UNSAFE URL DETECTION FOR SECURE WEB SEARCH

Gayathri G[1] Kokila V[2] Swetha B[3] Mrs. Vennila P[4]

[1]UG Student, CSE Dept, E.G.S Pillay Engineering College, Nagapattinam, India.
[2]UG Student, CSE Dept, E.G.S Pillay Engineering College, Nagapattinam, India.
[3]UG Student, CSE Dept, E.G.S Pillay Engineering College, Nagapattinam, India.
[4]Assistant Professor, CSE Dept, E.G.S Pillay Engineering College, Nagapattinam, India

[1]*Abstract*— **Malicious URL (or) malicious website is a common and serious threat to cyber security. Naturally, search engine becomes the backbone of information management. Nevertheless, the flooding of large number of malicious websites on search engine has posed tremendous threat to our users. Most of exiting systems to detect malicious websites focus on specific attack. At the same time, available browser extensions based on blacklist are powerless to countless websites. Therefore, it is essential that any data leaving the client side should be effectively masked such that the server cannot interpret any valuable information from the masked data. Here propose the first PPSB service. It provides strong security guarantees that are missing in existing SB services. In particular, it inherits the capability of detecting unsafe URLs, while at the same time protects both the user's privacy (search history) and blacklist provider's proprietary assets (the list of malicious URLs). In this work, proposed a model which encrypts the users' sensitive data to prevent privacy from both outside analysts and service provider. Also, completely supports selective aggregate functions for online user behavior analysis and guaranteeing differential privacy. Here AES (Advanced Encryption Standard) algorithm is used for encrypting users' online search behavior data. Authenticated users can only decrypt the history with the help of secret key sharing and verification process. User feedback system helps to intimate the any malicious URL to the administrator of the search engine. Administrator can update the malicious URL database frequently.**

***Index Terms*— Malicious URL Detection, Blacklist Creation, History encryption using AES, URL Recommendation, Key verification, History Access**.

I. INTRODUCTION

NETWORK SECURITY

Managing security is the process of understanding the risks and identifying how plenty danger is appropriate. Different stages of security are suitable for specific organizations. No network is fully secure, so don't intention for that level of protection. If you attempt to stay up-to-date on each new danger and every virus, you'll soon be anxiety and strain. Look for the main vulnerabilities that you can deal with together with your present assets.

Here present more benefits of computer networks and the Internet. Connecting your network to the Internet presents access to a full-size amount of facts and lets in you to share information on an first rate scale. However, the communal nature of the Internet, which creates so many advantages, also offers malicious users access to several targets. The Internet is only as secure because the networks it connects, so we all have a responsibility to ensure the safety of our networks.

Information safety is the process of securing statistics records from unauthorized access, use, change, tempering, or disclosure. With the extended use of electronic media in our personal lives as well as corporations, the opportunity of security breach and its main effect has elevated. The robbery of personal identity, credit card data, and other critical facts the use of hacked user names and passwords have grow to be common place these days. In addition, the theft of exclusive business statistics might also lead to loss of enterprise for industrial groups.

1.1 WEB SECURITY

The Internet is a risky location, with extremely good regularity, customers listen about websites becoming unavailable due to denial of provider attacks, or showing modified (and regularly unfavorable) information on their homepages. In different excessiveprofile cases, hundreds of thousands of passwords, electronic mail addresses, and credit score card details were leaked into the public area, exposing website customers to both private embarrassment and financial danger.

The purpose of internet site safety is to prevent those (or any) kinds of attacks. The more formal definition of website safety is the act/practice of protective websites from unauthorized access, use, modification, destruction, or disruption.

Effective internet site security requires layout attempt throughout the entire of the website: to your net utility, the configuration of the web server, your rules for growing and renewing passwords, and the patron-facet code. While all that sounds very ominous, the coolest information is that in case you're the usage of a server-aspect net framework, it will almost actually enable "with the aid of default" robust and properly-thought-out protection mechanisms in opposition to some of the extra not unusual assaults. Other attacks can be mitigated thru your net server configuration, for instance with the aid of permitting HTTPS. Finally, there is publically available vulnerability scanners gear that let you find out in case you've made any obvious mistakes.

1.2 COMMON ATTACKS/ VULNERABILITIES

A. Click jacking

In this attack, a malicious person hijacks clicks meant for a seen pinnacle-degree website online and routes them to a hidden web page below. This approach might be used, as an example, to show a valid financial institution web page however capture the login credentials into an invisible <iframe> managed via the attacker. Click jacking may also be used to get the person to click a button on a visible website online, but in doing so truely unwittingly click on a totally exclusive button. As a protection, your website can save you itself from being embedded in an iframe in another website online through placing the precise HTTP headers.

B. Denial of Service (DoS)

DoS is typically achieved by way of flooding a goal site with faux requests so that access to a site is disrupted for valid customers. The requests might also genuinely be numerous, or they'll individually consume large quantities of aid (e.g., gradual reads or uploading of large files). DoS defences generally paintings via figuring out and blocking off "terrible" visitors at the same time as allowing legitimate messages through. These defences are commonly located earlier than or within the net server (they're not part of the internet utility itself).

C. Directory Traversal (File and disclosure)

In this attack, a malicious person attempts to access elements of the web server files that they ought to not be capable of access. This vulnerability takes place while the person is capable of pass filenames that encompass record gadget navigation characters (for example,/../). The solution is to sanitize enter earlier than using it.

D. File Inclusion

In this type of attack, a consumer is able to specify an "accidental" report for show or execution in records handed to the server. When loaded, this report might be finished at the web server or the customer-side (leading to an XSS attack). The solution is to sanitize enter before the use of it.

E. Command Injection

Command injection attack permits a malicious person to execute arbitrary machine instructions on the host working system. The solution is to sanitize consumer enter earlier than it is probably used in device calls.

1.3 SAFE BROWSING

Malicious SB service issuer wants to recognize whether a person is journeying a specific web web page, e.g., some political information. One way to gain that is that the web browser sends all the visited URLs to a far off server, either inside the plaintext, hash cost or encrypted layout. However, this behavior may be detected by tracking and analyzing the browser, e.g., the usage of the taint evaluation technique. Specifically, as a way to track a particular URL the SB carrier issuer can insert the 32-bit hash prefixes of all its decompositions, e.g., c01e362f, after which push this newly up to date prefix filter out to the customers. Later, once a user visits the internet page (or comparable URLs that percentage a few decompositions), the matched hash prefixes might be sent to the far flung SB server. Based on the prior information of the prefix filter (i.e., the mappings among the hash prefixes and their corresponding URLs), the server can infer the URL (or area) navigated by means of the user. It gives strong protection ensures that are lacking in present SB services. In precise, it inherits the capability of detecting dangerous URLs, at the same time as on the identical time protects both the person's privacy (surfing records) and blacklist provider's proprietary belongings (the list of risky URLs). This approach has a few disadvantages along with; developing metadata of URLs fails while the server gets multiple prefixes for a URL and there may be a threat that other URLs may additionally have the equal hash prefixes this makes collision among URLs.

A malicious user would possibly leverage PPSB to degrade the consumer-facet consumer experience, like putting a number of faux or secure URLs or increasing the server-aspect delay. To cope with this capability difficulty, PPSB presents a flexible mechanism for customers to add or eliminate blacklist providers. Admin ought to add the fake URL and keyword to this blacklist storage. User can also allowed suggesting the malicious internet site info concerning black list. In this gadget malware detection machine makes use of a supervised machine gaining knowledge of technique for discovering malwares. The SVM classification with malware detection system extends the idea of signature primarily based detection system with a aggregate of conduct tracking approach. It utilizes static and dynamic evaluation of malwares with the aid of taking the run time traces of the executable. Image based malicious detection also provide to compare the image functions based totally on original internet site and malicious website. This version also affords seek records security which encrypts the users' sensitive statistics to save you privateness from both outside analysts and the aggregation provider. Also, completely helps selective combination functions for on-line consumer conduct analysis and ensuring differential privateness.

II. RELATED WORK

In [1], Mourtaji, et.al implemented a phishing detection mode that incorporates 37 features extracted from six different methods including the black listed method, lexical and host method, content method, identity method, identity similarity method, visual similarity method, and behavioral method. Furthermore, comparative analysis was undertaken between different machine learning and deep learning models which includes CART (decision trees), SVM (support vector machines), or KNN (K-nearest neighbors) and deep learning models such as MLP (multilayer perceptron) and CNN (convolutional neural networks).

Ahammad, et.al [2] presented a solution for detecting such websites with the help of machine learning algorithms focused on the behaviors and qualities of the suggested URL. The web security community has created blacklisting services to identify malicious websites. A variety of methods, such as manual reporting, and site analysis heuristics are used to create these blacklists. Due to their recentness, lack of evaluation, or incorrect evaluation, many malicious websites inadvertently escape blacklisting. To create a machine learning model for detecting whether a URL is malicious or not, algorithms such as Random Forests, Decision Trees, Light GBM, Logistic Regression, and Support Vector Machine (SVM) are used.

Yuan, et.al [3] proposed a parallel neural joint model algorithm for analysis and detection of malicious Uniform Resource Locator (URL). By detecting and analyzing malicious URL's characteristics, the semantic and visual information will be extracted. First, a visualization algorithm is used to realize the visualization of the URL mapping to a gray image with texture characteristics. Second, the lexical feature and character feature of URL are extracted and further processed through word vector technology. These extracted features are transformed into lexical embedding vectors and character embedding vectors.

In [4], Xu, Pingfan, et.al introduced a transformer-based malicious URL detection model, which has significant accuracy and outperforms current detection methods. Proposed approach to the transformer model design is not identical to the standard structure. The model design of Rudd and Abdallah from FireEye Inc. inspires the transformer model design of our approach [25]. In our solution, the transformer model is very similar to the design of OpenAI's GPT model, one of the famous variants of the Transformer model. Our Transformer model applied left to right (L-R) modeling and only contained the encoder part from the standard transformer model.

Butnaru, et.al [5] here proposed and evaluate a phishing detection engine, which uses supervised machine learning in order to detect phishing attacks based on a novel combination features that are extracted from the URL. This allows us to avoid any delays which stem from the computation of features that need access to third-party resources, such as access to WHOIS records. In summary, our work makes the following contributions: Train, optimise and evaluate a phishing detection engine which relies on supervised machine learning, based on features that stem from the URL. Our feature selection process includes features that have been proven suitable by the literature, coupled with new ones that we propose and evaluate. To the best of our knowledge, we are the first to use the Levenshtein distance as a similarity index feature for training a range of machine learning algorithms in this domain. And also, revisit the use of suggestive vocabulary, as one of our features. Evaluate the performance of our phishing detection engine over time by classifying active phishing attacks that were reported on PhishTank, without model retraining. Find that the performance of the classification is not affected by time, as well as it significantly outperforms the protection that is offered by GSB.

Odeh, et.al [6] presented the state of art techniques for detection of phishing websites using the ML techniques. This research identifies solutions to the website's phishing problem based on the ML techniques. The majority of the examined approaches are focused on traditional ML techniques. Random Forests (RFs), Support Vector Machines (SVMs), Naïve Bayes (NB), and Ada Boosting are the powerful ML models examined in the literature. This survey paper also identifies deep learning-based techniques to demonstrate better performance for detecting phishing websites compared to the conventional ML techniques.

[7] In this paper Butt, et.al implemented a deep learning technique on the benchmark datasets to identify the pattern of phishing URLs. We used gradient boosted decision trees algorithm to train our model and applied the regular deeply connected neural network layers in various sequences and Adam optimizer. The most found patterns will help the system to detect phishing URLs and avoid phishing. Proposed work consists of mainly two phases; the first phase has two substages: the URL search phase and the other is the feature extraction phase to identify the relevant features commonly found in phishing URLs using xgboost. After that, pass data to the pre-trained model to deeply connected neural network layers and optimizer. It can help us to differentiate phishing websites from real websites with the help of data mining algorithms. The primary purpose of this method is to classify a URL that is given as an input or not as a phished URL.

Tang, et.al [8] presented a state-of-the-art survey on methods for phishing website detection. It starts with the life cycle of phishing, introduces common anti-phishing methods, mainly focuses on the method of identifying phishing links, and has an in-depth understanding of machine learning-based solutions, including data collection, feature extraction, modeling, and evaluation performance. Since one user reported and verified the website as a phishing website, the URL will be added to the blacklists, which could be used to prevent other users from being disrupted. Heuristic strategies identify a phishing web page depending on a group of features extracted from the textual contents of the website and compare the features with the legitimate one. The idea of the approach is that the attackers usually deceive users by imitating well-known websites. The machine learning methods also depend on the features from the website, build the model to learn from a batch of data with structured features, and then predict if the new website is a phishing website. In the machine learning area, detecting phishing websites is a classification problem.

Purbay, et.al [9] presented a method to detect phishing URLs by examining different components of URLs using machine learning and deep learning techniques. We have discussed different supervised learning methods used for phishing URL detection based on lexical feature, WHOIS properties, PageRank, Traffic Rank details and page importance properties. A URL is a website address that represents the location of a website on a network and the means of gaining access to it. By accessing the URL, we connect to the database on the server, which stores all the details relating to the website, and it contains a webpage that displays them. URLs are divided into two categories: malicious and benign. Malicious URLs are used in URL phishing, while benign URLs are harmless and secured. A cybercriminal will create a site that looks like the real one, and all of its information will be identical to that of the absolute URL. The URL will appear as an advertisement on other websites, and the fraud will happen when the user enters their credentials. And another way is by sending the malicious URL to the user through email, and when the user tries to open the URL some nasty virus will be downloaded, this allows the cybercriminals to access the information to commit their crimes. Malicious and benign URLs look similar, so to distinguish them we need to extract some features from them. Detecting malicious URLs requires extracting some of their features from them, then comparing these features to determine whether the URL is malicious or benign.

Wazirali, et.al [10] proposed an efficient URLs Phishing detection technique. Our technique depends on Software Defined Network (SDN) technology, clustering and feature method, and Conventional Neural Network (CNN) algorithm. Feature selection technique is based on Recursive Feature Elimination (RFE) with Support Vector Machine (SVM) algorithm. The SDN is used to transfer the URLs phishing detection process out of the user's hardware to the controller layer, continuously train on new data, and then send its outcomes to the SDN-Switches. RFE-SVM and CNN are used to increase accuracy of phishing detection. Therefore, the proposal model does not require retrieving the content of the target website or using any third-party services. It captures the information and sequential patterns of URL strings without requiring a prior knowledge about phishing, and then uses the sequential pattern features to quickly classify the actual URL. .

III. EXISTING METHODOLOGIES

Phishing is the fraudulent activity to get sensitive records inclusive of usernames, passwords and credit score card info, frequently for malicious motives, with the aid of disguising as a sincere entity in an digital conversation. Phishing attack can be carried out in various form like Email phishing, Website phishing, spear phishing, Whaling, Tab napping, Evil dual phishing and many others. To avoid this phishing attack various anti-phishing applications have to be use. There are diverse anti phishing solutions inclusive of Blacklist, heuristic, visible similarity, machine learning techniques and many others. This is maximum usually used method in which list of phishing URL is stored in database after which if URL is found in database, it's miles referred to as phishing U and offers warning otherwise it's far referred to as legitimate. This technique is simple and quicker to put in force as it see URL is in db or not. But limitations are small trade in URL is sufficient to skip the list based totally technique and frequent replace of listing is important to counter new attack.

Phishing imitates the traits and functions of emails and makes it appearance similar to the original one. It appears much like that of the legitimate supply. The user thinks that this electronic mail has come from a genuine business enterprise or an business enterprise. This makes the consumer to forcefully go to the phishing website via the links given inside the phishing email. These phishing websites are made to mock the arrival of an unique organization internet site. The phishers force consumer to top off the personal facts through giving alarming messages or validate account messages and so forth so they replenish the desired information which can be utilized by them to misuse it. They make the situation such that the person isn't left with another option but to go to their spoofed website.

In the training phase, use the categorized facts wherein there are samples consisting of phish region and valid area. If we try this then category will now not be a hassle for detecting the phishing area. To do an operating detection version it's miles very critical to use records set inside the schooling section. We need to use samples whose lessons are regarded to us, which means that the samples that we label as phishing ought to be detected best as phishing. Similarly the samples which are classified as legitimate might be detected as legitimate URL. The dataset to be used for machine learning without a doubt consist those features. There so many machine learning algorithms and every set of rules has its very own operating mechanism which we have already seen in the previous work. The existing device uses any one of the best machine learning algorithms for the detection of phishing URL and predicts its accuracy. The present device has good accuracy but its miles still not the satisfactory as phishing assault is a totally crucial; we ought to find a fine approach to put off this. In the currently existing device, most effective one device mastering algorithm is used to be expecting the accuracy; the usage of most effective one algorithm isn't a terrific method to improve the prediction accuracy.

IV. MALICIOUS WEBSITE DETECTION WITH HISTORAY ENCRYPTION

A malicious user might leverage PPSB to degrade the client-side consumer experience, like putting a number of fake or safe URLs or growing the server-side postpone. To deal with this ability problem, PPSB gives a reliable mechanism for users to feature or put off blacklist providers. Admin should add the unsafe URL and keyword to this blacklist storage system. User also can allowed suggesting the malicious internet site info regarding black listing. In this proposed application malware detection system makes use of a supervised learning method for discovering malwares. The SVM based malware detection device extends the concept of signature based detection with a mixture of behavior tracking approach. It makes use of static and dynamic evaluation of malwares by way of taking the run time traces of the executable. Keyword based malicious detection also provide to compare the keyword capabilities based on unique website and malicious website. This version also gives seek data protection which encrypts the customers' access history data to prevent privateness from both outside analysts and the aggregation provider issuer. Also, completely helps selective combination functions for on line user conduct evaluation and making sure differential privacy.

In this system malware detection approach uses a supervised machine learning approach for discovering malwares. The SVM based malware detection system extends the concept of signature based detection system with a mixture of behavior tracking approach. Also, completely supports selective combination capabilities for on line user behavior evaluation and making certain differential privacy. A malicious user would possibly leverage PPSB (Privacy Preserving Safe Browsing) to degrade the consumer-aspect user revel in, like inserting a number of fake or secure URLs or increasing the server-side delay. To deal with this potential problem, PPSB affords a flexible mechanism for users to add or eliminate blacklist providers. Admin could add the faux URL and key-word to this blacklist storage. User also can allowed suggesting the malicious internet site information regarding black list.

The cause of internet site security is to prevent these (or any) sorts of attacks. Effective website protection calls for layout attempt throughout the entire of the website: in net application, the configuration of the internet server, user policies for creating and renewing passwords, and the patron-aspect code. The proposed assignment detects Malicious or Fake URLs to prevent the customers accessing from Unsafe URLs. Also offer comfortable encryption method to encrypt the consumer search records earlier than saved at the server.

METHODOLOGY

A. AES (Advanced Encryption Standard)

The Advanced Encryption Standard (AES) algorithm is a widely used encryption algorithm for securing data. The AES algorithm is a symmetric encryption algorithm, which means that the same key is used for both encryption and decryption. The key size can be 128 bits, 192 bits, or 256 bits, depending on the level of security required.

The AES encryption process involves the following steps:

Key Expansion: The AES algorithm expands the original key to create a set of round keys, which will be used in the encryption process.

Initial Round: In the first round, the AES algorithm performs a bitwise XOR operation between the input data and the first round key.

Rounds: The AES algorithm performs a series of rounds, each consisting of four operations: SubBytes, ShiftRows, MixColumns, and AddRoundKey.

SubBytes: In this step, the AES algorithm substitutes each byte of the input data with a corresponding byte from a pre-defined S-box.

ShiftRows: In this step, the AES algorithm shifts the rows of the input data matrix to the left by a certain number of bytes.

MixColumns: In this step, the AES algorithm performs a matrix multiplication on the columns of the input data matrix.

AddRoundKey: In this step, the AES algorithm performs a bitwise XOR operation between the output of the MixColumns step and the current round key.

Final Round: The final round of the AES algorithm is similar to the earlier rounds, except that it does not include the MixColumns step.

Output: The final output of the AES encryption process is the encrypted data.

To decrypt the data, the AES algorithm performs the same steps in reverse order, using the same key.

B. *Support Vector Machine (SVM)*

Support Vector Machine (SVM) is a supervised algorithm based on machine learning which can be used for both classification and regression problems. However, it's far ordinarily used in classification work. In this work, plot each data item as a point in n-dimensional space with the value of every feature being the count of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well. Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is best for segregates the two classes (hyper-plane/ line). The hyperplane is the line with the biggest margin to both groups.

Support Vector Machines has higher effectiveness in higher dimensional spaces. It is even very effective on data sets where number of dimensions is greater than the number of samples. This is mainly because of the kernel trick, which we talk about it later. Further advantages of Support Vector Machines are the memory efficiency, speed and general accuracy in comparison to other classification methods like k-nearest neighbor or deep neural networks.

Step1: Malicious URLs and keywords have been collected and stored on blacklist storage.

Step 2: For the collection of Malicious URLs number of features could be used like URL length, the number of dots, ip Address, SSL connection, at symbol(@) and dash symbol(-).

Step 3: The selected features identified from URL then stored on blacklist.

Step 4: User could enter the URL or Keyword for searching details.

Step 5: Input data classified with trained dataset with the help of SVM classifier.

Step 6: SVM classifier returns either an URL is phishing or non-phishing.

PROCEDURE

A. FRAMEWORK CONSTRUCTION

The detection of malicious URLs limits web-based attacks by preventing web users from visiting malicious URLs and warning web users prior to accessing content located at a malicious URL. Thus, malicious URL detection protects computing system hardware/software from computer viruses, prevents execution of malicious or unwanted software, and helps avoid accessing malicious URLs web users do not want to visit.

This proposed framework uses SVM classification models to detect a malicious URL and categorize the malicious URL as one of a phishing URL. The blacklist storage models by using a set of training data (unsafe URLs and keywords) and machine learning algorithms. The training data includes a known set of unsafe URLs and a known set of malicious keywords. This framework also supports URL encryption process, to avoid the unauthorized prediction of URL details.

B. USER REGISTRATION AND LOGIN

Users have to register with their name, password and Email id. These details will be saved in the database. The user have to login with the name and password. The entered data will be compared with the available data. If match found, the user can proceed. If no match found, the user have to re-enter the details again.

UNSAFE URL DETECTION

The verification of URLs and keywords is very essential in order to ensure that user should be prevented from visiting malicious websites. SVM mechanisms have been proposed to detect the malicious URLs. One of the basic features that a mechanism should possess is to allow the fake URLs that are requested by the client and prevent the malicious URLs before reaching the user. This is achieved by notifying the user that it was a malicious website. The techniques extract features associated with the known URLs, and use the machine learning algorithms to train the classification models to detect and categorize an unknown malicious URL. A database updation is performed every time the systems come across a new URL. Here, the new URL will be matched and tested with every previously known malicious URL in the black list. The update has to be made in black list whenever system comes across a new malicious URL. This also allows users to provide suggestions to add malicious URLs.

C. SEARCH URL ENCRYPTION

Once the login procedure is succeeded, the user can search details using URLs and keywords. The user will enter a URL or keyword in the search box and click the submit button. When the user clicks the search button, the request was processed and related details are shown to the user. Then the searched keyword and URL will be encrypted and stored in the intermediate. The users' search data will be encrypted using AES encryption algorithm.

D. SEARCH URL ENCRYPTION

Once the login procedure is succeeded, the user can search details using URLs and keywords. The user will enter a URL or keyword in the search box and click the submit button. When the user clicks the search button, the request was processed and related details are shown to the user. Then the searched keyword and URL will be encrypted and stored in the intermediate. The users' search data will be encrypted using AES encryption algorithm.

E. MALICIOUS URL SUGGESTION

In proposed work, the URL suggestion process could be implementing to enhance the performance of blacklist storage. When user finds any malicious URL during searching process, they will allow to suggestion process. Here user should send URL details to admin, to add blacklist storage. This frequent update in blacklist improves the performance of unsafe (or) malicious URL detection.

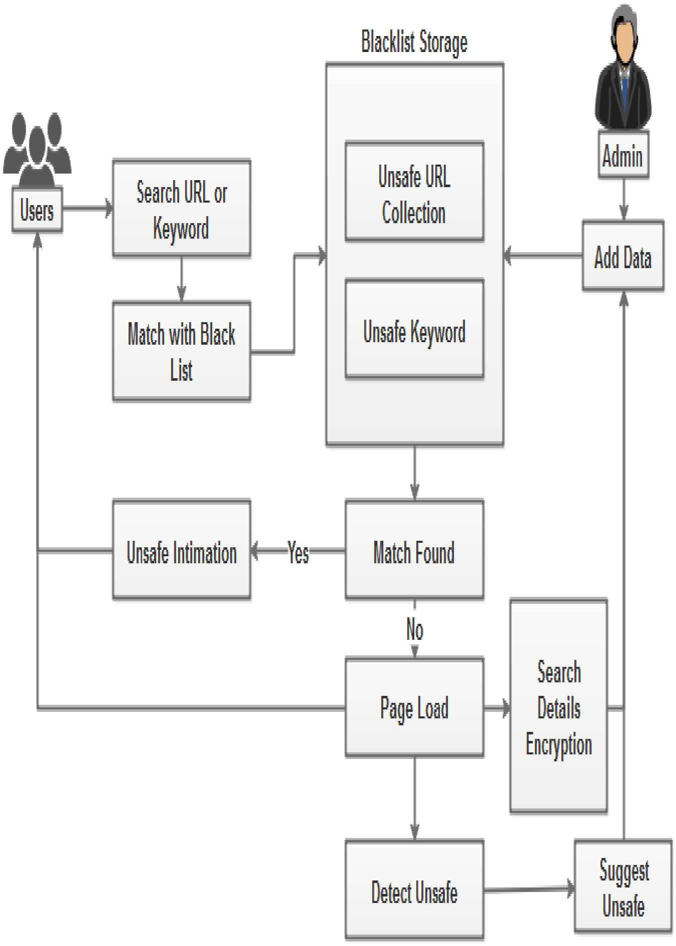

Fig 4.1: **Architecture for Proposed Work**

V CONCLUSION

In this proposed work, implement a Malicious URL Detection process using machine learning techniques. This focuses on detecting unsafe website URLs and keywords with the help of encrypted blacklist storage. According to few selected features can be used to differentiate between legitimate and malicious web pages. These selected features are many such as URLs and Keywords. In proposed work a service provider that owns a high-quality blacklist, which may be more frequently updated or simply contains more items. User also allowed to directly sharing blacklists with servers in an uncontrollable way could make these dataset be obtained by every user. With the help of efficient classification approach will detect the fake websites accurately and prevent the users from accessing that websites. This also provides the secure encryption approach avoid the unknown access of search history. The security is provided to the search data which has been stored in the database.

REFERENCES

[1] Mourtaji, Youness, Mohammed Bouhorma, Daniyal Alghazzawi, Ghadah Aldabbagh, And Abdullah Alghamdi. "Hybrid Rule-Based Solution For Phishing Url Detection Using Convolutional Neural Network." Wireless Communications And Mobile Computing 2021 (2021): 1-24.

[2] Ahammad, Sk Hasane, Sunil D. Kale, Gopal D. Upadhye, Sandeep Dwarkanath Pande, E. Venkatesh Babu, Amol V. Dhumane, And Mr Dilip Kumar Jang Bahadur. "Phishing Url Detection Using Machine Learning Methods." Advances In Engineering Software 173 (2022): 103288.

[3] Yuan, Jianting, Guanxin Chen, Shengwei Tian, And Xinjun Pei. "Malicious Url Detection Based On A Parallel Neural Joint Model." Ieee Access 9 (2021): 9464-9472.

[4] Xu, Pingfan. "A Transformer-Based Model To Detect Phishing Urls." Arxiv Preprint Arxiv:2109.02138 (2021).

[5] Butnaru, Andrei, Alexios Mylonas, And Nikolaos Pitropakis. "Towards Lightweight Url-Based Phishing Detection." Future Internet 13, No. 6 (2021): 154.

[6] Odeh, Ammar, Ismail Keshta, And Eman Abdelfattah. "Machine Learningtechniquesfor Detection Of Website Phishing: A Review For Promises And Challenges." In 2021 Ieee 11th Annual Computing And Communication Workshop And Conference (Ccwc), Pp. 0813-0818. Ieee, 2021.

[7] Butt, Muhammad Hassaan Farooq, Jian Ping Li, Tehreem Saboor, Muhammad Arslan, And Muhammad Adnan Farooq Butt. "Intelligent Phishing Url Detection: A Solution Based On Deep Learning Framework." In 2021 18th International Computer Conference On Wavelet Active Media Technology And Information Processing (Iccwamtip), Pp. 434-439. Ieee, 2021.

[8] Tang, Lizhen, And Qusay H. Mahmoud. "A Survey Of Machine Learning-Based Solutions For Phishing Website Detection." Machine Learning And Knowledge Extraction 3, No. 3 (2021): 672-694.

[9] Purbay, Madhurendra, And Divya Kumar. "Split Behavior Of Supervised Machine Learning Algorithms For Phishing Url Detection." In Advances In Vlsi, Communication, And Signal Processing: Select Proceedings Of Vcas 2019, Pp. 497-505. Springer Singapore, 2021.

[10] Wazirali, Raniyah, Rami Ahmad, And Ashraf Abdel-Karim Abu-Ein. "Sustaining Accurate Detection Of Phishing Urls Using Sdn And Feature Selection Approaches." Computer Networks 201 (2021): 108591.

[11] Jessica Scarpati, John Burke, “URL (Uniform Resource Locator (2021))” techtarget.com (Sep 2021).

[12] John Hughes, “What is a website URL? The 3most important parts explained” Themeisle.com (May 2022).

[13] Cho Do Xuan, Hoa Dinh Nguyen, “ Malicious URL Detection based on Machine Learning” International journal of advanced computer science and application (IJACSA) Vol. 11, No. 1, 2020.

[14] Kibreab Adane “Email and website based phishing attack: Examining online user’s security behaviour in cyber space environment” IEEE Xplore (Dec 2022).

[15] Ekta Gandotra, Deepak gupta “Chapter an efficient approach for phishing detection using machine learning” Springer.com (Jan 2021).

[16] Deepak Ram, Mohamed Thawfiq, S. Revathy & Others “URL Protection and bookmark hiding using AES algorithm” IEEE Xplore (JUL 2022).

[17] Sashank Dara “Advanced searchable Encryption over welldefined strings” ResearchGate.net (Aug 2014).

[18] S. Garera, N. Provos, and M. Chew,” A framework for detection and measurement of phishing attacks,” In proceeding of the ACM Workshop on Recurring Malcode, ACM, New York, NY, USA, Nov 2007.

[19] D. K. M. M. Gupta, ”Behind phishing: an examination of phisher modi operandi, ”In proceedings of the Usenix workshop on Largescale Exploits& Emergent Threats, DBLP, San Francisco, CA, USA, APRIL 2008.

[20] J. Ma, L. K. Saul, and S. Savage, “Beyond blacklists : Learning to detect Malicious websites from suspicious URL, “ In proceeding of the fifth ACM SIGKDD International conference on knowledge discovery and data mining, ACM, Paris, France, (JUL 2009).

[21] Jian Ting Yuan, Yipeng Liu, Long Yu, “ A Novel approach for malicious URL detection based on Joint model” doi.org (Dec 2021).