A Spoofed E-mail Countermeasure Method by Scoring the Reliability of DKIM Signature Using Communication Data

Kanako Konno*, Kenya Dan*, and Naoya Kitagawa†

*Department of Computer and Information Sciences, Faculty of Engineering,
Tokyo University Agriculture and Technology, Tokyo, 184-8588, Japan
†Division of Advanced Information Technology and Computer Science, Institute of Engineering,
Tokyo University Agriculture and Technology, Tokyo, 184-8588, Japan

E-mail:*{k_konno,kdan}@net.cs.tuat.ac.jp, †nakit@cc.tuat.ac.jp

***Abstract*—Recently, spoofed e-mails become sophisticated and the messages that cannot be identified by visual confirmation have been increasing. DKIM is one of the most famous methods as countermeasure of spoofed e-mail. However, since DKIM allows a signature by third party domains, a recipient server can be unable to determine whether the attached signature is the legitimate. To solve the problem, although a method using data of DKIM signature domain passed DKIM verification is proposed, this method only verifies the existence of the past deliveries, therefore it has an issue in reliability. To overcome the issue, in this paper, we propose a method to score the reliability of combinations of senders Header-From domain and DKIM signature domain by analyzing fluctuation of the existence of e-mail delivery obtained from the past communication data. By utilizing this method, the recipients can obtain the reliability score for each message including spoofed e-mail which was uniformly verified successfully in the ordinary DKIM verification, and can utilize our mechanism to spoofed e-mails countermeasure system.**

. ***Keywords*—Spoofed e-mail; DKIM; Sender Domain Authentication; spam mail; anti spam.**

I. Introduction

E-mail is one of the most utilized services all over the world as a convenient way to communicate with each other or notify various information from companies to customers. However, the rapid increase in spoofed e-mail is a serious problem. Spoofed e-mail is abused by attackers in order to steal personal and/or sensitive information or send malicious programs such as computer viruses.

The damage of business e-mail scams that are directly attributing to spoofed e-mails has been rapidly increasing. According to the newest statistics report of FBI, 22,314 companies were damaged, and the total financial damage is 3.1 billion US dollar from October 2013 to June 2016 [1].

Sender domain authentication is an effective method as countermeasure of spoofed e-mail. DomainKeys Identified Mail (DKIM) [2] is one of the most widely used methods among sender domain authentication. In DKIM, the sender adds the electronic signature generated from e-mail header and body to the e-mail header. Then, the recipient obtains the public key of the sender's domain, and verifies its e-mail. However, DKIM has a problem that can be unable to determine whether the signer of the attached signature is the legitimate. DKIM permits to sign by the domain different from the domain of Header-From domain (hereinafter, this address is called "RFC5322.From domain"). Although many legitimate senders utilize such a third party signature, many attackers can exploit this mechanism and send a large number of spoofed e-mails which succeed in DKIM verification.

As a framework to solve the problem of DKIM, Domain-based Message Authentication, Reporting, and Conformance (DMARC) [3] has been proposed. DMARC uses SPF and/or DKIM, and has the concept called "alignment". With this concept, DMARC verification gets failed when a sender uses a third party signature, thus the problem of DKIM to third party signature cannot be happened. However, DMARC is not a widely utilized framework at the moment. Actually, since many DKIM compliant domains use third party signature, the penetration rate of DMARC is only 29% as of 2016 [4].

At the present time, our research group previously reported a method to verify the reliability of DKIM signature by checking the past combination of RFC5322.From and DKIM signature domains [5]. However, this method simply collates the combination of these domains observed in the past communication, legitimate deliveries and suspicious deliveries are treated equally in this method.

In this paper, we propose a method to measure the reliability of combinations of sender's RFC5322.From domain and DKIM signature domain by score. Our approach measures the reliability of these domains' combination by scoring based on transition regarding existence of e-mail for each combination getting from the past communication data and statistical evaluation. In addition, our approach has the ability to detect moving of DKIM signing domain and reflect on scoring.

The paper is organized as follows. In section II, we describe several existing methods of countermeasure for spoofed e-

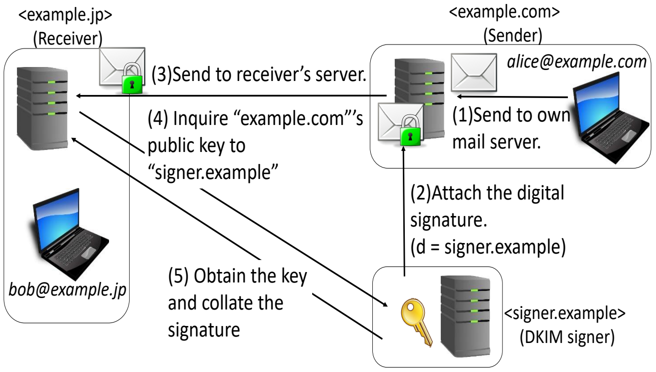
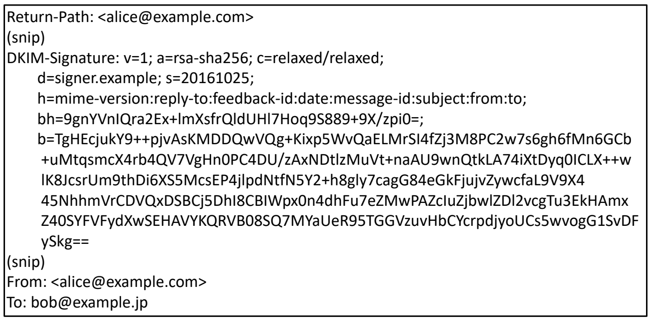

<example.jp>
(Receiver)
(3)Send to receiver's server.
<example.com>
(Sender)
alice@example.com
(1)Send to own mail server.
(4) Inquire "example.com"'s public key to "signer.example"
(2)Attach the digital signature.
(d = signer.example)
bob@example.jp
(5) Obtain the key and collate the signature
<signer.example>
(DKIM signer)

Fig. 1. Flow of DKIM verification.



```
Return-Path: <alice@example.com>
(snip)
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
      d=signer.example; s=20161025;
      h=mime-version:reply-to:feedback-id:date:message-id:subject:from:to;
      bh=9gnYVnIQra2Ex+lmXsfrQldUHl7Hoq9S889+9X/zpi0=;
      b=TgHEcjukY9++pjvAsKMDDQwVQg+Kixp5WvQaELMrSI4fZj3M8PC2w7s6gh6fMn6GCb
       +uMtqsmcX4rb4QV7VgHn0PC4DU/zAxNDtlzMuVt+naAU9wnQtkLA74iXtDyq0ICLX++w
       lK8JcsrUm9thDi6XS5McsEP4jlpdNtfN5Y2+h8gly7cagG84eGkFjujvZywcfaL9V9X4
       45NhhmVrCDVQxDSBCj5DhI8CBIWpx0n4dhFu7eZMwPAZcIuZjbwlZDl2vcgTu3EkHAmx
       Z40SYFVFydXwSEHAVYKQRVB08SQ7MYaUeR95TGGVzuvHbCYcrpdjyoUCs5wvogG1SvDF
       ySkg==
(snip)
From: <alice@example.com>
To: bob@example.jp
```

Fig. 2. Example of e-mail header.

TABLE I
PERCENTAGE OF DKIM SIGNATURE DOMAIN USING THIRD PARTY SIGNATURE

Month / Year	# of signature domains	# of Third party signature domains (%)
Feb. / 2017	266,136	209,454 (78.70%)
Jan. / 2017	258,051	206,518 (80.03%)
Dec. / 2016	270,922	218,544 (80.76%)

mail. In section III, we explain the design of our scoring system. Section IV shows evaluation of this method by applying actual communication data to the system. Finally, in section V, we present concluding remarks.

II. Existing methods

Although a large number of anti-spam methods have been proposed over the years such as Bayesian Filter [6] [7], SpamAssassin [8] [9] and so on, these methods are not a direct measure to spoofed e-mail. On the other hand, sender domain authentication is a method for direct countermeasure of spoofed e-mail. In this section, we describe four existing countermeasure methods against spoofed e-mail.

A. Sender Policy Framework (SPF)

Sender Policy Framework (SPF) [10] is a method to confirm whether the IP address of the sender's SMTP server is legitimate or not by checking SPF record. SPF record indicates a list of IP address of severs that the senders may use to send e-mails. The administrator of the sender domain explicitly publishes on the authoritative DNS server of their domain beforehand. The receiver queries the sender's DNS server for the SPF record that using sender's Envelope-From domain, then verifies whether the IP address of the sender's SMTP server is included in it. However, SPF has a problem that the verification cannot succeed for forwarded messages. When a message was forwarded, the original IP address of the SMTP server changes to the relay server's IP address which does not include the SPF record. As a result of this, there are many cases where valid mail fails the verification.

B. DomainKeys Identified Mail (DKIM)

DomainKeys Identified Mail (DKIM) is a method to authenticate by using the digital signature generated from the e-mail body and header.

Fig.1 shows an example of the flow of DKIM verification and Fig.2 shows an example of e-mail header.

At first, in order to utilize the DKIM mechanism, the sender domain ("example.com" in Fig.1) prepares a pair of a private key and a public key beforehand, and publishes the public key on their authoritative DNS server of the domain for DKIM verification ("signer.example" in Fig.1). Then, the sender domain ("example.com") generates the DKIM signature from the e-mail body and header using the private key, and attaches it to the e-mail header as the DKIM signature as shown by "b=" tag in Fig.2.

Next, the receiver ("example.jp" in Fig.1) inquires the public key to the sender specified domain authoritative DNS server that is shown in the "d=" tag of the DKIM signature ("signer.example" in Fig.1 and Fig.2). Then, the receiver obtains the hash value from the digital signature using the public key, and compares with the value of "bh=" tag of the DKIM signature. When these values are the same, the e-mail is passed the DKIM verification. With this mechanism, DKIM can verify correctly even forwarded messages unlike SPF.

As described above, DKIM signature domain do not need to relate with the name of sender's domain in DKIM mechanism. Moreover, the receiver cannot distinguish whether the DKIM signer is legitimate or not. As a result, spammer can send the spoofed e-mails with the DKIM signature using their own malicious domain, which can pass the DKIM verification.

TABLE I shows the percentage of DKIM signature domain using third party signature based on the number of domains that we have observed. As shown in the table, we can confirm that approximately 80% of the DKIM compatible domains utilize third party signature.

C. Domain-based Message Authentication, Reporting, and Conformance (DMARC)

M. Kucherawy and E. Zwicky have proposed a framework called Domain-based Message Authentication, Reporting, and Conformance (DMARC) [3] that can solve the third party signature problem of DKIM. This framework is a reporting and policy control mechanism which uses two sender domain

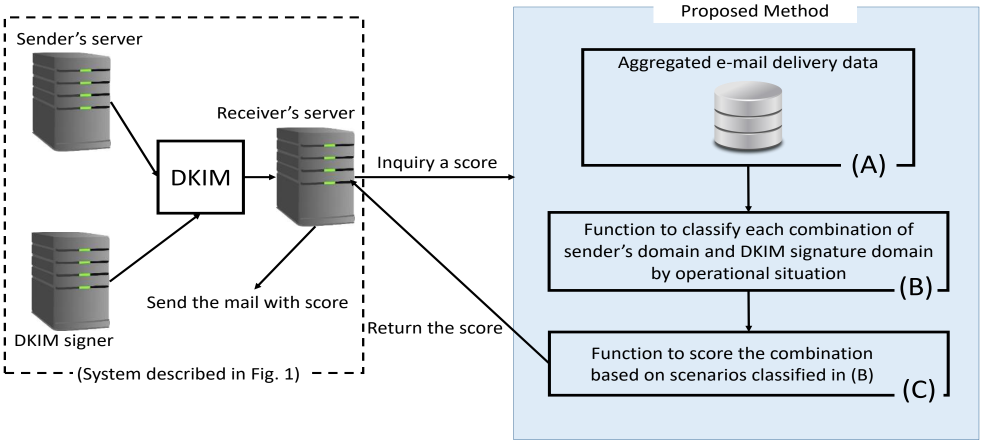

Sender's server
Receiver's server
DKIM
Inquiry a score
Send the mail with score
Return the score
DKIM signer
(System described in Fig. 1)
Proposed Method
Aggregated e-mail delivery data
(A)
Function to classify each combination of sender's domain and DKIM signature domain by operational situation
(B)
Function to score the combination based on scenarios classified in (B)
(C)

Fig. 3. Configuration of our proposed method.

authentication methods, SPF and DKIM. DMARC provides a mechanism for the sender domain's administrator to declare the policy in the "p=" tag of the DMARC record how the receiver handles the e-mail which fails SPF and/or DKIM verification. DMARC policy has the following three variations, "none (do nothing even if authentication failed)", "quarantine (quarantine the authentication failure e-mail)", and "reject (reject the authentication failure e-mail)". Moreover, DMARC sends the aggregate results of the DMARC verification to the e-mail address shown in "rua=" tag of DMARC record. The sender domain published the DMARC record as a TXT record that "_dmarc." is added to the beginning of own domain.

In addition, DMARC has a concept called "alignment", which means that the e-mail cannot pass the DMARC verification when the DKIM signature domain is different from the sender's RFC5322.From domain. In DKIM, the sender's RFC5322.From domain do not need the same as the DKIM signature domain. On the other hand, spammers can fraud the RFC5322.From address easily. As a countermeasure against the issue, with the concept of alignment, the receiver can check whether the RFC5322.From domain is correct or not.

Although DMARC is an effective method to compensate for the weakness of DKIM, it has a problem with low penetration rate, that is only 29% as of 2016 [4]. As one reason why DMARC does not spread is that many DKIM compliant domains are using third party signature. Actually, as mentioned in subsection II-B, the rate of DKIM signature domain utilizing third party signature is around 80% according to our observation (TABLE I).

D. Conventional DKIM signature verification method

We previously proposed a method for verifying the reliability of DKIM signature by checking the past delivery combination of e-mail sender's domain and DKIM signature domain [5].

Although this method provided a novel criterion to evaluation the reliability of the combination of e-mail sender's RFC5322.From domain and DKIM signature domain, this method only judges whether the combination is exist or not in the dataset including a lot of spamming deliveries data. Therefore, this method unable to evaluate the reliability of DKIM signature with high precision.

TABLE II
E-MAIL DELIVERY EXISTENCE (EXAMPLE)

Combinations of domains		Period					
DKIM domain	From domain	P1	P2	P3	P4	P5	P6
sign.example	example.jp	1	1	1	1	1	1
example.net	example.com	1	0	0	1	1	1

III. Design of our verification system

As described in subsection II-D, conventional DKIM signature verification method has an issue about reliability of determination accuracy. To overcome this issue, we propose a method to score the combinations of sender's domain and DKIM signature domain by checking the fluctuation of existence of e-mail delivery for each combinations obtained from past communication data.

Our approach consists three parts A, B, and C as shown in Fig.3. In this section, we describe the design of these three parts.

A. Aggregating e-mail delivery data

As preparation for scoring the reliability of DKIM signature, our mechanism aggregate the existence of e-mail delivery of several periods for each combinations of e-mail sender's domain and DKIM signature domain as shown in TABLE II.

This table shows whether e-mail was delivered or not for each combination of e-mail sender's domain and DKIM signature domain. In the table, "1" indicates that at least one e-mail was delivered, and "0" indicates that there was no delivery within the period. Our approach utilizes this existence pattern of e-mail deliveries as a basis, and scores by the scoring method described in the next subsections.

B. Classifying the e-mail delivery data

Although it is assumed that legitimate senders occasionally change the DKIM signing domain, they do not frequently change the signer in normal operation. Focusing on such a characteristic of DKIM signature, our method confirms the transition of the combinations of senders domain and DKIM signing domain for several periods, and classifies the delivery data into three assumed operational situation scenarios. With this approach, not only the suspicious combination of these domains but also changes of the DKIM signer can be detected. Thus, our approach can score the reliability of DKIM signature domain with high precision.

Subsequently, we explain about the classification of the operational situation scenarios for each combination of domains. Our method classified the combination data into the following three scenarios.

Scenario 1:
Combination including newly generated domains or new DKIM signature domains due to the migration of the signing.

Scenario 2:
Combination including domain which DKIM signer migration has been completed and is not currently used.

Scenario 3:
Combination of domains other than the above two scenarios.

The domain combinations that have recently started e-mail delivery but were not previously used are classified as Scenario 1. In other words, these combinations are expected to have moved the DKIM signing domain during the communication data collection period used by this method. In order to classify such domain combinations into Scenario 1, it is necessary to determine the length of the period without deliveries observation. If this length of the period setting sets too long, it will be delayed for the system to recognize the DKIM signer domain change. On the other hand, if the period sets too short, the system becomes difficult to distinguish between the change of DKIM signer and cases where there is no delivery by chance within the period. Based on these considerations, we decided to categorize the domain combinations delivered consecutively within the recent three months into Scenario 1.

Secondly, Scenario 2 indicates the opposite situation to Scenario 1. Although the domain combinations were used in the past, the combinations that are not currently used are classified as Scenario 2. In other words, this scenario indicates that the domain of the sender which was using this DKIM signature domain has changed the other signing domain. Our method classifies the domain combinations that are not delivered continuously for the recent three months, and that was delivered in the past, into Scenario 2 for the same reason described in Scenario 1.

Thirdly, Scenario 3 is other than above two scenarios. This scenario includes many cases, for example, case of DKIM-compliant e-mail is not delivered during the period, case of existence of delivery during the period is alternately, etc.

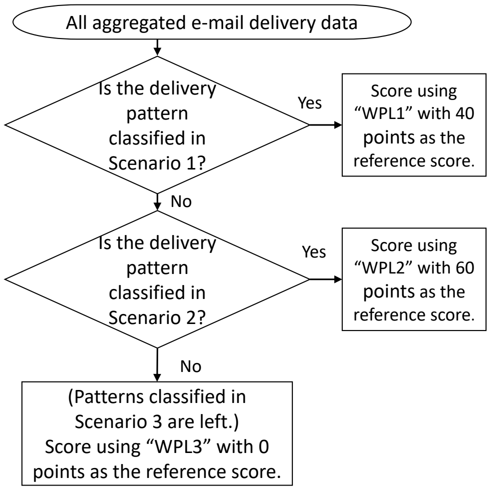

All aggregated e-mail delivery data
Is the delivery pattern classified in Scenario 1?
Yes
Score using "WPL1" with 40 points as the reference score.
No
Is the delivery pattern classified in Scenario 2?
Yes
Score using "WPL2" with 60 points as the reference score.
No
(Patterns classified in Scenario 3 are left.) Score using "WPL3" with 0 points as the reference score.

Fig. 4. Flow of scoring by our method.

C. Scoring method

In order to evaluate the reliability of the DKIM signature taking into account the difference in priority between the newer data and the older data, our scoring method applies different weights to each period of the delivery existence data. Our approach scores the combinations of sender's domain and DKIM signature domain from 0 points to 100 points for each three scenarios described in subsection III-B.

Fig.4 shows the flow of the scoring method. First of all, the system confirms the pattern of e-mail delivery is classified in Scenario 1. If the operational situation of the domain's combination is classified in Scenario 1, the scoring method gives 40 points to these combination as reference score, and adds "Weighted Point List 1" (hereinafter, this word is called "WPL1") to the reference score when e-mails were delivered in each period.

WPL1 is obtained as following. First, the system divides 60 points, which are 40 points (reference score) subtract from 100 points, by the number of periods (N). In other words, the system calculates the average of the score for N periods. Secondly, we set this average value as the middle value of WPL1 in N periods. Finally, the system generates an arithmetic progression having N items with this average value as the central value so that the total of the progression becomes 60 points, and the result of this flow is set as WPL1.

In this pattern, as discussed in subsection III-B, these combinations are expected to have moved the DKIM signing domain. However, it is known that spammers send spoofed e-mail acquiring new domains one after another [11]. Therefore, we define 40 points as the reference score that is a slightly lower score than middle score, and consider that it would be appropriate to gradually increase the score of reliability.

Then, the system confirms whether the pattern of e-mail delivery is classified to Scenario 2 described in subsection III-B. In this case, our mechanism subtracts "Weighted Point List 2" (hereinafter, this word is called "WPL2") from 60 points. WPL2 indicates the points to be subtracted when any

TABLE III
THE NUMBER OF COMBINATIONS
OF RFC5322.FROM DOMAIN AND DKIM SIGNERS DOMAIN

Month	Sep.	Aug.	Jul.	Jun.	May	Apr.
# of combinations	285,916	269,087	244,684	246,019	228,054	225,576

TABLE IV
WEIGHTED POINT LIST 1

Month	Sep.	Aug.	Jul.	Jun.	May	Apr.
Weight	13	12	11	9	8	7

TABLE V
WEIGHTED POINT LIST 2

Month	Sep.	Aug.	Jul.	Jun.	May	Apr.
Weight	7	8	9	11	12	13

TABLE VI
WEIGHTED POINT LIST 3

Month	Sep.	Aug.	Jul.	Jun.	May	Apr.
Weight	25	21	18	15	12	9

e-mails were not delivered in each period. WPL2 is obtained from inverting new and old of WPL1.

In this pattern, as mentioned in subsection III-B, these combinations are expected to have been used in the past, but not used currently. Since this scenario means the domain combinations which are not currently used, this method gradually lower the reliability score after evaluating the past delivery records. Thus, we set score of 60 points, that is a slightly higher score than middle score, as the reference scores.

Finally, in the case of the remaining patterns which classified in Scenario 3 described in subsection III-B, our mechanism adds “Weighted Point List 3” (hereinafter, this word is called “WPL3”) to 0 point of the reference score, and set that result as a score of the reliability.

WPL3 is obtained as following. First, the system divides 100 points, which are maximum score of our method, by N periods. In other words, the system calculates the average of the score for N periods. Secondly, we set this average value as the middle value of WPL3 in N periods. Finally, the system generates an arithmetic progression having N items with this average value as the central value so that the total of the progression becomes 100 points, and the result of this flow is set as WPL3.

In this pattern, unlike the above two patterns, it is difficult to estimate the DKIM signature domains condition. Therefore, the system adopts the simple additional point method without setting a reference score.

IV. EVALUATION

A. Creation of score dataset using actual communication data

We evaluated the proposed system using data for six months from April to September 2016. We set a period as a month in this evaluation, therefore, we divided the data into six periods by month. TABLE III shows the number of combinations of RFC5322.From domain and DKIM signing domain for each month used for this evaluation. As shown in the table, we evaluated using a sufficiently large size of dataset.

Based on the scoring method for each of the three scenarios described in subsection III-C, the weighted points of each month in each scenario to be used for scoring are obtained as shown in TABLE IV, V, and VI.

In this evaluation, when the delivery existence patterns of each month within a period are “100000”, “110000”, and “111000”, these are classified as Scenario 1 mentioned in subsection III-B. In the case of Scenario 1, the reference score is 40 points and the score shown in TABLE IV is added to the reference score. Therefore, when the delivery existence pattern is “100000”, the score is 53 points, similarly, if the pattern is “110000”, the score is 65 points, and the score in the case of “111000” is 76 points.

Secondly, if the delivery existence patterns of each month within a period are “000001”, “000011”, and “000111”, these are classified as Scenario 2. For Scenario 2, the score shown in TABLE V is subtracted from 60 points (reference score). Thus, the score of “000001”, “000011”, and “000111” are 13, 25, and 36 points, respectively.

Finally, the delivery existence patterns other than above two are classified as Scenario 3. In this case, the reference score is 0 point and the score of TABLE VI is added to the reference score. For instance, the score of “101010” is 55 points.

B. Example of scoring result by our method

In order to evaluate the trustworthiness of our proposed mechanism, we applied our method to e-mail deliveries which signed by one of the most famous e-mail service providing companies (hereinafter, this company is called “X”). The list of scoring results is shown in TABLE VII. In order to maintain the anonymity of each RFC5322.From domain, we assigned five symbols for each type of the domain.

The meaning of the symbols of RFC5322.From domain in TABLE VII is described as follows.

- **(A)** X’s own RFC5322.From domain (RFC5322.From domain same as the DKIM signature domain) and sub domains of X.
- **(B)** RFC5322.From domains of Internet Service Providers (ISPs) or Mail Service Providers (MSPs).
- **(C)** RFC5322.From domains of universities or governments.
- **(D)** RFC5322.From domains of companies excluding ISPs and MSPs.
- **(E)** Other than above (Most of these RFC5322.From domains are little known and/or suspicious.)

As shown in TABLE VII, the reliability score of all domains classified as type “A”, which is obviously legitimate domains, was 100 points. In contrast, the reliability score of all domains classified as type “E”, which is extremely low possibility of legitimate deliveries, was 18 points and less.

TABLE VII

THE SCORING RESULT LIST TO THE RFC5322.FROM DOMAINS SIGNED BY "X"

#	RFC5322.From domain	Score	#	RFC5322.From domain	Score	#	RFC5322.From domain	Score	#	RFC5322.From domain	Score
1	C	0	16	B	13	31	B	48	46	A	100
2	C	0	17	C	15	32	C	53	47	A	100
3	B	0	18	C	15	33	B	53	48	A	100
4	B	0	19	D	15	34	B	53	49	A	100
5	D	0	20	D	15	35	D	65	50	A	100
6	C	0	21	B	15	36	C	73	51	A	100
7	C	0	22	B	18	37	D	75	52	A	100
8	C	0	23	E	18	38	B	100	53	A	100
9	E	0	24	B	21	39	A	100	54	A	100
10	E	12	25	C	24	40	A	100	55	A	100
11	D	12	26	C	27	41	A	100	56	A	100
12	E	12	27	C	33	42	A	100	57	A	100
13	E	13	28	D	36	43	A	100	58	D	100
14	D	13	29	B	43	44	A	100	59	D	100
15	B	13	30	C	46	45	A	100	60	A	100

Moreover, the RFC5322.From domains of the group company of X, major ISPs and MSPs, for example a combination of No.31, No.33, and No.38, indicate that were used for forwarded messages. Our system scored higher for these domain as shown in TABLE VII.

On the other hand, the reliability score of RFC5322.From domains which are frequently used for spoofed e-mails for example a combination of No.3, No.4 and No.11, were 12 points and less as shown in TABLE VII.

In ordinary DKIM verification, not only 60 domains listed in TABLE VII but also any e-mail deliveries signed by domain X will be successfully verified. Next, in our conventional DKIM signature verification method described in subsection II-D, any e-mails signed by domain X other than the 60 RFC5322.From domains shown in TABLE VII will fail the verification. However, this conventional method does not consider the difference in reliability of these 60 domains and treats them equally. On the other hand, from the evaluation result described in above, we confirmed that our approach can accurately score the reliability of DKIM signature domains with high accuracy. Therefore, our method can be contributed as one method of countermeasure against spoofed e-mails.

V. CONCLUSION

In this paper, we proposed a countermeasure method against spoofed e-mail that scores the reliability of combinations of sender's RFC5322.From domain and DKIM signature domain by analyzing fluctuation of the existence of e-mail delivery obtained from the past communication data.

As a solution to the problem that DKIM cannot verify the validity of the signer, our method provided a new spoofed e-mail countermeasure method by our own reliability scoring method. By utilizing our mechanism, the recipients can obtain the reliability score for each message including spoofed e-mail which was uniformly verified successfully in the ordinary DKIM verification. In addition, receiving servers can be applied to an anti-spam/anti-spoofed mail system such as whitelist or blacklist using the score, and can be warned by notifying each recipient of the score.

From the evaluation result, we confirmed that our proposed method can provide appropriate score to the domain combination that all passed in conventional DKIM verification, and it can utilize as a novel spoofed e-mail countermeasure mechanism.

ACKNOWLEDGMENT

We would like to thank Mr. Shuji Sakuraba, who is the chairman of anti spam committee of The Internet Association Japan, for providing us communication data for analysis.

REFERENCES

[1] FBI (Federal Bureau of Investigation), "Public Service Announcement, Business E-mail Compromise: The 3.1 billion dollar scam" [online] Available: https://www.ic3.gov/media/2016/160614.aspx, 2016.

[2] D. Crocker, T. Hansen, and M. Kucherawy, "DomainKeys Identified Mail (DKIM) signatures", (STD 76), 2011.

[3] M. Kucherawy, and E. Zwicky, "Domain-based message authentication, reporting, and conformance (DMARC)", (RFC 7489), 2015.

[4] Return Path, "DMARC Intelligence Report" [online] Available: https://returnpath.com/wp-content/uploads/2016/02/DMARCIntelligenceReport_2016.pdf, 2016.

[5] B. Rienthong, N. Kitagawa, and N. Yamai, "Countermeasure of spoofed emails by checking the reliability of DKIM signature," IEICE Tech. Rep., vol.116, no.282, IA2016-49, pp.103-107, 2016.

[6] I.Androutsopoulos, J.Koutsias, K.V.Chandrinos, G.Paliouras, C.D.Spyropoulos, "An evaluation of Naive Bayesian anti-spam filtering," Proceedings of the workshop on Machine Learning in the New Information, pp.9-17, 2000.

[7] I.Androutsopoulos, J.Koutsias, K.V.Chandrinos, C.D.Spyropoulos, "An experimental comparison of naive Bayesian and keyword-based anti-spam filtering with personal e-mail messages," Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval, pp.160- 167, 2000.

[8] "The Apache SpamAssassin Project," [online] Available: http://spamassassin.apache.org/

[9] J.Mason, "Filtering spam with spamassassin," In HEANet Annual Conference, 2002.

[10] M. Wong, and W. Schlitt, "Sender Policy Framework (SPF) for authorizing use of domains in e-mail", (RFC 4408), 2006.

[11] S. Hao, M. Thomas, V. Paxson, N. Feamster, C. Kreibich, C. Grier and S. Hollenbeck, "Understanding the domain registration behavior of spammers," Proceedings of the 2013 ACM Conference on Internet Measurement Conference, pp.63-76, 2013.