Public Sender Score System(S3) by ESPs for Email Spam Mitigation with Score Management in Mobile Application

https://doi.org/10.3991/ijim.v14i17.16609

Lucky Kannan (✉), Jebakumar R
SRM Institute of Science and Technology, Chennai, India
lucky.de.knite@gmail.com

Abstract—Many businesses use email as a medium for advertising and they use emails to communicate with their customers. In the email world, the most common issue that remains unresolved even now is spamming or in other terms unsolicited bulk email. Currently, there is no common way to regulate the practices of an email sender. This proposed system is to formulate a protocol common for all the ESPs or inbox providers and a centralized system that will easily find the spammers and block them. By this method, the Email Service Providers (ESPs) or Inbox Providers need not wait for the sender behaviour and then take actions on the sender or sender domain or sender IP address. Instead, they can get the sender history of reputation from blockchain where the ESPs or Inbox Provider provides a score based on the emails they have received from the sender. The ESPs can get the Public Sender Score(S3) from the mobile application or web application which provides the score management user interface and APIs. The email marketers can also monitor their score through the application.

Keywords—Email, spam detection, sender score, reputation system, blockchain, mobile application, marketing email.

1 Introduction

The email industry has been functioning steadily for very long and they needed minimal or no change over the past years. There is a clear vision on the existence and role of email in the future. The email medium provides a means of formal communication with lower cost and faster than many other mediums present. However, there is an important issue of spamming that needs to be completely addressed yet. Email spamming is a problem with the user consent and not email content. Whether the Unsolicited Bulk Email message is an advertisement, a scam, a product offer, or pornography, the content is not as much relevant as the consent of the sender. If the email is being sent unsolicited and in bulk, then the email can be classified as spam. Most of the emails are sent unsystematically to unnecessary recipients that might be from improperly obtained or maintained lists of recipients or from bad senders. Most of the recipients are flooded with unnecessary promotional emails which is a nuisance

to the recipients and also to the recipient email system as the disk storage space, computational resources and network bandwidth of the recipient email systems is consumed. Events like this can lead to the annoyance of the recipients and they might permanently deny to accept these emails and as a result, the credibility of the business of the sender may get destroyed. So, the violators of the basic rules of marketing or newsletter have to be identified and make them stop their approach. Most of the spam identification mechanisms like Machine Learning are done by the email providers as in [1] after the mail is received by them which is found to be very effective. There are clickbait detection mechanisms like [2] that identifies spammers too. Some content differs from the subject and the spammers insert spam content in between the email content which can be identified by mechanisms like [3]. There are a lot of other mechanisms in place by the email providers in order to identify spam in the email content. This paper aims a different approach to stop the email spams at the origin even before the emails are sent out to the recipients.

In this paper, we propose a public Sender Score System (S3) that maintains the sender's reputation by a score which is updated by the inbox providers periodically. The public server is either a decentralized blockchain or centralised server with an authority. A sender score updation protocol has to be created which will be used by the inbox providers who will be updating or querying the score in the Sender Score System (S3). They are also used by the promotional email senders who will be querying the score to know their own reputation.

In Section II, the background information, the previous works to classify spam emails, and the sender verification processes and sender scores are discussed. Section III explains the research design and process diagram of the proposed public sender score system with explanation. Section IV gives conclusion of the proposed design. Finally, Section V presents future directions of improving the sender score protocol and management of the centralised or decentralised public sender score system.

2 Related Works

In this section, we will discuss the previous works on giving weightage for the email sender and calculating credibility of the emails using the sender information.

2.1 Sender email authentication

In order to have better inbox reach, the senders have recently adopted all the authentication mechanisms that were proposed which directly affects the authenticity and credibility of the emails sent by them. One of the mechanisms which is adopted worldwide is Sender Policy Framework (SPF) [4] which prevents impostors by blocking the email if the SPF check fails. SPF allows the domain owners to specify who can send emails from their domain which would reduce the risk of their domain being compromised.

Another mechanism is Domain Keys Identified Mail (DKIM) Signatures [5] through which the person that owns the signing domain can claim responsibility for a message

by associating the domain with the message. Using DKIM signature the recipient system can verify that the email is not tampered by checking the sign with the public key in the DKIM record of the domain and the content of the email. The recipient system can also identify the responsible domain that signed the email.

Another standardised mechanism called Domain-based Message Authentication, Reporting, and Conformance (DMARC) [6] is extremely powerful as a tool to stop email spoofing. DMARC policies can be set on who can send email for your domain based on DKIM and SPF. Along with SPF and DKIM, the DMARC policy in DNS allows the domain owner to set rules to reject or quarantine (junk folder) or do nothing to the emails from unknown sources.

2.2 Sender reputation

There are various methods that have been implemented over the years which are used to identify the bad senders. The sender domains and IP addresses have a sender reputation using which we can guess the quality of the emails from them.

a) **DMARC:** There are various ways in which the sender is notified about their reputation or email deliverability. Domain-based Message Authentication, Reporting, and Conformance (DMARC) method, which was discussed before, also allows the sender to get reports of their email delivery through an email which is specified in the DMARC DNS TXT record [3].
b) **Feedback Loops:** There are Feedback Loops (FBL) [7] [8] [9] provided by some Mailbox Providers which will be useful for the marketing Email Service Providers. A large volume sender can use the FeedBack Loop (FBL) to identify email campaigns in its traffic that are getting a high volume of complaints from users of the Mailbox Providers. The FBL is useful to ESPs to detect abuse of their services.
c) **Reputation System**: As Taylor [10] pointed out in the paper, there are reputation systems with the Mailbox Providers that calculate the reputation of the sender and regulate the emails by classifying bad emails under spam or drop them. The systems use the connecting IP address to represent the sender. They use the sender score and also send the message to a statistical spam filter and finally make a judgement on the emails.
d) **Postmaster Tools:** Postmaster Tools [11] are provided by the Mailbox Provided for the bulk email senders to check the deliverability of their emails and point out the issues in their problems in their deliverability. The tool is used to see if users are marking sender emails as spam, to see the prevention method for sender's emails from being blocked, to see the reason for sender's emails not being delivered, and to see if the emails are sent securely.
e) **Domain Blacklist, Whitelist and History**: Besides these different techniques, there are several other methods to check by considering their sending histories. Domain Name System White List (DNSWL) [12] is a central database that stores the senders with better sending history. Domain Name System Blacklist (DNSBL) [13] [14] is a central database which stores the senders list with bad sending history. WHOIS protocol [14] [15] is used to search for the Domain Name System (DNS) and Name

Server (NS) information of the sender. It provides more information about the age of the sender domain which helps in finding the credibility of the message sent by the sender.

All of this and many other methods [16-18] are present which influences the reputation of the sender which in turn affects inbox delivery in the recipient system. But in all cases, the sender reputation is maintained separately in each recipient system and there is no way to maintain a common reputation for the bulk sending system and their users yet.

3 Proposed System

There is a problem of blocking the bad users or sender from blocking as they might keep changing the sending service, when the sending service blocks them. So, they keep sending spam messages through a different service by switching instead of correcting their sending methods. This event is shown as a diagram in Fig. 1

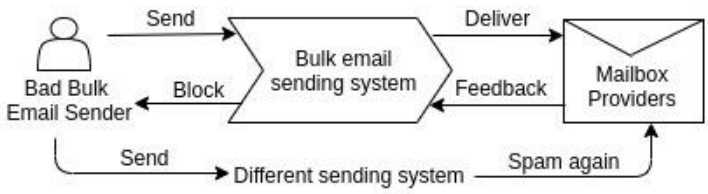

Fig. 1. Bad senders switching Bulk Sending System

By using the proposed system, this can be avoided by using the Public Sender Score System (S3). In this system, the senders of unsolicited bulk emails are blocked in the sending system instead of sending it to the recipients. The sending system checks S3 (Public Sender Score System) for a Sender Score and a sender score history, and if the sender score does not seem to be proper, then they will block the user from sending emails, until best practices are followed. By this way, the reputation of Sending System is not affected and the Recipient or Inbox Providing System does not receive the unsolicited emails. Even if the Inbox Providing system receives email, it can check the Sender Score in the Public Sender Score System (S3) and take required action on the email. This process is shown in the following illustration in Fig. 2

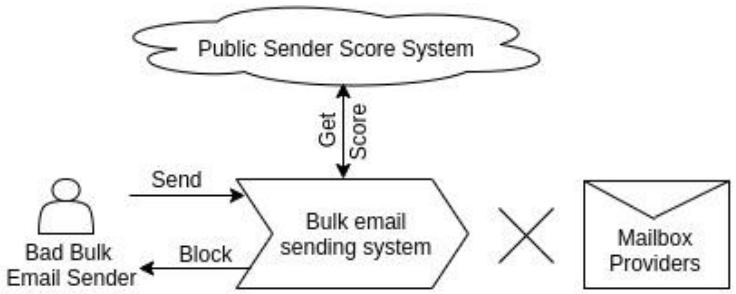

Fig. 2. Bad users blocked by Bulk Sending System

Here, the Public Sender Score System, shortly called as S3, can be either a centralised server or a decentralised blockchain. For this paper, in most of the cases, we will assume that S3 (Public Sender Score System) is a decentralised blockchain implementation which can be built using Ethereum [19]. There can be two blockchains in this architecture in which one is maintained by the mailbox providers and the other is maintained by the bulk email sending services. For this paper, in most of the cases, we will assume that there is one blockchain maintained by the email sending services.

Fig. 3 shows the common architecture of the proposed system with all the different elements involved in the system. It shows how the Public Sender Score System is used along with the other elements in email sending. It also shows where the Sender Score is updated and where the score is retrieved in the email sending ecosystem. Let us discuss the elements in the architecture.

3.1 Marketer or user

Marketer or user is the origin of the email. Marketer has the bulk list of emails to whom they send the emails. So, their reputation is directly proportional to the credibility of the emails. They should be solely responsible for any impact on the domain reputation, either positive or negative.

3.2 Bulk email service provider

Bulk Email Service Provider is an application or a service or a platform that helps the users to send out emails with ease. They would also make sure the right emails are sent to the right recipients. One way to make sure is through the feedback loops and postmaster tool provided by the recipient system. So, by the time the bad sender is identified, all the emails are probably sent and there is no reverting back. By the method of the Public Sender Score system, the sender's score and history is already present in S3 and the email providers are free to take a step on the potential spammer based on the score. By this way, the reputation of the service is saved. In case of a user allowed to send email, the bulk email service providers are free to update a new score to the S3 network based on the feedback from the recipient mailbox provider systems.

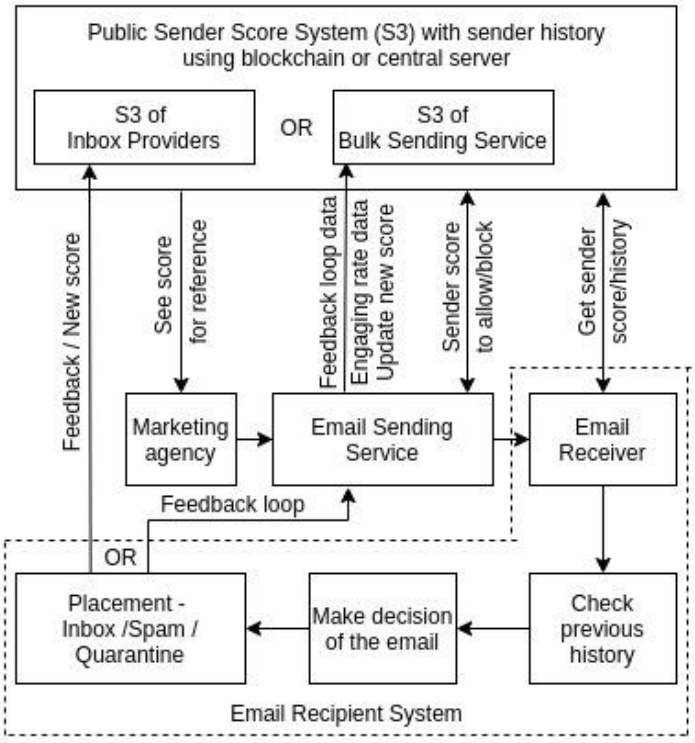

Fig. 3. Architecture of Proposed System

3.3 Email recipient system

Email Recipient System is the system that accepts the emails and puts them in the mailbox of the recipients. It makes the decision of keeping the incoming email, or quarantine email or discard email. They maintain the reputation for the incoming traffic which is notified through the feedback loop or postmaster tool to the senders. In addition to that, the recipient system can update the S3 system which contributes to taking action against users in a robust manner.

3.4 Public sender score system

Public Sender Score System is the system designed to hold the reputation score of the sender in a decentralised system called blockchain which can be implemented using Ethereum. The blockchain contains the updated sender score of the bulk email sending domains. The sender score is calculated by the bulk email senders on their own along with the help of the postmaster tools and feedback loops. New blocks are added with the added or subtracted sender score when the email senders find any steep change in the behaviour of their campaign. Those who have rights to modify the blockchain network are called updaters.

3.5 Authentication to public network

For now, the new authentication of a service to the public network is by the fellow bulk email sending services who form an association to maintain authentication. They will provide a separate authentication key which is used in making changes in the public network. In order to add a new block, the service needs its own authentication and the authentication of the domain shared by the user.

3.6 Transaction in public network

Transaction in the public blockchain network represents a change in the sender score made by the service in the blockchain. The transaction details are as in Fig. 4.

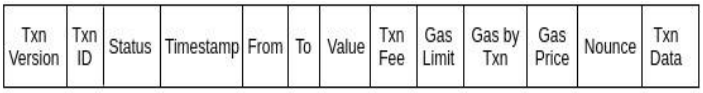

Fig. 4. Transaction details

Here the transaction version is the version of the syntax of transaction details. Txn ID uniquely represents the transaction. Block address or hash can also be used instead of Txn ID. The Status component says the status of the transaction, whether it is completed or not. The timestamp has the time of execution of transaction. The From is the update who updates sender score. The to is the user whose sender score is updated. The value contains the sender score to be updated. The Txn Data contains the change in the Sender Score and Sender Integrity Integer which will be discussed later. The Txn Fee and Gas are pertained to Ethereum. Txn Fee is the charge for the transaction and the Gas is the work that is done for a transaction to be updated. This transaction detail is added in a block and the block is added to the blockchain by a miner who gets awarded for the work.

3.7 Sender score policies

The bulk users or senders are given a sender score on their domain which is on a scale of 0 to 100 which interprets the reputation of the sender. For the bulk email service, there needs to be reputation maintained for the domain as well as the IP address. The measurement of the scale is as follows:

a) 100 - Entirely whitelisted. Sender is trusted blindly and any emails sent by the user is legitimate.
b) > 60 - Good sending practices
c) 60 (50 + 10) - Neutral sender with all authentication mechanisms in place
d) 50 - Neutral sender
e) < 50 - Bad sending practices
f) 0 - Entirely blacklisted. Sender is blocked blindly and any emails sent by the user is dropped.

There is a Sender Integrity Integer in the block that is present in the blockchain. Each of the binary bits in the Sender Integrity Integer is a flag that represents either true or false. Each digit represents as in the following:

1) Transport Security Layer
2) Sender Policy Framework
3) Domain Keys Identified Mail
4) Domain-based Message Authentication, Reporting, and Conformance
5) Age of domain is more than 5 years

For conformance of each of the above integrity, a score of 2 will be allowed to be added and so a total of 10 can be added in case all the bits are set and the integrity value is 31. The new undefined bits can be used to represent any other factor that can affect the reputation. Other than the integrity score, the updater can update a maximum score of +2 or -2 from the current score 3 times (3 blocks) for a sender in a period of 6 months.

4 Conclusion

The users are at a disposal of using the data updated in the S3 network for their own good and there are a lot of advantages in updating and using the data. The marketers or the users use them to identify their sending impact and give an insight about their email campaigns. They can use it to change their sending practices. The public score automatically urges the user to change the sending practice and makes them work towards improving reputation. The Bulk Email Sending services can use the data to assign the correct IP address or IP address pool in their services. They can even restrict the bad users to send minimum loads until their reputation increases and then increase their email load gradually. The S3 data can be used by the inbox provider to make a decision on landing the email in the right place. So, by this method, instead of worrying about how to identify the spam after sending, the unwanted senders are stopped even before sending unsolicited emails. Fig. 5 shows the impact of reduced bad sender practices with increase in the S3 implementation.

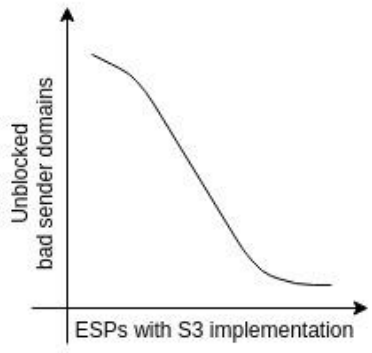

Fig. 5. Usage and impact

5 Future Work

S3 architecture has to be expanded and deeply designed to solve real-time problems of updating and retrieving the sender score. The scalability and performance has to be analysed and improved. A better authentication mechanism has to be in place to make only the authorised service update the related blocks in the Public Sender Score System (S3).

6 Acknowledgement

I am grateful for the enormous emotional support from my family and my fellow classmates in completing the research paper.

7 References

[1] Mohammed Almseidin, AlMaha Abu Zuraiq, Mouhammd Al-kasassbeh, "Phishing Detection Based on Machine Learning and Feature Selection Methods", International Journal of Interactive Mobile Technologies iJIM. Available: https://doi.org/10.3991/ijim.v13i12.11411

[2] Daoud M. Daoud, M. Samir Abou El-Seoud, "An Effective Approach for Clickbait Detection Based on Supervised Machine Learning Technique", International Journal of Online and Biomedical Engineering iJOE. Available: https://doi.org/10.3991/ijoe.v15i03.9843

[3] Mashail Shaeel Althabiti, Manal Abdullah, "CDDM: Concept Drift Detection Model for Data Stream", International Journal of Interactive Mobile Technologies iJIM. Available: https://doi.org/10.3991/ijim.v14i10.14803

[4] S. Kitterman, "Sender Policy Framework (SPF) for Authorizing Use of Domains in Email," Version 1, IETF RFC- 7208, April 2014. [Online]. Available: https://doi.org/10.17487/rfc7208

[5] D. Crocker, T. Hansen, M. Kucherawy, "DomainKeys Identified Mail (DKIM) Signatures," IETF RFC-6376, September 2011. [Online]. Available: https://doi.org/10.17487/rfc6376

[6] M. Kucherawy, E. Zwicky, "Domain-based Message Authentication, Reporting, and Conformance (DMARC)," IETF RFC-7489, March 2015. [Online]. Available: https://doi.org/10.17487/rfc7489

[7] Google, "Feedback Loop (FBL)" [Online]. Available: https://support.google.com/mail/answer/6254652?hl=en, [Accessed: April, 08, 2020]

[8] J. Falk, "Complaint Feedback Loop Operational Recommendations," IETF, RFC - 6449, November 2011. [Online]. Available: https://tools.ietf.org/html/rfc6449

[9] J. Falk, M. Kucherawy, "Battling spam: The evolution of mail feedback loops," Internet Computing, IEEE, vol. 14, no. 6, pp. 68–71, Nov 2010 https://doi.org/10.1109/mic.2010.133

[10] B. Taylor, "Sender reputation in a large webmail service," in CEAS, 2006.

[11] Google, "Postmaster Tool", [Online]. Available: https://support.google.com/mail/answer/6227174?hl=en&ref_topic=6259779, [Accessed: April, 10, 2020]

[12] J. Levine, "DNS Blacklists and Whitelists," IRTF, RFC 5782, February 2010. [Online]. Available: https://tools.ietf.org/html/rfc5782

[13] T. Sochor and R. Farana, "Improving efficiency of e-mail communication via spam elimination using blacklisting," in Telecommunications Forum (TELFOR), Nov 2013, pp. 924–927. https://doi.org/10.1109/telfor.2013.6716382

[14] L. Daigle, "WHOIS Protocol Specification," IETF, RFC- 3912, September 2004. [Online]. Available: https://tools.ietf.org/html/rfc3912.
[15] A. Newton, "Replacing the WHOIS protocol: Iris and the IETF's CRISP working group," Internet Computing, IEEE, vol. 10, no. 4, pp. 79–84, July 2006. https://doi.org/10.1109/mic.2006.86
[16] Upasana and S. Chakravarty, "A survey on text classification techniques for e-mail filtering," in Machine Learning and Computing (ICMLC), 2010 Second International Conference on February 2010, pp. 32–36 https://doi.org/10.1109/icmlc.2010.61
[17] Hang Hu, Peng Peng and Gang Wang, "Towards Understanding the Adoption of Anti-Spoofing Protocols in Email Systems", 2018 IEEE Cybersecurity Development (SecDev) 30 September - 2 October 2018 https://doi.org/10.1109/secdev.2018.00020
[18] Holly Esquivel, Aditya Akella and Tatsuya Mori, "On the Effectiveness of IP Reputation for Spam Filtering", 2010 Second International Conference on COMmunication Systems and NETworks (COMSNETS 2010), 5-9 Jan 2010. https://doi.org/10.1109/comsnets.2010.5431981
[19] "Ethereum," [Online]. https://www.ethereum.org/, [Accessed: March 29, 2020].
[20] *Andreas Solias, Agisilaos Chaldogeridis, Areti Batzikosta, Magdalini Tsolaki. (2020).* Tablet-Administered Screening Tests for the Detection of Major and Mild Cognitive Disorders – Preliminary Findings of a Comparative Study, International Journal of Interactive Mobile Technologies, 14(11), 200-223. https://doi.org/10.3991/ijim.v14i11.14629
[21] Wongkhamdi, T., Cooharojananone, N., & Khlaisang, J. (2020). E-Commerce Competence Assessment Mobile Application Development for SMEs in Thailand. International Journal of Interactive Mobile Technologies, 14(11), 48-75. https://doi.org/10.3991/ijim.v14i11.11358
[22] Mada' Abdel Jawad, Saeed Salah, Raid Zaghal," DSDV Extension to Enhance the Performance of Ad Hoc Networks in High Diverse-Velocity Environments". International Journal of Interactive Mobile Technologies (iJIM), Vol:14 No.06, April 2020. https://doi.org/10.3991/ijim.v14i06.11889

8 Authors

Lucky K is pursuing MTech in the field of Computer Science and Engineering in SRM Institute of Science and Tech, Chennai and has completed B.E., in Government College of Engineering, Salem. He is currently working as a Senior Software Engineer at Zoho Corp, Chennai. He is a Deliverability Engineer, and a developer in Email Marketing Campaign product. He has also developed the products Message Transfer Agent (MTA) and Email Validator, an email address cleanup service. Email: lucky.de.knite@gmail.com

Dr. R. Jebakumar M.E., PhD., is working as an Associate Professor in the Department of Computer Science and Engineering SRM Institute of Science and Technology, Kattankulathur, Chennai since June 2006. In 2015, he received Ph.D. in Information and Communication Engineering, in Anna University. Email: jebakumr@srmist.edu.in

Article submitted 2020-06-25. Resubmitted 2020-07-30. Final acceptance 2020-07-30. Final version published as submitted by the authors.