如何拦截内容类型根据字符集

目录

简介

本文描述如何写入和配置过滤器为了检测和采取在内容类型基于字符集的行动在思科电子邮件安全工具(ESA)。 以下文档可以用于检测在垃圾邮件消息看到的外国以语言为基础的字符。

背景信息

ESA管理员可能接收的邮件消息汇集包含字符基于外语不是他们的公司的合法邮件或域。 一种方式从ESA寻址，我们有三个选项：

3. 使用情况消息语言，写入过滤器。(此选项是AsyncOS电子邮件安全的10.0.0-203一新特性和更新。)

如何拦截内容类型根据字符集

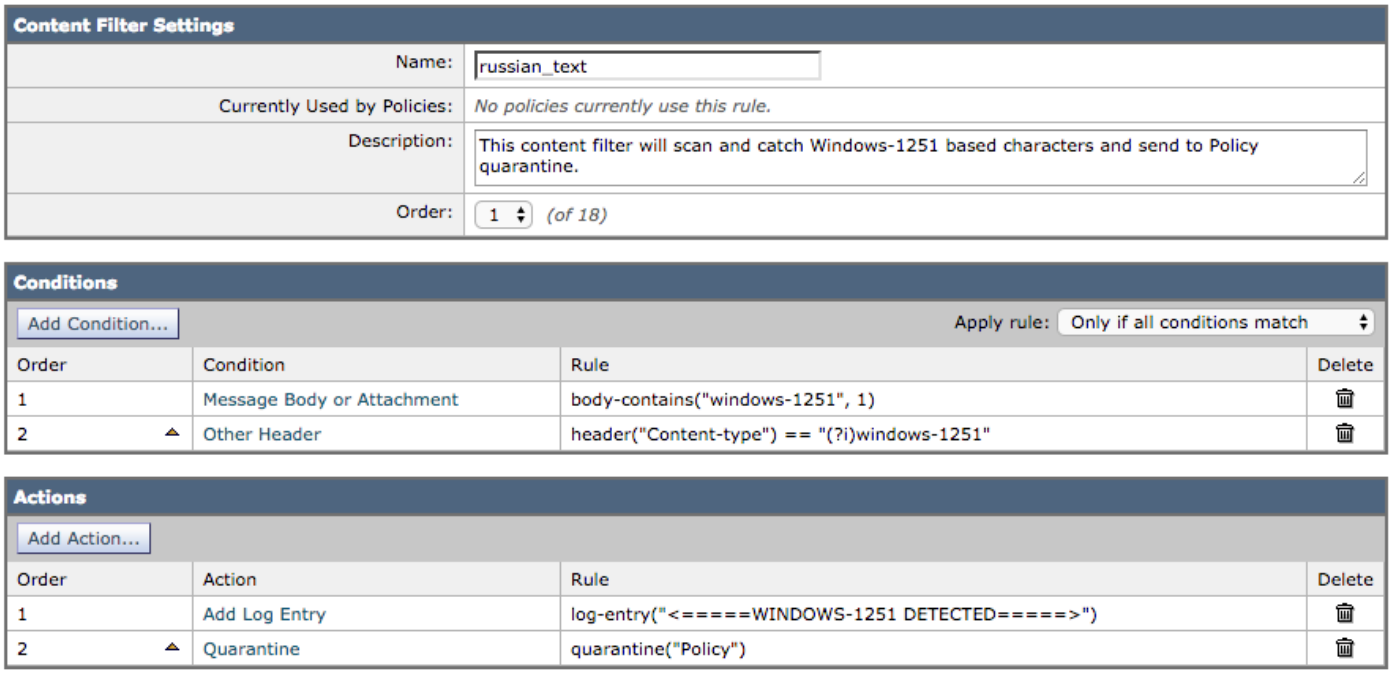
写入过滤器检测内容类型

第一个选项是为了配置的管理员能写入和过滤器，并且关联它对邮件策略，当必要时。

> **Note**: 写入和配置此过滤器作为消息过滤器可能是资源昂贵为了扫描电子邮件正文字符集的。

> **Note**:配置此，内容过滤器强烈建议，内容过滤器在反垃圾邮件扫描以后发生。 若需要然而，这可以写入和配置作为消息过滤器。

以下示例将考虑到邮件消息通过Windows-1251基于字符集包含俄罗斯(西里尔字母)基于字符。 写入，内容过滤器：



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")	

使用的测验电子邮件将包含以下在电子邮件的正文：

```
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.
```

使用如上所述配置的内容过滤器，邮件日志将记录类似于以下：

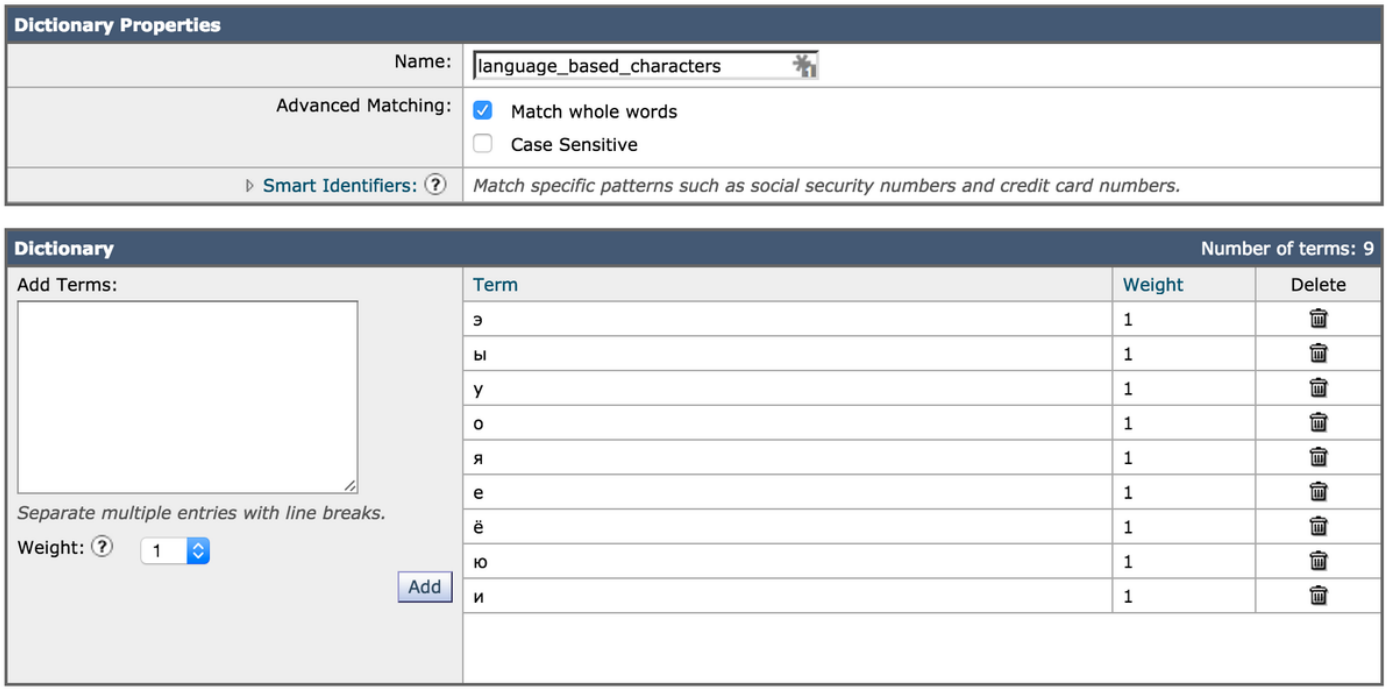
```
Thu Sep 10 14:50:09 2015 Info: Start MID 164993 ICID 266729
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 From: <end_user@test.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 RID 0 To: <recpient@my_co.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 using engine: SPF Verdict Cache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

可以使用其他语言和字符集。 请参阅References部分关于其他信息。

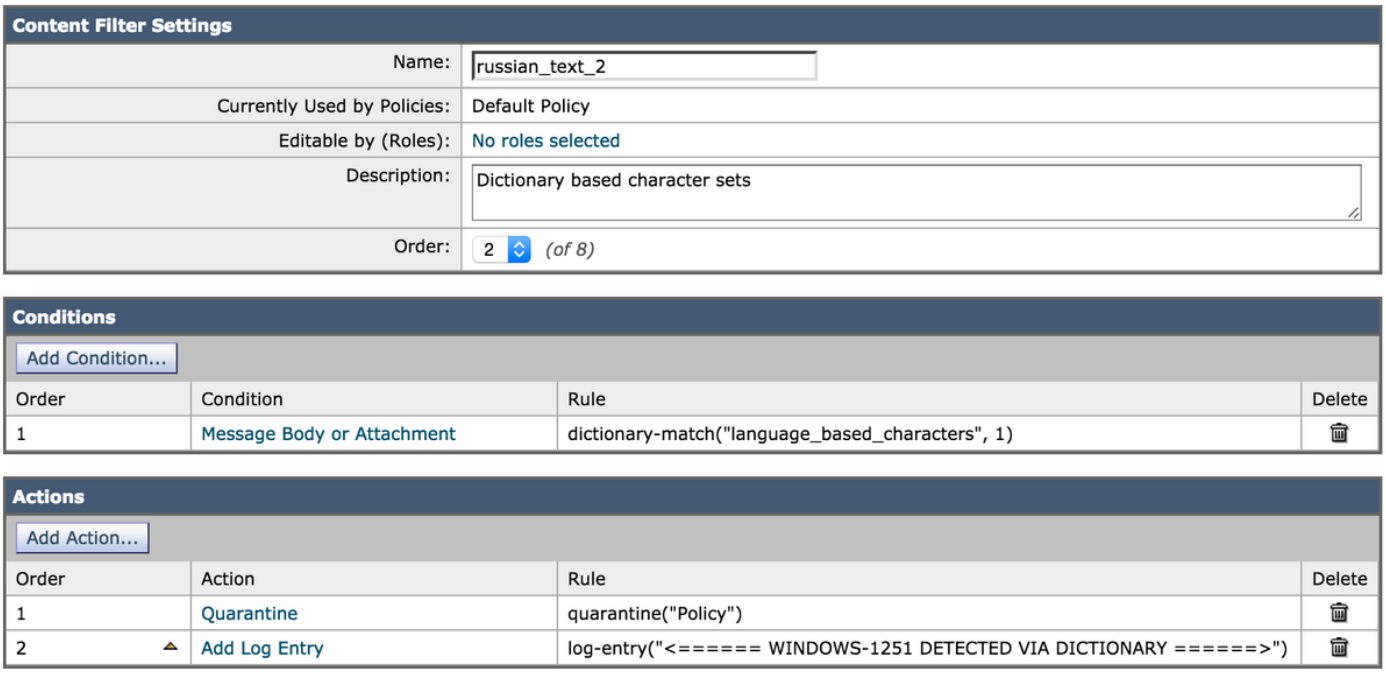
写入过滤器参考字符基于字典

第二个选项是添加字符集列表到字典文本文件和参考那在过滤器。

添加字符示例到字典：



字符当前分配到字典，并且字典被参考过滤器的情况项目：



使用测验电子邮件和上述一样，它包含以下在电子邮件的正文：

```
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.
```

使用如上所述配置的内容过滤器使用字典匹配情况，邮件日志将记录类似于以下：

```
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“消息语言”情况，写入内容过滤器

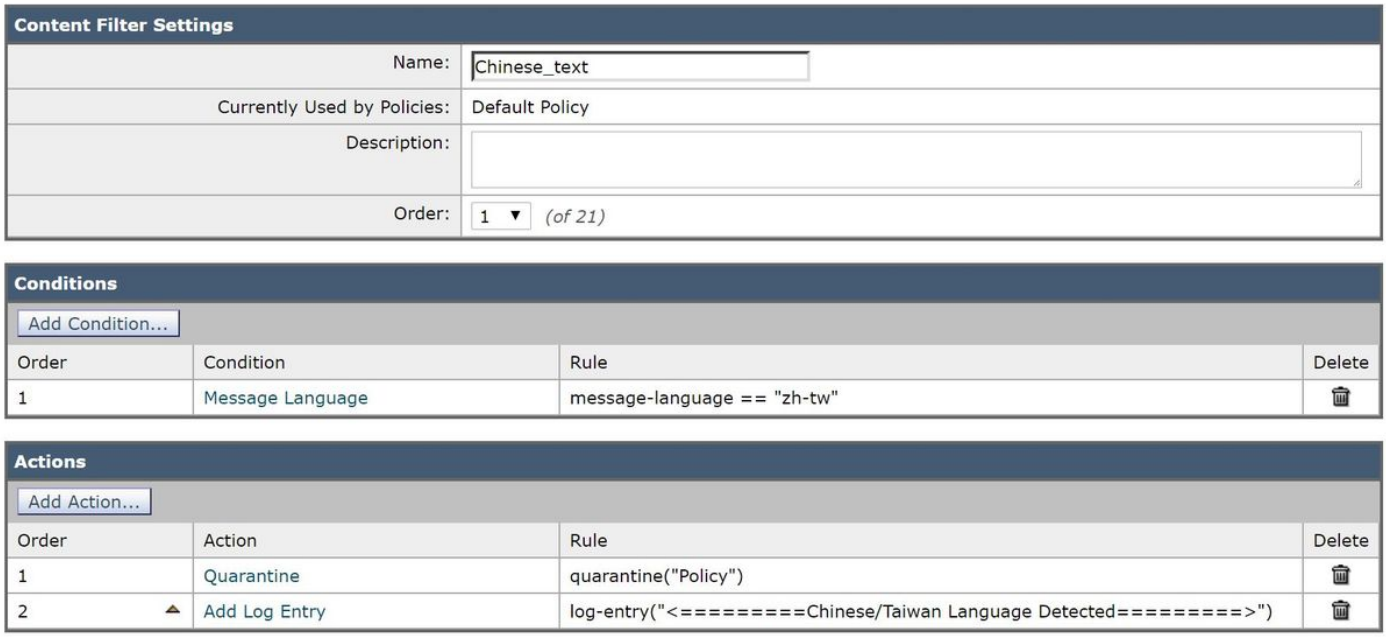
第三个选项是使用“消息语言”情况。ESA使用内置的语言检测引擎检测在消息的语言。设备解压缩主题和消息主题并且通过它到语言检测引擎。

语言检测引擎确定每个语言的可能性在解压缩的文本的并且通过它回到设备。设备考虑与高可能性的语言作为消息的语言。设备考虑消息的语言作为“未确定”在任一个下列的方案中：

- 如果ESA不支持检测的语言
- 如果设备无法检测消息的语言
- 如果解压缩的文本的总大小少于50个字节发送到语言检测引擎是。

Note:此选项是AsyncOS电子邮件安全的10.0.0-203一新特性和更新。

以下示例将考虑到包含中国/台湾基于字符集的邮件消息。 写入，内容过滤器：



使用如上所述配置的内容过滤器，邮件日志将记录类似于以下：

```
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
```

参考

- Microsoft提供字符集名称(*.NET命名*)在可以被参考，当写入和配置过滤器时的他们的代码表标识符。

  Note: ANSI代码表是不同的在不同的计算机或者可以为单个的计算机更改，导致数据损坏。对于最一致的结果，应用程序应该使用Unicode，例如UTF-8或UTF-16，而不是一个特定代码表。

- Mozillazine为内容类型提供详细细节：报头，外国字母，外来词，等等，在他们的外语垃圾邮件的条款

相关信息

- **Homoglyph提前网络钓鱼攻击**
- **思科电子邮件安全工具最终用户指南**
- **技术支持和文档 - Cisco Systems**