Copyright© NTT corp. All Right Reserved.

NTT

NTT Information Sharing Platform Laboratories
NTT 情報流通プラットフォーム研究所

SBMを利用したフィッシングサイト検知とその展望
ー集合知セキュリティという考え方ー

NTT情報流通プラットフォーム研究所
中山心太

© 2009 NTT Information Sharing Platform Laboratories

ウェブのインシデントはなぜ発生する？

Copyright© NTT corp. All Right Reserved.

- ウェブの発展速度は非常に急激
  - 情報漏えいなどは上位レイヤーにシフト
  - システムではなく、人間が攻撃される
    ex)フィッシング詐欺、ネットストーカー

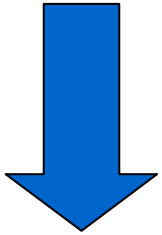
- セキュリティの進歩は非常に緩やか
  - 研究者
    - ウェブ開発者に比べて頭数がそれほど増えていない
    - 複合分野なので、できる人が少ない
  - 開発者
    - 人間的なレイヤーまでは守れない
    - 既存サービスのセキュリティを高めるくらいなら、新しいサービスを作りたい

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

集合知セキュリティとは何か？

- Web2.0的サービス
  - コンテンツがコンテンツを生み、コンテンツが価値を生む
  - コンテンツが増殖しまくり
  - コンテンツから情報が漏れまくり



Web2.0をセキュリティに応用

- 集合知セキュリティの考え方
  - コンテンツからセキュリティを生成する
  - 自動増殖するコンテンツ→自動で対応するセキュリティ
  - メンテナンスフリーでみんな幸せ

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

目次

- 背景
- フィッシングサイトの特徴
- 検索エンジンを用いたフィッシング検知手法
- SBMを用いたフィッシング検知手法
- SBMを用いた場合の未来シナリオ

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

背景

- フィッシングサイト
  - 既存ウェブサイトをコピーするだけで簡単に作れる
  - 英語圏の途上国が多い
  - 100ドルの被害額は、途上国では大金

- 全米被害額
  - 2006年　28億ドル
  - 2007年　32億ドル

- 日本はあまり多くは無い
  - 日本語のウェブサイトを作るコストが高い
  - 日本円をマネーロンダリングする環境が無い

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

フィッシングサイトの特徴

- 正規サイトのコピー
  - 同じ文言を含む
- 平均寿命は三日
  - APWG(http://www.antiphishing.org/)の統計より
- 検索エンジンからの評価が低い
  - 出来たばかり、被リンクが少ない

- 正規サイトは検索エンジンの評価が高い
  - 長期間運営、被リンクが多い

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

目次

- 背景
- フィッシングサイトの特徴
- 検索エンジンを用いたフィッシング検知手法
- SBMを用いたフィッシング検知手法
- SBMを用いた場合の未来シナリオ

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

検索エンジンを使った検知

- フィッシングサイトは正規サイトのコピー
  - 正規サイトと同じ文言(会社名、製品名、サービス名)を含む

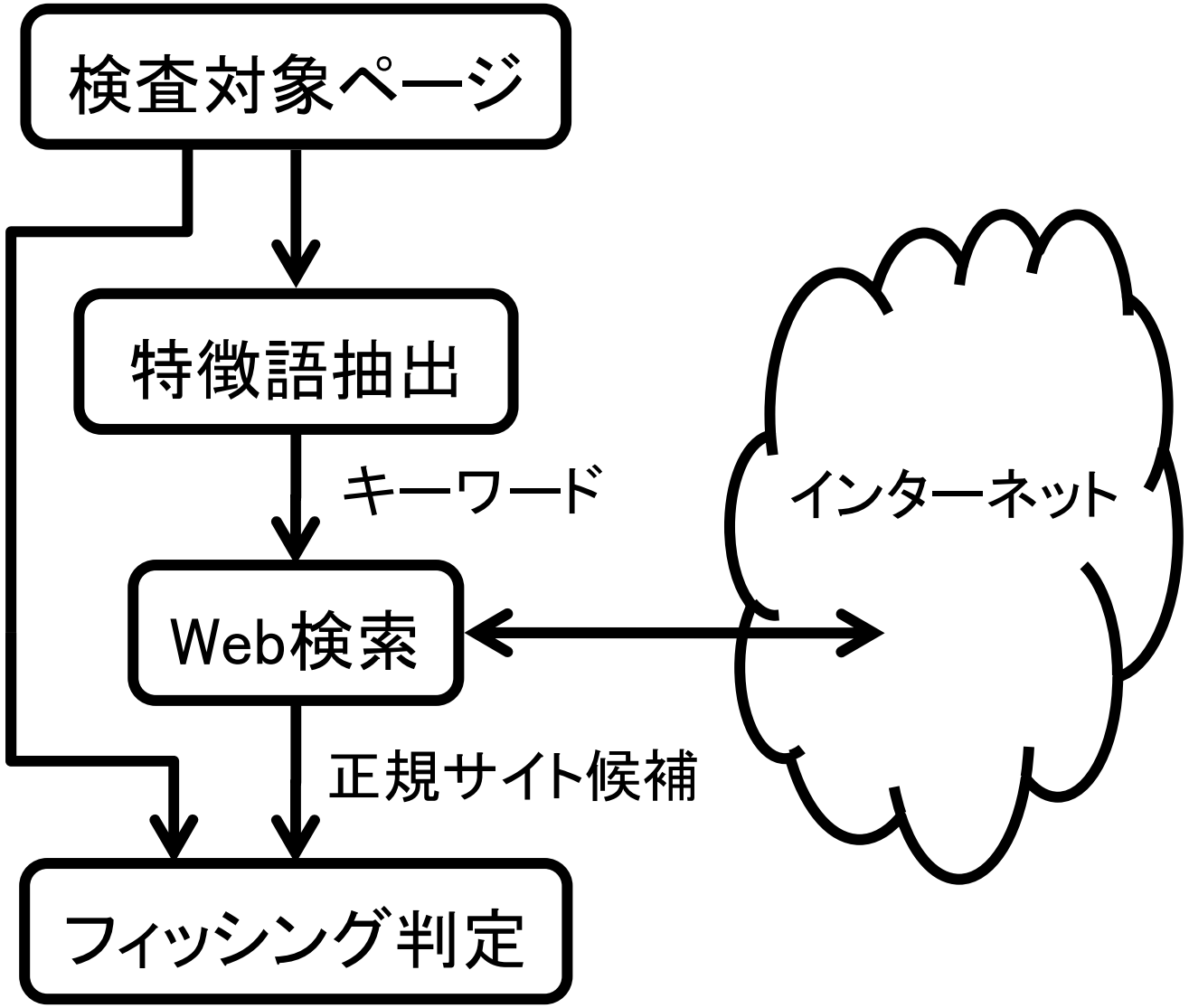
- 検査対象ページを自然言語処理して、特徴語抽出
- 特徴語で検索すると、正規サイトが出てくる

- 検査対象ページのドメインと、検索結果のドメインを比較
  - 不一致ならばフィッシングサイト
  - 一致なら正規サイト

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

検索エンジンを使った検知



© 2009 NTT Information Sharing Platform Laboratories

検索エンジン方式の特徴

Copyright© NTT corp. All Right Reserved.

- 認知プロセスを模倣
  - 自然言語処理と検索エンジンを利用
  - 集合知
- メンテナンスフリー
  - 検索エンジンをホワイトリストとして利用
- 高い検知率(論文から引用)
  - 正規サイト 92.4%
  - フィッシングサイト 97.1%

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

検索エンジン方式の問題点

- 正規サイトの誤検知率が高い
  - 自然言語処理の失敗確率が高い
  - 誤った単語を検索キーワードとしてしまう
  - 見当違いな検索結果を得てしまう

- 検索エンジンの性質に頼った方式
  - 検索エンジンのアルゴリズムが不明
  - 検索エンジンのクロール間隔が今より短くなったら、フィッシングサイトが検索エンジンに登録されてしまう

そうだ、ソーシャルブックマークを使おう！！

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

目次

- 背景
- フィッシングサイトの特徴
- 検索エンジンを用いたフィッシング検知手法
- SBMを用いたフィッシング検知手法
- SBMを用いた場合の未来シナリオ

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

仮説

・検索エンジンの評価　∝　SBMのブックマーク数

・正規サイトは検索エンジンの評価が高い
　→SBMのブックマーク数が多い
・フィッシングサイトは検索エンジンからの評価が低い
　→SBMのブックマーク数が少ない

SBMのブックマーク数を調べることで、
フィッシングサイト判定が可能なのではないか？

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

実験

- あるURLがSBMに何件登録されているか調べる
  - SBMサービスはDeliciousを利用
- 正規サイトのサンプル
  - Alexaから北米の銀行のURL 1309件
- フィッシングサイトのサンプル
  - Phishtankから最新のフィッシングサイト 200件

© 2009 NTT Information Sharing Platform Laboratories

実験結果

Copyright© NTT corp. All Right Reserved.

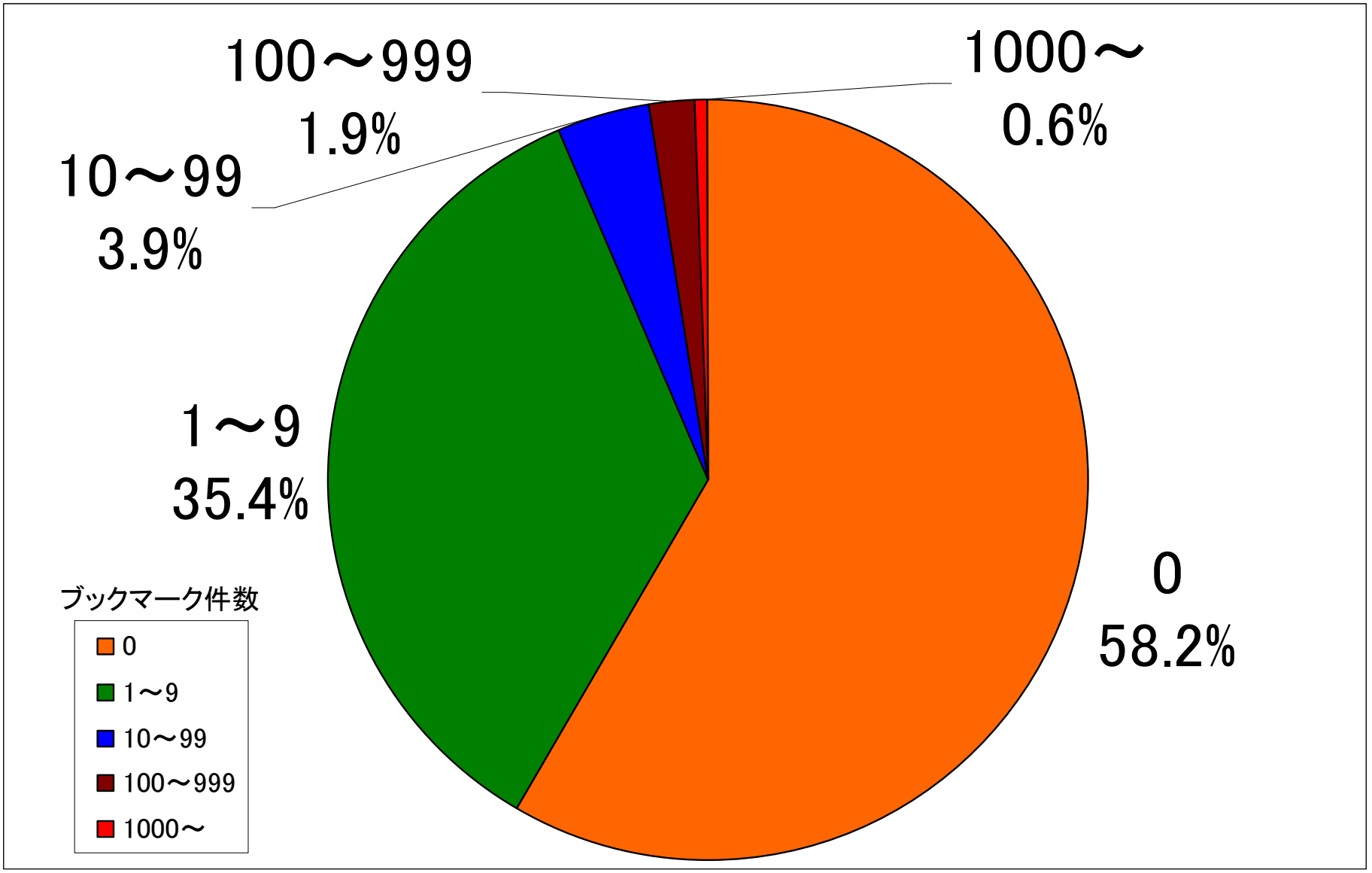
・正規サイトのうち、41.8%は
SBMにブックマークされていた

・フィッシングサイトは1件も
SBMにブックマークされていなかった

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

実験結果



© 2009 NTT Information Sharing Platform Laboratories

考察

Copyright© NTT corp. All Right Reserved.

- 正規サイトはSBMによくブックマークされているが、
  フィッシングサイトはSBMにあまりブックマークされない
  - 仮説は部分的には正しい
  - 既存ツールの検知率より低いので、
    判断の補助材料として利用可能なのではないか？

- 考察
  - 利用した母集団には地方銀行が多い
    - ほとんど更新されていない
  - SBMと地方銀行は相性が悪いのではないか
    - 誰が好き好んで地方銀行のサイトをブックマークするんだろう？

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

目次

- 背景
- フィッシングサイトの特徴
- 検索エンジンを用いたフィッシング検知手法
- SBMを用いたフィッシング検知手法
- SBMを用いた場合の未来シナリオ

© 2009 NTT Information Sharing Platform Laboratories

未来シナリオ1

Copyright© NTT corp. All Right Reserved.

・SBMに登録されていると、正規サイトだと判定される

・多くのサイトがSBMユーザーを向いた記事作り

→ウェブ上の情報発信速度の加速

・SBMの利用思想の変化

－従来は「自分のためにブックマークする」
　便利だから使えというが、実感するのは使った後

－これからは「みんなのためにブックマークする」

・SBM利用者の拡大

© 2009 NTT Information Sharing Platform Laboratories

未来シナリオ2

Copyright© NTT corp. All Right Reserved.

・SBMは誰でも登録可能、どんなURLでも登録可能

・フィッシングサイト製作者が、
　フィッシングサイトを登録可能

SBMオワタ＼(^o^)／

でも・・・

© 2009 NTT Information Sharing Platform Laboratories

未来シナリオ3

Copyright© NTT corp. All Right Reserved.

・SBM事業者は、
正規サイトとフィッシングサイトの両方のURLを得る

・SBM事業者は、SBMのURLに対してフィッシング検査
→ブラックリストを効率的に生成
（たとえば、検索エンジンを使うアルゴリズムとかで）

・従来はフィッシングサイトのURLさえ手に入らなかった
・SBMをフィッシング検知に使うことで、
フィッシング製作者から、URLを提供させることが可能

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

フィッシングサイト製作者の視点

- フィッシングサイトをSBMに登録しない
  - –フィッシングサイトが無条件で黒判定
- フィッシングサイトをSBMに登録する
  - –フィッシングサイトのURLをSBM事業者に知られる
  - –内容からフィッシングサイトだと判断され、ブラックリストに登録される
- フィッシングサイト製作者はどちらを選んでも、作ったフィッシングサイトが黒判定

© 2009 NTT Information Sharing Platform Laboratories

Copyright© NTT corp. All Right Reserved.

まとめ

- 集合知セキュリティの概念の紹介
- 検索エンジンを使ったフィッシング検知手法の紹介
- SBMを使ったフィッシング検知手法の検討
  - SBMに登録されてるなら正規サイト
- SBMを使ったフィッシング検知手法の未来シナリオ
  - SBMを向いたウェブサイト製作が拡大
  - SBMの利用者の拡大
  - SBMにフィッシングサイトが登録される
  - SBMのURLからブラックリストを生成、配布

© 2009 NTT Information Sharing Platform Laboratories