ダブルバウンスメールを活用した悪性メール対策の有効性

笠間 貴弘[†1a)] 神宮 真人[†1†2] 清水 雄介[†3] 井上 大介[†1]

概要：メールは依然としてサイバー攻撃における主要な攻撃経路の一つとして利用されており，標的型攻撃メールと呼ばれる特定の攻撃対象宛てに送る正規のメールとの区別が困難な巧妙なメールから，不特定多数に送信するフィッシングメールや広告目的のメールまで様々な悪性メールが存在する．ダブルバウンスメールは送信元および宛先メールアドレスが共に存在しないメール等によって発生するエラーメールであり，典型的には送信元メールアドレスを詐称しランダムな宛先メールアドレスに送られる悪性メールが原因で発生する．こうした悪性メールはボットネットなどを利用して大規模に送信されるケースがあるため，ダブルバウンスメールの分析を通じてボットネットの特定を行うといった既存研究が存在する．ダブルバウンスメールの発生源となったメールは確度高く悪性メールと判断できるため，当該メールに含まれる送信元メールアドレスや本文中の URL、添付ファイルといった各種情報を組織内の実ユーザに届く悪性メール対策に活用できる可能性がある．そこで本稿では，ダブルバウンスメールと実際に組織内の実ユーザに届いた悪性メールの関連性を調査し，ダブルバウンスメールの情報を活用した悪性メール対策の有効性について検証を行う．

キーワード：Malicious Email Detection，Double Bounce Email

Effectiveness of Malicious Email Detection based on Double Bounce Emails

Takahiro Kasama[†1a)] Masato Jingu[†1†2] Yusuke Shimizu[†3] Daisuke Inoue[†1]

Abstract: Email is still one of the most popular attack vectors in cyber attacks including advanced persistent threat, drive-by-download attack, and phishing. Bounce email is an error mail that is returned to the sender because it cannot be delivered for some reason, and double bounce email occurs when a bounce email also cannot be delivered. Double bounce emails typically occur due to emails which do not have any valid sender and recipient addresses. This kind of email is often sent from spamming botnet, and can be regarded as malicious email. Thus some studies analyzed them for identifying spamming botnets and spam campaigns. In this paper, we investigate whether the information such as sender address, URL in the body, and attached file obtained from double bounce emails is effective for detecting and blocking malicious emails which are delivered to existing users in the organization.

Keywords: Malicious Email Detection, Double Bounce Email

1. はじめに

電子メール（以下，単にメールという）はインターネット初期から現在に至るまで主要なコミュニケーションツールの一つと利用し続けられている．一方で，その利用の手軽さから悪用されることも多く，不特定多数のユーザに一方的に広告目的で送りつけられるメールから，ユーザの ID やパスワードを盗んだり直接的に金銭を振り込ませるフィッシング目的のメール，マルウェアを直接添付したりドライブ・バイ・ダウンロード攻撃の悪性 URL を含んだマルウェア感染を拡げるためのメール，特定の組織や個人を攻撃対象とした正規のメールと判断しづらい巧妙な標的型攻撃メールまで多種多様な悪性メールが存在する．Symantec のレポート[1]によれば，2015 年におけるスパムメール（悪性メール）数の割合はメール全体の 50%を超えているほか，情報処理推進機構からマルウェア添付のメールに対する注意喚起[2]が度々公表されるなど，未だに主要な攻撃経路の一つとなっており，効果的な悪性メール対策の実現は重要な課題となっている．

Send Mail Transfer Protocol（SMTP）では，メール送信時に送信元および宛先メールアドレスを自由に指定することが可能である．悪性メールにはランダム生成した存在しないメールアドレスや既に失効されている古いメールアドレス宛てに送信されるものがあり，このようなメールは宛先不明としてエラーメール（バウンスメール）が宛先メールサーバから送信元メールアドレス宛てに返信される．このとき，送信元を特定されないように存在しない架空のメールアドレスに送信元メールアドレスを詐称していると，バウンスメール自体が再び宛先不明のエラーとなり，バウンスメールの送信元に対して再度エラーメール（ダブルバウンスメール）が返信される（図 1 参照）．ダブルバウンスメールとなったメッセージはメールサーバによって破棄される．通常，このような送信元および宛先メールアドレスが存在しないメールが送信されることは，正規利用の範囲で

†1 国立研究開発法人 情報通信研究機構
National Institute of Information and Communications Technology
†2 株式会社日立システムズ
Hitachi Systems, Ltd.
†3 NTT アドバンステクノロジ株式会社
NTT Advanced Technology Corporation

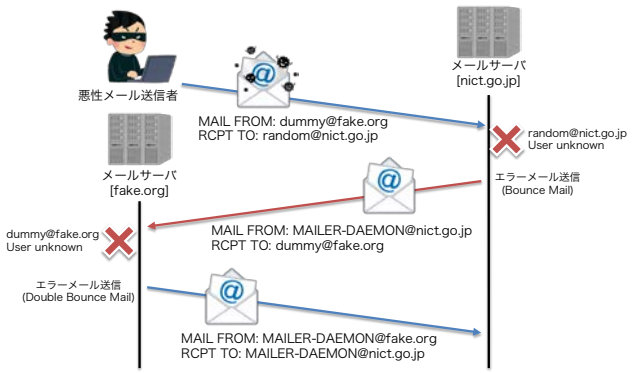

図 1　ダブルバウンスメールの概要
(送信元を存在しない dummy@fake.org に詐称し，存在しない random@nict.go.jp に悪性メールを送信した場合の流れ)

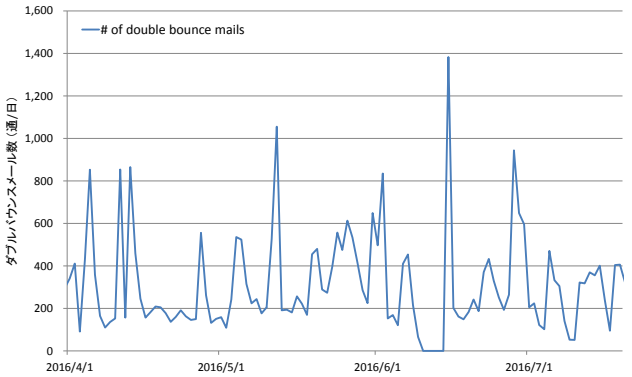

図 2　ダブルバウンスメール数の推移

は稀であり，ダブルバウンスメールの発生原因となったメールは確度高く悪性メールと判断できる．また，このような悪性メールはメール送信機能を備えたボットネットなどを通じて大規模に送信されるケースがあるため，ダブルバウンスメールの分析を通じて，悪性メール送信に用いられているボットネットの特定や攻撃キャンペーンの特徴を分析する研究が行われてきた．

一方，ダブルバウンスメールから得られる送信元メールアドレスや本文中の URL，添付ファイル等といった各種情報が組織内の実際のユーザに届く悪性メール全体の検知においてどの程度効果があるのかについては十分な調査が行われていない．そこで本稿では，ダブルバウンスメールと実際に組織内の実ユーザに届いた悪性メールの関連性を調査し，ダブルバウンスメールから取得できる各種情報を悪性メール全般のフィルタリングに用いた際の有効性について検証を行う．

本稿の構成は以下の通りである．まず 2 章で観測したダブルバウンスメールの統計情報およびそこから抽出した各種情報について分析を行う．3 章では，実ユーザ宛ても含む組織に届いた悪性メールの統計を示し，4 章でダブルバウンスメールから抽出した各種情報を利用したフィルタリングの効果について調査を行う．5 章で関連研究について述べ，6 章でまとめとする．

表 1　ダブルバウンスメール観測結果の統計

総ダブルバウンスメール (2016/04/01-07/21)	34,812	(通)
送信元メールアドレス数	24,688	(メールアドレス)
メールタイトル数	7,259	(タイトル)
URLを含むメール数	16,954	(通)
URL数（パラメータまで含む）	45,178	(URL)
URL数（パラメータ除外）	44,180	(URL)
URL数（ホスト名まで）	4,397	(URL)
添付ファイルを含むメール数	14,656	(通)
ファイル数（ファイル名で集計）	11,050	(個)
ファイル数（md5で集計）	6,628	(個)
転送経路のIPアドレス数	38,251	(アドレス)
2つ以上の経路で現れるIPアドレス数	1,863	(アドレス)

2.　ダブルバウンスメールの観測結果

本章では，情報通信研究機構（以下，NICT という）のメールサーバで観測したダブルバウンスメールについて分析を行う．

まず，2016 年 4 月 1 日から 2016 年 7 月 21 日までの期間で観測されたダブルバウンスメール数の推移を図 2 に示す．日によってばらつきはあるものの平均して 1 日約 300 通程度のダブルバウンスメールが観測されており（6 月 15 日のピークはシステム側の不具合で 6 月 10～15 日までのメールデータが合算されて集計されたことによる），上記期間（112 日間）の観測によって計 34,812 通のダブルバウンスメールが得られた．表 1 にダブルバウンスメールの発生源となったメール（以下，単にダブルバウンスメールという）から取得した送信元メールアドレスや添付ファイルなどの情報を示す．以下，それぞれの情報について詳細を分析する．

2.1　送信元メールアドレス

まず，ダブルバウンスメールの FROM ヘッダより，送信元メールアドレスの情報を抽出した．その結果，計 34,812 通から 24,688 メールアドレスが得られ，全体の約 7 割のダブルバウンスメールにおいて異なる送信元メールアドレスが用いられていることがわかった．基本的にこれらは全て送信元を詐称する目的で設定された存在しないメールアドレスである．

攻撃者による送信元メールアドレスの詐称の傾向を把握するために，抽出した送信元メールアドレスにおけるドメインの統計を表 2 に示す．表 2 を見ると，トップレベルドメインでの集計では最も多いのが com ドメインで 4 割だが，jp ドメインも約 2 割を占めており，日本のメールアドレスに詐称しているケースが多いことがわかる．実際にドメイン全体で見ると，“so-net.jp”や“softbank.ne.jp”，“docomo.ne.jp”など実際の日本のメールサービスのアドレスに詐称しているケースが多い．また，メールアドレスのローカルパート（アカウント名）については，24,688 メールアドレス中 23,590 アドレスで異なるアカウント名となっており，ランダムな英数字列ではなく人名のように見えるアカウント名が全体的に多いほか，後半に適当な数字を

表 2　ダブルバウンスメールの送信元メールアドレスの統計

	ドメイン（全て）	%	ドメイン (セカンドレベル+トップレベル)	%	ドメイン (トップレベル)	%
1	aa2.so-net.jp	1.6	ne.jp	10.3	com	41.5
2	airtelbroadband.in	1.5	so-net.jp	3.5	jp	18.7
3	softbank.ne.jp	1.3	co.jp	1.7	net	6.4
4	i.softbank.jp	1.3	airtelbroadband.in	1.5	pl	3.0
5	docomo.ne.jp	1.2	com.br	1.5	in	2.6
6	da2.so-net.jp	1.1	softbank.jp	1.3	br	2.0
7	yahoo.co.jp	1.0	com.tr	1.2	org	1.9
8	ttnet.com.tr	1.0	co.uk	1.2	uk	1.4
9	ab.auone-net.jp	0.9	com.ar	1.0	vn	1.3
10	gmail.com	0.8	auone-net.jp	0.9	tr	1.2

表 3　ダブルバウンスメールから抽出された URL の統計

	URL (パス名とパラメータ除外)	%	ドメイン (セカンドレベル+トップレベル)	%	ドメイン (トップレベル)	%
1	http://6url.ru	13.4	6url.ru	15.4	com	30.3
2	http://www.w3.org	8.1	w3.org	5.9	ru	16.8
3	http://funkyimg.com	5.0	funkyimg.com	1.9	org	11.6
4	http://fs5.directupload.net	1.7	directupload.net	1.9	top	7.6
5	https://support.google.com	1.7	postimg.org	1.8	download	3.9
6	http://comajci.org	1.6	google.com	1.5	date	3.6
7	http://ipaem.com	0.8	ac.cr	1.4	net	3.3
8	http://biltekharita.com.tr	0.8	oakleyc.top	0.8	bid	3.2
9	http://starkesod.com	0.7	comajci.org	0.8	pro	2.5
10	http://grupocodarco.com	0.6	com.tr	0.7	faith	1.9

付けているアカウント名（例：takahiro_kasama_12345）が多数観測された．このようなアカウント名のばらつきは悪性メール送信側のブラックリストによるフィルタリングを回避する意図によりものと推測できる．

2.2 メールタイトル

送信元メールアドレスにおいてはばらつきがある一方，メールタイトルについては計 34,812 通から 7,259 タイトルのみが得られ，異なる悪性メールでも受信側の興味が惹かれやすいタイトルが使いまわされる傾向が観測された．さらに，7,259 タイトルの中には，例えば“Bill N-6C5010”や“Bill N-78ACBF”のように請求書の送付を装ったタイトルで異なる ID が含まれているようなタイトルも多数存在していたため，ランダムに見える部分を除外した場合のタイトルのパターンはより少ない．メールタイトルおよび本文の言語としては英語のメールが約 85%で大半を占めており，残りは中国語と日本語のメールが同数程度であった．

2.3 URL

フィッシングサイトへの誘導やドライブ・バイ・ダウンロードによってマルウェア感染を試みる場合，メール本文中に URL を含める必要がある．ダブルバウンスメールより URL を抽出した結果，全メールの約半数にあたる 16,954 通が URL を含んでおり，計 45,178 種類の URL（パラメータまで含めたユニーク数）が抽出された．なお，パラメータを除外すると 44,180 種類，パス名を除外してホスト名までにした場合 4,397 種類にまで減少した．

表 3 に抽出された URL の統計を示す．表 3 を見ると，トップレベルドメイン単位で見ると，com ドメインが約 3 割を占めており，org ドメインや net ドメインといった初期のトップレベルドメインが多く観測された．一方国別コードトップレベルドメインに関しては，ロシアドメイン（ru）のみが全体の約 2 割と多数観測されており，その他は top ドメインや download ドメインなどの新しいトップレベルが多数用いられる傾向が見られた．

セカンドレベルドメインを含んだ統計や，URL（ホスト名まで）の統計を見ると，短縮 URL サービスの 6url.ru が全体の 13～15%を占めている．このような短縮 URL サービスの利用は，本来のアクセスさせたい URL を隠すことが可能でありブラックリストを回避する目的があると考えられる．しかしながら，短縮 URL サービスの URL が含まれていること自体が悪性メール判定に利用される場合もあるため，こうした短縮 URL サービスの利用が攻撃成功確率を必ずしも上げるとは言えない．

一方，URL（ホスト名まで）の統計で 6 位から 10 位にまでに位置しているサイトについて悪性メールから抽出した実際の URL を調べると，パス名が 6 文字程度の固定長の英数字列になっており（例：http://comajci.org/a1b2c3），実際にアクセスすると短縮 URL と同様にリダイレクトによって別サイトに移動することから，一見，6url.ru と同様の一般的に利用されている短縮 URL サービスを利用した URL のように見える．しかしながら，それぞれのサイトのトップページ（例：http://comajci.org）にアクセスしてみると，短縮 URL サービスを提供しているとは思えないまったく関連のない Web ページが表示されることがわかった．このことから，これらのケースでは攻撃者が正規の Web サーバを改ざんし，当該サーバ上で短縮 URL と同じ仕組みを動作させることで中継サーバとして悪用し，悪性メールに利用

表 4　ダブルバウンスメールの添付ファイルの統計

No.	拡張子	割合 (%)	VirusTotalの結果		
			検知 (%)	正常 (%)	結果なし (%)
1	zip	91.1	98.7	0.2	1.1
2	rar	3.0	57.6	31.3	11.1
3	jpg	1.9	0	5.7	94.3
4	docx	1.3	0.7	97.1	2.2
5	docm	1.0	100	0	0
6	gz	0.6	0	0	100
7	pdf	0.2	14.8	18.5	66.7
8	xlsx	0.2	0	95.8	4.2
9	xls	0.2	9.1	72.7	18.2
10	doc	0.1	26.7	66.7	6.7

している可能性が高いと推測できる。

なお, その他の上位に位置する URL は, http://www.w3.org は XHTML の名前空間の宣言で記載される URL，http://funkyimg.com と http://fs5.directupload.net は画像ファイルのホスティングサービスの URL，https://support.google.com は google のアドレスに詐称した場合等にエラーの説明用にエラーメッセージに記載される URL である．

2.4 添付ファイル

ダブルバウンスメールの添付ファイルを抽出した結果，14,656 通に添付ファイルが存在し，ファイル名で集計すると 11,050 種類，ハッシュ値（md5）で集計すると 6,628 種類の添付ファイルが得られた．表 4 にハッシュ値で集計した際の各添付ファイルの拡張子の統計と各ファイルのハッシュ値を基に VirusTotal [3]で各種ウイルス対策ソフトによる検知結果を検索した結果を示す．表 4 では，1 種類以上のウイルス対策ソフトによって検知されていたものは「検知」，ウイルス対策ソフトが 1 つも検知していないものは「正常」，過去に投稿されていないものは「結果なし」に分類している．なお，今回は各ファイルの VirusTotal への投稿は行っておらず過去に投稿されたことのあるファイルの検知結果を検索した結果のみを示している．

表 4 を見ると，添付ファイルの 9 割以上が zip 形式のファイルであり，pdf や doc 形式といった悪性メールでの攻撃によく用いられる文書形式のファイルが直接添付されているケースは稀であった．最も添付されていた zip 形式のファイルについては，その 98.6%が VirusTotal に投稿済みであり悪性判定がされていた. zip ファイルで圧縮されていた中身のファイル形式としては，exe 形式は 0.5%未満と非常に少なく，JavaScript／JScript ファイル（拡張子 .js）が全体の 85.6%と最も多く, Windows スクリプトファイル（拡張子 .wsf）が全体の 12.5%と，2 種類のスクリプトファイルがほとんどを占めていた．これらのスクリプトファイルの中身は基本的に外部のサーバから実行可能ファイルをダウンロード実行させるものになっていた．

その他にも，マクロ機能が有効な文書ファイルに付く拡張子の docm 形式のファイルは全て悪性判定されていた．一方，同じ文書ファイルであっても，docx 形式のファイルはほぼ全てが正常判定されており，PDF ファイルについては 3 分の 1 が過去に VirusTotal に投稿されていないなど，ファイル形式によって傾向が異なっていた．

2.5 メール転送経路

メールの Received ヘッダにはメールの転送経路情報，つまり悪性メールの中継経路として利用される中継サーバ（オープンリレーサーバ）の情報が含まれている．しかしながら，メールのヘッダ情報は容易に偽装できるため，先行研究[4]の手法を基に複数のダブルバウンスメールからメール転送経路情報を抽出することで，詐称されている可能性の低い中継サーバの IP アドレス情報を抽出した．

抽出の流れは以下の通りである．

1) 各ダブルバウンスメールの Received ヘッダからメール転送経路の IP アドレス情報を抽出する（なお, NICT アドレスは除外する）
2) 2 つ以上の異なる経路において同一 IP アドレスが現れる場合, 詐称されている可能性が低い経路情報として当該 IP アドレスを抽出する
3) 2)で抽出された各 IP アドレスに対して，逆引きアドレスの登録確認および MX レコードとして当該 IP アドレスが登録されているかを確認する．登録が確認できた IP アドレスについては正規のメールサーバと判断し除外する

ステップ 1) で Received ヘッダから抽出した結果 38,251 IP アドレスが得られ，ステップ 2) で 2 つ以上の異なる経路に現れたものに限定することで 1,863 IP アドレスまで絞り込まれた．そしてステップ 3) の正規メールサーバ確認によって 7 IP アドレスが除外され，最終的に 1,856 IP アドレスが中継サーバの IP アドレスとして抽出された．

各 IP アドレスについて GeoIP データベース[Geo]で国を調べたところアメリカが 20.2%で最も多く，続いて中国が 9.1%，日本が 7.0%となっており，計 100 ヶ国に広く分布していた．また，これらの IP アドレスについて，スパムメール対策組織の Spamhaus が管理するブラックリスト[5]で検索した結果，約半数の 910 IP アドレスが悪性アドレスとして登録されていた．

3. 組織内の実ユーザ宛ての悪性メールの統計

本章では，ダブルバウンスメールに限らず，NICT 全体に届いた悪性メールについて調査を行う．なお，悪性メールの判定には我々の所有する 3 種類の市販のセキュリティアプライアンスを用いた．3 種類のセキュリティアプライ

表 5　各セキュリティアプライアンスによる悪性メール検知結果の統計

	アプライアンスL	アプライアンスF	アプライアンスT	
検知メール数	15,547	69,095	23,955	(通)
送信元メールアドレス数	11,243	63,477	11,688	(メールアドレス)
メールタイトル数	3,437	8,931	9,010	(タイトル)
URLを含むメール数	-	483	11,573	(通)
URL数（全体）	-	59	6,029	(URL)
URL数（パラメータ除外）	-	53	5,243	(URL)
URL数（ホスト名まで）	-	48	4,000	(URL)
添付ファイルを含むメール数	15,547	69,095	15,132	(通)
ファイル数（md5で集計）	3,762	30,025	2,874	(個)

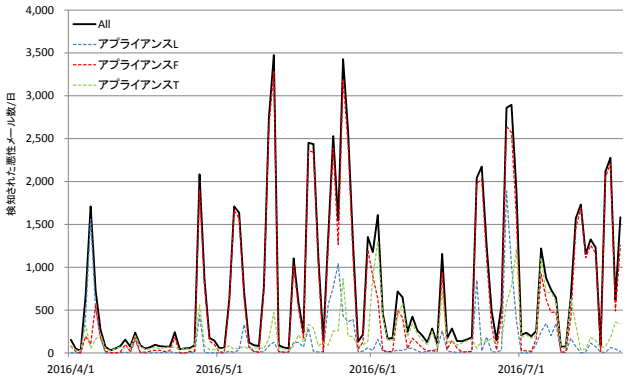

図 3　機構内アドレスに届いた悪性メール数の推移

アンス（ここではアプライアンス L，F，T とする）は悪性メールの検知機能を備えており，機構内に流れる全てのトラフィックを入力とし，自動的にメールを再構築して検知を行うようになっている．

まず，図 3 にダブルバウンスメールの観測期間と同じ 2016 年 4 月 1 日から 7 月 21 日までに各アプライアンスで検知された悪性メール数の推移を示す．「All」は 3 つのアプライアンスのうちどれか 1 つ以上のアプライアンスで検知された悪性メールの統計である．なお，機構内のトラフィックは複数地点で観測を行っている都合上，トラフィックの重複等によって同一のメールが重複して抽出・検知される場合があるため，検知メール数は各メールの Message-ID ヘッダのユニーク数で集計をしている．図 3 を見ると，日毎に検知される悪性メール数は多い日で 3000 通以上，少ない日で数十通とかなり検知数に波があることがわかる．また，各アプライアンスの検知数にも大きな差あり，期間中に 3 つのアプライアンスによって検知された悪性メールの総数は計 85,384 通であった．

各アプライアンスにおける検知結果の統計を表 5 に示す．

3 種類のアプライアンスはいずれもトラフィックからメールの添付ファイルを自動で抽出し，動的解析によって悪性判定を行う機能を備えるほか，各種シグネチャ等による判定なども行っている．それぞれのアプライアンスの検知結果を見ると，アプライアンス L は，期間中 15,547 通のメールを悪性メールとして検知した．アプライアンス L ではメールの悪性判定では添付ファイルの動的解析結果が主な判定基準であり，URL に関しては実際にユーザがアクセスした際の Web トラフィックなどから悪性判定を行う仕組みとなっている．そのため，検知した悪性メールは全て添付ファイル付きのメールであり，URL に関してはシステムの出力からも情報が得られず空欄となっている．アプライアンス F は検知数が一番多く，他の 2 つのアプライアンスの約 2 倍以上の期間中 69,095 通のメールを悪性メールとして判定した．こちらもアプライアンス L と同様に，メールの判定においては添付ファイルの動的解析による判定が主であると推測され，検知されたメールは全て添付ファイルを含むメールであった．最後に，アプライアンス T は期間中 23,955 通のメールを悪性メールとして判定した．アプライアンス T は他の 2 つのアプライアンスとは異なり，添付ファイルの含まれないメールも多数検知しており，URL や送信元アドレスのシグネチャもしくはレピュテーションによる判定機能の影響が大きいと推測できる．

3 種類のアプライアンスの結果を見ると，検知数にかなりの差が見られ，各々のアプライアンスでのみ検知されたメールも相当数あるなど，アプライアンス毎の判定エンジンの差異がかなり現れている．

4. ダブルバウンスメールを活用した悪性メール対策の有効性

本章では，ダブルバウンスメールから取得できる各種情報（送信元メールアドレスや本文中の URL，添付ファイル等）を，実際のユーザに届く悪性メール全体の検知に利用した際の有効性について検証を行う．特に今回は最もシンプルな活用方法としてダブルバウンスメールから抽出した各種情報をブラックリストとして用いた際の有効性を検証する．

まず 4.1 節において，ダブルバウンスメール以外の悪性メールの検知に対する有効性を確認するために，3 章で示した各種セキュリティアプライアンスで検知された悪性メールを用いて検証する．次に 4.2 節において，NICT に一定期間中に届いた全てのメールを用いて誤検知の影響を検証する．

4.1 アプライアンスが検知した悪性メールとの突合

3 章で示した，各セキュリティアプライアンスが検知し

表 6　ダブルバウンスメールの情報を利用した悪性メールの検知結果

	アプライアンスL	アプライアンスF	アプライアンスT	
各アプライアンスによる検知メール数	15,547	69,095	23,955	(通)
ダブルバウンスメールを除いた検知メール数	11,957	55,165	23,955	(通)
マッチングで検知できた総メール数	6,114 (51.1%)	36,419 (66.0%)	14,910 (62.2%)	(通)
送信元メールアドレスで一致	707	583	2,019	(通)
タイトルで一致	2,609	17,894	10,622	(通)
URLで一致	-	170	6,565	(通)
添付ファイル（ハッシュ値）で一致	4,898	27,812	4,364	(通)
メール転送経路（IPアドレス）で一致	-	12,134	-	(通)

た悪性メールについて，ダブルバウンスメールから得られた情報を基にどの程度検知可能か検証を行う．

最初に，各セキュリティアプライアンスは NICT 内の全てのトラフィックを入力としているため，表 5 で示した検知メールの中には NICT に届いたダブルバウンスメールも含まれている可能性がある．そこで，各アプライアンスが検知したメールからダブルバウンスメールを除外した上で，残りの悪性メールに対して 2 章で得られたダブルバウンスメールに関する以下の情報をブラックリストとして用いて検証を行った．なおメールタイトルについては，返信や転送時に付く“Re:”や“Fwd:”といったもののみのタイトルは除外してマッチングを行った．

- 送信元メールアドレス
- メールタイトル
- 本文中の URL
- 添付ファイル（ハッシュ値）
- メール転送経路（IP アドレス）

各情報でマッチングできたメール数およびどれか 1 つ以上の情報で検知可能であった総メール数を表 6 に示す．

表 6 を見ると，ダブルバウンスメールの情報を活用することで，ダブルバウンス以外の悪性メールについても半数から約 7 割弱が検知できていることがわかる．検知に用いた各情報で見ると，タイトルと添付ファイル（ハッシュ値）で検知できたメールの割合が多く有効性が高いことがわかる．また，URL に関しても一見アプライアンス F が検知したメールとのマッチング数が少ないように見えるが，これは元々アプライアンス F で検知された URL を含む悪性メールが 483 通と少ないためであり，実際にはそれらの 3 分の 1 以上が検知できている．メール転送経路についても，検知したメールの Received ヘッダの情報が得られたアプライアンス F の結果のみではあるが，約 2 割の悪性メールを検知できている．

一方，送信元メールアドレスのマッチングは他の情報と比較してあまり効果が無かった．これは 2.1 節で示したダブルバウンスメールの送信元アドレスと同様に，アカウント名の部分を毎回変えるなどして異なる送信元メールアドレスに詐称されている悪性メールが多いことで検知数が少なかった可能性が高い．

表 7　正規メールに対する誤検知の検証結果

総メール数（2016/07/31-08/06）	266,095	(通)
アプライアンスで検知されたメールを除外	255,755	(通)
送信元メールアドレスで検知	1,154 (0.45%)	(通)
検知された送信元メールアドレス数	110	(メールアドレス)
タイトルで検知	2,351 (0.92%)	(通)
検知されたタイトル数	137	(タイトル)
添付ファイル（ハッシュ値）で検知	39 (1.5×10^{-4}%)	(通)
検知された添付ファイル数	17	(個)
メール転送経路（IPアドレス）で検知	2,249 (0.88%)	(通)
検知された転送経路	292	(アドレス)

4.2 正規メールを用いた誤検知の検証

次に，誤検知の影響を明らかにするために，2016 年 7 月 31 日から 8 月 6 日までの 1 週間で NICT に届いた計 266,095 通のメールを用いて検証を行った．

まず観測されたメールから正規のメールだけを抽出するため，前述の 3 つのセキュリティアプライアンスで検知された悪性メールを除外した結果，255,755 通が正規メールとして得られた．この正規メールに対して，4.1 節と同様にダブルバウンスメールから得られた各種情報とマッチングを行った結果を表 7 に示す．

表 7 を見ると，タイトルおよびメール転送経路情報による誤検知が他と比較して多くなっているが，それでも 1% 未満の誤検知率に収まっていることがわかる．特に添付ファイルを用いた場合には非常に低い誤検知率となっているまた，実際に誤検知の原因となった各種情報のユニーク数は誤検知されたメール数に対して 10 分の 1 程度と少なく，特定の少数の情報による誤検知が頻発していることが考えられる．

実際に，各々の誤検知された情報を確認したところ，送信元メールアドレスについては，NICT 内のある部署のシステムからの自動通知メールがダブルバウンスメールとなっており，当該メールの送信元メールアドレスがブラックリストに含まれた結果として誤検知が多数発生していることがわかった．当該アドレスを除外すると送信元メールア

ドレスによる誤検知率は 0.45%から 0.32%まで減少した．メールタイトルについては，メールサーバからのエラーメッセージに用いられるタイトルがダブルバウンスから抽出され多数の誤検知が発生していた．それらのメールタイトルを除外したところメールタイトルによる誤検知率は 0.92%から半分以下の 0.43%まで減少した．転送経路の IP アドレスについては, とある学会のメールサーバの IP アドレスによる誤検知が全体の 3 分の 1 を占めていた．これは学会のメールサーバから過去に存在していたと思われる NICT のメールアドレスに転送されたメールがダブルバウンスとなったことが原因であり, 当該 IP アドレスを除外したところ誤検知率は 0.88%から 0.6%まで減少した．

実際のフィルタリングの運用時には，ユーザから誤検知の報告が頻発するものについては適時ブラックリストから除外することで対応が可能であると考えられる．

4.3 結果まとめ

4.1 節の結果より，ダブルバウンスメールから得られる各種情報を用いることで実際のユーザに届く悪性メールの半数以上が検知できており，フィルタリングとしての一定の有効性があると考えられる．ダブルバウンスメールから得られる情報は, 例えば経路情報の IP アドレスであれば半数以上が既存のブラックリストに含まれておらず，添付ファイルについても VirusTotal に投稿されていない新たなファイルが含まれるなど，既存の対策でカバーされていない情報が得られていると言える．また，URL に関しては改ざんされた Web サイトが悪用されている可能性が高い独自短縮 URL のような URL が多数観測されていた．こうした正規サーバを悪用した URL では, ブラックリストやレピュテーションでの検知が難しくなるため，ダブルバウンスの観測結果を用いることで悪用された改ざんサイトを早期に発見し，フィルタリングに活用できる可能性がある．

一方，誤検知に関しても，得られた情報をそのまま用いた場合でも 1%未満の誤検知率となっており，誤検知が発生した情報も少なかったため，それらの原因を調査しブラックリストから除外することで誤検知を相当数減らすことができた．

5. 関連研究

悪性メールをフィルタリングする最も単純な方法は送信元の IP アドレスでのフィルタリングである．ある IP アドレスがブラックリストに含まれるか否か DNS プロトコルを利用して確認可能なブラックリストは DNSBL[6]などと呼ばれ, 多くのブラックリストが公開されている. また，初めてメールを送信してきたサーバに対して一時的なエラーコードを応答しリトライを促すグレイリスティングと呼ばれる方式もある．これは，正規のサーバであれば再送を行ってくるが，悪性メールの送信サーバの場合は再送を行わないケースが多いという経験則に基づくフィルタリング手法である．他にもボットに感染したエンドユーザのコンピュータが直接メールを送信するようなケースを防ぐために, 送信元 IP アドレスを逆引きして得られるホスト名の特徴からフィルタリングを行う S25R[7]などもある．

一方，送信元メールアドレスを詐称した悪性メールを検知する技術として送信ドメイン認証が存在する．DNS を利用した方法として SPF（Sender Policy Framework）[9]が，電子署名を利用した方法として DKIM（Domain Keys Identified Mail）[10]が，それら両方を利用した DMARC（Domain-based Message Authentication, Reporting & Conformance）[11]などが存在する．これらの手法によってフィルタリングを行うには，送信側と受信側の両方の対応が必要となるが, IIJ のレポート[12]によると 2016 年 1 月から 3 月までの期間において IIJ のメールサービスで受信したメールのうち，送信側で SPF を導入している割合は 77.4%，送信側で DKIM を導入している割合は 20.1%，DMARC の導入割合は 13.1%となっており，一定割合に普及しているものの全てのメールサーバが対応するにはまだ時間がかかると思われる．

直接的なフィルタリングの他に，収集した悪性メールの分析を通じて攻撃傾向を把握したり，悪性メール送信に利用されるボットネットを特定したりする研究も数多く存在する．志村ら[13]は，スパムトラップによって収集した添付ファイル付き悪性メールの分析を行っている．スパムトラップは過去に利用されていたが現在は誰も利用していないメールアドレスに届くメールを悪性メールとして収集する手法であり，主にメールサービスプロバイダで行われている．基本的な考え方はダブルバウンスメールで悪性メールを観測する手法と同じであるが，ダブルバウンスメールでは送信元アドレスが詐称されたものだけが観測されるという点で，正規メールが紛れ込む可能性がより低いと言える．中里ら[4]は，複数のダブルバウンスメールの情報から偽装されている可能性の低いオープンリレーサーバを抽出する手法を提案し，ダークネット観測結果と突合を行うことでボット判定などを行った．Kim ら[14]は，DNS シンクホールの効果を高めるために，ダブルバウンスメールに含まれる URL を自動的に抽出・分析し, 悪性判定された URL のドメインを DNS シンクホールに自動的に追加するシステムを構築している．Song ら[15]は，悪性メールのクラスタリングを目的とした K-means 法を改良した O-means 法を提案し，ダブルバウンスメールを用いて精度評価を行うほか，ダークネット観測およびハニーポットの観測結果と突合し，ボットネットの特定を行っている[16]．Stringhini ら[17]は，既知のスパムキャンペーンにおけるスパム送信ボットの特徴を基に，メールサーバのログから同一のボットネットに属する未知のスパム送信ボットを抽出する手法を

提案しているほか，メール送信クライアントの SMTP 実装の差異による特徴を学習し，正規のメールクライアントとスパム送信ボットを識別する手法を提案している[18].

6. まとめ

本稿では，組織に届くダブルバウンスメールから得られる送信元メールアドレス，URL，添付ファイルといった各種情報を，実ユーザに届く悪性メールフィルタリングに利用する有効性について検証を行った．検証の結果，ダブルバウンスメールの情報を単純にブラックリストとして利用した場合でも全悪性メールの約半数から 7 割弱の悪性メールを検知でき，一定の有効性があることが明らかになった．また，誤検知に関しても 1%未満の低い誤検知率に抑えることができている．ダブルバウンスメールの観測・利用は，囮のメールアドレスを準備したり，そのメールアドレスに悪性メールが届くようにインターネット上で広告したりする必要がなく導入が容易であり，分析の結果，ダブルバウンスメールから得られた情報には既存のブラックリスト等に含まれない IP アドレスや添付ファイル，URL といった情報も含まれていたことからも，既存の対策手法を補完する効果が期待でき，自組織に届く悪性メールフィルタリングの一つとして有効であると考えられる．

参考文献

[1] Symantec, “Security Response Publications,” https://www.symantec.com/security_response/publications/monthlythreatreport.jsp

[2] 独立行政法人情報処理推進機構, “2015 年 12 月の呼びかけ：「ウイルス感染を目的としたばらまき型メールに引き続き警戒を」,” https://www.ipa.go.jp/security/txt/2015/12outline.html

[3] “VirusTotal - Free Online Virus, Malware and URL Scanner,” https://www.virustotal.com/en/

[4] 中里純二, 班涛, 島村隼平, 衛藤将史, 井上大介, 中尾康二, “メール転送経路に着目したスパムメール分析,” 電子情報通信学会 信学技報, Vol. 113, No. 502, ICSS2013-75, pp. 101-106, 2014 年 3 月.

[5] “DNSBL Usage Terms - The Spamhaus Project,” https://www.spamhaus.org/organization/dnsblusage/

[6] “DNSBL Information - Spam Database and Blacklist Check,” http://www.dnsbl.info/

[7] 浅見秀雄, “Selective SMTP Rejection (S25R)方式,” http://www.gabacho-net.jp/anti-spam/anti-spam-system.html

[8] M. Wong and W. Schlitt, “Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1” https://tools.ietf.org/html/rfc7208

[9] S. Kitterman, “A Methodology for Multipurpose DNS Sinkhole Analyzing Double Bounce Emails,” Proc. 20th International Conference on Neural Information Processing (ICONIP2013), pp. 609-616, 2013.

[10] D. Crocker, T. Hansen, and M. Kucherawy, “DomainKeys Identified Mail (DKIM) Signatures,” https://www.ietf.org/rfc/rfc6376.txt

[11] M. Kucherawy and E. Zwicky, “Domain-based Message Authentication, Reporting, and Conformance (DMARC),” https://tools.ietf.org/html/rfc7489

[12] 櫻庭秀次, “Internet Infrastructure Review (IIR) Vol.31 メッセージングテクノロジー「迷惑メール最新動向」,” http://www.iij.ad.jp/company/development/report/iir/pdf/iir_vol31_message.pdf

[13] 志村正樹, 畑田充弘, 森達哉, 後藤滋樹, “スパムトラップを用いたマルウェア添付スパムメールの分析,” 情報処理学会コンピュータセキュリティシンポジウム 2015 論文集, 2015.

[14] H. S. Kim, S. S. Choi, and J. Song, “A Methodology for Multipurpose DNS Sinkhole Analyzing Double Bounce Emails,” Proc. 20th International Conference on Neural Information Processing (ICONIP 2013), 2013.

[15] J. Song, D. Inoue, M. Eto, H. C. Kim, and K. Nakao, “O-means: An Optimized Clustering Method for Analyzing Spam Based Attacks,” IEICE Transactions, Vol. 94-A, No. 1, pp. 245-254, 2011.

[16] J. Song, J. Shimamura, M. Eto, D. Inoue, and K. Nakao, “Correlation Analysis between Spamming Botnets and Malware Infected Hosts,” Proc. 11th IEEE/IPSJ International Cymposium on Applications and the Internet (SAINT2011), pp. 372-375. 2011.

[17] G. Stringhini, T. Holz, B. S. Gross, C. Kruegel, and G. Vigna, “BOTMAGNIFIER: Locating Spambots on the Internet,” Proc. 20th USENIX Security Symposium, 2011

[18] G. Stringhini, M. Egele, A. Zarras, T. Holz, C. Kruegel, and G. Vigna, “B@bel: Leveraging Email Delivery for Spam Mitigation,” Proc. 21st USENIX Security Sympoium, 2012.