Protection System for Analysis of External Link Placing

Ivan Liminovych¹, Vadym Poltorak², Nadia Kushnir³, Bohdan Zhurakovskyy¹, and Sergiy Obushnyi⁴

¹ Olimp Digital, 17b Mala Berdnychivska str., Zhytomyr, 10014, Ukraine
² Igor Sikorsky Kyiv Polytechnic Institute, 37 Beresteiskiy pros., Kyiv, 03056, Ukraine
³ Zhytomyr Polytechnic State University, 103 Chudnivska str., Zhytomyr, 10005, Ukraine
⁴ Borys Grinchenko Kyiv Metropolitan University, 18/2 Bulvarno-Kudriavska str., Kyiv, 04053, Ukraine

Abstract
In the rapidly evolving digital landscape, effective Internet marketing strategies, particularly Search Engine Optimization (SEO), have become crucial for businesses. A novel system for optimizing SEO through strategic link placement on external internet resources is presented. At its core, an innovative algorithm analytically evaluates potential link-hosting platforms, focusing on key SEO metrics for optimal search engine visibility and authority. Emphasizing security, the system utilizes the Laravel PHP framework to guard against cyber threats, safeguarding user data integrity. This system offers a blend of enhanced SEO efficacy and robust security measures, revolutionizing online marketing through precise analysis, advanced security, and flexible user management.

Keywords
Search engine optimization, link analysis, security, data protection.

1. Link Analysis System

In our time, when digital technologies and information systems play a crucial role in shaping business strategies, the significance of Internet marketing and the effective use of SEO tools is rapidly increasing [1, 2].

SEO is a critically important aspect of modern digital marketing, playing a significant role in shaping business strategies. This process includes a series of actions aimed at increasing a website’s visibility in search engine results like Google. The primary goal of SEO is to increase organic traffic to the site by ensuring high positions in search results for key queries [3, 4]. Such positioning is important as it ensures high brand visibility, attracts the target audience, and enhances the competitiveness of the enterprise in the digital space [5].

Scientific research in the field of digital marketing confirms that effective SEO use not only contributes to increased traffic and improved online presence but also serves as a tool for creating a better user experience through content and website structure optimization [6]. This, in turn, leads to increased brand trust and long-term strengthening of the company's market position. Thus, SEO is not just a marketing tool but a comprehensive strategy that contributes to the holistic development of a business in the digital economy [5–7].

1.1. Link Analysis System Goal

An SEO optimization strategy that includes placing external links becomes a key element in increasing a website’s visibility in search engines. This is important because a high position in search results can significantly increase web resource traffic, as well as its authority and recognition in the digital world. The importance of choosing quality and relevant external resources for links cannot be underestimated, as they affect the site’s
perception by search engines and its final placement in the search output. Attention must be paid to the context in which links are placed and their relevance to the overall site theme to avoid a negative impact on the SEO rating. Improperly chosen resources can lead not only to loss of potential traffic but also negatively affect the website's reputation, causing skepticism among users and search engines [7].

In the context of developing a system for effective analysis of SEO parameters, the use of an algorithmic approach to evaluate potential platforms for placing links is important. Such a system includes a comprehensive analysis of key aspects: from the authority and popularity of the domain, its history, and traffic, to the analysis of content published on the site, its quality, uniqueness, and relevance to the target audience [8]. This comprehensive approach allows for a deeper dive into the context of link placement, assessing not only their potential impact on SEO but also their overall alignment with the brand or company's marketing strategy. The analysis may also include checking for negative SEO factors, such as spam links or artificially excessive optimization, which can lead to a site's rating decline. Using such a system minimizes the risks associated with choosing unreliable or ineffective resources for link placement, thereby ensuring more stable and effective growth in search engine rankings.

It's worth mentioning a functionally similar system like LinkChecker Pro [9].

LinkChecker Pro is a backlink management tool designed for various users, including link-building specialists, SEO experts, project managers, and business owners. It provides the ability to monitor and manage the status of backlinks, Google indexing, rel attributes, robots meta tags, and other parameters. The tool also provides notifications and reports on changes in the link status, access to detailed analytics, and the ability to conduct bulk analysis. The application interface is shown in Fig. 1.

The program offers several tariff plans, starting from $21 per month for small businesses and marketing beginners, to plans for large agencies and enterprises, which can be customized according to their needs. LinkChecker Pro users note its advantages such as reducing stress, automating routine tasks, improving team management, and supporting only quality backlinks.

1.2. Link Analysis Basics

The innovation of the proposed system lies in its comprehensive approach to analyzing and selecting resources for link placement. The use of advanced technologies to obtain various SEO metrics from internet resources where links are planned to be placed ensures its relevance and effectiveness in the long term. The system can analyze these metrics and inform about the advisability or inadvisability of placing links on resources. This approach not only enhances the effectiveness of SEO strategies but also prevents financial losses due to the use of inappropriate resources that will not provide the expected result [10].

The link analysis consists of two stages—checking page indexing indicators for search engines and taking statistical page indicators. The first stage includes indicators such as:

- Whether the page is found: relates to the HTTP response status that the page returns when trying to access it, whether the page exists and loads successfully.
- Whether the page has the <meta name="robots" content="noindex, nofollow"> tag (or individually noindex, and nofollow): noindex instructs search engines not to index the page, i.e., not to add it to their search results, nofollow indicates that search engines should not follow the links on this page [11].
- Whether the page has an x-robots status noindex: This is an HTTP header that can also be used to instruct search engines. The X-Robots-Tag header with a noindex value instructs not to index this page.
- Whether the page is closed in robots.txt: this is a file on the server that gives instructions to web crawlers about which pages or sections of the site they should or should not visit. If a page is "closed" in

![Figure 1: LinkChecker Pro interface](image)
robots.txt, it means that web crawlers are instructed not to visit this page.

- Whether the page corresponds to the canonical tag: this tag is used to indicate the main version of the page if there are duplicates or very similar versions.
- Whether the page is indexed in Google: this means whether the page has been added to Google's index and can appear in Google's search results.

These page indicators generally relate to the indexing and accessibility of the page for search engines. They include detecting pages, checking for meta tags that control indexing, checking for compliance with canonical tags, and Google indexing status.

The second stage involves obtaining indicators related to search engine optimization and website authority. Each of them plays an important role in determining the popularity, reliability, and visibility of the website in search engines. These indicators include:

- Domain Rating (DR): This indicator reflects the authority of the domain based on the quality and quantity of external links pointing to it. A higher DR rating indicates that the website has more high-quality inbound links, which can positively affect its position in search engines [12].
- Referring Domains (RD): This is the number of unique domains that link to the website. This indicator is important because search engines consider sites with a large number of inbound links from different domains to be more authoritative and reliable [13].
- Pages: This is the total number of web pages on the site. A larger number of pages can mean more content for indexing by search engines, which can increase the site's visibility.
- Organic Keywords (OK): This indicator reflects the number of keywords for which the site appears in search queries in search engines. A higher number of organic keywords can indicate greater visibility and attractiveness of the site [14].
- Organic Traffic (OT): This is the number of website visits obtained through free search queries. High organic traffic is an indicator of the effectiveness of SEO strategies and the site's popularity among users [15].
- URL Rating (UR): This indicator is similar to DR but assesses the authority of a specific web page, not the entire domain. It is based on the quality and quantity of links pointing to this page [16].
- Trust Flow (TF): This metric measures the reliability of a site based on the quality of the links pointing to it. Sites with high TF are usually associated with high-quality and reliable sources [17].
- Citation Flow (CF): This indicator assesses the influence of a website based on the number of links pointing to it, regardless of the quality of these links. A higher CF indicates a greater number of links but not necessarily high quality [18].

To obtain data related to search engine optimization and website authority, APIs from Ahrefs and Majestic were used. Both of these tools are integral in the analysis of SEO and the evaluation of websites.

Ahrefs is a leading tool for backlink analysis and SEO auditing. It provides extensive information on backlinks, organic search, keywords, and other factors influencing SEO. Ahrefs is also instrumental in determining Domain Rating (DR) and URL Rating (UR), which are key indicators of a domain's and individual web pages' authority [19].

Majestic specializes in backlink analysis and boasts one of the largest link databases in the world. This tool provides detailed insights into metrics such as Trust Flow (TF) and Citation Flow (CF). Trust Flow evaluates the trustworthiness of a website based on the quality of links pointing to it, while Citation Flow assesses a website's influence based on the number of links pointing to it, regardless of their quality [20].

By using these tools, a comprehensive understanding of a website's SEO standpoint can be obtained. This includes an evaluation of the quality of backlinks, domain authority, and page influence, as well as an understanding of the impact of organic traffic and keywords on the overall visibility of the website in search engines.
1.3. Link Analysis Algorithm for Evaluating Donor

After these indicators are collected, their analysis takes place. Fig. 2 shows a flowchart of the link analysis algorithm for evaluating donor sites. The algorithm starts with a list of donors and uses a series of metrics to classify domains as "Bad Donor", "Good Donor", or "Potential Spam". Key steps of this algorithm include:

- Checking Domain Rating. If DR is less than or equal to 10, the domain is classified as a "Bad Donor".
- For domains with DR greater than 10, the number of Referring Domains is analyzed. If RD is less than or equal to 50, the domain is also classified as a "Bad Donor".
- Organic Traffic is then assessed. If OT is less than or equal to 500, the domain is classified as a "Bad Donor".
- If OT is more than 500, it is checked whether Referring Domains are greater than Organic Traffic. If so, the domain may be classified as "Potential Spam".
- If RD is less than or equal to OT, the analysis continues with Organic Keywords and URL Rating.
- If OK is less than or equal to 5, Trust Flow and Citation Flow are checked.
- If TF is greater than or equal to 10, and CF is less than or equal to TF, the domain is classified as a "Good Donor".
- If UR equals 0, the domain is classified as "Potential Spam".
- If the OK is more than 5, the domain is considered a "Good Donor".

The proposed flowchart demonstrates a systematic approach to evaluating websites to determine their value as link donors.

Only those donors that receive the status of "Good Donor" at the output of the algorithm are economically feasible.

Additionally, the system has sufficient functionality for reporting, namely:

- Domain reporting: Each domain can have many different link donors, each with its own price. The system allows for exporting a report by domain to understand when and how money was spent on it.
- User reporting: Each user will have a fee for placing a link on the resource.

Managers will be able to receive reports on users for a month or a year.

- Project reporting: Links can be divided by projects. There is functionality for obtaining reports on the use of financial resources on links related to a given project.
- Obtaining reports on the use of financial resources on links related to a given project.

![Figure 2: Algorithm for analyzing page statistical indicators](image-url)
2. Protection of Information in the System

Given the functionality and innovation of the proposed system, as well as the near absence of similar comprehensive solutions in one package, it is potentially at risk of unauthorized attempts at appropriation. Therefore, security was one of the key factors in choosing the development tools.

Information protection in web systems is a critical aspect, particularly in an era where data breaches and cyber attacks are increasingly common. Effective security measures are essential to safeguard sensitive data, maintain user trust, and comply with legal standards. A robust web system security strategy encompasses several layers, including secure coding practices, data encryption, authentication and authorization controls, regular security audits, and adherence to best practices in cybersecurity.

Firstly, secure coding practices are fundamental in preventing common vulnerabilities such as SQL injection, Cross-Site Scripting (XSS), and Cross-Site Request Forgery (CSRF). Developers must be vigilant in sanitizing user inputs, validating data, and using secure APIs [21]. Additionally, implementing data encryption both in transit (using SSL/TLS protocols) and at rest ensures that sensitive information, such as personal user data and payment details, is protected from unauthorized access. Encryption acts as a critical barrier, making it extremely challenging for attackers to exploit intercepted data [22].

Furthermore, robust authentication and authorization mechanisms are crucial [23]. Multi-factor Authentication (MFA) adds a layer of security beyond traditional password-based systems. Implementing Role-Based Access Control (RBAC) ensures that users can access only the information and functionalities that are necessary for their role, minimizing the risk of internal data breaches. Regular security audits and penetration testing can identify and address potential vulnerabilities in the system while staying updated with the latest cybersecurity trends and threats ensuring that the system remains resilient against new types of attacks.

With this in mind, the PHP framework Laravel was selected for the development of the system.

Laravel includes many built-in features for ensuring security and data protection. Here are some key security aspects in Laravel:

- Protection against SQL Injections: Laravel uses Eloquent Object-Relational Mapping (ORM) and PHP Data Objects (PDO), which automatically sanitize query parameters, reducing the risk of SQL injections.
- Protection against Cross-Site Scripting (XSS): Laravel's Blade templating engine automatically escapes output data, preventing the insertion of malicious scripts.
- Protection against Cross-Site Request Forgery (CSRF): Laravel has built-in CSRF protection, namely CSRF tokens, which ensure requests are sent from a trusted form.
- Hashing and Encryption of Passwords: Laravel uses Bcrypt and Argon2 algorithms for password hashing, and provides tools for secure data encryption and decryption.
- Middleware for Security: Laravel allows the use of middleware classes to implement various levels of access and authorization, controlling access to different parts of the application.
- Rate limiting: Laravel provides capabilities for limiting the frequency of requests, which can help prevent DoS attacks.
- Content Security Policy (CSP): The framework offers the ability to integrate a content security policy to prevent the use of malicious resources.
- Logging and Auditing: Laravel provides extensive logging capabilities, allowing for tracking and analysis of user and system actions.
- Updates and Support: Laravel is actively maintained and regularly updated, including security enhancements.
- Data Validation: The framework offers powerful tools for validating input data, and reducing risks associated with incorrect data.

The Laravel framework includes special classes that provide an organized way to manage authorization, especially when it comes to models or resources—Policies. These classes allow you to easily define what actions
a user can perform on a specific resource. Policies define methods that correspond to certain actions (e.g., view, create, update, delete). Each Policy is usually associated with a specific Eloquent model. Laravel automatically detects the Policy that corresponds to a certain model based on standard naming conventions. Policies can be easily integrated with Laravel Middleware, allowing for control of access to different parts of the application. Policies can be used in conjunction with Gates, another Laravel mechanism for access control [24, 25].

2.1. User Roles in the System

Additionally, the Spatie/Laravel-Permission library was used for the development of the user roles system. It is one of the most popular third-party libraries for managing roles and permissions in Laravel. Key features of the Spatie/Laravel-Permission library include:

- The ability to create and manage roles and permissions in the database.
- Roles and permissions can be easily assigned to users, allowing for flexible access management.
- Support for Middleware classes to protect routes and controllers based on permissions and roles.
- Integration with the Blade templating engine, controlling UI elements display based on user permissions.
- Efficient caching to optimize database queries.
- Customizability to tailor to the specific needs of the application.

The proposed system has three user roles:

- Administrator: Has unlimited access to all system resources.
- Manager: Has access to all system resources except managing access and roles. Can create, edit, and delete users, as well as determine their salary for placing links on the resource. Managers have access to all types of reporting.
- Outreach: Only has access to links and projects. Can add and edit them, but not delete them. They only see links assigned to them or unassigned. They have a limited number of domains to work with, by default—50. Can only view reporting related to their pay.

Using Policies in combination with the Spatie/Laravel-Permission library provides significant flexibility and control over authorization in the application. This allows for clearly defined authorization logic that is easily maintainable and scalable, enhancing the security and scalability of the proposed system.

Administrators can create new methods for accessing different components of the system directly from the admin panel.

Furthermore, the system does not have functionality for unauthorized users. If a user navigates to the site via a link, they are redirected to the login page. Without authorization, they will not have access to the system’s resources.

2.2. System Interface Overview

When working with large data volumes, the question of optimizing system performance becomes especially important. In modern web application development, various JavaScript libraries like Vue.js and React.js are widely used, which significantly facilitate the creation of interactive web interfaces. However, in the context of working with large data, using these libraries can lead to additional loads on the client side of the web application, as they load the page and require additional resources for processing JavaScript code [26].

Considering this, a decision was made to avoid using large JavaScript libraries in the development of the client side of the web application. Instead, the web application is designed to maximize the use of built-in HTML and CSS capabilities, which provide a sufficient level of interactivity without heavy loading on the browser. This approach reduces page load times, as the absence of large JavaScript frameworks reduces the overall load on the server [27].

Moreover, this approach improves the interface’s responsiveness to user actions, as browsers do not need to perform complex JavaScript operations, which can be critical for web applications processing large data volumes. It also ensures better compatibility with various browsers and devices, as standard HTML and CSS technologies have broad support without needing to account for the specifics of particular JavaScript libraries.
Forgoing the use of Vue.js, React.js, or similar frameworks requires more work on interface optimization and may limit some interactive capabilities. However, in the context of working with large data, this is a justified compromise. This approach ensures faster user response, better overall performance, and reduced load on the client side, which is critically important for systems processing large volumes of data. The interface of the page with the list of domains of the proposed system is presented in Fig. 3.

The domain list contains essential information about the links, specifically: the domain; the user it is assigned to; the projects of this domain; results of the first stage of checking for each parameter; expenses related to the domain, and other additional information, as well as buttons for actions with the domain (edit, view, scan, and delete).

Above the table with the list of domains, there are buttons for bulk addition of domains using an Excel spreadsheet or a standard list of domains in the specified format, as well as a button for scanning all domains. The form for importing domains varies depending on the user’s role. Both the administrator and manager have the functionality to assign a link to an outricher and to disable the check for domain duplicates (Fig. 4). Instead of checking for duplicates, domains that are already in the system will be added to the projects selected in the form. The outricher does not have access to this functionality. Domains added by them are automatically assigned to that user. The form for importing domains by an outricher is shown in Fig. 5.

Also, the interface shown in Fig. 3 has a search field for the domain and filters for the results of the first stage of domain verification. For convenience, the filter block can be collapsed or expanded. Domain filtering occurs asynchronously, i.e., without reloading the page. This significantly enhances the ease of working with the system.

2.3. Important Features of the System

Each domain can be added to several projects. This is because a single website can host several different articles on various topics with links to promoted resources. For each project, specific parameters such as Donor, Acceptor, Anchor, the status of work with the domain, and the placement price are filled out separately.

In the context of SEO, the term “Donor” refers to the web address from which an external link leads to another site—the “Acceptor.” The Anchor, or anchor text, refers to the visible, clickable text in a hyperlink leading to the Acceptor. This text is usually highlighted in color (often blue) and underlined to distinguish it from regular text on the page. Anchor text helps search engines understand the context and content of the linked page. If the anchor text is closely related to the content on the target page, it improves
SEO, as it indicates higher relevance and quality of information.

The status of work with the domain involves setting one of the stages of working with this domain for a given project by the person responsible for the domain. There are several main statuses, namely: "new," "in progress," "negotiation," "recommended for placement," "refusal," "error," and "closed deal." The status "closed deal" can only be set by a manager or administrator, as it is not available to other users. This is because only users with these roles make the final decision on placing links on the resource. After the user working with the donor has set the status "recommended for placement", the manager receives a notification about this and decides to place or not to place the link on this resource.

After placing the link on the resource and setting the corresponding status in this project, the user responsible for the donor is credited with a salary, and the costs for placement and salary are recorded in the overall expenses for the domain.

All links that have received the status "closed deal" undergo periodic automatic checks by the system. The link on which the link is placed, i.e., the Donor, undergoes the first stage of verification, as well as a check for the presence of the Acceptor link. If the link fails any of the checks, the manager receives a notification, and the Donor gets the status "error". To implement such functionality, CRON was used.

CRON is a program for UNIX-like operating systems, used for the scheduled execution of tasks (commands or scripts) at a specific time. It operates as a daemon, continuously executing tasks according to instructions specified in the crontab file. This file contains a schedule of tasks specified through specific time intervals. The format of the schedule in crontab includes five fields representing minutes, hours, days of the month, months, and days of the week, allowing flexibility in planning. Each task in the crontab file usually points to a specific script or command to be executed.

In the Laravel framework, a concept based on CRON is used for managing scheduled tasks. Laravel introduces a wrapper over CRON called "task scheduling". It allows developers to easily define the execution schedule of various tasks within the application's code. Using this mechanism, tasks such as cache clearing, sending emails, and gathering statistics can be easily set up. Laravel requires only one CRON entry on the server, which calls the artisan schedule: run command every minute. This command checks if there are tasks scheduled for execution and launches them according to the defined schedule. This approach greatly simplifies the management of scheduled tasks in the Laravel environment, making the process more intuitive and integrated with the rest of the framework's functionality.

All domains in work undergo the first stage of verification weekly. This is necessary to update information about the activity of the domain.

Certain projects are targeted at specific countries or regions, so when choosing domains that can potentially be donors for links, websites that only operate in their country are sometimes selected. Also, certain sites may be blocked in the territory of a country important for the planned link placement. Therefore, the system contains functionality for selecting a proxy server when performing an analysis.

A proxy server acts as an intermediary node between the user and target internet resources, serving as a mediator in transmitting requests. The use of a proxy involves a special mechanism: user requests are first directed to the proxy server, which then forwards them to the specified internet resource. Proxy servers are also used to bypass geographic restrictions and optimize performance through caching.

With the Laravel framework and the HTTP client Guzzle, integration with proxy servers becomes particularly significant. Guzzle, as a tool for creating HTTP requests in PHP, provides extensive capabilities for setting up a proxy. This is achieved through the 'proxy' parameter in Guzzle client configuration settings, where the URL of the proxy server or a set of configurations for different protocols can be specified. This approach effectively manages the routing of requests through a proxy.

### 3. Discussion

In this work, a cutting-edge system for optimizing internet marketing is presented, particularly in the context of SEO and the management of external links. The proposed
system is distinguished by its high level of innovation, providing users with the ability to effectively analyze and select the most optimal platforms for link placement based on a variety of SEO parameters.

One of the key advantages of this system is its high level of security. Thanks to integration with the PHP framework Laravel, the system provides robust protection against threats such as SQL injections, XSS, CSRF, and others, which is critically important for ensuring the integrity and security of user data. This integration not only guarantees security but also ensures a high level of flexibility and adaptability of the system.

The system also stands out for its ability to precisely analyze and classify potential donor sites, categorizing them as "Good Donor," "Bad Donor," and "Potential Spam." This allows users to avoid unreliable or ineffective resources, focusing instead on those that provide the maximum return on SEO campaigns.

Compared to existing tools, such as LinkChecker Pro, this system offers a significantly wider range of functionalities, making it more effective in solving complex tasks related to internet marketing. Considering its high level of security, comprehensive analytical approach, and flexibility of use, this system has the potential to become an indispensable tool for companies seeking to efficiently manage their online marketing strategies.

References


[9] Link Checker Pro. URL: https://linkchecker.pro


[14] Organic Keywords: What Are They & How to Find Them. URL:
https://www.semrush.com/blog/organic-keywords/

[15] Organic Traffic: Why it is Important, how to Generate it and Reasons Behind a Drop. URL: https://baresquare.com/blog/organic-traffic-why-it-is-important-how-to-generate-it-and-reasons-behind-a-drop

[16] What is URL Rating (UR)? URL: https://ahrefs.com/seo/glossary/url-rating


[18] What Is Citation Flow (CF) in SEO? URL: https://www.pageonepower.com/search-glossary/citation-flow


