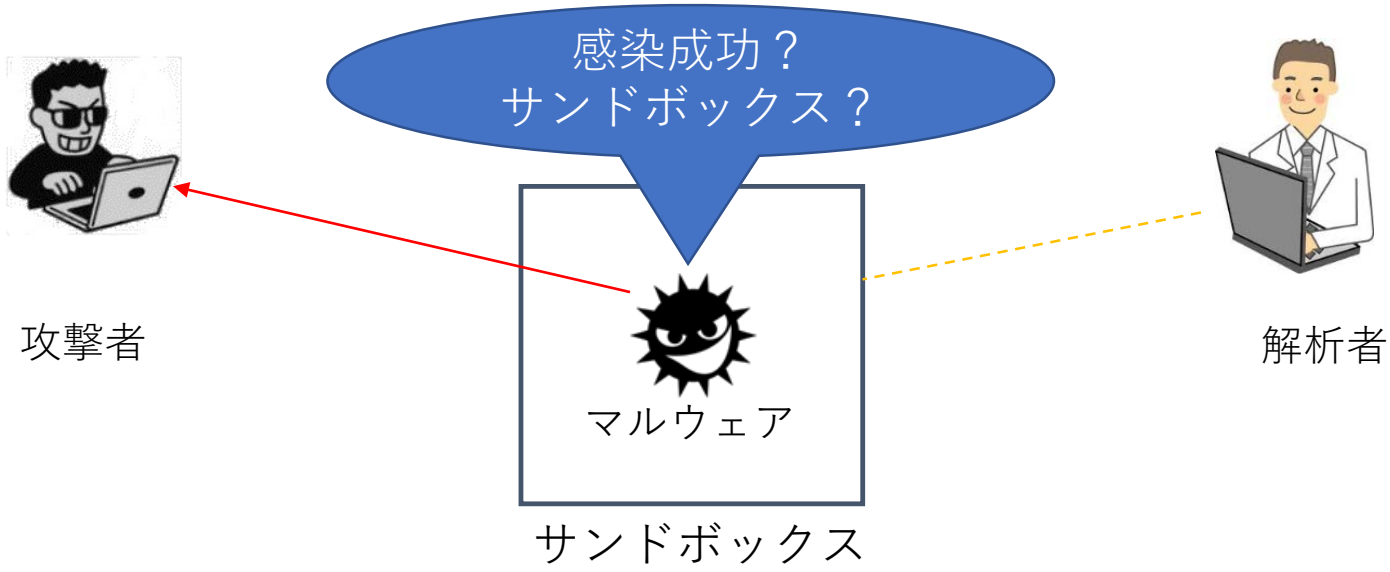
田辺瑠偉, 石井攻, 横山日明, 吉岡克成, 松本勉, “標的組織の内部情報を有する攻撃者を前提としたセキュリティアプライアンス評価,” 情報処理学会コンピュータセキュリティシンポジウム2016, セッション3F4, 2016 より

こんなところで使われています(2)
オンラインファイル解析サービス

- ファイルをWebから投稿するとマルウェア解析結果を返してくれるサービス
- バックエンドに**サンドボックス**が動作し投稿ファイルが自動的に解析される
- 有名なサービス(VirusTotal)は一日100万件近いファイル投稿がある

攻撃側と防御側の戦い

- 攻撃側はサンドボックスを検知して無害な振りをするマルウェアを使用して検知や解析を逃れようとする.
- 防御側はサンドボックスであることが見破られないように工夫する.
- これまで攻撃側、防御側の立場で多くの研究



関連研究・報告

- サンドボックスを検知するマルウェアの報告.
  - 仮想マシンの検知[2][4][5][7]
    - バックドアポート，プロセス，レジストリ
  - デバッガの検知[3][4]
  - タイミングベース[4][5]
  - エミュレータの検知[5]
  - ユーザ操作の有無[2][6]
    - マウスイベント，ダイアログボックス

[2]FireEye： Hot knives through butter, https://www.fireeye.com/content/dam/legacy/resources/pdfs/fireeye-hot-knives-through-butter.pdf,
[3] X. Chen , J. Andersen , Z. M. Mao, M. Bailey and J. Nazario : Towards an Under- standing of Anti-virtualization and Antidebugging Behavior in Modern Malware, The 38th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2008), pp.177–186 (2008)
[4] T. Holz and F. Raynal : Detecting Honeypots and other Suspicious Environ- ments, Proc. 2005 IEEE Workshop on Information Assurance and Secrity, pp.29–36 (2005).
[5] T. Raffetseder, C. Kruegel and E. Kirda : Detecting System Emulators, Proc. 10th Information Security Conference (SC), LNCS Vol.4779, pp.1–18 (2007)
[6] Understanding and fighting evasive malware, http://www.rsaconference.com/writable/presentations/file_upload/hta-w10-understanding-and-fighting-evasive-malware_copy1.pdf
[7] G.N.Barbosa, R.R.Branco. (2014), Prevalent characteristics in modern malware, http://www.blackhat. com/docs/us-14/materials/us-14-Branco-Plevalent-Characteristics-In-Modern-Malware.pdf

関連研究

サンドボックス検知を行うマルウェアへの対策.

- 解析検知をするマルウェアを逆に特定する研究[8][9]
- 実マシン上での解析[10]
- 実マシンの情報に置き換える[11]

[8] D. Kirat, G. Vigna, C. Kruegel, “Barecloud: bare-metal analysis-based evasive malware detection”, 23rd USENIX conference on Security Symposium (SEC'14). USENIX Association, Berkeley, CA, USA, 287-301.
[9] M. Lindorfer, C. Kolbitsch, P. M. Comparetti, “Detecting Environment-Sensitive Malware”, 14th international conference on Recent Advances in Intrusion Detection(RAID’11),338-357,2011.
[10] D. Kirat, G. Vigna, C. Krugel, “Barebox: Efficient malware analysis on bare-metal”, Annual Computer Security Applications Conference (ACSAC), 2011, 403-412
[11] A. Vasudevan and R. Yerraballi, “Cobra: Fine-grained Malware Analysis using Stealth Localized-executions”, IEEE Symposium on Security and Privacy, 2006

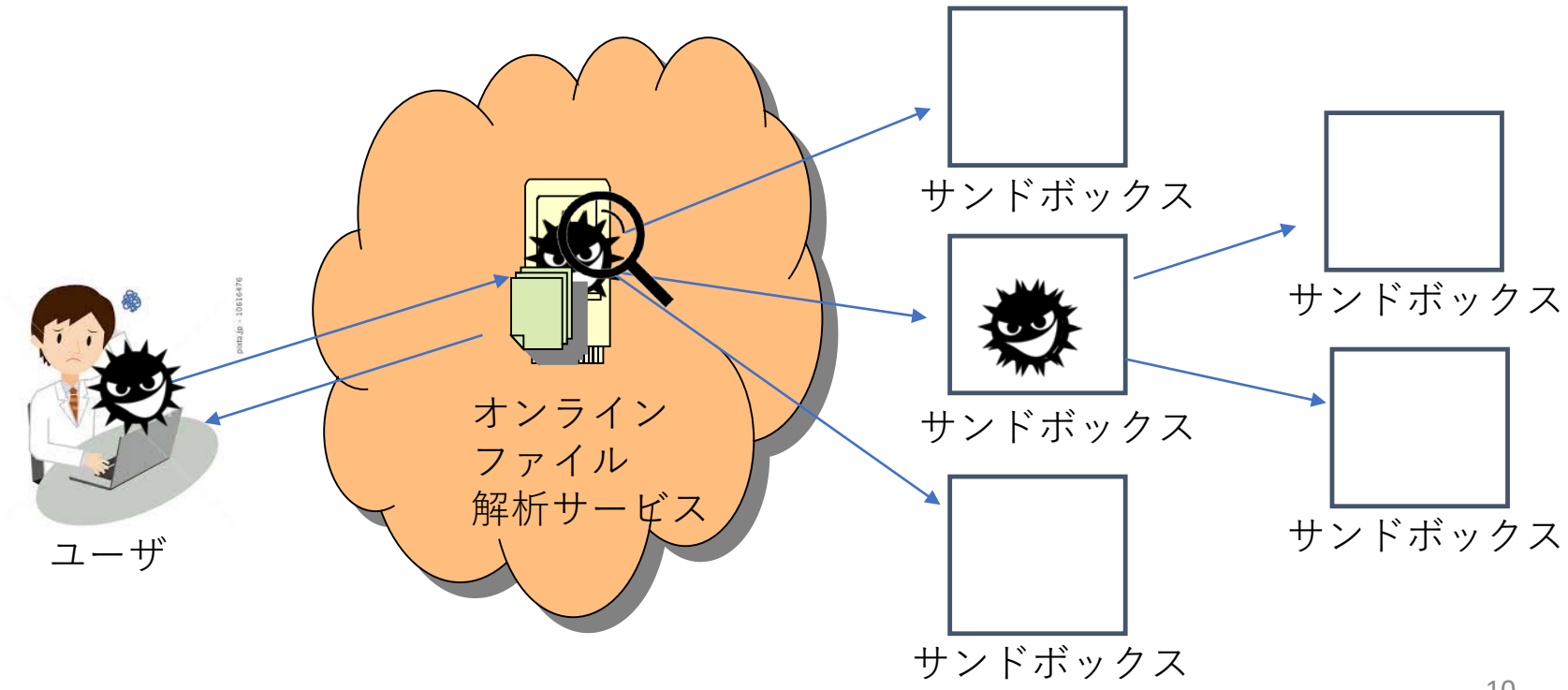
我々が着目したサンドボックスの特徴

サンドボックスは様々な理由で一般ユーザの環境とは異なる特徴を持ち，これらによって解析が検知される恐れがある

- ハードウェア
  - サンドボックスによっては，割り当てられるメモリ等のリソースが限られている.
- スナップショット
  - サンドボックスはスナップショットで状態をマルウェア感染前に復元するため，ファイルアクセス履歴などのユーザの操作履歴が乏しい.
- 環境構成
  - サンドボックスにはデスクトップやOSの設定が初期状態でいわゆる一般ユーザらしさが無い可能性がある.
- ユーザ操作
  - 実行中のマシンの操作が少ないあるいは全く無い.

サンドボックス情報収集ツールSandPrint

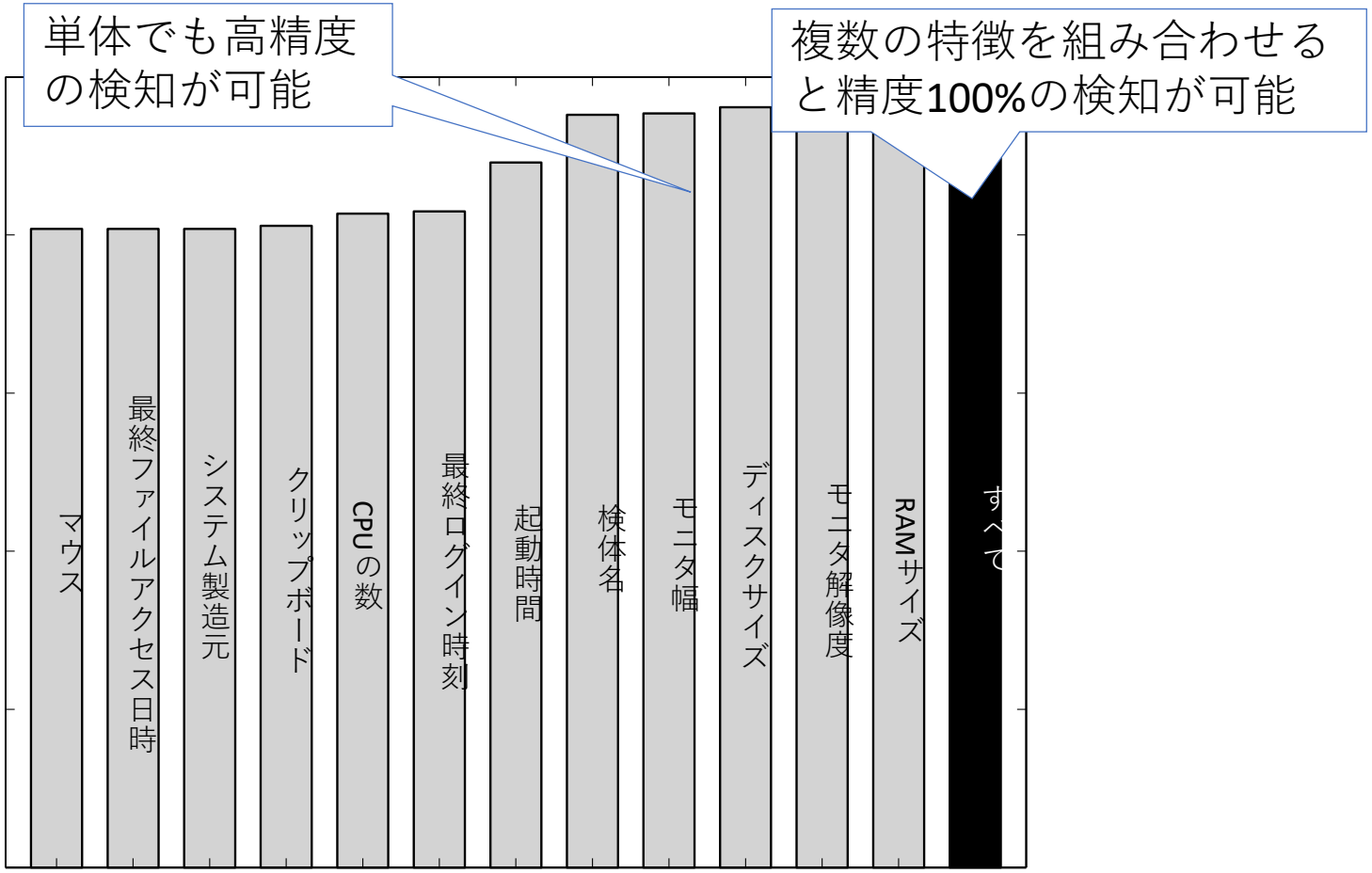
- サンドボックス固有の特徴を調査するために
  サンドボックス情報収集ツール SandPrint を作成.
- オンラインファイル解析サービスに投稿.



結果

- SandPrintにより，合計で2666件の実行環境内の情報に関するレポートを収集した.
- 20サービス中11のマルウェア検査・解析サービスに投稿したSandPrintからの通信を確認.
- 合計 33 カ国の 395 IP アドレスから通信を観測.
- サービスに投稿された SandPrint が共有され,様々な研究機関などで動的解析されていることが確認できた.

高精度でサンドボックス検知が可能



サンドボックス固有な特徴により高精度なサンドボックス検知が可能.

セキュリティアプライアンスの検知

実際に販売されている3つのセキュリティアプライアンスで同様に高精度のサンドボックス検知が可能だった

→サンドボックス構築時にはこれらの特徴に着目した検知に注意する必要があり、セキュリティベンダはこのようなサンドボックス検知への対策を講じる必要があるといえる．

RAID2016投稿の裏話

- 論文投稿時は製品名・サービス名は全て実名で投稿(ただし、出版時には製品名を匿名化する予定であることを記載)

- プログラム委員会から「この対応では不十分」との指摘．適切な「Responsible Disclosure」対応をしなければ採録とできない旨のコメントを受ける

- シェパード(論文添削監視者：お世話役)は、イリノイ大学のMichael Bailey准教授であり、**メンロレポートの著者の一人**だった

Responsible Disclosureと論文採録までの流れ

2016/6/4 論文が条件付き採録となりResponsible Disclosureが採録要件となる

2016/6/10 シェパードと著者の各組織(横浜国大、NICT、ザールラント大)が一同にSkypeミーティング。今後のDisclosureの手順について提案し、承認を得る

2016/6/10-20 Disclosureの準備(研究内容と指摘する問題点の説明文、連絡先情報等の確認)

2016/6/21 Disclosureの開始(セキュリティベンダ３社、オンライン解析サービスオペレータ２０組織が対象)

Responsible Disclosureと論文採録までの流れ

2016/6/29 シェパードへの中間報告(この時点でセキュリティベンダ３社全て、解析サービスオペレータ20社のうち11社から返信あり)

2016/7/4　返信の無い解析サービスオペレータ9社に他のチャネル(組織の問い合わせ窓口メールなど)から再度連絡

2016/7/9シェパードとプログラム委員会により採録の判定を受ける

2016/7/18 返信の無い解析サービスオペレータがある国のNational CERTに連絡

2016/9/19-21 RAID2016開催(論文の公表)

スケジュール

90日間

３つのセキュリティベンダと２０のオンライン解析サービスのオペレータに研究内容と脆弱性について通知(セキュリティ情報提供用メール、または、Webフォームより)

14日間返答がない場合

別のチャネル(各企業の問い合わせ用アドレス等)で連絡

14日間返答がない場合

通知先組織がある国のNational CERTにメール連絡

RAID 2016

論文内での記述(約1ページ)

7.1 Ethical Considerations

Our research may seem offensive in the sense that we reveal fingerprints of malware sandboxes that adversaries can use to evade them. Note, however, that the information we presented can be gathered by any other person reproducing our (conceptually simple) fingerprinting method. We thus consider the information shown in this paper as public knowledge. Still, we present data only in aggregated form and refrain from revealing any internals of particular sandboxes.

Using our insights, sandbox operators can
ysis systems. For example, we have shown th
features that are inherent to the snapshot o
be possible to find artifacts that can identif
nificantly harder to build a classifier that wo
more people randomize characteristics. We h
ticularly characteristic of sandboxes, giving
to significantly improve the stealthiness of th

7.2 Responsible Disclosure

Organizations developing sandboxes and/or a
by our research results and we thus consider
responsible disclosure process. To notify these
90 days prior to the publishing date of this pa
and including hints on how to protect against potential adversaries in the future. We used direct contacts whenever possible and available. Alternatively, we resorted to contact details stated on the organization's websites, notably including Web-based contact forms. If we did not receive a response after 2 weeks, we retried to contact the organization, if possible using alternative communication channels (e.g., using generic email addresses like `info@organization.com` or email addresses found in the WHOIS database for the organization's website domain). If we did not hear back from the organization after 4 weeks, we contacted the national CERT(s) that are in the same country as the affected organization in order to notify the party via the CERT as trusted intermediary.

We handed to each organization an executive summary of our research results as well as a full description of our research methodology (i.e., a copy of this paper in the pre-print version). We made sure to highlight the implications of our work with respect to future operations of the sandbox and/or appliance. We also specified our contact details of both research institutions, including physical address, phone number, and the email address of a representative for the research activities. We allowed the organizations to download the latest version of SandPrint and its source code. Such auxiliary data is helpful to build protection mechanisms against sandbox-evasive programs similar to SandPrint. We also remove all organizations' names when referring to individual sandboxes/services.

開示先組織からの反応

- **最終的に１８の組織から研究に対する**
  ポジティブな意見と反応が得られた。

- ネガティブな反応はなかった。

- SandPrintのソースコードを７つの組織に提供した

所感

- Responsible Disclosure等の適切に対応すれば、脆弱性指摘に対するベンダの反応はポジティブなものが多い(友好的な反応を示した組織のうちの１社は以前に匿名化なしに同社製品の脆弱性を詳細暴露した研究者を訴えた実績があった)

- Responsible Disclosureには手間と時間が掛かるので、脆弱性研究を発表する場合は、**時間に余裕をもって対応を計画**する必要がある(投稿前にやるべき)

- 必要以上に倫理問題を意識して保守的な研究を行うよりも、世界の動向は**必要な責任を果たして社会への恩恵を高める**研究を評価する傾向にあると考える。(メンロレポートの考え方)