**SECURING DATA WITH ENCRYPTED K-NN DEFENDING AGAINST STATISTICAL ANALYSIS**

**Preface**

**TABLE OF CONTENTS**

1. SOFTWARE DEVELOPMENT LIFE CYCLE

2. PLATFORM KNOWLEDGE

3. ABOUT THE PROJECT

3.1 ABSTRACT

3.2 SCOPE OF THE PROJECT

3.3 EXISTING SYSTEM

3.3.1 DISADVANTAGES

3.4 PROPOSED SYSTEM

3.4.1 ADVANTAGES

4. BOTTOM LINE AND FUTURE ENHANCEMENT

5. HARDWARE AND SOFTWARE REQUIREMENTS

**1. SDLC (Software Development Life Cycle)**

The Software Development Life Cycle is a systematic process for building software that ensures the quality and correctness of the software built. SDLC process aims to produce high-quality software which meets customer expectations. The software development should be completed within the pre-defined time frame and cost.

**SDLC Phases**

The entire SDLC process is divided into the following stages:



Phase 1: Requirement collection and analysis

Phase 2: A feasibility study

Phase 3: Design

Phase 4: Coding

Phase 5: Testing

Phase 6: Installation/Deployment

Phase 7: Maintenance

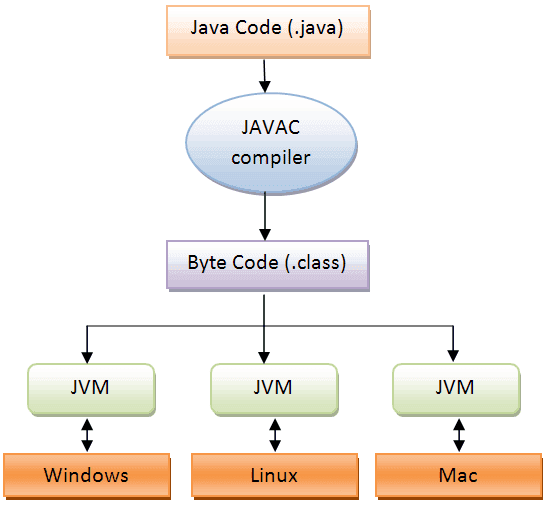
**2. PLATFORM KNOWLEDGE**

**Introduction to java**

Java programming language was originally developed by Sun Microsystems which was initiated by James Gosling and released in 1995 as a core component of Sun Microsystems' Java platform. Initially, the language was called "Oak" but it was renamed "Java" in 1995. The primary motivation of this language was the need for a platform-independent language. Finally, Java is for Internet Programming whereas C was for System Programming.

**Java architecture**

Java is a high-level Object-oriented programming language. A program written in a high-level language cannot be run on any machine directly. First, it needs to be translated into that particular machine language. The javac compiler does this thing, it takes the java program (.java file containing source code) and translates it into machine code (referred to as byte code or .class file). Java Virtual Machine (JVM) is a virtual machine that resides in the real machine (your computer) and the machine language for JVM is byte code. JVM executes the byte code generated by the compiler and produces output. JVM is the one that makes the java platform independent.



**SECURING DATA WITH ENCRYPTED K-NN DEFENDING AGAINST STATISTICAL ANALYSIS**

**4. ABOUT THE PROJECT**

**4.1 Abstract :**

In the article we proposed the Encryption and decryption using cryptographic algorithm AES(Advance Encryption Standard) is used to secure the data on the educational details and for important educational certificate in an organization or educational department, AES relies on substitution-permutation network principle which means it is performed using a series of linked operations which involves replacing and shuffling of the input data others involve shuffling bits around (permutations). The implementation of cryptographic algorithms is used to achieve the prevention of significant data loss and to avoid malevolent situations, which will provide an optimal learning environment. Cryptography is used to limit data leakage and ensure data security. The Key generator has been using to generate different keys randomly with help of RNG(Random Number Generator) at each time while encrypting the data, It has the symmetric key chipper block is used to generate the key as range in 128 or 192 or 256 bit of key value. A Key Schedule algorithm is used to calculate all the round keys from the key. So the initial key is used to create many different round keys which will be used in the corresponding round of the encryption. The mode of operation will handles the process in series of sequential message in an blocks, there are various encryption modes such as ECB, CBC, CFB, OFB, CTR etc., This application has contains the CBC(Cipher Blocking Chain) mode of to improve the vulnerability, through this multiple blocks has been encrypted as parallel process. It also avoids the bit flipping attacks from the foreign object.

**4.2 Scope of the project**

The purpose of the Advance Encryption Standard Algorithm is used to makes a symmetric block of cipher data (secret keys), it will be implemented in software and hardware to throughout the world to encrypt sensitive data. A 256-bit encryption key is significantly more difficult for brute-force attacks to guess than a 128-bit key, because the latter takes so long to guess, even with a huge amount of computing power, it is unlikely to be an issue for the foreseeable future. AES encryption refers to the process of concealing electronic data using an approved 128-bit, 192-bit, or 256-bit. When hackers want to access a system, they will aim for the weakest point. This is typically not the encryption of a system, regardless of whether it's a 128-bit key or a 256-bit key. Users should make sure the software under consideration does what they want it to do, that it protects user data in the way it's expected to and that the overall process has no weak points and makes it in decipherable to hackers and other individuals attempting to access your data without authorization. AES algorithm works, a sender sends a plaintext file to an encryption server, where an AES secret key is used to encrypt it, or turn it into cipher text. The receiver of the encrypted file can only read the contents of the file once that same secret key is used to decrypt the contents of the file.

**4.3 Existing System**

In an existing, the educational organization is run as manual certificate distribution and it have more chance to malpractice, frauds or it will destroyed in any global disasters, so that student and organization will may suffer to handle the situation, the overall organization will may face the critical risk situation through these things. In Future also the student has a difficulty to face the problems for them to recover the difficult situation and problems will happens on joining to any interested jobs. The educational department also has responsibilities by providing the students certificate and to securing the details of the students. \

**4.3.1 Disadvantages**

* In traditional life student facing fraudulent, malicious activities on the important certificates. Natural disaster will also damage student data.
* The missing of data is also more possible, So that educational department may face a big problem and it causes the popularity among people.
* It has a high possibility to duplicate the important files or data as manually and as well as systemized.
* The security is lacking on the traditional situation, along with that educational department aspects will be reduced among the people.

**4.4 Proposed System**

A Cryptographic algorithm will organize a better solution from the existing system to make the secure environment to fulfill the needed requirements, The usage of the cryptography is to reduce the data leakage and provide data security, the cryptographic algorithm implementation is used to achieve the prevention of huge data loss and it also avoid the malicious situation it will make excellent educational environment. The AES algorithm is used for the encryption, which is one of the more secured algorithm with large size of key value is used. The key is generated to decrypt the encrypted data, here the CBC(cipher block cryptography) is used for the data protection. It make the digitalized student records has been provided and it has secured with cryptography algorithm to store securely in the database. The better data security will secure the massive data and prevent the data loss or malicious activities from other interpreters. Here the algorithm implementation will make better way to handle negative situations

**4.4.1 Advantage**

* The data loss has been prevented from the natural disaster and sudden unhandled situation will be recovered from the database.
* The AES algorithm will improve the security level and the key generation is randomly generated it will prevent the data loss from the foreign objects
* It reduce the manual process and improve the digital wellbeing with highly secured algorithm implementation it has been stored in database
* The Encryption process will handle all types of data

**4.5 BOTTOM LINE AND FUTURE ENHANCEMENT**

Our proposed model has implemented the cryptography algorithm has been implemented and it has been useful for the educational department. It also describes the course of action to manage the risk on data loss and malicious activities, It plays a vital role on securing the data, risk management and it helps to avoid the data duplication, The cryptography implementation is used to secure the important and massive data storage with the algorithm implementation it helps the educational department to handle the student by securing the details. Thus our proposed model makes the great impact and satisfies required need in educational industry. In future it has been enhanced and applied with experimented for an effective needed situations.

**5. HARDWARE AND SOFTWARE REQUIREMENTS**

**Hardware requirements:**

* Processor : Intel (R) Pentium (R)
* Speed : 1.6 GHz and Above
* RAM : 4 GB and Above
* Hard Disk : 120 GB
* Monitor : 15’’ LED SVGA
* Input Devices : Keyboard, Mouse

**Software requirements:**

* Operating system : Windows 7 / 8 / 8.1 / 10
* Coding Language : JAVA / J2EE
* Java Version : JDK v8
* IDE : Eclipse Oxygen / Neon
* Database : MySQL v5.1
* Database Tool : HeidiSQL v11.0
* Application Server : Apache Tomcat 8.x / 9.x