


NICTER Dataset 2017

笠間 貴弘

国立研究開発法人 情報通信研究機構
サイバーセキュリティ研究所
サイバーセキュリティ研究室

Darknet → Dataset

● ダークネットトラフィックデータ

- /20(約4千アドレス)のダークネットトラフィック
- 観測期間は2011年4月1日から現在まで6年間以上
- NONSTOP上で提供 (pcap+DB)

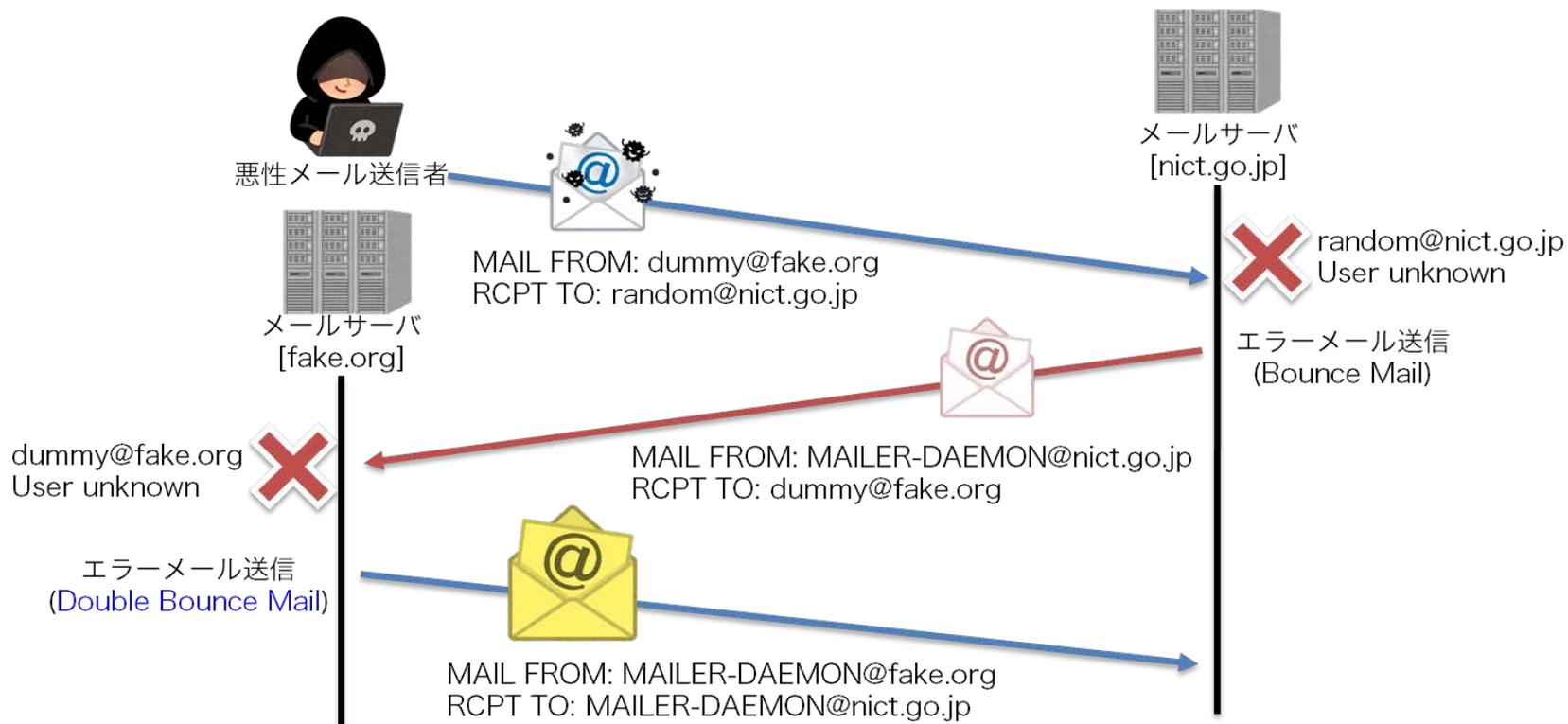
● スпамメールデータ (New!)

- NICTのメールサーバに届いたダブルバウンスメール
- 観測期間は2015年1月1日から現在まで2年間以上
- NONSTOP上で提供 (メールファイル)

ダブルバウンスメールとは？

● エラーメールの一種

- 主に送信元/宛先アドレスが存在しない場合に発生する
- ほぼ全て悪性メール（宛先ランダム+送信元詐称）



ダークネット観測とは？

- **ダークネット：未使用のIPアドレス空間**

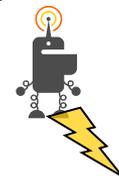
- ✓ 正常な通信は“基本的に”届かない

- **実際は大量の通信が届く**

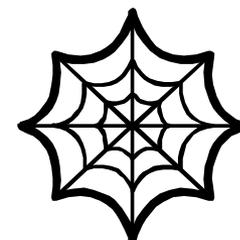
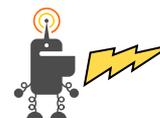
- ✓ マルウェアによるスキャン
- ✓ DDoS攻撃の跳ね返り
- ✓ リフレクション攻撃の準備活動
- ✓ etc.

- **ダークネットの観測によって
パンデミックの兆候が分かる**

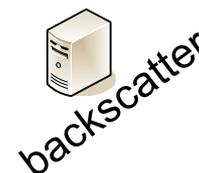
- ✓ パンデミック：マルウェアの大量感染



scan



Darknet



backscatter

よくある誤解（その1）

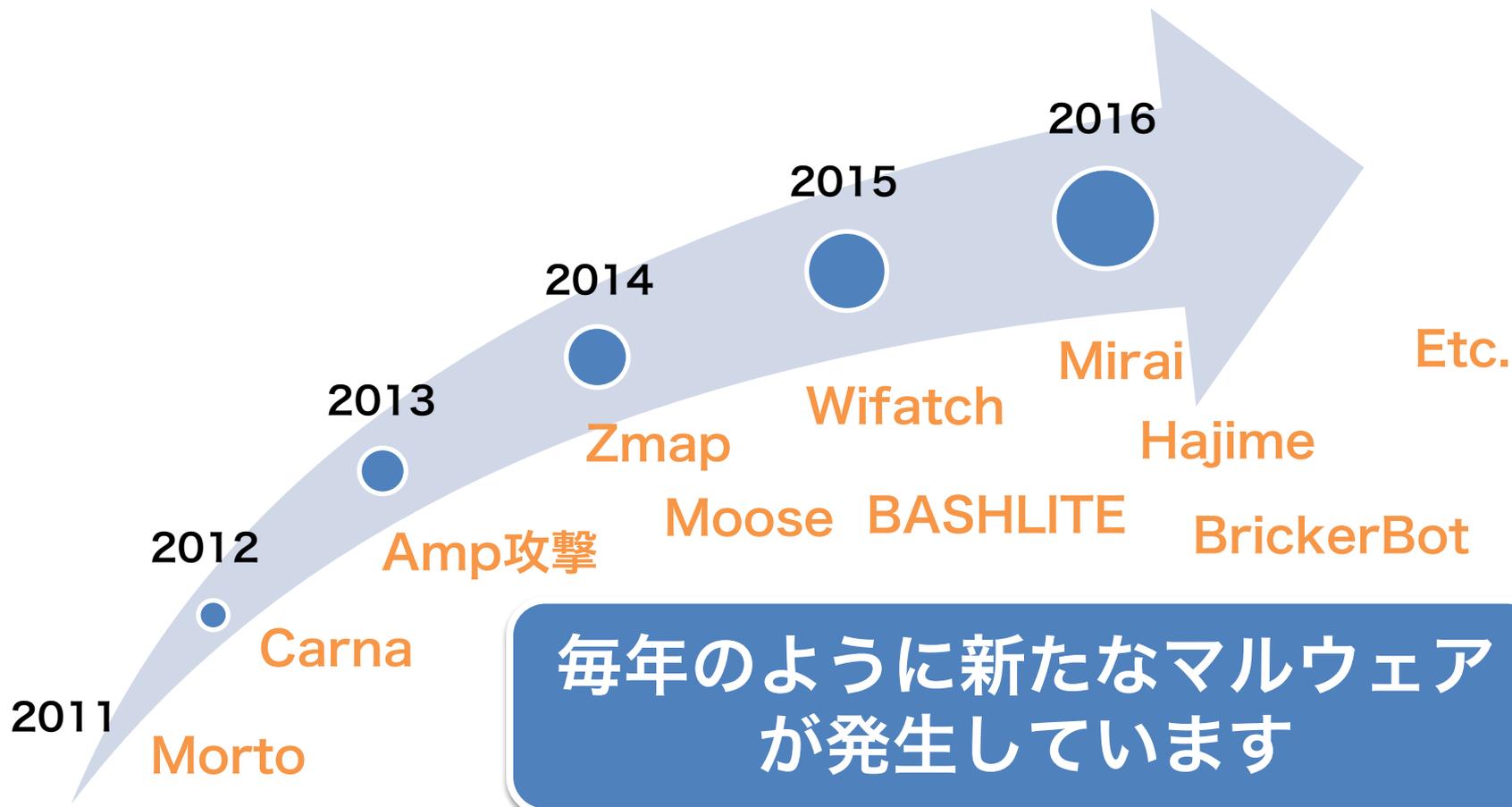
- ワームとか古いし今時スキャンとか飛んでこないでしょ

10年間観測し続けていますが
基本的にずっと増加傾向です



よくある誤解（その2）

- スキャンしてるのなんて昔のConfickerだけでしょ



よくある誤解（その3）

- ダークネット使った研究とか枯れ果ててるでしょ

USENIX Sec'14 The 23rd USENIX Security Symposium, August 2014.

USENIX WOOT'15

NDSS'14

SIGCOMM'14

NDSS'17

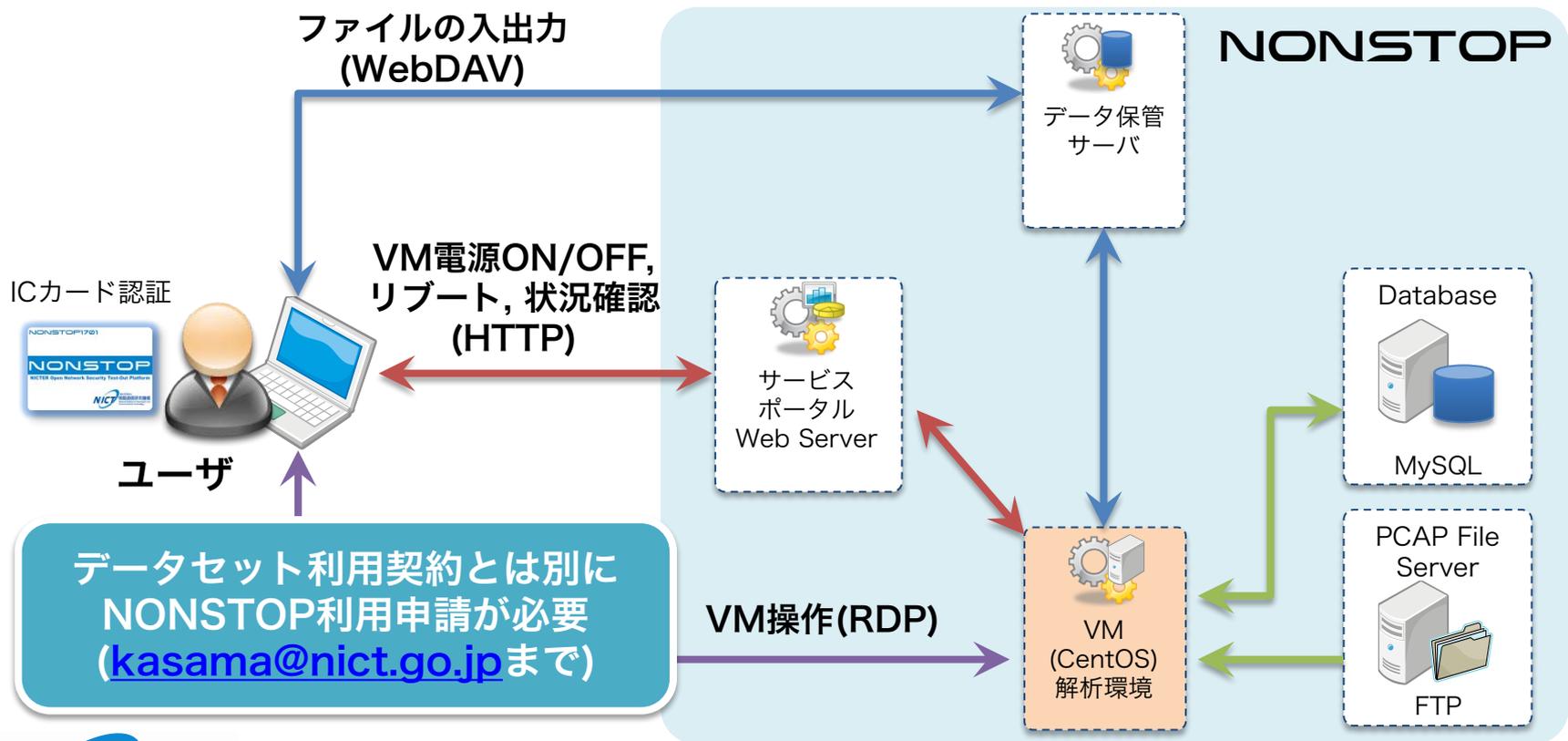
IMC'15

ダークネットデータを使った論文は
難関国際会議も含めて多数発表されています

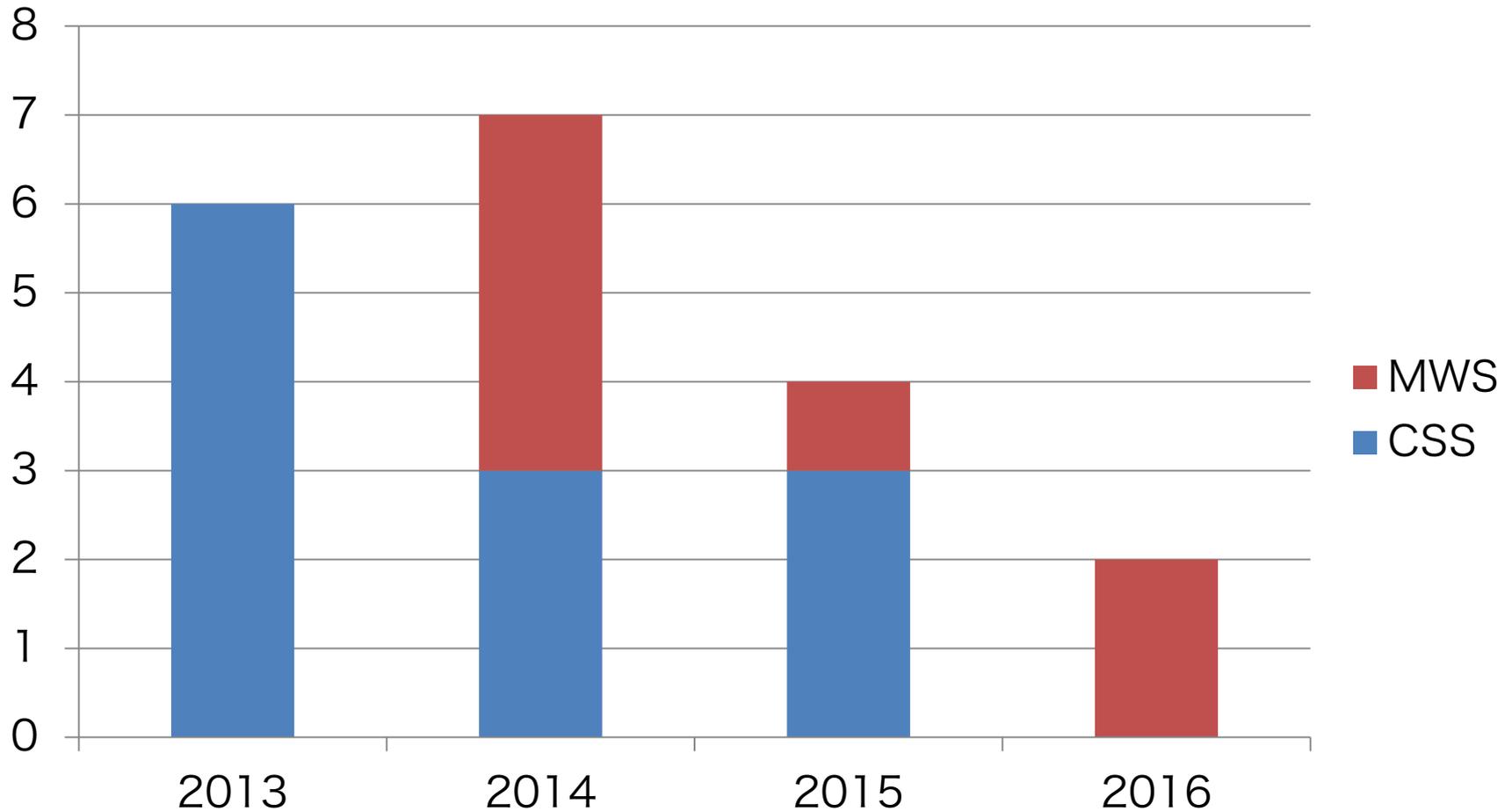
paper uniquely exploits passive monitoring and analysis of a newly deployed network telescope IP address space in a first attempt ever to build broad notions of real CPS maliciousness. Specifically, we approach this problem by inferring, investigating, characterizing, and reporting large-scale probing activities that these systems have been undergoing large-scale transformations with the infusion of new "smart" cyber-based technologies to improve their efficiency and reliability. These transitions are being driven by continual advances and cost-

NONSTOPって？

- NICTが持つサイバーセキュリティ情報を遠隔から安全に利用してもらうための環境



NICTER Darknet 利用実績



※MWSのデータセット契約枠としては0件

NICTER Dataset まとめ

- **今年度提供するデータは2種類：**
 - ダークネットトラフィック
 - スパムメールデータ
- **データセットはNONSTOP上で提供：**
 - データにアクセスできるVM環境をユーザ毎に用意
 - 利用申請はNICT笠間(kasama@nict.go.jp)まで
- **メリット：**
 - リアルタイムかつ継続的な長期間のデータセット提供
 - 加工されていない生データなので用途は自由
- **需要(実績)が無いなら(いずれ)供給は止まります**
 - *積極的な論文投稿を期待しています！*