ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering and Technology
Vol. 6, Issue 4, April 2021

A Novel Hybrid Algorithm For Securing Data Over Wireless Transfer With Machine Learning Disease Prediction

[1] C. Anitha,[2] Priyadharshini.K.S

[1] AP / CSE

[1][2] Department of Computer Science And Engineering, RVS Technical Campus, Coimbatore-641402, Anna University, Chennai, India.

Abstract *Web-page recommendation plays an important role in intelligent Web systems. Useful knowledge discovery from Web usage data and satisfactory knowledge representation for effective Web-page recommendations are crucial and challenging. This paper proposes a novel method to efficiently provide better Web-page recommendation through semantic-enhancement byinte grating the domain and Web usage knowledge of a website. A number of effective queries have been developed to query about these knowledge bases. Based on these queries, a set of recommendation strategies have been proposed to generate Web-page candidates. The recommendation results have been compared with the results obtained from an advanced existing Web Usage Mining (WUM) method. The experimental results demonstrate that the proposed method produces significantly higher performance than the WUM method.Many kinds of research have shown that people are more likely to trust each other with the same attitude toward similar things. In this paper, we consider seeking and accepting sentiments and suggestions in E-commerce systems somewhat implies a form of trust between consumers during shopping. Following this view of point, an E-commerce system reviews mining oriented sentiment similarity analysis approach is put forward to exploring users' similarity and their trust. We divide the trust into two categories, namely direct trust, and propagation of trust, which represents a trust relationship between two individuals. The direct trust degree is obtained from sentiment similarity, and we present an entity-sentiment word pair mining method for similarity feature extraction.*

1. INTRODUCTION

Web Personalization

The process of providing the information that is related to user's current page is known as web personalization.

Web usage mining

It focus on determining patterns from user's browsing behaviour and use this pattern to personalize the browsing of the current user.

Web semantic mining

Semantic mining where the documents are clustered together based on their content and the user's recommended pages from the cluster to which the current document belongs.

1.1 HISTORY

Different efforts have been made to address the problem of information overload on the Internet. Web recommendation systems based on web usage mining try to mine users' behavior patterns from web access logs, and recommend pages to the online user by matching the user's browsing behavior with the mined historical behavior patterns. In this paper we propose effective and scalable technique to solve the web page recommendation problem. We use distributed learning automata to learn the behavior of previous users' and cluster pages based on learned pattern. One of the challenging problems in recommendation systems is dealing with unvisited or newly added pages. As they would never be recommended, we need to provide an opportunity for these rarely visited or newly added pages to be included in the recommendation set. By considering this problem, and introducing a novel Weighted Association Rule mining algorithm, we present an algorithm for recommendation purpose. We employ the HITS algorithm to extend the recommendation set. We evaluate proposed algorithm under different settings and show how this method can improve the overall quality of web recommendation.

Web Personalization

The process of providing the information that is related to user's current page is known as web personalization.

Web usage mining

It focus on determining patterns from user's browsing behaviour and use this pattern to personalize the browsing of the current user.

Web semantic mining

Semantic mining where the documents are clustered together based on their content and the user's recommended pages from the cluster to which the current document belongs.

II. LITERATURE REVIEW

2.1 In the past few years many researchers devoted their work to overcome these issues. Web access sequence (WAS) in Web usage data can be represented approaches based on tree structure and probabilistic model [1].These approaches learn from the training datasets to build the transition links between Web-pages. By using these approaches, given the current visited Web-page (referred to as a state) and k previously visited pages (the previous k states), the Web-page(s) that will be visited in the next navigation step can be predicted. The performance of these approaches depends on the sizes of training datasets. The bigger the training dataset size is, the higher the prediction accuracy is. However, these approaches make Web-page recommendations solely based on the Web access sequences learnt from the Web usage data. Therefore, the predicted pages are limited within the discovered Web access sequences, i.e., if a user is visiting a Web-page that is not in the discovered Web access sequence, then these approaches cannot offer any recommendations to this user. We refer to this problem as "new-page problem" in this study.

2.2 Depending on the domain of interest in the system, we can reuse some existing o ontologies or build a new ontology, and then integrate it with Web mining. For example, ontology concepts are used to semantically enhance Web logs in a Web personalization system [29]. In this system, ontology is built with the concepts extracted from the documents, so that the documents can be clustered based on the similarity measure of
the ontology concepts. Then, usage data is integrated with the ontology in order to produce semantically enhanced navigational patterns. Subsequently, the system can make recommendations, depending on the input patterns semantically matched with the produced navigational patterns.

2.3 The literature review covers the background, latest development of and related techniques for semantic-enhanced recommender systems. Web mining (WM) is the process of discovering useful knowledge from Web data. Depending on different types of Web data, appropriate mining techniques are selected. There are three main broad categories of Web mining.
»Web content mining (WCM) is used to mine Web content, such as HTML or XML documents.
»Web structure mining (WSM) focuses on Web structure, such as hyperlinks on Web-pages.
»Web usage mining (WUM) is applied to Web usage data, such as Web logs or clickstreams, from a website.

2.4 Domain ontology is commonly used to represent the semantics of Web-pages of a website. It has been shown that integrating domain knowledge with Web usage knowledge enhances the performance of recommender systems using ontology-based Web mining techniques [4-6]. Integrating semantic information with Web usage mining achieved higher performance than classic Web usage mining algorithms [5]. However, one of the big challenges that these approaches are facing is the semantic domain knowledge acquisition and representation. How to effectively construct the domain ontology is an ongoing research topic.

III. EXISTING SYSTEM:

It focus only to improve the quality of web personalization but none of them have experimented with considering the time spent on the web pages. Personalization based only on the web usage mining has the shortcoming of not taking the context of the web page into account.They proposed an approach ISTS that can exploit two factors from online social network. The sentiment orientation in friends posts about certain items and the trust relations between friends and a promising methodology to handle the trust mechanism for P2P network.

It focus only to improve the quality of web personalization but none of them have experimented with considering the time spent on the web pages. Personalization based only on the web usage mining has the shortcoming of not taking the context of the web page into account.They proposed an approach ISTS that can exploit two factors from online social network. The sentiment orientation in friends posts about certain items and the trust relations between friends and a promising methodology to handle the trust mechanism for P2P network.

ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 6, Issue 4, April 2021

DRAWBACKS OF EXISTING

- A single encryption algorithm is insufficient to transmit data securely.
- The system can be easily hacked.
- Cannot secure highly confidential data.
- Does not perform real time disease prediction

IV.PROPOSED SYSTEM

The system uses the information from web usage mining, web semantics and time spent on web pages to improve the recommendations. It focus on a system which evaluates user's interest based on combination of other user's browsing pattern, the content of web pages and the time spent by users on web pages.

ADVANTAGES:

A medical application is developed to perform diagnosis and to predict if the patient is diabetic or normal using machine learning algorithm.

An end to end data security is provided using hybrid secure algorithm. Ensures security over the wireless data transfer server to the database

It can be used in all applications where data is transferred wirelessly Bank and hospital details can be secured with utmost security

ARCHITECTURE DIAGRAM:

The architectural diagram is shown in Figure 4.1

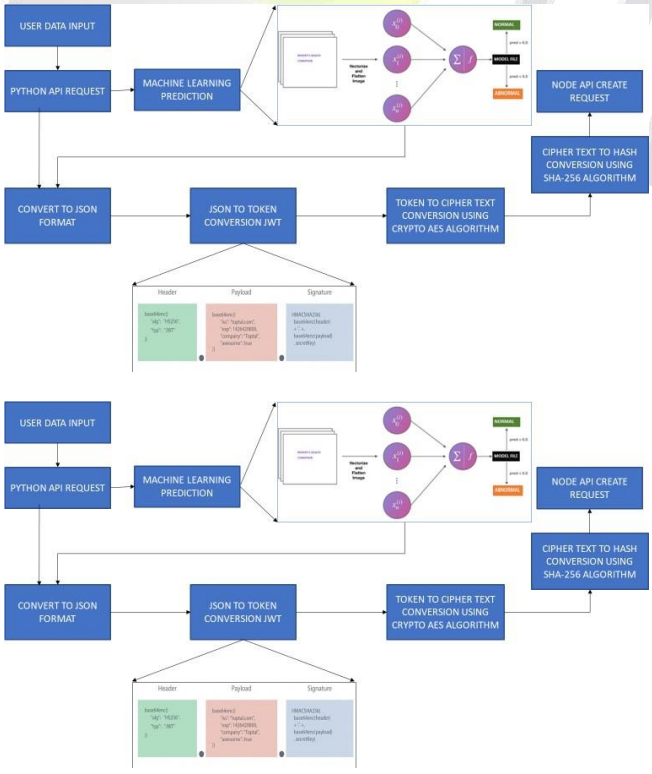
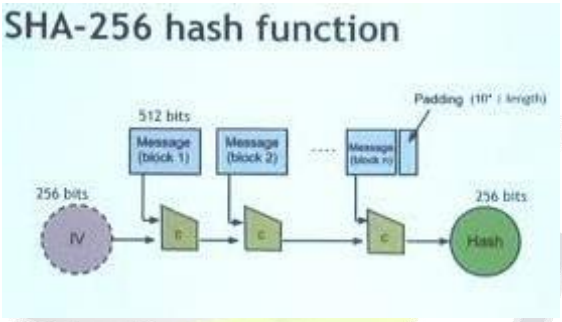


Figure 4.1 User-end & back-end system architecture
Figure 4.2 Disease prediction using ML

WORKING:

The working of the total model from the above diagrams can be explained as follows. The medical data to be secured is given through a web application which is converted to the JSON format, that is to be secured and encrypted is first converted to a token

using a HS-256



algorithm, with an inbuilt secret key. This token is then encrypted to a cipher text using a crypto AES algorithm, with an inbuilt secret key. These two secret keys from the JWT and crypto AES algorithm are converted to a hash key using a blockchain algorithm SHA-256 algorithm. This hash key is then embedded with the cipher text which is nothing but the end output of the encryption process. This medical data is sent from a reactJS web application to a mongoDB database through API which is developed using nodeJS. A machine learning algorithm logistic regression is used to predict if the patient is diabetic or normal based on the input medical data. The data secured through block chain can never be decrypted it can only be encrypted using the hash key of another member of a network. Thus, the project has successfully ensured secure data transfer over wireless communications.

SIGNIFICANCE OF SHA-256 ALGORITHM

SHA-256 is one of the most secure hashing functions on the market. The US government requires its agencies to protect certain sensitive information using SHA-
256. While the exact details of how SHA-256 works are classified, we know that it is built with a Merkle-Damgård structure derived from a one-way compression function itself created with the Davies-Meyer structure from a specialized block cipher.

Three properties make SHA-256 this secure. First, it is almost impossible to reconstruct the initial data from the hash value. A brute-force attack would need to make 2256 attempts to generate the initial data. Second, having two messages with the same hash value (called a collision) is extremely unlikely. With 2256 possible hash values (more than the number of atoms in the known universe), the likelihood of two being the same is infinitesimally, unimaginably small. Finally, a minor change to the original data alters the hash value so much that it's not apparent the new hash value is derived from similar data; this is known as the avalanche effect.

Figure 4.7 SHA-256 Structure

FUTURE WORK

In the coming future, we review the application of the project to determine technology in the fields that require data security and privacy. In the banking and medical field, they are more chance to develop or convert this project in many ways. Thus, this project has an efficient scope in coming future where data can be transmitted securely over a wireless communication.

CONCLUSION:

This project is used to provide a solution to perform diagnosis for diabetics and securely transmit data through a wireless communication medium using a web application developed using a javascript framework reactJS. By this project, we can protect data from unauthorized access. Thus, this project provides an affordable and efficient means to protect and preserve data.

REFERENCES

[1] Borja BORDEL, Diego MARTIN, Ramón ALCARRIA and Tomás ROBLES ," A Blockchain -based Water Control System for the Automatic Management of Irrigation Communities",2019

[2] Caixue Zhou," Comments on "Light-weight and Robust Security-Aware D2D- assist Data Transmission Protocol for Mobile-Health Systems"",[Vol:8937,2017]

[3] CaioDavi, André Pastor, Thiego Oliveira, Fernando B. de Lima Neto, Ulisses Braga-Neto, Abigail W. Bigham,MichaelBamshad, Ernesto T. A. Marques, Bartolomeu Acioli-Santos," Dengue E protein detection using graphene oxide integrated tapered optical fiber sensor",[Vol:6781,2018]

[4] CaioDavi, Andre Pastor, Thiego Oliveira, Fernando B. de Lima Neto, Ulisses Braga-Neto, Abigail W. Bigham, Michael Bamshad, Ernesto T. A. Marques, Bartolomeu Acioli-Santos,"Severe Dengue Prognosis Using Human Genome Data and Machine Learning"[2020]

[5] Mauro Mangia, Member, IEEE,, Alex Marchioni, Student Member, IEEE,, Fabio Pareschi, Member, IEEE,RiccardoRovatti, Fellow, IEEE,, Gianluca Setti, Fellow,
IEEE,"Chained Compressed Sensing: A Block-Chain-inspired Approach for Low- cost Security in IoT Sensing",[2019]

[6] Mary Subaja Christo, AnigoMerjora A, ParthaSarathy G, Priyanka C and Raj Kumari M," An Efficient Data Security in Medical Report using Block Chain Technology", April 4-6, 2019

[7] Y. Mustapha Kamil, Member, IEEE, M. H. Abu Bakar, Member, IEEE, M. H. Yaacob, Senior Member, IEEE, A. Syahir, H. N. Lim, M. A. Mahdi, Senior Member, IEEE," A New Intelligence-Based Approach for Computer-Aided Diagnosis of Dengue Fever"[Vol:2341-6781,2019]

[8] Shuai Wang , Liwei Ouyang, Yong Yuan , Senior Member, IEEE, Xiaochun Ni, Xuan Han,and Fei-Yue Wang , Fellow, IEEE," Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends",[Vol:26,2019]

[9] VadrevuSree Hari Rao, Senior Member, IEEE, and Mallenahalli Naresh Kumar," Real-Time Dengue Prediction using Naive Bayes Predicator in the IoT",[Vol:6785,2018]

[10] Junqin Huang, Linghe Kong, Senior Member, IEEE, Guihai Chen, Min-You Wu, Xue Liu, Senior Member, IEEE, Peng Zeng," Towards Secure Industrial IoT: Blockchain System with Credit-Based Consensus Mechanism",[Vol: 1551- 3203,2019]

[11] Junqin Huang, Linghe Kong, Senior Member, IEEE, Guihai Chen," Novel Hybrid CMOS/Memristor Implementation of the AES Algorithm Robust against