QKits Electronics

User Guide

SearchAbuse Log Scanner for Magento 2
Module: QKits_SearchAbuseLogScanner
Version: 1.0.0
SKU: QKMOD-SL
Platform: Magento 2.4.x
Website: https://modules.qkits.com
Support: support@qkits.com

Table of Contents

1. Overview

The SearchAbuse Log Scanner analyzes your server's access logs to detect and block security threats that go beyond search abuse. While the base SearchAbuse module protects your search functionality, the Log Scanner watches all HTTP traffic for malicious patterns including vulnerability scanning, admin probing, webshell attempts, and brute-force attacks.

The Log Scanner can run automatically via cron (default: every 15 minutes) or on-demand from the admin panel. Detected threats can be automatically banned using SearchAbuse's progressive penalty system and .htaccess blocking.

2. Threat Detection Categories

The Log Scanner identifies threats across nine categories. Each flagged IP is tagged with all matching categories, giving you a clear picture of the attack type.

Category	What It Detects
Search Abuse	SQL injection, XSS, shell commands, and blocked terms in search queries
Webshell Probe	Attempts to access known backdoor files (shell.php, c99.php, r57.php, etc.)
CMS Probe	WordPress, Joomla, Drupal paths on a Magento site (wp-admin, wp-login, etc.)
Admin Probe	Scanning for admin panels, phpMyAdmin, Adminer, setup wizards
Config/Env Probe	Hunting for .env, .git, config.php, database dumps, backup files
Suspicious UA	Known scanner user agents (sqlmap, nikto, masscan, zgrab, etc.)
High Frequency	Excessive requests per IP within the lookback window
Error Scanner	High ratio of 403/404 errors indicating automated vulnerability scanning
Path Traversal	Directory traversal attempts (../, etc/passwd, proc/self)

3. Admin Log Scanner Page

Access the Log Scanner from the admin sidebar: Search Abuse > Log Scanner.

The page displays available log files from your configured log directory. Both live log files (transfer.log, access.log) and archived/zipped logs (transfer-2026-03-01.log.zip) are listed with their file size and last modified date.

Select one or more log files and click Scan Selected Files to analyze them. The scanner processes all selected files and merges the results, deduplicating IPs that appear across multiple files.

4. Running Scans

4.1 Admin Panel Scan

From the Log Scanner page, select log files and click Scan. Admin scans always process the entire file regardless of the lookback window setting. This lets you scan historical logs for patterns.

4.2 Cron Scan

When enabled, the cron job runs on the configured schedule (default: every 15 minutes). Cron scans only look back the configured number of minutes (default: 60) to avoid re-processing old data. The cron job automatically bans flagged IPs if auto-ban is enabled.

4.3 CLI Scan

Run a scan from the command line:

```
bin/magento qkits:searchabuse:scan --lookback 60
```

Options:

- --lookback [minutes] : How far back to scan (default: 60, use 0 for full file)
- --auto-ban : Automatically ban flagged IPs
- --dry-run : Show what would be banned without banning

5. Reading Scan Results

After a scan completes, the results page shows:

5.1 Summary Bar

Total requests analyzed, lines scanned, unique IPs, and number of abusive IPs found.

5.2 Flagged IPs Table

Each flagged IP is shown with:

- IP address
- Color-coded threat category tags (e.g. red for Webshell, orange for Admin Probe)
- Total request count
- Error count and error rate percentage
- User agent string
- Sample URLs showing the suspicious requests
- Ban status (already banned, whitelisted, or available to ban)

5.3 Banning from Results

Select IPs from the results table using the checkboxes and click Ban Selected IPs. Whitelisted IPs are automatically skipped. Banned IPs are immediately added to SearchAbuse's ban list and .htaccess if auto-update is enabled.

6. Auto-Ban and Cron Scanning

When both auto-ban and cron scanning are enabled, the Log Scanner operates fully automatically:

- Cron runs every 15 minutes (configurable)
- Scans the last 60 minutes of access log entries (configurable)
- Flags IPs matching 2+ threat categories (configurable threshold)
- Automatically bans flagged IPs using SearchAbuse's progressive penalty system
- Skips already-banned and whitelisted IPs
- Updates .htaccess Deny rules immediately if enabled

The auto-ban minimum categories setting controls how aggressive the auto-ban is. The default of 2 means an IP must match at least two different threat categories before being auto-banned. Set to 1 for maximum protection (any single detection triggers a ban), or higher for fewer false positives.

7. Configuration Reference

All settings are under Stores > Configuration > QKits > Search Abuse Protection > Log Scanner.

Setting	Default	Description
Enable Log Scanner Cron	Yes	Run automated scanning on schedule
Log File Path	(empty)	Absolute path to server access log
Auto-Ban Abusive IPs	Yes	Automatically ban flagged IPs
Auto-Ban Min Categories	2	Minimum threat categories to trigger ban
Request Frequency Threshold	20	Requests per IP to flag as high frequency
Error Count Threshold	10	Minimum 403/404 errors to flag
High Freq Error Rate (%)	20	Error rate filter for high frequency IPs
Error Scanner Rate (%)	50	Error rate to flag as scanner
Flag Suspicious UA Alone	No	Whether UA alone triggers flagging
Lookback Window (min)	60	How far back cron scans
Cron Schedule	*/15 * * * *	Cron expression for scan frequency

8. CLI Commands

The Log Scanner adds one CLI command:

```
bin/magento qkits:searchabuse:scan
```

Option	Description
--lookback [min]	Scan last N minutes (0 = full file)
--auto-ban	Ban flagged IPs automatically
--dry-run	Show results without banning
--file [path]	Scan a specific file instead of configured path

9. How Detection Works

The Log Scanner parses each line of the access log using the Combined Log Format (IP, timestamp, request method, URL, status code, user agent). Each request is checked against pattern libraries for all nine threat categories.

IPs are scored per-category. An IP that triggers multiple categories is more likely to be a genuine threat. The auto-ban minimum categories threshold (default: 2) prevents banning IPs that only trigger a single low-confidence detection.

The High Frequency and Error Scanner categories use rate-based detection with configurable thresholds. A customer browsing 200 pages with a 0.5% error rate will not be flagged, but a scanner hitting 200 URLs with a 95% error rate will.

All detections respect the SearchAbuse whitelist. Whitelisted IPs are never flagged or banned, even if they match threat patterns. Always whitelist your own IPs, monitoring services, and payment processor callback IPs.

10. Best Practices

- **Whitelist first.** Before enabling auto-ban, add your office IP, monitoring services, and payment gateways (Stripe, PayPal) to the whitelist.
- **Start with min categories = 2.** This prevents false positives from single-category detections like a customer with a slightly unusual user agent.
- **Review results before enabling auto-ban.** Run a few manual scans first to understand what the scanner detects in your specific traffic patterns.
- **Scan archived logs.** Use the admin panel to scan historical log files and identify IPs that have been attacking your site for weeks or months.
- **Monitor var/log/searchabuse.log.** The scanner logs all cron activity, bans, and any errors to this file.
- **Keep the lookback window reasonable.** The default 60 minutes is good for most sites. Shorter windows scan faster but may miss slow attackers.