IJRSET October 2015 Volume 2, Issue 8 Pages: 14-18

ISSN 2394-739X

International Journal for Research in Science Engineering and Technology

A NOVEL FUZZY CLASSIFICATION PROTOCOL OVER ENCRYPTED DATA IN THE CLOUD.

[1] Mrs. R. UMA MAHESWARI M.C.A, [2] Mrs. P.SHANTHI M.C.A, M.Phil.

[1] Assistant Professor, [2] Associate Professor

[1] Department Of Computer Science, [2] Department Of Information Technology,

[1] Nirmala College for Women, [2] Sri Jayendra Saraswathy Maha Vidyalaya College of Arts and Science,

[1, 2]Coimbatore – 641 005

Abstract: -

Data mining is the investigation venture of the "Learning Discovery in database". It is an interdisciplinary subfield of software engineering and the computational procedure of finding examples in huge data sets Classification is a data mining strategy used to anticipate bunch participation for data occasions. The data is scrambled utilizing the Knapsack cryptosystem calculation. In past strategy utilizing the paillier cryptosystem calculation. The scrambled data is ordered semantically utilizing the fluffy rationale calculation. In past technique we are utilized KNN classifier. This order calculation gives the low Efficiency is low contrast with the Fuzzy rationale. We are utilizing numerous protected conventions for actualizing grouping calculation. Enhancing the effectiveness of SMIN convention is vital in first stride for development of the entire classifier. In the KNN the SMIN convention productivity is less in this way, the effectiveness of the KNN is less contrast with the fluffy rationale classifier. The effectiveness is less in light of the fact that the conventions utilized as a part of the current strategy The issue of registering nth deposit raunchy is accepted to be computationally troublesome.. The classification is critical in the data mining. For high privacy we go for fluffy classifier. The expense and proficiency of the fluffy classifier is high .

Keywords: - [Fuzzy classification, Classification protocol, cloud security]

1. INTRODUCTION

Data mining systems are the consequence of a long procedure of examination and item improvement. This advancement started when business data was initially put away on PCs, proceeded with upgrades in data get to, and all the more as of late, produced advances that permit clients to explore through their data continuously. Data mining takes this developmental process past review data access and route to forthcoming and proactive data conveyance. Data mining gets its name from the likenesses between hunting down significant business data in a huge database for instance, discovering connected items in gigabytes of store scanner data and mining a mountain for a vein of profitable metal. Both procedures oblige either filtering through a colossal measure of material, or astutely testing it to discover precisely where the worth lives. Given databases of adequate size and quality, data providing so as to mine innovation can produce new business opportunities these capacities. Established rationale just allows recommendations having an estimation of

truth or lie. The idea of whether 1+1=2 is a flat out, changeless and scientific truth. Be that as it may, there exist certain recommendations with variable answers, for example, requesting that different individuals recognize shading. The idea of truth doesn't fall by the wayside, yet rather on a method for speaking to and aggregating so as to think over incomplete information when managed, every single conceivable result into a dimensional range.

Both degrees of truth and probabilities reach somewhere around 0 and 1 and henceforth may appear to be comparable at first. For instance, let a 100 ml glass contain 30 ml of water. At that point we may consider two ideas: unfilled and full. The significance of each of them can be spoken to by a certain fluffy set. At that point one may characterize the glass as being 0.7 void and 0.3 full. Note that the idea of vacancy would be subjective and in this manner would rely on upon the onlooker or planner. Another fashioner may, just as well, plan a set participation capacity where the glass would be viewed as full for all qualities down to 50 ml. It is crucial to understand that fluffy rationale utilizes truth degrees as a numerical model of the unclearness marvel while likelihood is a scientific model of obliviousness.

The Japanese were the first to use fluffy rationale for down to earth applications. The primary eminent application was on the fast prepare in Sendai, in which fluffy rationale had the capacity enhance the economy, solace, and accuracy of the ride.[7] It has additionally been utilized as a part of acknowledgment of manually written images in Sony pocket PCs; flight help for helicopters; controlling of tram frameworks so as to enhance driving solace, exactness of stopping, and force economy; enhanced fuel utilization for auto mobiles; single-catch control for clothes washers; programmed engine control for vacuum cleaners with acknowledgment of surface condition and level of ruining; and expectation frameworks for ahead of schedule acknowledgment of quakes through the Institute of Seismology Bureau of Metrology, Japan

2. PAILLIER CRYPTOSYSTEM

The Paillier cryptosystem, named after and invented by Pascal Paillier in 1999, is a probabilistic asymmetric algorithm for public key cryptography. The problem of computing *n*-th residue classes is believed to be computationally difficult. The decisional composite residuosity assumption is the intractability hypothesis upon which this cryptosystem is based. The scheme is an additive homomorphism cryptosystem; this means that, given only the public-key and the encryption of m_1 and m_2, one can compute the encryption of $m_1 + m_2$.

The Paillier cryptosystem is an additive homomorphism and probabilistic public-key encryption scheme whose security is based on the Decisional Composite Residuosity Assumption [4]. Let Epk be the encryption function with public key pk given by (N, g), where N is a product of two large primes of similar bit length and g is a generator in Z ∗ N2 . Also, let Dsk be the decryption function with secret key sk. For any given two plaintexts a, b ∈ ZN , the Paillier encryption scheme exhibits the following properties:

1) Homomorphism Addition Dsk(Epk(a + b)) = Dsk(Epk(a) ∗ Epk(b) mod N 2);

2) Homomorphism Multiplication Dsk(Epk(a ∗ b)) = Dsk(Epk(a) b mod N 2);

3) Semantic Security - The encryption scheme is semantically secure.

Briefly, given a set of cipher texts, an adversary cannot deduce any additional information about the plaintext(s). For succinctness, we drop the mod N2 term during homomorphism operations

2.1 KEY GENERATION

1. Choose two large prime numbers *p* and *q* randomly and independently of each other such

that $\gcd(pq, (p-1)(q-1)) = 1$. This property is assured if both primes are of equal length.[1]

2. Compute $n = pq$ and $\lambda = \text{lcm}(p-1, q-1)$.

3. Select random integer g where $g \in \mathbb{Z}^*_{n^2}$

4. Ensure n divides the order of g by checking the existence of the following modular multiplicative inverse: $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n$,

where function L is defined as $L(u) = \frac{u-1}{n}$.

Note that the notation $\frac{a}{b}$ does not denote the modular multiplication of a times the modular multiplicative inverse of b but rather the quotient of a divided by b, i.e., the largest integer value $v \geq 0$ to satisfy the relation $a \geq vb$.

- The public (encryption) key is (n, g).
- The private (decryption) key is (λ, μ).

If using p,q of equivalent length, a simpler variant of the above key generation steps would be toset $g = n+1, \lambda = \varphi(n)$, and $\mu = \varphi(n)^{-1} \bmod n$, where $\varphi(n) = (p-1)(q-1)$.

2.2 ENCRYPTION

1. Let m be a message to be encrypted where $m \in \mathbb{Z}_n$

2. Select random r where $r \in \mathbb{Z}^*_n$

3. Compute ciphertext as: $c = g^m \cdot r^n \bmod n^2$

2.3 DECRYPTION

1. Let c be the cipher text to decrypt, where $c \in \mathbb{Z}^*_{n^2}$

2. Compute the plaintext message as: $m = L(c^\lambda \bmod n^2) \cdot \mu \bmod n$

As the original paper points out, decryption is "essentially one exponentiation modulo n^2."

3. FUZZY LOGIC

Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1. By contrast, in Boolean logic, the truth values of variables may only be 0 or 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false.[1] Furthermore, when linguistic variables are used, these degrees may be managed by specific functions. Any axiomatizable fuzzy theory is recursively enumerable. In particular, the fuzzy set of logically true formulas is recursively enumerable in spite of the fact that the crisp set of valid formulas is not recursively enumerable, in general. Moreover, any axiomatizable and complete theory is decidable. It is an open question to give supports for a "Church thesis" for fuzzy mathematics, the proposed notion of recursive innumerability for fuzzy subsets is the adequate one. To this aim, an extension of the notions of fuzzy grammar and fuzzy Turing machine should be necessary. Another open question is to start from this notion to find an extension of Gödel's theorems to fuzzy logic. It is known that any Boolean logic function could be represented using a truth table mapping each set of variable values into set of values $\{0, 1\}$. The task of synthesis of boolean logic function given in tabular form is one of basic tasks in traditional logic that is

solved via disjunctive (conjunctive) perfect normal form.

Fuzzy logic and probability address different forms of uncertainty. While both fuzzy logic and probability theory can represent degrees of certain kinds of subjective belief, fuzzy set theory uses the concept of fuzzy set membership, i.e., *how much* a variable is in a set (there is not necessarily any uncertainty about this degree), and probability theory uses the concept of subjective probability, i.e., *how probable* is it that a variable is in a set (it either entirely is or entirely is not in the set in reality, but there is uncertainty around whether it is or is not). The technical consequence of this distinction is that fuzzy set theory relaxes the axioms of classical probability, which are themselves derived from adding uncertainty, but not degree, to the crisp true/false distinctions of classical Aristotelian logic.

4. DATA SETS

Dataset and Experimental Setup For our investigations, we utilized the Car Evaluation dataset from the UCI KDD chronicle [34]. It comprises of 1,728 records (i.e., n ¼ 1; 728) and six traits (i.e., m ¼ 6). Likewise, there is a different class quality and the dataset is classified into four unique classes (i.e., w ¼ 4). We encoded this dataset trait astute, utilizing the Knapsack encryption whose key size is shifted in our investigations, and the scrambled data were put away on our machine. Taking into account our FUZZY convention, we then executed an arbitrary inquiry over this scrambled data. For whatever remains of this segment, we don't talk about the execution of Alice since it is an one-time cost. Rather, we evaluate and break down the two's exhibitions stages in FUZZY independent

CONCLUSION

To ensure client protection, different classification procedures have been proposed over the previous decade. The current strategies are not pertinent to outsourced database situations where the data dwells in encoded structure on an outsider server. This paper proposed a novel Fuzzy classification convention over scrambled data in the cloud. Our convention ensures the data's secrecy, client's information inquiry, and shrouds the data access examples furthermore assessed the execution of our convention under diverse parameter settings. Since enhancing the effectiveness is a critical first stride for enhancing the execution of our Fuzzy convention, the multifaceted nature time of the proposed strategy utilizing KD-trees or other hashing systems . Such endeavors are best left to be done later on.

REFERENCES

[1] “Solving the Problem of the K Parameter in the KNN Classifier Using an Ensemble Learning Approach”,2014

[2], “Efficient Interactive Brain Tumor Segmentation As Within-Brain Knn Classification”,2014.

[3] “Feature Selection for High-Dimensional Genomic Microarray Data”, 2002.

[4] “ “No Free Lunch” Theorems Applied to the Calibration of Traffic Simulation Models , 2014.

[5] “Automatic Classification of Asymptomatic and Osteoarthritis Knee Gait Patterns Using Kinematic Data Features and the Nearest Neighbor Classifier” , 2008.

[6] “Another Move Toward the Minimum Consistent Subset: A Tabu Search Approach to the Condensed Nearest Neighbor Rule” , 2001

[7] “k-Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data” , 2015

[8] T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification,".

[9] Y. Hamamoto, S. Uchimura, and S. Tomita, "A Bootstrap Technique for Nearest Neighbor Classifier Design.

[10] E. Alpaydin, "Voting Over Multiple Condensed Nearest Neoghbors," .

[11] K. Q. Weinberger and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," .
[12] A. Kataria and M. D. Singh, "A Review of Data Classification Using K-Nearest Neighbour Algorithm," .
[13] N. Bhatia and A. Vandana, "Survey of Nearest Neighbor Techniques,".
[14] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN Model- Based Approach in Classification," .
[15] G. Gates, "The Reduced Nearest Neighbour Rule," .
[16] M. Kubat and M. Jr, "Voting Nearest-Neighbour Subclassifiers,".