Sender Authentication Technology Deployment and Authentication Identifiers

In this report we will present an overview of spam trends for week 14 through week 26 of 2012. The ratio of spam has dropped 2.5% since the previous report, but phishing-based net banking fraud is on the rise. We will take a closer look at the identifiers used in sender authentication technology to prevent these crimes.

2.1 Introduction

In this report we discuss the latest trends in spam and email-related technologies, and summarize various activities in which IIJ is engaged. In this volume we focus on data for the period of 13 weeks from week 14 of 2012 (April 2 to April 8, 2012) to week 26 (June 25 to July 1, 2012), which corresponds to the 1st quarter for many Japanese companies.

In "Trends in Email Technologies," we examine the deployment status of sender authentication technology using data from IIJ's email services and other survey results. We will also continue on from our previous technical discussion of DMARC, and look at future issues.

2.2 Spam Trends

In this section, we will report on spam trends, focusing on historical ratios of spam detected by the Spam Filter provided through IIJ's email services and the results of our analysis concerning spam sources.

2.2.1 Spam Ratios Lower but Threat Increasing

Figure 1 shows spam ratio trends over the period of one year and three months (65 weeks), including the current survey period and the same period for the previous year. The average spam ratio for the current survey period was 44.7%. This is a slight drop of 2.5% compared to the previous report (Vol.15). It also represents a 5.5% decrease from the average for the same period the previous year (Vol.12). This demonstrates that the actual volume of spam has decreased significantly since the first IIR was published in 2008. However, according to information*1 released by the National Police Agency last year, there has been an increase in the number of phishing incidents in which emails fraudulently posing as financial institutions are sent to induce recipients to input their ID and password into fake internet banking sites. The resulting financial damages reportedly amounted to approximately 300 million yen in illegal remittances over a period of about eight months from late March last

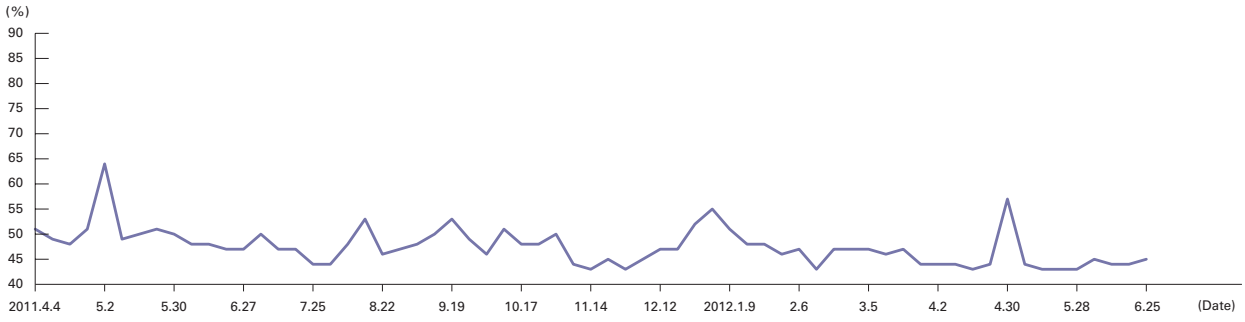

Figure 1: Spam Ratio Trends

*1 "Status of Violations of the Unauthorized Computer Access Law related to Internet Banking" (http://www.npa.go.jp/cyber/warning/h23/111215_1.pdf) (in Japanese).

year. Recent media reports also indicate that net banking fraud has been on the rise again since June of this year. It seems that phishing techniques that led users to fraudulent sites listed in emails were also used in these incidents. To prevent this fraud, it is necessary to first check the sender authentication results and authenticated domain name when mail is received, and determine whether or not the mail can be trusted.

2.2.2 Japan Once Again No.2 Source

Figure 2 shows our analysis of regional sources of spam over the period studied. China (CN) was once again the number one source of spam in this survey, accounting for 20.7% of total spam. China has been the top source of spam in Japan for the past six quarters in a row. Japan (JP) was 2nd at 13.2%. Though in the previous report (Vol.15) Japan dropped to 3rd place, it has now climbed back to 2nd. The United States (US) was 3rd at 11.5%, trading places with Japan. The Philippines (PH) was 4th at 9%, maintaining its high ratio from the previous report. These four regions totaled 54.4%, accounting for over half of all spam sent. India (IN) was 5th at 4.5%, Vietnam (NV) 6th at 3.3%, and Hong Kong (HK) 7th with the same 3.3% ratio as 6th place.

2.3 Trends in Email Technologies

Here we will examine a variety of technological trends relating to email. In this report we look at the deployment status of sender authentication technology by also referencing data for services other than IIJ's email services. We also continue our discussion of the new DMARC*2 message authentication mechanism that uses the SPF*3 and DKIM*4 sender authentication technologies.

2.3.1 Deployment Status of Sender Authentication Technology on IIJ Services

Figure 3 shows SPF authentication result ratios for email received during the current survey period (April to June 2012). The ratio of authentication results showing "none," indicating that the sender domain has not implemented SPF (no SPF record declared), was 33.8%. This is a 2.7% decrease from the last survey, indicating a corresponding increase in the ratio of email that could be authenticated. In other words, the sender SPF deployment ratio for mail received increased to approximately 66.2% in the current survey period.

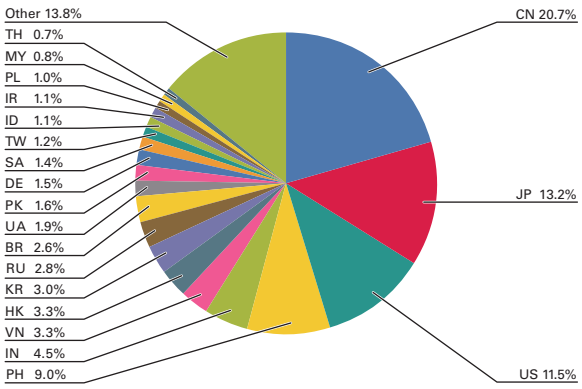

Figure 2: Regional Sources of Spam

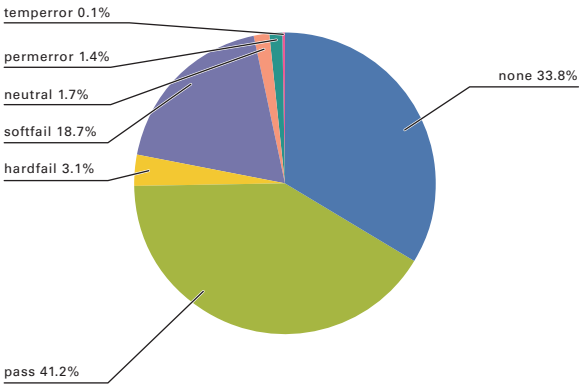

Figure 3: SPF Authentication Result Ratios

*2 DMARC: Domain-based Message Authentication, Reporting & Conformance.
*3 SPF: Sender Policy Framework, RFC4408.
*4 DKIM: DomainKeys Identified Mail (DKIM) Signatures, RFC6376.

Figure 4 shows trends in this deployment ratio for the period of 35 months between August 2009 and June 2012. Although there are some fluctuations the ratio is increasing steadily, demonstrating that deployment is progressing at a good rate. The most recent data is June 2012 at 69.6%, which is an increase of about 27% over the figure for August 2009 when the survey started.

2.3.2 Deployment Status of Sender Authentication Technology According to the WIDE Project

The WIDE project*5, through a collaborative research contract with JPRS*6, has measured the deployment ratio of sender authentication technology since April 2005. Surveys were conducted monthly until May 2011, and since then measurement has been carried out biannually. Figure 5 shows trends in these measurement results for each domain registration type. As of the latest data for May 2012, the average SPF implementation ratio was 43.9%.

The gap between this and the 66.2% ratio for IIJ email services comes down to fact that the WIDE project only covers "jp" domains managed by JPRS, and surveys SPF deployment ratios (the number of domains declaring an SPF record, to be exact) for registered domains. Put more simply, this is due to the difference between static measurement, and dynamic measurement of actual mail volume.

Even more noteworthy in Figure 5 is that implementation rates are rising sharply for "go.jp" domains used by government institutions. The latest measurement is approximately 114%. This value may seem odd, but it is not an error. This is because the ratio of domains declaring an SPF record is measured using the domains utilized for mail as a parameter. For "go.jp" domains an SPF record such as the one below is declared even if they are not utilized for mail, to prevent them from being exploited in email spoofing.

```
example.jp  TXT  "v=spf1 -all"
```

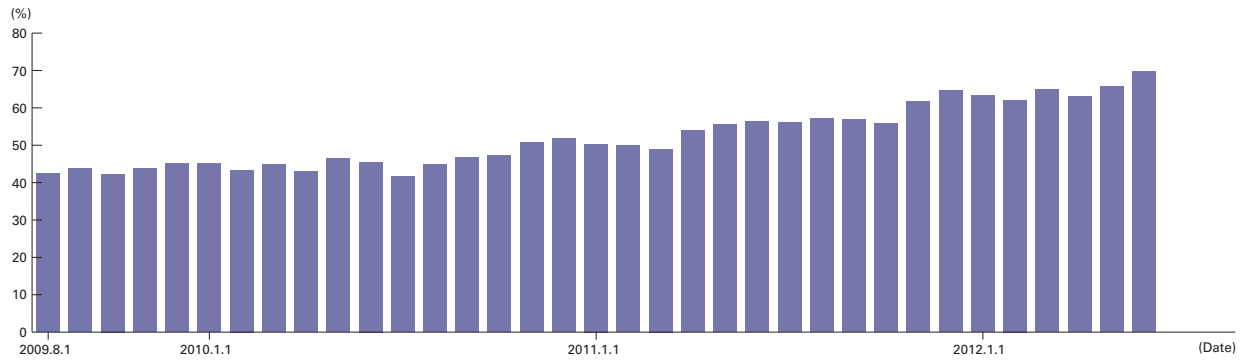

Figure 4: Trends in SPF Implementation Ratios

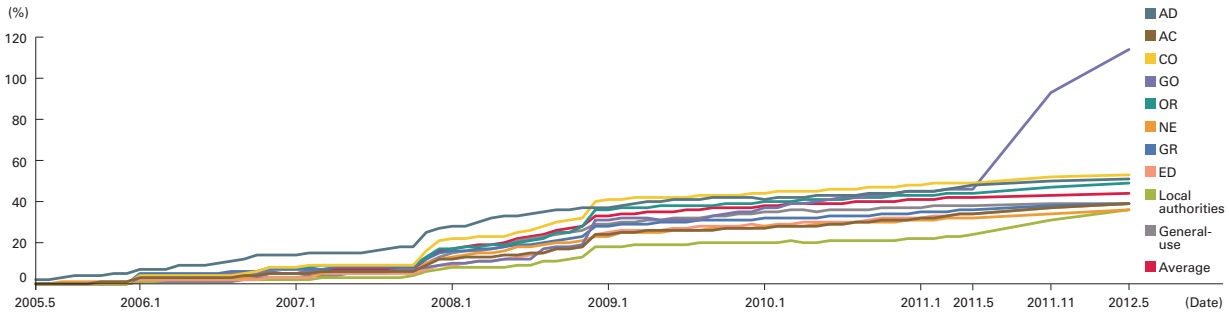

Figure 5: Trends in SPF Deployment Ratios for each JP Domain Registration Type

*5 WIDE: Widely Integrated Distributed Environment (http://www.wide.ad.jp/).
*6 JPRS: Japan Registry Services (http://jprs.co.jp/en/).

This SPF record does not specify information for a legitimate mail server, so all match the "-all" mechanism. The "-" is a qualifier for matches, so the authentication result will always be either fail or hardfail. In other words, the SPF record will cause authentication to inevitably fail when this domain name is used. This means that even if a third party appropriates the sender information for the corresponding domain, recipients can detect it as fraudulent.

2.4 The Relationship Between DMARC and Authentication Identifiers

In our previous report we gave an overview of the purpose and mechanism of DMARC. This time we discuss the identifiers used in authentication, including those for existing sender authentication technologies. DMARC is compatible with the SPF and DKIM sender authentication technologies. The SPF authentication identifier is the reverse-path (RFC5321.MailFrom) domain used as sender information for mail delivery*7. This sender information is submitted to the receiving MTA before the message body itself is received, and is not usually presented to the mail recipient. Consequently, it may not be possible to trust the sender information displayed in MUA (email software) or mobile phones, even if mail has been authenticated using SPF. For this reason, it is necessary to check the domain listed in the header*8 that indicates the authentication results. DKIM authenticates the SDID (Signing Domain Identifier) indicating the signer. This is the identifier after the "d=" parameter in the "DKIM-Signature" header for signature information. Basically, this means that DKIM is actually technology for authenticating the signer who created the signature data, rather than the mail sender. That is why DKIM ADSP*9 was created to indicate signing practices for sent mail and guidelines for when the desired authentication is not possible on the recipient side. See IIR Vol.6*10 for more information on DKIM ADSP. Table 1 below shows each of the authentication identifiers, including for DMARC. We can see that, just for those shown here, four authentication technologies and three different identifiers are used in mail authentication.

2.4.1 DMARC Identifier Alignment

The identifier used for DMARC is Identifier Alignment. Basically, this is an identifier authenticated using either SPF or DKIM, which has also been verified as the same as or closely related to the RFC5322.From domain in the sender information of the mail headers. This means use of DMARC requires an identifier authenticated with either SPF or DKIM, and if either of these authentication processes fail or cannot be carried out, it is also treated as a failure under the DMARC mechanism. Next, we will examine the relationship between the identifier authenticated using SPF or DKIM, and the RFC5322.From identifier. If both domains are the same, there is no problem. However, in cases such as large organizations that frequently use subdomains, or when mail delivery is sorted by purpose using subdomains, it can be more convenient to consolidate email on a representative domain due to the difficulty of setting up a DKIM signature or preparing DMARC records for each subdomain. In DMARC records it is possible to specify whether to be strict or relaxed with authenticated identifiers for SPF and DKIM individually to determine Identifier Alignment. When strict alignment is used, the authenticated identifier domain must match the RFC5322.From domain exactly. When relaxed alignment is used, only the organization domain for the authenticated identifier must match. The organizational domain is conceptually the top domain or parent domain. There are a number of different patterns for domain names, such as when a TLD (Top Level Domain) like ".com" or ".org" is used, or when there is an attribute indicating domain type before the ccTLD as with Japan (.jp). That means it is difficult to accurately define an organizational domain, and this is an issue that needs to be examined in the future. Table 2 provides some examples of organizational domains.

Authentication Technology	Authentication Identifier
SPF	reverse-path (envelope from, RFC5321.MailFrom)
DKIM	SDID (d=)
DKIM ADSP	Author Domain (RFC5322.From)
DMARC	Identifier Alignment (RFC5322.From)

Table 1: Relationship Between Authentication Technologies and Authentication Identifiers

Authentication Domain	Organizational Domain
foo.example.com	example.com
foo.bar.example.co.jp	example.co.jp

Table 2: Organizational Domain Examples

*7 The system for mail delivery is defined in SMTP (Simple Mail Transfer Protocol, RFC5321).
*8 Message Header Field for Indicating Message Authentication Status, RFC5451.
*9 DomainKeys Identified Mail (DKIM) Author Domain Signing Practices (ADSP), RFC5617.
*10 IIR Vol.6 "2.3 Trends in Email Technologies" (http://www.iij.ad.jp/en/company/development/iir/pdf/iir_vol06_EN.pdf).

2.5 DMARC Issues

DMARC makes it possible to consolidate authentication domains with a certain degree of flexibility by providing the option of setting the alignment mode. This means that if you organize domain transfers suitably using subdomains, the DMARC system can be employed even if you outsource mail delivery or use hosting services. It also means that, as long as SPF or DKIM authentication is successful, that can be used as an identifier. This makes it effective in situations where, as with mail forwarding, DKIM authentication is possible even if SPF authentication fails.

However, DMARC is not compatible with mailing lists, which is currently a common form of mail use. Table 3 shows some common examples of the relationship between authentication indicators and authentication results for mailing lists.

SPF and DKIM authentication is successful for mailing lists when both are authenticating the domain for the mailing list's administration address. However, for mailing lists the RFC5322.From domain is typically the domain for the person who first posted in a list thread. This causes authentication to fail as DMARC Identifier Alignment is not possible. This issue is being discussed at DMARC.org, but there is unfortunately no clear-cut solution in sight.

Authentication Technology	Authentication Result	Authentication Identifier
SPF	pass	Mailing list administration domain
DKIM (no re-sign)	fail *11	Mail author
DKIM (re-sign)	pass	Mailing list administration domain
DMARC	fail	Mail author

Table 3: Relationship Between Authentication Identifiers and Authentication Results for Mailing Lists

2.6 Conclusion

Many engineers involved in mail tend to believe it preferable to pass on mail that is received via the mail delivery process without changing it if possible. This also applies to mail headers, and until now even if a "Received" header was added during mail delivery, the existing header was normally not deleted. Functionality has also been expanded for new technologies such as DKIM by adding a new "DKIM-Signature" header. For this reason mail headers have gradually become larger, while the process of viewing the full mail header has become more difficult due to MUA only displaying the minimum necessary information. Mail headers have a comparatively flexible structure, which makes them easy to use when adding functions in this way. On the other hand, this makes it slightly harder to properly analyze and change existing headers. For mail delivery systems, rather than endlessly thinking up ways to preserve data that is ultimately not referenced, at some point it may be necessary to perform a thorough overhaul, while maintaining proper balance as we have done until now.

Author:
Shuji Sakuraba
Mr. Sakuraba is a Senior Engineer in the Strategic Development Center at the Application Development Department of the IIJ Product Division. He is engaged in the research and development of messaging systems. He is also involved in various activities in collaboration with external related organizations for securing a comfortable messaging environment. He is a M3AAWG member and JEAG board member. He is a member of the Anti-Spam mail Promotion Council (ASPC) and administrative group, as well as chief examiner for the Sender Authentication Technology Workgroup. He is also a member of Internet Association Japan's Anti-Spam Measures Committee. He is a member of the Ministry of Internal Affairs and Communications' Unsolicited Mail Measure Working Group.

*11 When the headers or body text used in digital signatures are not altered, authentication may be successful in some cases.