Journal of Al-Qadisiyah for Computer Science and Mathematics
Vol. 3 No.1 Year 2011

Page 210- 220

Build a Spam Filter at the E-mail on the Server Side

Ahmed Abdulrudah abbas
Kufa University
Education College for girls
Computer Department

Abstract

The spam has now become a significant security issue and a massive drain on financial resources. In this paper, We present a spam filter, which works at the server level. The proposed filter is a combination of antispam solution. The task of proposed filter is to minimize the ability of the spammers to distract the network by the spams. That is done by blocking the spam message at the server level. A server based solution is normally more advantageous than protecting e-mail users individually. Such a solution gives more control to administrators.

1 - Introduction

A Spam mail is an Internet mail that is sent to a group of recipients who have not requested it [1]. The spam or unsolicited mails have already caused many problems such as filling mailboxes, engulfing important personal mail, wasting network bandwidth, consuming users' time and energy to sort through it, and crashing mail-servers.

People who send spams, called spammers, don't do it for fun, but to earn money. Spamming is an actual business process, and like most businesses, the goal is to make a profit. Spammers profit when the spam they send equals the revenue from sales generated from a spam campaign, less the cost of sending spam [2].

The spammers have several features to hide their tracks. Most bulk mailers do not use the mail server of their ISP, but instead connect to the destination mail server directly or use a so-called open relay. This way, the spammer avoids to be detected by his ISP [3].

Any one when he see this statistics of the spam in the figure (1) he will feel the danger of the spam

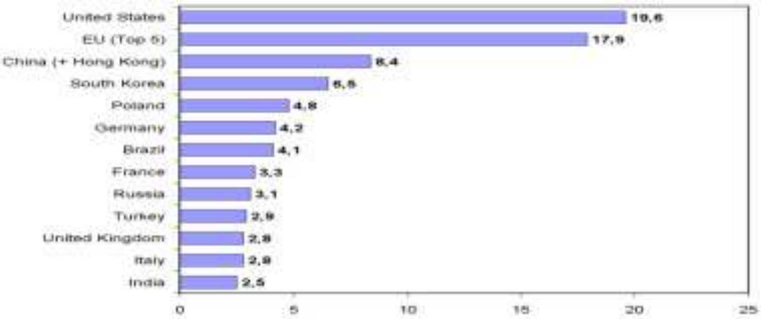


Figure (1) E-mail spam relayed by country 2007 (% of total)

Also the percentage of the spam from 8/2/2008 to 7/2/2009 illustrate this problem as in figure (2). [4]

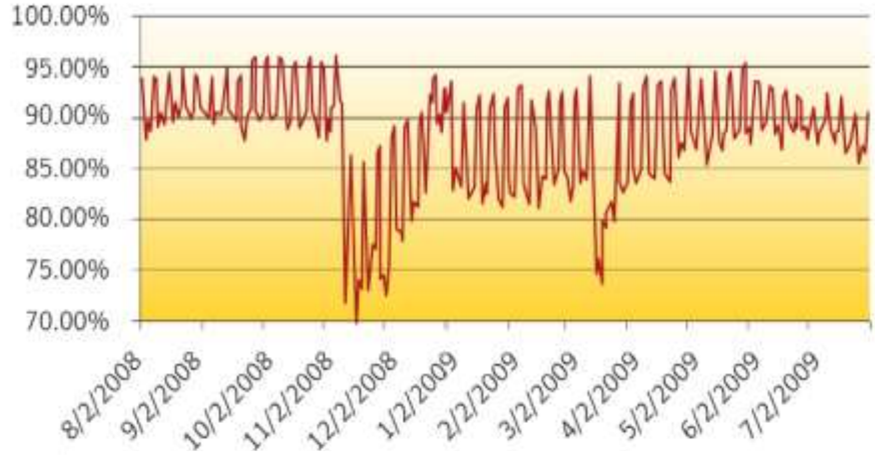


Figure (2) spam percentage

A machine is considered an open relay if it allows unauthorized users to e-mail to a third party that is, neither the person sending the mail nor the person receiving the mail are within domains for which the e-mail system is a mail server [5]. To make the tracking even harder when an open relay is used, most bulk mailers add so-called bogus received headers to the spam message (in front of the real received headers added by the SMTP protocol). By adding these bogus headers they hope to redirect any tracking to a site in the fake header [3].

2 E-mail Server Filter Importance

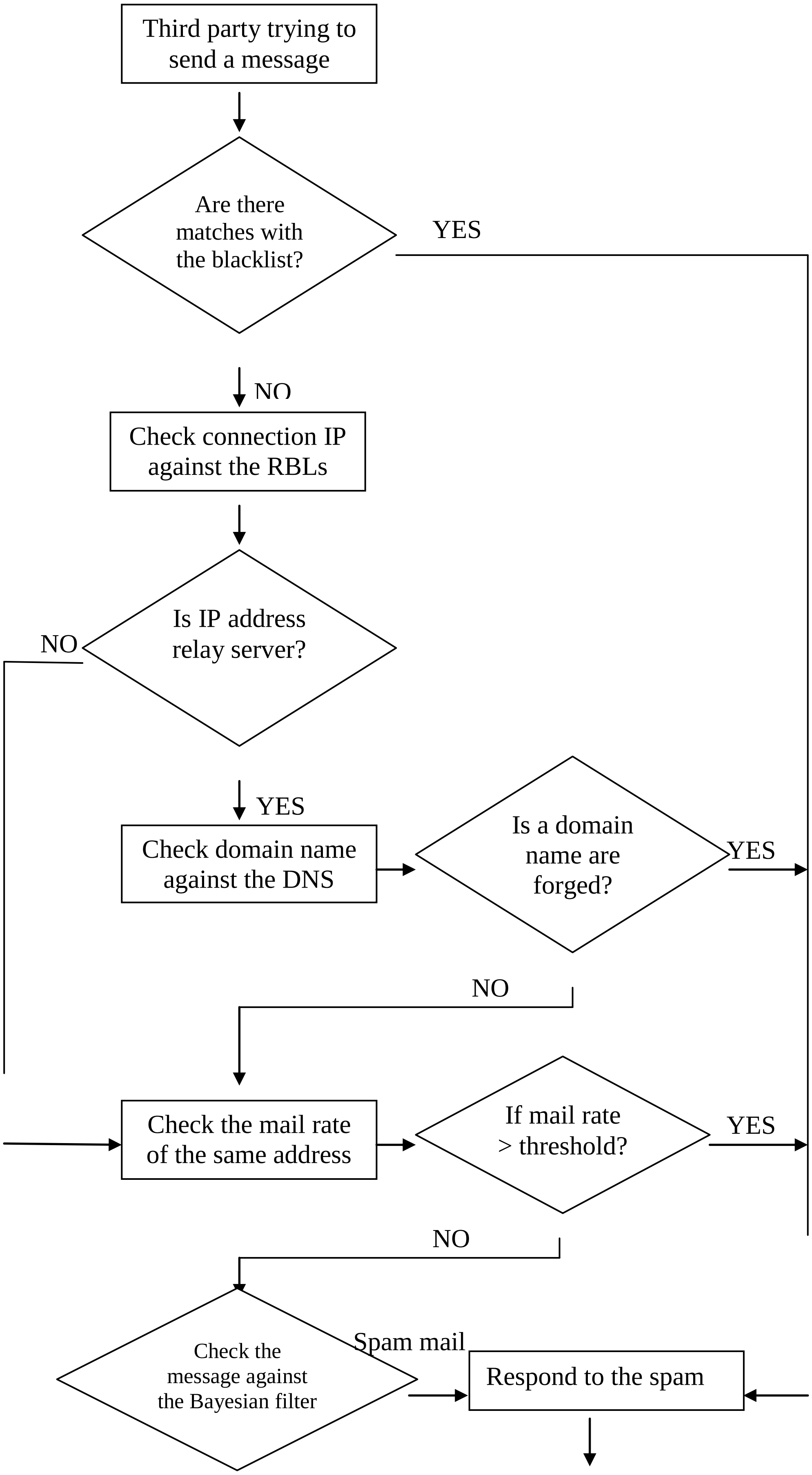
The increased volume of spam consumes an organization's messaging infrastructure, requiring more server’s capacity and network bandwidth, in other words, it has a direct impact on the availability of e-mails [2]. Some filters can be put in place to eliminate unsolicited e-mails sent in bulk because these waste bandwidth and can slow down service for users.

The cost of purchasing software to protect users individually can be higher than protecting them indirectly by protecting the server. Server-based solutions give administrators more control. Even if a company purchases anti-spam software for all of its employees, it is not guaranteed that they will use it correctly. Furthermore, employees who do want to benefit from their anti-spam software will have to spend time tuning their spam filters. Some might not tune them correctly; therefore, spam messages will continue to appear in their mail box or, ever worst, legitimate e-mails could be lost. Client-based solutions do not prevent an organization's networks from being taxed by unwanted e-mails [6]

3 - A Proposed Filtering Spam Structure

The proposed filter concerns with an e-mail server side. The e-mail servers have different work features than other e-mail parts. They can connect with other servers to receive the incoming messages and get the resource IP of the delivering servers, thus the filter can check the source whether it has been trusted or not.

The proposed filter consists of many stages as shown in figure (3). Each stage has its special mechanism to handle the spam. Thereafter each stage is discussed.


Third party trying to send a message
Are there matches with the blacklist?
YES
NO
Check connection IP against the RBLs
Is IP address relay server?
NO
YES
Check domain name against the DNS
Is a domain name are forged?
YES
NO
Check the mail rate of the same address
If mail rate > threshold?
YES
NO
Check the message against the Bayesian filter
Spam mail
Respond to the spam

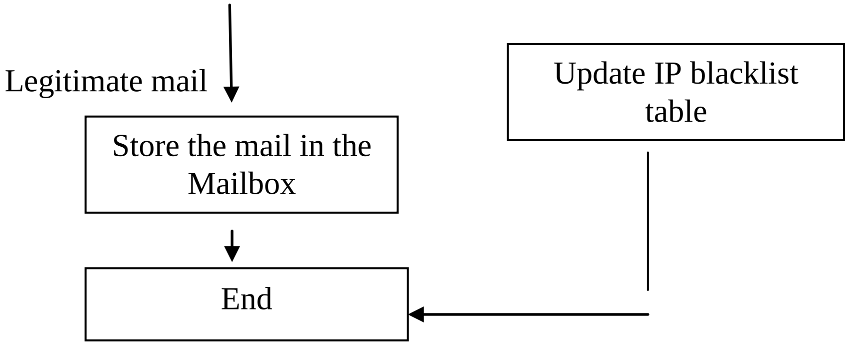


Figure (3) The Proposed Filter

3.1 IP address Blacklist

Blocking the email from certain domains known to be used by spammers can yield good results. It is a simple mechanism to stop the spam by the sender IP address. Every connection that has an unaccepted IP address will be considered as a spam mail.

The receiving server will get the IP address of the sending mail server from the SMTP HELO command, and checks it against the IP addresses in the Blacklist. If a match is found the sender will be considered a spammer and the connection will be disabled, otherwise the filter passes the mail to the next mechanism.

3.2 Real Time Blackhole List (RBL)

This technique, commonly referred to as RBL (real-time black-hole lists), checks the incoming IP address against Black Lists to verify that the sending server is not listed as an open mail relay that spammers can use to relay their unsolicited e-mails. The RBL contains a list of open relay IP addresses maintained by third-party organizations. One of the most reliable databases of server addresses, is maintained ORDB.org [7]

The RBL works like the DNS lockup. It takes the IP address of the sender and checks it with the RBL. If the sender IP address is an open relay server, then the filter will consider the sender as a not trusted one. Here the filter will send the IP address to the DNS Lookup to check whether it has been forged or not. If there has been no positive result from the RBL, that is the IP address has not been an open relay server, the filter considers the sender as trusted and its address is real. Nevertheless the filter will send the IP address to the mail rate control mechanism.

3.3DNS lockup

This technique verifies that the domain name of the sender has not been spoofed. The proposed filter extracts the domain name from the "from:" filed of the address or sender ID-address. The receiving server will get the host name of the sending mail server from the "from:" filed of the header or sender ID-address, performs a simple DNS query and compares the connected IP address with the retrieved IP addresses list to check if there is a match with an IP of the retrieved IP address list.

If the domain name had been forged, then the proposed filter will consider the message or the connection as a spam, and will reject this connection. Thereafter it adds this IP address to the IP blacklist. Else if the domain name has been correct the filter will implement the mail rate control mechanism as the next stage.

This technique can identify if the sending mail server is a legitimate one and has a valid host name. This will eliminate the majority of spam sent by mail servers connected to the Internet using a dial-up connection, as well as most ADSL and cable connections, simply because they are not registered in any domain name server (DNS) as a qualified host.

3.4 Mail Rate control

The proposed filter checks the behavior of the sender to stop who is trying to send a huge number of mails.

Rate mail controls can allow only a certain number of connections from the same e-mail address during a specified time. For example, a rate control time can be set of to 30 minutes with only a certain number of connections to be allowed in that given time period. If the administrator sets this parameter to 50 connections, this stage will block any correspondence after the first 50 connections that come from a single e-mail address within a given 30 minute time period.

The proposed filter also considers the "to:" field as input through the rate mail stage, because the sender can put many recipients' addresses in this field in one message.

3.5 Bayesian filter

The proposed filter uses probabilistic reasoning to decide whether or not a message is spam. This filter bases its choices on the Baye's rule, which is useful for calculating the probability of one event when one knows another event is true. In our case, the rule is used to determine the probability that an e-mail is spam given that it contains certain words. What makes Bayesian filters different from other filters is that they learn. To decide the probability that an e-mail is spam based on the words that it contains the filter needs to know about the e-mails that a user receives.

For the implementation of the Bayesian filter it is required to learn with a set of labeled messages. There are two stages carried out by the Bayesian filter:

i. Training stage

This stage is called training or learning stage. This stage is focused on gathering the information, concerning both spam and legitimate mails. At this stage the filter extract the tokens (words) of the labeled mail by an operation called tokenization that will responsible on extract tokens from the mails, and store them in tables. Two tables will be used, one for tokens of spam mails and other for tokens of legitimate mails. When an e-mail is declared as a spam, the spam table is updated by incrementing the frequency counts for each word contained in that e-mail. Legitimate mail counts are incremented similarly. The count number of spam and non spam e-mails is also recorded for use in the test stage. We can get a list of spam mails from some dependable location in the web to learn filter with it. Also the unlabeled message when it is labeled by the filter will be considered as input to learn with at the test stage. This process is illustrated by the following algorithm:

Given an e-mail message X, labeled with C_j . . . Where j= {spam, legitimate}

```
1- Break X to tokens { $x_1 \dots x_n$ }, each token represents a word.

2- For each token $x_i$ .
   If xi exists in the table of type $C_j$ , then freq [ $x_i$ ] = freq [ $x_i$ ] + 1.
   Else, freq [ $x_i$ ] = 1.
3- Increment the e-mail count of type $C_j$ , count [ $C_j$ ] =count [ $C_j$ ] +1.
```

ii. Testing stage

In the test stage the collected information about spam and non spam will be used as vectors to find the probability that the incoming mail is spam or not. This process is implemented by the following steps:

1- Compute the probability for each x_i considering the training information from the training stage:

$$\Pr [x_i | C_j] = \text{freq} [x_i] / \text{total} [x_e] \quad \text{- - - - - - - - - - - - -} \ (1)$$

Where the freq (x_i) represents the frequency of a particular word in the incoming message and the total (x_e) represents the total frequencies for all words in the training information for all labeled C_j messages.

2- Compute the probability Pr [X| C_j]:

Pr [X| C_j] = Pr [x_1 | C_j] Pr [x_2 | C_j] . . . . Pr [x_n | C_j] - - - - - - (2)

$$= \prod_{i=1}^{k} \Pr\left[x_i \middle| C_j\right]$$ - - - - - - - - - - - - - - - - - - (3)

3- Calculate Pr [C_j] which represents a probability of a message being a spam or non spam on the frequency of spam or normal e-mails:

Pr [C_s] = (count [C_s]) / (count [C_s] +count [C_n]) - - - - - - - - (4)
Pr [C_n] = (count [C_n]) / (count [C_s] +count [C_n]) - - - - - - - - (5)

Where (count [C_s]) is the count of spam mails, (count [C_n]) is the count of non spam mails.

4- For an unlabeled message, X, evaluate the quantities Pr [C_n |X] and Pr [C_s |X], where C_n denotes the class of normal e-mail messages and *Cs* is the class of Spam e-mail messages:

Pr [C_j |X] = Pr [X| C_j] Pr [C_j] / Pr[X] - - - - - - - - - - - - (6)

Pr[X] represents the estimated data about the incoming message, but it will be ignored because it has no effect on the first and second steps.

5- Label the message X as normal, if:

Pr [C_n |X] > Pr [C_s |X] - - - - - - - - - - - - - - - (7)

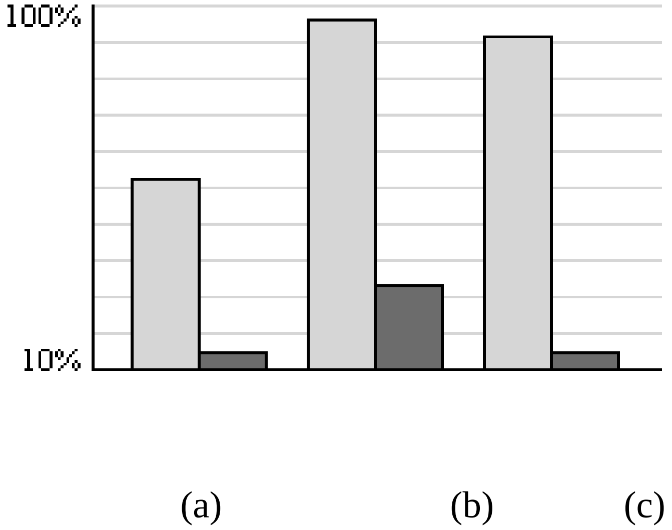
Else it is labeled as spam.

Bayesian formula in the spam filter needs the same number of features (tokens) to implement the equations (2) and (3). In the test stage the Bayesian may have tokens founded only in one class and may not appears in the training stage. Therefore, to equalize the number of parameters considered by the equations we suggest values to these tokens. This is done by suggestion that these tokens have appeared for one time only and assign them the lowest probability.

4 Filter Efficiency and Results

The detection of spam introduces two sources of misclassification: false positive where a non spam e-mail is classified as spam and false negatives where spam slips through incorrectly identified as non spam.

Filter efficiency depends fundamentally on two factors. The first factor is spam detection percentage, and the second factor is the misclassification factors. Our proposed filter efficiency is observed against these two factors as shown in figure (4).



░ Represent the Spam Detected Percentage

■ Represent the Misclassification Percentage

Figure (4) Filter Efficiency versus Different Layers Combinations (a)

statistical filter layer, (b) combining the Blackhole list layer to the precious layer, (c) combining new filter layers consisting of DNS lookup and mail rate control.

Test have been made to filter efficiency by inserting a set of spam and legitimate e-mails to soma layer of the filter. Thereafter the test has been done with same sample of e-mails after combining the previous filters layer with another layer. In each test, filter responses measured as shown in figure (4.a) shows the filter response while testing the statistical filter layer as shown this layer expose 62.4% spam detection percentage and 2% false positive. However, figure (4.b) shows an improvement in filter efficiency after combining the Blackhole list layer to the precious layer with 96.6% spam detection percentage. But, an efficiency degradation occur in the second factor , where the false positive has increased to 20%.

Finally, figure (4.c) show a significant improves in filter efficiency with 93.2% spam detection percentage and 2.3% false positive percentage. These results are obtained after combining new filter layers consisting of DNS lookup and mail rate control, which shown an improvement in spam detection .

5 - Conclusion

In this paper an important tool to protect the mail server from the spam has been presented. Combination of statistical , blackhole , and combining new filter layers consisting of DNS lookup and mail rate control have been tested. A performance measures has been carried out with a set of gathering mails the results are used to evaluate the performance of the different layers of the filter.

Several techniques have been combined , which prove better efficiency in detecting the spam and exhibiting false positive. Some of these techniques depend on validate the legitimacy of the sender. Whiles , the other analyze the contents of the e-mail to classify the mails as spam or legitimate .The proposed filter exposes graceful results in eliminating the spam ,or at least reduces the spam rate at the server side .

References

[1] Le Zhang, Jingbo Zhu, Tianshun Yao , "**An Evaluation of Statistical Spam Filtering Techniques**", Natural Language Processing Laboratory Institute of Computer Software & Theory Northeastern University, China , 2004. homepages.inf.ed.ac.uk/s0450736/paper/2004-spameval.pdf

[2] Sahami et al**," Spam: A Security Issue**", Ciphertrust whitepaper, December 2003.www.ciphertrust.com/partnerconnection/collateral/whitepapers/spam_security_issue.pdf

[3] Flavio D. Garcia, Jaap-Henk Hoepman, Jeroen van Nieuwenhuizen, "**SPAM FILTER ANALYSIS**", University of Nijmegen, Netherlands, 2004. www.cs.ru.nl/~flaviog/publications/spam-filter.pdf

[4] http://news.softpedia.com/newsImage/Spam-Accounted-89-of-All-Emails- in- July-2009-3.png/

[5] Paul Wolfe, Charlie Scott, Mike Erwin, "**Anti-Spam Toolkit**", Publisher: Brandon A. Nordin, ISBN 0-07-223167-x, 2004.
searchexchange.techtarget.com/searchExchange/downloads/Chapter05.pdf

[6] L. Pelletier, J. Almhana, V. Choulakian, "**Adaptive Filtering of SPAM**",University of Moncton, E1A 3E9 , 2004.
www.umoncton.ca/greti/papers/ Adaptive Filtering of Spam.pdf

[7] Walter Daelemans, Jakub Zavrel, Ko van der Sloot, **"TiMBL: Tilburg Memory-Based Learner",** Reference Guide, Technical Report, Tilburg University, December 31, 2004.
http://ilk.uvt.nl/downloads/pub/papers/ilk0402.pdf

بناء مرشح لرسائل الدعاية للبريد الألكتروني في جانب الخادم

احمد عبد الرضا عباس
جامعة الكوفة
كلية التربية للبنات
قسم الحاسبات

الخلاصة

لقد أصبحتْ رسالةُ الدعاية الآن قضية أمنية هامّة وهي تسبب هدر هائل على المصادرِ الماليةِ. في هذه الورقةِ قدمنا مرشِح رسالةِ دعاية، والذي يَعْملُ في مستوى الخادمَ. إنّ المرشِحَ المُقتَرَحَ هو مجموعة حَلِّ ضِدِّ الرسائل الدعائية. إنّ مهمّةَ المرشِحِ المُقتَرَحِ أَنْ تُقلّلَ قدرةَ مرسلي الرسائل الدعائية لصَرْف انتباه عمل الشبكةَ برسائلِ الدعاية. وهذا يحصل من خلال مَنْع رسالةِ الدعاية في مستوى الخادمَ. الحل الأساسي للخادم هو عادة يفيد في حِماية مستعملي البريد الإلكتروني بشكل منفرد. مثل هذا الحل يعطي سيطرة أكثر إلى مدراء إدارة الشبكة.