Como obstruir o tipo de conteúdo baseou jogos de caracteres

Índice

Introdução

Este documento descreve como escrever e configurar um filtro a fim detectar e tomar a ação em jogos de caracteres baseados tipo de conteúdo na ferramenta de segurança do email de Cisco (ESA). O seguinte documento pode ser usado para detectar os caráteres baseados na linguagem estrangeiros considerados em mensagens do Spam.

Informações de Apoio

Os administradores ESA podem receber um influxo das mensagens do correio que contêm as línguas estrangeiras baseadas caráter que não são correio legítimo para sua empresa ou domínios. Uma maneira de endereçar do ESA, nós temos três opções:

3. Escreva um filtro usando a língua da mensagem da circunstância. (Esta opção é uns novos recursos para a Segurança 10.0.0-203 do email de AsyncOS e mais novo.)

Como obstruir o tipo de conteúdo baseou jogos de caracteres

Escreva um filtro para detectar o tipo de conteúdo

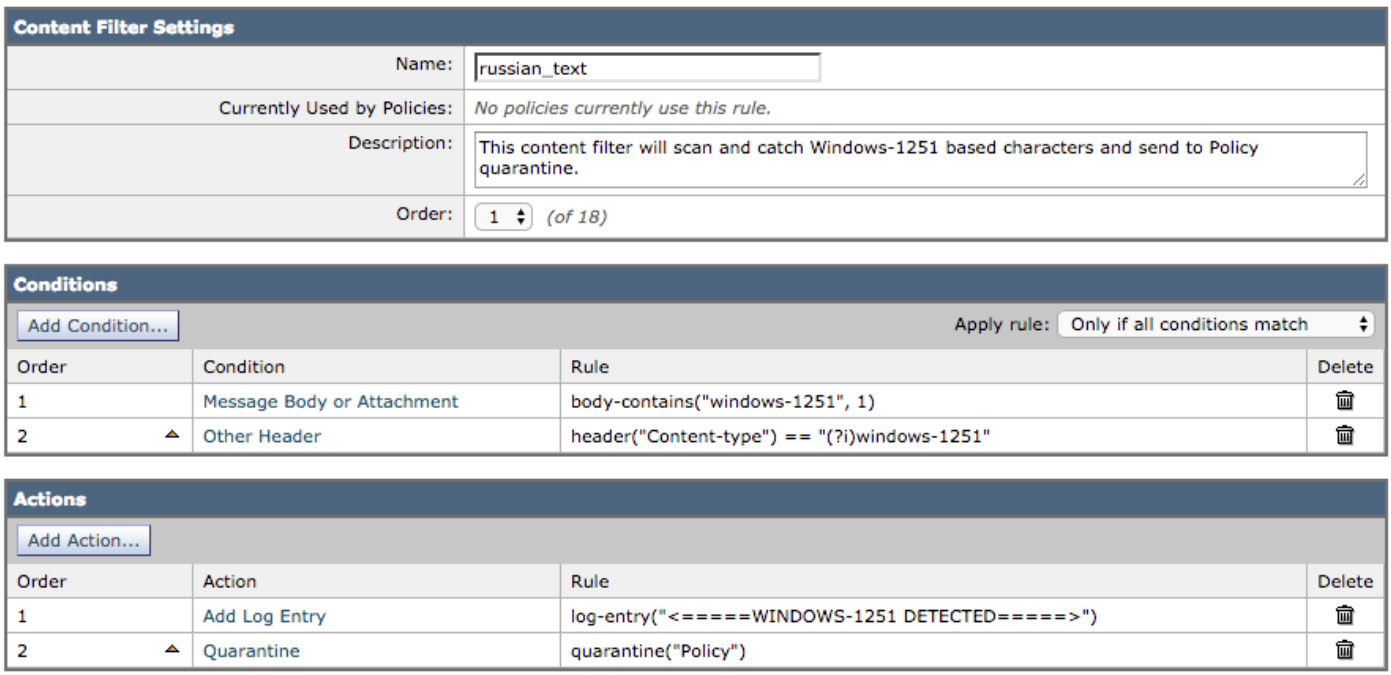
A primeira opção é para que o administrador escreva e configure um filtro, e associa-o a uma política do correio, como necessário.

> **Note**: Escrever e configurar este filtro como um filtro da mensagem podem ser recurso-caras a fim fazer a varredura do corpo dos email para os jogos de caracteres.

> **Note**: Configurando isto como um filtro satisfeito é sugerido fortemente, como os filtros satisfeitos ocorrem após a exploração do anti-Spam. Contudo, isto pode ser escrito e

configurado como um filtro da mensagem, se necessário.

O exemplo seguinte levará em consideração uma mensagem do correio contém caráteres baseados (cirílicos) do russo através do jogo de caracteres baseado Windows-1251. Escrito como um filtro satisfeito:



Content Filter Settings

Name:	russian_text
Currently Used by Policies:	*No policies currently use this rule.*
Description:	This content filter will scan and catch Windows-1251 based characters and send to Policy quarantine.
Order:	1 (of 18)

Conditions

Add Condition... Apply rule: Only if all conditions match

Order	Condition	Rule	Delete
1	Message Body or Attachment	body-contains("windows-1251", 1)	
2	Other Header	header("Content-type") == "(?i)windows-1251"	

Actions

Add Action...

Order	Action	Rule	Delete
1	Add Log Entry	log-entry("<=====WINDOWS-1251 DETECTED=====>")	
2	Quarantine	quarantine("Policy")	

O email do teste usado conterá o seguinte no corpo do email:

```
Russian uses , , , , o, , , , ,  as vowels. You could create a message filter set to "Matches
any of the following" that test whether "Body" "contains" "", "Body" "contains" "" and so forth
until you covered all of the vowels. Ssince English also uses "a" , "e" , "o", and "y" letters
don't test for them. The reason for "Matches any of the following" is to logically OR them - you
want the action to take place if any of those letters are found.
```

Com o filtro do índice configurado como acima, os logs do correio gravariam similar ao seguinte:

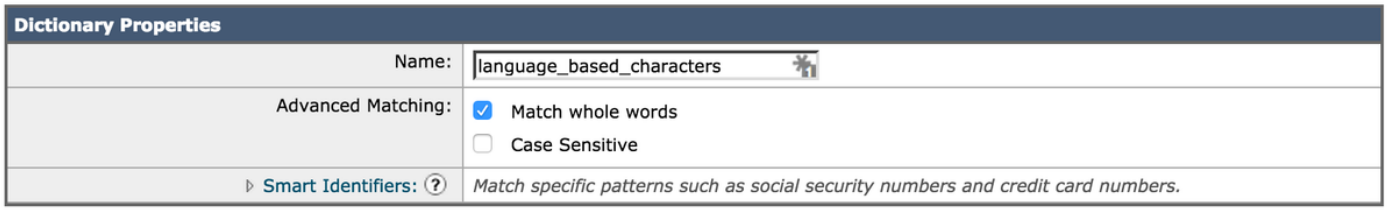
```
Thu Sep 10 14:50:09 2015 Info: Start MID 164993 ICID 266729
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 From: <end_user@test.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 RID 0 To: <recpient@my_co.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 using engine: SPF Verdict Cache using cached verdict
Thu Sep 10 14:50:09 2015 Info: MID 164993 Message-ID '<7A961F85-A5F1-413F-87CB-
C31D2E5605EC@my_co.com>'
Thu Sep 10 14:50:09 2015 Info: MID 164993 Subject 'russian test'
Thu Sep 10 14:50:09 2015 Info: MID 164993 ready 2302 bytes from <end_user@test.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 matched all recipients for per-recipient policy
DEFAULT in the inbound table
Thu Sep 10 14:50:09 2015 Info: MID 164993 AMP file reputation verdict : CLEAN
Thu Sep 10 14:50:09 2015 Info: MID 164993 using engine: GRAYMAIL negative
Thu Sep 10 14:50:09 2015 Info: MID 164993 Custom Log Entry: <====== WINDOWS-1251 DETECTED
======>
Thu Sep 10 14:50:09 2015 Info: MID 164993 quarantined to "Policy" (content filter:russian_text)
Thu Sep 10 14:50:09 2015 Info: Message finished MID 164993 done
```

Outros línguas e jogos de caracteres podem ser usados. Veja por favor a seção de referências para a informação adicional.

Escreva um filtro para prover um dicionário baseado caráter

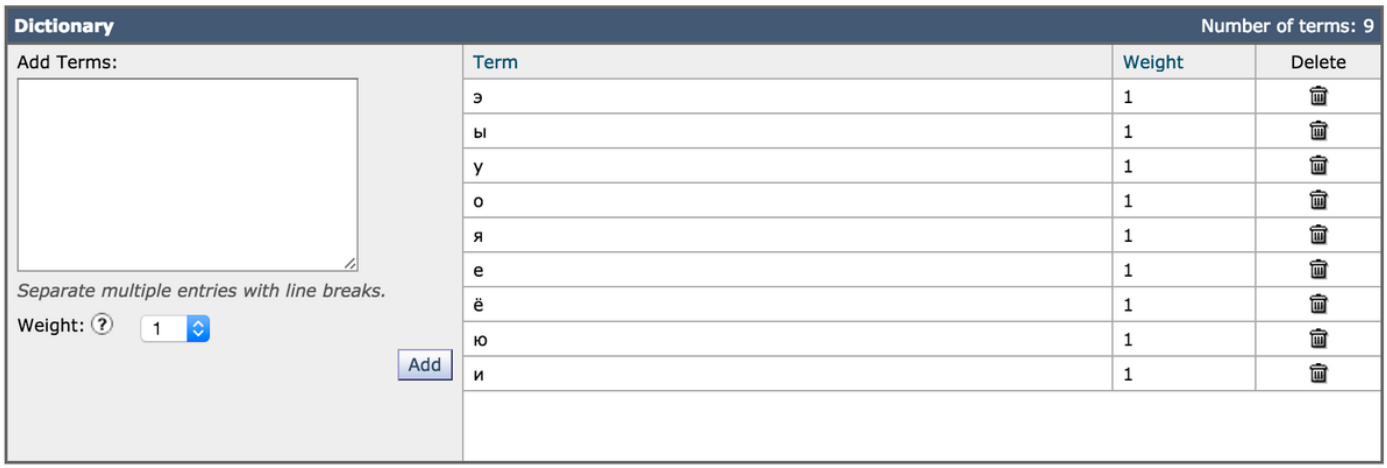
A segunda opção é adicionar a lista de jogos de caracteres a um arquivo de texto do dicionário e referir aquela no filtro.

Exemplo de adicionar os caráteres ao dicionário:



Dictionary Properties

Name:	language_based_characters
Advanced Matching:	☑ Match whole words ☐ Case Sensitive
▷ Smart Identifiers:	*Match specific patterns such as social security numbers and credit card numbers.*



Dictionary — Number of terms: 9

Add Terms:

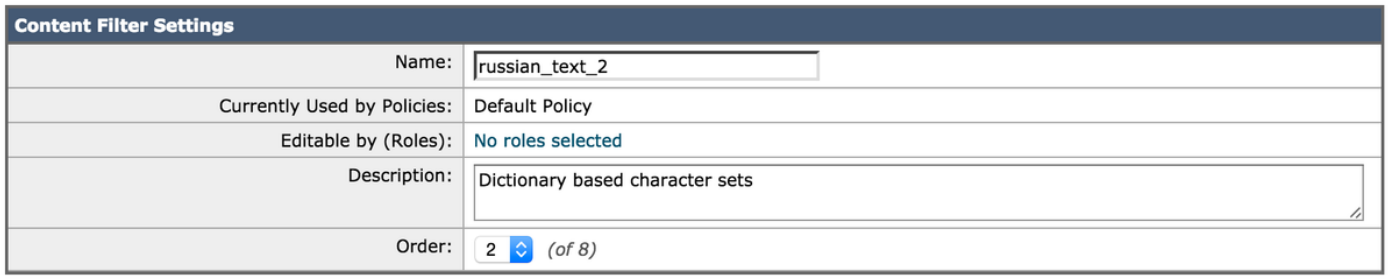
Separate multiple entries with line breaks.

Weight: 1

Add

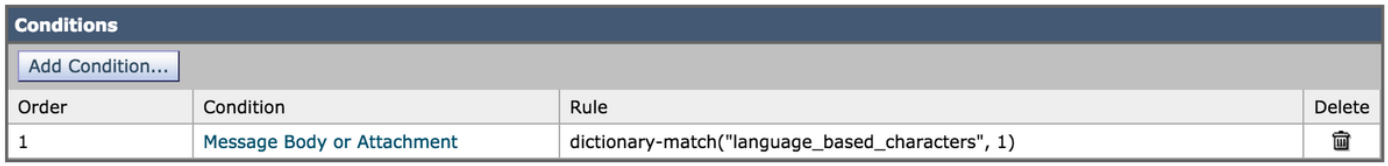
Term	Weight	Delete
э	1	
ы	1	
у	1	
о	1	
я	1	
е	1	
ё	1	
ю	1	
и	1	

Os caráteres são atribuídos agora ao dicionário e o dicionário próprio é provido nos artigos da circunstância para o filtro:



Content Filter Settings

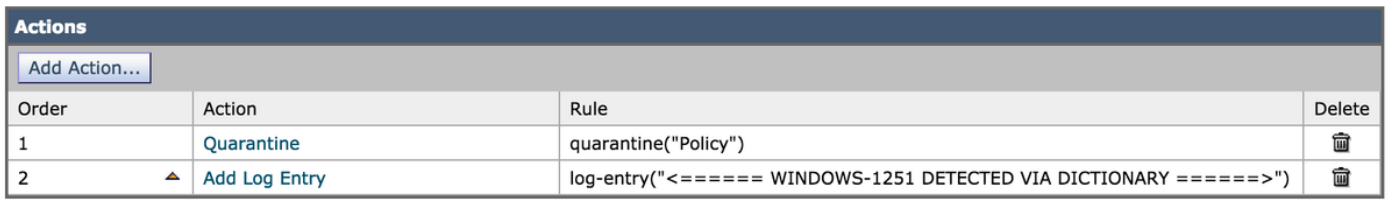
Name:	russian_text_2
Currently Used by Policies:	Default Policy
Editable by (Roles):	No roles selected
Description:	Dictionary based character sets
Order:	2 *(of 8)*



Conditions

Add Condition...

Order	Condition	Rule	Delete
1	Message Body or Attachment	dictionary-match("language_based_characters", 1)	



Actions

Add Action...

Order	Action	Rule	Delete
1	Quarantine	quarantine("Policy")	
2	Add Log Entry	log-entry("<====== WINDOWS-1251 DETECTED VIA DICTIONARY ======>")	

Usando o mesmo email do teste que acima, contém o seguinte no corpo do email:

```
Russian uses , , , , o, , , , ,  as vowels. You could create a message filter set to "Matches
any of the following" that test whether "Body" "contains" "", "Body" "contains" "" and so forth
until you covered all of the vowels. Ssince English also uses "a" , "e" , "o", and "y" letters
don't test for them. The reason for "Matches any of the following" is to logically OR them - you
want the action to take place if any of those letters are found.
```

Com o filtro do índice configurado como acima usando a condição do fósforo de dicionário, os logs do correio gravariam similar ao seguinte:

```
Thu Sep 10 15:26:08 2015 Info: Start MID 164995 ICID 266737
Thu Sep 10 15:26:08 2015 Info: MID 164995 ICID 266737 From: <end_user@test.com>
Thu Sep 10 15:26:08 2015 Info: MID 164995 ICID 266737 RID 0 To: <recpient@my_co.com>
Thu Sep 10 15:26:08 2015 Info: MID 164995 using engine: SPF Verdict Cache using cached verdict
Thu Sep 10 15:26:08 2015 Info: SPF Verdict Cache cache status: hits = 6, misses = 4, expires =
1, adds = 4, seconds saved = 0.50, total seconds = 0.85
Thu Sep 10 15:26:08 2015 Info: MID 164995 Message-ID '<BCC88307-EB91-476E-8732-
334E9EE84EC8@my_co.com>'
Thu Sep 10 15:26:08 2015 Info: MID 164995 Subject 'russian test 3'
Thu Sep 10 15:26:08 2015 Info: MID 164995 ready 2316 bytes from <end_user@test.com>
Thu Sep 10 15:26:08 2015 Info: MID 164995 matched all recipients for per-recipient policy
DEFAULT in the inbound table
Thu Sep 10 15:26:08 2015 Info: MID 164995 AMP file reputation verdict : CLEAN
Thu Sep 10 15:26:08 2015 Info: MID 164995 using engine: GRAYMAIL negative
Thu Sep 10 15:26:08 2015 Info: MID 164995 Custom Log Entry: <====== WINDOWS-1251 DETECTED VIA
DICTIONARY ======>
Thu Sep 10 15:26:08 2015 Info: MID 164995 quarantined to "Policy" (content
filter:russian_text_2)
Thu Sep 10 15:26:08 2015 Info: Message finished MID 164995 done
```

Escreva um filtro satisfeito usando da “a condição da língua mensagem”

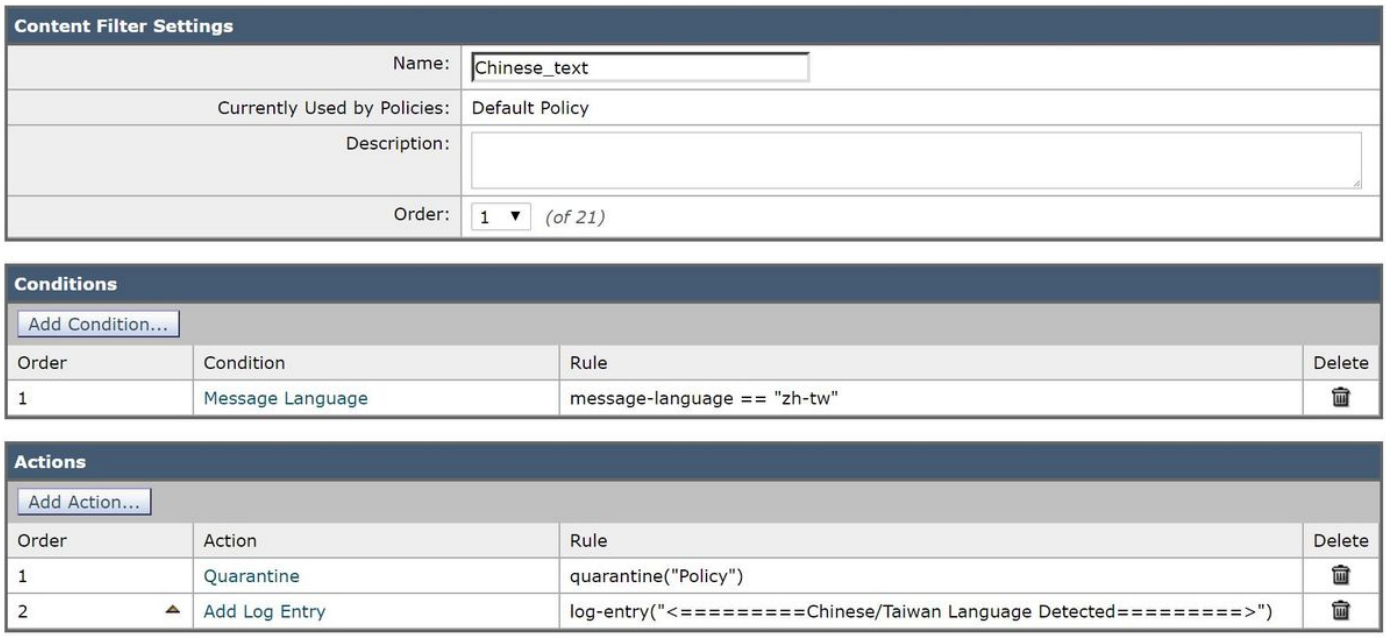
A terceira opção é usar da “a condição da língua mensagem”. O ESA usa o motor incorporado da detecção da língua para detectar a língua em uma mensagem. O dispositivo extrai o assunto e o corpo da mensagem e passa-o ao motor da detecção da língua.

O motor da detecção da língua determina a probabilidade de cada língua no texto extraído e passa-a de volta ao dispositivo. O dispositivo considera a língua com a probabilidade a mais alta como a língua da mensagem. O dispositivo considera a língua da mensagem como “indeterminado” em uma das seguintes encenações:

- Se a língua detectada não é apoiada pelo ESA
- Se o dispositivo é incapaz de detectar a língua da mensagem
- Se o tamanho total do texto extraído enviado ao motor da detecção da língua é menos do que bytes dos 50 pés.

Note: Esta opção é uns novos recursos para a Segurança 10.0.0-203 do email de AsyncOS e mais novo.

O exemplo seguinte levará em consideração uma mensagem do correio que contenha o chinês/o jogo de caracteres baseado Taiwan. Escrito como um filtro satisfeito:



Com o filtro do índice configurado como acima, os logs do correio gravariam similar ao seguinte:

```
Tue Feb 28 06:53:18 2017 Info: Start MID 481 ICID 27
Tue Feb 28 06:53:18 2017 Info: MID 481 ICID 27 From: <end_user@test.com>
Tue Feb 28 06:53:18 2017 Info: MID 481 ICID 27 RID 0 To: <recipient@my_co.com>
Tue Feb 28 06:53:18 2017 Info: MID 481 Subject 'Chinese text test'
Tue Feb 28 06:53:18 2017 Info: MID 481 ready 1047 bytes from <end_user@test.com>
Tue Feb 28 06:53:18 2017 Info: MID 481 matched all recipients for per-recipient policy DEFAULT
in the inbound table
Tue Feb 28 06:53:18 2017 Info: MID 481 interim verdict using engine: CASE spam negative
Tue Feb 28 06:53:18 2017 Info: MID 481 using engine: CASE spam negative
Tue Feb 28 06:53:18 2017 Info: MID 481 interim AV verdict using Sophos CLEAN
Tue Feb 28 06:53:18 2017 Info: MID 481 antivirus negative
Tue Feb 28 06:53:18 2017 Info: MID 481 using engine: GRAYMAIL negative
Tue Feb 28 06:53:18 2017 Info: MID 481 Message language: 'Chinese/Taiwan'
Tue Feb 28 06:53:18 2017 Info: MID 481 Custom Log Entry: <==========Chinese/Taiwan Language
Detected==========>
Tue Feb 28 06:53:18 2017 Info: MID 481 Outbreak Filters: verdict negative
Tue Feb 28 06:53:18 2017 Info: MID 481 quarantined to "Policy" (content filter:Chinese_text)
Tue Feb 28 06:53:18 2017 Info: Message finished MID 481 done
```

Referências

- Microsoft fornece nomes do jogo de caracteres (*nome do .NET*) em seus identificadores da página de código que podem ser providos ao escrever e ao configurar filtros.

  Note: As páginas de código ANSI podem ser diferentes em computadores diferentes, ou podem ser mudadas para um único computador, conduzindo ao corrompimento de dados. Para os resultados os mais consistentes, os aplicativos devem usar Unicode, tal como UTF-8 ou UTF-16, em vez de uma página de código específica.

- Mozillazine fornece detalhes detalhados para o tipo de conteúdo: encabeçamento, letras estrangeiras, palavras estrangeiras, e mais, em seu artigo para o Spam da língua estrangeira

Informações Relacionadas

- Homoglyph avançou ataques do phishing
- Cisco envia por correio eletrónico guias do utilizador final da ferramenta de segurança
- Suporte Técnico e Documentação - Cisco Systems