Blockieren von auf Inhaltstypen basierenden Zeichensätzen

Inhalt

Einführung

In diesem Dokument wird beschrieben, wie Sie einen Filter schreiben und konfigurieren, um inhaltsbezogene Zeichensätze auf der Cisco E-Mail Security Appliance (ESA) zu erkennen und entsprechende Maßnahmen zu ergreifen. Das folgende Dokument kann verwendet werden, um fremdsprachige Zeichen zu erkennen, die in Spam-Nachrichten sichtbar sind.

Hintergrundinformationen

ESA-Administratoren können einen Zustrom von E-Mail-Nachrichten empfangen, die zeichenbasierte Fremdsprachen enthalten, bei denen es sich nicht um legitime E-Mails für ihr Unternehmen oder ihre Domänen handelt. Eine Möglichkeit, auf die ESA zuzugreifen, besteht aus drei Optionen:

3. Schreiben Sie einen Filter mithilfe der Bedingung Message Language. (Diese Option ist eine neue Funktion für AsyncOS Email Security 10.0.0-203 und höher.)

Blockieren von auf Inhaltstypen basierenden Zeichensätzen

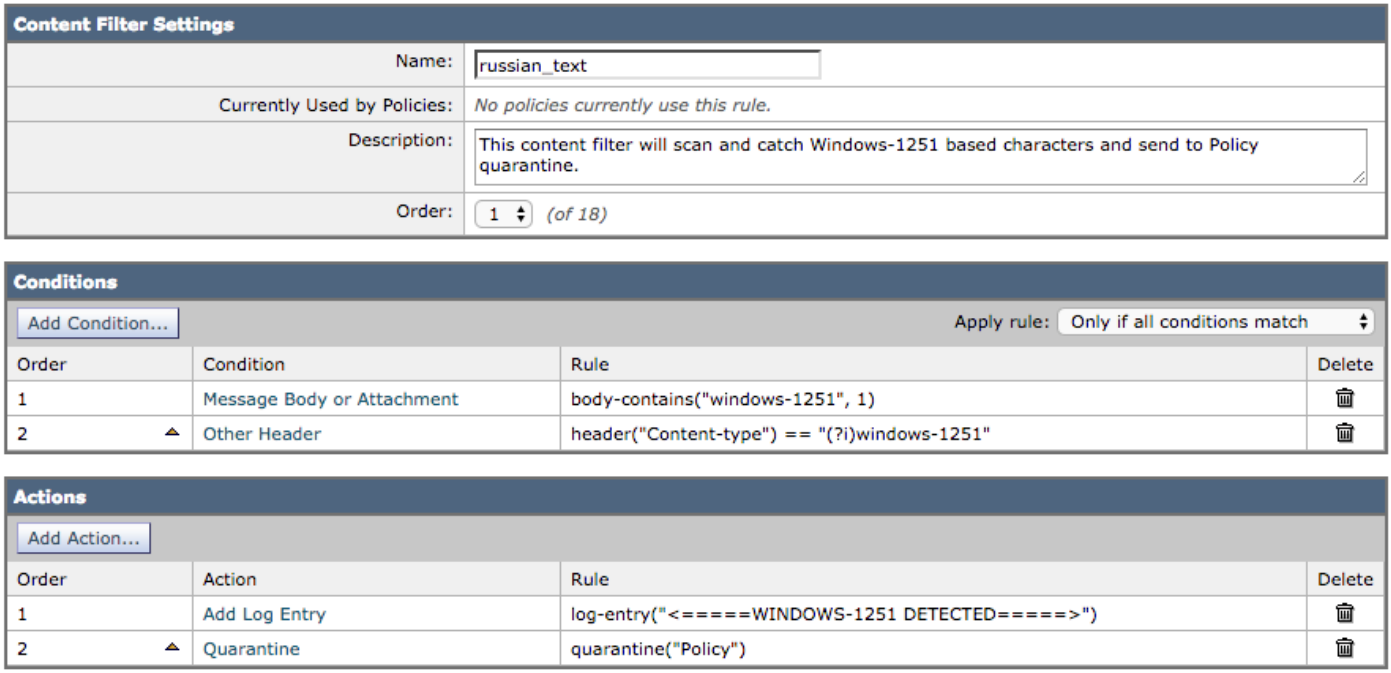
Schreiben eines Filters zum Erkennen des Inhaltstyps

Die erste Option besteht darin, dass der Administrator einen Filter schreibt und konfiguriert und ihn nach Bedarf einer Mail-Richtlinie zuordnet.

> **Hinweis**: Das Schreiben und Konfigurieren dieses Filters als Nachrichtenfilter kann ressourcenaufwändig sein, um den Text der E-Mails nach den Zeichensätzen zu durchsuchen.

Hinweis: Die Konfiguration als Content-Filter wird dringend empfohlen, da Content-Filter nach dem Anti-Spam-Scannen auftreten. Dies kann jedoch bei Bedarf als Nachrichtenfilter geschrieben und konfiguriert werden.

Im folgenden Beispiel wird eine E-Mail-Nachricht mit auf Russisch (kyrillisch) basierenden Zeichen über den Windows-1251-basierten Zeichensatz berücksichtigt. Geschrieben als Inhaltsfilter:



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")	

Die verwendete Test-E-Mail enthält den folgenden Text im E-Mail-Text:

```
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.
```

Wenn der Content-Filter wie oben konfiguriert konfiguriert ist, werden die E-Mail-Protokolle ähnlich wie folgt aufgezeichnet:

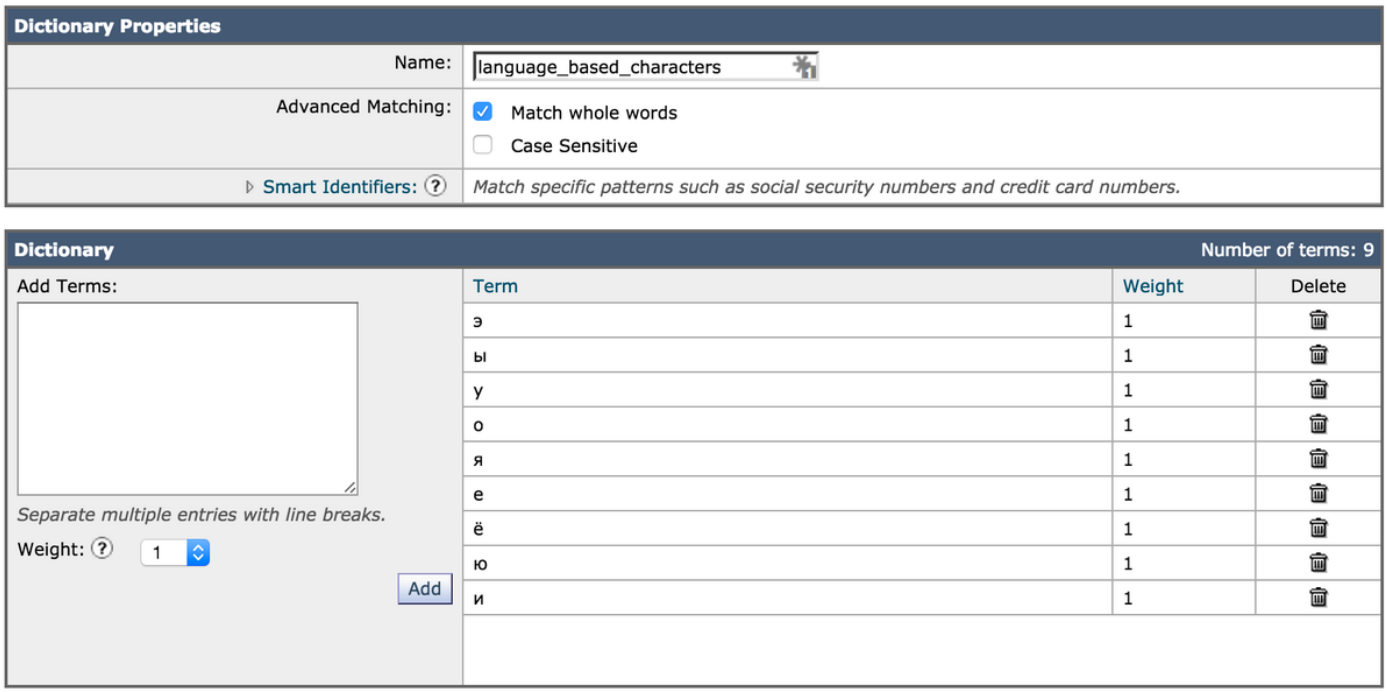
```
Thu Sep 10 14:50:09 2015 Info: Start MID 164993 ICID 266729
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 From: <end_user@test.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 ICID 266729 RID 0 To: <recpient@my_co.com>
Thu Sep 10 14:50:09 2015 Info: MID 164993 using engine: SPF Verdict Cache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

Es können auch andere Sprachen und Zeichensätze verwendet werden. Weitere Informationen finden Sie im Abschnitt Referenzen.

Schreiben eines Filters, um auf ein zeichenbasiertes Wörterbuch zu verweisen

Die zweite Option besteht darin, einer Wörterbuchtextdatei eine Liste mit Zeichensätzen hinzuzufügen und auf diese im Filter zu verweisen.

Beispiel für das Hinzufügen von Zeichen zum Wörterbuch:



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:

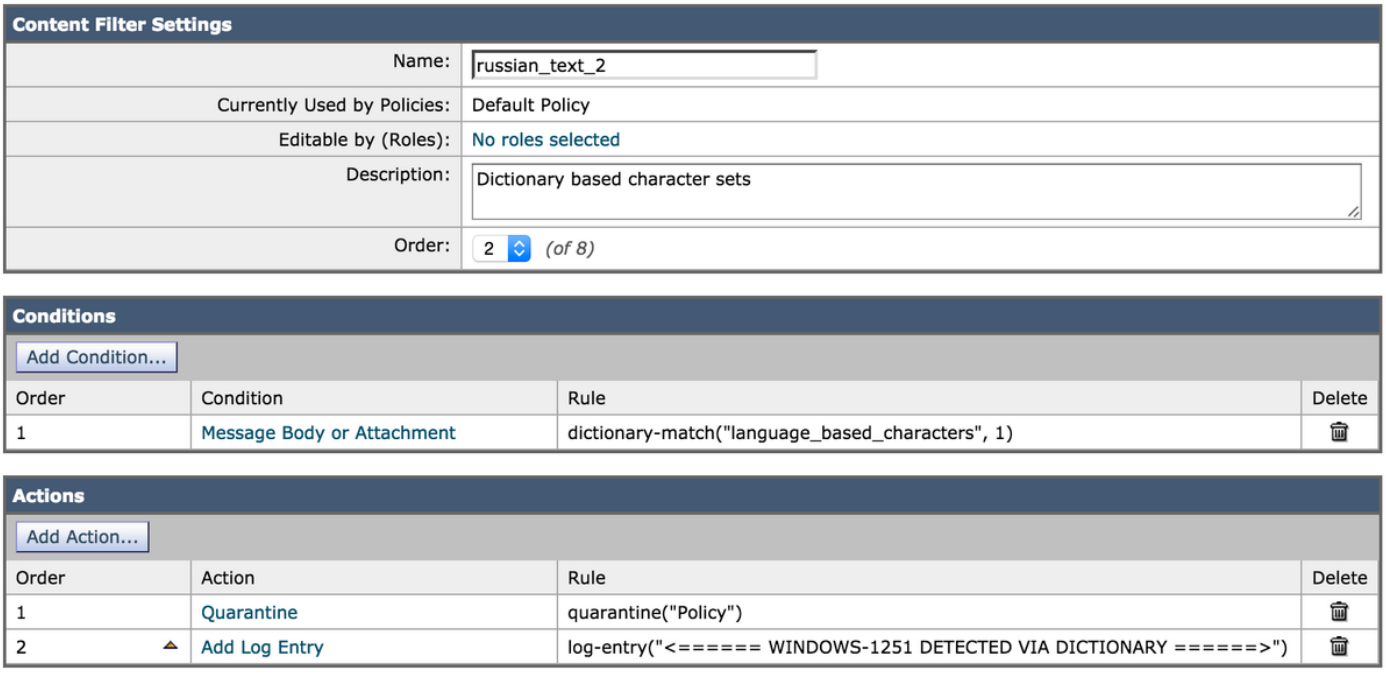
Separate multiple entries with line breaks.

Weight: 1

Add

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

Die Zeichen werden nun dem Wörterbuch zugewiesen, und das Wörterbuch selbst wird in den Bedingungselementen für den Filter referenziert:



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 ======>")	

Die E-Mail enthält unter Verwendung der oben angegebenen Test-E-Mail den folgenden Text:

```
Russian uses , , , , о, , , , ,  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.
```

Wenn der Content-Filter wie oben mit der Bedingung für die Wörterbuchübereinstimmung konfiguriert wurde, werden die Mail-Protokolle ähnlich wie folgt aufgezeichnet:

```
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
```

Erstellen eines Inhaltsfilters mit der Bedingung "Message Language" (Sprache der Nachricht)

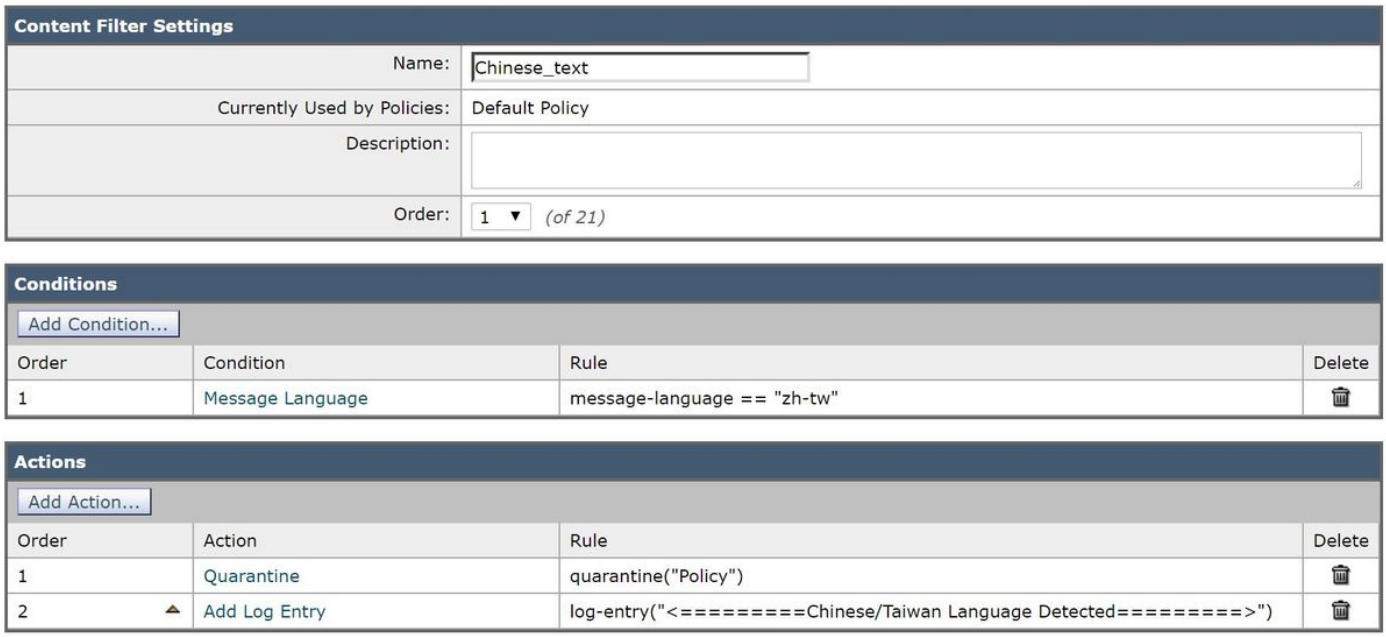
Die dritte Option ist die Verwendung der Bedingung "Message Language" (Sprache der Nachricht). Die ESA verwendet die integrierte Spracherkennungs-Engine, um die Sprache in einer Nachricht zu erkennen. Die Appliance extrahiert das Betreff und den Nachrichtentext und leitet ihn an die Spracherkennungs-Engine weiter.

Die Spracherkennungs-Engine bestimmt die Wahrscheinlichkeit für jede Sprache im extrahierten Text und übergibt ihn an die Appliance. Die Appliance betrachtet die Sprache mit der höchsten Wahrscheinlichkeit als die Sprache der Nachricht. Die Appliance betrachtet die Sprache der Nachricht in einem der folgenden Szenarien als "unbestimmt":

- Wenn die erkannte Sprache von der ESA nicht unterstützt wird
- Wenn die Appliance die Sprache der Nachricht nicht erkennen kann
- Wenn die Gesamtgröße des extrahierten Texts, der an die Spracherkennungs-Engine gesendet wird, weniger als 50 Byte beträgt.

  Hinweis: Diese Option ist eine neue Funktion für AsyncOS Email Security 10.0.0-203 und neuere Versionen.

Das folgende Beispiel berücksichtigt eine E-Mail-Nachricht, die einen auf Chinesisch/Taiwan basierenden Zeichensatz enthält. Geschrieben als Inhaltsfilter:



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=========>")	

Wenn der Content-Filter wie oben konfiguriert konfiguriert ist, werden die E-Mail-Protokolle ähnlich wie folgt aufgezeichnet:

```
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
```

Referenzen

- Microsoft stellt Zeichensatznamen (*.NET-Name*) in der Codepage-Bezeichner auf die beim Schreiben und Konfigurieren von Filtern verwiesen werden kann.

  Hinweis: ANSI-Codepages können auf verschiedenen Computern unterschiedlich sein oder für einen einzelnen Computer geändert werden, was zu Datenbeschädigungen führt. Für die konsistentesten Ergebnisse sollten Anwendungen statt einer bestimmten Codeseite Unicode verwenden, z. B. UTF-8 oder UTF-16.

- Mozillazin enthält detaillierte Informationen zum Inhaltstyp: Überschrift, ausländischen Buchstaben, ausländischen Wörtern und mehr, in ihrem Artikel für Spam in Fremdsprachen

Zugehörige Informationen

- Homoglyph Advanced Phishing-Angriffe
- Cisco Email Security Appliance - Benutzerhandbücher
- Technischer Support und Dokumentation - Cisco Systems