

**SOLUCIÓN PARA EL MEJORAMIENTO DEL SERVICIO DE CORREO ELECTRÓNICO
EN LA RED DE DATOS DE LA UNIVERSIDAD DEL CAUCA.**

Trabajo de desarrollo

**DAVID FERNANDO ANDRADE SOLANO
VICTOR ANDRES CASTRO DUEÑAS**



ANEXO C

**APLICACIÓN DE METODOLOGÍA OPENBRR PARA LA EVALUACIÓN DE
HERRAMIENTAS *ANTISPAM***

Director: Guefry Agredo Méndez M.Sc.

**UNIVERSIDAD DEL CAUCA
FACULTAD DE INGENIERÍA ELECTRÓNICA Y TELECOMUNICACIONES
DEPARTAMENTO DE TELECOMUNICACIONES
Grupo I+D Nuevas Tecnologías en Telecomunicaciones - GNTT
POPAYÁN, 2012**

TABLA DE CONTENIDO

Introducción	1
1. Evaluación rápida	1
2. Criterios de evaluación para la escogencia del software	3
3. Ponderación de los criterios de evaluación.....	3
4. Selección y ponderación de Métricas por cada criterio.....	5
5. Evaluación de las métricas de cada criterio	12
6. Resultados de la evaluación	21
REFERENCIAS BIBLIOGRÁFICAS.....	23

LISTA DE TABLAS

Tabla 1. Evaluación rápida de herramientas <i>antispam</i>	2
Tabla 2. Puntuación de criterios.....	4
Tabla 3. Métricas de Funcionalidad	6
Tabla 4. Métricas de Calidad	7
Tabla 5. Métricas de Rendimiento.....	7
Tabla 6. Métricas de Comunidad	7
Tabla 7. Métricas de Documentación	8
Tabla 8. Métricas de Soporte	8
Tabla 9. Métricas de Escalabilidad.....	8
Tabla 10. Evaluación Métrica de Efectividad en el rechazo de <i>spam</i>	9
Tabla 11. Evaluación Métrica de Efectividad en la aceptación de correo legítimo	9
Tabla 13. Evaluación Métricas de Calidad	10
Tabla 14. Evaluación Métricas de Rendimiento	10
Tabla 15. Evaluación Métricas de Comunidad	10
Tabla 16. Evaluación Métricas de Documentación.....	11
Tabla 17. Evaluación Métricas de Soporte.....	11
Tabla 18. Evaluación Métricas de Escalabilidad	11
Tabla 19. Calificación para Criterios	12
Tabla 20. Evaluación de Funcionalidad.....	17
Tabla 21. Puntuación asignada al porcentaje obtenido	17
Tabla 22. Resultado de Evaluación de Funcionalidad.....	18
Tabla 23. Evaluación de Calidad.....	18
Tabla 24. Evaluación de Rendimiento.....	19
Tabla 25. Evaluación de Comunidad	19
Tabla 26. Evaluación de Documentación	20
Tabla 27. Evaluación de Soporte	20
Tabla 28. Evaluación de Escalabilidad.....	21
Tabla 29. Puntuación BRR.....	21

LISTA DE FIGURAS

Figura 1 Prueba de Spamassassin con mensajes <i>spam</i>	13
Figura 2 Prueba de Dspam con mensajes <i>spam</i>	14
Figura 3 Prueba de Spamassassin con mensajes legítimos	15
Figura 4 Pruebas de Dspam con mensajes legítimos	16

ANEXO C

APLICACIÓN DE METODOLOGÍA OPENBRR PARA LA EVALUACIÓN DE HERRAMIENTAS *ANTISPAM*

Introducción

En la selección de software para realizar un proceso detallado y justificado acerca de la escogencia de las herramientas que se van a utilizar. Se utilizó la metodología de Puntuación para la Preparación de Negocios Abiertos (OpenBRR, *Open Business Readiness Rating*) [1] para la evaluación de las herramientas *antispam*, esta metodología define ciertos pasos los cuales se ven evidenciados en este anexo y en el cual se muestran las diferentes tablas de como se realizó la evaluación de estas herramientas.

1. Evaluación rápida

La metodología OpenBRR, define que se debe realizar en primera instancia una evaluación rápida [1], la cual consiste en reducir el número de herramientas de forma rápida evaluando algunas características de las mismas. Algunas características a evaluar son definidas por la metodología y otras se consiguen según al contexto del proyecto desarrollado. De acuerdo con lo anterior, las características escogidas para la evaluación rápida se presentan a continuación.

- a) Evaluar el tipo de licencia, el sistema debe ser de código abierto (*Open Source*)
- b) La herramienta se debe poder instalar en Linux.
- c) El sistema tiene una comunidad que lo respalde.
- d) Existen soluciones a errores encontrados. Se han publicado nuevas versiones o nuevos paquetes en el último año.
- e) El software debe funcionar en el lado del servidor SMTP y no en el lado del cliente.
- f) Existe documentación disponible suficiente y clara del sistema.
- g) El sistema es conocido o utilizado. Existencia de empresas clientes o usuarios a los que se puede referenciar.
- h) El sitio del sistema tiene un diseño profesional.
- i) La versión compilada sea gratuita

En esta etapa se realiza una lista de las posibles herramientas, y la aplicación de los filtros planteados en el punto anterior. La forma de filtrar los sistemas será la siguiente:

- Se selecciona la herramienta a evaluar
- Se evaluará si cumple con los atributos de comparación que figuran en el punto anterior. La evaluación se realizará siguiendo el orden indicado según su importancia. En caso de encontrar que una herramienta no cumple con un atributo, o la información disponible no es suficiente para determinar si ésta cumple o no con el requisito, se descarta de la lista sin tener la necesidad de seguir evaluando los siguientes atributos. En la Tabla 1 se muestran las herramientas y la evaluación de las características de forma rápida.

Tabla 1. Evaluación rápida de herramientas *antispam*

Software	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
Mailfilter	✓	✓	✓	✓	X	?	?	?	?
QSF	✓	✓	X	?	?	?	?	?	?
BMF	✓	✓	✓	✓	X	?	?	?	?
Spamoracle	✓	✓	✓	X	?	?	?	?	?
Bsfilter	✓	✓	✓	✓	X	?	?	?	?
Spamfighter.	X	?	?	?	?	?	?	?	?
Cloudmark DesktopOne	X	?	?	?	?	?	?	?	?
Spamassassin	✓	✓	✓	✓	✓	✓	✓	✓	✓
MailWasher	X	?	?	?	?	?	?	?	?
ChoiceMail One	X	?	?	?	?	?	?	?	?
VIPRE-Internet-Security	X	?	?	?	?	?	?	?	?
Dspam	✓	✓	✓	✓	✓	✓	✓	✓	✓
Spam Bully	X	?	?	?	?	?	?	?	?
SpamEater	X	?	?	?	?	?	?	?	?
ASSP	✓	X	?	?	?	?	?	?	?
Policy patrol	X	?	?	?	?	?	?	?	?
Spambayes	✓	✓	✓	✓	X	?	?	?	?
Hackerfas	✓	✓	✓	✓	✓	✓	✓	✓	X

Con base en el cuadro de comparación mostrado en la Tabla 1 y de acuerdo con el cumplimiento de los parámetros aplicados para la eliminación de herramientas, las que cumplieron con todos los requisitos fueron:

- Spamassassin
- Dspam

A continuación se realiza la evaluación de estas herramientas de acuerdo con la metodología OpenBRR, empezando por escoger los criterios con los cuales se llevará a cabo esta evaluación y de esta manera realizar la selección de las herramientas.

2. Criterios de evaluación para la escogencia del software

La metodología propone 12 criterios de evaluación [1], pero también sugiere que la inclusión de todas las categorías es desfavorable, por lo tanto se deben escoger 7 o menos de los criterios que se muestran a continuación. Los criterios son escogidos de acuerdo al contexto del proyecto.

- Funcionalidad
- Usabilidad
- Calidad
- Seguridad
- Rendimiento
- Escalabilidad
- Arquitectura
- Soporte
- Documentación
- Adopción
- Comunidad
- Profesionalismo

Los criterios para realizar la evaluación de las herramientas *antispam* se escogieron dada la importancia que estos tienen para seleccionar un software que sea completamente funcional a largo plazo para el servicio de correo de la Institución, estos criterios son los siguientes:

- Funcionalidad
- Calidad
- Rendimiento
- Comunidad
- Documentación
- Soporte
- Escalabilidad

3. Ponderación de los criterios de evaluación

De acuerdo con la metodología BRR hay que asignar una puntuación a cada criterio dependiendo de su importancia, la suma de esta puntuación debe de ser un total del 100%. En caso de que se escojan 7 criterios, aquel que se encuentre en el cuarto lugar de importancia, la metodología BRR aconseja que su puntuación sea del 15% y las demás se ponderan alrededor de éste.

En la Tabla 2, se muestra los porcentajes asignados a cada criterio para la evaluación de las herramientas *antispam*.

Tabla 2. Puntuación de criterios

Posición	Categoría	Peso
1	Funcionalidad	25,00%
2	Calidad	20,00%
3	Rendimiento	20,00%
4	Comunidad	15,00 %
5	Documentación	10,00%
6	Soporte	5,00%
7	Escalabilidad	5,00%
	Total	100%

A continuación se realiza la descripción de cada uno de los criterios tenidos en cuenta, el orden de importancia de estos criterios fue escogido de acuerdo a las recomendaciones del documento oficial de OpenBRR.

Funcionalidad: la herramienta debe cumplir con las funcionalidades solicitadas para contribuir con el mejoramiento del sistema de correo en cuanto al rechazo de *spam*, esta categoría se considera la mas importante, por lo tanto es la que va a tener mayor ponderación entre todas las categorías.

Calidad: qué tan bien respondan las herramientas en la realización de sus funciones, la herramienta debe brindar garantías de buen funcionamiento asegurando un grado de calidad en sus operaciones.

Rendimiento: tiempo de respuesta de la herramienta en la realización de sus funciones.

Comunidad: tener el respaldo de una comunidad activa que pueda ofrecer una pronta respuesta a los problemas que presente el software.

Documentación: documentación necesaria para poder realizar las instalaciones y configuraciones de las herramientas.

Soporte: respuestas obtenidas en un tiempo prudente de acuerdo a las inquietudes y a la presencia de problemas por parte de los usuarios.

Escalabilidad: crecimiento que pueda tener la herramienta sin alterar la calidad de sus funciones. Este criterio es el de menor ponderación ya que no es muy influyente para la selección de herramientas, además, no difiere mucho entre las herramientas que se van a evaluar.

Luego de conocer las categorías tenidas en cuenta para realizar la comparación, se deben escoger las métricas para cada una de éstas.

4. Selección y ponderación de Métricas por cada criterio

Cada uno de los criterios escogidos esta compuesto por unas métricas para realizar su evaluación, estas métricas se encuentran definidas en el documento oficial de la metodología OpenBRR [1] y a cada una de estas se le asigna un porcentaje de acuerdo a la importancia que tiene para la evaluación del criterio, la suma de los porcentajes de todas las métricas por cada criterio debe dar como resultado 100%.

A las métricas de funcionalidad se les realiza un tratamiento diferente de las métricas de los otros criterios, estas se escogen de acuerdo al proyecto que se esté desarrollando y de las funcionalidades del software a evaluar.

Para la elección de las métricas de funcionalidad se tomó como referencia lo descrito en [2] la cual expresa que la eficiencia de una herramienta *antispam* se puede medir siguiendo lo enunciado en el siguiente apartado:

“En un entorno de funcionamiento real, podemos medir la calidad del filtro mediante dos indicadores:

- *El porcentaje de mensajes no spam calificados como spam erróneamente (“tasa de falsos positivos”).*
- *El porcentaje de mensajes spam filtrados correctamente (“precisión”).*

El parámetro más importante desde el punto de vista del usuario real es el porcentaje de falsos positivos. En efecto, si un mensaje Spam evade el filtro es fácil hacer click en el botón de borrar. Sin embargo, si se marca con etiqueta de Spam a un mensaje normal, el usuario no lo verá y éste quedará perdido.”

También se observó que en la referencia [3] para evaluar las herramientas *antispam*, se analizan los resultados aplicando la siguiente formula:

*“Tasa de detección de spam = (spam marcado como spam/total de spam)*100
Tasa de falsos positivos= (genuinos marcados como spam/total de genuinos)*100”*

Según las citas anteriores, los principales parámetros que se tienen en cuenta para evaluar una herramienta *antispam* son la efectividad en el rechazo de *spam* y la detección

de falsos positivos de acuerdo al número total de mensajes no deseados y legítimos que llegan al servidor.

Para los otros parámetros se tomaron referencias a partir del conocimiento y la experiencia por parte del jefe del área de Servidores y Servicios de Internet de la Universidad del Cauca, el director y también del estudio y aprendizaje de los desarrolladores de este proyecto acerca del *spam*, como por ejemplo que existen redes colaborativas que pueden complementar el filtro *antispam* y otros parámetros que dependen más de los métodos utilizados para el control de correo no deseado como el manejo de redes bayesianas y de métodos heurísticos. El parámetro de que la herramienta fuera de software libre fue seleccionado, ya que el desarrollo de este proyecto se dirige a ofrecer una mejora basada en herramientas de software libre.

Para la puntuación de las métricas de este criterio se usa una escala de 1 a 3, en donde 1 es la de menor importancia y 3 la de mayor importancia.

A continuación, se especifican en diferentes tablas las métricas tenidas en cuenta para cada uno de los criterios escogidos con su respectiva descripción.

Tabla 3. Métricas de Funcionalidad

Funcionalidad	Descripción	Ponderación
Efectividad en el rechazo de <i>spam</i>	Cantidad de mensajes que sean detectados como <i>spam</i> .	3
Efectividad en la aceptación de correo legítimo	Cantidad de falsos positivos detectados por la herramienta, estos son correos legítimos detectados como <i>spam</i> .	3
Manejo de redes bayesianas	Consisten en que la herramienta se encarga de revisar el contenido del mensaje y determinar a través de métodos probabilísticos y de unas políticas establecidas si se cataloga como <i>spam</i> , basado en un entrenamiento previo.	2
Manejo de métodos heurísticos	Consisten en buscar patrones que suelen repetirse en el contenido de los mensajes.	2
Uso de redes colaborativas	Son redes que complementan el funcionamiento de las herramientas <i>antispam</i> , estas son bases de datos que se utilizan para la detección y notificación de <i>spam</i> .	2
Software libre	Que la herramienta sea de software libre ya que el desarrollo del proyecto se basa en herramientas de software libre	1

Tabla 4. Métricas de Calidad

Calidad	Descripción	Ponderación
Número de errores abiertos en los últimos 6 meses	Esta mide la calidad de uso del producto	20%
Número de errores corregidos en los últimos 6 meses (en comparación con el número de errores abiertos)	Mide que tan rápido son corregidos los errores	30%
Cantidad de errores críticos/P1 abiertos.	Mide la gravedad de los problemas de calidad encontrados	30%
Promedio de tiempo en que se solucionan los errores críticos / P1 en los últimos 6 meses	Mide la capacidad de respuesta para solucionar problemas críticos	20%

Tabla 5. Métricas de Rendimiento

Rendimiento	Descripción	Ponderación
Pruebas de rendimiento e informes de referencia disponibles	Mide si hay alguna prueba de rendimiento realizada y puntos de referencia publicados, típicamente en comparación a otras soluciones equivalentes.	50%
Ajuste de rendimiento y configuración	Mide si hay alguna documentación o herramienta para ayudar a afinar el componente por desempeño (información de CPU, disco, red)	50%

Tabla 6. Métricas de Comunidad

Comunidad	Descripción	Ponderación
El volumen medio de la lista de correo general en los últimos 6 meses	La lista de correo general es el lugar donde la comunidad se ayuda.	50%
Número de contribuyentes de código en los últimos 6 meses	Los contribuyentes de código por lo general promueven la construcción de la comunidad en torno al proyecto. Cuanto mayor sea el número de contribuyentes de código, mejor será el apoyo de la comunidad.	50%

Tabla 7. Métricas de Documentación

Documentación	Descripción	Ponderación
Existencia de varios tipos de documentación	Una buena documentación debe incluir la documentación de varios grupos de usuarios en varios formatos.	50%
Marco de contribución de usuario	Las mejores guías a menudo provienen de muestras de usuarios. Quiere decir, realimentarse de personas que han usado los productos.	50%

Tabla 8. Métricas de Soporte

Soporte	Descripción	Ponderación
El volumen promedio de la lista de correo general en los últimos 6 meses	La lista de correo general, es el primer lugar donde la gente va por ayuda gratuita.	50%
Calidad de apoyo profesional	El apoyo profesional que ayude a afinar para el despliegue local y la solución de problemas es siempre deseable.	50%

Tabla 9. Métricas de Escalabilidad

Escalabilidad	Descripción	Ponderación
Despliegue de referencia	Mide si el software es escalable y probado en uso real a través de un despliegue en el mundo real.	50%
Diseñado para escalabilidad	Mide si el componente fue diseñado con escalabilidad en mente ¿se ejecuta en un ambiente de grupo? Puede resolver H/W problemas de rendimiento?	50%

Cada una de las métricas tiene sus propias puntuaciones para realizar la calificación de las herramientas, a continuación se muestra como se realiza la puntuación de las métricas para cada uno de los criterios escogidos.

Las métricas de funcionalidad se evaluarán de la siguiente manera:

- La evaluación de la primera métrica que es la efectividad en el rechazo de *spam* se realizara como se muestra en la Tabla 10.

Tabla 10. Evaluación Métrica de Efectividad en el rechazo de *spam*

Efectividad en el rechazo de <i>spam</i>	
Porcentaje de Rechazo (%)	Puntuación
0% - 30%	1
30% - 70%	2
70% - 100%	3

- La métrica de efectividad en la aceptación de correo legítimo (falsos positivos), se evalúa como se muestra en la Tabla 11.

Tabla 11. Evaluación Métrica de Efectividad en la aceptación de correo legítimo

Efectividad en la aceptación de correo legítimo	
Reconocimiento de falsos positivos	Puntuación
No	3
Si	1

- Para las métricas Manejo de redes bayesianas, Manejo de métodos heurísticos y el Uso de redes colaborativas, por ser características de funcionamiento adicionales que pueden presentar o no las herramientas, según lo establecido por OpenBRR la evaluación se realiza premiando a la herramienta que cumpla con estas métricas, de tal forma que se suma el valor del peso que se le dio a la métrica, o castigándola restándole este mismo valor en caso de que no cumpla. La evaluación de la métrica de Software Libre se hace de la misma manera.

Para los otros criterios, de acuerdo con la metodología OpenBRR, la evaluación se realiza como se muestra a continuación en cada una de las tablas.

Tabla 12. Evaluación Métricas de Calidad

Calidad					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
Número de errores abiertos en los últimos 6 meses	< 50	50-100	100-500	500-1000	>1000
Número de errores corregidos en los últimos 6 meses (en comparación con el número de errores abiertos)	>75 %	60% - 70%	45% - 60%	25% - 45%	< 25%
Cantidad de errores críticos/P1 abiertos.	0	1-5	5-10	10-20	>20
Promedio de tiempo en que se solucionan los errores críticos / P1 en los últimos 6 meses	<1 semana	1-2 semanas	2-3 semanas	3-4 semanas	>4 semanas

Tabla 13. Evaluación Métricas de Rendimiento

Rendimiento					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
Pruebas de rendimiento e informes de referencia disponibles	Si, con buenos resultados		Si		No
Ajuste de rendimiento y configuración	Si, frecuente		Si, un poco		No

Tabla 14. Evaluación Métricas de Comunidad

Comunidad					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
El volumen medio de la lista de correo general en los últimos 6 meses	> 720 mensajes por mes	300 - 720 msjs por mes	150 - 300 msjs por mes	30 - 150 msjs por mes	< 30 msjs por mes
Número de contribuyentes de código en los últimos 6 meses	> 50	20 – 50	10 - 20	5 - 10	< 5

Tabla 15. Evaluación Métricas de Documentación

Documentación					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
Existencia de varios tipos de documentación	Instalar / implementar, usuario, administración, optimización, mejora, desarrollo, la documentación está disponible en múltiples formatos (pdf, html sola, de varios archivos html).	Instalar / desplegar, usuario, administrador, guías de actualización disponibles en varios formatos.	Instalar / implementar y guía de usuario disponible	Sólo basado en texto, existe documentación de instalación	Sin la documentación adecuada. Un archivo README no cuenta
Marco de contribución de usuario	Las personas pueden contribuir, y las contribuciones son editadas/filtradas por expertos		A las personas se les permite contribuir		Los usuarios no pueden contribuir.

Tabla 16. Evaluación Métricas de Soporte

Soporte					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
El volumen promedio de la lista de correo general en los últimos 6 meses	> 720 mensajes por mes	300 - 720 msjs por mes	150 - 300 msjs por mes	30 - 150 msjs por mes	< 30 msjs por mes
Calidad de apoyo profesional	Instalación + solución de problemas + Integración / Apoyo personalizado		Solamente Soporte de Instalación.		Sin soporte profesional.

Tabla 17. Evaluación Métricas de Escalabilidad

Escalabilidad					
Métricas	Calificación				
	Excelente	Muy bueno	Aceptable	Pobre	Inaceptable
Despliegue de referencia	Si, con publicaciones de los usuarios		Si		No
Diseñado para escalabilidad	Si, extensivo		Si, un poco		No

A excepción del criterio de funcionalidad, cada uno de los criterios escogidos para la evaluación tiene asociadas unas métricas para realizar la comparación. Para esto se utiliza una escala que asocia unos valores de 1 a 5 a una calificación, como se observa en la Tabla 19.

Tabla 18. Calificación para Criterios

Valor	Concepto
1	Inaceptable
2	Pobre
3	Aceptable
4	Muy bueno
5	Excelente

5. Evaluación de las métricas de cada criterio

Luego de saber como se realiza la evaluación de los criterios escogidos con sus respectivas métricas, el siguiente paso es evaluar las herramientas que cumplieron con todas las características tenidas en cuenta en la evaluación rápida.

Primero se realiza la evaluación para las métricas de funcionalidad.

Para verificar el cumplimiento de la primera métrica que es la **Efectividad en el rechazo de spam** se realizan pruebas de las herramientas con el fin de verificar cual es la que mejor cumple con este parámetro, la que mejor cumpla es la que mayor cantidad de *spam* rechace. Para graficar los resultados de estas pruebas se utilizó el programa llamado **Mailgraph**¹ el cual genera una gráfica basándose en el sistema de registro de mensajes del servidor de correo, mostrando la cantidad de correo que se recibió en un periodo de tiempo y cuales de estos son detectados como *spam*.

A continuación se evidencian las pruebas realizadas con los filtros **Spamassassin** y **Dspam**.

➤ Prueba con Spamassassin

Para esta prueba se enviaron 484 correos *spam* para observar cuántos de estos son rechazados. La Figura 1 muestra la cantidad de mensajes *spam* enviados y cuántos de estos fueron rechazados.

¹ <http://mailgraph.schweikert.ch/>

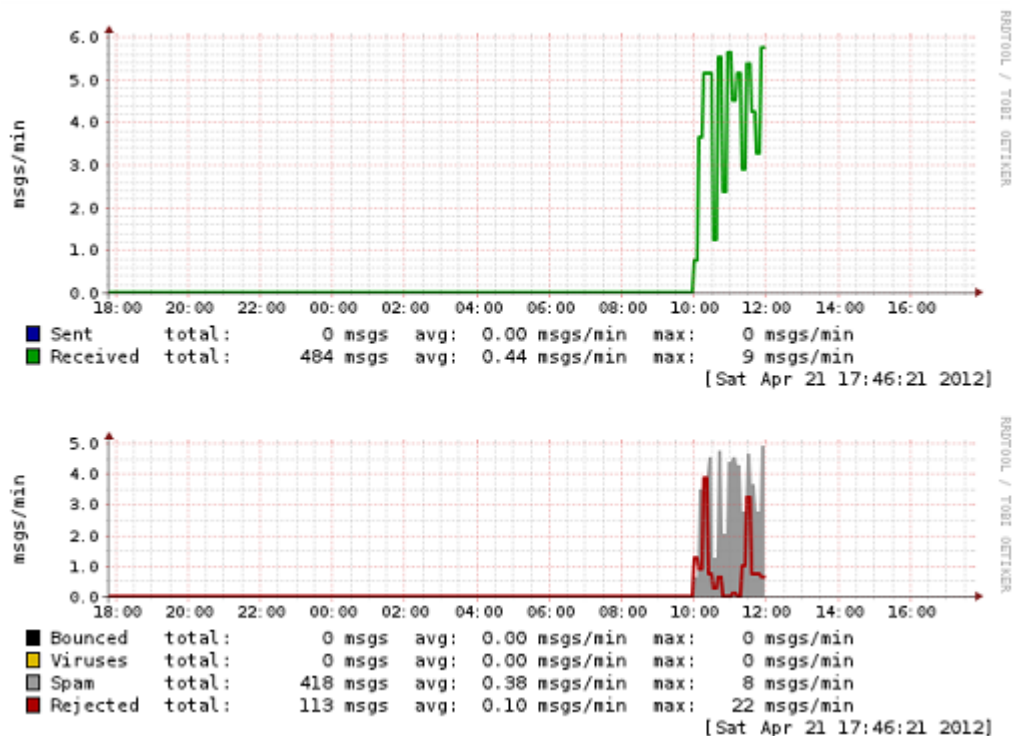


Figura 1 Prueba de Spamassassin con mensajes *spam*

La primera parte de la gráfica representa con una línea verde los mensajes recibidos por minuto y la hora en que estos se recibieron, se observa que el filtro recibió 484 mensajes.

En la segunda parte, el color gris representa los mensajes clasificados como *spam*, se observa que **Spamassassin** detectó 418 mensajes como *spam*, de los 484 recibidos, esto quiere decir que el filtro rechazó el 86% de los mensajes *spam*. La línea roja representa otras conexiones por parte de otros servidores hacia este servidor de correo. Como este servidor tiene configurado los filtros de reputación, estos rechazan estas conexiones de servidores desconocidos.

➤ Prueba con Dspam

Para esta prueba se llevó a cabo el envío de 484 mensajes *spam*, para observar que acciones realizaba el **Dspam** al recibirlos, esto se muestra en la Figura 2.

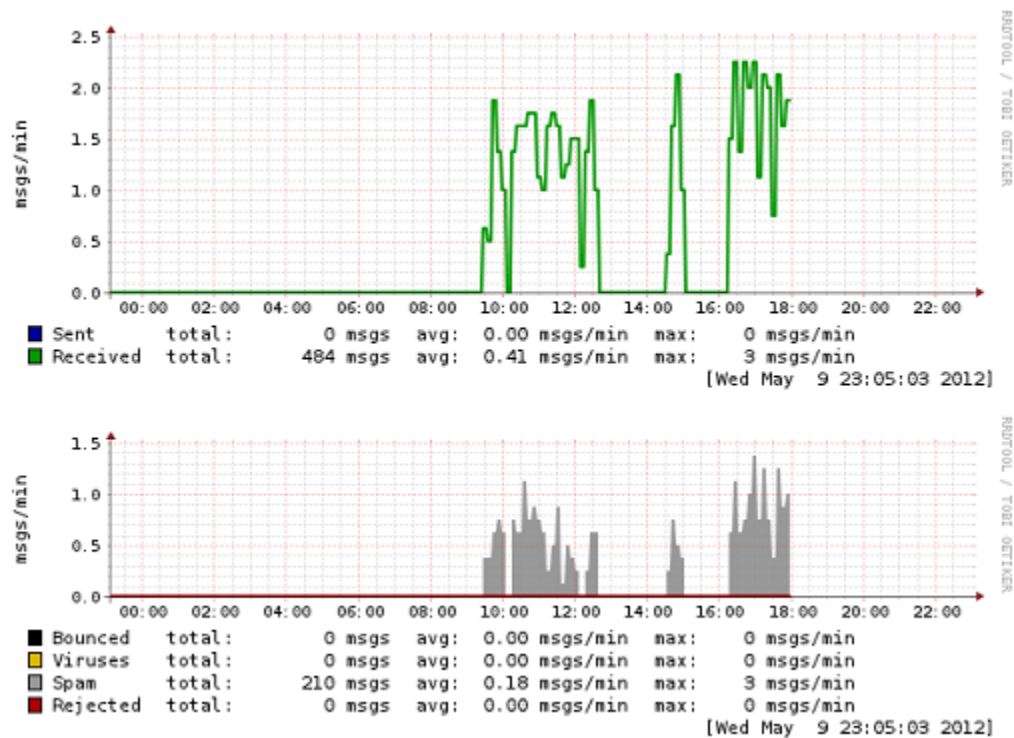


Figura 2 Prueba de Dspam con mensajes *spam*

Se puede observar que de los 484 mensajes recibidos que se representan con la línea verde, 210 fueron detectados como *spam*, esto quiere decir que el 44% de los mensajes fueron detectados como *spam*.

En cuanto a la evaluación, la herramienta *antispam* que mejor cumple con este parámetro es **Spamassassin**, esta herramienta tuvo un rechazo de correo *spam* del 86% de acuerdo al número de mensajes *spam* enviados, mientras que **Dspam** solamente tuvo un rechazo del 44% en el rechazo de mensajes no solicitados, esto quiere decir que la efectividad en el rechazo de mensajes no deseados por parte de **Spamassassin** es mucho mayor a la efectividad que presenta **Dspam** realizando esta misma función.

La segunda métrica evaluada en el criterio de funcionalidad es la **Efectividad en la aceptación de correo legítimo**, para esta se realizan pruebas de las herramientas con el fin de verificar cual es la que mejor cumple con este parámetro. La que mejor cumple es la que menor número de falsos positivos identifique. Para graficar los resultados se utilizó **Mailgraph**².

A continuación se evidencian las pruebas realizadas con los filtros **Spamassassin** y **Dspam**.

² mailgraph.schweikert.ch

➤ Prueba con Spamassassin

Luego de probar con correo *spam*, se realizaron pruebas con correo legítimo para observar cuántos de estos deja pasar y observar si se detecta alguno como *spam*, el resultado se muestra en la Figura 3.

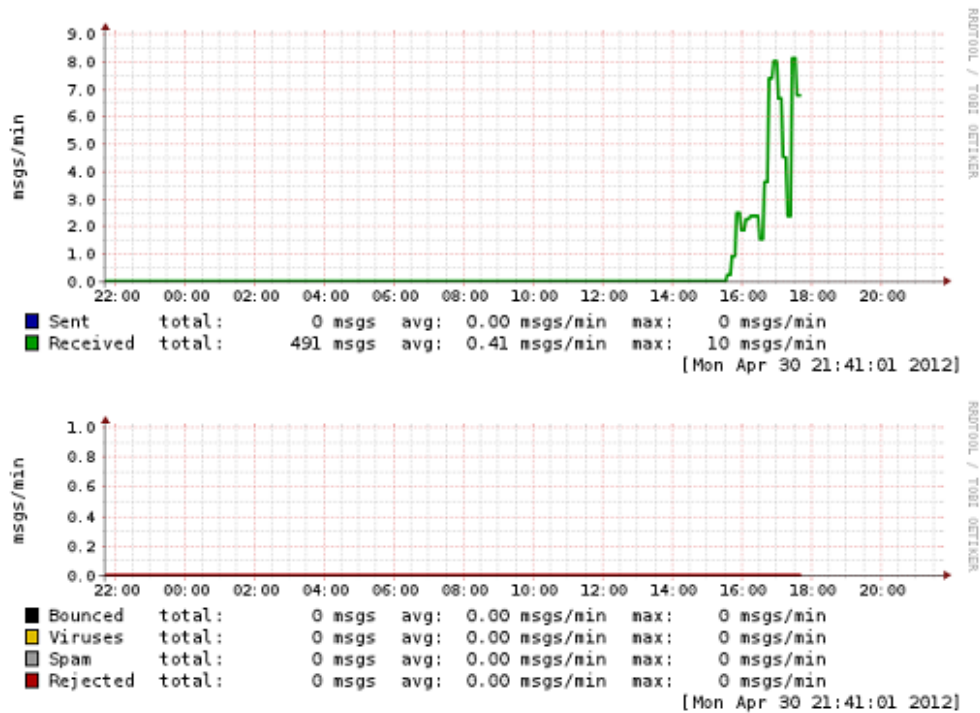


Figura 3 Prueba de Spamassassin con mensajes legítimos

La línea verde representa los mensajes recibidos por el filtro, se recibieron 491 mensajes. En la segunda parte de la gráfica se observa que ninguno de los mensajes fue detectado como *spam*. Esto quiere decir que no hubo falsos positivos.

➤ Prueba con Dspam

De los 491 mensajes recibidos representados con la línea verde, **Dspam** reconoció un falso positivo, esto quiere decir que de los supuestos 491 mensajes legítimos, esta herramienta detectó uno (1) de ellos como *spam*.

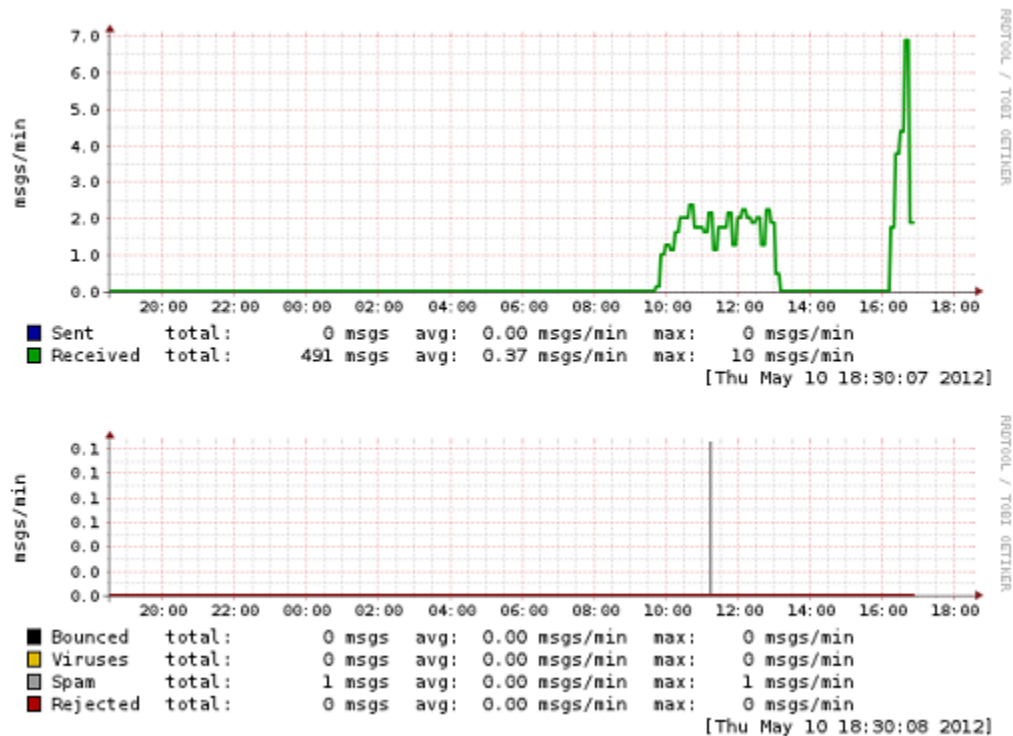


Figura 4 Pruebas de Dspam con mensajes legítimos

En la Figura 4 se muestra que se enviaron 491 mensajes legítimos que fueron recibidos por el filtro **Dspam**, para verificar que acciones realizaba.

En cuanto a la evaluación, la herramienta *antispam* que mejor cumple con este parámetro es **Spamassassin**, el número de falsos positivos de esta herramienta fue igual a cero o sea de un 0%, mientras que **Dspam** detectó un falso positivo que equivale a un 0.2%. Esto quiere decir que la efectividad en aceptación de correo legítimo por parte de **Spamassassin** es mayor que la de **Dspam**.

La siguiente métrica es el **Manejo de redes bayesianas** y tanto **Spamassassin** como **Dspam** manejan este método, pero la única que cumple con la métrica del **Manejo de métodos heurísticos** es **Spamassassin**. Las dos herramientas cumplen con el **Uso de redes colaborativas** como complemento para la detección de correo no deseado y también cumplen con la última métrica ya que ambas son de **Software Libre**.

A continuación se observan una serie de tablas, cada una de ellas muestra la evaluación de cada uno de los criterios y sus métricas en las herramientas.

Tabla 19. Evaluación de Funcionalidad

Funcionalidad	Peso	Puntuación	
		Spamassassin	Dspam
Efectividad en el rechazo de <i>spam</i>	3	3	2
Efectividad en la aceptación de correo legítimo	3	3	1
Manejo de redes bayesianas	2	2	2
Manejo de métodos heurísticos	2	2	-2
Uso de redes colaborativas	2	2	2
Software libre	1	1	1
Total	13	13	6
Porcentaje total (%) = Puntuación Total / Peso total		100.00 %	46.15 %

Al porcentaje obtenido como resultado de la evaluación del criterio de funcionalidad se le asigna una puntuación, como se muestra en la Tabla 21.

Tabla 20. Puntuación asignada al porcentaje obtenido

Porcentaje obtenido	Puntuación	Descripción
Menor a 65 %	1	Inaceptable
Entre 65 % y 80 %	2	Pobre
Entre 80 % y 90 %	3	Aceptable
Entre 90 % y 96 %	4	Muy Bueno
Mayor a 96 %	5	Excelente

De acuerdo a esa puntuación (Puntuación total no ponderada) se obtiene el resultado final de la puntuación de cada una de las herramientas evaluadas de acuerdo al peso que se le dio al criterio de funcionalidad que para este caso fue del 25%, esto se muestra en la Tabla 22.

Tabla 21. Resultado de Evaluación de Funcionalidad

Funcionalidad (25%)	Spamassassin	Dspam
Puntuación total no Ponderada (PNP)	5	1
Calificación Ponderada= PNP * 25%	1.25	0.25

La Tabla 23, muestra la evaluación del criterio de Calidad, La puntuación de las métricas de este criterio se obtuvo de acuerdo a las características de las herramientas teniendo en cuenta algunas referencias, para **Spamassassin** se tomo como referencia a [4] y para **Dspam** se utilizó como referencia a [5] y [6].

Tabla 22. Evaluación de Calidad

Calidad (20%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
Número de errores abiertos en los últimos 6 meses (desde 1/1/2009 hasta 30/6/2009)	20%	5	1	5	1
Número de errores corregidos en los últimos 6 meses (en comparación con el número de errores abiertos)	30%	4	1.2	3	0.9
Cantidad de errores críticos/P1 abiertos. (Se toman como críticos errores de prioridad 9 en escala del 1 al 10),	30%	4	1.2	4	1.2
Promedio de tiempo en que se solucionan los errores críticos / P1 en los últimos 6 meses	20%	4	0.8	1	0.2
Calificación No Ponderada (CNP)			4.2		3.3
Calificación Ponderada= CNP * 20%		0.84		0.66	

En la Tabla 24, se observa la evaluación de Rendimiento. Para la información de pruebas de rendimiento de **Spamassassin** se tomó como referencia a [8] y para **Dspam** a [9].

Tabla 23. Evaluación de Rendimiento

Rendimiento (20%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
Pruebas de rendimiento e informes de referencia disponibles	50%	3	1.5	3	1.5
Ajuste de rendimiento y configuración	50%	5	2.5	5	2.5
Calificación No Ponderada (CNP)			4		4
Calificación Ponderada= CNP * 20%		0.8		0.8	

La Tabla 25, muestra la evaluación del criterio Comunidad. Para **Spamassassin** se tomó como referencia a [10] para la lista de correo general y [11] para el número de contribuyentes. Para conseguir la información de **Dspam** se tomó como referencia a [12] y [13].

Tabla 24. Evaluación de Comunidad

Comunidad (15%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
El volumen medio de la lista de correo general en los últimos 6 meses	50%	4	2	1	0.5
Número de contribuyentes de código en los últimos 6 meses	50%	4	2	2	1
Calificación No Ponderada (CNP)			4		1.5
Calificación Ponderada= CNP * 15%		0.6		0.225	

En la Tabla 26, se muestra la evaluación del Criterio de Documentación, para evaluar este criterio, la información de **Spamassassin** se obtuvo de [16], [17] y [18], la información de documentación de **Dspam** se obtuvo de [19] y [20].

Tabla 25. Evaluación de Documentación

Documentación (10%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
Existencia de varios tipos de documentación	50%	5	2.5	5	2.5
Marco de contribución de usuario	50%	5	2.5	5	2.5
Calificación No Ponderada (CNP)			5		5
Calificación Ponderada= CNP * 10%		0.5		0.5	

La Tabla 27, muestra la evaluación del criterio de soporte, para obtener la información sobre **Spamassassin** se tomó como referencia a [21], [22], [23] y [24] y para **Dspam** se utilizó como referencia a [25] y [26].

Tabla 26. Evaluación de Soporte

Soporte (5%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
El volumen promedio de la lista de correo general en los últimos 6 meses	50%	5	2.5	4	2
Calidad de apoyo profesional	50%	5	2.5	5	2.5
Calificación No Ponderada (CNP)			5		4.5
Calificación Ponderada= CNP * 5%		0.25		0.225	

La Tabla 28, muestra la evaluación de Escalabilidad, para obtener la información de este criterio, para **Spamassassin** se tomo como referencia [28] y para **Dspam** a [19].

Tabla 27. Evaluación de Escalabilidad

Escalabilidad (5%)	Ponderación	Spamassassin		Dspam	
		Puntuación no Ponderada (PNP)	Puntuación Ponderada (PP)	PNP	PP
Despliegue de referencia	50%	3	1.5	3	1.5
Diseñado para escalabilidad	50%	5	2.5	5	2.5
Calificación No Ponderada (CNP)			4		4
Calificación Ponderada= CNP * 5%		0.2		0.2	

6. Resultados de la evaluación

El resultado final de la evaluación de cada criterio en cada una de las herramientas y la puntuación BRR se muestra en la Tabla 29.

Tabla 28. Puntuación BRR

Posición	Categoría	Spamassassin	Dspam
1	Funcionalidad	1.25	0.25
2	Calidad	0.84	0.66
3	Rendimiento	0.8	0.8
4	Comunidad	0.6	0.225
5	Documentación	0.5	0.5
6	Soporte	0.25	0.225
7	Escalabilidad	0.2	0.2
Puntuación BRR		4.44	2.86

De acuerdo a los resultados obtenidos se puede realizar un análisis que permite llegar a las siguientes conclusiones:

- En cuanto a funcionalidad se puede decir que la herramienta que mayor puntuación tuvo en este criterio es **Spamassassin**, esto se debe a que es la herramienta que presenta mayor rechazo de mensajes *spam* y porque no tuvo detección de falsos positivos. A **Spamassassin** y **Dspam** se les premia por usar redes colaborativas y a **Dspam** se le castiga por no manejar métodos heurísticos.
- De acuerdo a la evaluación y resultado de calidad, se puede observar que las dos herramientas presentan una buena calificación debido a que los errores que estas han presentado, se les ha dado una rápida solución.
- Se puede observar que la documentación y el soporte que tienen las herramientas permiten tener una alta puntuación, ya que estas tienen diversos tipos de documentación en los cuales los usuarios se pueden apoyar y además de esto los grupos de trabajo de cada herramienta brindan un gran soporte en cuanto a la solución de problemas.
- En cuanto a escalabilidad, todas las herramientas fueron pensadas en que fueran escalables, por lo tanto tienen una puntuación similar.
- La puntuación del criterio de Comunidad, varía entre las dos herramientas debido a que el número de contribuyentes en el desarrollo de las herramientas en los últimos meses difiere bastante.

De acuerdo a la puntuación BRR, la herramienta con mayor calificación es **Spamassassin**, por lo tanto es el filtro de contenido que se va a tener en cuenta para ser puesto en funcionamiento en el desarrollo de este proyecto.

REFERENCIAS BIBLIOGRÁFICAS

- [1] openbrr.org, «Business Readiness Rating for Open Source,» 2005.
- [2] Grupo de Nuevas Actividades Profesionales, «Correo Electronico y Spam,» [En línea]. Available: <http://es.scribd.com/doc/19049881/20/Medidas-de-lucha-contr-el-Spam>.
- [3] P. R.Thomas, «Anti_spam Comparison Report,» [En línea]. Available: https://www.westcoastlabs.com/downloads/productTestReport_0061/Anti-spam_Comparative_Report.pdf.
- [4] «Debian Package Tracking System,» [En línea]. Available: <http://packages.qa.debian.org/s/spamassassin.html>.
- [5] [En línea]. Available: <http://bugs.debian.org/cgi-bin/pkgreport.cgi?src=dspam>.
- [6] «sourceforge.net,» 2012. [En línea]. Available: http://sourceforge.net/tracker/?atid=1126467&group_id=250683&func=browse.
- [7] «www.policyd.org,» [En línea]. Available: <http://devlabs.linuxassist.net/projects/policyd/>.
- [8] «Apache Spamassassin Project,» [En línea]. Available: http://spamassassin.apache.org/tests_3_3_x.html.
- [9] «packages.debian.org,» [En línea]. Available: <http://packages.debian.org/source/testing/dspam>.
- [10] «wiki.apache.org,» [En línea]. Available: <http://wiki.apache.org/spamassassin/MailingLists>.
- [11] «sourceforge.net,» [En línea]. Available: http://sourceforge.net/project/memberlist.php?group_id=250683.
- [12] «sourceforge.net,» [En línea]. Available: http://sourceforge.net/mail/?group_id=250683.
- [13] «sourceforge.net,» [En línea]. Available: http://sourceforge.net/project/memberlist.php?group_id=250683.
- [14] «lists.policyd.org,» [En línea]. Available: <http://lists.policyd.org/mailman/listinfo>.
- [15] [En línea]. Available: <http://devlabs.linuxassist.net/projects/policyd>.
- [16] «spamassassin.apache.org,» [En línea]. Available: <http://spamassassin.apache.org/doc.html>.
- [17] «wiki.apache.org,» [En línea]. Available: <http://wiki.apache.org/spamassassin/WeLoveVolunteers>.
- [18] «packages.qa.debian.org,» [En línea]. Available: <http://packages.qa.debian.org/s/spamassassin.html>.
- [19] «dspam.nuclearelephant.com,» [En línea]. Available: <http://dspam.nuclearelephant.com/>.
- [20] [En línea]. Available: http://rtc.umn.edu/dspam/I_am_DSPAM_Summer_2011_Newsletter.pdf.
- [21] [En línea]. Available: <http://wiki.apache.org/spamassassin/MailingLists>.
- [22] [En línea]. Available: <http://gmane.org/plot->

- rate.php/plot.png?group=gmane.mail.spam.spamassassin.general&plot.png.
- [23] [En línea]. Available: <http://svn.apache.org/repos/asf/spamassassin/trunk/CREDITS>.
- [24] [En línea]. Available: http://egrojgomez.blogspot.com/2008_12_19_archive.html.
- [25] [En línea]. Available:
http://sourceforge.net/mailarchive/forum.php?forum_name=dspam-user.
- [26] [En línea]. Available: <http://sourceforge.net/projects/dspam/support>.
- [27] [En línea]. Available: http://lists.policyd.org/mailman/listinfo/users_lists.policyd.org.
- [28] [En línea]. Available: <http://lwn.net/Articles/87818/>.
- [29] «lists.policyd.org,» [En línea]. Available: <http://lists.policyd.org/mailman/listinfo>.
- [30] «lists.policyd.org,» [En línea]. Available: <http://lists.policyd.org/mailman/listinfo>.