Comment bloquer le type satisfait a basé des jeux de caractères

Contenu

Introduction

Ce document décrit comment écrire et configurer un filtre afin de détecter et agir sur les jeux de caractères basés de type satisfait sur l'appliance de sécurité du courrier électronique de Cisco (ESA). Le document suivant peut être utilisé pour détecter les caractères basés sur un langage étrangers vus dans les messages spam.

Informations générales

Les administrateurs ESA peuvent recevoir un afflux des messages qui contiennent les langues étrangères basées par caractère qui ne sont pas messagerie légitime pour leur société ou domaines. Une manière d'adresser de l'ESA, nous avons trois options :

3. Écrivez un filtre utilisant le langage de message de condition. (Cette option est une nouvelle caractéristique pour la sécurité du courrier électronique 10.0.0-203 d'AsyncOS et plus nouveau.)

Comment bloquer le type satisfait a basé des jeux de caractères

Écrivez un filtre pour détecter le type satisfait

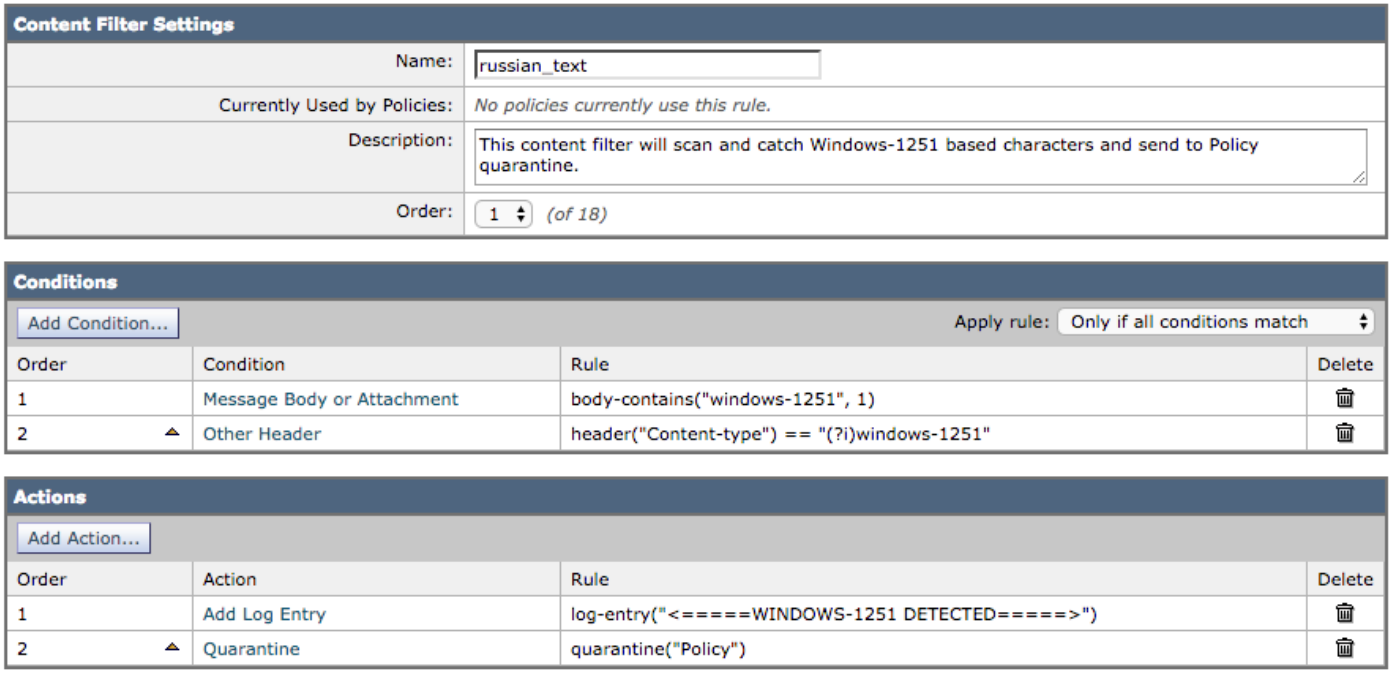
Le premier choix est pour que l'administrateur écrive et pour configure un filtre, et l'associe à une stratégie de messagerie, comme nécessaire.

> Remarque: L'inscription et configurer de ce filtre comme filtre de message peuvent être ressource-chères afin de balayer le corps des emails pour les jeux de caractères.

> Remarque: Configurer ceci comme un filtre satisfait est fortement suggéré, comme les filtres

satisfaits se produisent après la lecture d'anti-Spam. Cependant, ceci peut être écrit et configuré comme filtre de message, si nécessaire.

L'exemple suivant prendra en considération un message contiennent les caractères basés (cyrilliques) russes par l'intermédiaire du jeu de caractères basé par Windows-1251. Écrit comme filtre satisfait :



L'email de test utilisé contiendra le suivant dans le corps de l'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.
```

Le filtre de contenu étant configuré comme ci-dessus, les logs de messagerie enregistreraient semblable à ce qui suit :

```
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
```

D'autres langages et jeux de caractères peuvent être utilisés. Veuillez voir la section de références pour information les informations complémentaires.

Écrivez un filtre pour mettre en référence un dictionnaire basé par caractère

La deuxième option est d'ajouter la liste de jeux de caractères à un fichier texte de dictionnaire et de se rapporter à cela dans le filtre.

Exemple d'ajouter les caractères au dictionnaire :

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	

Les caractères sont maintenant assignés au dictionnaire et le dictionnaire lui-même est mis en référence dans les éléments de condition pour le filtre :

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 ======>")	

Utilisant le même email de test comme ci-dessus, il contient le suivant dans le corps de l'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.
```

Le filtre de contenu étant configuré comme ci-dessus utilisant l'état de correspondance de

dictionnaire, les logs de messagerie enregistreraient semblable à ce qui suit :

```
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
```

Écrivez un filtre satisfait utilisant l'état « de langage de message »

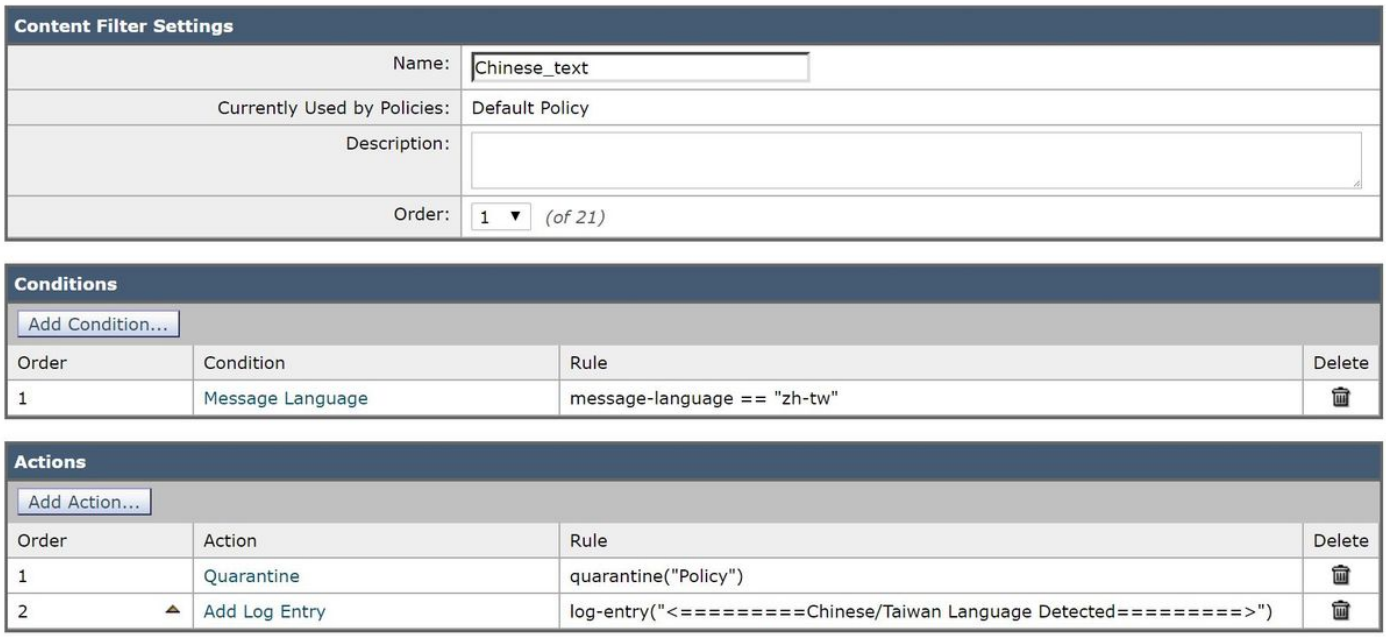
La troisième option est d'utiliser l'état « de langage de message ». L'ESA utilise l'engine intégrée de détection de langage pour détecter le langage dans un message. L'appliance extrait le sujet et le corps du message et le passe à l'engine de détection de langage.

L'engine de détection de langage détermine la probabilité de chaque langage dans le texte extrait et la passe de nouveau à l'appliance. L'appliance considère le langage avec la probabilité la plus élevée comme langage du message. L'appliance considère le langage du message en tant que « indéterminé » dans un des scénarios suivants :

- Si le langage détecté n'est pas pris en charge par ESA
- Si l'appliance ne peut pas détecter le langage du message
- Si la taille totale du texte extrait envoyé à l'engine de détection de langage est moins de 50 octets.

  Remarque: Cette option est une nouvelle caractéristique pour la sécurité du courrier électronique 10.0.0-203 d'AsyncOS et plus nouveau.

L'exemple suivant prendra en considération un message qui contient le chinois/le jeu de caractères basé par Taïwan. Écrit comme filtre satisfait :



Content Filter Settings

Name:	Chinese_text
Currently Used by Policies:	Default Policy
Description:	
Order:	1 (of 21)

Conditions

Add Condition...

Order	Condition	Rule	Delete
1	Message Language	message-language == "zh-tw"	

Actions

Add Action...

Order	Action	Rule	Delete
1	Quarantine	quarantine("Policy")	
2	Add Log Entry	log-entry("<=========Chinese/Taiwan Language Detected=========>")	

Le filtre de contenu étant configuré comme ci-dessus, les logs de messagerie enregistreraient semblable à ce qui suit :

```
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
```

Références

- Microsoft fournit des noms de jeu de caractères (*nom de .NET*) dans leurs identifiants de page de code qui peuvent être mis en référence en écrivant et en configurant des filtres.

  Remarque: Les pages de code d'ANSI peuvent être différentes sur différents ordinateurs, ou peuvent être changées pour un ordinateur unique, menant à la corruption des données. Pour les résultats les plus cohérents, les applications devraient utiliser Unicode, tel qu'UTF-8 ou UTF-16, au lieu d'une page de code spécifique.

- Mozillazine fournit les détails en profondeur pour le type de contenu : en-tête, lettres étrangères, mots étrangers, et plus, en leur article pour le Spam de langue étrangère

Informations connexes

- Homoglyph a avancé des attaques par phishing
- Guides d'utilisateur final d'appareils de sécurité du courrier électronique de Cisco
- Support et documentation techniques - Cisco Systems