E-learning system using cryptography and data mining techniques

Vijaya Patil[1], Aditi Vedpathak[2], Pratiksha Shinde [3], Vishakha Vatandar[4], Prof. Surekha Janrao[5]

[1,2,3,4] Student, Dept. of computer Engineering, Terna Engineering College, Maharashtra, India
[5]Professor, Dept. of computer Engineering, Terna Engineering College, Maharashtra, India

---

Abstract - E-learning is one amongst the foremost wide used technique for the education. In this paper, we have introduce e-learning platform in which system provides security to learning material and for candidate as well. We are aiming to implement the application of data mining in e-learning system that has specific needs, mainly the need to take into consideration learner's specific behaviour. Security is provided by Elliptic curve cryptography algorithm and content is filtered using Decision tree algorithm.

Key words - Cryptography, data mining techniques, e-learning.

I. INTRODUCTION

Many companies, institutions and organizations have adopted e-learning as a promising solution to provide on-demand learning for their employees and students. They applied several technologies to deliver text, audio, streaming video used in learning and teaching. There are several benefits of e-learning, such as reducing costs, time and improving performance of learning.

E-learning systems facilitate a spread of tasks associated with learning like self-study, guided learning, tutoring, communication etc. Privacy may be a natural concern at an equivalent time that trust is a crucial think about associate e-learning surroundings. Cryptography algorithm is used for authorizing the users into the E-Learning system. It generates key for each users when they are entering their information in the registration page. E-Learning system allows only authorized person to enter into the website to access the services. Data mining can be used to classify the learners' information and to find out the more useful courses for that learners which could be more beneficial to them. Classification technique is used for prediction of learner's behaviour.

II. LITERATURE SURVEY

In [1] R. Krishnaveni, V. Pandiyaraju published paper A Secure E-Learning System and its Services which has introduced a system which use elliptic curve cryptography algorithm for securing the system and they sugested support vector machine algorithm for filtered web search engine.

In [2] Aski and Torshizi proposed The use of classification Techniques to enrich e-learning environments. This paper discuss about four classification method which were used to classify the student's information and found out which method gives accurate result.

In [3] Gaurav Yadav and Apana Majare published paper A comparative study of performance analysis of various encryption algorithms. This paper proposed some recent existing cryptography techniques and their performance according to different parameters used in algorithm. This comparison analysis shows which algorithm is best suited in which environment.

In [4] Yong Chen We He published paper Security risk and protection in online learning. This paper will offer necessary insights and tips so that online learning providers can become proactive and knowledgeable as they mitigate the security risks found e-learning environments.

As more and more people are studying online more attention and efforts are needed from e-learning providers to prevent possible security breaches in online learning. There are lots of issue comes with e-learning like, the attackers steal the user's authentication or the information are eavesdropped in the insecure communication, unauthorized users alter or modify the content of the information by executing malicious codes. Unstructured inform-ation chokes the educational system without providing any articulate knowledge to its actors. Hence there is a need for secure e-learning which uses data mining techniques which would help to avoid the above problems.

III. PROPOSED SYSTEM

A. Proposed Framework

The proposed system has three main entities Admin, Student and teacher. Student can enter into the system by doing registration process and select specific course. In next step student recommended for study material as well as related courses. Student can access the e-learning platform any where and learn the course. Student can ask query to the teacher in the discussion forum. Teachers can add or delete study material. Teachers design the course and gives solutions to the students query in discussion forum. The most crucial task of course management and

content management done by admin. Admin can add students/Teachers or delete students/Teachers. In this way all this entities interact with proposed system. The overall system architecture is represented in the figure 1.

By studying literature survey in proposed system we have observed following points:

- Proposed system contains huge bulk of database as course details, learner information, and study materials.
- For classifying these database data mining concepts gives better idea. Like decision tree algorithm helps for proper recommendation.
- System has smart interaction between student and teacher.
- System provides security to learning material and candidate similarly.
- System reaches up to some extends to produce higher data, higher expertise victimization digital learning.

B. Methodology

Proposed system has two phases 1) Security 2) Data Mining

1) Security:

Cryptography*:*

Cryptography is the practice and study of techniques for secure communication in the presence of third parties called adversaries. More generally, cryptography is about constructing and analysing protocols that prevent third parties or the public from reading private messages. There are many cryptography algorithms which can be led to the security. One of them is Elliptic curve cryptography algorithm.

Elliptic Curve Cryptography*:*

In this system user security provided by Elliptic Curve Cryptography algorithm. Elliptic-curve cryptography (ECC) is an approach to public-key cryptography based on the algebraic structure of elliptic curves over finite fields. ECC requires smaller keys compared to non-ECC cryptography to provide equivalent security. ECC is the best suited in constrained environments. The advantages like speed and smaller Keys are especially important in environments where at least one of the following resources is limited: processing

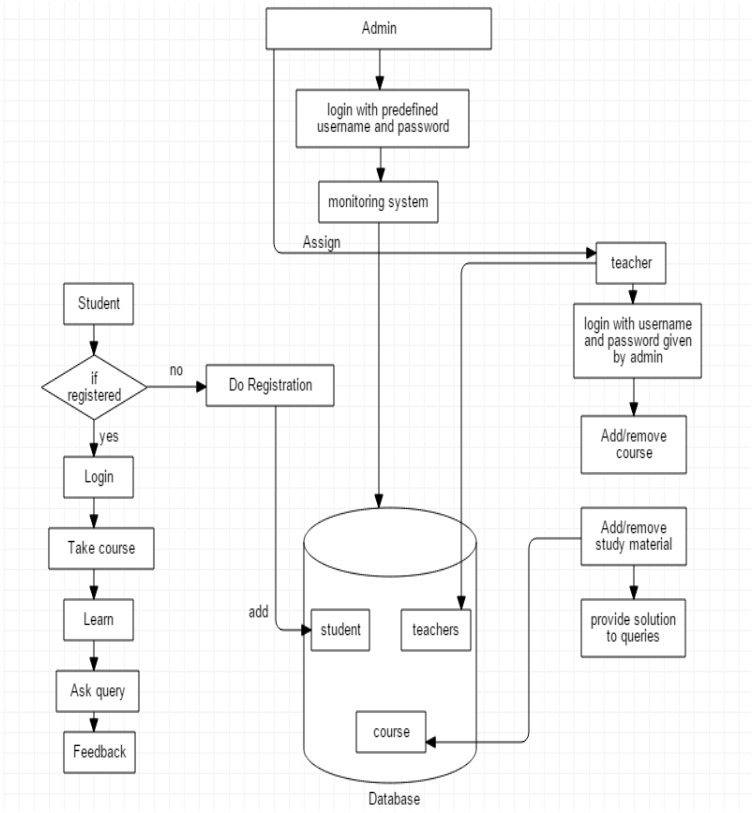

Fig. 1: System framework

power, storage space, band width, or power consumption. This advantage is because its inverse operation gets harder, faster, against increasing key length than do the inverse operations in Diffie Hellman and RSA.

Table 1 is referred from R. Krishnaveni and V. Pandiyaraju[1]. This table shows a comparison of the RSA and ECC cryptographic operations performed by an SSL server. These micro benchmarks highlight ECC's performance advantage over RSA for different security levels. ECC's performance advantage increases even faster than its key-size advantage as security needs increase.

Encryption

1. Define a Curve.
2. Generate public private Key pair using that curve, for both sender and receiver.
3. Generate a Shared secret key from the key pair.
4. From that shared secret key, generate an encryption key.
5. Using that encryption key and symmetric encryption algorithm, encrypt the data to send.

	ECC 160	RSA 1024	ECC 192	RSA 1536	ECC 224	RSA 2048
Ops/ Sec	271.3	114	268.5	36.4	195.5	17.8
Performance	2.4:1		7.4:1		11.4:1	
Key Size Ratio	1:6.4		1:8		1:9.1	

Table 1: Measured Performance of public-key algorithm

Decryption

The sender can either share the curve with receiver or sender and receiver can have the same use for the same curve type. Also, sender can share its public key with receiver.

Generate public private Key pair using the same curve for that curve. For receiver.

1. Regenerate a shared secret key using private key of receiver and public key of sender.
2. From that shared secret key, generate an encryption key.
3. Using that encryption key and symmetric encryption algorithm, decrypt the data.

ECC is an asymmetric cryptography algorithm. The algorithm for ECC is given as:

- Suppose Alice wants to send to Bob an encrypted message. Both agree on a base point, B.
- Alice
  Private Key = a
  Public Key = PA = a * B
- Bob
  Private Key = b
  Public Key = PB = b * B
- Alice takes plaintext message, M, and encodes it onto a point, PM, from the elliptic group
- Alice chooses another random integer, k from the interval [1, p-1]

The cipher text is a pair of points

PC = [(kB), (PM+ kPB)]

- To decrypt, Bob computes the product of the first point from PC and his private key, b **b * (kB)**
- Bob then takes this product and subtracts it from the second point from PC
- **(PM + kPB) – [b(kB)] = PM + k(bB) – b(kB) = PM**
- Bob then decodes PM to get the message, M.

2) Data Mining

The amount of saved data in data warehouses has been rapidly increased. So one needs a useful method for automatic and intelligent organization of large data collection. Data mining is used to discover and show some knowledge in an understandable form. The aim of data mining is description and prediction. There are many strategies in data mining which can be led to the prediction. One of them is classification.

Classification

Classification is based on available features that leads to new data description and causes a better understanding of each class in a Database, so classification can prepare a model to describe the proper class for any given data. In other words by using classification, we can predict that which given data would belong to which predefined class. Different statistical techniques are used for classification functions like; Bayesian, Neural Network, Decision Tree and Support Vector Machine.

Decision Tree

Decision tree algorithm adopts a super incumbent model to create a tree structure from the given data set where each node represents attributes test or conditions and final leaf node represents the test results or classes. The construction of decision tree is done by divide and conquer strategy. The attribute used here should be of categorical and if it is of continuous values then it had to be converted to discrete values before starting the process. Initially all samples are on the single root node and then the remaining nodes are created based on the attribute partitioning condition. This is a recursive process and the stopping condition for this are:

- If all the samples in each node belongs to same class
- If there are no more samples to be portioned
- If there is no further division to be made in the given sample

The algorithm for decision tree is given as:

Decision tree method:

1. Initially create a new node N.
2. If all the data samples are in same category C, then return N and mark it as class C.
3. If attribute list of node is empty, the return N and label with class C whichever is major in that set of samples in the node.
4. Using attribute selection method find out the best Splitting criteria.
5. Use Splitting criteria to mark N.
6. In attribute list remove the splitting attributes.

7. Create new node N with the resultant set of attributes and continue the process.
8. Return N

In this method all the attributes should be classified hence the continuous values have to be discretized before starting the process. This method is easy to understand, execute and validate. Validation of the algorithm can be done using simple statistical tests.

Statistical technique of the Classification	Prediction Precision
Decision Tree	80.113%
Simple Bayesian	80.113%
Support Vector Machine	78.409%
K Nearest Neighbour	72.443%

Table 2: Comparison Result different algorithms

Table 2 is referred from Aski and Torshizi[2]. We can see that when we use the Decision Tree or Simple Bayesian techniques for our classifications, we would acquire more precision in our learners' studying result predictions.

CONCLUSION

In this paper, Efficient E-Learning system is provided to the learners. Elliptic curve cryptography algorithm provides secured authentication. Recommendation of courses is obtained by Decision Tree algorithm to classify the content based on category and subcategories which are mentioned by learners.

REFERENCES

[1] R. Krishnaveni, V. Pandiyaraju "A Secure E-Learning System and its Services", International Journal of Advanced Computational Engineering and Networking, ISSN: 2320-2106 Volume- 1, Issue- 4, June-2013.

[2] Aski, B. A. and Torshizi, H. A. Islamic Azad University, Iran. "The Use of Classification Techniques to Enrich e-Learning Environments", Online proceedings of the university of Salford Fifth Education in a Changing Environment conference, September-2009.

[3] Gaurav Yadav and Aparna Majare "A comparative Study of Performance Analysis of Varipous Encryption Algorithms", International Conference on Emanations in Modern Technology and Engineering 2017.

[4] Yong Chen We He "Security Risks and protection in Online Learning: A Survey", The international Review of Research in Open and Distributed Learning, December-2013.

[5] Priyanka Pradhan, R.B.Kulkarni International conference on Electrical, Electronics, and Optimization Techniques-2016 "Secure e-learning using data mining techniques and concepts".