Content Filter

- Content Filter
- Description of Content Filter Categories
- Setting up HTTPS Filtering

Content Filter

Setting up content filtering and troubleshooting.

Content filtering on the SafeUTM server is implemented based on web traffic data received from the web traffic proxying module. Thus, the content filter allows you to efficiently block access to various internet resources.

The mechanism of content filtering consists in checking the affiliation of the address requested by a website or website page user and its presence in prohibited resource lists. The lists in their turn are divided into categories for easier administration.

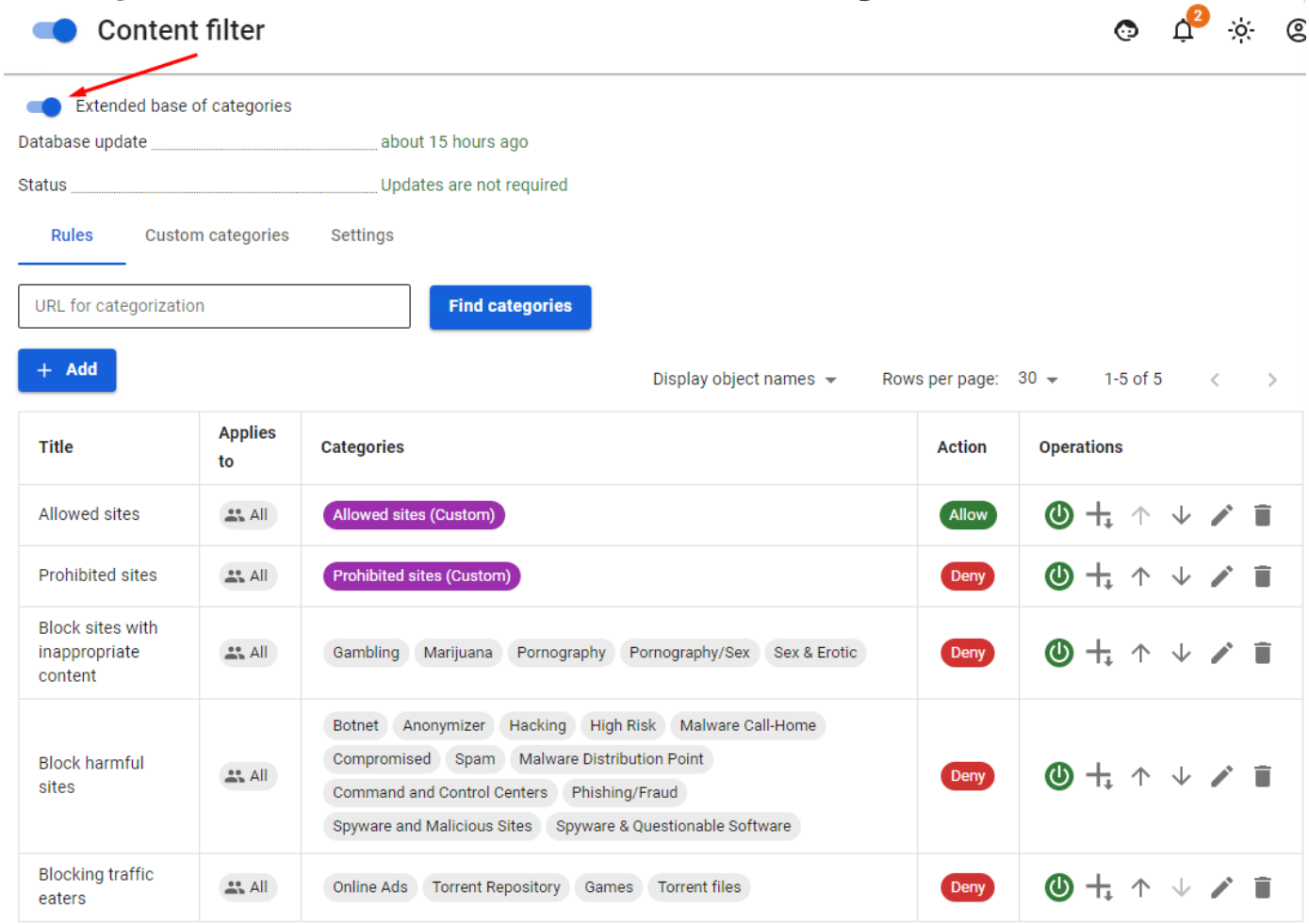
> The content filtering module only works with an active subscription to updates in the Enterprise edition.

> HTTPS sites without traffic decryption are filtered by domain only (not by full URL), **Files** category rules cannot be applied to them either. Create rules for decrypting HTTPS traffic of necessary categories in order to fully filter HTTPS.

Content filter settings and categories

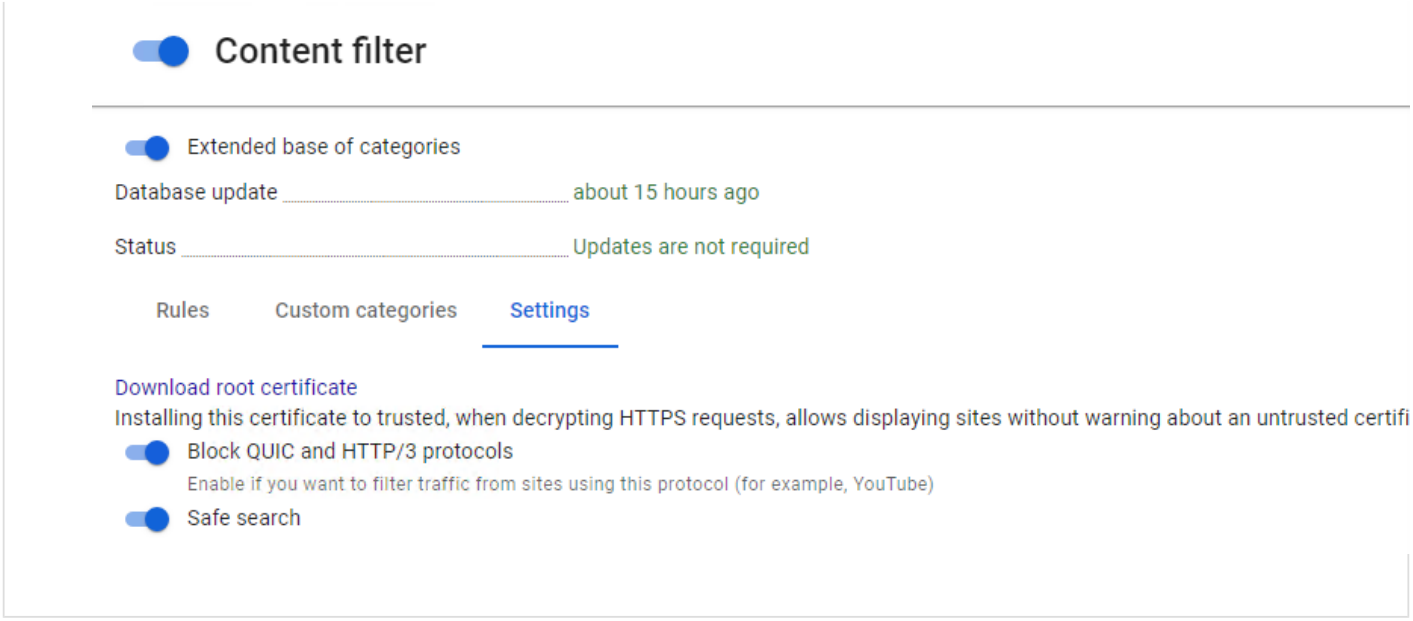
Content Filtering Setup

Go to **Traffic Rules -> Content Filter** and activate the extended content filter database by switching the slide to **Enabled** next to **Extended base of categories**.



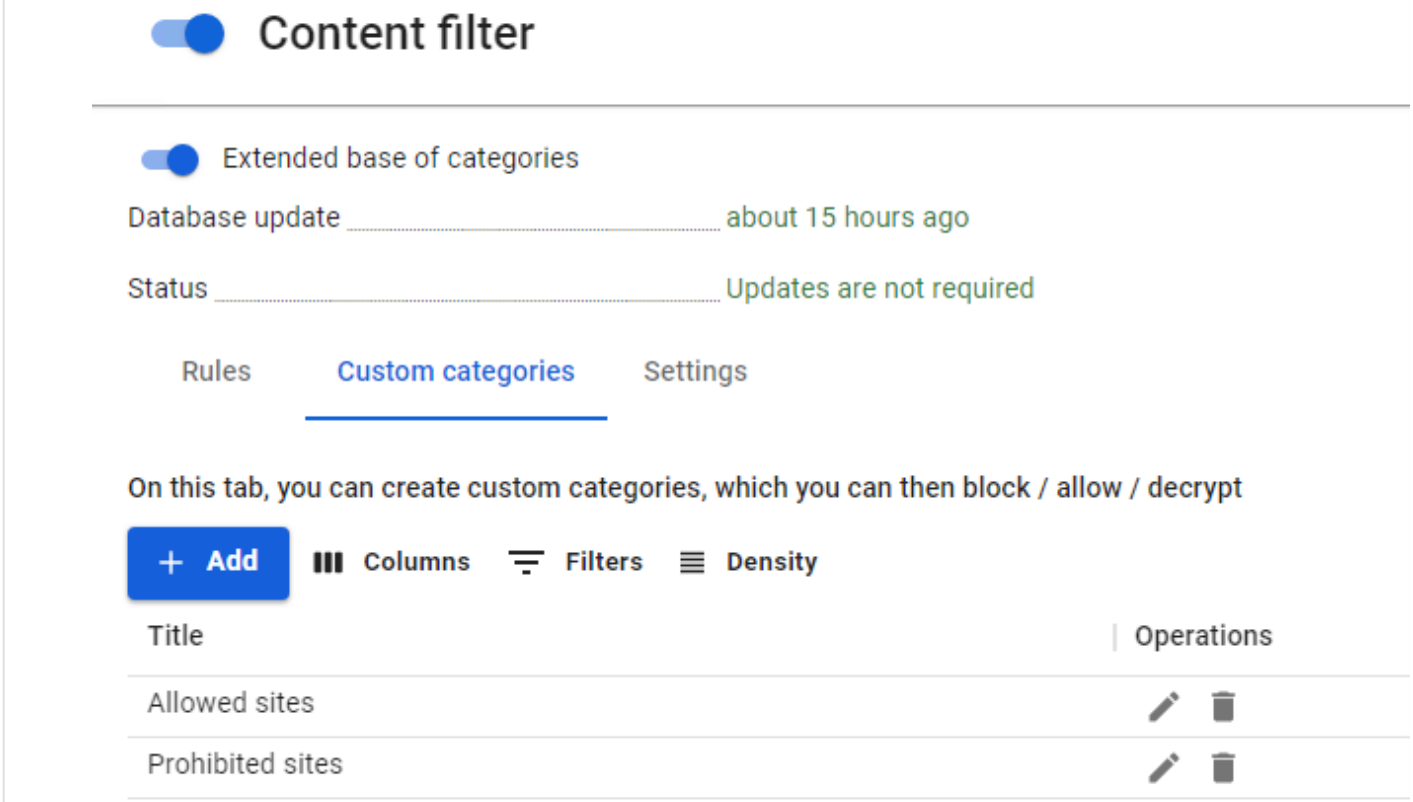
You can configure additional filtering options in the **Settings** tab:

- **Block QUIC and HTTP/3 protocols.** An experimental protocol used by Chrome browser for access to some resources (e.g. YouTube). It is recommended to be blocked as filtering of resources working under this protocol will not be possible otherwise.
- **Safe search.** Forcibly enables safe search in search engines (Google, Yandex, YouTube, Yahoo, Bing). **In order for this function to work, you need to enable HTTPS filtering by certificate substitution for these resources**.



Content Filtering Categories

- **Extended base of categories**. Over 140 categories including millions of URLs automatically updated by the server. The status of updates and database usage can be viewed in the **Settings** tab in the Content Filtering section. These categories only work with an active subscription to updates in commercial editions.
- **Custom categories**. You can create your own rules in the tab with the same name.



- **Special**. Includes four categories – all queries, all categorized queries, all non-categorized queries, and queries with direct access by IP addresses.

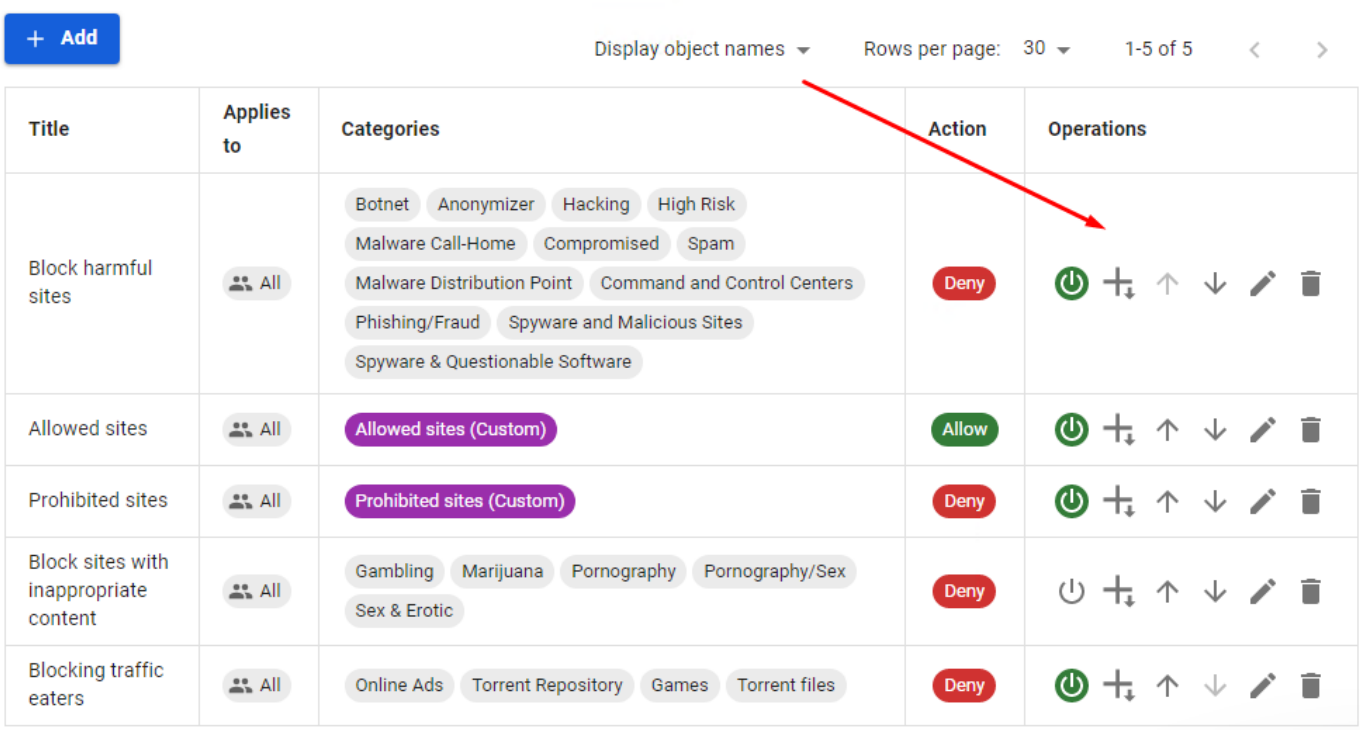
- **Files**. Eight defined categories of files blocked by extension and MIME type. Preset file groups (Executable Files, Archives, Video Files, Audio Files, Flash video, Active-X, Torrent files, and Documents) cannot be edited. Filtering HTTPS traffic for these types of categories is only possible when it is decrypted.

Applying Filtering

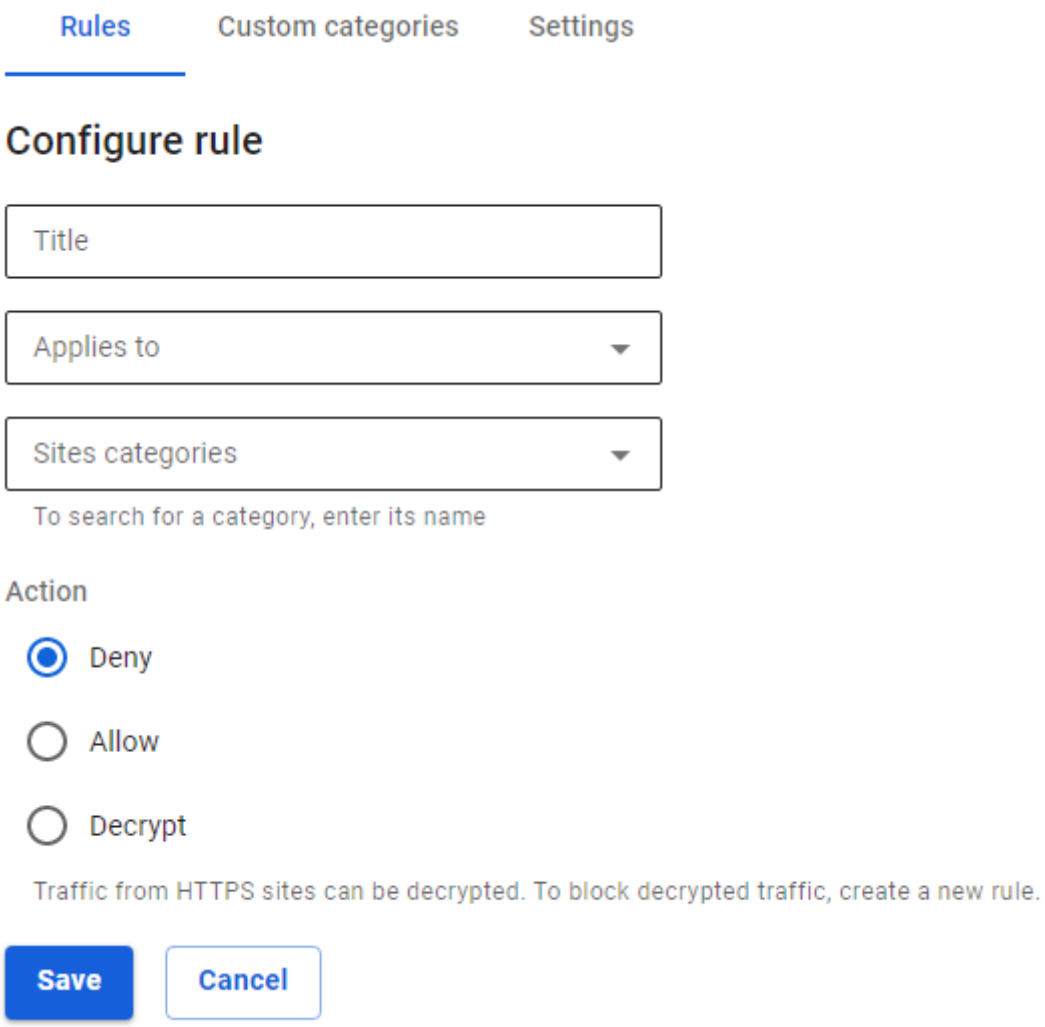
Applying Filtering Rules to Users

The rules are applied from top to bottom according to the order in the table until the first match. Thus, if the higher-level rule allows a certain resource for a specified user group, the lower-level rules will not be applied to it. This way more flexible filtering settings can be created, excluding desired users by higher-level rules from blocking rules. HTTPS decryption rules apply in a similar way.

Rules can be enabled, disabled, changed in priority, edited, and deleted in the **Operations** column. Content filtering rules are applied immediately after they are created and enabled.

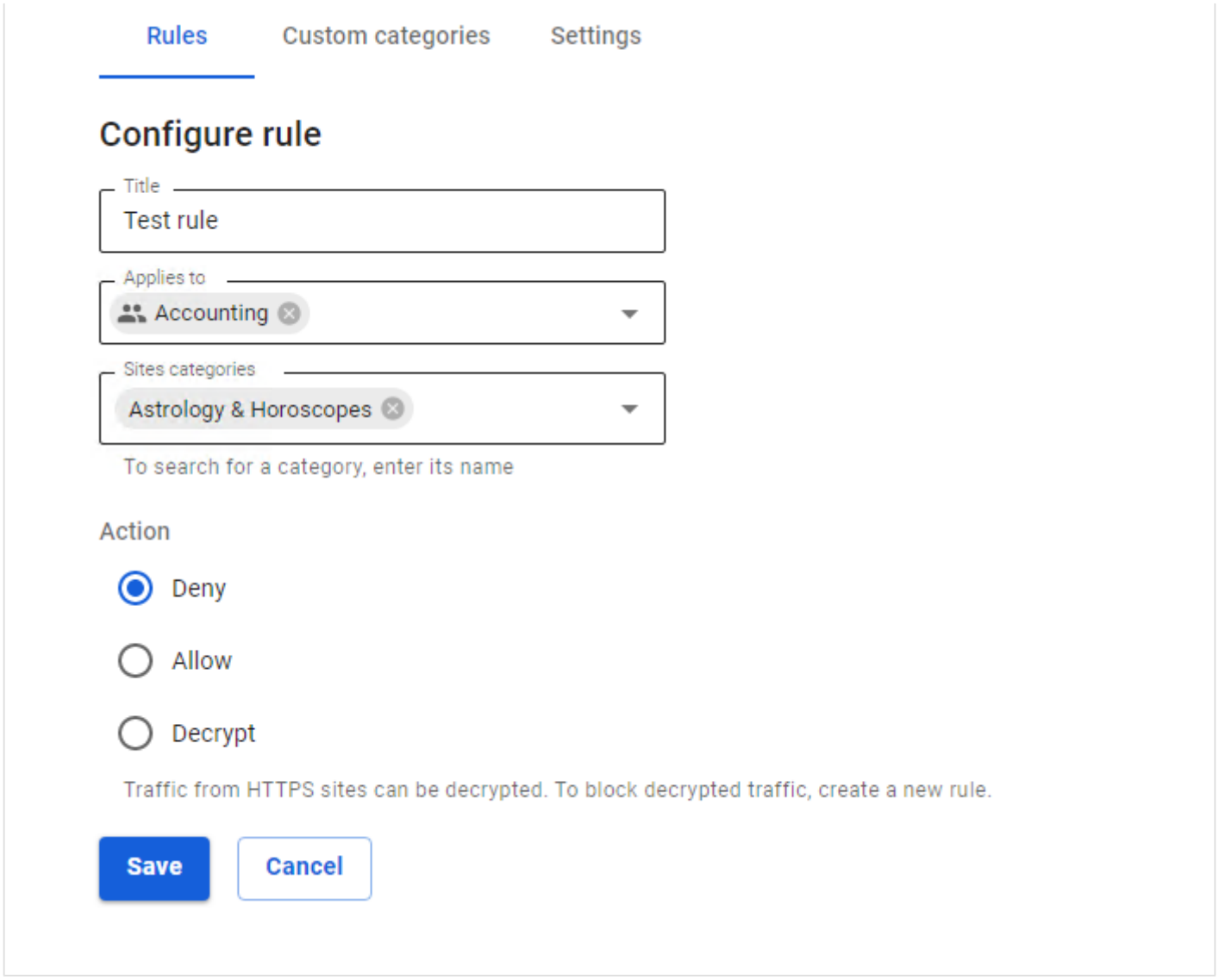


To create a new rule, click on **Add** in the left corner above the table.



Fill in the following fields:

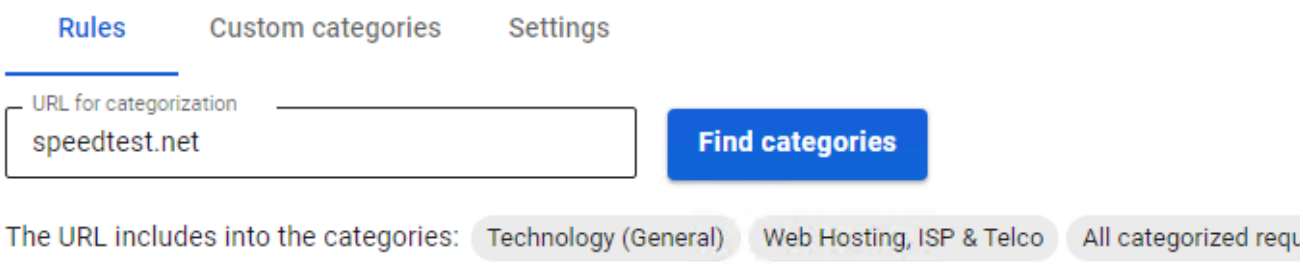
- **Title** – the rule name in the list. Maximum 42 characters.
- **Applies to** – you can select objects of the following types: user, user group, IP address, IP address range, subnet, list of IP addresses, or a special object **Quota Exceeded** (users who exceed traffic quota fall in this object).
- **Sites Categories** – user, special, and advanced web-resource categories.
- **Action** – the action of this rule towards web requests. You can prohibit, allow or decrypt HTTPS traffic.



Diagnostics

If content filtering rules are not working, check the following parameters in the settings:

1. The IP address of the user's computer must correspond to their address in authorization (section **Monitoring** – **Authorized users**), and the user must be in the group to which the rule applies.
2. The IP address of the user and the resource to which they access must not be included in the proxy server exceptions.
3. Check if the resource to which you are accessing is categorized correctly in the field **URL for Categorization** in the **Rules** tab.



If the site is incorrectly categorized, please use the **SafeDNS feedback form**.

4. VPN functions or plug-ins are not used in the user’s browser or computer; third-party proxy servers are not set.

Description of Content Filter Categories

The article describes in detail the categories of queries to web resources.

Special categories

- **All queries** - all queries to web resources fall under this category.
- **All categorized queries** - all queries to web resources categorized by built-in or custom categories fall under this category.
- **All non-categorized queries** - all queries to web resources that have not been categorized by built-in or custom categories fall under this category.
- **Direct access by IP** - queries to web resources by IP address (http://84.201.128.105 /).

Extended categories

	Category	Description
1	Abortion	Websites, that discuss abortions from medical, legal, historical, and other points of view
2	Abortion - Pro-Choice	Websites advocating the legal right to choose whether or not to have an abortion
3	Abortion - Pro Life	Websites condemning the use of abortion
4	Advocacy Groups & Trade Associations	Websites about industrial shopping groups, lobbyists, unions, professional organizations, and other associations, including communities of like-minded people
5	Agriculture	Websites about science, art, and business related to agriculture (production of grain crops, lifting livestock, products, services, etc.).
6	Alcohol	Websites calling for alcohol consumption (or justifying its use), as well as sites that sell alcoholic beverages, including beer, wine, etc.

7	Anonymizer	Websites designed to bypass network filters. Such resources can be used by company employees to visit prohibited websites
8	Architecture & Construction	Websites about the construction, design of buildings and structures, architecture, as well as organizations or services related to design, construction, and construction design
9	Art	Websites about fine art
10	Arts	Websites about art in general
11	Astrology & Horoscopes	Websites about astrology, horoscopes, as well as predictions on stars or zodiac signs
12	Atheism & Agnosticism	Websites leading anti-religious propaganda or questioning religious, spiritual, metaphysical, or supernatural views
13	Auctions & Marketplaces	Websites about sales of goods and services through ads, online auctions, or other non-traditional channels
14	Banking	Websites of banks and other credit institutions, including websites of Internet banks. This category does not include sites of organizations offering brokerage services
15	Biotechnology	Websites about studies in the field of genetics, as well as sites of research institutes and organizations working in the field of biotechnology
16	Botnet	Websites or compromised web resources on which the software used by hackers for spam mailings and the implementation of various Internet attacks is launched
17	Business/Services	Websites about business and services.
18	Businesses & Services (General)	Websites about business and services. This category includes resources that are not subject to more accurate categorization than business and services
19	Cars/Transportation	Websites about vehicles, including sales, promotion, discussion, manufacturers, and online stores

20	Cartoons, Anime & Comic Books	Websites with animation, cartoon TV shows, and comics
21	Catalogs	Websites with grocery lists and catalogs without the ability to make an online purchase
22	Chat	Online chats
23	Chat/IM	Online chats and messengers
24	Child Abuse Images	Websites with images of physical or sexual violence against children
25	Child Inappropriate	Materials inappropriate for children: tasteless, cruel (including, in relation to animals), toilet humor, etc.
26	Command and Control Centers	Internet servers used to manage botnets
27	Community Forums	Websites of forums, news groups, archives of mailing lists, announcement boards, and similar community resources
28	Community Sites	Social networks, as well as websites of various online communities
29	Compromised	Websites that were compromised by attackers and look like official websites, but actually contain malicious code
30	Computers & Technology	Websites about it, software, Internet, and computers
31	Content Servers	Websites that do not contain navigation elements and are usually used to place images or other media content in order to increase productivity and scalability
32	Contests & Surveys	Websites about online competitions, sales, and lotteries that are created to study consumer preferences, and can also be used as an element of various marketing activities
33	Coupons	Websites offering the acquisition of discount coupons
34	Criminal Skills	Websites providing information on how to commit illegal activity, such as theft, murder, creation of a bomb, opening locks, etc.

35	Criminal Skills/Hacking	Websites providing information about computer hacks
36	Dating & Relationships	Websites about acquaintances, marriage, etc.
37	Download Sites	Websites with software catalogs, including shareware, paid, free, and open-source software
38	Education	Websites related to learning
39	Educational Institutions	Websites of schools, universities, and other educational institutions
40	Educational Materials & Studies	Websites on which academic publications, magazines, research results, curricula, as well as online courses, textbooks, etc. are posted.
41	Entertainment and Videos	Websites with video and entertainment
42	Entertainment News & Celebrity Sites	Websites about news and gossip about celebrities, television shows, films, and show business in general
43	Entertainment Venues & Events	Websites about cultural institutions such as theaters, cinemas, nightclubs, festivals, etc.
44	Fashion & Beauty	Websites about fashion and beauty, including sites related to fashion and containing information about clothes, jewelry, cosmetics, and perfume
45	File Repositories	File sharing sites
46	Finance	Websites, that discuss economic issues, investment strategies, pension, and tax planning
47	Finance (General)	Websites, that discuss economic issues. This category includes resources that are not subject to more accurate categorization than finances.
48	Fitness & Recreation	Websites about fitness and other recreational activities
49	Food & Restaurants	Food sites: from restaurants and cafes to recipes and cooking tips
50	Gambling	Websites calling for participation in gambling (lotteries, casinos, etc.)
51	Games	Websites about computer games, as well as sites with online games

52	Gay, Lesbian, or Bisexual	Websites that discuss non-traditional sexual orientation
53	Government	Websites of state organizations
54	Government Sponsored	Websites of state organizations, including police, fire services, election commissions, research, and programs sponsored by the state
55	Hacking	Websites containing information or utilities that can be used to make online hacks
56	Hate Speech	Websites calling for extremism, discrimination on sexual, racial, religious, and other signs
57	Health	Websites about personal health
58	Health & Medical	Websites about personal health, medical services, medical equipment, procedures, mental health, hospitals, and clinics
59	High Risk	High threat sites
60	Hobbies & Leisure	Websites containing information about various crafts and hobbies, such as collecting, aircraft modeling, etc.
61	Home & Office Furnishings	Websites that include information about furniture manufacturers, and retail stores for the sale of furniture, tables, chairs, cabinets, etc.
62	Home, Garden & Family	Websites about family relationships and the house, including information about education, internal decoration, landscaping, cleaning, family, etc.
63	Home/Leisure	Websites about house and leisure
64	Humor	Websites containing humorous information, such as comics, jokes, funny pictures
65	Illegal Drugs	Websites about narcotic substances, including improper use of drugs
66	Image Search	Websites and search engines used to search for images and return results containing miniatures of the latter
67	Information Security	Websites of organizations providing information security services

68	Instant Messenger	Websites of instant messages, as well as websites used for advertising instant messengers on them
69	Insurance	Websites about all types of insurance, including medical, state, property insurance, etc.
70	Internet Phone & VoIP	Websites allowing to make calls via the web or programs of software products that are designed to make calls via the Internet
71	Job Search	Websites about the search for work, including recruiting agencies
72	Kid's Pages	Websites designed for young children (up to 10 years old), including games and entertainment pages
73	Legislation, Politics & Law	Websites on legislation, politics, parties, elections, their results, and opinions
74	Lingerie, Suggestive & Pinup	Websites with photos and videos that depict women in sexy provocative clothes, for example, in lingerie
75	Literature & Books	Websites on which literature is presented, including fiction and documentary novels, poems, and biographies
76	Login Screens	Websites that are used for a single authentication and access to a wide variety of services. For example, systems such as Yahoo or Google
77	Malware Call-Home	Malicious event when active malware on a computer attempts to contact a remote “home” server
78	Malware Distribution Point	Websites with viruses, exploits, and other malicious programs
79	Manufacturing	Websites about business related to industrial production
80	Marijuana	Websites on which information about marijuana, its cultivation, or smoking is presented, including sites about the legal use of marijuana, for example, in medicine
81	Marketing Services	Websites of advertising and marketing agencies
82	Mature	Mature content

83	Medium Risk	Websites that pose the average threat
84	Military	Websites sponsored by the armed forces and other state military institutions
85	Miscellaneous	Websites that cannot be unambiguously attributed to any of the categories
86	Mobile operators paid sites	Paid websites of mobile operators
87	Mobile Phones	Websites of mobile phone manufacturers, including sites selling mobile phones and accessories for them
88	Motorized Vehicles	Websites about transport with engine
89	Music	Websites about music. Internet radio, files in mp3 format, information about music groups, clips, etc.
90	Nature & Conservation	Websites with environmental information, ecology, etc.
91	News	News web resources. Online publications of newspapers, magazines, news feeds
92	No Content Found	Websites with unrecognizable content, which does not allow categorization of them
93	No Known Risk	Websites that do not pose threats and do not fall into other categories
94	Non-profits	Websites of nonprofit organizations
95	Non-traditional Religion & Occult	Websites about religions that are not in the mainstream or not included in the top10 of world religions (folk religions, mysticism, cults, and sects)
96	Nudity	Websites containing erotic materials (partial or complete exposure), excluding pornographic materials
97	Nutrition & Diet	Websites with information about healthy diets, weight loss, weight loss programs, and food allergies
98	Online Ads	Web pages strictly about advertising, banners, or popup windows with advertising

99	Online Financial Tools & Quotes	Websites containing information about financial quotes, as well as tools for financial analysis and budget planning, such as mortgage calculators, software for tax reporting, etc.
100	Online Information Management	Websites about programs for managing personal information, for example, applications for managing tasks, calendars, address books, etc.
101	Online Shopping	Online stores and other sites offering things online
102	Online Stock Trading	Websites of brokerage companies that carry out online securities, etc.
103	Parked	Websites that are used as “plugs” for acquired but not used domain names
104	Parks, Rec Facilities & Gyms	Websites about parks and other zones intended for wellness activities, such as swimming, skateboarding, mountaineering, etc.
105	Pay to Surf	Websites of companies offering to view advertising in their specialized applications
106	Peer-to-Peer	Peering networks sites
107	Personal Pages & Blogs	Personal and lifestyle content
108	Personal Storage	Websites for storing personal files
109	Personal Webpages	Personal pages, including blogs and other means of exchange of news, opinions, and information about the author, as well as home and family pages
110	Pets & Animals	Websites containing information, products, and services for pets
111	Pharmacy	Websites containing information about drugs (including legal narcotic substances), as well as their use
112	Philanthropic Organizations	Websites with information about charitable institutions and other non-profit philanthropic organizations

113	Phishing/Fraud	Websites used for fraud are also known as phishing. Usually, they look like official web pages of financial or other institutions with the aim of unauthorized access to confidential information, for example, CVV codes of bank cards
114	Photo Sharing	Websites on which users can place digital photos, as well as search for images, exchange them, etc.
115	Physical Security	Websites related to products and services regarding security, with the exception of computer security
116	Piracy & Copyright Theft	Websites that provide access to an illegal content, for example, pirated software (Warez), pirated films, music, etc.
117	Politics and Law	Websites on politics and legislation
118	Pornography	Websites containing images or videos with a demonstration of sexual intercourse or a naked body
119	Pornography/Sex	Websites containing images or videos with a naked body
120	Portal Sites	Web resources that provide access to custom-made personal portals, including yellow pages and other catalogs
121	Possible Risk	Websites with the possibly risky contents
122	Private IP Address	Websites served on private IP addresses reserved for use within organizations and houses
123	Product Reviews & Price Comparisons	Websites designed to help customers compare shops, products, and prices, but not selling online
124	Profanity	Websites containing episodic or serious swearing or blasphemy
125	Professional Networking	Websites of social networks focused on professionals and building business relations
126	R-Rated	Websites, the content of which must contain material intended only for an adult audience. There can be sexual topics or training materials

127	Real Estate	Websites about real estate issues (purchase, sale, rent, etc.)
128	Redirect	Websites that redirect visitors to other resources
129	Reference Materials & Maps	Websites containing reference materials and data sets: atlases, dictionaries, encyclopedias, census, etc.
130	Religion	Websites about a specific religion
131	Religions	Websites about the main world religions, as well as general religious and theological topics
132	Remote access	Websites that provide remote access to private computers and networks, intranet resources (files and web applications)
133	Restaurants	Restaurant websites
134	Retirement Homes & Assisted Living	Websites about nursing houses and thematic communities, including patient care and hospice assistance
135	School Cheating	Websites with answers to tests, ready-made works, step-by-step solutions to problems, and similar resources
136	Search Engines	Search systems that search for websites, news groups, pictures, and other content
137	Self-help & Addiction	Websites offering information and assistance in alcohol, drug, gaming dependencies, as well as eating disorders (anorexia, etc.)
138	Sex & Erotic	Websites offering products and services related to sex, but not containing naked nature and other candid images
139	Sex Education & Pregnancy	Websites with teaching materials and clinical explanations about sex, safe sex, pregnancy, childbirth, etc.
140	Shipping & Logistics	Websites on stock management, including transportation, warehouse, distribution, storage, execution, and delivery of orders
141	Shopping	Online shopping and purchases

142	Sites from the list of the Ministery of Justice	Additional blocklist for specific regions
143	Social and Affiliation Organizations	Websites of social and affiliated organizations
144	Social Networking	Websites of social networks - communities in which people are "friends"
145	Software, Hardware & Electronics	Site about computer equipment, software, peripherals, data networks, and electronics, as well as manufacturers of appropriate goods and services
146	Spam	Websites advertised using spam
147	Sport and Recreation	Websites about training and competitions in martial arts: boxing, struggle, fencing, etc.
148	Spyware & Questionable Software	Websites with software sending information to the central server, including spy software and keyboard spies
149	Spyware and Malicious Sites	Websites that are spying, sending information about the visitor to a special address
150	Streaming & Downloadable Audio	Storage sites broadcasting music or other audio content (can consume the entire available bandwidth of the company's channel)
151	Streaming & Downloadable Video	Storage sites broadcasting videos, including streaming (can consume the entire available bandwidth of the company's channel)
152	Supplements & Compounds	Websites containing information about vitamins and other substances of unregulated turnover
153	Swimsuits	Websites containing images of people in bathing suits. Images of costumes themselves do not fall into this category
154	Technology (General)	Websites about web design, standardization on the Internet (for example, RFC), protocol specifications, news, and other wide discussions of technology

155	Television & Movies	Websites about television shows and films, including reviews, programs, plots, discussions, trailers, marketing, etc.
156	Text Messaging & SMS	Websites designed to exchange short text messages (SMS) between a web page and a mobile phone
157	Tobacco	Websites about tobacco products (cigarettes, cigars, vapes, etc.)
158	Torrent Repository	Websites that host torrent files that allow you to download potentially large files via P2P networks
159	Toys	Websites of toy manufacturers, as well as marketing resources and online toy stores
160	Translator	Dictionaries and translators from foreign languages
161	Travel	Websites about travel information and tourism, as well as online orders of plane tickets, hotels, cars, etc.
162	Unknown Sites	Websites without category
163	Unreachable	Websites that display errors such as “The connection time expired”, “the address is not found”, etc.
164	Violence	Websites about dubious actions, such as violence and aggression
165	Weapons	Websites about weapons
166	Web Hosting, ISP & Telco	Websites offering web-hosting services, blogging blogs, Internet providers, and telecommunication companies
167	Web-based Email	Services providing web access to mailboxes
168	Web-based Greeting Cards	Websites that allow users to send and accept postcards online
169	Wikis	Websites and resources of communities creating information documents available for editing for all participants

Setting up HTTPS Filtering

HTTPS traffic filtering provides the possibility for further processing of websites accessible via HTTPS.

Filtering is implemented by several methods:

- Analysis of Server Name Indication (SNI) headings. Thanks to this method, it is possible to analyze the domain to which the client connects without certificate substitution and interference with HTTPS traffic. Domains specified in the certificate are also analyzed.
- SSL-Bump Method. Filtering occurs by substitution on the fly of the certificate used to sign the requested site. The original site certificate is substituted by a new one signed by the SafeUTM root certificate instead of a certificate authority. Thus, the traffic transmitted over a secure HTTPS connection becomes available for processing by all modules provided by SafeUTM, namely by the content filter (it is possible to categorize the full URL of the query and MIME-type content), ClamAV, as well as external ICAP services.

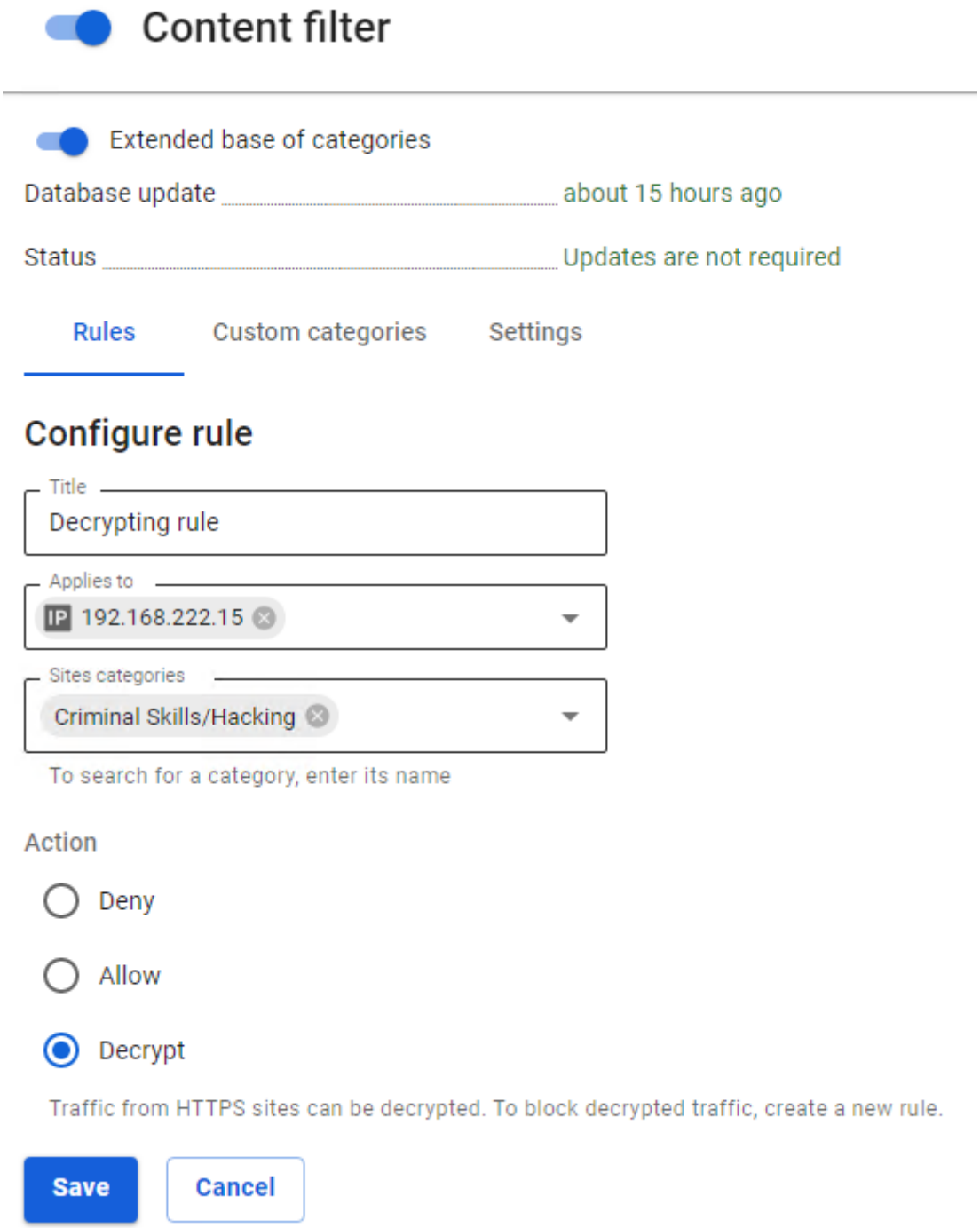
> The specifics of implementing HTTPS traffic filtering with certificate substitution require configuring both sides of the connection – the SafeUTM server and each user’s workstation in the local network.

Setting up SafeUTM Server

By default, the server performs HTTPS filtering without certificate substitution by analyzing SNI and domains in the certificate.

HTTPS traffic decryption is configured in **Rules -> Content Filter -> Rules** using the rules created by the admin with the action **Decrypt**.

An example of a decryption rule can be seen below:

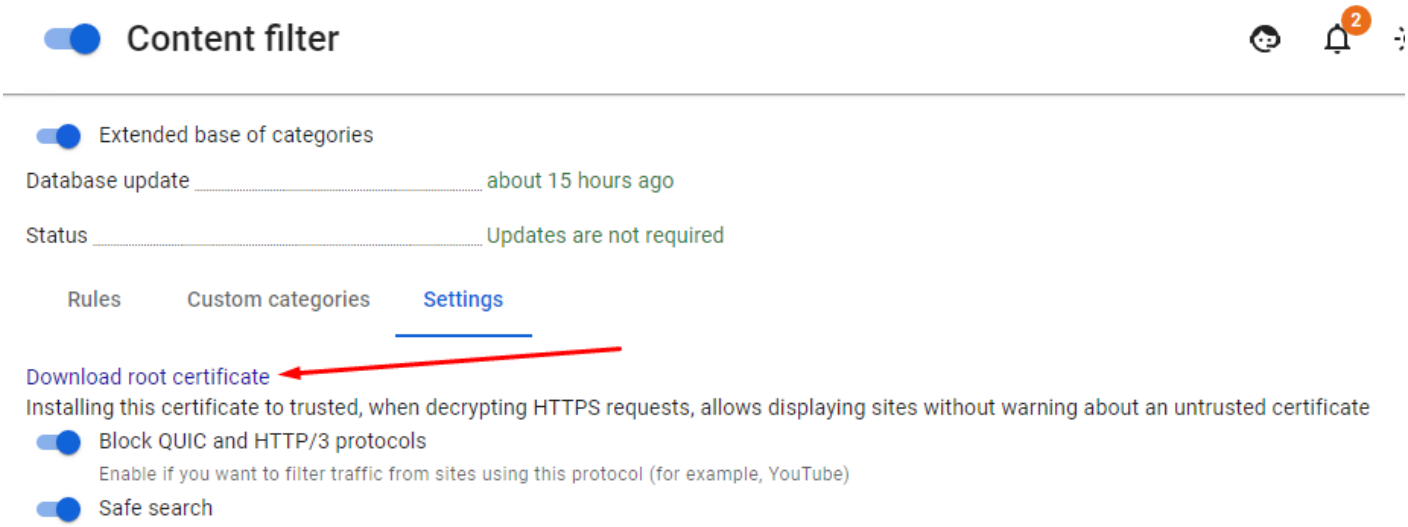


Setting up the User's Workstation

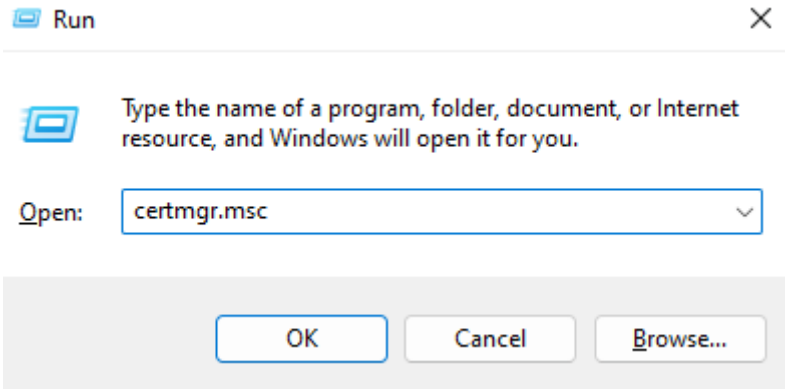
When the HTTPS traffic decryption option is enabled, the browser and other network software (for example, antiviruses, IM clients, etc.) on the user's workstation will require explicit confirmation to use a substitute certificate created and issued by the SafeUTM server. To improve the user's convenience, the SafeUTM server's root certificate should be installed in the workstation's operating system and made trusted. The root SSL certificate is available for download from the section **Traffic Rules -> Content Filter -> Settings**.

In order to install the root certificate, you need to follow these steps:

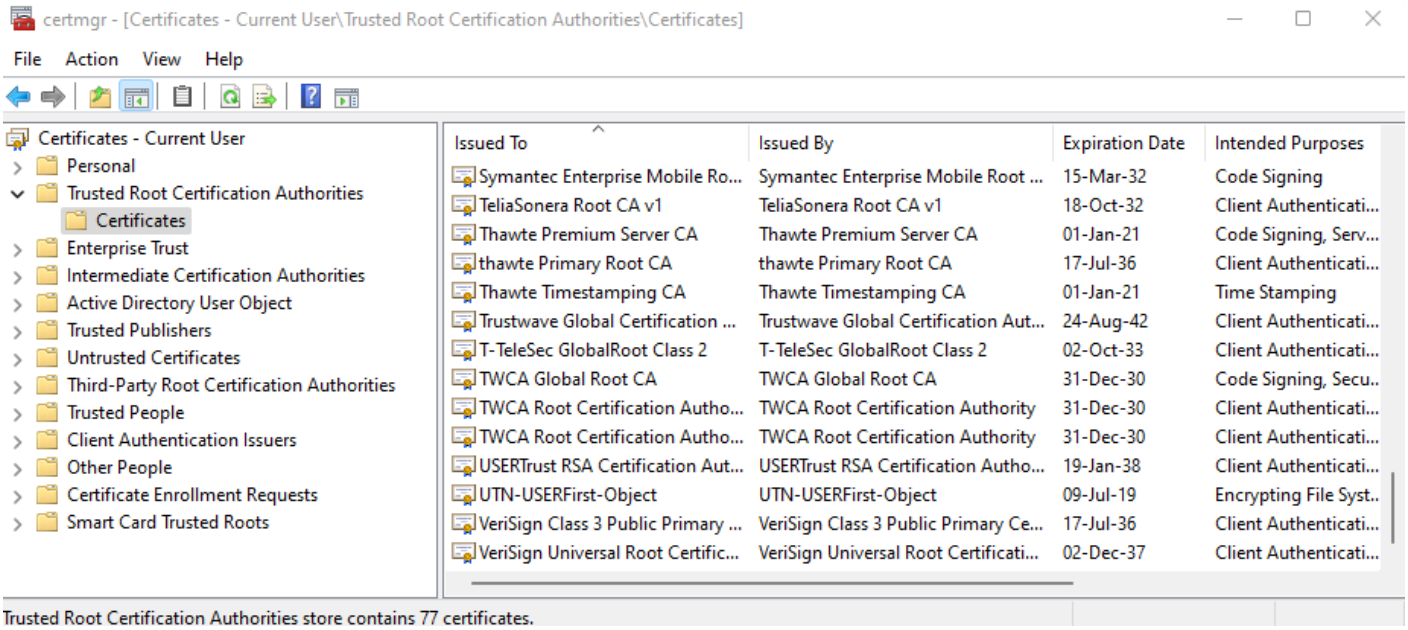
1. Download the root SSL certificate by opening the SafeUTM web interface section **Traffic Rules -> Content Filtering -> Settings**:



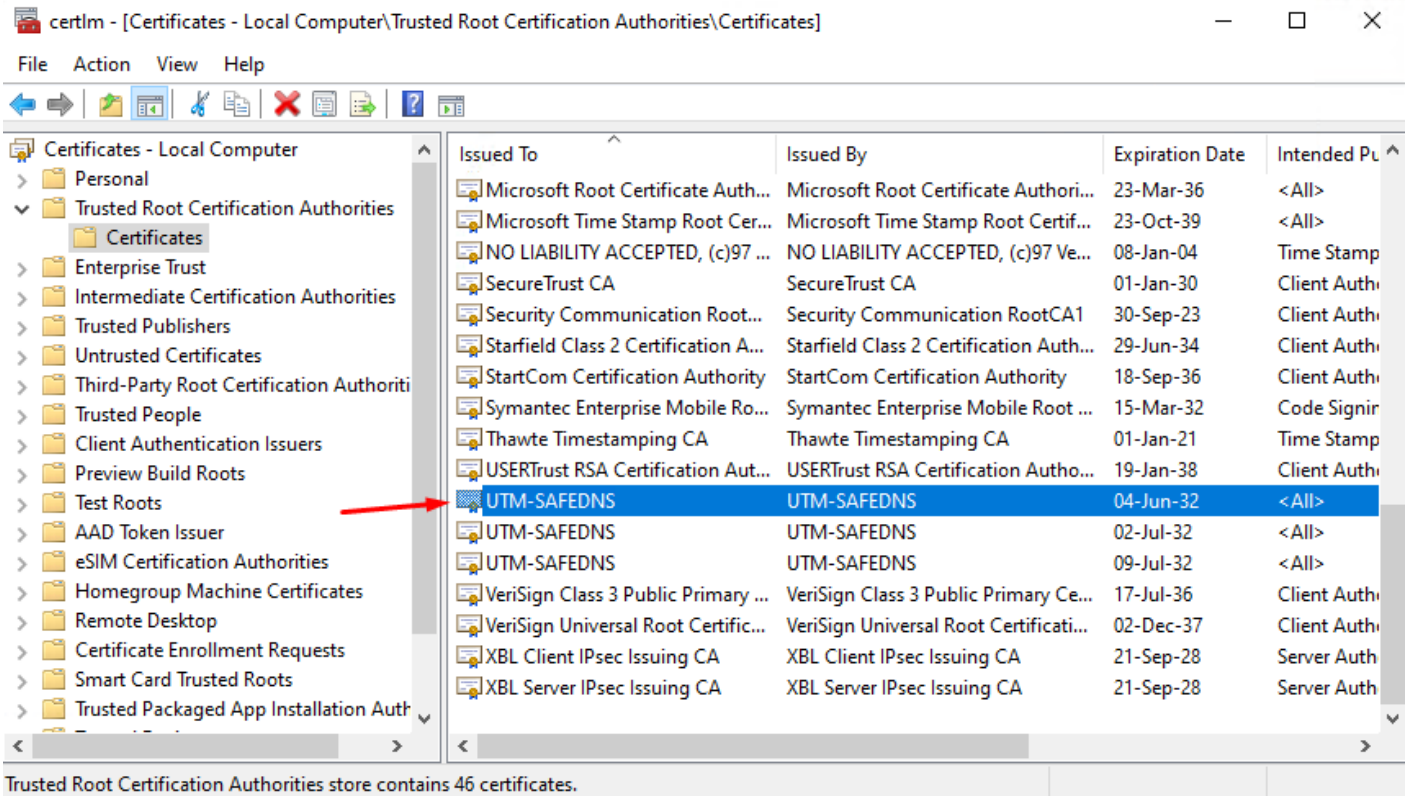
2. Open the certificate management center on the workstation **Start -> Run** by executing the command **certmgr.msc** in the dialog:



3. Select the section **Trusted Root Certificates -> Certificates**:



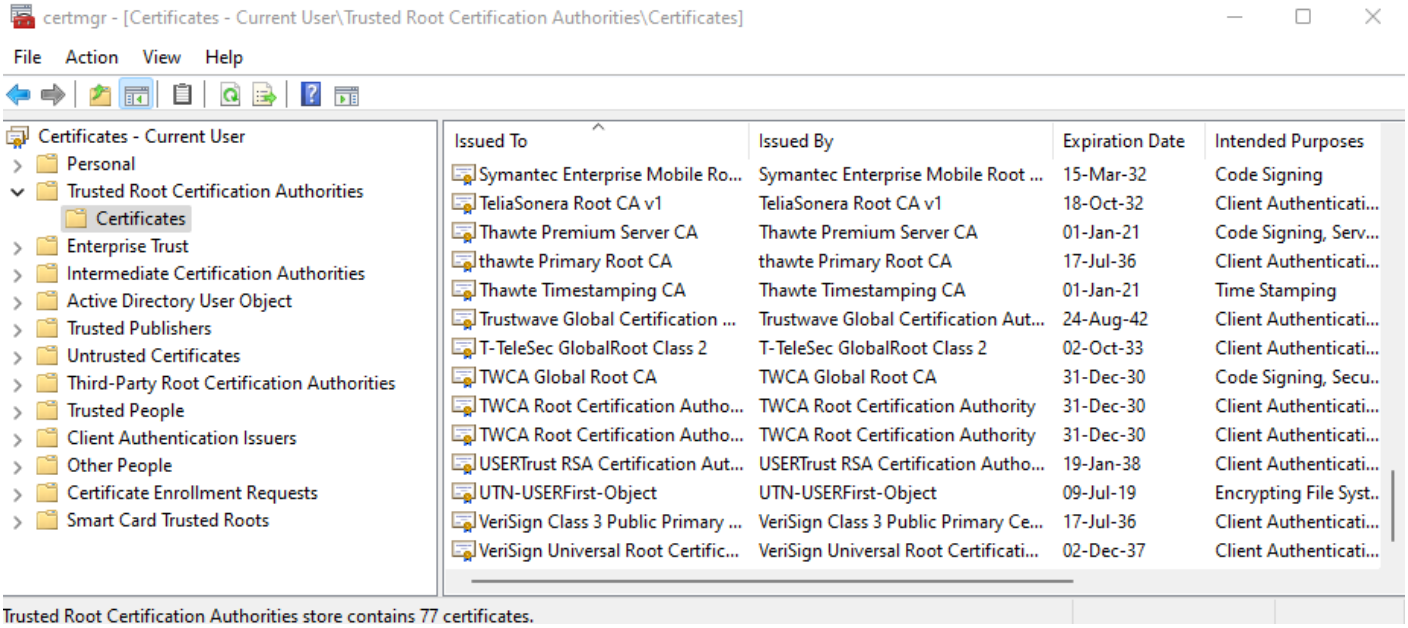
4. In the right part of the window, right-click and select action **All Tasks -> Import...** The Certificate Import Wizard will open. Follow the wizard's instructions to import the SafeUTM server's root certificate. The imported certificate will appear on the list in the right part of the window:



Adding Certificate via Microsoft Active Directory Domain Policies

In networks where users are managed using Microsoft Active Directory, you can install a SafeUTM certificate for all users automatically using Active Directory. To do this, follow these steps:

1. Download the root SSL certificate by opening the SafeUTM interface section **Access rules -> Content Filtering -> Settings**:



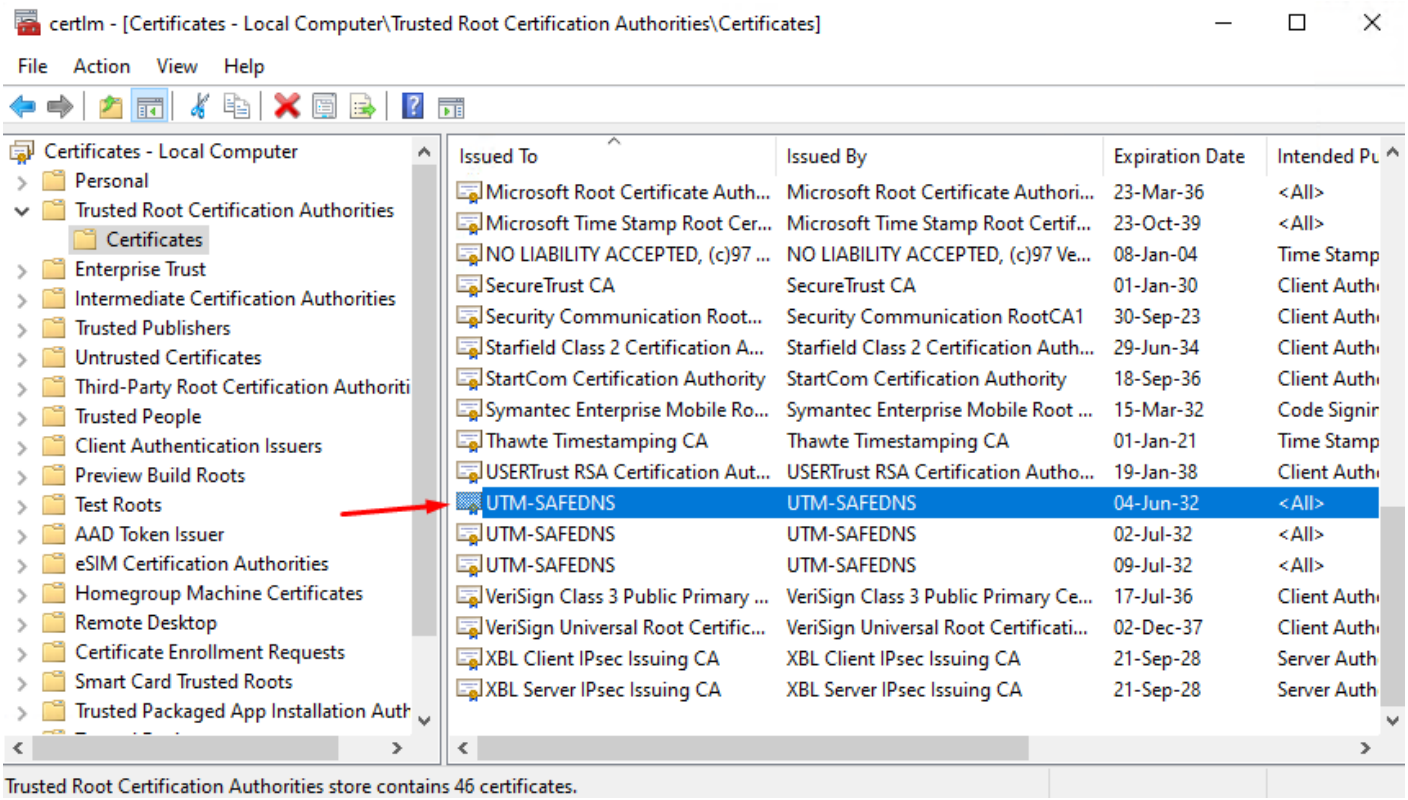
2. Log in to the domain controller with administrator privileges.

3. Launch the group policy management snap-in by executing the command **gpmc.msc**.

4. Find **the domain policy** used on users' computers in **Group Policy Objects** (Default Domain Policy in the screenshot). Right-click on it and select **Change**.

5. In the group policy management editor that opens, select: **Computer Configuration -> Policies -> Windows Configuration -> Security Settings -> Public Key Policies -> Trusted Root Certificate Authorities**.

6. Right-click on the list that opens, select **Import...** and import the SafeUTM key.



7. After restarting workstations or executing the command **gpupdate /force** on them, the certificate will appear in the local certificate stores and the required level of trust will be established for it.

Possible Problems and Troubleshooting

- Some browsers, such as Mozilla Firefox, do not use the system certificate store, in which case it is necessary to add a SafeUTM certificate to the browser’s trusted certificates. In Firefox, you can also set the parameter **security.enterprise_roots.enabled** (in **about:config**) to **true** for trusting system certificates.
- If the local machine uses an antivirus that checks HTTPS traffic using certificate substitution, sites may not open because of double certificate substitution. HTTPS traffic check must be disabled in the antivirus settings.
- With SNI filtering enabled, the server will not allow non-HTTPS traffic through the HTTPS port. Thus, problems with programs trying to do this may occur. In order for them to run, is necessary to allow bypassing the proxy server to the resources they require.
- When blocking HTTPS resources, in order to display the blocking page, it is necessary to set up trusting the UTM root SSL certificate even if only SNI filtering is enabled, as in case the resource opened via HTTPS is blocked, SSL-bumping with UTM SSL-certificate replacement will be applied for the possibility of replacing the resource content with the page about it being blocked by the server.