

Sensitive Data Detection

The following topics explain sensitive data detection and how to configure it:

- Sensitive Data Detection Basics, on page 1
- Global Sensitive Data Detection Options, on page 2
- Individual Sensitive Data Type Options, on page 3
- System-Provided Sensitive Data Types, on page 4
- License Requirements for Sensitive Data Detection, on page 4
- Requirements and Prerequisites for Sensitive Data Detection, on page 5
- Configuring Sensitive Data Detection, on page 5
- Monitored Application Protocols and Sensitive Data, on page 6
- Selecting Application Protocols to Monitor, on page 7
- Special Case: Sensitive Data Detection in FTP Traffic, on page 8
- Custom Sensitive Data Types, on page 8

Sensitive Data Detection Basics

Sensitive data such as Social Security numbers, credit card numbers, driver's license numbers, and so on may be leaked onto the Internet, intentionally or accidentally. The system provides a sensitive data preprocessor that can detect and generate events on sensitive data in ASCII text, which can be particularly useful in detecting accidental data leaks.

Global sensitive data preprocessor options control how the preprocessor functions. You can modify global options that specify the following:

- whether the preprocessor replaces all but the last four credit card or Social Security numbers in triggering packets
- which destination hosts on your network to monitor for sensitive data
- how many total occurrences of all data types in a single session result in an event

Individual data types identify the sensitive data you can detect and generate events on in your specified destination network traffic. You can modify default settings for data type options that specify the following:

- a threshold that must be met for a detected data type to generate a single per-session event
- the destination ports to monitor for each data type

- the application protocols to monitor for each data type

You can create and modify custom data types to detect data patterns that you specify. For example, a hospital might create a data type to protect patient numbers, or a university might create a data type to detect student numbers that have a unique numbering pattern.

The system detects sensitive data per TCP session by matching individual data types against traffic. You can modify the default settings for each data type and for global options that apply to all data types in your intrusion policy. The system provides predefined, commonly used data types. You can also create custom data types.

A sensitive data preprocessor rule is associated with each data type. You enable sensitive data detection and event generation for each data type by enabling the corresponding preprocessor rule for the data type. A link on the configuration page takes you to a filtered view of sensitive data rules on the Rules page, where you can enable and disable rules and configure other rule attributes.

When you save changes to your intrusion policy, you are given the option to automatically enable the sensitive data preprocessor if the rule associated with a data type is enabled and sensitive data detection is disabled.



Tip The sensitive data preprocessor can detect sensitive data in unencrypted Microsoft Word files that are uploaded and downloaded using FTP or HTTP; this is possible because of the way Word files group ASCII text and formatting commands separately.

The system does not detect encrypted or obfuscated sensitive data, or sensitive data in a compressed or encoded format such as a Base64-encoded email attachment. For example, the system would detect the phone number (555)123-4567, but not an obfuscated version where each number is separated by spaces, as in (5 5 5) 1 2 3 - 4 5 6 7, or by intervening HTML code, such as (555)-123-4567. However, the system would detect, for example, the HTML coded number (555)-123-4567 where no intervening codes interrupt the numbering pattern.

Global Sensitive Data Detection Options

Global sensitive data options are policy-specific and apply to all data types.

Mask

Replaces with Xs all but the last four digits of credit card numbers and Social Security numbers in the triggering packet. The masked numbers appear in the intrusion event packet view in the web interface and in downloaded packets.

Networks

Specifies the destination host or hosts to monitor for sensitive data. You can specify a single IP address, address block, or a comma-separated list of either or both. The system interprets a blank field as `any`, meaning any destination IP address.

Global Threshold

Specifies the total number of all occurrences of all data types during a single session that the preprocessor must detect in any combination before generating a global threshold event. You can specify 1 through 65535.

Cisco recommends that you set the value for this option higher than the highest threshold value for any individual data type that you enable in your policy.

Note the following points regarding global thresholds:

- You must enable preprocessor rule 139:1 to detect and generate events and, in an inline deployment, drop offending packets on combined data type occurrences.
- The preprocessor generates up to one global threshold event per session.
- Global threshold events are independent of individual data type events; that is, the preprocessor generates an event when the global threshold is reached, regardless of whether the event threshold for any individual data type has been reached, and vice versa.

Individual Sensitive Data Type Options

At a minimum, each custom data type must specify an event threshold and at least one port or application protocol to monitor.

Each system-provided data type uses an otherwise inaccessible `sd_pattern` keyword to define a built-in data pattern to detect in traffic. You can also create custom data types for which you use simple regular expressions to specify your own data patterns.

Sensitive data types display in all intrusion policies where Sensitive Data Detection is enabled. System-provided data types display as read-only. For custom data types, the name and pattern fields display as read-only, but you can set the other options to policy-specific values.

Table 1: Individual Data Type Options

Option	Description
Data Type	Specifies the unique name for the data type.
Threshold	Specifies the number of occurrences of the data type when the system generates an event. You can specify 1 through 255. Note that the preprocessor generates one event for a detected data type per session. Note also that global threshold events are independent of individual data type events; that is, the preprocessor generates an event when the data type event threshold is reached, regardless of whether the global event threshold has been reached, and vice versa.
Destination Ports	Specifies destination ports to monitor for the data type. You can specify a single port, a comma-separated list of ports, or `any`, meaning any destination port.
Application Protocols	Specifies up to eight application protocols to monitor for the data type. You must activate application detectors to identify application protocols to monitor. Note that, for Classic devices, this feature requires a Control license.
Pattern	Specifies the pattern to detect. This field is only present for custom data types.

Related Topics

Activating and Deactivating Detectors

System-Provided Sensitive Data Types

Each intrusion policy includes system-provided data types for detecting commonly used data patterns such as credit card numbers, email addresses, U.S. phone numbers, and U.S. Social Security numbers with and without dashes.

Each system-provided data type is associated with a single sensitive data preprocessor rule that has a generator ID (GID) of 138. You must enable the associated sensitive data rule in the intrusion policy to generate events and, in an inline deployment, drop offending packets for each data type that you want to use in your policy.

The following table describes each data type and lists the corresponding preprocessor rule.

Table 2: System-Provided Sensitive Data Types

Data Type	Description	Preprocessor GID:SID
Credit Card Numbers	Matches Visa®, MasterCard®, Discover® and American Express® fifteen- and sixteen-digit credit card numbers, with or without their normal separating dashes or spaces; also uses the Luhn algorithm to verify credit card check digits.	138:2
Email Addresses	Matches email addresses.	138:5
U.S. Phone Numbers	Matches U.S. phone numbers adhering to the pattern `(\d{3}) ?\d{3}-\d{4}`.	138:6
U.S. Social Security Numbers Without Dashes	Matches 9-digit U.S. Social Security numbers that have valid 3-digit area numbers, valid 2-digit group numbers, and do not have dashes.	138:4
U.S. Social Security Numbers With Dashes	Matches 9-digit U.S. Social Security numbers that have valid 3-digit area numbers, valid 2-digit group numbers, and dashes.	138:3

To reduce false positives from 9-digit numbers other than Social Security numbers, the preprocessor uses an algorithm to validate the 3-digit area number and 2-digit group number that precede the 4-digit serial number in each Social Security number. The preprocessor validates Social Security group numbers through November 2009.

License Requirements for Sensitive Data Detection

FTD License

Threat

Classic License

Protection, or as indicated in a procedure.

Requirements and Prerequisites for Sensitive Data Detection

Model Support

Any.

Supported Domains

Any

User Roles

- Admin
- Intrusion Admin

Configuring Sensitive Data Detection

Because sensitive data detection can have a high impact on the performance of your system, Cisco recommends that you adhere to the following guidelines:

- Choose the No Rules Active default policy as your base intrusion policy.
- Ensure that the following settings are enabled in the corresponding network analysis policy:
  - **FTP and Telnet Configuration** under **Application Layer Preprocessors**
  - **IP Defragmentation** and **TCP Stream Configuration** under **Transport/Network Layer Preprocessors**.

Before you begin

For classic devices, this procedure requires the Protection or Control license.

Procedure

Step 1 Choose **Policies** > **Access Control heading** > **Intrusion**

Step 2 Click **Snort 2 Version** next to the policy you want to edit.

If **View** ( ) appears instead, the configuration belongs to an ancestor domain, or you do not have permission to modify the configuration.

Step 3 Click **Advanced Settings** in the navigation panel.

Step 4 If **Sensitive Data Detection** under **Specific Threat Detection** is disabled, click **Enabled**.

Step 5 Click **Edit** ( ) next to **Sensitive Data Detection**.

Step 6 You have the following choices:

- Modify the global settings as described in Global Sensitive Data Detection Options, on page 2.
- Choose a data type in the **Targets** section, and modify the data type configuration as described in Individual Sensitive Data Type Options, on page 3.
- If you want to inspect custom sensitive data, create a custom data type; see Custom Sensitive Data Types, on page 8.

Step 7 Add or remove application protocols to monitor for a data type; see Monitored Application Protocols and Sensitive Data, on page 6.

Note
To detect sensitive data in FTP traffic:

- Ensure that the file policy is enabled for the access control policy.
- You must add the `Ftp data` application protocol.

Step 8 Optionally, to display sensitive data preprocessor rules, click **Configure Rules for Sensitive Data Detection**.

You can enable or disable any of the listed rules. You can also configure sensitive data rules for any of the other actions available on the Rules page, such as rule suppression, rate-based attack prevention, and so on; see Intrusion Rule Types for more information.

Step 9 To save changes you made in this policy since the last policy commit, click **Policy Information** in the navigation panel, then click **Commit Changes**.

If you enable sensitive data preprocessor rules in your policy without enabling sensitive data detection, you are prompted to enable sensitive data detection when you save changes to your policy.

If you leave the policy without committing changes, changes since the last commit are discarded if you edit a different policy.

What to do next

- If you want to generate intrusion events, enable Sensitive Data Detection rules 138:2, 138:3, 138:4, 138:5, 138:6, 138:>999999, or 139:1. For more information, see Intrusion Rule States, Global Sensitive Data Detection Options, on page 2, System-Provided Sensitive Data Types, on page 4, and Custom Sensitive Data Types, on page 8.
- Deploy configuration changes; see Deploy Configuration Changes.

Related Topics

Special Case: Sensitive Data Detection in FTP Traffic, on page 8

Monitored Application Protocols and Sensitive Data

You can specify up to eight application protocols to monitor for each data type. At least one detector must be enabled for each application protocol you select. By default, all system-provided detectors are activated. If no detector is enabled for an application protocol, the system automatically enables all system-provided detectors for the application; if none exist, the system enables the most recently modified user-defined detector for the application.

You must specify at least one application protocol or port to monitor for each data type. However, except in the case where you want to detect sensitive data in FTP traffic, Cisco recommends for the most complete coverage that you specify corresponding ports when you specify application protocols. For example, if you specify HTTP, you might also configure the well-known HTTP port 80. If a new host on your network implements HTTP, the system monitors port 80 during the interval when it is discovering the new HTTP application protocol.

In the case where you want to detect sensitive data in FTP traffic, you must specify the `FTP data` application protocol; there is no advantage in specifying a port number.

Related Topics

Activating and Deactivating Detectors

Special Case: Sensitive Data Detection in FTP Traffic, on page 8

Selecting Application Protocols to Monitor

You can specify application protocols to monitor in both system-provided and custom sensitive data types. The application protocols you select are policy-specific.

Before you begin

For classic devices, this procedure requires the Control license.

Procedure

Step 1 Choose **Policies** > **Access Control heading** > **Intrusion**.

Step 2 Click **Snort 2 Version** next to the policy you want to edit.

If **View** ( ) appears instead, the configuration belongs to an ancestor domain, or you do not have permission to modify the configuration.

Step 3 Click **Advanced Settings** in the navigation panel.

Step 4 If **Sensitive Data Detection** under **Specific Threat Detection** is disabled, click **Enabled**.

Step 5 Click **Edit** ( ) next to **Sensitive Data Detection**.

Step 6 Click the name of a data type under **Data Types**.

Step 7 Click **Edit** ( ) next to the **Application Protocols** field.

Step 8 You have the following choices:

- To add application protocols for monitoring, choose one or more application protocols from the **Available** list, then click right arrow (>). You can add up to eight application protocols for monitoring.
- To remove an application protocol from monitoring, choose it from the **Enabled** list, then click left arrow (<).

Step 9 Click **OK**.

Step 10 To save changes you made in this policy since the last policy commit, click **Policy Information** in the navigation pane, then click **Commit Changes**.

If you leave the policy without committing changes, changes since the last commit are discarded if you edit a different policy.

What to do next

- Deploy configuration changes; see Deploy Configuration Changes.

Related Topics

Special Case: Sensitive Data Detection in FTP Traffic, on page 8

Special Case: Sensitive Data Detection in FTP Traffic

You usually determine which traffic to monitor for sensitive data by specifying the ports to monitor or specifying application protocols in deployments.

However, specifying ports or application protocols is not sufficient for detecting sensitive data in FTP traffic. Sensitive data in FTP traffic is found in traffic for the FTP application protocol, which occurs intermittently and uses a transient port number, making it difficult to detect. To detect sensitive data in FTP traffic, you **must** include the following in your configuration:

- Specify the `FTP data` application protocol to enable detection of sensitive data in FTP traffic.

  In the special case of detecting sensitive data in FTP traffic, specifying the `FTP data` application protocol does not invoke detection; instead, it invokes the rapid processing of the FTP/Telnet processor to detect sensitive data in FTP traffic.

- Ensure that the FTP Data detector, which is enabled by default, is enabled.
- Ensure that your configuration includes at least one port to monitor for sensitive data.
- Ensure that the file policy is enabled for the Access Control Policy.

Note that it is not necessary to specify an FTP port except in the unlikely case where you only want to detect sensitive data in FTP traffic. Most sensitive data configurations will include other ports such as HTTP or email ports. In the case where you do want to specify only one FTP port and no other ports to monitor, Cisco recommends that you specify the FTP command port `23`.

Related Topics

The FTP/Telnet Decoder
Activating and Deactivating Detectors
Configuring Sensitive Data Detection, on page 5

Custom Sensitive Data Types

Each custom data type you create also creates a single sensitive data preprocessor rule that has a Generator ID (GID) of 138 and a Snort ID (SID) of 1000000 or greater, that is, a SID for a local rule.

You must enable the associated sensitive data rule to enable detection, generate events and, in an inline deployment, drop offending packets for each custom data type that you want to use in your policy.

To help you enable sensitive data rules, a link on the configuration page takes you to a filtered view of the intrusion policy Rules page that displays all system-provided and custom sensitive data rules. You can also display custom sensitive data rules along with any custom local rules by choosing the local filtering category on the intrusion policy Rules page. Note that custom sensitive data rules are not listed on the intrusion rules editor page (**Objects** > **Intrusion Rules**).

Once you create a custom data type, you can enable it in any intrusion policy in the system. To enable a custom data type, you must enable the associated sensitive data rule in any policy that you want to use to detect that custom data type.

Data Patterns in Custom Sensitive Data Types

You define the data pattern for a custom data type using a simple set of regular expressions comprised of the following:

- three metacharacters
- escaped characters that allow you to use the metacharacters as literal characters
- six character classes

Metacharacters are literal characters that have special meaning within regular expressions.

Table 3: Sensitive Data Pattern Metacharacters

Metacharacter	Description	Example
`?`	Matches zero or one occurrence of the preceding character or escape sequence; that is, the preceding character or escape sequence is optional.	`colou?r` matches `color` or `colour`
`{n}`	Matches the preceding character or escape sequence n times.	For example, `\d{2}` matches `55`, `12`, and so on; `\l{3}` matches `AbC`, `www`, and so on; `\w{3}` matches `a1B`, 25C, and so on; `x{5}` matches `xxxxx`
`\`	Allows you to use metacharacters as actual characters and is also used to specify a predefined character class.	`\?` matches a question mark, `\\` matches a backslash, `\d` matches numeric characters, and so on

You must use a backslash to escape certain characters for the sensitive data preprocessor to interpret them correctly as literal characters.

Table 4: Escaped Sensitive Data Pattern Characters

Use this escaped character...	To represent this literal character...
\?	?
\{	{
\}	}
\\	\

When defining a custom sensitive data pattern, you can use character classes.

Table 5: Sensitive Data Pattern Character Classes

Character Class	Description	Character Class Definition
\d	Matches any numeric ASCII character 0-9	0-9
\D	Matches any byte that is not a numeric ASCII character	not 0-9
\l (lowercase “ell”)	Matches any ASCII letter	a-zA-Z
\L	Matches any byte that is not an ASCII letter	not a-zA-Z
\w	Matches any ASCII alphanumeric character Note that, unlike PCRE regular expressions, this does not include an underscore (_).	a-zA-Z0-9
\W	Matches any byte that is not an ASCII alphanumeric character	not a-zA-Z0-9

The preprocessor treats characters entered directly, instead of as part of a regular expression, as literal characters. For example, the data pattern 1234 matches `1234`.

The following data pattern example, which is used in system-provided sensitive data rule 138:4, uses the escaped digits character class, the multiplier and option-specifier metacharacters, and the literal dash (-) and left and right parentheses () characters to detect U.S. phone numbers:

```
(\d{3}) ?\d{3}-\d{4}
```

Exercise caution when creating custom data patterns. Consider the following alternative data pattern for detecting phone numbers which, although using valid syntax, could cause many false positives:

```
(?\d{3})? ?\d{3}-?\d{4}
```

Because the second example combines optional parentheses, optional spaces, and optional dashes, it would detect, among others, phone numbers in the following desirable patterns:

- `(555)123-4567`
- `555123-4567`
- `5551234567`

However, the second example pattern would also detect, among others, the following potentially invalid patterns, resulting in false positives:

- `(555 1234567`
- `555)123-4567`
- `555) 123-4567`

Consider finally, for illustration purposes only, an extreme example in which you create a data pattern that detects the lowercase letter `a` using a low event threshold in all destination traffic on a small company network. Such a data pattern could overwhelm your system with literally millions of events in only a few minutes.

Configuring Custom Sensitive Data Types

You cannot delete a data type if the sensitive data rule for that data type is enabled in any intrusion policy.

Procedure

Step 1 Choose **Policies** > **Access Control heading** > **Intrusion**

Step 2 Click **Snort 2 Version** next to the policy you want to edit.

If **View** ( ) appears instead, the configuration belongs to an ancestor domain, or you do not have permission to modify the configuration.

Step 3 Click **Advanced Settings** in the navigation panel.

Step 4 If **Sensitive Data Detection** under **Specific Threat Detection** is disabled, click **Enabled**.

Step 5 Click **Edit** ( ) next to **Sensitive Data Detection**.

Step 6 Click **Add** ( ) next to **Data Types**.

Step 7 Enter a name for the data type.

Step 8 Enter the pattern you want to detect with this data type; see Data Patterns in Custom Sensitive Data Types, on page 9.

Step 9 Click **OK**.

Step 10 Optionally, click the data type name, and modify the options described in Individual Sensitive Data Type Options, on page 3.

Step 11 Optionally, delete a custom data type by clicking **Delete** ( ), then **OK** to confirm.

Note
If the sensitive data rule for that data type is enabled in any intrusion policy, the system warns that you cannot delete the data type. You must disable the sensitive data rule in affected policies before attempting the deletion again; see Setting Intrusion Rule States.

Step 12 To save changes you made in this policy since the last policy commit, click **Policy Information** in the navigation panel, then click **Commit Changes**.

If you leave the policy without committing changes, changes since the last commit are discarded if you edit a different policy.

What to do next

- Enable the associated custom sensitive data preprocessing rule in each policy where you want to use that data type; see Setting Intrusion Rule States.
- Deploy configuration changes; see Deploy Configuration Changes.

Related Topics

Editing Custom Sensitive Data Types, on page 12

Editing Custom Sensitive Data Types

You can edit all fields in custom sensitive data types. Note, however, that when you modify the name or pattern field, these settings change in all intrusion policies on the system. You can set the other options to policy-specific values.

Procedure

Step 1 Choose **Policies** > **Access Control heading** > **Intrusion**

Step 2 Click **Snort 2 Version** next to the policy you want to edit.

If **View** ( ) appears instead, the configuration belongs to an ancestor domain, or you do not have permission to modify the configuration.

Step 3 Click **Advanced Settings** in the navigation panel.

Step 4 If **Sensitive Data Detection** under **Specific Threat Detection** is disabled, click **Enabled**.

Step 5 Click **Edit** next to **Sensitive Data Detection**.

Step 6 In the **Targets** section, click the name of the custom data type.

Step 7 Click **Edit Data Type Name and Pattern**.

Step 8 Modify the data type name and pattern; see Data Patterns in Custom Sensitive Data Types, on page 9.

Step 9 Click **OK**.

Step 10 Set the remaining options to policy-specific values; see Individual Sensitive Data Type Options, on page 3.

Step 11 To save changes you made in this policy since the last policy commit, click **Policy Information** in the navigation panel, then click **Commit Changes**.

If you leave the policy without committing changes, changes since the last commit are discarded if you edit a different policy.

What to do next

- Deploy configuration changes; see Deploy Configuration Changes.