如何阻止基於內容型別的字符集

目錄

簡介

本文檔介紹如何編寫和配置篩選器，以便檢測基於內容型別的字符集並對思科郵件安全裝置(ESA)上的字符集執行操作。以下文檔可用於檢測垃圾郵件中出現的基於外語字元。

背景資訊

ESA管理員可能會收到大量包含基於字元的外語郵件的郵件，這些郵件對於他們的公司或域來說不是合法的郵件。 從ESA解決這個問題的方法有三種：

3. 使用條件消息語言編寫過濾器。（此選項是AsyncOS電子郵件安全10.0.0-203及更高版本的新功能。）

如何阻止基於內容型別的字符集

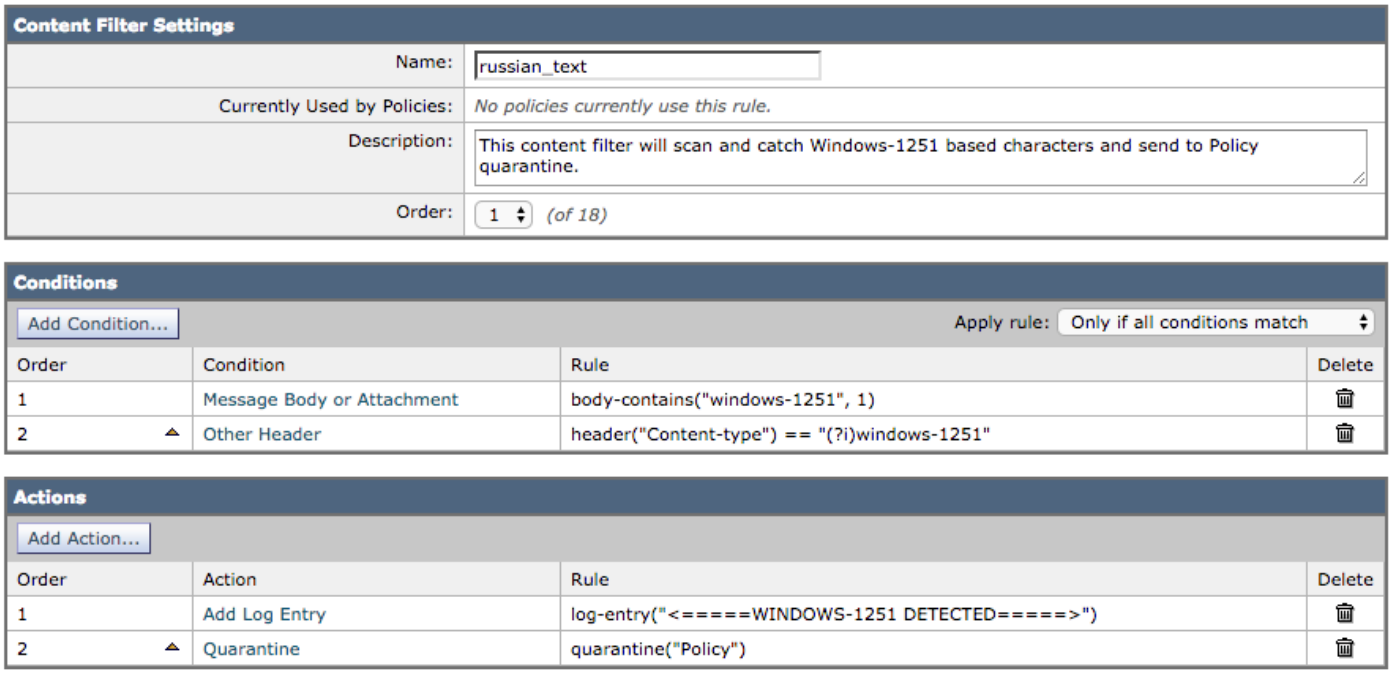
寫入篩選器以檢測內容型別

第一個選項供管理員編寫和配置過濾器，並根據需要將其與郵件策略相關聯。

> **注意**：編寫和配置此過濾器作為郵件過濾器可能需要耗費大量資源，以便掃描郵件正文中的字符集。

> **附註**：強烈建議將此項配置為內容過濾器，因為內容過濾器會在反垃圾郵件掃描之後發生。但是，如果需要，可以將其寫入並配置為消息過濾器。

以下示例將考慮通過基於Windows-1251的字符集傳送包含俄語(西里爾文)字元的郵件。 寫入內容過濾器：



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")	

使用的測試電子郵件將在郵件正文中包含以下內容：

```
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.
```

按照上述方式配置內容過濾器後，郵件日誌記錄將類似於以下內容：

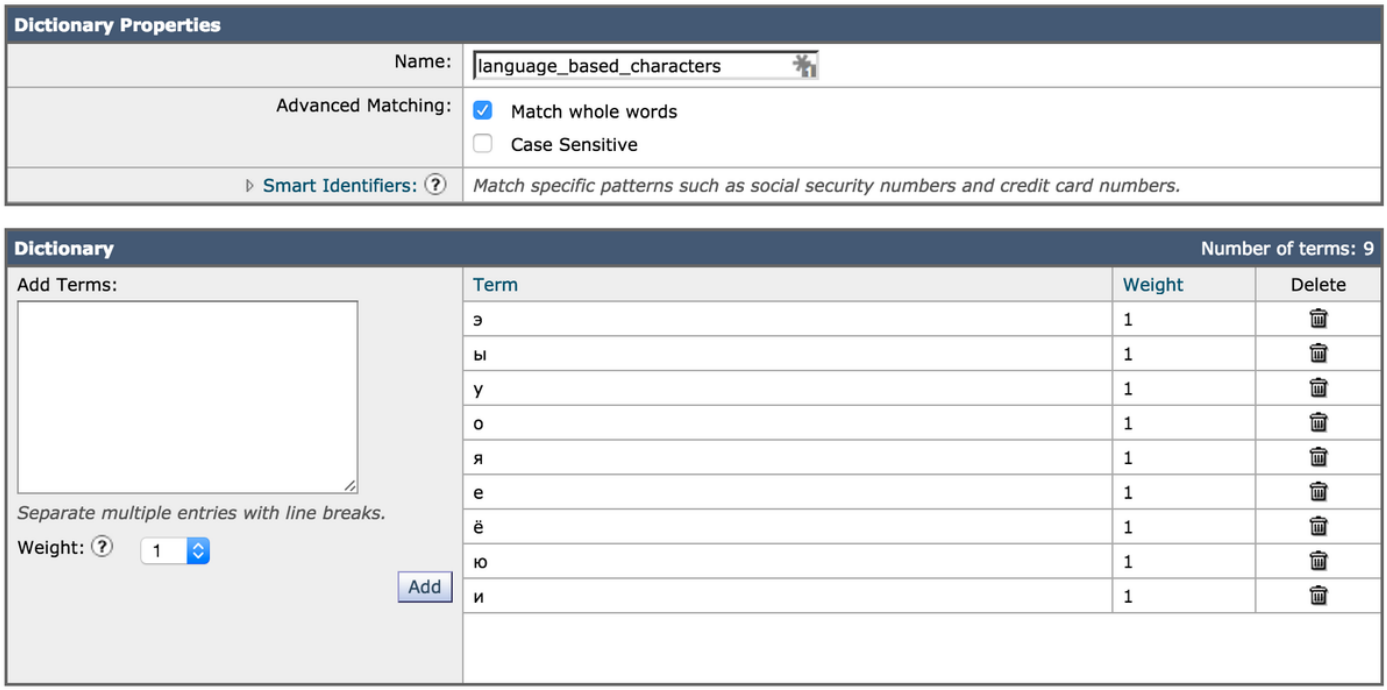
```
Thu Sep 10 14:50:09 2015 Info: Start MID 164993 ICID 266729
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 From: <end_user@test.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 RID 0 To: <recpient@my_co.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 using engine: SPF Verdict Cache 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
```

可以使用其他語言和字符集。 請參見參考部分以瞭解更多資訊。

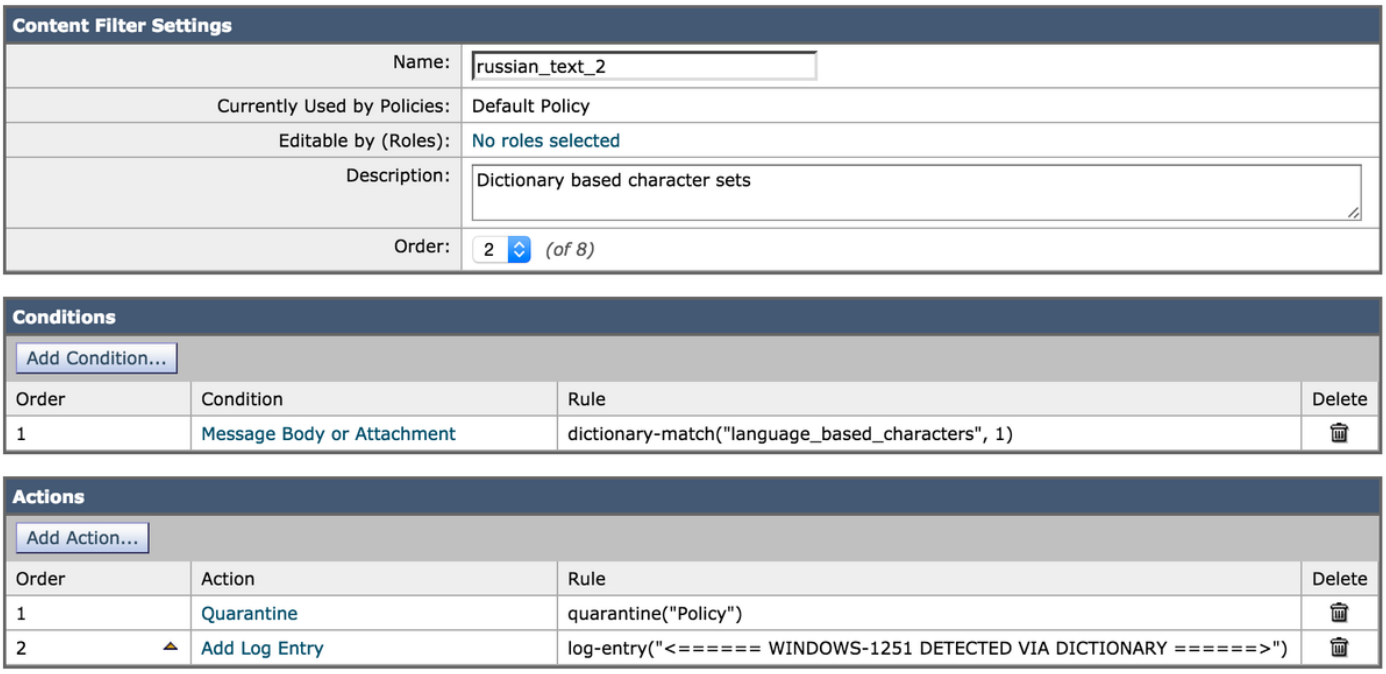
寫入篩選器以引用基於字元的字典

第二個選項是將字符集清單新增到字典文本檔案，並在過濾器中引用該清單。

將字元新增到字典的示例：



現在，字元將分配給詞典，並且詞典本身在過濾器的條件項中引用：



使用與上面相同的測試電子郵件，該電子郵件正文中包含以下內容：

```
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.
```

使用詞典匹配條件按上述方式配置內容過濾器後，郵件日誌記錄將類似於以下內容：

```
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
```

使用「消息語言」條件編寫內容過濾器

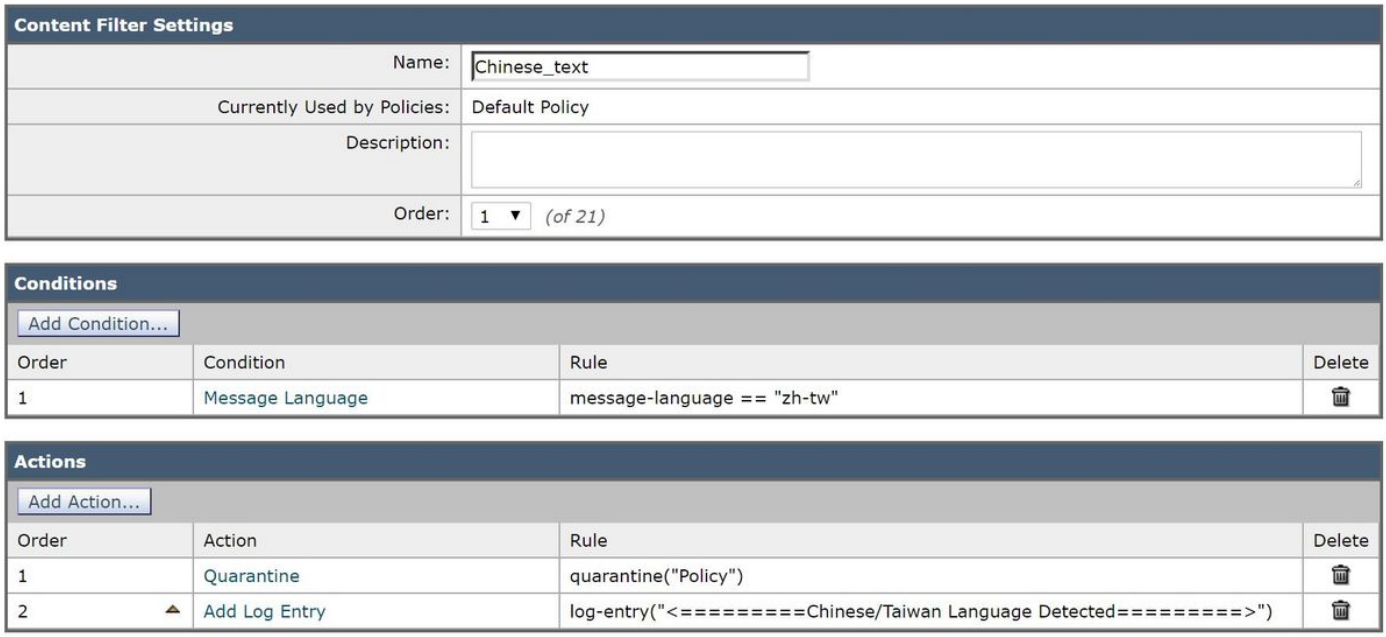
第三個選項是使用「消息語言」條件。ESA使用內建語言檢測引擎來檢測消息中的語言。裝置提取主題和郵件正文，並將其傳遞到語言檢測引擎。

語言檢測引擎確定提取的文本中每種語言的概率並將其傳回裝置。裝置將概率最高的語言視為消息的語言。在以下情況之一中，裝置將消息的語言視為「未確定」：

- 如果ESA不支援檢測到的語言
- 如果裝置無法檢測消息的語言
- 如果傳送到語言檢測引擎的提取文本的總大小小於50位元組。

附註：此選項是AsyncOS電子郵件安全10.0.0-203及更高版本的新功能。

以下示例將考慮包含基於中文/台灣字符集的郵件。 寫入內容過濾器：



按照上述方式配置內容過濾器後，郵件日誌記錄將類似於以下內容：

```
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
```

參考資料

- Microsoft提供字符集名稱(*.NET名稱*在其中) 內碼表識別符號 在寫入和配置過濾器時可以參考的。

  注意:ANSI內碼表可能在不同電腦上不同，或者可以針對單個電腦進行更改，從而導致資料損壞。為了獲得最一致的結果，應用程式應使用Unicode，如UTF-8或UTF-16，而不是特定的內碼表。

- 莫齊拉津 提供Content-type的深層詳細資訊：在文章中的標題、外來字母、外來單詞等 外語垃圾郵件

相關資訊

- **同形高級網路釣魚攻擊**
- **思科電子郵件安全裝置最終使用手冊**
- **技術支援與文件 - Cisco Systems**