

<https://info.nodo50.org/Usermin-Filtros-avanzados-usando-expresiones-regulares.html>



Usermin. Filtros avanzados usando expresiones regulares

- Nodo50 - Manuales -

Fecha de publicación en línea: Sábado 24 de mayo de 2008

Copyright © Nodo50 - Todos derechos reservados

Podemos crear un filtro *Basado en expresión regular*, que nos permite una gran flexibilidad a la hora de seleccionar correos para procesarlos con un filtro. Muy útil para el spam que cambia de destinatario y/o de asunto, o para reglas que no encajen con las otras opciones mas "normalitas"

Como hacer el filtro

Se crea igual que cualquier otro filtro, con la salvedad de que en la parte de **Condición** hay que seleccionar **Basado en expresión regular**, que es la última del bloque **Condición**

[<https://info.nodo50.org/local/cache-vignettes/L400xH29/expresion-regular-detalle-56806.png>]

A la derecha nos aparece un recuadro donde tendremos que poner la expresión regular, Debajo del recuadro a rellenar nos aparece un **tick** que si dejamos en blanco significa que se buscará la coincidencia en las cabeceras (*From* , *Subject* ...) y si la marcamos se aplicará al *cuerpo*, es decir el texto del mensaje.

¿Que es una expresión regular?

Una expresión regular es una frase normal, pero alguno, o incluso todos, de los caracteres tiene un significado especial que permite variaciones a la hora de hacer las comprobaciones. El mejor ejemplo es el *, en cualquier buscador en GNU/Linux, Windows o Mac puedes buscar todos los archivos que empiecen por *HOLA* buscando *HOLA**. Aunque como veremos el * funciona un poco diferente aquí.

Pueden llegar a ser muy complejas, pero eso no significa que necesitemos explotar todo su potencial para resolver nuestro problema, aquí os explicamos lo más básico, pero es lo necesario para hacer un filtro de correo más complejo de lo normal.

La teoría

Las expresiones regulares se utilizan mucho en informática, en nuestro contexto consideraremos que lo que escribimos en ese cuadro es una expresión regular que se va a aplicar línea por línea hasta encontrar una coincidencia o llegar al final del correo electrónico.

Primero hay que conocer los caracteres especiales y su significado. Para ello hay que usar un anglicismo, *machear* a partir de ahora hará referencia a la frase, o conjunto de frases, que coinciden con nuestra expresión regular. Por ejemplo *HOLA.** machea *HOLA*, *HOLA1*, *HOLAadr3*, *HOLA7* ...

Es importante destacar que se tiene en cuenta la diferencia entre mayúsculas y minúsculas. y que estos caracteres especiales se pueden combinar entre sí.

Caracter	Significado	Ejemplo
<i>^expr</i>	machea el principio de la línea	<i>^a</i> machea <i>adios</i> , <i>aula</i> , <i>a</i> ...
<i>expr\$</i>	machea el final de la línea	<i>a\$</i> machea <i>petra</i> , <i>hola</i> , <i>a</i>

Usermin. Filtros avanzados usando expresiones regulares

.	Machea cualquier carácter, una sola vez	<i>a.b</i> machea <i>abb acb a3b aAb ...</i> pero no <i>a34b</i>
*	El carácter anterior puede repetirse 0 o más veces	<i>ab*</i> machea <i>a, ab, abb, abbb ...</i>
+	El carácter anterior puede repetirse 1 o más veces	<i>ab+</i> machea <i>ab, abb, abbb ...</i>
?	El carácter anterior puede aparecer o no	<i>ab?</i> machea <i>a y b</i>
[<i>expr</i>]	Machea cualquier carácter que se indique dentro de los corchetes, acepta que se le indiquen rangos, como [<i>a-z</i>] que incluye todas las letras minúsculas de la <i>a</i> a la <i>z</i>	[<i>abc</i>] Machea <i>a o b o c</i>
[^ <i>expr</i>]	Machea cualquier carácter que NO esté dentro de los corchetes	[^ <i>abc</i>] Machea cualquier carácter que no sea <i>a o b o c</i>
<i>expr1</i> <i>expr2</i>	Machea lo coincidente con <i>expr1</i> o lo coincidente con <i>expr2</i>	^ <i>A ^a</i> Machea las expresiones que empiecen por <i>a</i> y las que empiecen por <i>A</i>
(<i>expresión</i>	Agrupar el efecto de los operadores <i>*, +, ? y </i> y lo aplica a todo <i>expresión</i>	<i>A(ab)*</i> machea <i>A, Aab, Aabab, Aababab ...</i>
{ <i>n</i> }	Espera encontrar lo mismo que para el paréntesis número <i>n</i>	(<i>ABC</i>){ <i>x</i> } <i>1</i> machea <i>AxA, BxB o CxC</i> pero nada más

La práctica

Todo lo anterior nos proporciona mucha flexibilidad a la hora de hacer un filtro pero, ¿cómo empezar?. Lo primero es identificar los elementos que son hijos. Por ejemplo, si quiero buscar frases que coincidan con:

Querido Pepe:

Donde *Pepe* puede ser cualquier nombre, empezaré con lo más fácil que es darse cuenta de que termina en *;*, por lo que una fácil sería:

:\$

Aunque esto machearía casi cualquier cosa. También podemos ir a lo seguro:

^Querido

Aunque para esto puede que no necesitamos meternos en expresiones regulares, ya que es casi crear un filtro que busque por las cabeceras que empiecen por *Querido*. (La única diferencia reside en que las expresiones regulares las puedo aplicar al cuerpo)

Sabemos que todos los nombres propios empiezan por una letra mayúscula [*A-Z*] y un montón de letras minúsculas [*a-z*], que pueden repetirse cuantas veces quieran en distintas combinaciones [*a-z*]*, por lo que *Pepe* se puede "traducir" por [*A-Z*][*a-z*]*, y finalmente nuestra expresión sería:

Querido [A-Z][a-z]*:

Un ejemplo simple pero útil, cuando buscamos una frase muy concreta que puede estar en medio de un texto:

Hola Jose, necesito que me prestes la chaqueta para esta tarde

Si estoy buscando todo lo relacionado con *necesito que me prestes*

.*necesito que me prestes.*

También pueden combinarse:

Hola [A-Z][a-z]*, .*necesito que me prestes.*

De distintas maneras dependiendo de las necesidades, si queremos interceptar un mensaje que venga con un saludo en plan *Querido* o bien alguien que nos pide cosas:

Querido [A-Z]([a-z]*:|.*necesito que me prestes.*

Terminando

Una vez que hemos descubierto cual es la expresión regular que más se ajusta a lo que queremos la introducimos en su cajita correspondiente, y en **Acción si la condición se cumple** seleccionamos lo que más nos interesa, *Descartar*, *Reenviar*, *Guardar en una carpeta ...*

Si estas empezando con expresiones regulares **te recomendamos que Guardes en una carpeta** por si acaso tu filtro machea algún correo legítimo (y te recordamos que la carpeta debe estar creada antes de crear el filtro [https://info.nodo50.org/local/cache-vignettes/L400xH229/crear-filtro-3fa18.png])

Y hacemos click en **Crear** para terminar y nos devuelve a la página general de filtros, donde podremos ver nuestro nuevo filtro

[https://info.nodo50.org/local/cache-vignettes/L400xH50/filtro-6d58c.png]