02
2021

Journal of Computer Science Research

BILINGUAL PUBLISHING CO.
Powered Global Academics Since 1984

AOSCI Asia-Pacific Science Citation Index
creative commons
CNKI中国知网 www.cnki.net 中国知识基础设施工程
Google scholar
Crossref
MyScienceWork

Volume 3 Issue 2 · April 2021·ISSN: 2630-5151 (Online)

ISSN 2630-5151
02
9 772630 515216

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Volume 3 Issue 2 • April 2021 • ISSN 2630-5151 (Online)

Journal of Computer Science Research

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Volume 3 | Issue 2 | April 2021 | Page 1-51
Journal of Computer Science Research

Contents

ARTICLE

REVIEW

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

ARTICLE

Integration of Expectation Maximization using Gaussian Mixture Models and Naïve Bayes for Intrusion Detection

Loka Raj Ghimire Roshan Chitrakar*

Department of graduate study, Nepal College of Information Technology, Nepal

ARTICLE INFO

Article history
Received: 27 February 2021
Accepted: 17 March 2021
Published Online: 20 April 2021

Keywords:
Anomaly detection
Clustering
EM classification
Expectation maximization (EM)
Gaussian mixture model (GMM)
GMM classification
Intrusion detection
Naïve Bayes classification

ABSTRACT

Intrusion detection is the investigation process of information about the system activities or its data to detect any malicious behavior or unauthorized activity. Most of the IDS implement K-means clustering technique due to its linear complexity and fast computing ability. Nonetheless, it is Naïve use of the mean data value for the cluster core that presents a major drawback. The chances of two circular clusters having different radius and centering at the same mean will occur. This condition cannot be addressed by the K-means algorithm because the mean value of the various clusters is very similar together. However, if the clusters are not spherical, it fails. To overcome this issue, a new integrated hybrid model by integrating expectation maximizing (EM) clustering using a Gaussian mixture model (GMM) and naïve Bays classifier have been proposed. In this model, GMM give more flexibility than K-Means in terms of cluster covariance. Also, they use probabilities function and soft clustering, that's why they can have multiple cluster for a single data. In GMM, we can define the cluster form in GMM by two parameters: the mean and the standard deviation. This means that by using these two parameters, the cluster can take any kind of elliptical shape. EM-GMM will be used to cluster data based on data activity into the corresponding category.

1. Introduction

Recently, through their networks, many organizations have encountered heavy network use. The large technological expansion that followed these networks, however, gave them different threads. Such threads include many types of malicious programs that affect network efficiency or unauthorized network access to data. This has encouraged work to strengthen and develop new ways of addressing and mitigating these threats. Any unauthorized operation on a computer network constitutes a network intrusion [1].

Intrusion detection is a "species of security technology that can collect information from some of the network or computer system's key points and attempt to analyze it to assess whether there is a violation of the security policy or a suspicion of the computer system's network attack." Intrusion detection methods are classified into two groups according to the different objects for intrusion detection. One is called the identification of anomaly that is used to detect the unknown intrusion. And the other is called detection of abuse, which is used to detect the identified intrusion.

Mixed intrusion detection techniques have been fo-

**Corresponding Author:*
Roshan Chitrakar,
Department of graduate study, Nepal College of Information Technology, Nepal;
Email: roshanchi@gmail.com

 DOI: https://doi.org/10.30564/jcsr.v3i2.2922

cused on to resolve shortcomings in anomaly detection and misuse detection methods. The anomaly detection model and signature detection system can be paired with three different strategies: anomaly detection followed by misuse recognition, identify anomalies and misuse concurrently, and misuse identification accompanied by anomaly detection [1].

While new technologies in intrusion detection and research have been suggested, the accuracy and detection rate as well as the false alarm rate have still to be improved. The proposed method provides high detection and precision compared to previous attack detection with low false alarm rate by using a hybrid model.

2. Related Work

Dorothy Denning first described intrusion detection in 1987 [2]. According to him, "network intrusion can be detecting by monitoring network activity in terms of data and then the system can generate alerts and responses before the infringement". Instantaneity is one of the key features of intrusion. Snort IDS applied the rule-based intrusion detection method [3, 4]. Rule-based detection system has quick detection characteristics, but it has a big problem. It cannot detect other than pre-defined types of attack. Since intruders will frequently change their technique of attack, which is often riskier. For this case, this approach cannot adopt itself so that it has not been suitable in new types of attacks. It also has a higher false alarm rate.

Intrusion detection using data mining technique requires extensive data collection in advance. Large quantities of data limit the rate of online detection [5]. Conventional intrusion detection methods are being developed using data mining [6–7] and common file analyzed [8]. In differential analyzes performed by Fisher, an et al. [9] used the approach of combining the minimum scatter class with a traditional support vector (SVM) analysis and then implemented a minimum scatter support class vector (WCS-SVM) analysis, which is better than traditional SVM. Kabir et al [10] suggested a vector based intrusion detection method (LS-SVM) that supports the least squares, called (LS-SVM) method. The new method of improved decision mapping for intrusion detection was introduced by M. Gudadhe, Al. [11] to develop an intermediate classifier for multiple decision makers. Sufyan et al. [12] used backpropagation models for artificial neural networks to detect intrusion, encouraging intrusion detection system to adapt more effectively respond to new environments and new attack types. The vast scale of the network data set takes time and effort for manual tagging. The classification of the dataset is therefore subject to clustering methods [13]. The Ymeans clustering algorithm [14] surmounts two disadvantages of K-means clustering. This is dependency and deterioration of k-means by splitting the set of data automatically into a correct number of clusters. The k-means clustering algorithm is a simple algorithm that solves the complexity of previous clustering algorithms. Traditional SOM algorithm has some disadvantages like, not providing accurate result while clustering. This has been overcome by integration of SOM and k-means [15]. One of the major problems in clustering is to determine the cluster center and number of clusters. High speed, high detection can be achieved by the parallel clustering integration algorithm [16] for IDS. The ANN classifier [17] has a good performance in the detection of intrusion. Research in [18–20] uses a mixed learning approach to have a higher detection. Shah et al. [21] compared directly to the Snort intrusion detection system and the machine learning detection performance and found the better performance in machine learning detection system.

Sheng Yi Jang et al. proposed a clustering-based intrusion detection method [22] wherein clusters consist of unlabeled datasets and have been classified as normal or abnormal by their external factors. This method's time complexity is linear with the dataset size and number of attributes. A method for anomaly detection by clustering regular user behavior is proposed by Sang Hyum Oh et al. [23] to model a user's typical behavior using the clustering algorithm. Clustering prevents statistical analysis causing inaccuracy. Therefore, the user's daily habits are more reliable than the statistical analysis.

Tasi and Lin use K-Means clustering in K-clusters to cluster data instances [24]. Next the study trains the latest dataset consisting only of cluster centers with support vector machine (SVM). They managed to achieve a high precision rate for nearly all types of attacks. This approach provides a high rate of detection but comes with a high false alarm rate.

The new approach of the IDS based on the Artificial Neural Network (ANN) with the clusters ANN and Fuzzy FC-AN Network, is suggested by Gang, Jin Xing and Jian [25]. Before a similar ANN model is trained, fuzzy clustering is carried out to formulate different models to produce different training subsets. A fuzzy module of aggregation is then used to sum the result. The subset of the training set is less complex with the use of fuzzy clusters that help the ANN learn from each subset more effectively and to detect low frequency attacks such as U2R and R2L attacks. Nevertheless, in contrast with the Naïve Bayes approach, this approach results in a lower detection rate for probing attacks.

DOI: https://doi.org/10.30564/jcsr.v3i2.2922

Shaohua et al. [26] suggested detection of intrusion based on Fuzzy SVMs (FSVM) to improve classification accuracy. The clustering algorithm's aim is to build a new training set using cluster centers. This new set will then be trained to get a support vector with FSVM. Although their findings have shown that the accuracy rate has been improved by this approach, it is not an adequate percentage.

Amiri et al. [27] used a feature selection method to improve the performance of existing classifiers by eliminating unimportant features like SVM with heavy computational challenges for large datasets. The authors have recently introduced the support vector machine of an improved least squire called PLSSVM. PLSSVM performs well in the classification of normal records and probes but misses many dynamic attacks that are very similar to normal behavior, such as DOS and U2R.

Horng [28] suggested hierarchical clustering of SVM-based IDS BIRCH as a pre-processing step and a basic feature selection method to remove unimportant features. The hierarchical clustering algorithm enhances SVM's efficiency while the simple selection of features allows the SVM model to properly classify some data. As this method was unable to differentiate between R2L and Normal data, the percentage of predictions for this class dropped dramatically.

In terms of classification accuracy and AUC, Huang, Lu and Ling [29] performed a comparative study of Naïve Bayes, Decision tree and SVM. They found that both Naïve Bayes and SVM have very similar predictive accuracy as well as similar AUC scores are produced.

Roshan Chitrakar and Huang Chauanhe proposed a hybrid anomaly detection approach using K-medoids clustering and support vector machine classification [30]. Since there may be too many support vectors in the case of using a high dimensional kernel, this also reduces the training speed, KMeans / Medoids needs a large sample and can only handle spherical shape.

S. Varuna and Dr. P. Natesan proposed an integrated model of K- Means clustering and Naïve Bayes classification for intrusion detection [1]. The integrated algorithm improved the detection rate for the normal, Probe, R2L and U2R attacks, but it does not meet the requirements for DOS.

This paper is organized as follows: Section 3 describes the proposed work and the implementation details. Section 4 contains the results and discussion.

3. Proposed Model

In this research an integrated model has been proposed. This is the integration of Expectation Maximization using Gaussian Mixture Model clustering and Naïve Bayes classifier. Data are clustered and formed five clusters with outlier. The purpose of clustering is to label the data with enhancing the accuracy and performance of model by improving capacity of parallel processing of the model. Thus, clustered data with outlier are then classified using Naïve Bayes classifier.

3.1 Description of Dataset

Each dataset record reflects a 41-feature network connection. Among them, 7 are nominal features, 34 are continuous features and a label. Label indicates that the data is either in normal status or in one of the 39 identified attack status. The NSL-KDD data can be categorized as either a standard class or one of four attack classes, i.e. remote to local, denial of service, Users to root and Probe classes.

Table 3.1 lists the number of instances in the training and testing data set of every type of attack group and the total number of instances in each data set.

Table 3.1 Size and Distribution of Training and Test Data Based on Attack Class

Attack Class	Training data size	Test data size
Normal	67343	9711
Prbe	11656	2421
Remote to local	995	2754
Denial of services	45927	7456
User to root	52	200
Total	125973	22542

3.2 Feature Scaling and Selection

There are 41 attributes in the NSL-KDD dataset. In this analysis, 14 common and basic characteristics, also known as traditional characteristics, are used.

3.3 Conceptual Model Diagram

The proposed model consists of three sub modules. These are data preprocessing module, clustering and classification module with outlier detector and decision module. In the first module all the functionalities of data preprocessing such as feature selection, feature scaling, data encoding is performed. In the second module, data are cluster to the appropriate number of clusters with outlier detection. Thus, clustered data with outlier are then classified using Naïve Bayes classifier. The third module is a decision-making module.

 DOI: https://doi.org/10.30564/jcsr.v3i2.2922

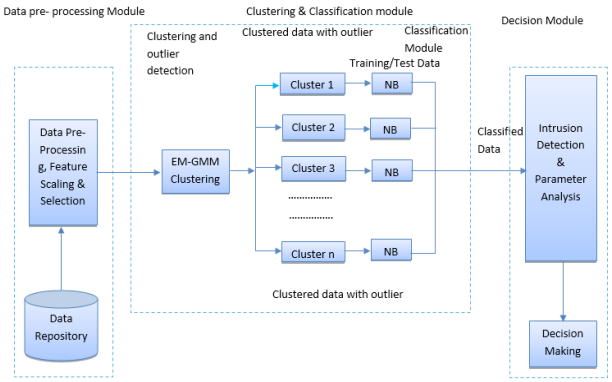


Figure 1. Conceptual Model Diagram of Proposed method

3.4 Algorithm

Algorithm 1 Data Clustering

Input: Dataset
Output: K number of clusters with outlier
Initialization:
1: Randomly choose μ_k, Σ_k, π_k
2: Specify k
3: Choose an initial random gaussian parameter θ
4: E step
5: Estimate the value of the latent variables Υ_k
6: Compute P (Z_i = k |X_i, θ);
7: M step
8: Update gaussian parameters μ_k, Σ_k, π_k
9: if
10: log-likelihood value converges
11: Stop
12: Else
13: Compute Υ_k and update μ_k, Σ_k, π_k
14: Assign data to appropriate cluster
15: End

3.5 Outlier Detection

Outer detection is the method of detecting the pattern in data that did not expect property. Following is the process of outlier detection:

• Randomly choose data in the dataset and measure the distance of the data to all other data. If the distance between the data and certain data is below the radius that we already set, assign that certain data as a neighbor, then assign the data and its neighbors as 1 cluster.

• Do as in previous step but the data is replaced by its neighbors. Neighbors of the neighbor are in the same cluster with previous data. Do this step until all detected neighbor is chosen.

• When all detected neighbor is chosen, construct a new cluster using data that has not been chosen. The new clusters are formed as in steps first and second.

The data that are not part of any cluster considered as an outlier.

Table 3.5 Outlier Statistics

Outlier = TRUE	32194	25.55%
Outlier = FALSE	125973	74.45%

3.6 Clustering

Clustering is a non-supervised approach to machine learning, but it can be used to maximize the precision of the supervised machine learning algorithm and cluster the data point into similar groups.

The purpose of clustering is to create dataset sub-population based on clustering results and to develop separate cluster classification models. Clustered membership can be considered as a feature in the classification and may have more details from these features. That improves the parallel processing capability of the model and manages data skews. That increases the accuracy of the classification.

3.6.1 EM Clustering

The expectation maximization (EM) clustering algorithm measures probabilities of cluster membership based on one or more distributions of probabilities. The goal of the clustering algorithm is then, given the (final) clusters, to maximize the overall likelihood or probability of the results.

Each gaussian j (j=1,2…..k) is defined in the EM clustering by its own μ and σ^2 as:

$$P_X\left(\frac{x}{\mu^j}, \sigma_j^2\right) = N(X; \mu^j, \sigma_j^2) = \frac{1}{\left(2\pi\sigma_j^2\right)^{\frac{d}{2}}} e^{\frac{-\left|x-\mu^j\right|^2}{2\sigma_j^2}} \quad (1)$$

Where,
μ is mean
σ is standard variable
x- μ is the distance between two points.
Each gaussian component has a mixture weight that indicates the likelihood.

3.7 Maximum Likelihood Estimation

In construction of a Bayesian classifier the class-conditional probability density functions need to be determined. The initial model selection can be done for example by visualizing the training data, but the adjustment of the model parameters requires some measure of goodness, i.e., how well the distribution fits the observed data. Data likelihood is such a goodness value. Assume that there is a set of independent samples X = {X_1, ... …X_N} drawn from

DOI: https://doi.org/10.30564/jcsr.v3i2.2922

a single distribution described by a probability density function P (x; θ) where θ is the PDF parameter list. The likelihood function can be written as:

$$L(X;\theta) = \prod_{n=1}^{N} p(x_n;\theta) \quad (2)$$

Equation (2) indicates the probability of X due to its distribution parameters θ. The goal is to calculate $\hat{\theta}$ which optimize the likelihood.

$$\hat{\theta} = \arg\max_{\theta} L\ (X;\theta) \quad (3)$$

This function is generally not explicitly maximized, but rather the logarithm as:

$$L(X;\theta) = \ln L(X;\theta) = \sum_{n=1}^{N} \ln p(x_n;\theta) \quad (4)$$

This is due to easier to handle logarithm function analytically. The limit can be identified analytically according to P (x; θ) by setting the derivatives of the log-like function to zero and θ resolution. A gaussian PDF can be used which leads to the estimation of intuitive mean and variance but the research approach is generally intractable. In this case, the iterative method, such as the EM algorithm, is used in practice [1]. Maximizing the likelihood in certain situations will lead to unique estimates, which is the main issue of highest probability methods. The function of classifying vector in K classes is recalled by Gaussian mixture model. If different classes are treated as distinct (i.e., class samples don't say anything about other courses), the k class-conditional PDF estimation problem can be divided into K separate estimation problems.

3.8 Gaussian Mixture Probability Density Function

In a single dimensional bell-shaped curve, the Gaussian probability density function is defined by two parameters; mean (μ) and variance (σ^2). But, for D dimensional space it is in matrix form as:

$$N(x;\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}\,|\Sigma|^{1/2}} e^{\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right]} \quad (5)$$

Where,

Σ is a matrix of covariance

μ is the mean vector.

Gaussian surfaces are μ-centered hyperellipsoids.

The gaussian mixture model (GMM) consists of a mixture of several Gaussian distributors, thus representing different subclasses within a class. The probability density fction is also known as the weighted sum of Gaussian.

$$P(x;\theta) = \sum_{c=1}^{C} \propto_c\ N(x;\mu_c,\Sigma_c) \quad (6)$$

were

α_c is the component weight c, $0 < \alpha_c < 1$ for all components, and PC c=1

α_c is the list of parameters whose value is equal to 1.

$$\theta = \{\alpha_1, \mu_1, \Sigma_1, ..., \alpha_C, \mu_C, \Sigma_C\} \quad (7)$$

defines a fundamental Gaussian density.

3.9 Basic EM Estimation

Suppose, X is all good features of sample and Y is all unknown features of sample, then the expectation (E) step of the EM algorithm is

$$Q(\theta;\theta^i) \equiv E_Y[\ln L(X,Y;\theta) \mid X;\theta^i] \quad (8)$$

Where θ^i is the previous distribution parameter estimate and θ is the distribution-descriptive estimation variable for the new estimate. L is the probability function which determines the likelihood of the data, including the unknown attribute Y marginalized in relation to the current distribution estimate defined by θ^i. Maximization step (M) is to optimize Q for θ and set steps are repeated until the conditions of convergence have been met.

$$\theta^{i+1} \leftarrow \arg\max_{\theta} Q\ (\theta;\theta^i) \quad (9)$$

It is proposed in [14] that the convergence parameters

$$Q(\theta^{i+1};\theta^i) - Q(\theta^i;\theta^{i-1}) \le T \quad (10)$$

with a correctly chosen T and in [18] that

$$\|\theta^{i+1} - \theta^i\| \le e \quad (11)$$

The EM algorithm begins with an initial distribution parameter guess, which ensures that the log-likelihood will increase on each iteration up to converge. Convergence results in a local or global limit, but it can also lead to specific estimates, especially for Gaussian mixture distributions with arbitrary matrices. The definition and implementation of the general EM algorithm for the Gaussian mixture model can be found in [6,1,4]. One of the main problems of EM algorithm is to initialize. The selection of theta (θ) defines where the algorithm converges or reaches the space parameter boundary that generates singular, insignificant results. Many solutions use random multiple starts or a clustering initialization algorithm [7]. The Gaussian mixtures implementation of the EM algo-

 DOI: https://doi.org/10.30564/jcsr.v3i2.2922

rithm as follows:

Let,

X is incomplete data

Y is knowledge of component that produced each sample X_n

For each X_n, a binary vector is assigned as:

$y_n = \{y_{n,1}, \ldots, y_{n,c}\}$

where,

$y_{n,c} = 1$, if component c or zero otherwise was generated in the sample.

The maximum probability of data log is

$$\ln L(X, Y; \theta) = \sum_{n=1}^{N} \sum_{c=1}^{C} y_{n,c} \ln\left(\propto_c p\left(x_n \mid c; \theta\right)\right) \quad (12)$$

The propose of E step is to calculate conditional expectancy for the whole log-like data, Q-function is produced by X and θ^i is current parameters estimation. As the whole data log-like function in L (X, Y; θ) is straightforward to the missing Y. Conditional expectation W simply needs to be determined and placed in ln L (X, Y; θ). That's why

$$Q(\theta, \theta^i) \equiv E \ln L(X, Y; \theta) | X, \theta^i = \ln L(X, W; \theta) \quad (13)$$

Where:

W elements have been defined as

$$\omega_{n,c} \equiv E [y_{n,c} | X, \theta^i] = Pr[y_{n,c} = 1 | x_n, \theta^i] \quad (14)$$

The estimate is determined using the Bayes law

$$\omega_{n,c} = \frac{\propto_c^i p(x_n \mid c; \theta^i)}{\sum_{j=1}^{C} \propto_j^i p(x_n \mid j; \theta^i)} \quad (15)$$

Where $\propto_c^i$ is the probability of a priori, and $\omega_{n,c}$ is the likelihood of posteriori of $Y_{n,c} = 1$ after observing X_n. In other words, "$\omega_{n,c}$ is the probability that X_n was produced by component c" [21].

If the M-step is used to evaluate the distribution parameters for C-component Gaussian mixture, with Arbitrary covariance matrices the following formulas will be used:

$$\propto_c^{i+1} = \frac{1}{N} \sum_{n=1}^{N} \omega_{n,c} \quad (16)$$

$$\mu_c^{i+1} = \frac{\sum_{n=1}^{N} x_n \omega_{n,c}}{\sum_{n=1}^{N} \omega_{n,c}} \quad (17)$$

$$\sum_{c}^{i+1} = \frac{\sum_{n=1}^{N} \omega_{n,c} \left(x_n - \mu_c^{i+1}\right)\left(x_n - \mu_c^{i+1}\right)^T}{\sum_{n=1}^{N} \omega_{n,c}} \quad (18)$$

previous numbers are now x θ^{i+1}. Unless the convergence criterion (Equations 10 or 11) is met, i ← i + 1 and Equations 15-18 new models are being tested again. [15]

weight α_c of the item is the sample portion of the element. The conditional PDF variable is estimated with the preliminary parameter estimates, and later the likelihood is determined for each sample point of c. The mean μ component is calculated in the same way as a covariant matrix Σ_c. The samples are evaluated according to the probability of the variable and the sample average and covariance matrix are calculated.

Table contains classification statistics, the number of instances transmitted into each cluster, and the proportion of instances from each cluster's total data.

Table 3.9 Clustered Instances

	No. of instances	% of instances
Cluster 1	45108	36%
Cluster 2	34025	27%
Cluster 3	13432	11%
Cluster 4	27394	22%
Cluster 5	6013	5%

3.10 Classifier

A classifier may adjust a number of parameters to the function. This is known as training. The samples in the training are labelled in supervised learning and the training algorithm tries to reduce the training set's classification error. Unsupervised learning does not label samples, but the training algorithm recognizes clusters and classes. The training samples are not also classified in reinforcement learning, but the training algorithm uses input to inform whether or not to identify a sample properly [40].

3.10.1 Bayesian Classification

Bayesian classification and its decisions are based on the probability theory and on the idea that the most likely or lowest risk i.e, expected cost is chosen. Suppose there is a classification task in which to assign functional vectors to K various classes. A vector function is labelled with $x = [X_1, X_2, \ldots\ldots, X_D]$ T. Where, D is the dimension of a vector. Probability that a feature vector x belongs to class ω_k is $p\left(\frac{\omega_k}{x}\right)$, and this is referred to as a posteriori

DOI: https://doi.org/10.30564/jcsr.v3i2.2922

probability. The vector's classification is based on the subsequent probabilities or decision risks determined from the probabilities. The conditional probability can be determined by Bayes formula as

$$P\left(\frac{\omega k}{x}\right) = \frac{P\left(x/\omega k\right)P\left(\omega k\right)}{P\left(x\right)} \quad (19)$$

where $P\left(\frac{X}{\omega_k}\right)$ is the probability density function of class ω_k in the feature space and $P(\omega_k)$ is the a priori probability. That gives the likelihood class before any characteristics are calculated. When previous probabilities are not known, they can be calculated in the training set according to the class proportions.

$$P(x) = \sum_{i=1}^{k} P\left(\frac{x}{\omega i}\right) P(\omega i) \quad (20)$$

It's just a factor in scaling to ensure that later probabilities are actual probabilities, that is, their sum is 1. Choosing the lowest retrograde likelihood class will illustrate the minimum pbability of error [1,4]. However, if the costs of making various types of error are not consistent, a risk function can be used which calculates the expected cost with the following probabilities and selects the lesser-risk class. The main problem in the Bayesian classification is the class-conditional density function $p\left(\frac{x}{\omega K}\right)$. The function defines the dispersion of feature vectors within a specific class, i.e., the class model. It is always unclear in reality, except for certain artificial classification activities. With a variety of methods, the distribution can be calculated in the training set.

3.11 Unit of Results

The model performance is calculated based on the following parameters and unit.

3.11.1 Accuracy of Classification

It is the proportion of correctly classified.

$$\text{Classification accuracy} = \frac{\frac{TP}{TN}}{\left(TP+TN+FP+FN\right)}$$

3.11.2 Sensitivity (True Positive Fraction)

It is the percentage of the number of properly identified attack.

$$\text{Sensitivity} = \frac{TP}{\left(TP+FN\right)}$$

3.11.3 Specificity (True Negative Faction)

It is the percentage properly categorized.

$$\text{Specificity} = \frac{TN}{\left(TP+FN\right)}$$

3.11.4 False Alarm Rate (FAR)

It is the percentage of the number of normal connections in correctly classified.

$$\text{False alarm rate (FAR)} = \frac{FP}{\left(TN+FP\right)}$$

3.11.5 Detection Rate (Precision)

It is the rate of detection of total anomaly from the total flow of packets in the network.

$$\text{Detection rate (DR)} = \frac{TP}{\left(TP+FP\right)}$$

Where,

True positive (TP) = Attacks that are correctly detected as attack.

True negative (TN) = Normal data that are correctly detected as normal.

False positive (FP) = Normal data that are incorrectly detected as attack.

False negative (FN) = Attack that are incorrectly detected as normal.

4. Results and Discussion

Based on obtained result, the overall accuracy in compared with different algorithms. The obtained result is illustrated in the following table.

Table 4. Result Comparison of Different Algorithm

Attack Class	K-NN	C4.5	SVM	DSSVM	K means with NB	Proposed method
Normal	98.3	97.0	97.7	98.4	74.11	97.48
DoS	97.0	96.8	97.2	97.2	86.05	81.65
Probe	79.4	84.3	86.1	87.5	92.48	97.13
R2L	6.5	3.0	7.2	6.3	32.02	95.17
U2R	11.8	4.4	9.2	3.1	19.0	73.66

From the above discussion, it is cleared that low frequency attack (probe, R2L, U2R) detection rate is improved in the integrated models. In proposed model this rate is significantly improved. Also, the detection rate for

 DOI: https://doi.org/10.30564/jcsr.v3i2.2922

normal class also improved in competitive ratio with the existing algorithms.

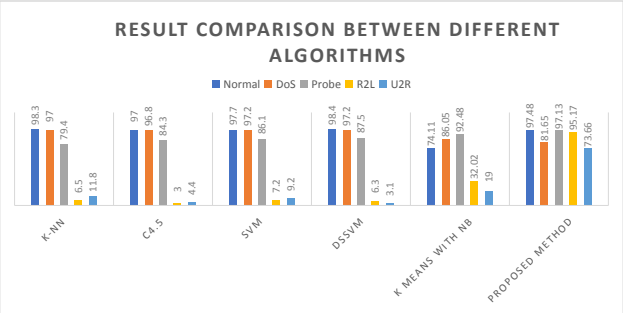


Figure 2. Result comparison between different algorithm with proposed model

From the above comparison chart with various models, overall performance is beaten by EM GMM with naïve Bays (proposed method) for low frequency attack i.e, R2L and U2R. Also, the performance for Prob is better than other models but the performance of DoS class is higher in other intrusion detection systems.

Except DoS, overall performance of proposed model is better than integration of K-means clustering with Naïve Bayes.

In this paper, we tried to simulate proposed model with various parameters with different ratio of training/testing model and calculate different matrices based on the obtained result. These metrics are objective measurements that are calculated mathematically defined algorithms. The comparison table for the experimental result is shown above in the table.

5. Conclusions

The research work observed with overall performance winner as integration of Expectation Maximization clustering with Naïve Bayes classifier for intrusion detection over Integration of K-Means clustering and Naïve Bayes classifier is considered to be best in terms of precision, sensitivity, specificity, and false alarm rate for the different types of attack class such as Probe, R2L, U2R and normal. It is shown that clustering plays a supportive role for classification by parallel computation so that the computation capacity of the model is improved. Since the whole dataset is clustered in a K number of clusters and compute parallelly, it can be used as real time/online computation with full efficiency computation on large data.

As the overall result of this model is significantly improved in different attack classes such normal, probe, R2L and U2R. But the other intrusion detection system has a higher detection rate for DoS attack.

Further improvement can be done in a number of ways. Firstly, the overall accuracy of DoS can be improve. Next improvement can be done in reducing the computation time at outlier detection.

References

[1] S. Varuna, Dr. P. Natesan "An Integration of K-Means Clustering and Naïve Bayes Classifier for Intrusion Detection." 2015 3rd international conference on signal processing, communication and networking " ICSCN. 978-1-4673-6823-0/15. 2015 IEEE.

[2] D. E. Denning, "An intrusion-detection model," IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222-232, 1987.

[3] W. Parkand S. Ahn, "Performance Comparison and Detection Analysis in Snortand Suricata Environment," Wireless Personal Communications, vol.94, no.2, pp.241-252, 2016.

[4] R. T. Gaddam and M. Nandhini, "An analysis of various snort based techniques to detect and prevent intrusions in networks: Proposal with code refactoring snort tool in Kali Linux environment," in Proceedings of the 2017 International Conference on Inventive Communication and Computational Technologies, ICICCT2017, pp.10-15, India, March 2017.

[5] C.-T. Huang, R. K. C. Chang, and P. Huang, "Signal Processing Applications in Network Intrusion Detection Systems," EURASIP Journal on Advances in Signal Processing, vol. 2009, Article ID 527689, 2 pages, 2009.

[6] U. Adhikari, T. H. Morris, and S. Pan, "Applying Non-Nested Generalized Exemplars Classification for Cyber-Power Event and Intrusion Detection," IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 3928-3941, 2018.

[7] R. Taormina and S. Galelli, "A Deep Learning approach for the detection and localization of cyber-physical attacks on water distribution systems," Journal of Water Resources Planning & Management, vol.144, no.10, Article ID 04018065, 2018.

[8] F. Raynal, Y. Berthier, P. Biondi, and D. Kaminsky, "Honeypot forensics,"in Proceedings of the Proceedings from the Fifth Annual IEEE System, Man and Cybernetics Information Assurance Workshop, SMC, pp.22-29, USA, June 2004.

[9] W. J. Anand M. G. Liang, "A new intrusion detection method based on SVM with minimum within-class scatter," Security and Communication Networks, vol.6, no. 9, pp. 1064-1074, 2013.

[10] E. Kabir, J. Hu, H. Wang, and G. Zhuo, "A novel statistical technique for intrusion detection systems," Future Generation Computer Systems, vol. 79, pp. 303-318, 2018.

DOI: https://doi.org/10.30564/jcsr.v3i2.2922

[11] M. Gudadhe, P. Prasad, and K. Wankhade, "A new data mining based network intrusion detection model," in Proceedings of the 2010 International Conference on Computer and Communication Technology, ICCCT-2010, pp. 731-735, India, September 2010.

[12] S. T. Al-Janabi and H. A. Saeed, "A Neural Network Based Anomaly Intrusion Detection System," in Proceedings of the 2011 Developments in E-systems Engineering (DeSE), pp. 221-226, Dubai, United Arab Emirates, December 2011.

[13] K. D. Denatious and A. John, "Survey on data mining techniques to enhance intrusion detection," in Proceedings of the International Conference on Computer Communication and Informatics, pp. 1-5, 2012.

[14] Y.Guan, A. A. Ghorbani, and N. Belacel, "Y-means: A clustering method for intrusion detection," in Proceedings of the CCECE 2003 Canadian Conference on Electrical and Computer Engineering: Toward a Caring and Humane Technology, pp. 1083-1086, Canada, May 2003.

[15] H.-B. Wang, H.-L. Yang, Z.-J. Xu, and Z. Yuan, "A clustering algorithm use SOM and K-means in intrusion detection," in Proceedings of the 1st International Conference on E-Business and E-Government (ICEE'10), pp. 1281-1284, May2010.

[16] H. Gao, D. Zhu, and X. Wang, "A Parallel Clustering Ensemble Algorithm for Intrusion Detection System," in Proceedings of the 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES), pp. 450–453, Hong Kong, China, August 2010.

[17] Akashdeep, I. Manzoor, and N. Kumar, "A Feature Reduced Intrusion Detection System Using ANN Classifier," Expert Systems with Applications, vol. 88, pp. 249-257, 2017.

[18] Z. Muda, W. Yassin, M.N. Sulaiman, and N.I. Udzir, "Intrusion detection based on K-Means clustering and Naïve Bayes classification," in Proceedings of the 7th International Conference on Information Technology in Asia (CITA '11), pp. 1-6, IEEE, July 2011.

[19] M. Ishida, H. Takakura, and Y. Okabe, "High-performance intrusion detection using OptiGrid clustering and grid-based labelling," in Proceedings of the 11th IEEE/IPSJ International Symposium on Applications and the Internet, SAINT 2011, pp. 11-19, Germany, July 2011.

[20] H. Om and A. Kundu, "A hybrid system for reducing the false alarm rate of anomaly intrusion detection system," in Proceedings of the 2012 1st International Conference on Recent Advances in Information Technology, RAIT-2012, pp. 131-136, India, March 2012.

[21] S. A.R. Shah and B. Issac, "Performance comparison of intrusion detection systems and application of machine learning to Snort system," Future Generation Computer Systems, vol. 80, pp. 157-170, 2018.

[22] J. S. Yi., X. song, H. Wang, J.-J. Han and Q.-H. Li, "A clustering-based method for unsupervised intrusion detections." Pattern recognition letters 27, no. 7 (2006): 802-810.

[23] Oh, S. Hyum, and W. S. Lee. "An anomaly intrusion detection method by clustering normal user behavior." Computer and security 22, no.7 (2003): 596-612.

[24] C.F. Tasi and C.Y. Lin 2010. "A triangle area-based nearest neighbors approach to intrusion detection." Pattern recognition, 43(1): p.222-229.

[25] W. Gang, H. Jinxing and M. Jian 2011. "A new approach to intrusion detection using artificial neural networks and fuzzy clustering. Expert systems with applications, 376: p.6255-6232.

[26] Shaohua, D. Hongle, W. Naiqi, Z. Wej and S. Jiangyi, 2010. "A cooperative network intrusion detection based on fuzzy SVMs. Journals of networks, 5: p. 475-483.

[27] F. Amiri, F. Mohammad, R. Y. Caro, L. Azadeh, S. and Y. Nasser 2011. "Mutual information-based feature selection for intrusion detection system." Journal of network and computer applications, 34: p.1184-1199.

[28] S.J. Horng 2011 "A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert systems with applications. 38(1) :P.399-408.

[29] J. Huang, J. Lu, C. X. Ling, "Comparing Naïve Bayes, Decision trees, and SVM with AUC and accuracy." The third international conference on data mining 2003.

[30] R. Chitrakar and H. Chauanhe "Anomaly detection using support vector machine classification with K-medoids clustering". 978-1-4673-2590-5/12. 2012 IEEE.

[31] F. Kelly. "The mathematics of traffic in networks." The Princeton companion to mathematics, 1(1):862-870, 2008.

[32] Z.Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir "K-Means clustering and Naïve Bayes classification for intrusion detection." Journal of IT in Asia Vol 4 (2014).

[33] V.-E. Neagoe, V.C.-Berbentea "Improved Gaussian mixture model with Expectation Maximization for clustering of remote sensing imagery." 978-1-5090-

 DOI: https://doi.org/10.30564/jcsr.v3i2.2922

3332-4/4/16. 2016 IEEE.

[34] A. Reddy, M. Ordaway-West, M. Lee, M. Dugan, J. Whitney, R. Kahan, B. Ford, J. Muedsam, A. Henslee, & M. Rao "Using Gaussian Mixture models to detect outliers in seasonal univariate network traffic." DOI 10.1109/SPW.2017.9 IEEE computer society 2017.

[35] E. A. Shams and A. Rizaner, "A novel support vector machine based intrusion detection system for mobile adhoc networks," Wireless Networks, pp.1-9, 2017.

[36] W. Shang, L. Li, M. Wan, and P. Zeng, "Industrial communication intrusion detection algorithm based on improved one-class SVM," in Proceedings of the World Congress on Industrial Control Systems Security, WCICSS 2015, pp. 21-25, UK, December 2015.

[37] T. Jan, "Ada-Boosted Locally Enhanced Probabilistic Neural Network for IoT Intrusion Detection," in Proceedings of the Conference on Complex, Intelligent, and Software Intensive Systems, pp. 583-589, Springer, 2018.

[38] O. Osanaiye, K.-K. R. Choo, and M. Dlodlo, "Distributed denial of service (DDoS) resilience in cloud: review and conceptual cloud DDoS mitigation framework," Journal of Network and Computer Applications, vol.67, pp.147-165, 2016.

[39] H. Li, "Research and Implementation of an Anomaly Detection Model Based on Clustering Analysis," Journal of Beijing Information Science & Technology University, pp. 458-462, 2010.

[40] R. O. Duda, P.E. Hart, and D.G. Stork. Pattern Classification. John Wiley & Sons, Inc., 2nd edition, 2001.

DOI: https://doi.org/10.30564/jcsr.v3i2.2922

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

REVIEW

Enhancing Primary School Teaching through Virtual Reality

Vasileios Drakopoulos* **Panagiotis-Vlasios Sioulas**

Department of Computer Science and Biomedical Informatics, Faculty of Science, University of Thessaly, Lamia, Greece

ARTICLE INFO

Article history
Received: 5 January 2021
Accepted: 22 February 2021
Published Online: 18 April 2021

Keywords:
Virtual reality
360° videos
Religion
Primary school

ABSTRACT

In this day and age, the usage of computers as well as Internet combined with mobile devices is an integral part of our routine especially for adolescents and younger children. Thus, it puts forward a multitude of challenges and advances for educational institutions. The purpose of this article is to explore the current use of virtual reality in order to support teaching and learning along with presenting a teaching proposal concerning the utilisation of CoSpace Edu software on the subject of Religious Affairs.

1. Introduction

In development platforms, Augmented Reality, or AR for short, as well as Virtual Reality, or VR for short, were referred to as the "fourth wave." In both business and educational contexts, personal computers, Internet, and mobile apps, AR and VR applications are now taking their place. AR and VR have, as if their predecessors, changed the way we connect and interact with people and the world around us [1].

Advances in the field of technology and computing have brought about the need for a shift in the paradigm of teaching. Therefore, educators should take advantage of the familiarisation of youngsters with gaming and applications [2] so as to incorporate VR in the classroom. If we were to provide a definition of VR, we could highlight that it is the experience through which although users enter a virtual world comprised of 3D objects through the use of a headset attached to a computer or mobile device, they still preserve their physical presence in the real world [3]. What is more, within such a simulated environment differentiated feedback is generated and it could be auditory, visual, haptic and sensory. In order for an application to be regarded as VR, it should exploit 3D, real or fictional models of objects. Another trait is that head movement and adjustment of view on the part of the user are prerequisites.

The origins of contemporary VR date back to 2012 when Oculus Rift was initiated and introduced into the market in order to offer a more cost-effective high-quality Head-Mounted Display (HMD) to the potential consumers. Therefore, prior to the official launch, a variety of models was developed in order for a number of applications to be developed [4]. According to[5], the use of a head-mounted device dwindles the user's surrounding reality and provides a shift to another environment.

**Corresponding Author:*
Vasileios Drakopoulos,
Department of Computer Science and Biomedical Informatics, Faculty of Science, University of Thessaly, Lamia, Greece;
Email: vdrakop@uth.gr

 DOI: https://doi.org/10.30564/jcsr.v3i2.2777

2. VR in Education

Technologies aiding teaching and learning have been in the spotlight for the past decade. Except their application in gaming, increased interest has been shown in the use of such technology in educational contexts. On the grounds that it opens up a window to a whole new array of differentiated learning experiences [6]. Virtual reality is broadly applicable. It is currently being applied to areas of education including natural science, technology training, history, architecture and medicine. In education, virtual reality has found a new area in which to showcase its full potential [7]. The learning methodologies that have the greatest effect on current educational systems are those that present students with a specific problem they have to solve using acquired theoretical knowledge or through improving students' capacities that are non-existent or underdeveloped until that moment.

The particular situation can be programmed through virtual reality technology with several variables and environments on which the student can act. Applications can be customised to suit every subject, knowledge area, population segment or geography [8]. Access to information would be more inclusive thanks to those kinds of technologies. Students struggling to meet certain learning goals with a poor success rate should now be able to effectively reach the goals.

Another major area where virtual reality is providing a more than significant value is in the representation of abstract concepts [9]. Laboratories completely simulated through this technology allow interaction between the student and the devices [10]. Taking this analysis further, the cost savings in space would be huge. The underutilized space within the centers would be significantly reduced and would be replaced by "multi-laboratory" room in which, according to the subject, one laboratory or another could be accessed [11].

Undoubtedly, the merit of VR technology use in comparison to more conventional methods is that learners are able to accurately depict and illuminate characteristics or processes that they would otherwise not be able to recognize or recount with more conventional methods. In other words, VR can help the learners experience while simultaneously observe various perspectives of an object which may not have even been considered let alone seen in the past.

3. Positive Aspects

A number of studies have been carried out concerning the utilisation and practicality of VR in both education and training. Duncan, Miller and Jiang [12] highlight that VR worlds may provide space for joint work, entertainment and socialization within an educational or learning context. Gilbert [13] claims that more often than not learners reckon science subjects to be incomprehensible and complex; therefore, they require a depth of understanding and visualization skills that can be provided through VR environments.

The issue of misconception can be tackled through the use of visualisation technologies such as virtual reality. Despite the indisputable merits of mobile and VR technologies, it is vital that we investigate the positive and negative aspects of using them in educational environments [14]. Distractions can be eliminated through the exploitation of VR which can be extremely beneficial for students with disorders such as anxiety disorder, impulse-control disorder or attention-deficit disorder while at the same time may address them by capturing their interest [15].

Moreover, virtual and augmented reality allow students to interact and learn in environments beyond their physical reach [16]. For example, students experiencing the aforementioned disorders would be able to take a virtual tour of well-known museums globally, explore the prehistoric world of dinosaurs or even carry out experiments without any fear of injury.

3.1 Use of Examples

Throughout primary and secondary education contexts, virtual reality is one of the mostly used developments. Programmes such as Google Expeditions and Google Earth allow for virtual visits of student to landmarks, museums, places of interest globally without any need for students to leave the classroom. Another example of such use of virtual technology is Mind and Anatomy 4D through which learners are given the insights to explore the brain and body organs while EON Experience provides content for teaching and learning History and Science. The use of VR can simulate events and submerge learners in the virtual world. As a result, they perceive events, points, traits and differentiations that would otherwise be difficult to comprehend and consolidate [1]. VR could act as a trigger for motivating learners that would be indifferent towards a subject and build a positive outlook in using VR in their learning process [17]. In other words, VR triggers students' involvement in the learning process since they are challenged and urged to interact, explore and manipulate objects, visualize with precision which could in no way be feasible in a conventional educational environment [18].

3.2 Elements in Learning

There appear to be three elements in learning that act

DOI: https://doi.org/10.30564/jcsr.v3i2.2777

as catalysts; motivation, clear-cut goals and sufficient practice. If fulfilled, education may become a totally intriguing and captivating experience. The teaching and learning process tends to detach from the traditional classroom and move towards utilizing a virtual environment involving computer-generated 3D models barely exploited in the past partly owing to a number of constraints such as familiarization of educators with this type of technology, cost-efficiency or even hardware limitations [19]. However, recent progress in the field of technology has resolved such problems, creating new educational opportunities that are less costly and more efficient in the long run [20].

Although it seems to be implausible to radically alter the way teaching is performed in the classroom, VR is undoubtedly going to enable teachers to enrich their learning with entertaining and enticing experiences that will make learning and consequently teaching more appealing to both counterparts. Apart from primary and secondary education, higher education institutions can benefit from using VR. Instead of having undergraduate and postgraduate students read about a topic in textbooks, they could become members of a virtual laboratory. The latter allows students to investigate a scientific phenomenon and perform learning by doing; through a kind of virtual hands-on learning. It goes without saying that being outside the classroom helps students acquire practical skills rather than just read instructions and eventually triggers emotional reactions. Through their emotional reactions, students will be able to recollect and be highly motivated to keep up with learning through their virtual reality classroom.

The collaborative learning approach supported by the computer refers to the use of computers as cognitive "artifacts" which can promote active and collaborative knowledge building. This approach focuses on the role that computers can play in student learning, from mediating face-to - face learning to providing environment for virtual learning [21]. A successful virtual learning environment can be described as an environment in which students are able to build their own knowledge, challenged to be active agents who are interdependent and perceive and experience the virtual learning environment as supporting collaborative learning [22].

3.3 Gaming Effects

On the other side of the spectrum, a number of studies emphasize on the beneficial effects of the use of games in education, especially that of interactivity [23]. In this respect, games engage students' interest and urge them to apply what they have learnt in the game context. This experiential type of learning is boosted through the use of virtual reality games that provide an elevated sense of interactivity and engagement [23]. Such technology allows three kinds of experiences; the first being experience in size. This means that virtual reality devices directly place the student in stereoscopic 3D within the environment allowing them to experience size differently than in any ordinary game [24]. This allows them to see information in their correct scaling. Secondly, transduction refers to interface devices to present information being used to present information that is difficult to perceive by the senses [24]. These include changes in the sound to indicate distance or changes in motion speed to indicate movement in different terrains. Finally, reification involves the process of representing objects or events that have no physical form into perceptible objects.

3.4 Future of Virtual Reality in Education

Unquestionably the development of technologies linked to virtual and augmented reality will be driven by the entertainment world. More efficient development and processing engines will be developed, and new interaction devices will amplify the potential use of human senses [25]. Two networks with very different characteristics are clearly identified in relation to the education sector, to concentrate on and make the most of the technologies in question. The future concerns the creation of frameworks that allow fast configuration of these environments without having to follow up with a new architecture from scratch if a new solution is to be introduced. Exploring a motor or a human body inside a theoretical description gives the consumer the ability to select and push each feature of the model at will is a very clear example. In addition to seeing the model, students will also receive all the information around each part of the model. In the field of technology, the future will be through the introduction of more senses within the experience [26].

4. Methodology – Research Part

Research in the international literature focuses on learning outcomes related to characteristics such as discomfort, user sickness, the impact of immersion on students' interest and involvement, student motivation, and the development of critical analysis [27]. The research method that will be followed is the quantitative study between subjects as it is necessary to control the results of both this application and the use of the textbook. This model is considered suitable as it provides us with flexibility in possible teaching interventions. The number of students is small (10) as the application will "run" in a

 DOI: https://doi.org/10.30564/jcsr.v3i2.2777

school with a small number of students, so the result will be classified in a single case.

The steps that need to be taken to make the research successful initially include teaching the application to the students followed by the implementation of the teaching scenario on their part [28]. The main purpose of the research is to investigate the extent to which a Virtual Reality application can help students to understand and evaluate the religious elements presented to them in comparison to the textbook. Having studied several researches that have integrated 360° applications [29-31], we came across two crucial aspects the results of which will be examined at a later stage. The questions that arise are mainly how students can better understand the data presented to them through a virtual reality application using 360° video compared to the textbook as well as their impressions and attitudes regarding the use of virtual reality in the lesson of Religion.

In order to create the courses, especially designed pamphlets and additional material was opted for. The sections of the booklet were separated using titles and symbols at the beginning of each section. They were also divided into three main parts; theoretical part, activity and material for further study. Along with the leaflets, the students were given additional material [32]. The teaching model chosen is the model of constructivism which represents reality in a variety of ways leading to consolidation through social experiences leading to discovery learning, in which students develop skills or discover various ideas and principles. According to this model in the first stage of engagement, students are given the opportunity to be involved in the learning process through open - ended, be interested in the teaching unit through small activities in order to connect previous and current knowledge. In the exploration stage, students are provided with a common basis for activities in which current concepts are identified. In the explanation stage, the students' attention is focused on specific aspects of the previous two stages and provides opportunities to prove that they have understood the views they have developed so far through the implementation of activities. In the expansion phase, students explore what they have learned thoroughly and apply any new knowledge acquired to additional activities. In the evaluation stage, the students as well as the teacher evaluate the progress and the advancement of their knowledge.

5. Creating Lessons

5.1 360° Video

Among the various fully-immersed virtual reality learning environments, 360° videos are most often used as they have wide availability and low construction costs [33]. These videos are multi-directional panoramic videos that allow the user to rotate and tilt their point of view in a continuous sphere. They can be displayed on mobile devices or other devices, such as mobile devices (Google Cardboard) or on device screens exclusively for OP (Oculus Rift).

In terms of teaching approach, the topic and objectives of learning should be identified. It should be decided on what elements or aspects of teaching the use of this video is required and then which objects and scenes to be photographed or recorded should be defined. Many studies have explored their potential with different research questions such as the involvement of students, the development of their skills [34], the ability to solve complex problems [35], the cognitive load that create in students and the effects on their interests. The use of videos seems to provide opportunities for the development of teachers' abilities through virtual experiences that examine the effectiveness of their teaching methods or for the potential training of teachers [36]. Video 360-degree allows students to watch a scene in any direction they want. This helps the student experience essentially the world captured on the 360-degree video. The picture they see moves in unison on mobile devices as the students switch and turn the screen left and right or up and down. Students navigate videos through laptops and desktops by clicking and dragging onto the image, or by running their finger across the screen on screen-sensitive devices. Students can commit to this new form of storytelling by creating their own 360-degree videos. A 360-degree video creation involves a gadget with at least two lenses, one in the front and one in the reverse. The software stitches the two videos together to create a 360 degree view [37].

In educational contexts these videos can be used to teach a concept or skill by meeting specific learning objectives which are summarised as follows: experience of dealing with unexpected events, presentation of a scenario containing details that learners can consult as many times as needed, experience of a hidden treasure scenario with a certain number of hidden objects and finally the ability to explore a more interactive environment.

5.2 Teaching Religious Affairs as School Subject using 360° and CoSpaces Edu

In this section we briefly present how a religious affairs task can be combined with the use of 360° video. For the purposes of this paper, it is not advisable to provide a detailed presentation of a lesson plan other than capture its basic axes. Finally, some screenshots of the application created with the software mentioned above are presented.

DOI: https://doi.org/10.30564/jcsr.v3i2.2777

Web-based VR authoring tools such as InstaVR, WondaVR, and CoSpaces make creating original VR artefacts with little or no programming experience feasible. This study chose CoSpaces because of its simple visual, drag-and-drop interface and built-in support for use in educational settings [1]. Virtual reality systems promote situated learning through the immersive experience of interactive objects, environments and processes [38]. CoSpaces allows for the creation of virtual 3D worlds that can be explored using smartphones, tablets, and PCs with the ability to take advantage of the VR viewer" [38,39]. CoSpaces uses a visual programming editor similar to the Scratch programming environment, to specify code for modelling and animating simple virtual worlds. In primary and secondary education, students have used CoSpaces for digital storytelling, creating virtual art exhibitions [39] and recreating historical scenes [1]. Creating VR lessons accelerates learning by allowing students to apply their own subject-matter knowledge.

CoSpaces Edu[1] is a tool to create virtual tours using 360° photos. This tool was chosen as it is considered easy to use by all age groups of students. The main goal of this lesson is to urge students to do a significant amount of research and data collection, organize the information they gathered, and decide how to best present it. Moreover, the advantages involve giving students the opportunity to discover and explore destinations outside the classroom without having to go out of the school grounds, motivate them to use new technologies constructively while enhancing creativity.

In the first section, an introduction is given and the students are asked if they have ever attended a tour of a museum, sports field, city or attraction in general, and then are presented with a virtual tour created with the specific software. In the next step, students are informed about how they should 'move' during the lesson and what the final result should be. In the second phase, we divide the students into groups of two or three people depending on the equipment of the school and direct them to look for suitable 360° type images. In addition, we ask them to write down on a piece of paper the points they consider important and write a text about them. In the last stage, the students proceed under the supervision of the teacher in the construction of the virtual tour. Finally, students can visualise their projects in VR or AR mode and see them using a VR headset.

5.3 Application

The virtual reality application was based on the CoSpaces application. All material used was uploaded to this application. 360° images were added with interaction points and connection between the scenes was made. The audio files used were processed with VLC open source software vlc. Additional functions of the application required, programmed in a script language similar to scratch. Yamanda [40] states that virtual reality applications allow users to look and / or move in any direction, but some of the students may consider that this constant operation makes them tired. This does not apply to the application as its size is such that it does not tire the students. In the final stages of development, video files were converted for better playback on mobile devices, as well as configuring build-in options for better application performance on low-end devices. Finally, the files were exported so that the application could be "run" on virtual reality devices such as Oculus Rift, Oculus Go or Google Cardboard.

The application consists of three 360° videos, each depicting an Orthodox church, a Catholic church, a mosque and a Buddhist temple. Moreover, in each video there is a figure that plays the role of the guide in each temple, conveying information about each religion and the architecture of each temple. Students can also interact with specific points marked with icons that contain additional information.

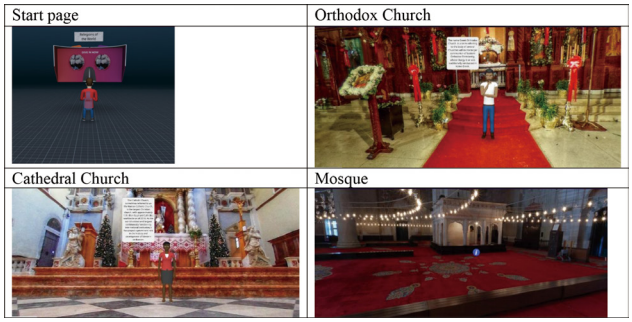


Figure 1. Application

5.4 Worksheet

Worksheets that focus on the analysis of religious - historical data are made up of forms that enable students to read and analyse documents in terms of their physical attributes and content. Studying the documents with such questions, students get the opportunity to combine their interpretation of the document with the scientific connections they've made and put it down in writing. That way we explore what students have grasped from the text and follow the mental processes they go through during the whole task of analysing it. Then, the worksheet that is to be given to the students is presented, which will be completed based on what they have been taught from the textbook but also through the information they had retrieved via the application [41].

1. https://cospaces.io/edu/

 DOI: https://doi.org/10.30564/jcsr.v3i2.2777

Table 1. Worksheet

Comparing World Religions			
	Christianity	**Islam**	**Buddhism**
Profit or Founder	Who was an important missionary that spread Christianity?	Who founded the religion of Islam?	Who founded the religion of Buddism?
Monotheistic or Polytheistic	Is the religion monotheistic or polytheistic?	Is the religion monotheistic or polytheistic?	Is the religion monotheistic or polytheistic?
Population	How many people practice Christianity?	How many people practice Islam?	How many people practice Buddhism?
Location	Where was the first Christian church located?	Where did Islam begin?	Where did Buddhism begin?
Major Beliefs	What is the main Christian belief?	Summarize the Five Pillars of Islam.	What are the most important virtues of Buddhism?
Holy Text	Which is the sacred text that Christians accept?	Which book do Muslims consider sacred?	There is a book that is considered sacred?
Place of Worship	Where do Christians worship?	Where do Muslims worship?	Where do followers of Buddhism worship?

6. Conclusions - Future Plans

Taking everything into consideration, the use of virtual or augmented reality in classrooms could be a groundbreaking feature that could shift the direction of contemporary teaching and learning. We should take into account the benefits and the potential limitations of such implementation so as to plan and set the goals and objectives for the lessons. The ultimate prospect is to engage learners' interest and make them get involved in order to acquire the knowledge that they initially regarded as incomprehensible or even unattainable.

The process of creating original VR scenes will be new for most students and will provide a useful way to apply knowledge gained by researching VR applications. Creating VR also offers a new way for students to use their mobile devices. Researching panoramic camera apps, using a 360-camera connected to a smartphone over Bluetooth, designing or running the CoSpaces app with Google Cardboard, all incorporated the use of mobile devices to this project. The entire project introduced VR as a relevant current technology, and as the goal for student-created multimedia. Students will also learn how VR scenes can change how they experience virtual worlds, as well as how they apply their knowledge gained in those virtual worlds, to the real world.

The 360^{o} videos are an original teaching method that allows students to immerse themselves in virtual learning environments of authentic 3D photos and videos [42]. These videos are more immersive than traditional ones because users have the ability to look at all the points and explore different parts of the scenes. In addition to students, their use can greatly benefit teachers in evaluating their teaching methods [43] as well as training potential teachers [44]. Regarding the lesson of religion, their use enables students to attend a place of worship exactly as it is. In addition, they have the ability to better understand the various religions and their symbols in relation to simple text reading or 2D image observation. Instructors can also create digital material for the lesson using a 360^{o} camera or use an application such as Google Expeditions.

The authors' immediate plans involve implementing the educational scenario in a primary school and in the first phase to analyse the data in detail and then statistically display them.

References

[1] Andone, D., and Frydenberg, M. Creating virtual reality in a business and technology educational context [C]. In tom Dieck Claudia, M. and Jung, T. (Eds) Augmented Reality and Virtual Reality: The Power of AR and VR for Business, Springer International Publishing, 2019; 147-159.

[2] Stojšić, I., Ivkov Džigurski A., Maričić, O., Ivanović Bibić, L., and Đukičin Vučković, S. Possible application of virtual reality in geography teaching [J]. Journal of Subject Didactics. 2016; 1(2): 83-96.

[3] Lessick, S., and Kraft, M. Facing reality: the growth of virtual reality and health sciences libraries [J]. Journal of the Medical Library Association. 2017; 105(4): 407-417.

[4] Smutny, P., Babiuch, M., and Foltynek, P. A Review of the Virtual Reality Applications in Education and Training [C]. In 20th International Carpathian Control Conference (ICCC). 2019: 1-4.

[5] Peña, J.G.V., and Tobias, G.P.A.R. Space Rift: an oculus rift solar system exploration game [J]. Philippine IT Journal. 2014; 7(1): 55-60.

[6] Moro, C., Štromberga, Z., and Stirling, A. Virtualisation devices for student learning: Comparison between desktop-based (Oculus Rift) and mobile-based (Gear VR) virtual reality in medical and health science education. Australasian Journal of Educational Technology. 2017; 33(6).

[7] Fernandez, M. Augmented virtual reality: How to improve education systems [J]. Higher Learning Research Communications. 2017; 7(1): 1-15.

[8] Falloon, G. Using avatars and virtual environments in learning: What do they have to offer? [J] British Journal of Educational Technology. 2010; 41(1): 108-122.

DOI: https://doi.org/10.30564/jcsr.v3i2.2777

[9] Curcio, I.D.D., Dipace, A., and Norlund, A. Virtual realities and education [J]. Research on Education and Media. 2016;8(2),.

[10] Hoffmann, M., Meisen, T., and Jeschke, S., Shifting virtual reality education to the next level - Experiencing remote laboratories through mixed reality [C], In Frerich S. et al. (Eds) Engineering Education 4.0. Springer, Cham, 2016; 235-249.

[11] Lindgren, R., Tscholl, M., Wang, S., and Johnson, E. Enhancing learning and engagement through embodied interaction within a mixed reality simulation [J]. Computers & Education. 2016; 95: 174-187.

[12] Duncan, I.M.M., Miller, A.H.D., and Jiang, S. A taxonomy of virtual worlds usage in education [J]. British Journal of Educational Technology. 2012; 43(6): 949-964.

[13] Gilbert, J.K., Models and modelling: Routes to more authentic science education [J]. International Journal of Science and Mathematics Education. 2004; 2: 115-130.

[14] Martín-Gutiérrez, J., Mora, C.E, Añorbe-Díaz, B., and González-Marrero A. Virtual technologies trends in education [J]. Journal of Mathematics. 2017; 13(2): 469-486.

[15] Frost, M., Goates, M.C., Cheng, S., and Johnston, J. Virtual reality: A survey of use at an academic library [J]. Information Technology and Libraries. 2020; 39(1).

[16] Siegle, D., Seeing Is Believing: Using Virtual and Augmented Reality to Enhance Student Learning [J]. Gifted Child Today. 2018; 42(1): 46-52.

[17] Mikropoulos, T.A., Chalkidis, A., Katsikis, A., and Emvalotis, A. Students' attitudes towards educational virtual environments [J]. Education and Information Technologies. 1998; 3: 137-148.

[18] Martín-Gutiérrez, J., and Fernández, M.D.M. Applying augmented reality in engineering education to improve academic performance & student motivation [J]. The International Journal of Engineering Education. 2014. 30(3): 625-635.

[19] Sin, L.H. Enhancing learning environment using augmented reality technology [C]. In Siti Aishah Hashim Ali et al. (Eds) Ice 2019 Conference Proceedings. 2019; 78-90.

[20] Trelease, R.B. From chalkboard, slides, and paper to e-learning: How computing technologies have transformed anatomical sciences education [J]. Anatomical sciences education, 2016; 9(6): 583-602.

[21] Sligte, H., Best of The Netherlands: International Computer Supported Collaborative Learning-projects in education. Old dreams and current realities. In J. Theo Bastiaens (Ed.), Proceedings of EdMedia + Innovate Learning. Amsterdam, Netherlands: Association for the Advancement of Computing in Education (AACE). 2019; 1033-1040.

[22] Zhang, B., Robb, N., Eyerman, J., and Goodman, L. Virtual worlds and gamification to increase integration of international students in higher education: An inclusive design approach [J]. International Journal of E-Learning & Distance Education, 2017; 32(2).

[23] Griffiths, M.D., The educational benefits of videogames [J]. Education and Health, 2002; 20(3): 47-51.

[24] Youngblut, C., Educational uses of virtual reality technology [M]. Institute for Defense Analyses, 1998.

[25] Inoue, S., Makino, Y., and Shinoda, H. Active touch perception produced by airborne ultrasonic haptic hologram [C]. In 2015 IEEE World Haptics Conference (WHC). 2015; 362-367.

[26] Jara, C., Candelas-Heridas, F. A., Fernández, M., and Torres, F. An augmented reality interface for training robotics through the web [J]. Communication. 2009; 189-194.

[27] Won, M., Mocerino, M., Tang, K-S, Treagust, D.F., and Tasker, R. Interactive Immersive Virtual Reality to Enhance Students' Visualisation of Complex Molecules. In Schultz M., Schmid S., Lawrie G. (Eds) Research and Practice in Chemistry Education. Springer, Singapore. 2019; 51-64.

[28] Blanco, P., Windmiller, G., Welsh, W., and Hauze, S., Lessons learned from teaching astronomy with virtual reality [C]. In G. Schultz, J. Barnes, and Linda Shore (Eds) Advancing Astronomy for All: ASP 2018 ASP Conference Series, Vol. 524, proceedings of a conference held (10-13 October 2018) 2019; 159.

[29] Duanmu, F., Mao, Y., Liu, S., Srinivasan, S., and Wang, Y. A Subjective Study of Viewer Navigation Behaviors When Watching 360-Degree Videos on Computers [C]. In 2018 IEEE International Conference on Multimedia and Expo (ICME). 1-6.

[30] Calvert, J., Abadia, R., and Tauseef, S.M., Design and Testing of a Virtual Reality Enabled Experience that Enhances Engagement and Simulates Empathy for Historical Events and Characters [C]. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 868-869.

[31] Lau, K.W., and Lee, P.Y., Exploring the Use of a Stereoscopic 360 Degree Learning Environment for Business Education [J]. International Journal of Information Education Technology. 2019; 9(2): 110-114.

[32] Gödde, M., Gabler, F., Siegmund, D., and Braun, A., Cinematic Narration in VR - Rethinking Film Conventions for 360 degrees [C]. In Chen J., Fragomeni G. (Eds) Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry.

 DOI: https://doi.org/10.30564/jcsr.v3i2.2777

VAMR 2018. Lecture Notes in Computer Science, vol 10910. Springer; 184-201.

[33] Bessa, M., Melo, M., Narciso, D.G., Barbosa, L., and Vasconcelos-Raposo, J. Does 3D 360 video enhance user's VR experience? An evaluation study [C]. In Proceedings of the XVII International Conference on Human Computer Interaction. 2016; 1-4.

[34] Sun, F.-R., Pan, L.-F., Wan, R.-G., Li. H., and Wu, S.-J. Detecting the effect of student engagement in an SVVR school-based course on higher level competence development in elementary schools by SEM [J]. Interactive Learning Environments, 2021; 29: 3-16.

[35] Wu, J., Guo, R., Wang, Z., and Zeng, R. Integrating spherical video-based virtual reality into elementary school students' scientific inquiry instruction: effects on their problem-solving performance [J]. Interactive Learning Environments, 2019.

[36] Theelen, H., van den Beemt, A., and den Brok, P. Developing preservice teachers' interpersonal knowledge with 360-degree videos in teacher education [J]. Teaching and Teacher Education, 2020; 89: (102992).

[37] Talbot, C. Using Augmented Reality to enhance teaching and learning [C]. European Library Automation Group conference in Palma, Majorca, May 2012.

[38] Greenwald, S., Kulik, A., Kunert, A. Technology and applications for collaborative learning in virtual reality. In Computer-Supported Collaborative Learning Conference, CSCL. Vol. 2 2017.

[39] Bertolini, M., Scali, F., Poletti, G., Guerreschi, A., Fontana, F., and Hohenstein, U.T. Virtual Portable Art: un percorso virtuale per le pietre incise di Riparo Tagliente [J]. Sezione di Museologia Scientifica e Naturalistica, 2018; 13: 120-122.

[40] Yamand-Rice, D., Mushteq, F., Woodgate, A, Bosmans, D., Douthwaite, A., Douthwate, I., Harris, W., Holt, R., Kleeman, D., Marsh, J., Milovidov, E., Mon Williams, M., Parry, B., Riddler, A., Robinson, P., Rodrigues, D., Thompson, S., and Whitley, S. Children and virtual reality: Emerging possibilities and challenges [R]. 2017.

[41] Ayva, O. Developing students' ability to read, understand and analyse scientific data through the use of worksheets that focus on studying historical documents [J]. Procedia-Social Behavioral Sciences, 2012; 46: 5128-5132.

[42] Huang, H.L., Hwang, G.J., and Chang, C.Y. Learning to be a writer: A spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses [J]. British Journal of Educational Technology, 2019; 51(4): 1386-1405.

[43] Walshe, N., and Driver, P. Developing reflective trainee teacher practice with 360-degree video [J]. Teaching and Teacher Education, 2019; 78: 97-105.

[44] Roche, L., and Gal-Petitfaux, N. Using 360o video in physical education teacher education [C]. In P. Resta & S. Smith (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference (3420-3425). Austin, TX, United States: Association for the Advancement of Computing in Education (AACE), 2017.

DOI: https://doi.org/10.30564/jcsr.v3i2.2777

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

REVIEW

A Review of Consensus Protocols in Permissioned Blockchains

Nenad Zoran Tomić*

University of Kragujevac, Serbia

ARTICLE INFO

Article history
Received: 26 February 2021
Accepted: 8 March 2021
Published Online: 20 April 2021

Keywords:
Permissioned blockchain
Consensus protocols
Byzantine Fault Tolerance
Crash fault tolerance

ABSTRACT

Consensus protocols are used for the distributed management of large databases in an environment without trust among participants. The choice of a specific protocol depends on the purpose and characteristics of the system itself. The subjects of the paper are consensus protocols in permissioned blockchains. The objective of this paper is to identify functional advantages and disadvantages of observed protocol. The analysis covers a total of six consensus protocols for permissioned blockchains. The following characteristics were compared: security, trust among participants, throughput and scalability. The results show that no protocol shows absolute dominance in all aspects of the comparison. Paxos and Raft are intended for systems in which there is no suspicion of unreliable users, but only the problem of a temporary shutdown. Practical Byzantine Fault Tolerance is intended for systems with a small number of nodes. Federated Byzantine Fault Tolerance shows better scalability and is more suitable for large systems, but can withstand a smaller number of malicious nodes. Proof-of-authority can withstand the largest number of malicious nodes without interfering with the functioning of the system. When choosing a consensus protocol for a blockchain application, one should take into account priority characteristics.

1. Introduction

Blockchain technology enables distributed management of large databases. Its functioning was explained for the first time in the Bitcoin cryptocurrency manifesto, in late 2008 [1]. Given its characteristics, blockchain soon came out of the shadow of cryptocurrencies and found application in the broader electronic business context. This technology allows participants to execute transactions in order to enter new data in the public ledger. A transaction is any instruction that leads to a change in the state of the system. The public ledger consists of a series of blocks, which contain records of performed transactions [2]. The content of each block depends on the content of the previously entered blocks, because the new state depends on the previous state and the changes brought by the transactions. Data is entered into the public ledger without third party mediation [3]. As there is no trust among participants, it is necessary to provide a mechanism by which the entered data will be checked and confirmed. This mechanism is called a consensus protocol.

Blockchain technology itself is new, but its foundations are previously known technologies and methods, such as asymmetric cryptography, timestamping, Merkle tree, hash functions and smart contracts. Asymmetric cryptography is used to sign executed transactions. Transactions are timestamped to avoid double spending by creating confusion about the order. Hash functions are used to

**Corresponding Author:*
Nenad Zoran Tomić,
University of Kragujevac, Serbia;
Email: ntomic@kg.ac.rs

 DOI: https://doi.org/10.30564/jcsr.v3i2.2921

prevent subsequent changes to the contents of blocks embedded in the public ledger. Therefore, the hash value of the previous block is entered in each new block, which prevents the change of their content [4]. Blockchain technology enables the implementation of smart contracts, as an electronic document that is executed on the basis of a programming code [5]. Consensus algorithms themselves are not a new technological solution. Their foundations were laid by Lamport (1978) and Schneider (1990) in the desire to formulate algorithms tolerant to a certain kind of faults [6] [7].

The subjects of the paper are consensus protocols in permissioned blockchains. The objective of this paper is to identify the functional advantages and disadvantages of observed protocols. The paper will be divided into three sections. The section one will present the key features of blockchain technology. Special attention will be addressed to the difference of the blockchain systems according to the degree of openness for participants. The section two analyzes the principles of functioning of the observed protocols individually. The final section identifies advantages and disadvantages of all protocols through a comparative analysis of their key characteristics.

2. Blockchain Characteristics

There are three types of participants in blockchain systems: nodes, full nodes, and miners. Nodes are participants that can send or receive transactions, but do not participate in consensus building nor keep a copy of the public ledger. In addition to participating in transactions, full nodes also store a copy of the public ledger. Miners are full nodes that participate in consensus building and embed new blocks in the public ledger. Generally, participants do not know each other and do not trust each other. This means that each blockchain must have a built-in protocol for reaching consensus in trustless environment where there is no third party to confirm data authenticity [8]. The protocol determines which participants can create blocks, how consensus is reached and whether there is a reward.

When performing a transaction, the sender applies the selected hash function to it and signs the resulting record using a private key. The signature authenticates the sender. Miners need to confirm the integrity of the transaction and the participant who sent it. This means that the digital signature should correspond to the sender's signature, i.e. hash value of the transaction should correspond to the one signed by the user. After confirmation, one of the miners (depending on the algorithm) packs the transactions into a block and suggests a new block to the other miners [9]. The new block contains the hash value of the previous block, the timestamp, and a list of included transactions. Other miners check whether the size of the block is within the allowed values, whether it follows the previous block according to the timestamp, as well as all hash values.

Sharing a public ledger of transactions among participants and signing transactions creates the conditions for overcoming the problem of mistrust. After reaching a consensus, the block is embedded in the public ledger, which is available to everyone and shows the current state of the system. This eliminates the need for an intermediary in transmitting and storing data [10]. Once recorded in the public ledger, transactions are irreversible. Merkle tree technology is used to connect blocks, so any attempt to change the content of a previously performed transaction leads to a change in the content of all subsequent blocks. Attempting to systemically change a series of blocks would require enormous computing power regardless of consensus protocol, an investment that can hardly be justified by benefits.

According to their openness to participants, blockchain systems are divided into permissioned blockchains and permissionless blockchains. They differ fundamentally in terms of access to the system and the role that the user can perform. Permissionless blockchains have open access. Each user can become part of the network and act as a node, full node or miner, as all roles are available [11]. Because of these characteristics, these blockchain systems are often referred to as public.

In permissioned blockchains, there is a clear separation of roles. Miners are always known and predetermined [12]. There are differences in terms of the capabilities that other users may have. For some systems, membership is open, but nodes can only send and receive transactions. For others, each user must receive a special invitation to become a node. In such systems, all users are known and identified in advance. Due to these characteristics, such systems are referred to as private or consortium blockchains in the literature. However, it should be borne in mind that higher centralization compared to permissionless blockchains should not mean that one institution is the full owner of the system. For any business application in which there is a trusted institution, it is better to use some other database technology than blockchain.

3. Types of Consensus Protocols in Permissioned Blockchains

3.1 Byzantine Fault Tolerant Protocols

Most of the protocols for reaching consensus in permissioned blockchain systems are based on solving the problems of Byzantine generals. The problem describes

DOI: https://doi.org/10.30564/jcsr.v3i2.2921

difficulties in reaching an agreement in conditions of mutual distrust among decision makers [13]. One can imagine that several divisions of Byzantine army attack the enemy city. A unilateral attack of a single division cannot lead to victory. But if the generals reach a consensus on the timing of the simultaneous attack, the city will be conquered. The problem is that generals cannot communicate directly, but solely through couriers. There may be two problems.

The first problem is the possibility that some of the generals are traitors and deliberately send contradictory messages. Another problem is the possibility for some of the couriers to change the content of the messages, either because they are traitors, or because the enemy intercepts them and replaces them with their own couriers. Therefore, it is necessary to devise a mechanism for reaching consensus, so that:

a. All loyal generals adopt the same plan (traitors, if any, can do what they want).

b. Traitors cannot lead loyal generals into adopting a wrong plan.

The problem that participants in permissioned blockchains face is similar to the problem of Byzantine generals. There are a finite number of known participants, but it is not possible to say with certainty which of them are loyal and which are traitors. Therefore, consensus algorithm must be able to allow decision-making even when some participants are unreliable. In addition to this type, crash faults also occur, when, due to technical, or problems of some other nature, the decision-making process is slowed down or stopped. Therefore, consensus algorithms in permissioned blockchains are divided into crash fault tolerant and Byzantine fault tolerant.

3.1.1 Practical Byzantine Fault Tolerance

Practical Byzantine fault tolerances was formulated by Castro & Liskov (2002) [14]. The algorithm is designed to work in asynchronous systems and to provide liveness and safety. Liveness is reflected in the fact that some consensus will certainly be reached. Safety refers to the ability to reach a valid consensus in a situation where at most $(n\text{-}1)/3$ nodes act maliciously, with n being the total number of nodes participating in the decision-making. If we denote the faulty nodes with f, then the total number of nodes must be $n = 3f+1$.

To prevent misrepresentation and confusion, each node signs messages with its own secret key. Also, each message has an authentication code, and when sent, it is compressed using the hash function. Each node communicates with all other nodes in the system. Nodes can identify each other based on the signature and check if the message was changed during transmission. Before the consensus-building process begins, the nodes are divided hierarchically, with one chosen as the leader and the others as the backup. The role of the nodes changes before each new round of decision-making on a round robin basis. One round of consensus-building consists of four phases. In the first phase, the client sends a message to the leader wanting to change the state of the system. In the second phase, the leader forwards the message to the backup nodes. Backup nodes consider the content of the message and send a response in the third phase. In the last phase, the client collects $f+1$ identical responses from the backup. The selected response represents the attitude of the entire system towards the message sent by the client.

The key advantage of pBFT in relation to all permissionless blockchains protocols is lower computational complexity, and, thus, lower electricity consumption. Also, the throughput is higher than with the mentioned systems. However, pBFT is intended for systems with a small number of participants. Increasing the number of nodes exponentially increases the volume of communication, so application in permissionless blockchains would lead to congestion. Of the known blockchain platforms, *Hyperledger Fabric* and *Zilliqa* use pBFT.

3.1.2 Delegated Byzantine Fault Tolerance

Delegated Byzantine fault tolerance (dBFT) is a modification of the basic form of pBFT (Coelho et al., 2020) [15]. It was proposed during the creation of the NEO blockchain, which, in addition to the cryptocurrency of the same name, offers a code for creating smart contracts. The GAS token is used to execute smart contracts, which users who own NEO cryptocurrency receive as a kind of dividend.

The use of dBFT overcomes the problem of excessive communication due to the increase in the number of nodes. Any user who owns a NEO can vote for one of the nodes, which then become delegates. In order for a node to become a delegate, it is necessary to positively identify itself, to have a stable internet connection and appropriate computer equipment, and to invest 1000 GAS units [16]. The speaker is then randomly selected from among the delegates. The speaker selects the transactions to be included in the new block and sends the proposal to the delegates for confirmation. The block needs to be confirmed by at least 2/3 of the delegates. Otherwise the proposal is rejected and a new random speaker is elected who repeats the process.

The dBFT protocol is criticized for the increased level of centralization. Although in theory this should not be the case, the NEO cryptocurrency has shown that all delegates are also members of the founding consortium. It

 DOI: https://doi.org/10.30564/jcsr.v3i2.2921

can be assumed that this was not the idea of the founders, but it should also be borne in mind that the selection of delegates is a great challenge. On the one hand, if no are set, participants can delegate themselves. On the other hand, setting criteria too high reduces the number of delegates and the system becomes centralized. In that case, the blockchain does not fulfill its basic decentralization premise among users who do not trust each other.

3.1.3 Federated Byzantine Fault Tolerance

Federated Byzantine fault tolerance or the Federated Byzantine Agreement (FBA) is also a modification of the basic form of pBFT. In order to overcome the problem of extensive communication, the FBA implies that nodes exchange messages only with nodes they trust [17]. The *Stellar* and *Ripple* cryptocurrencies use the FBA variants.

A set of nodes that one node trusts is called a quorum slice, while a quorum is a set of nodes that trust each other. Based on the mutual trust with other nodes, one node can be part of multiple quorums at the same time. That node represents the quorum intersection. Nodes communicate only within the quorum. Each node collects transactions performed after the last block was embedded and those performed before which so far have not become part of the blocks. The nodes then declare on the validity of the proposed transactions. When they receive the required percentage of positive votes, the transactions are included in the block, which is embedded into the public ledger.

In the Ripple protocol, a quorum is a unique node list (UNL). Making a decision implies a positive response from at least 80% of the nodes involved. This means that the maximum number of malicious nodes can be $f \leq (n - 1) / 5$, where n is the number of nodes within a particular UNL. This leads to a strong correctness of the system. In case the number of malicious nodes is higher, the system has a defense plan. As long as their number is $f \leq (4n + 1) / 5$, the system exhibits weak correctness, i.e. it is not capable of validating true transactions, but can prevent malicious nodes from embedding fake transactions [18]. Unlike the previous two protocols, which ensured finality without the possibility of forking, in FBA, the possibility of forking depends on the size of the quorum. If the size of individual quorums does not have a lower limit, then forking is possible. However, even with the lower limit of $0.2 * n_{total}$, where n_{total} is the total number of nodes in the system, forking becomes impossible if the quorums partially overlap.

3.1.4 Proof-of-authority

Proof-of-authority (PoA) is a consensus protocol based on the identification of nodes and their reputation. It was created in 2015 based on the proof-of-stake (PoS), with the proviso that in this case the nodes do not invest the coins they own, but their own reputation [9]. Nodes that embed blocks in the public ledger are called validators. In order to become a validator, node's identity must not only be publicly known, but also confirmed by the notary service. The point of the system is that in case of any undesirable behavior, the validator gains a negative reputation and loses the opportunity to participate in further block creation. Nodes with a satisfactory reputation change in the role of validator on a round robin basis.

Although PoA can also be used in permissionless blockchains, its design is extremely centralistic, because a small group of validators embed blocks. Also, PoA does not ensure anonymity, as one of the key features of a permissionless blockchain. The advantage is that, unlike PoS, it does not favor rich individuals. In practice, PoA is used on the *PoA Network* and *Vechain* trading platforms, as well as on the *Quorum* blockchain platform.

3.2 Crash Fault Tolerant Protocols

The second class of consensus protocols deals with the problems of the imperfect environment in which systems operate. A metaphorical representation of the problems addressed by these protocols is the fictitious local assembly on the Greek island of Paxos. Deputies constantly enter the assembly hall and leave, while staying in the building for a period of time that is not foreseen in advance, nor is it the same for each deputy. Despite the occasional absence of deputies, it is necessary to provide conditions for unhindered legislative decisions. Thus, crash fault tolerant protocols ensure functioning in situations where individual participants are unavailable during a certain period. The assumption is that there are no participants who are malicious, i.e. that Byzantine faults are not a problem that the system encounters.

3.2.1 Paxos

Paxos is not an individual protocol, but a whole family of protocols, created with the aim of solving the problem of absence of participants [19]. The basic system assumptions are as follows: processors perform operations at any arbitrary speed; during work they may experience a crash fault; after recovery, processors can rejoin the protocol; during operation, processors do not attempt to trick the system in any way. In terms of communication, one processor can send messages to any other processor, which takes an arbitrarily long time to be delivered. Messages can be lost, but they are not modified by third parties.

DOI: https://doi.org/10.30564/jcsr.v3i2.2921

Consensus can be reached when the total number of nodes is $n = 2F + 1$, where F is the maximum number of nodes that can experience a simultaneous crash fault. In other words, most of the total number of nodes must always be present.

Nodes are called acceptors in Paxos protocols. They are divided into quorums, as subsets of a set of acceptors, with any two quorums having to be intersected by at least one acceptor [20]. The size of the quorum is such that it includes the majority of the total number of acceptors (and in the event that the acceptors are assigned a certain voting weight, the quorum must include more than half of the votes). The client sends a message requesting a change of state from the system. In order for a decision to be made, all quorum acceptors must receive the same message. The proposer coordinates the distribution of messages and decision-making (especially in case of disagreement). Each system has one prominent proposer, called a leader. Finally, learners execute the decisions made (change the state of the system) and send a response to the client.

When a proposer receives a request from a client, they formulate a message, assign it a certain number p, and distribute it to the acceptors within the quorum. The number p serves only to determine the chronological order of the message and must be greater than the numbers of all messages that the acceptors have processed so far. This activity is called *Prepare*. Upon receipt, the acceptors check the number p and if that number is greater than the numbers of all previous messages, they send a positive response. This process is called Promise. If they have accepted a message in the previous period, they shall also submit to the proposer the number of that message q (where $q < p$) and the value of w accepted by that message. Since Paxos is intended for asynchronous systems, participants may be at different stages of the process, so it is possible that the proposer does not know that the message q is accepted at all. In case the number of the message p is less than the number of any of the previous messages, the acceptors can send a negative response to the proposer, or not respond at all. This concludes the first phase.

If it receives a sufficient number of positive responses, the proposer in the second phase submits to the acceptors the full content of the message p, which, in addition to the number, also contains the proposed value $v = x$. In case the acceptors inform them that the message q has been accepted in the meantime, the proposer can decide to modify the proposal before sending, and for the proposed value to be $v = y$. A request for verification and acceptance is sent to the acceptors with the message. Acceptors accept the message (p, v) if in the meantime they have not promised to consider only the message under the number r; $r > p$. [21]. When they accept the message, they inform the proposer and the learners (although the role of the learner is often played by the proposer itself).

This case is a general form of the Paxos algorithm. As it is a family of protocols, this means that the general case has a large number of modifications into specific forms. Individual cases vary significantly in terms of purpose, and, therefore, have different performance, which further complicates the comparison of this protocol with other ones.

3.2.2 Raft

Raft is a consensus protocol with resilience to crash faults, proposed by Ongaro & Ousterhout (2014) [22]. It is a modification of the Paxos algorithm, in which nodes can exist in one of three states: leader, followers or candidates. Time is divided into arbitrarily long sections, called terms. During one term, the same node always plays the role of the leader. They receive requests for transactions from clients, check them and index them, in order to maintain insight into chronological order. The transactions are then packed in blocks, which are forwarded to followers. The task of the followers is to send the leader an acknowledgment of the correctness of the block and to replicate the block, i.e. to take over the information it contains.

When requests for new entries stop arriving, the node waits for a certain period, known as an election timeout. This means that the old leader has lost its role and that elections for a leader for the next term are taking place. The node becomes a candidate, gives itself a vote and sends messages to the rest of the network asking for support for its candidacy. If it gets the support of most other nodes, it will be elected leader in the coming term. It then sends a message to all other nodes informing them of its election. Also, it can happen that, while waiting for votes, the node receives a message from another node claiming to have been elected leader. In that case, the node compares its term index with the term index of the node that claims to be elected. If its term index is lower, it must recognize the other node as the leader. However, if its term index is higher, it can reject the message of the node claiming to be the leader and continue to collect votes. A higher term index value indicates that this node is aware of the last changes that have occurred. If no candidate collects enough votes, the whole process is repeated with a randomly assigned short pause. The node assigned the shortest pause will initiate a new vote.

The key difference between Paxos and Raft algorithms is that Raft allows only the best updated nodes to take on the role of the leader, while with the Paxos algorithm it can be any node [23]. The well-known blockchain plat-

 DOI: https://doi.org/10.30564/jcsr.v3i2.2921

forms, *R3 Corda* and *Quorum*, use Raft as a consensus protocol.

4. Comparing Protocols

Permissionless blockchains are easy to compare. They all have the same purpose, but show significant differences in terms of required investments, throughput and scalability. In addition, a large number of papers compare these protocols in some respect. The situation with permissioned blockchains is somewhat more complex. First of all, the systems are divided into two large groups, which differ according to the level of security they provide. Differences in purpose undoubtedly lead to differences in performance. However, the problem that exists in the literature is the uneven terminology in this area. Some papers that generally compare consensus protocols for permissioned blockchains actually compare the performance of the platforms applying them [24].

Since none of the observed protocols are based on intensive computation, no high initial investment is necessary. Protocols are not competitive as proof-of-work and related protocols are, so there is no energy load. They still differ a lot in scalability and throughput. Also, there is a significant difference in terms of mutual trust between nodes. Having in mind the representative research in this field, the comparison was made on the basis of the following characteristics: security, mutual trust, throughput and scalability. Table 1 presents the results of the comparison.

Table 1. Comparative analysis of the presented consensus protocols in permissioned blockchains

Protocols Characteristics	pBFT	dBFT	FBA	PoA	Paxos	Raft
Security	Byzantine if f < 33.3%	Byzantine if f < 33.3%	Byzantine if f < 20%	Byzantine if f < 49%	Only from crash fault	Only from crash fault
Mutual trust	Based on node selection	Nodes choose who to trust	Flexible trust	Based on identity	Complete in terms of good intentions	Complete in terms of good intentions
Throughput	Moderate	High	High	Low	Moderate	Moderate
Scalability	Limited	High	High	Low	Limited	Limited

Source: author, according to the broad literature

The first four presented protocols protect the system from Byzantine faults. They differ from each other in the proportion of malicious nodes that the system can tolerate without losing normal functioning. Paxos and Raft assume that all nodes have already been checked and known, and that they are not characterized by malicious behavior. Therefore, they do not provide protection against Byzantine faults, but only from a temporary crash fault. Thus, the level of security they guarantee is significantly lower.

Nodes do not know each other in permissionless blockchains, so there is no basis for trust. With permissioned blockchains, the nodes do not have to know each other either, but the very fact that the membership is permissioned gives some degree of trust. Thus, with pBFT, mutual trust is based on the fact that the nodes have been allowed to enter the system. That is why each node communicates with all other nodes, because it trusts their decisions. With dBFT and FBA, the situation is somewhat different, so the nodes choose who they trust. This is especially pronounced with FBA and the so-called flexible trust, meaning that, regardless of the total number of nodes, the participants make a decision only with those who they trust. With PoA, trust is based on a previously irrevocably established identity and the responsibility it entails. With the Paxos and Raft protocols, nodes have complete mutual trust in terms of goodwill, but they do not trust each other in terms of prompt response to tasks.

One of the key problems of permissionless blockchains is that, due to protocol design in a trustless environment, they cannot achieve throughput at the level of payment card companies. In order for a single protocol to be considered an adequate basis for a financial or other business system, it must allow the processing of a large number of transactions. All protocols analyzed in this paper show significantly high throughput, but there are differences between them. Thus, with pBFT, on a sample of 10,000 transactions on the Hyperledger Fabric platform, a throughput of about 200 transactions per second (TPS) was proven [25]. dBFT is said to be able to support around 4000 TPS [26], though specific data on the NEO cryptocurrency blockchain shows that this number is 4 times lower. Also, the data shows throughput of about 4000 TPS with FBA [27], though theoretical assumptions claim that this number could be over 10,000. PoA has a slightly lower throughput, of about 80 TPS.

There are large differences in terms of scalability of the observed protocols. It was shown that a pBFT network with 40 nodes takes only 4 seconds to confirm a block of 10,000 transactions, but that with an increase in the number of nodes to 200, than time increases to 26 seconds and continues to grow exponentially [28]. From the above it can be concluded that pBFT is a good solution for small and controlled systems, but that due to the need for all nodes to communicate with each other it is not a good choice for large systems. It was stated that the NEO blockchain can process a block within 20 seconds, when the number of nodes ranges from 7 to 1024 [26]. Thus, with the growth of

DOI: https://doi.org/10.30564/jcsr.v3i2.2921

the system, dBFT becomes a better choice than pBFT, but it also has scalability limits. FBA is designed to maintain high scalability despite the involvement of thousands of nodes. However, this has not been confirmed in practice, because the Ripple and Stellar cryptocurrency blockchains have only 130 and 136 nodes, respectively. At the same time, they maintain a very stable scalability. PoA has a stable scalability which means 5-8 seconds required to validate the block even in case of increasing the number of nodes to over 1000. The problem that Paxos and Raft face in this context is the need for all communication to go through one node (leader), which leads to a bottleneck.

5. Conclusions

Comparing consensus protocol performance has three limitations. First of all, although all the observed protocols are intended for permissioned blockchains, not all have the same purpose. The clearest clasification is into those protocols that protect the system from Byzantine faults and those that protect only from crash faults. However, there are differences within these groups as well. Thus, pBFT is intended for systems with a small number of nodes, while FBA is intended for systems with a large number of nodes. If performed outside the intended environment, these systems can show significant deviations in terms of performance. Comparison under conditions that favor one group or one separate protocol yields results that do not reflect their actual usability.

Furthermore, observed protocols were not tested evenly in practice. It is worth mentioning that dBFT and PoA actually have limited application, and that their performance should therefore not be taken as completely reliable. Some protocols can be analyzed in parallel on the examples of a larger number of cryptocurrencies and/or business platforms, while others have an application that is sufficient only for conditional results. Finally, most protocols in practice show deviations from the theoretically stated optimal results. This problem is related to the previous one and speaks of insufficient testing in different conditions and on different platforms.

Having in mind the above and the results of the comparison, it can be concluded that no protocol shows absolute dominance in all aspects of the comparison. When choosing a consensus protocol for a blockchain application, one should take into account priority characteristics.

References

[1] Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system [R]. 2008, retrieved from https://bitcoin.org/bitcoin.pdf.

[2] Zheng, Z., Xie, S., Dai, H.N., Chen, X. & Wang, H. An overview of blockchain technology: architecture, consensus and future trends [C], IEEE 6th international congress on Big data, Honolulu, HI, 2017: 557-564 https://doi: 10.1109/BigDataCongress.2017.85.

[3] Belotti, M., Božić, N., Pujolle, G. & Secci,S. A vademecum on blockchain technologies: when, which, and how [J], IEEE Communication, Surveys & Tutorials, 2019, 21(4): 3796-3838.

[4] Oliveira, M.T., Reis, L.H.A., Medeiros, D.S.V., Carrano, R.C., Olabarriaga, S.D. & Mattos, D.M.F. Blockchain reputation-based consensus: A scalable and resilient mechanism for distributed mistrusting applications [J], Computer Networks, 2020,179: 107367. https://doi.org/10.1016/j.comnet.2020.107367.

[5] Szabo, N. Formalizing and Securing Relationships on Public Networks [J]. First Monday, 1997, 2(9). https://doi.org/10.5210/fm.v2i9.548.

[6] Lamport, L. Time, Clocks and the Ordering of Events in a Distributed System [J]. Communications of the ACM, 1978, 21(7): 558-565.

[7] Schneider, F. Implementing Fault-Tolerant Services Using the State Machine Approach: A Tutorial [J]. ACM Computing Surveys, 1990, 22(4): 299-319.

[8] Bamakan, S.M.H., Motavali, A. & Bondarti, A.B. A survey of blockchain consensus algorithms performance evaluation criteria [J], Expert Systems with Applications, 2020, 154: 113385. https://doi.org/10.1016/j.eswa.2020.113385.

[9] Ismail, L. & Materwala, H. A Review of Blockchain Architecture and Consensus Protocols: Use Cases, Challenges, and Solutions [J], Symmetry 2019, 2019, 11: 1198. https://doi:10.3390/sym11101198.

[10] Zheng, Z., Xie, S., Dai, H.N. & Wang, H. Blockchain challenges and opportunities: A survey [J]. International Journal of Web and Grid Services, 2018, 14(4): 352-375.

[11] Lin, I.C. & Liao, T.C. A survey of blockchain security issues and challenges [J], International Journal of Network Security, 2017, 19(5): 653-659. https://doi:10.6633/IJNS.201709.19(5).01.

[12] Wang, X., Zha, X., Ni, W., Liu, R.P., Guo, Y.J., Niu, X. & Zheng, K.. Survey on blockchain for internet of things [J], Computer Communication, 2019, 136: 10-29. https://doi.org/10.1016/j.comcom.2019.01.006.

[13] Lamport, L., Shostak, R., & Pease, M. The Byzantine Generals Problem [J]. ACM Transactions on Programming Languages and Systems, 1982, 4(3): 382-401. https://doi.org/10.1145/357172.357176.

[14] Castro, M., Liskov, B. Practical Byzantine Fault Tolerance and Proactive Recovery [J]. ACM Transac-

tions on Computer Systems, 2002, 20 (4): 398-461. https://doi:10.1145/571637.571640.

[15] Coelho, I. M., Coelho, V. N., Araujo, R. P., Yong, Q. W. & Rhodes, B. D. Challenges of PBFT-Inspired Consensus for Blockchain and Enhancements over Neo dBFT. Future Internet, 2020, 12(8): 129. https://doi:10.3390/fi12080129.

[16] Manoppo, M. Delegated Byzantine Fault Tolerance Consensus Mechanism [J], Medium, 2018, June 15.

[17] Yoo, J., Jung, Y., Shin, D., Bae, M. & Jee, E. Formal Modeling and Verification of a Federated Byzantine Agreement Algorithm for Blockchain Platforms [C], 2019 IEEE International Workshop on Blockchain Oriented Software Engineerin, 2019,. https://doi.org/10.1109/IWBOSE.2019.8666514.

[18] Schwartz, D., Youngs, N. & Britto, A. The Ripple Protocol Consensus Algorithm [R], 2018, retrieved from: https://ripple.com/files/ripple_consensus_whitepaper.pdf.

[19] Lamport, L. The part-time parliament [J], ACM Transactions on Computer Systems. 1998, 16(2): 133-169. https://doi:10.1145/279227.279229.

[20] García-Pérez Á., Gotsman A., Meshman Y. & Sergey I. Paxos Consensus, Deconstructed and Abstracted [M]. In: Ahmed A. (ed.) Programming Languages and Systems. ESOP 2018. Lecture Notes in Computer Science, 2018, 10801: 912-939, https://doi.org/10.1007/978-3-319-89884-1_32.

[21] Lamport, L. Paxos made simple [J], ACM SIGACT News. 2001, 32(4): 51-58.

[22] Ongaro, D. & Ousterhout, J. In Search of an Understandable Consensus Algorithm [C]. Proceedings of the 2014 USENIX Annual Technical Conference (USENIX ATC 14), Philadelphia, PA, USA, 2014, 305-319.

[23] Howard, H. & Mortier, R. Paxos vs Raft: Have we reached consensus on distributed consensus? [C], 7th Workshop on Principles and Practice of Consistency for Distributed Data (PaPoC '20), April 27, 2020, Heraklion, Greece, 2020, https://doi.org/10.1145/3380787.3393681.

[24] Polge, J., Robert, J. & Le Traon, Y. Permissioned blockchain frameworks in the industry: A comparison [J], ICT Express, in press, 2020, https://doi.org/10.1016/j.icte.2020.09.002.

[25] Nasir, Q., Qasse, I.A., Abu Talib, M. & Nassif, A.B. Performance analysis of hyperledger fabric platforms [J], Security and Communication Networks, 2018, vol. 2018. https://doi.org/10.1155/2018/3976093.

[26] Xian, M. NEO White paper [R], 2018, retrieved from: https://github.com/neo-project/docs/blob/3c5530197768d98a1abf7eeed0119a8f4e99e7cc/enus/whitepaper.md.

[27] McCaleb, J., Crain, B.F., Couture, S. & Roy, M. Soundcloud [R], 2017, retrieved from https://soundcloud.com/epicenterbitcoin/eb-128.

[28] Jalalzai, M.M., Busch, C. & Richard, G.G. Proteus: A scalable bft consensus protocol for blockchains [C], 2019 IEEE International Conference on Blockchain, 2019, 308-313.

DOI: https://doi.org/10.30564/jcsr.v3i2.2921

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

ARTICLE

Voting System Based on Blockchain

Zihan Guo Xiang He* Peiyan Zou

North China University of Science and Technology, Tangshan, China

ARTICLE INFO

Article history
Received: 13 January 2021
Accepted: 1 February 2021
Published Online: 18 April 2021

Keywords:
Blockchain
Voting system
Decentralization
Data cannot be tampered with

ABSTRACT

Online ballot box system has the advantages of high efficiency and environmental protection, but the existing network voting technology still has a lot of matter. Almost all electronic voting system could be proved to be intrusion. The administrator of the system could tamper with the data for benefit, and the system may be attacked by hackers. The safety and fairness of the existing network voting system depend entirely on the safety and credibility of the website itself, but these cannot guarantee the fairness of voting. Make full use of blockchain technology, so that voting, even if there are malicious participants, but also to ensure the correctness and safety of the vote. The introduction of block chain technology, block chain has decentralized, data tampering and other characteristics. P2P network is applied in the block chain layer to construct a distributed database, digital signature algorithm and encryption technology are used to ensure that the data cannot be tampered with, consensus network algorithm is used to ensure the consistency of the data in the network, and timestamp technology is applied to save the data blocks in a chain structure connected end to end. It paper focuses on the implementation of P2P network networking mode, node block synchronization, data and block verification mechanism and consensus mechanism to ensure data consistency in the network layer of block chain layer. Using time stamp, Merkle tree, asymmetric encryption and other technologies to design data blocks and use chain structure to store data blocks. Combined with the characteristics of blockchain, a fair and transparent voting system is constructed. Model aims to apply the block chain technology to the voting scenario and design a secure block chain voting architecture. It system is designed and developed based on the block chain system. It makes full use of its decentralization, removes the dependence of electronic voting on trusted third parties, and protects the privacy of voters and candidates. Data cannot be tampered with. Once the data are stored in the block chain, it cannot be tampered with. It provides a real and credible database.

1. Introduction

In order to indicate the origin of voting with blockchain matter, the following background is worth mentioning.

1.1 Discussion of Voting Issues

On the surface a lot of people might think voting is an easy thing to do, just cast your vote, count up the results and publish it. But it is not as simple as you think. Voters

**Corresponding Author:*
Xiang He,
Electronic information engineering, North China University of Science and Technology, Tangshan, China;
Email: 2216839099@qq.com

 DOI: https://doi.org/10.30564/jcsr.v3i2.2797

must cast only one vote. We must prevent people from voting twice. With the rapid development of network technology, the traditional questionnaire survey method has fallen behind. It has a complicated voting procedure and takes a long time to complete the user's vote. The task of counting votes is very heavy, and it is easy to make statistical mistakes and is easy to be manipulated. Modern society is the network information age, and the use of network technology could improve people's work efficiency, save the cost of human and material resources, promote the development of society [5].

1.2 The Dilemma of Voting in the United States

With a crucial election looming in the United States, little progress has been made in ensuring the integrity of the voting system. Existing voting systems leave plenty of room for doubt: it is possible to mimic voters in the first place (though surveys have found this to be a negligible proportion in the us);Postal ballots could be changed or stolen; Election officials may miscount; And almost all electronic voting machines have proven to be breakable.

1.3 Blockchain Technology Solve the Matter

However, the booming blockchain technology in recent years may well resist the corruption of the authorities and hackers [1]. Blockchain is a new distributed infrastructure and computing paradigm, which USES ordered chain data structure to store data, USES consensus algorithm to update data, and USES cryptography technology to guarantee data safety [11,12,13,14]. Came up with the block chain to provide the basis for the emergence of a new electronic voting system, voting system based on block chain could be ruled out the possibility of manipulation, the decentralized, distributed network structure is suitable for the voting system, voting centers do not need special maintenance and management of sorting system and network, to ensure the transparency of the network, also to prevent the malicious vote or tamper with the vote fraud, cheat cheating, its anonymity, some voters' personal information could be hidden, protect personal privacy [3]. Voters could also verify and track their votes.

1.4 The Matter Requirements

1) How to use block chain technology to construct an underlying design or a set of algorithms to solve the matter of online voting?

2) Evaluate the possible matter of blockchain technology in solving voting matter and try to improve.

3) Cybersafe and voting experts agree that blockchain is unnecessarily complex and no more secure than other network voting. Can blockchain technologies be combined with other technologies to reduce complexity and improve safety?

2. The Description of the Matter

2.1 For the Use of Blockchain Technology to Construct an Underlying Design or a set of Algorithms to Solve the Matter of Online Voting

To meet the requirements of individuals, enterprises, institutions and governments, design a fair, fair and transparent voting system. The voting system has the following three sub-indicators:

1)Ensure that voting data cannot be usurped;

2) The voting data could be traced and verified;

3)Vote anonymously to ensure the privacy of voters;

Blockchain is decentralized. Data cannot be tampered with, and it is safe and reliable. Once the data is stored in the blockchain, it cannot be tampered with. Blockchain is a real and reliable database that cannot be tampered with. Therefore, blockchain technology is the best solution to guarantee the fairness and safety of the voting system.

2.2 Analysis of Blockchain Technology in Solving Voting Matter

Blockchain technology is communal in many fields because it is distributed and untamable. Despite the advantages of the technical principle, it still requirements to be developed by human beings. Since it is a human software product, it is inevitable that there will be safety holes, which may ultimately make products using blockchain technology vulnerable. From a legal perspective, identification verification is difficult, time-consuming and often inaccurate, voters are sometimes wrongly singled out as ineligible to vote, and unqualified people are sometimes allowed to vote and repeat votes occur. Recounts (and recount requests) are common because ballots are counted inefficiently and conflicting agreements make auditing slow and difficult. Lawsuits proliferate when candidates seek an advantage. Whenever this happens, public confidence in the electoral system and results will lead over time to a lack of confidence in elected officials and governments, which may lead to lower voter turnout in the future. Voters' feelings, prejudices, confusion, misunderstandings and fears about the new voting technology are likely to be inflamed and used for political gain. It will create more confusion and mistrust, slowing the procedure needed to construct the necessary legal support around the newer voting forms. It runs counter to the direction we need to move towards a safe and more accurate electoral system.

In addition, other challenges, such as the current block-

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

chain technology does not well support the required speed and breadth, are also a question of whether a national blockchain election could be held. Also, blockchain is only as good or bad as its ecosystem, inputting wrong information and blockchain will do a great job storing it! While blockchains are immutable, the voting tools voters might use may not be. Voting from smartphones sounds great, but smartphones are one of the least secure electronic devices on the market. Someone could change your vote before the phone is sent to the blockchain. Therefore, there are still many matters and drawbacks in applying blockchain technology to solve the voting matter, and we could only construct the model under some ideal conditions [6].

2.3 How to Reduce the Complexity of Blockchain Technology and Improve the Safety through the Combination of Other Technologies

The electronic voting procedure includes: the voting initiator initiates the voting, the voting certification registration center completes the distribution and screening of the votes, the voters who are distributed the votes conduct the voting, after the voting is completed, the counting center completes the counting of the votes and the publication of the results. As voting results often affect personal interests, participants in online electronic voting may become dishonest due to the matter of being falsely recognized, coerce, bribe-taking, and existence of internal ghosts, etc. These dishonest participants may tamper, falsify, replay and deny voting data, thus affecting the correctness of voting results. In particular, the third party may have a ghost, so that its credibility is difficult to be guaranteed. In the whole voting system, only the voting initiator is credible and honest when he/she initiates the voting. However, through the introduction of blockchain technology, the decentralized registration and certification center is realized, the dependence of electronic voting on trusted third parties is removed, and the privacy of voters and candidates is protected. At the same time, the project provides an electronic voting system scheme based on verifiable secret sharing and other technologies, which could effectively guarantee the safety of electronic voting. Blockchain technology is one of the research hotspots in the field of network safety and fintech. It is a kind of point-to-point decentralized data sharing technology, and the data stored in it do not need to rely on a trusted third party. In other words, the data stored in it cannot be tampered with, forged, refuted, impersonated, and reproduced, so as to effectively guarantee the data safety and automatically form a credit relationship between nodes [2]. Up to now, blockchain technology has successfully protected the safety of more than $100 billion of digital assets such as bitcoin, and its safety has been fully verified [10]. If the blockchain technology could be effectively utilized, it may not need to rely on trusted third parties to complete safety authentication [4]. Based on the advantages of blockchain technology, if blockchain technology could be combined with artificial intelligence, big data and other technologies, the complexity of blockchain technology could be greatly reduced.

3. Models

In order to achieve the decentralization of the blockchain, the data cannot be tamper-proof, safety and credibility requirements, the blockchain layer will use the P2P network to construct a distributed database, using digital signature algorithms and encryption technology to ensure data non-tamperable, consensus algorithm to ensure the network The consistency of the data, and the use of timestamp technology to save the data block in a chain structure connected end to end As shown in Figure 1, the blockchain layer is divided into a network layer and a storage layer [11,12,13,14]. The network layer realizes the construction of the network, and the verification mechanism and the consensus mechanism ensure the safety and consistency of the data. The design focus is on constructing the P2P network, realizing the verification mechanism and the consensus mechanism. The storage layer encapsulates the data block, and the data block is stored in a chain of end to end. The design focus is to construct the data block structure and the chain of the block by using timestamp, hash function, Merkle tree, asymmetric encryption and other techniques to storage. The blockchain layer will be based on maven, with Java as the development language.

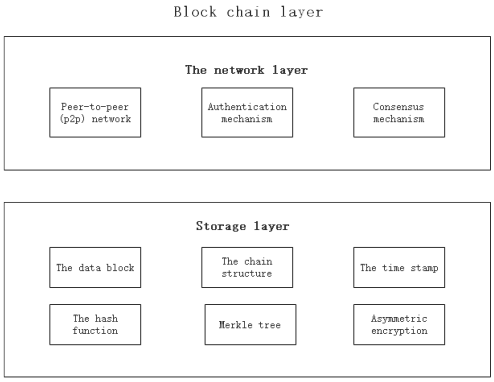


Figure 1. Block chain

3.1 Block Chain Layer Design

3.1.1 Node P2P Network Construction

The blockchain system is a distributed, decentralized system with nodes that are peer-to-peer, autonomous, and free to join. Therefore, this design uses a peer-to-peer network as the networking mode. P2P networks have the ad-

 DOI: https://doi.org/10.30564/jcsr.v3i2.2797

vantages of decentralization, scalability, decentralization, and robustness. They could organize nodes participating in data check and accounting, so that the system could run stably under decentralization. There is no central node in the P2P network, and the nodes are interconnected through a flat topology. Each node has the same functionality and provides network services. Each node has the functions of discovering new nodes, synchronizing blocks, applying layer network routing, verifying block data, and propagating block data.

Constructing a P2P network is the initialization procedure of the blockchain layer. If a new node joins the P2P network for the first time, the IP addresses of other nodes in the network are required. DNS seed A DNS server that provides the IP address of a node on a P2P network to help discover nodes. Therefore, the DNS seed method is used to join the P2P network, and the TCP protocol is used, and the port 8333 is used. The joining procedure is as follows.

1）Connection seed node

2）Receive node IP address list

3）The nodes in the connection list

4）When one or more connections are established, the node sends its own IP address to its neighbors. Neighboring nodes will forward IP addresses, allowing more nodes to receive IP addresses, ensuring a more stable connection.

The specific flow chart of constructing a P2P network is as follows.

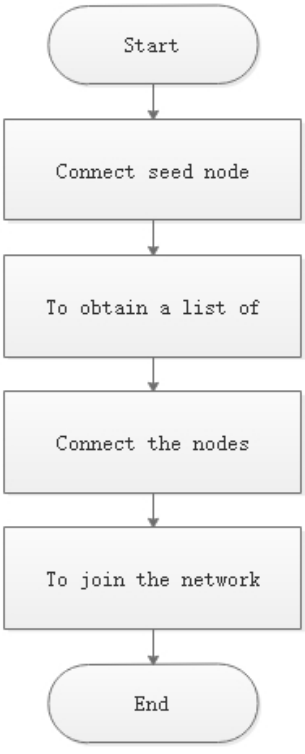


Figure 2. P2P Network flow chart

3.1.2 Block Synchronization

When a node joins a P2P network, there is no blockchain data or the data is incomplete. You need to request data blocks from other nodes. It procedure is called block synchronization and is the initialization procedure of the blockchain layer. The node will first load and verify the local data block first, which is divided into block verification and Block data verification. First verify whether the previous block is on the main chain, and after completing the block verification, perform block data verification. After verifying the local data block, the node performs block synchronization.

The block synchronization procedure is as follows:

1) Request data block information from the node that has established the connection, determine whether the node times out when it is timed, and request other nodes if it times out;

2) The requested node will first send the block height, and the requesting node compares the received block height with the local block, such as the height of the block, and issues a request to acquire the data block;

3) Verify the validity of the data block;

4) Continue to send the synchronization request until the node with the highest block height is found, and the procedure of requesting the data block synchronization is as shown in the figure.

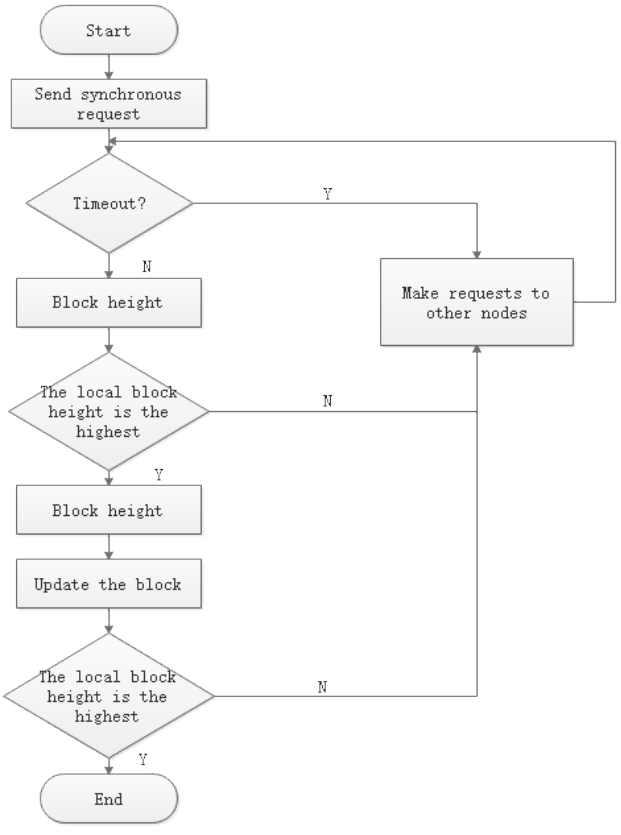


Figure 3. Block synchronization flow chart

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

3.1.3 Data Verification Mechanism

The verification mechanism guarantees that the data cannot be tampered with. Each node in the P2P network continuously receives data. After receiving the data, the node will verify the validity of the data at the first time. The node verifies the data structure and digital signature, and only the data that meets all the conditions is valid. If the data is invalid, it is discarded. If the data is valid, the valid data is stored in the data pool. The node that obtains the billing rights broadcasts the block to the entire network, and other nodes verify the validity of the block, including the random number in the block header, the timestamp, and the data in the block body. If the block is valid, it is stored in the blockchain. Otherwise, discard.

It design uses an elliptic curve digital signature algorithm (ECDSA) [14]. The private key is d and the public key is (E, P, n, Q).

1) Choose a random integer k, between [1, n-1]

2) Calculate kP = (x1, y1) and r = x1 mod n. If r = O, skip to step 1. Otherwise, continue to the next step.

3) Calculate s = k-1 {h(m) + dr} mod n (h is the hash algorithm). If s = O, skip to step l. Otherwise, continue to the next step.

4) Signature information m is Cr, s)

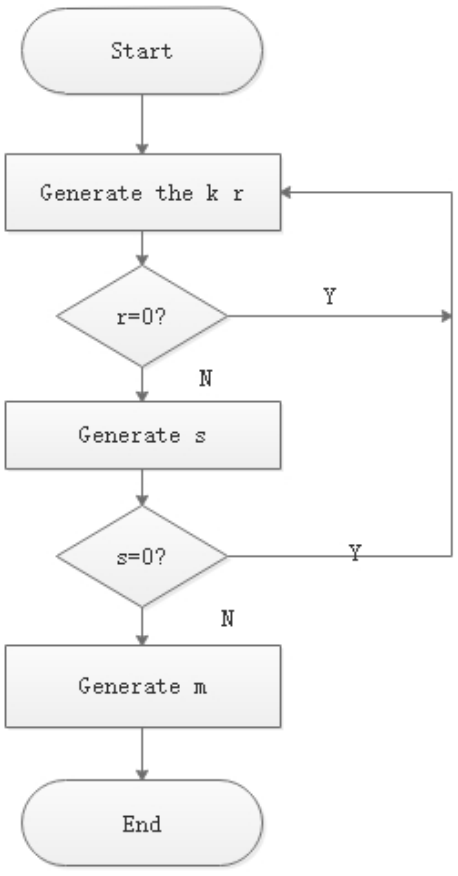


Figure 4. Generate signature flow chart

1) Get the public key (E, P, n, Q)

2) Verify that r and s are integers and are in the interval [1, n-1]

3) Calculate w = s-1 mod n and h(m)

4) Calculate ul = h(m)w mod n and u2 = rw mod n

5) Calculate u1P + u2Q = (x_0, y_0) and v = x0 mod n

6) Signature verification if and only if V = f

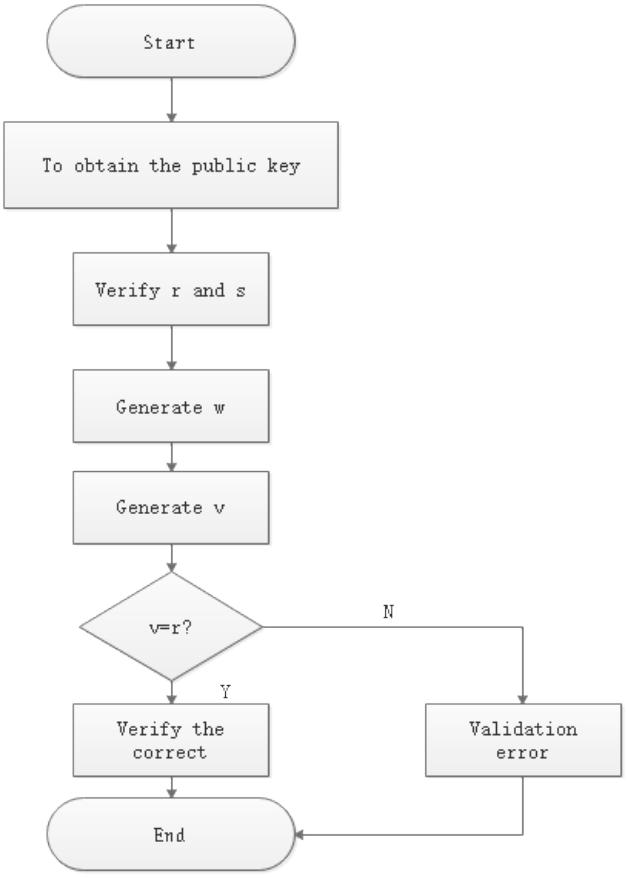


Figure 5. Verify signature flow chart

The core code verification core code is as follows:

public static Boolean verify (byte [] data, ECDSA Signature, byte [] pub) {// Digital signature verification

if (FAKE_SIGNATURES)

return true;

if (Secp256kl Context.is Enabled ()) {

try {return NativeSecp256kl.verify(data, signature.encodeToDER(), pub);}

catch (NativeSecp256klUtil.AssertFailException e) {

return false;}}

ECDSASigner signer= new ECDSASigner ();

ECPublicKeyParameters params = new

ECPublicKeyParameters (CURVE.getCurve().decodePoint(pub), CURVE);

signer.init (false, params);

try {return signer.verifySignature(data, signature.r, signature.s); }

 DOI: https://doi.org/10.30564/jcsr.v3i2.2797

```
catch (NullPointerException e) {
signatures. Those signaturesthread. log.error("Caught
NPE inside bouncy castle", e);
return false;}}
```

3.1.4 Consensus Mechanism

The consensus mechanism provides guarantees for the consistency of data in the blockchain and is the key to maintaining the blockchain's operation. The general model of the consensus mechanism in the computer field is: in a reliable distributed system with channels, how could the system ensure that other nodes are not affected by malicious nodes and could reach a correct consensus on a certain matter in the case of a malicious node? The entire system operates reliably and reliably.

The consensus mechanism in the blockchain is embodied in that when a node collects a certain amount of valid data, multiple nodes package the data into blocks, and how the system assigns the accounting rights in the case that the node may be attacked. Which node reaches a consensus and allows the blockchain to run reliably and reliably. A well-behaved consensus algorithm could select the appropriate node. The node broadcasts its packaged block data to the whole network. After other nodes verify the validity, the block could be stored in the blockchain.

The workload proof mechanism (Proof Of Work, Pow) is simple, easy to implement, and fault tolerant (Allow 50% of nodes in the network to be attacked). It design uses the workload proof mechanism (Proof of Work, Pow). Before the node packs the data into blocks, it requirements to find a random number so that the hash value of each element of the block header is not greater than the target hash (the target hash is generated by a specific algorithm), this raises the threshold for packing blocks. The first node that finds the conditional random number will obtain the accounting rights of the block and broadcast it to the whole network. After most nodes verify the validity, it will be stored in the blockchain. The stronger the computing power, the greater the probability of finding the first random number.

The random number search procedure of the Pow consensus mechanism is as follows:

1) Node verifies data finiteness and stores valid data in the data pool;

2) Calculate the Merkle root of the block data pool, and replenish the block header data, set the random number to zero Add a random number to 1;

3) Calculate the double SHA256 hash of the current block header. If the hash value is not large, find a random number that satisfies the condition, package the data into blocks, and broadcast to the entire network; otherwise, repeat this step. Until you find a random number that satisfies the condition;

4) If a random number that satisfies the condition is not found for a period of time, update the data pool and timestamp, and then calculate the Merkle root and continue to find the random number.

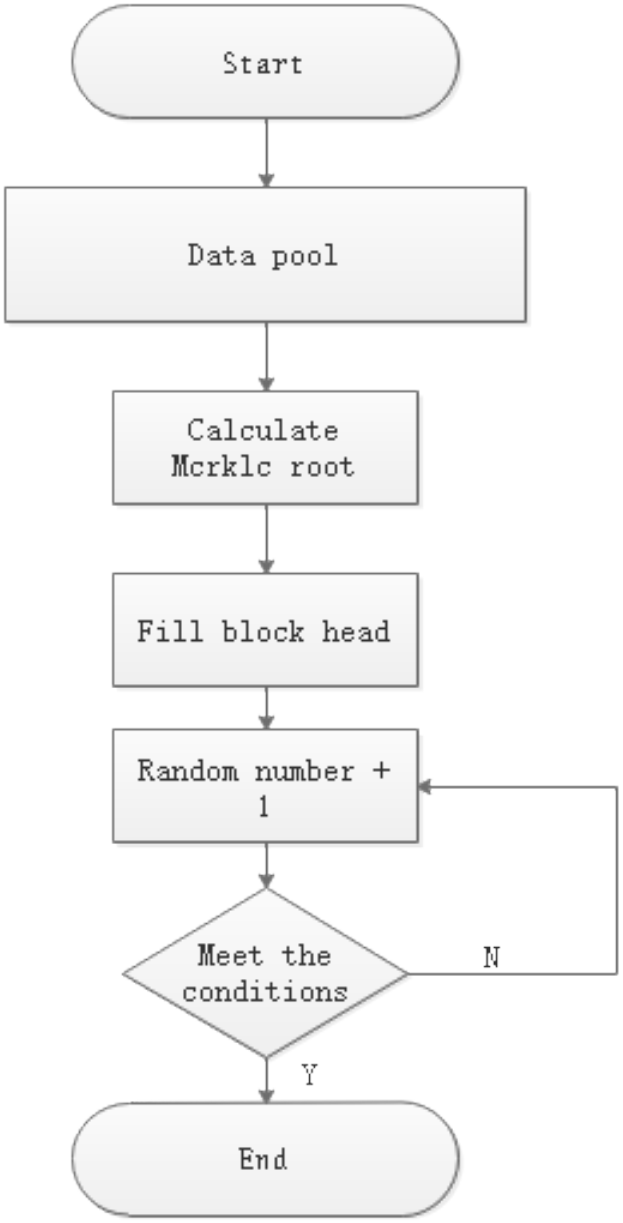


Figure 6. Random number search flow chart

3.2 Blockchain Storage Layer Design

3.2.1 Data Blocks

In order to realize the chain structure storage of the data base time stamp and to quickly verify the validity of the data, the data block adopts the structure as shown in the figure, and each data block is divided into a block header

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

and a block body. The version number is included in the block header, the previous block hash value, timestamp, random number, this block target hash value (Bits) and Merkle root. The valid data and corresponding quantities generated during the block creation procedure is saved in the block body. Valid data is generated by the hash of the Merkle tree to produce a unique Merkle root, which is stored in the block header.

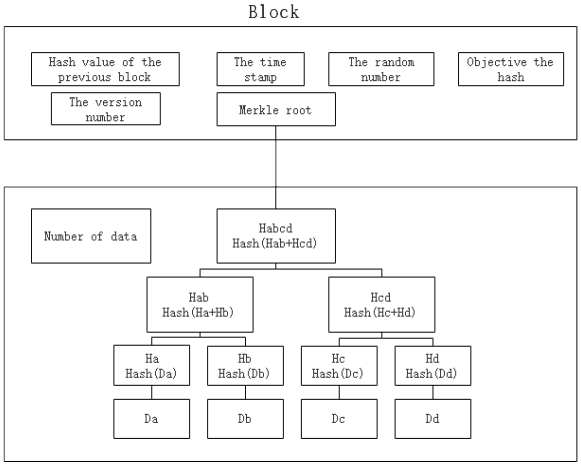


Figure 7. Block structure diagram

The data block corresponding class should contain relevant information, and could implement functions such as creating new blocks, adding valid data, verifying data, and verifying the Merkle root.

3.2.2 Timestamp

Using time stamping techniques in digital information, you could verify the exact time at which digital information is generated. Time stamping techniques provide reliable verification of the integrity and presence of digital information over a certain period of time or before a certain point in time.

In order to prove the existence of data in the blockchain, it is necessary to know the exact time of data block generation. Therefore, the node that first finds the random number and obtains the accounting right must add the timestamp to the block header, and record the data block write. The exact time of the blockchain. By using the timestamp technique, data blocks arranged in chronological order are formed in the blockchain. The time stamp technology itself is not a new technology, but its application in the blockchain technology could be used as a proof of the existence of data blocks, forming a basis for realizing the irreversible change of blockchain data.

3.2.3 Hash Function

Hash functions are characterized by unidirectionality, timing, fixed length, and randomness, making them ideal for validating data. The system will use digital signature algorithm and encryption technology to ensure that the data cannot be tampered. The selection of the hash function will affect the safety of the system. SHA256 and RIPEMD160 are based on MD4's improved hash function, which is more complex than the MD4 and MD5 algorithms and therefore more secure. It design will use two hash functions, SHA256 and RIPEMD160. Among them, SHA256 is used more, with thousands of block header hash values, block data, blockchain address generation, etc., while RIPEMD160 uses only 1000 to generate blockchain addresses. The blockchain layer will use the SHA256 hash function to procedure the data, that is, the SHA256 hash function is used twice to generate a 256-bit (32-byte binary value).

3.2.4 Merkle Tree

When performing block verification, if the data in the current block are verified one by one, the efficiency is very low. To do this, you could use Merkle Tree to quickly verify the block data. The Merkle Tree is a Hash Tree that verifies the integrity of block data in a short period of time, ensuring that data in the blockchain network are not lost and modified, and that nodes are not sending fake blocks. The Merkle tree usually consists of block data, the Merkle root, and all subtrees from the block data to the Merkle root. The procedure of generating a Merkle Tree is as follows:

1) Block data generates a hash value under the action of a hash function;

2) Combine two adjacent hashes into one string and generate a hash value under the hash function;

3) It recursion until only the last Merkle root remains.

Merkle Tree is mainly used to verify the integrity of block data in a short time. The Merkle root is the only feature of all leaf node values (block data). As long as you verify that the Merkle roots are equal, you could know if the block data have changed. If the block has N block data, the algorithm complexity of locating the tamper data is only logN.

The core code for generating the Merkle Tree is as follows:

```
private List<byte []> constructMerkleTree()
        {ArrayList<byte[]> tree = new ArrayList<>O;
        for (Transaction t: transactions)
        {tree.add(t.getHash().getBytes());}
        int levelOffset = O; // Offset
        for (int levelSize = transactions.size(); levelSize >
1; levelSize = (levelSize + 1) / 2)
            {for (int left= O; left< levelSize; left+= 2)}//
```

When it is singular, the left child and the right child are the same

```
int right= Math.min(left + 1, levelSize - 1);
byte[] leftBytes = Utils.reverseBytes(tree.get(levelOffset + left));
byte[] rightBytes = Utils.reverseBytes(tree.get(levelOffset + right));
tree.add(Utils.reverseBytes(hashTwice(leftBytes, 0, 32, rightBytes, 0, 32))); }
levelOffset+= levelSize;}
return tree;}
```

3.2.5 Chain Structure

In order to provide traceability and verification of blockchain data, blocks could be stored in a chained structure of Figure 8. The current block contains the hash value of the previous block. If the node does not know the hash value of the previous block, it cannot generate a new block. All data blocks in the blockchain are chained into a chain by block hash values, and the longest chain (main chain) is always stored in the blockchain, from the creation block to the newly generated block. When a new block is stored in the blockchain, it will be linked behind the main chain.

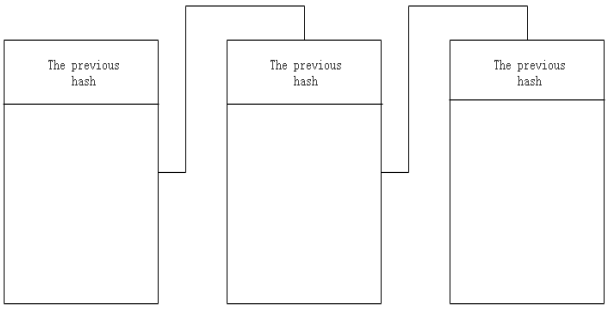


Figure 8. Chain structure diagram

The blockchain stores the data blocks in a chain structure on a blockchain with timestamps in the data blocks. It increases the time dimension of the data and makes it easy to trace and validate the data. With the previous block hash value, you could locate the previous block and verify that the previous block was modified. Through the "block ten chain structure", the blockchain could detect the tampering of any data in time. The blockchain provides a time-series, recordable record that could be viewed as a database that is not tamper-proof and authentic.

3.2.6 Asymmetric Encryption

For data safety and ownership verification, the system uses asymmetric encryption. Asymmetric encryption technology will play an important role in application scenarios such as digital signature and information encryption at the blockchain layer. In the digital signature scenario, the sender in the blockchain encrypts the information with its own private key, and sends the public key together with the private key to the blockchain network. The node decrypts the information with the public key, thereby verifying the information. ownership. In an information encryption scenario, the sender in the blockchain will encrypt the information with the recipient's public key, and the recipient decrypts the information with the private key. RSA relies on thousands of prime factorizations, so it is impossible to theoretically measure the confidentiality of RSA. The RSA key is too long, and the large number operation causes the encryption and decryption procedure to be slow. The elliptic curve encryption algorithm provides a shorter key than RSA, which is characterized by higher safety and faster procedureing. Therefore, this design will use an elliptic curve encryption algorithm, and the mechanism for generating asymmetric cryptographic public and private keys is shown in the figure.

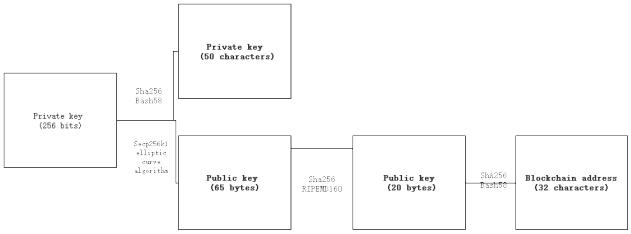


Figure 9. Asymmetric encryption

The procedure of generating the public and private key and the blockchain address is as follows:

1) Generate a 256-bit binary-shaped random number as the private key. The number of private keys will be as high as 2^{256}, making it impossible to traverse the private key. It could be generated by calling the random number generator of the base operating system.

2) For ease of memory, use the SHA256 hash function and Base58 encoding to convert a 256-bit binary private key into a private key of 50 characters, such as 6KXl j299sB bbwGnAbz3FPUDJ47r5 z2SCC Td7ytuX2QN-PyrnWjss

3) A public key of 65 bytes in length is generated by a private key of 256-bit binary form by the Secp256kl elliptic curve algorithm. Such as 075ab939e3aa7a4a9aa03c5bdf22ca2dea249366c0fa-88dec78aclcc2c335a72245a306218bee3c692fc540b-5277c02552b8b5626db790526e4109c4e5934b

4) Use the SHA25 hash function and the RIPEMD 160 hash function to compress a 65-byte public key into a 20-byte hash.

5) For ease of memory, a hash value of 20 bytes is

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

converted to a 33-character blockchain address using the SHA256 hash function and Base58 encoding. Such as 1HwgZoSa4ffzDijk5eaExNo6aKkZJ xpkw.

It is important to note that the procedure of generating a public key from a private key is irreversible, that is, the private key cannot be reversed by the public key. The procedure of generating a blockchain address by a public key is also irreversible. The role of the blockchain address is that the blockchain address could be used for reception when the user does not want to expose their public key.

4. Conclusions

4.1 Conclusions of the Matter

Blockchain has the advantages of decentralized, secure and reliable, and data cannot be tampered with. It paper studies the principle of block chain and use of P2P network technology, the consensus algorithm, timestamp technique, the chain structure, the digital signature algorithm and the encryption technology to construct the underlying block chain, on the basis of general block chain technique is applied to voting scenario, implements a data safety of the voting system, to ensure its vote fair, fair and credible [8].

4.2 Methods Used in Our Models

1) Hash function

Hash function is a kind of mathematical function, which is widely used in computer science and cryptography. In the field of cryptography, Hash functions could compress any length of input into a short, fixed output, known as a Hash value. Hash function has the characteristics of underactivity, timing, length and randomness. Cryptography has a wide range of applications, through a single hash function to compress arbitrary length of digital information into a fixed length digital digest. Digital digest is also known as digital fingerprint. Digital abstract has the characteristics of fixed length, same information abstract and different information abstract. Therefore, the digital digest could be used to verify whether the data have changed, that is, data integrity.

2) Asymmetric encryption

Encryption and decryption use the same secret key. Encryption method is called symmetric encryption. The encryption method is called asymmetric encryption. The secret keys of asymmetric encryption are produced in pairs. The public secret key is called public key, while the private key is called private key. If you encrypt the information with a private key, you could only decrypt the information using a public key. If the message is encrypted with a public key, only the private key could be used to decrypt the message. The safety of asymmetric encryption is guaranteed by the algorithm and the secret key. The strength of the algorithm is high and the secret key is highly confidential (unlike symmetric encryption, which requires the exchange of secret keys), so the safety of asymmetric encryption algorithm is more guaranteed. However, the speed of asymmetric encryption algorithm is relatively slow, which is not suitable for scenarios that need to encrypt halo data, including digital signature or secret key negotiation. Common asymmetric encryption algorithm RSA, ECC (elliptic curve encryption algorithm) knapsack algorithm, Ellamae, Rabin, Diffie - Hellman [. RSA is dependent on the quality of large number factorization, how to measure the confidentiality of the RSA in theory. The RSA key is too long, a large operation and lead to the encrypted the decryption procedure is slow. Elliptic curve encryption algorithm, and secret key provided by the shorter than RSA, has the characteristics of higher safety and faster procuring speed.

3) Digital signature

Digital signature is a combination of digital digest technology and asymmetric encryption technology, which provides a strong guarantee for the integrity of digital information and the authenticity of sender's identity.

4) Timestamp technology

Timestamp technology is the application of authoritative time source and digital signature technology. Using the timestamp technique in digital information could verify the exact time when the digital information was generated. Timestamp technology could provide reliable verification of the integrity and existence of digital information in a certain time period or before a certain time point.

5) Merkle tree

Merkle tree is a Hash tree, usually a binary tree. The hash value of Merkle tree leaf node data, the non-leaf node is the hash value of two adjacent hash values merged into a string, and the Merkle root is finally generated. Merkle tree could verify the integrity of the data in a short time, that is, if the data have been changed. The Merkle root is the unique characteristic of all leaf nodes. By verifying that the Merkle root is equal, you could know whether the data of the leaf node has been changed. In a distributed system, Merkle tree could quickly verify whether the data changes during transmission, greatly reducing the computational complexity.

6) P2P network

Peer-to- Peer (P2P) [25l, is a network formed in the application layer.

 DOI: https://doi.org/10.30564/jcsr.v3i2.2797

P2P network is different from client/server model. There is no concept of client or server in P2P network. There is no central node, only peer nodes. Each node seeks both the service and the provider of the service. A node in a P2P network acts as both a client and a server. There are no limits on the number, scope, time, or space of nodes in a P2P network, and each node is free to join or leave.

All nodes in P2P network share the pressure of server in traditional way. The more nodes join the network, the more stable the whole network will be, providing higher quality services. P2P networks return power to users and decentralization.

7) Distributed storage

Traditional storage systems generally store data centrally in centralized storage servers, but the storage resources of centralized storage servers are very limited and cannot meet the requirements of storing large-scale data. But the distributed storage system adopts the expandable system structure, and stores the data in several nodes, so that the capacity, stability and expansibility of the system are greatly improved.

8) Consensus mechanism

In any decentralized distributed system, the nodes of the participating system are equal in status and lack of trusted central nodes. When decisions are divided, the matter of how the nodes could reach consensus arises. Consensus on transactions is one of the core challenges of distributed systems.

9) Blockchain

There are many definitions of blockchain. It is difficult to intuitively understand the real meaning of blockchain from the definition alone. We could first understand the general meaning of blockchain from the perspective of application.

Traditional network storage system could adopt centralized storage or distributed storage to store data. Distributed storage system could solve the matter of limited space of centralized storage system. However, no matter centralized storage or distributed storage system, such as the storage system is attacked by hackers or improper right management, it will cause the modification or loss of data. Blockchain is a real and reliable database that cannot be tampered with. Once the data are stored in the database, it cannot be tampered with. Blockchain is a distributed database based on P2P network. There is no central node, and each node in the P2P network stores exactly the same data. Any malicious modification of data by any node will not affect the correctness of the entire network data. One of the important features of blockchain is that data cannot be tampered with.

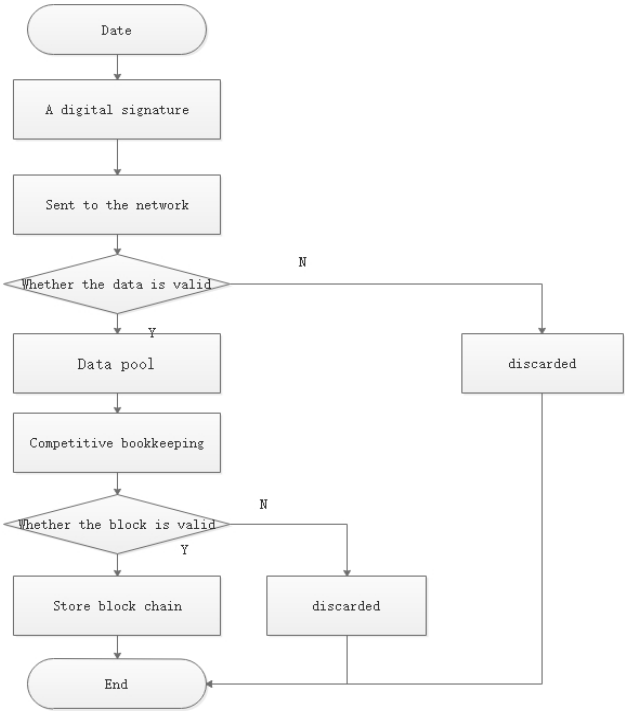


Figure 10. Blockchain data storage flow chart

4.3 Application of Our Models

On the basis of completing the design of the underlying blockchain, the application layer adopts the BIS architecture, and the web application layer mainly implements the system functions. According to the analysis of the system function, the registration module, the homepage module, the voting module, the new voting module, the voting result query module, and the voting history query module are designed. The web application layer will be developed using the framework of Spring + Spring Boot + My bits, and the database is MySQL.

The application layer adopts the MVC (Model View Controller) design pattern and is divided into a view layer, a business logic layer, and a data access layer. As shown.

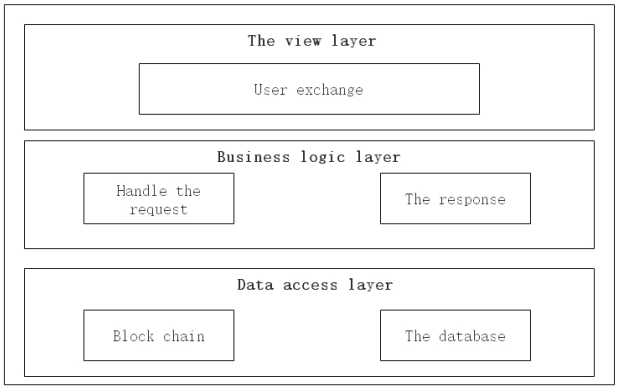


Figure 11. Web application layer architecture

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

1) View layer: The view layer provides the ability to interact with the user. A good user experience is an important indicator of system design. It design uses the Free Marker template engine, which uses a "template + data = output" model with no business logic. The template is only responsible for populating the data in the page, and finally the combination of the template and the data is presented in front of the user. Separate the view layer from the business logic to improve development efficiency. The front-end framework uses bootstrap, which provides comprehensive documentation, rich components, and ease of use to quickly and easily customize the pages you need.

2) Business logic layer: The business logic layer is the key to implementing system functions, procuring requests from the view layer and returning responses. For example, store the information entered by the user into the database. The ajax technology enables the view layer and business logic to asynchronously interact with data, providing a better user experience.

3) Data access layer: The data access layer is responsible for data access. Normal data (such as user name, password, etc.) are directly stored in the database, and blockchain data (voting records) are stored in the blockchain. When accessing normal data, you need to design the database and write sq. statements. When accessing the blockchain data, the underlying blockchain interface requirements to be called to implement access in the blockchain of data.

5. Future Work

5.1 Model Advantage

Voting is the preferred democratic expression in any situation where multiple people are required to make decisions. But when the voting system rises to the national level, its disadvantages gradually appear. Where power struggles exist, fraud and corruption are hard to root out. In addition, in the United States and other western countries with a large population and a large number of votes, the existing paper voting method has disadvantages such as low efficiency, high cost and low transparency. Blockchain offers voters an updated system to address these issues.

Since the emergence of blockchain technology, many people believe that it has the potential to reform the voting system because of its features such as non-tampering and high transparency. Citizens could vote through smart devices and record the data on the blockchain, which cannot be tampered with, so as to ensure safety and save time and capital cost. The characteristics of distributed and distrust could guarantee the personal privacy of voters and the full realization of public opinion [7].

5.2 Limitations of the Model in the Voting Domain

First, safety. Skepticism about the nascent blockchain technology is also based on distrust of electronic voting systems. Voting through blockchain still requires electronic devices, which poses the risk of hacking and affects confidentiality and fairness.

Second, unnecessary. Blockchain technology does have big advantages in voting, but these advantages are not irreplaceable. Opponents of blockchain voting still insist that paper ballots have irreplaceable advantages, and that blockchain technology's advantages in voting could be replaced by other ways.

Third, high costs. Blockchain technology is still in the early stage of implementation. If blockchain technology really wants to completely overturn the voting method, it will need a lot of promotion costs.

References

[1] Yuan Yong, Wang Feiyue. Current situation and Prospect of blockchain technology[[J]. Acta Automatics Sinica,2016,42(4):481-494.

[2] Swan M. Blockchain: Blueprint for a New Economy. USA:O'Reilly Media Inc., 2015.

[3] Ding Wei. Block chain based instrument data management system. China Instrumentation, 2015, (10): 15-17.

[4] Zhao He, Li Xiao-Feng, ZhanLi-Kui, Wu Zhong-Cheng. Data integrity protection method for microorganism sampling robots based on blockchain technology. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2015, 43(Zl): 216-219.

[5] Swan M.Blockchain thinking: the brain as a decentralized autonomous corporation.IEEE Technology and Society Magazine, 2015, 34(4): 41-52.

[6] Sarr Idrissa, Naacke Gueye Ibrahima. Blockchain-based model for social transactions procedureing[C]. DATA 2015-4th International Conference on Data Management Technologies and Applieations. 2015:309-315.

[7] H.Watanabe, S. Fujimura, A. Nakadaira, etc. Blockchain contract: Securing a blockchain applied to smart contracts[C]. IEEE International Conference on Consumer Electronics (ICCE)□2016:467-468.

[8] Zhu Yan, Gan Guohua, Deng Di and other key technologies of blockchain security research [J]. Infor-

 DOI: https://doi.org/10.30564/jcsr.v3i2.2797

mation security research□ 2016,(12):1090-1097.
[9] Shen Xin, Pei Qingqi, Liu Xuefeng. Overview of blockchain technology [J]. Journal of network and information security， 2016,(11) :11-20.
[10] Jia Liping. Theory, practice and influence of bitcoin [J] international financial research, 2013, (12) :14-25.
[11] National Institute of Standards and Technology (NIST), Secure hash standard. Federal Information Procuring Standards Publication (FIPS PUB)180, May 1993.
[12] National Institute of Standards and Technology (NIST) Computer Systems Laboratory, Secure hash standard. Federal Information Procuring Standards Publication (FIPS PUB)180-1, April 1995.
[13] National Institute of Standards and Technology (NIST) Computer Systems Laboratory, Secure Hash Standard. Federal Information Procuring Standards Publication (FIPS PUB)180-2, August 2002. http://csrc.nist.gov/publications/fips/fips 180-2/flips 180-2.pdf.
[14] Rivets R. The MD4 message digest algorithm. In Advances in Cryptology, Crypto'90 volume 537 of LNCS, pages 303-311.Springer-Verlag, 1991.

Technical Report

We focus on the design of P2P network design, node block synchronization, data and block verification mechanism and ensure the consistency of data consensus mechanism. Then the time stamp, Merkle tree and asymmetric encryption are used to design the data block.

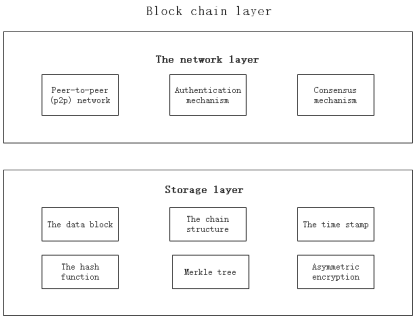


Figure 12. Blockchain layer

In order to satisfy the system centralization, the information cannot be tampered. Open and transparent requirements, combined with the underlying block chain to store data characteristics. We divide the system into view layers. Business logic layer and data access layer. Combined with the block chain characteristics of the business layer architecture design.

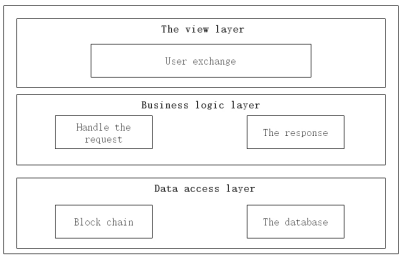


Figure 13. Application layer architecture design

A data secure voting system is designed by taking full advantage of the decentralized blockchain, untamable data, and secure and reliable data.

DOI: https://doi.org/10.30564/jcsr.v3i2.2797

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Computer Science Research
https://ojs.bilpublishing.com/index.php/jcsr

ARTICLE

Computerized FDTD Method for Longitudinal Optical Phonon Energy on Semiconductor Hybrid Structure for High Power Devices Fabrication

Phyo Sandar Win[1] Hsu Myat Tin Swe[1] Hla Myo Tun[2*]

1.Department of Electronic Engineering, Technological University (Taungoo), Bago, Myanmar
2.Faculty of Electrical and Computer Engineering, Yangon Technological University, Yangon, Myanmar

ARTICLE INFO

Article history
Received: 25 April 2021
Accepted: 17 May 2021
Published Online: 19 May 2021

Keywords:
FDTD
Semiconductor structure
Computer simulation
Computer programming
MATLAB

ABSTRACT

The research problem in this study is the longitudinal optical phonon energy on metal/semiconductor interface for high performance semiconductor device. The research solution is to make the software model with finite difference time domain (FDTD) solution for transmission and reflection pulse between metal and semiconductor interface for carrier dynamics effects. The objective of this study is to find the quantum mechanics understanding on interface engineering for fabricating the high performance device for future semiconductor technology development. The analysis was carried out with the help of MATLAB. The quantum mechanical spatial field on metal-semiconductor stripe structure has been analyzed by FDTD techniques. This emission reveals a characteristic polar radiation distribution of electric dipoles and a wavelength independent of the structure size or the direction of emission; consequently, it is attributed to thermally generate electric dipoles resonating with the longitudinal optical phonon energy. Phonon energy occurs lattice vibration of material by the polarization of light, if the material has rigid structure reflect back the incident light. So, high reflective metal-semiconductor structure always use as photodectors devices in optical fiber communication. No lattice vibration material structure has no phonon effect, so this structure based devices can get high performance any other structure based devices. The transmission and reflection coefficient of metal-semiconductor GaN/Au layer structure compare with GaN/Ti and GaN/Pt structure. Parallel (P) and transverse (S) polarization of light incident on a metal-semiconductor nanolayer structure with IR wavelength. Efficient use of the layer by layer (LbL) method to fabricate nanofilms requires meeting certain conditions and limitations that were revealed in the course of research on model systems.

1. Introduction

The FDTD method is one of the most widely used methods in electromagnetic simulation. Finite-difference time-domainor Yee's method is a numerical analysis technique used for modeling computational electrodynamics and it has recently been applied to the simulation of the Schrödinger equation. In principle, the approaches can be divided into three groups: (i) crystal defects (nitrogen vacancies, nitrogen interstitials, gallium vacancies, etc.)

**Corresponding Author:*
Hla Myo Tun,
Faculty of Electrical and Computer Engineering, Yangon Technological University, Yangon, Myanmar;
Email: hlamyotun@ytu.edu.mm

 DOI: https://doi.org/10.30564/jcsr.v3i2.3139

mediating magnetic ordering over the crystal from one magnetic center to another; (ii) presence of free charge carriers in the materials and co-doping with donors and (iii) coupling of the separated magnetic moments with magnetically neutral impurities (e.g., oxygen) which results in overall magnetic order in the scheme [1].

Recently, semiconductor researches are more attractive than other research fields. Semiconductor devices employ the charge of the carriers to achieve the desired functionality. Since it is a time-domain method, FDTD solutions can cover a wide frequency range with a single simulation run, and treat nonlinear material properties in a natural way. This technique allows for the simulation of laser excitation dynamics as well as for the determination of energy eigenstates [2].

Semiconductor devices have attracted great attention due to the quantization of carriers in three dimensions, leading to discrete spectra. Among other things, they present the possibility of studying the details interaction of particles in a controlled environment. The advent of measurement techniques, such as single-electron capacitance spectroscopy (SECS), has made possible determination of the energy of individual particles. The finite-difference time-domain method is arguably the simplest both conceptually and in terms of implementation of the full-wave techniques used to solve problems in electromagnetic. The FDTD method requires the discretization of time and space.

2. Materials and Methods

2.1 Implementation of Absorption by Fresnel Equation

The samples are undoped metal/semiconductor stripe structure. In this metal-semiconductor layer stripe structure, materials, thickness, refractive index (n) and extinction coefficient (k) values are shown in simulation results respectively.

In this paper, p-polarized light is incident on metal layer and then is compared with the reflectivity values for three types of metal-semiconductor structure to get the less phonon effect high performance layer structure.

The Fresnel equations describe the reflection and transmission of light when light is incident on an interface between different optical media. When light travelling in a denser medium strikes the surface of a less dense medium (that is, $n_1 > n_2$) beyond a particular incidence angle known as the critical angle, all light is reflected and $R_s = R_p = 1$.

P-polarized light:

$$r_p = \frac{n_i \cos(\theta_t) - n_t \cos(\cos\theta_i)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)} \quad \text{Equation (1)}$$

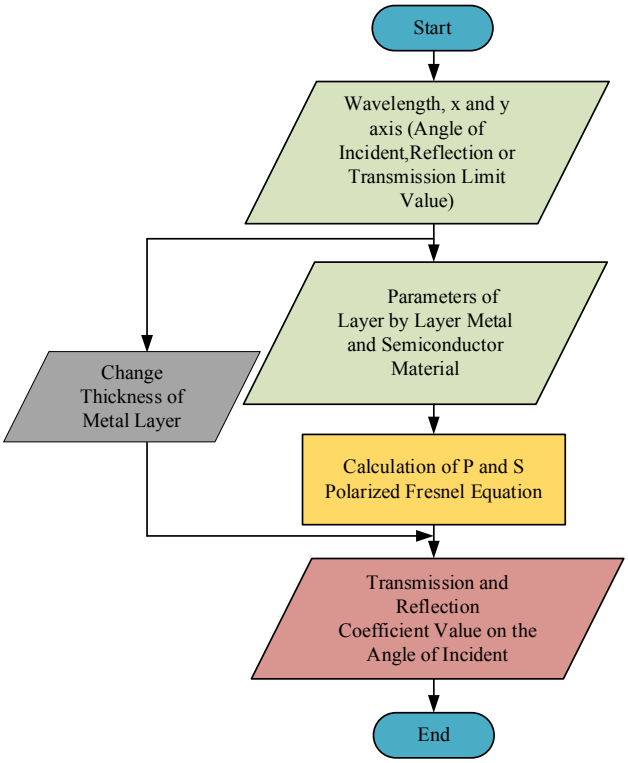


Figure 1. Flowchart of Transmission and Reflection Coefficient

$$t_p = \frac{2n_i cos(\theta_i)}{n_i cos(\theta_t) + n_t cos(\theta_i)} \quad \text{Equation (2)}$$

S-polarized light:

$$r_s = \frac{n_i cos(\theta_i) - n_t cos(\theta_t)}{n_i cos(\theta_i) + n_t cos(\theta_t)} \quad \text{Equation (3)}$$

$$t_s = \frac{2n_i cos(\theta_i)}{n_i cos(\theta_i) + n_t cos(\theta_t)} \quad \text{Equation (4)}$$

For both polarization,

$$n_i sin(\theta_i) = n_t sin(\theta_t) \quad \text{Equation (5)}$$

nall is a software for the simulation of surface plasmon resonance curves, transmission and reflection coefficient curves by polarizations based on the Fresnel formalism. Absorption value can be obtained by substration the value of reflection and transmission value from one.

2.2 Implementation of Reflection Analysis on FDTD Measurement

Figure 2 illustrates the flowchart of reflection analysis

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

between metal and semiconductor interface with MUR and PML boundary. At first, the boundary condition of MUR and PML with suitable courant factor is initialized. After that, the permittivity, permeability, speed of light and wavelength of free space parameter and specific time steps for FDTD is declared. The physical parameters for metal layer and semiconductor layers are specified. The two-dimensional finite difference time domain (2D FDTD) algorithm is initialized.

And then the boundary condition of MUR and PML is calculated. Finally, the reflection pluses for boundary condition of MUR and PML are displayed.

2.3 Implementation of Transmission Analysis on FDTD Measurement

Figure 3 illustrates the flowchart of transmission analysis between metal and semiconductor interface with MUR and PML boundary.

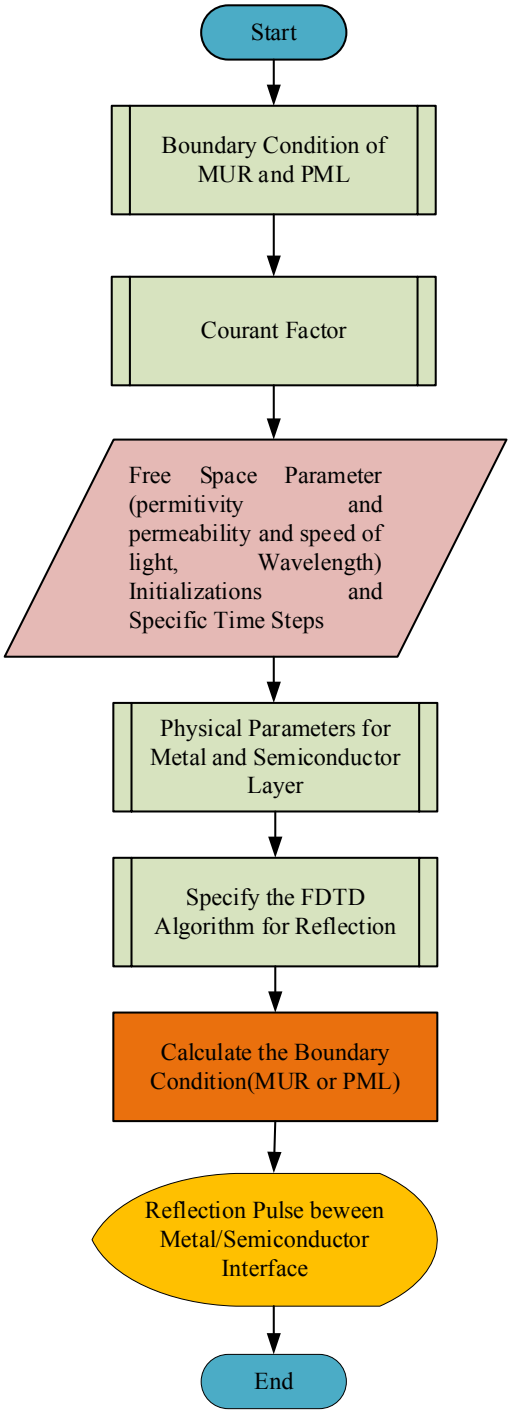


Figure 2. Flowchart for Reflection Analysis

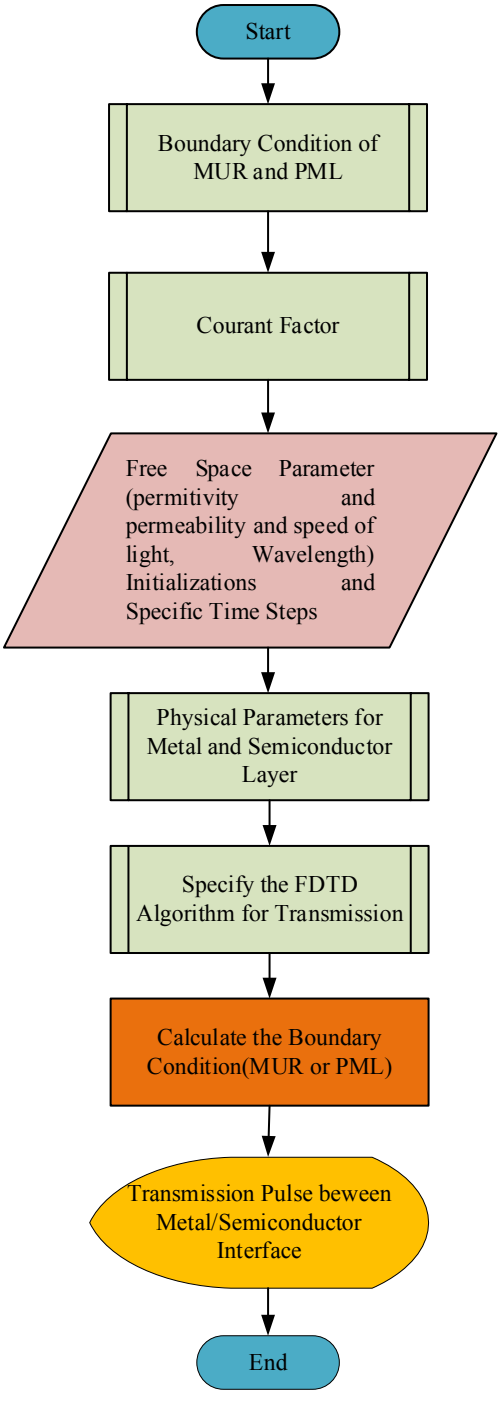


Figure 3. Flowchart for Transmission Analysis

 DOI: https://doi.org/10.30564/jcsr.v3i2.3139

At first, the boundary condition of MUR and PML with appropriate courant factor is reset. Afterward, the permittivity, permeability, speed of light and wavelength of free space parameter and specific time steps for FDTD is professed. The corporeal parameters for metal layer and semiconductor layers are quantified. The two-dimensional finite difference time domain (2D FDTD) process is modified. And then the boundary condition of MUR and PML is premeditated. Lastly, the transmission pluses for boundary condition of MUR and PML are exhibited.

2.4 Implementation of Electric Field and Electric Flux Density FDTD Measurement

All material is made up of charged particles. The material may be neutral overall because it has as many positive charges as negative charges. Nevertheless, there are various ways in which the positive and negative charges may shift slightly within the material, perhaps under the influence of an electric field. The resulting charge separation will have an effect on the overall electric field. Because of this it is often convenient to introduce a new field known as the electric flux density D which has units of Coulombs per square meter (C/m^2). The D field ignores the local effects of charge which is bound in a material [3]. In free space, the electric field and the electric flux density are related by

$$D = \varepsilon_0 E \qquad \text{Equation (6)}$$

Gauss's law states that integrating D over a closed surface yields the enclosed free charge.

$$\oint_s D.ds = Q_{enc} \qquad \text{Equation (7)}$$

Where S is the closed surface, ds is an incremental surface element whose normal is directed radically outward, and Q_{enc} is the enclosed charge. Taking S to be a spherical surface with the charge at the center, it is simple to perform the integral.

$$\oint_s D.ds = \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \varepsilon_0 \frac{Q_1}{4\pi\varepsilon_0 r^2}\hat{a}_r.\hat{a}_r r^2 \sin\theta d\varphi d\theta = Q_1$$

Equation (8)

The samples were undoped (u-) metal/semiconductor stripe structures. The semiconductor films with the thickness of 0.1 μm were grown by a metal organic vapour phase epitaxy on n-type doped (n-) (100) conventional substrate. The stripe width values of semiconductor and metal were variable micro-meter (μm). The geometric configurations for FDTD measurement on the metal/semiconductor stripe structure is shown in Figure 4.

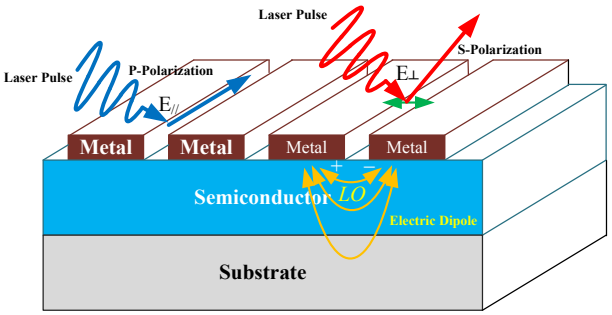


Figure 4. Geometric Configurations

The result is actually independent of the surface chosen, but the integral is especially easy to perform for a spherical surface. The integral in Equation (7) is always equal to the enclosed charge as it does in frees pace. However, things are more complicated when material is present. Two large parallel plates carry uniformly distributed charge of equal magnitude but opposite sign. The dashed line represents an integration surface S which is assumed to be sufficiently far from the edges of the plate so that the field is uniform over the top of S. This field is identified as E_0. The fields are zero outside of the plates and they are tangential to the sides of S within the plates. Therefore, the only contribution to the integral is from the top of S. The result of the integral HS .E0 .ds is the negative charge enclosed by the surface. Figure 5 shows the results of FDTD simulation for the simulation model of the electric field.

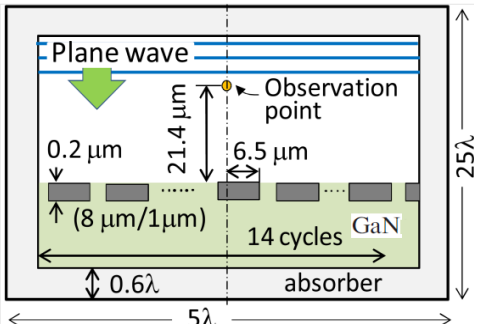


Figure 5. Simulation Model Using the FDTD Method

When a light wave packet with 10 cycles of oscillation (approximately 1 ps duration) and a low electric field amplitude of 3×10^3 V/cm, is incident on the sample for (a) polarization direction E parallel stripe and (b) E perpendicular stripe. The geometric condition including the observation point of the electric field is exhibited in the inset of Figure 5.

2.5 Quantum Mechanical Model

The conventional device modelling pilots to two im-

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

portant inaccuracies pertains to the carrier concentration near the semiconductor surface. Initially, the dividing of the conduction band into quite a few discrete eigenvalues is not measured. That goes ahead to an over evaluation of the surface charge as the energy difference between those discrete eigenvalues and the fermi-level is superior than the one from the bottom of the conduction band to the fermi-level. Next, the conventional models do not regard as that the shape of the wave function diminishes the carrier concentration near the semiconductor surface as well. Therefore, a meticulous approach to imitate the carrier concentration has to make sure of both effects by contributing the rough calculation for the wave function and the actual band structure of semiconductor devices.

2.5.1 Approach to Wave Approximation

The first quantum mechanical effects by a diminution of the density of states near the semiconductor interface affecting an exponential shape function is called wave function approximation. This pursues an approach proposed by [4],

$$SD(i) = SD\left(1 - e^{-(i-i_0)^2/\lambda_{Thermal}}\right) \quad \text{Equation (9)}$$

Where 'i' is the distance to the semiconductor interface and 'i_0' is an offset to match the nonzero carrier concentration near the surface stanching from the finite barrier height. $\lambda_{Thermal}$ is the thermal wavelength conscientious for the lessening of the quantum mechanical effects with increasing distance from the semiconductor interface.

$$\lambda_{Thermal} = \frac{\sqrt{2mkT}}{h} \quad \text{Equation (10)}$$

If that improvement is utilized the qualitative carrier distribution near the semiconductor interface in physically powerful inversion which is duplicated quite well but devoid deliberation of band structure effects is not the issue in the threshold level region [5].

2.5.2 Approach to Energy Band Structure Approximation

Near the surface of the lowest eigenenergy is connotation higher than the band edge thus reasoning an over evaluation of the charge when the conventional imitation approach is utilized. The essential initiative of the current model is to substitute the effective band edge by the first discrete energy level. This appears realistic as quantum mechanical computations confirm that regularly more than 95% of the carriers are in that energy band. The band edge at the semiconductor surface is set to:

$$E_{g,Semicon\ Surface}^{QMM} = E_{g,Semicon\ surface}^{Conventional} + \Delta E_g \quad \text{Equation (11)}$$

Whereas $E_{g,Semicon\ Surface}^{QMM}$ is the developed bandgap energy which is utilized in the Boltzmann stistics, $E_{g,Semicon\ Surface}^{Conventional}$ is the bandgap in accordance with the material specification and ΔE_g is the applied modification. The current model attaches the band edge $E_{g,Semicon\ Surface}^{QMM}(i)$ surrounded by the device to the value of $E_{g,Semicon\ Surface}^{QMM}$ as long as $E_{g,Semicon\ Surface}^{QMM} > E_{g,Semicon\ Surface}^{Conventional}(i)$..

he accurate computation of the first energy level is numerically expensive and ssitates the explanation of the Schrodinger equation and estimation is utilized. The offset ΔE_g is estimated subsequent aestablishment of Van Dort et al [6-9] which reads as

$$\Delta E_g = \frac{13}{9}\beta\left(\frac{\in}{4qkT}\right)^{1/3}\left|E_{semicon\ surface}\right|^{2/3} \quad \text{Equation (12)}$$

reas $|E_{semicon\ surface}|$ is the magnitude of the electric field at the semiconductor interface and $\in$ is the permittivity of the semiconductor, β=4.1×10^{-8} eVcm is an empirical constant.

3. Results and Discussions

3.1 Winspall Analysis on Metal-Semiconductor Nanolayer Structure

Winspall is a software for the simulation of surface plasmon resonance curves, transmission and reflection coefficient curves by polarizations based on the Fresnel formalism. A laser beam is reflected from the surface of the material layer and the reflected light is collected as a function of the angle of incidence.

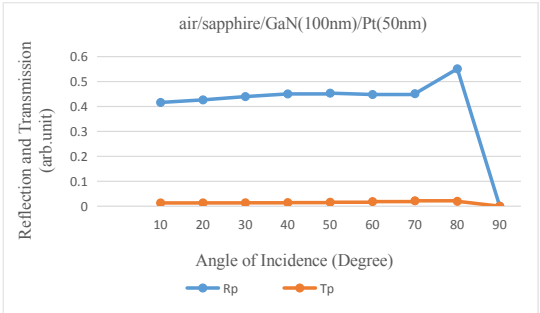


Figure 6. Reflection and Transmission Coefficient, λ=690 nm,Air(n=1,k=0, Thickness=0)/Sapphire(n=1.76, k=0,Thickness=100 nm)/GaN(n=2.365,k=0, Thickness=100 nm)/Pt(n=2.51,k=4.43,Thickness=50 nm)

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

Figure 6 shows the transmission and reflection coefficient value for air/sapphire/GaN /Platinum (Pt) structure. Reflection coefficient value of nearly 0.42 (arb.unit) and transmission coefficient value of 0.015 (arb.unit) obtain p-polarized light incident on this structure I. Nearly 0.42 reflection coefficient value is obtained along the incident angle 10 to 70 degree but the greater the angle of 70 degree, the higher the reflection coefficient value as shown in Figure 6.

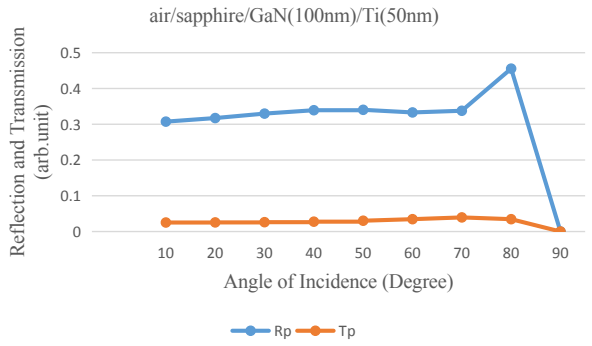


Figure 7. Reflection and Transmission Coefficient, λ=690 nm,Air(n=1,k=0, Thickness=0)/Sapphire(n=1.76,k=0,Thickness=100nm)/GaN(n=2.365, k=0, Thickness=100 nm)/Ti(n=3.03,k=3.65,Thickness=50 nm)

Figure 7 shows the transmission and reflection coefficient value for air/sapphire/GaN/Titanium (Ti) structure. Reflection coefficient value of nearly 0.31 (arb.unit) and transmission coefficient value of 0.03 (arb.unit) obtain p-polarized light incident on this structure II.

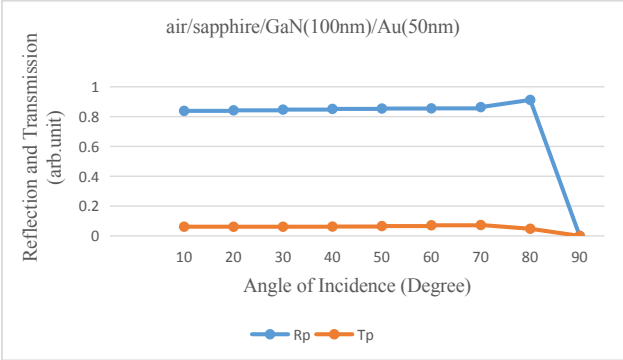


Figure 8. Reflection and Transmission Coefficient, λ=690 nm, Air (n=1,k=0, Thickness=0)/Sapphire(n=1.76,k=0,Thickness=100 nm) /GaN (n=2.365, k=0, Thickness=100 nm)/Au(n=0.16,k=3.80,Thickness=50 nm)

Nearly 0.31 reflection coefficient value is obtained along the incident angle of 10 to 70 degree but the greater the angle of 70 degree, the higher the reflection coefficient value as shown in Figure 8. Lower reflection coefficient value and a little higher transmission coefficient value in structure II compared with the structure I. So, structure II has high phonon effect than structure I.

Figure 8 shows the transmission and reflection coefficient value for air/sapphire/GaN/Gold(Au) structure. Reflection coefficient value of nearly 0.85 (arb.unit) and transmission coefficient value of 0.06 (arb.unit) obtain p-polarized light incident on this structure III. Nearly 0.85 reflection coefficient value is obtained along the incident angle of 10 to 70 degree but the greater the angle of 70 degree, the higher the reflection coefficient value as shown in Figure 8.

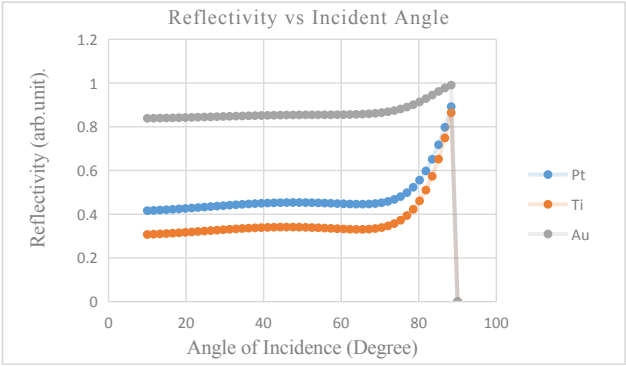


Figure 9. Comparison of Reflectivity for Au, Pt and Ti Metal

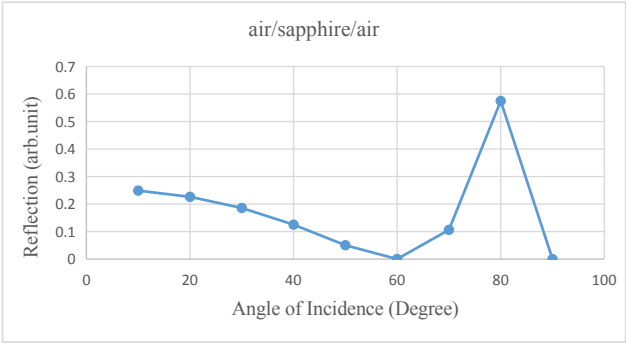


Figure 10. Reflection Coefficient, λ=632.8 nm, Air(n=1,k=0,Thickness=0) /Sapphire(n=1.76,k=0,Thickness=100 nm)

Highest reflection coefficient value is obtained in this structure III compared with the other two structures as shown in Figure 9. Structure III has less phonon effect than the other two structures, so this structure is very suitable to use in high performance metal-semiconductor based optical semiconductor devices. The following results show the step by step layer structure for structure III.

In Figure 10, layer 1 is air (n=1, k=0) so thickness is zero. Layer 2 is 100 nm thick Al_2O_3 sapphire substrate (n=1.77, k=0) with the wavelength of 632.8 nm. Reflection coefficient is zero about 58 degree to 63 degree and then more and more light is reflected until 90 degree.

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

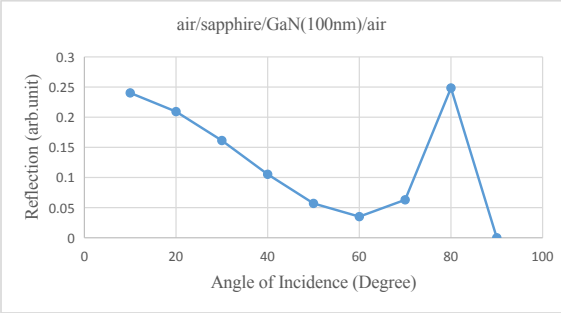


Figure 11. Reflection Coefficient, λ=632.8 nm, Air (n=1, k=0, Thickness=0) /Sapphire(n=1.76, k=0,Thickness=100 nm)/ GaN (n=2.37966,k=0, Thickness=100 nm)

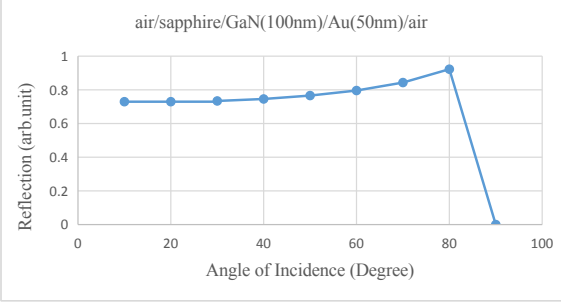


Figure 12. Reflection Coefficient, λ=632.8 nm, Air(n=1,k=0,Thickness=0) /Sapphire(n=1.76,k=0,Thickness=100 nm)/GaN(n=2.3796,k=0,Thickness=100 nm)/ Au((n=0.18104,k=3.0681,Thickness=50 nm)

In Figure 11, a 100 nm thick GaN layer is added on top of the substrate. Zero reflection coefficient value is lost because GaN layer is coated onto the substrate. Lower reflection coefficient value is obtained about nearly 70 degree but reflection coefficient value is steadily high from about 75 degree to 90 degree. Highest reflection coefficient value can get near the 90-degree p-polarized incident light. In Figure 12, a 50 nm thick gold (Au) layer is now coated on the layer of GaN. Higher reflection coefficient value is obtained in this structure.

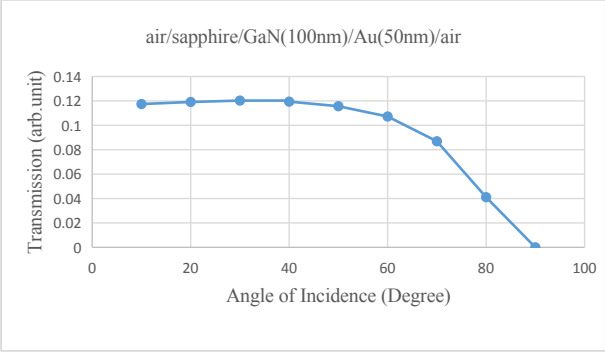


Figure 13. Transmission Coefficient, λ=632.8 nm, Air (n=1, k=0, Thickness=0) /Sapphire(n=1.76, k=0,Thickness=100 nm)/GaN(n=2.3796,k=0,Thickness =100 nm)/ Au(n=0.18104,k=3.0681,Thickness=50 nm)

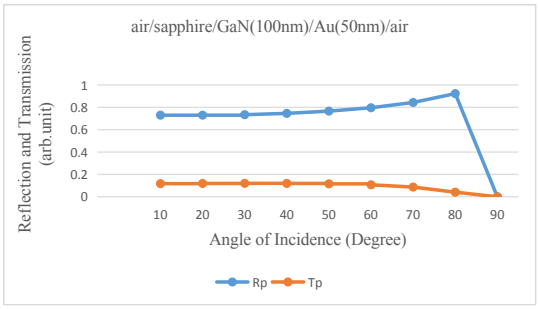


Figure 14. Transmission and Reflection Coefficient, λ=632.8 nm,Air(n=1,k=0, Thickness=0)/Sapphire(n=1.76,k=0,Thickness=100 nm)/GaN(n=2.37966, k=0,Thickness=100 nm)/ Au(n=0.18104,k=3.0681,Thickness=50 nm)

In Figure 13 shows transmission coefficient of the air / sapphire/ GaN / Au/ air structure. Transmission coefficient value is nearly 0.12 until 45 degree and then transmission coefficient is gradually decreased to zero until 90 degree. The higher the reflection value, the lower the transmission value, as shown in Figure 12 and Figure 13. Figure 14 illustrates transmission and reflection coefficient of p-polarization light in air/substrate/GaN/Au/air structure. According to Figure 14, reflection coefficient is more and more increased until total internal reflection. At that time transmission coefficient is closed to zero above 65 degree.

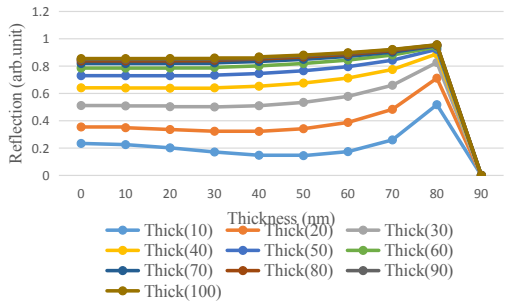


Figure 15. Reflection Coefficient, λ=632.8 nm, Air(n=1,k=0,Thickness=0)/ Sapphire(n=1.76,k=0,Thickness=100 nm)/GaN(n=2.37966,k=0,Thicknes s=100 nm)/ Au(n=0.18104,k=3.0681,Thickness=10 nm to 100 nm)

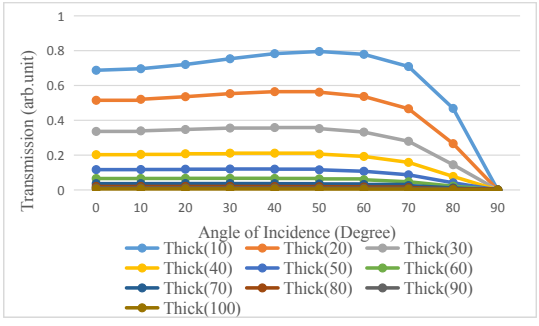


Figure 16. Transmission Coefficient, λ=632.8 nm, Air (n=1,k=0,Thickness=0) /Sapphire (n=1.76,k=0,Thickness=100 nm)/GaN(n=2.37966,k=0,Thicknes s=100 nm)/ Au(n=0.18104, k=3.0681,Thickness=10 nm to 100 nm)

 DOI: https://doi.org/10.30564/jcsr.v3i2.3139

Figure 15 shows the reflection coefficient values for top layer gold (Au) thickness are changed from 10 nm to 100 nm. The greater the thickness, the higher the reflection coefficient is obtained. According to the simulation result, top metal layer thickness value of 100 nm (0.1um) obtains high reflectivity value. Figure 16 illustrates the value of transmission coefficient for air/sapphire/GaN/Au/air structure. The greater the thickness, the lower the transmission coefficient is obtained. So, above 0.1 um thick of metal layer should be used in high performance metal-semiconductor based semiconductor devices.

3.2 Transmission, Reflection and Absorption Analysis on P-Polarized Incident Angle

Reflection, Transmission and Absorption values for air/sapphire/GaN/Au/air structure along the angle of incidence 10 to 90 degrees is shown in Figure 17.

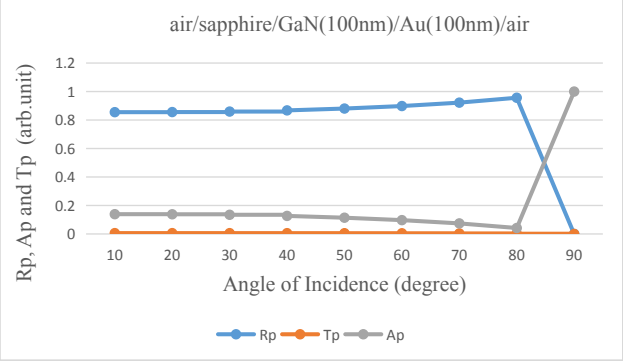


Figure 17. Reflection, Transmission and Absorption Analysis

When the reflection value is high, lower transmission and absorption value are obtained in this structure. Low absorption is less phonon energy for this structure. Along the angle of incidence, 80 degree incident light is the highest reflection value otherwise this condition gets low absorption rate.

3.2.1 Analysis on Thickness of Materials

The p-polarization of light 80 degree is incident on this M-S structure. The thickness of Au (10 nm to 100 nm) metal changes result is shown in Figure 18. 100 nm thick Au metal has high reflection value. Transmission value is zero and absorption value is nearly 0.03. So, 100 nm or upper thickness of Au metal layer can be used in high performance device (for less phonon effect).

The semiconductor GaN (10 nm to 100 nm) layer thickness changes result is shown in Figure 19.

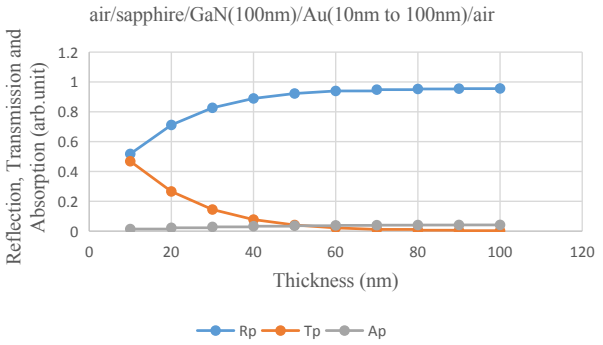


Figure 18. Reflection, Transmission and Absorption Analysis of Metal Layer

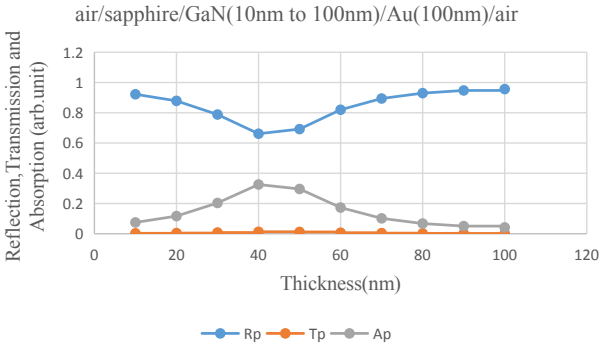


Figure 19. Reflection, Transmission and Absorption Analysis of Semiconductor Layer

Inversely proportional of reflection and absorption values can be seen in this result. Transmission value is nearly zero value in this condition. So, 100 nm or upper thickness of GaN semiconductor layer should be used in high performance device (for less phonon effect).

3.2.2 Absorption Analysis on IR Wavelength

Figure 20 illustrates the absorption coefficient value in visible wavelength region (400-700 nm).

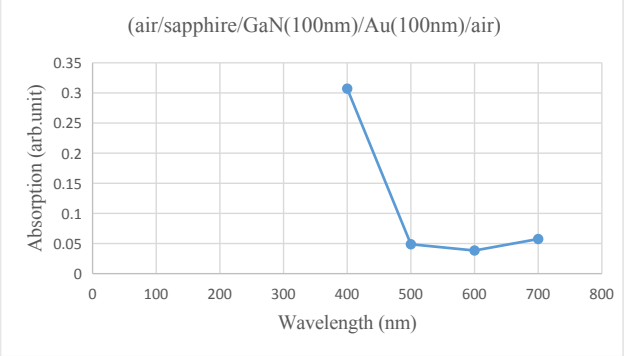


Figure 20. Absorption Analysis on 400-700 nm Wavelength

In 400 nm, 80-degree incident light, absorption coeffi-

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

cient value is about 0.31. In 500 nm, absorption coefficient value is about 0.05. In 600 nm, absorption coefficient value is about 0.04. In 700 nm, absorption coefficient value is about 0.0625. Less absorption rate gets about 600 nm-650 nm wavelength.

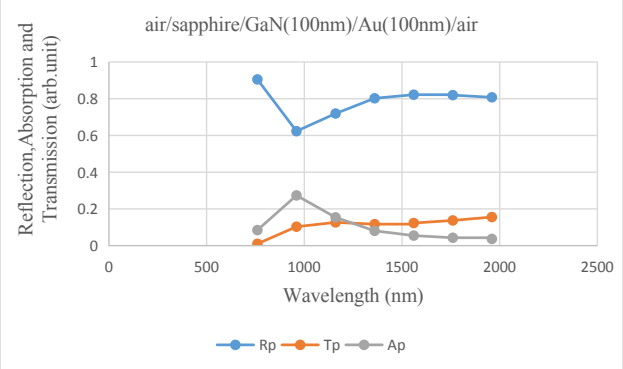


Figure 21. Absorption Analysis on Short Wave Wavelength

Figure 21 shows reflection, absorption and transmission value in short wave wavelength (760 nm-1960 nm). Reflection (R_p) value is inversely proportional to absorption (A_p) value. But transmission (T_p) value is almost inversely proportional absorption (A_p).

Reflection, absorption and transmission value in medium wave (2000 nm-4000 nm) is shown in Figure 22.

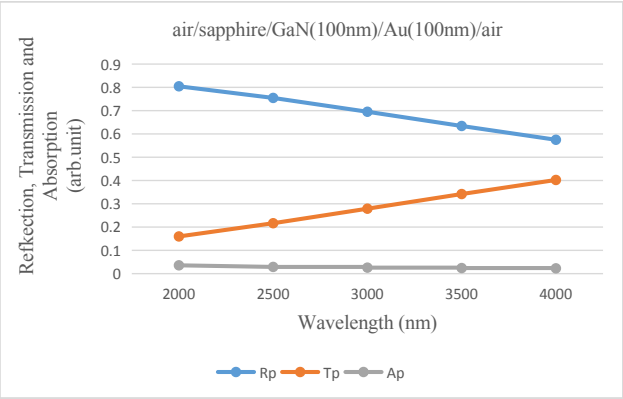


Figure 22. Absorption Analysis on Medium Wave Wavelength

Medium wave reflection (R_p) value is lower than short wave R_p value. Medium wave transmission (T_p) value is higher than short wave T_p value. Reflection (R_p) is inversely proportional to transmission (T_p) value. And Medium wave absorption (A_p) value is low in this condition.

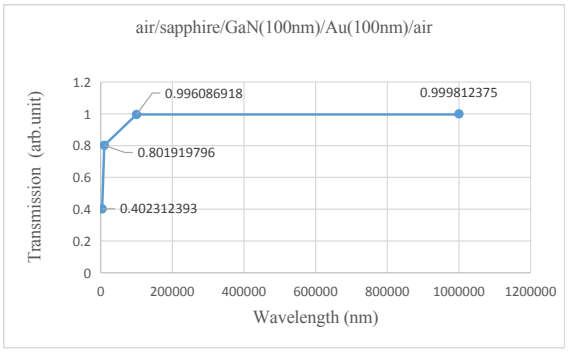


Figure 23. Transmission on Long Wave Wavelength

Figure 25 describes the reflection, absorption and transmission value in long wave wavelength (4000 nm-1000000 nm). Long wave reflection (R_p) value is lower than medium wave R_p value. Long wave transmission (T_p) value is higher than medium wave T_p value.

And long wave absorption (A_p) value is lower than medium condition and absorption value is nearly close to zero in this wavelength.

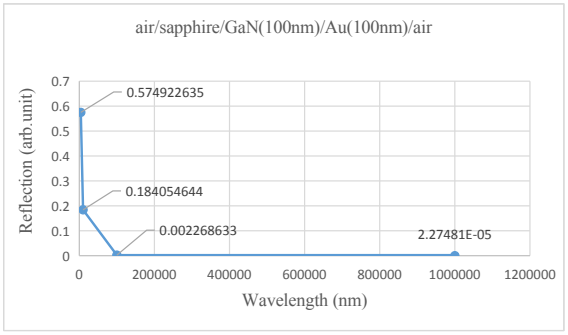


Figure 24. Reflection on Long Wave Wavelength

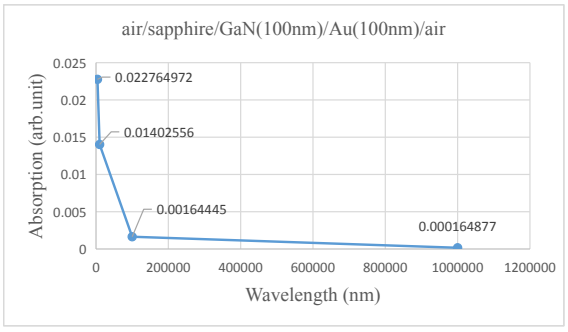


Figure 25. Absorption Analysis on Long Wave Wavelength

Absorption coefficient value is very important for optical semiconductor devices because heat in these devices is occurred in absorption of light in materials.

3.3 Transmission, Reflection and Absorption Analysis on S-Polarized Incident Angle

Reflection, transmission and absorption values for air/

 DOI: https://doi.org/10.30564/jcsr.v3i2.3139

sapphire/GaN/Au/air structure along the s-polarized angle of incidence 10 to 90 degrees is shown in Figure 26. When the reflection value is high, the transmission value is low in this structure. Absorption value is higher than transmission value for this condition.

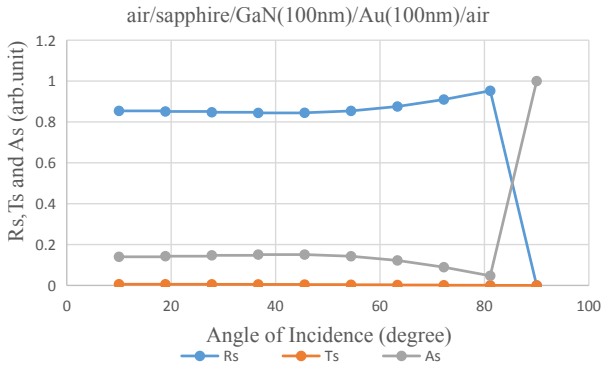


Figure 26. Reflection, Transmission and Absorption Analysis

Along the s-polarized angle of incidence, 81 degree incident light is the highest reflection value otherwise this condition gets low absorption rate.

3.3.1 Analysis on Thickness of Materials

The 80 degree s-polarization of light is incident on this metal-semiconductor structure. The thickness of Au (10 nm to 100 nm) metal result is shown in Figure 27. 100 nm thick Au metal has high reflection value. Transmission value is zero and absorption value is nearly 0.14. So, s-polarization of 100 nm or upper thickness of Au metal layer gets more absorption value than p-polarization of Au metal layer.

Inversely proportional of reflection and absorption values can be seen in this result. Transmission value is nearly zero value in this condition. So, 100 nm or upper thickness of GaN semiconductor layer should be used in high performance device (for less phonon effect).

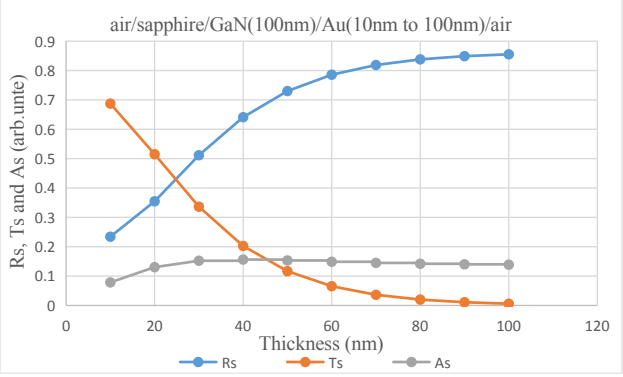


Figure 27. Reflection, Transmission and Absorption Analysis of Metal Layer

The semiconductor GaN (10nm to 100 nm) layer thickness changes result is shown in Figure 28.

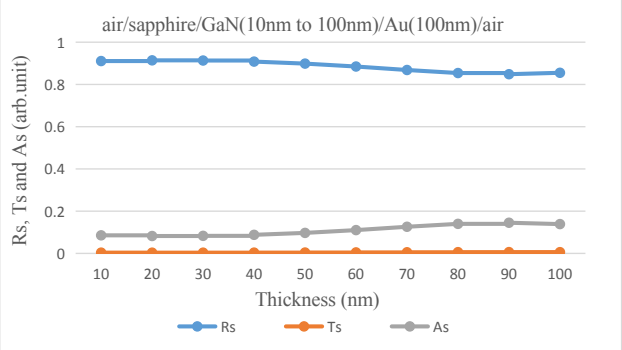


Figure 28. Reflection, Transmission and Absorption Analysis of Semiconductor Layer

3.3.2 Absorption Analysis on IR Wavelength

Figure 29 shows reflection, absorption and transmission value in short wave wavelength (760 nm-1960 nm).

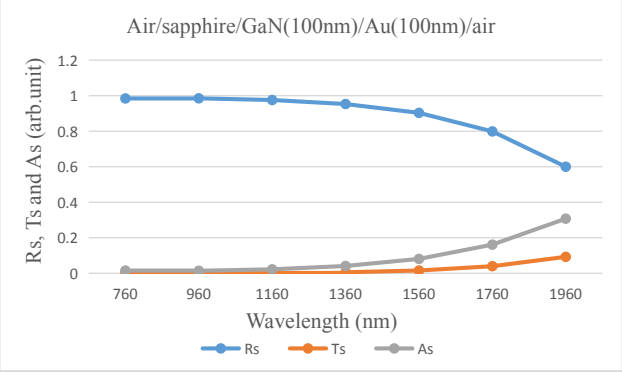


Figure 29. Absorption Analysis on Short Wave Wavelength

Reflection (R_s) value is inversely proportional to absorption (A_s) value. Transmission (T_s) value is lower than absorption (A_s) in s-polarized light. Reflection, absorption and transmission value in medium wave (2000 nm-4000 nm) is shown in Figure 30.

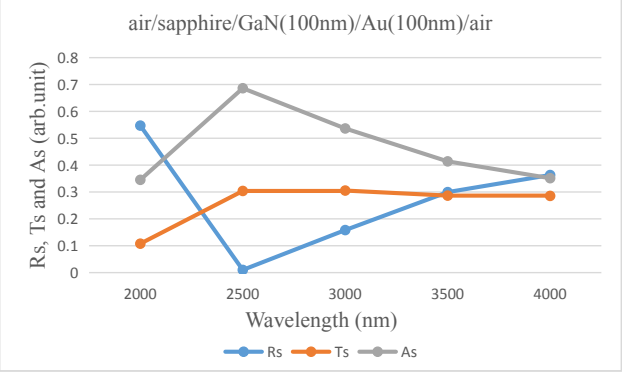


Figure 30. Absorption Analysis on Medium Wave Wavelength

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

Medium wave reflection (R_s) value is lower than short wave R_s value. Medium wave transmission (T_s) value is higher than short wave (T_s) value. Reflection (R_s) is inversely proportional to absorption (A_s) value. And Medium wave absorption (A_s) value is high in this condition.

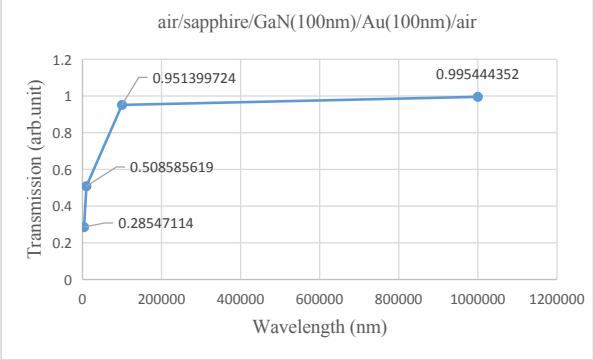


Figure 31. Transmission on Long Wave Wavelength

Figure 31, 32 and 33 describe the transmission, reflection, and absorption value in long wave wavelength (4000 nm-1000000 nm).

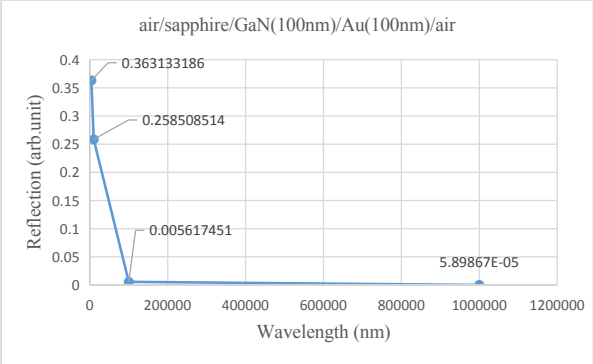


Figure 32. Reflection on Long Wave Wavelength

Long wave reflection (R_s) value is lower than medium wave R_s value. Long wave transmission (T_s) value is higher than medium wave T_s value. And long wave absorption (A_s) value is lower than medium condition and absorption value is nearly close to zero in this wavelength.

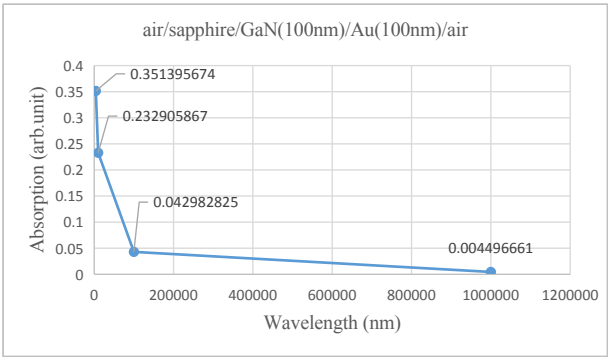
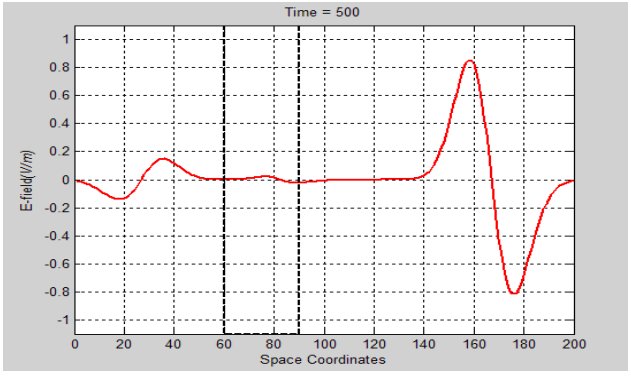


Figure 33. Absorption Analysis on Long Wave Wavelength

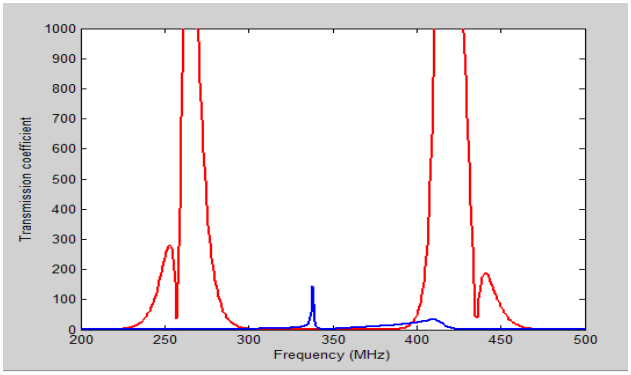
Absorption coefficient value is very important for optical semiconductor devices because heat in these devices is occurred in absorption of light in materials.

3.4 FDTD Absorbing Boundary Condition for Semiconductor Quantum Devices

The FDTD simulation for e-field and transmission coefficients for time steps =500 is illustrated in Figure 34. There are many simulation approaches for transmission and reflection condition for electromagnetic energy through the medium like PML boundary or MUR boundary. The simulation results show that the FDTD analysis on absorbing boundary condition for semiconductor quantum devices can be focused on the optical properties of the devices for high performance condition.



(a)



(b)

Figure 34. FDTD Simulation for (a) E-field and (b) Transmission Coefficients for Time Steps = 500

Figure 35 shows the reflection pulse at metal-semiconductor interface with MUR ABC boundary. Depending on the boundary condition, the reflection center is occurred at 250 microns in x-coordinate in MUR analysis with FDTD simulation between metal and semiconductor interface.

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

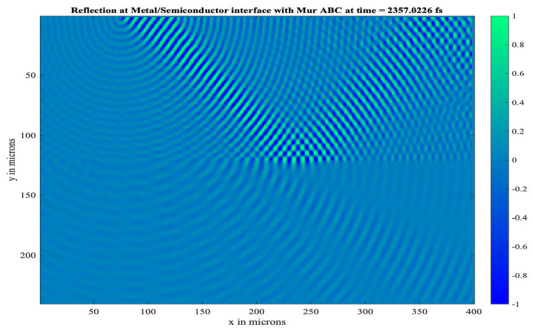


Figure 35. Reflection Pulse at Metal/Semiconductor Interface with MUR ABC Boundary

The reflection pulse can be detected 400 microns in x-coordinate and it is the best detection point for the carrier dynamics effects for interface engineering. Figure 36 shows the transmission pulse at metal-semiconductor interface with MUR ABC boundary.

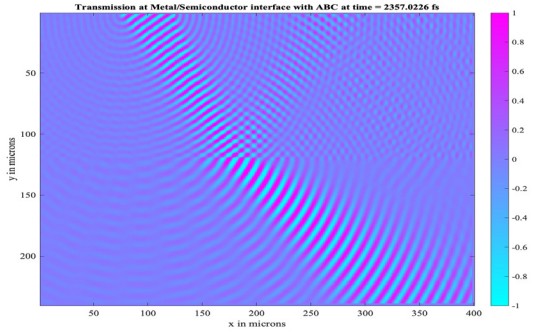


Figure 36. Transmission Pulse at Metal/Semiconductor Interface with MUR ABC

Resting on the boundary condition, the transmission midpoint is occurred at 180 microns in x coordinate in MUR investigation with FDTD simulation between metal and semiconductor interface. The reflection pulse can be detected 350 microns in x coordinate and it is the finest detection point for the carrier dynamics effects for interface engineering.

Figure 37 shows the reflection pulse at metal-semiconductor interface with PML boundary.

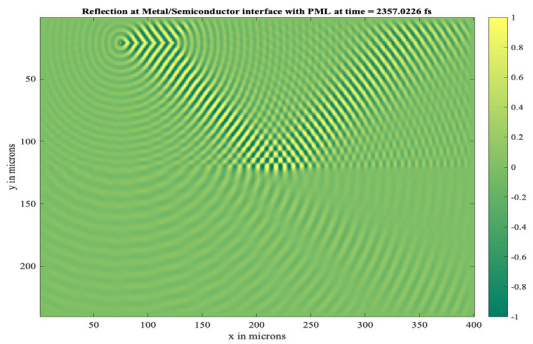


Figure 37. Reflection Pulse at Metal/Semiconductor Interface with PML Boundary

Based on the boundary condition of PML, the reflection center is transpired at 225 microns in x coordinate in PML analysis with FDTD simulation between metal and semiconductor interface. The reflection pulse can be identified 350 microns in x coordinate and it is the paramount recognition point for the carrier dynamics effects for interface engineering.

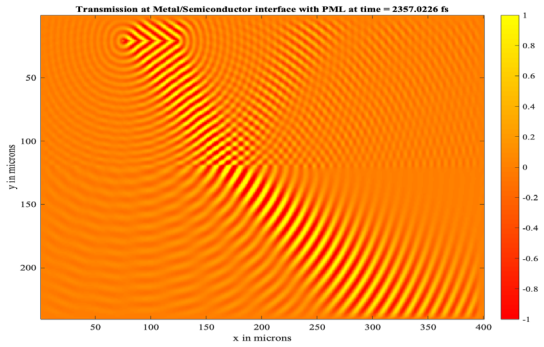


Figure 38. Transmission Pulse at Metal/Semiconductor Interface with PML Boundary

Figure 38 shows the transmission pulse at metal-semiconductor interface with PML boundary. Established on the boundary condition of PML, the transmission focus is emerged at 160 microns in x coordinate in PML analysis with FDTD simulation between metal and semiconductor interface.

3.5 Discussions

Finite-Difference Time-Domain analysis on absorbing boundary condition for solving a time-dependent Schrödinger equations are studied. The reflectance and transmittance energy from Au/GaN interface, the electric field and electric flux density value are obtained at the metal-semiconductor interface. The measurement on reflection value of gold, platinum and titanium metal is coated on GaN semiconductor with 100 nm thickness. One dimensional FDTD method is one of the solutions for observing the energy spectrum in a semiconductor material. These simulation researches tend to fabricate the photodetectors or other device to get the better performance of metal-semiconductor based optical semiconductor devices. This study affects to find the new solution for less phonon energy (or) small absorption measurement on metal-semiconductor layer structure. That practice can solve a discretized Schrödinger equation in an iterative progression. The comprehensive FDTD method is experienced by simulating a particle stirring in free space and then trouncing an energy potential. Numerical results correspond to attain based on the outcomes from the simulation results.

DOI: https://doi.org/10.30564/jcsr.v3i2.3139

The short, medium and long IR wavelengths of the Au/GaN samples were studied by affecting the winspall technique. The absorption value for Au/GaN sample was described with p-polarization and s-polarization of IR light. P-polarization light absorption value is smaller than s-polarization light absorption value. But reflection value is high in both p-polarization and s-polarization. Less absorption value device structure can protect the rising of heat in processing of this device. So, Au/GaN structure is less phonon effect structure for fabricating of metal-semiconductor based semiconductor device such as photodetectors. The transmittance and reflectance spectra on three metals is confirmed by winspall techniques which depend on the absorption spectrum. The transmittance and reflectance spectrum for interface of metal-semiconductor are also described with finite-difference time domain method by Mur's and PML (perfectly matched layer) boundary conditions. Perfectly matched layer boundary condition can early detect the sharp signal than the Mur's absorbing boundary condition. Conduction band energy or electron concentration of these structure models are compared with the Schrodinger model and conventional model. The junction capacitances due to dipole in the transition region are also illustrated under reverse bias condition.

4. Conclusions

A comprehensive FDTD method has developed with various kinds of absorbing boundary condition for solving the 1D time dependent Schrödinger equation and obtains a more relaxed condition for stability when central difference calculations are presented in new physical results, the power of the FDTD technique must be borne in mind. As an explicit space-domain technique, one can avoid difficulties associated with constructing single-particle orbital that is used in computations based on Slater determinate. A Slater-determinant-based calculation will require computing the single-particle orbital for each potential chosen. Of course, the trade-off is in the size of the spatial mesh chosen for calculations.

Low absorption based device structure is suitable for high performance device because this structure can fabricate less phonon effect device structure. The simulation results have been conducted by using MATLAB language for analysis. A comparison of the phonon energy with the p-polarization and s-polarization as measured by absorption and reflection value shows that the short, medium and long wave IR wavelength. In short wave, s-polarized IR light absorption value is higher than p-polarized IR light absorption value. In medium wave, s-polarized IR light absorption value is also higher than p-polarized IR light absorption value. Long wave absorption result is also the same above phenomenon. But reflection or transmission value is always high along the whole IR wavelength.

References

[1] Lee, Y. S. 2009. "Principles of Terahertz Science and Technology."

[2] Jacobsen, F.2008. "An Efficient Realization of Frequency Dependent Boundary Conditions in an Acoustic Finite-Difference Time-Domain Model." vol. 316: 234-247.

[3] Chatterjee, S.2004. "Excitonic Photoluminescence in Semiconductor Quantum Wells: Plasma versus Excitons." Physical Review Letters, vol.92, issue.6.

[4] Koch, S. W. 2011. "Semiconductor Quantum Optics." Cambridge University Press.

[5] Li, W. 2006. "Post Growth Thermal Annealing of GaN Grown by RF Plasma." Journal of Crystal Growth, vol.227 : 415-419.

[6] Khromov, S. 2013. "Doping Effects on the Structural and Optical Properties of GaN." Science and Technology Dissertation, no. 1554, Sweden: Thin Film Physics.

[7] Bengtsson, M. 2008. "Finite Difference Time-Domain Simulations of Exciton-Polariton Resonances in Quantum Dot Arrays." Optical Society of America, vol.16, no. 7 : 4507-4519.

[8] Lee, Y. S. 2009. "Principles of Terahertz Science and Technology."

[9] Moini, M. 2008. "On the Numerical Solution of One Dimensional Schrödinger Equation with Boundary Conditions Involving Fractional Differential Operators." International Journal of Engineering Science, vol.19 : 21-26.

 DOI: https://doi.org/10.30564/jcsr.v3i2.3139

BILINGUAL
PUBLISHING CO.
Pioneer of Global Academics Since 1984

Tel:+65 65881289
E-mail:contact@bilpublishing.com
Website:www.bilpublishing.com