European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 07, Issue 07, 2020

Iot Enabled Privacy Preserving In Federated Hybrid Smart Grid Cloud Environment Using Spark

Dr. Shafali Jain[1], Dr.A.Gnanasekar[2], Dr S Hasan Hussain[3], Dr.C.Bala Subramanian[4], S. Stewart Kirubakaran[5]

[1]Professor and Head, Department of Electrical and Electronics Engineering, Sagar Institute of Research and Technology, Bhopal, Madhya Pradesh -462041.
[2]Associate Professor, Department of Computer Science and Engineering, R.M.D.Engineering College, Kavaraipettai- 601206, Tamilnadu, India.
[3]Associate Professor, Department of Computer Science and Engineering, Syed Ammal Engineering College, Ramanathapuram- 623502, Tamilnadu, India.
[4]Assistant Professor, Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Virudhunagar- 626126, Tamilnadu, India.
[5]Assistant Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, SIMATS, Thandalam-602105,Tamilnadu, India.

Abstract: Association rule data mining provides to integrate large volume of dataset and extract information for making effective decisions. The extracting useful information is found large number of disjoint dataset with multiple cloud federations. Cloud federation is a collection of shared data and hosted from multiple cloud user environment. In this case the organizations are sharing data and adopt changing environments. The privacy preserving is an important factor and allows the organization to control the cloud environments while accessing the data. In the paper, we use association rule mining in federated cloud environment using spark model for preserving privacy. In this framework set of user can allows to access the private dataset and encrypt the dataset for providing privacy. In our proposed model we consider confidentiality, privacy and integrity major factors. We propose privacy preserving identity access mode with attribute based access control mechanism in cloud environment. The user can get access from set of policies granted by cloud environment. The spark model is used for testing the experiments and compares the results with existing methods.

Key terms: IoT, Cloud Services, Privacy Preserving, Smart Grid, Federations, Spark model

1. INTRODUCTION

The cloud federation is collaboration between various organizations that gives organizations access to hosted data. Fixed cloud platforms for other federal members to take advantage of a specific problem or specific business opportunity. However, the availability of cloud federations is very important Adopting solutions that enable secure data sharing between them Participating organizations. In particular, it is necessary Allowing identity and access management systems Federal agencies to authenticate customers from other affiliates to determine the companies and the permissions within them shared data.
The research team also suggested some identities Access management systems for federal clouds. However, these systems have several limitations. First, they need trust agreements between federations Companies to verify users' identities and map their access Manage

policies on shared data. These solutions are not ideal contributions for dynamic and Cloud Federations Controls on the organizations. Above all, there must be partner organizations Users' identities can be checked and access control enforced policies on shared data without establishing trust agreements between them.

Second, it provides better access to specific solutions and Limit policies, partner with potential competitors. Finally, different factors evaluate policies and implement the resulting decisions. However, it is harmful users may damage access control policies or their evaluation Engine to gain unauthorized access to shared data Federal. To this day, cyber-attacks rob integrity no policy decisions and no implementation points Reported, but attacks as witnesses are still possible Certification authority's on user identification characteristics. However, they do not protect the privacy of features, which convey strategic information that companies are not ready for accepting services

One of the main goals of big data is to gather useful knowledge from large datasets in practice Data Mining Techniques. An organization is effective mining for their data locally, the demand for better data will increase mining results encourage companies to share their data for mutual benefit. It is clear that data mining produces more accurate and useful results when applied in different ways datasets collected from multiple organizations. This is the process commonly known as Cooperative Data Mining (CDM). CDM can effectively address many of today's applications in multiple domains such as social networks and healthcare Finance, Construction, Cyber security, Biology, Physics. It requires the collection, maintenance, integration and analysis of joint datasets. Similarly, it could be customer data shared in insurance companies to improve their risk evaluation strategies. However, due to increasing privacy concerns about data and increased government control Privacy and companies may not share them confidential data with other organizations.

A solution Contribution to protecting data privacy is often argued Data mining is "data mining that protects privacy". However, Lack is a major drawback of existing solutions scalability and limited applicability. Also, the majority they compromise security to increase consumption Information. Companies want to source their data analytics work in our cloud environment. Effectively address their financial and performance needs. Since the cloud is an unreliable third party, data is often encrypted and then our source. In such a case, very limited work Data analytics are performed in the cloud to protect privacy. In this paper has following sections, section 2 discuss about various related works, section 3 gives models and procedure for implementing privacy model, section 4 describes proposed model with simulation and section 5 explain results and discussions.

Related Work

Cheng et al, one-to-one item mapping is suggested using a substitute technology to add duplicate items to modified each transaction dataset. In particular, the data owner attaches duplicate objects to dictionary, and then randomly maps each item to a subset values. As a result, the server could not easily find the distribution in the correct item sets in the database. Wong et al, this approach has two limitations. First, the possibility of adding counterfeit goods to it every transaction is the same. I.e., fraudulent transactions appear in a large database leading to similar frequencies cryptography analysis. Second, duplicate items were added to transaction database is different for existing objects and so on some calculations are true duplicate items can be eliminated and found. Myth et al, An approach based on K-support is suggested anonymity to protect each sensitive species with k – 1 item in the database with similar support value. Each the modified item will not be detected from at least k – 1 Items.

Wait et al, the pseudo-taxonomy tree was used to limit their approach Incidents of counterfeit goods. K-support suggests a model that extends the concept of anonymity to k-privacy, where

each modified element is at least k - 1 identifiable from other item sets to achieve K-privacy, they first used the One to One mapping method to replace plain objects, and then Items are divided into k-size groups for k-privacy. Sater et al, duplicate transactions have been added as in the previous steps. Recently, Yi et al, Data is considered a problem the owner encrypts his / her database and stores it on a DB server.

The Consumer Association Rule Mining Task is sent $n \geq 2$ are semi-standard servers, which enforce the association rule enter the encrypted results by collection of cloud data into the encrypted data User. They suggested three different solutions for data retrieval based on canonity, K-support and K-privacy techniques. Chung et al, privacy is specific solutions depend on the supplied Cryptocurrency system to obtain item privacy and transaction privacy Database privacy.

However, their solutions are valuable the information is therefore considered insecure. The proposed protocol addresses such security gaps. Federation Supervision includes runtime monitoring and offline auditing, FRM and FSA components, respectively. The FRM is formed by a set of distribution probes Controls all interactions between the federated during runtime Clouds; If any error is found, security Increases the alert for the administration console. FSA and more these interventions is audited offline using machine learning Techniques. Federal management supports this work Management of cloud associations.

There is an IWM component Optimized workload strategies based on computing charge Requests for implementation on members' cloud systems. The Provides administration console instead of FAM component to control, regulate and monitor the condition of the Confederacy. Through the RI component, all administrative data of important Confederation, e.g. SLA and access control policies are stored Make the registry accessible to components accordingly.

The Identity and Access Control Manager agrees our identity and access control management system. The system offers the following features: Unknown identities. The system allows access to shared data while protecting federal companies' customers Identification features also for Access Control Manager. The Depending on the anonymity of user identities (pseudo-) can be achieved Identity tokens without explicit submission of identity features. Identity Manager gives users identification tokens Contains structured information. In particular, each Username, identity attribute name, And a meaningful and secure Pederson commitment the identity attribute value of the user signed by the Identity Manager.

Problem Statement and System Architecture

This section provides an overview of what we are referring to Identification and access control system for cloud federation. As this system is intended to be part of spark Platform, we will first introduce FaaS and its platform. FaaS is the equivalent of a service for security enabled clouds Creation and maintenance of cloud associations. To this extent, It is supported by the spark platform and its graphical description is reported in Figure 1 allows you to federalize, verify cloud data and services.

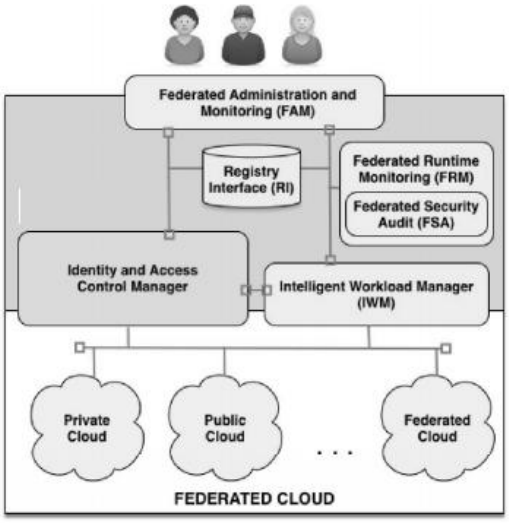


Figure 1: Identify access control model for implementing privacy preserving – Smart Grid

Secure provisioning and access through trusted identity management and access control functionality. In particular, the platform components reported in the image may be logical Recognized as Federation Monitoring and Federation Management Identity and Access Control Manager. Here we report that specific entities are a specific component Architecture.
Data Owner is readymade Federated Company Share data. Data Request User in the Cloud Federation To access shared data. The Data Owner Program is a simple web-based application. It gives the data owner an interface to share data with define access control policies that control access such data. The Identity Provider Application (IDP) is cloud based application that consists of three parts and runs on the Spark Enclave. It gets recognition features from data request then generates identification tokens. Relay IDP is user-space application Network connectivity from $Encl_{IdP}$ to CI_{dP} Data Request Application. Interface to data allows the app to verify Verification for $Encl_{IdP}$ is also sent to $Encl_{IdP}$ data the applicant's identity characteristics and its request to issue identity tokens.
In our model, the following two entities are considered, (i). Set of users are represented U1, U2,..Uq and (ii). Cloud environments are represented as C1 and C2. The transactional database represented as Ti with respect to q users. Access Control Manager Application (ACM) such as Identity provider application, ACM launched $Encl_{IdP}$, An application that has the following components: $Enclock_M$ is a program running in the Spark enclave Symmetry is responsible for encrypting federal data using key encryption. It creates for each feature the provision in the Access Control Policy is a special secret, this is called Conditional Subscription Secret (CSS). The key is used to rebuild the federated decryption Information. Enclock clearly communicates CMS to data Applicant running the collaborative association mining with data EnclR of the request application through RelayACM. That too Keeps a record of all deliveries of CSS. CACM is an excellent deal for storing AC access control Policies and instructions defined by the data owner Encrypted data (simply, hash) is stored in ACM.

Proposed Model with Implementation
Data Privacy - Specific protocol guarantees the contents of the user's database are never disclosed for other users and for the Federated Cloud. Also, the final output is known only to the participants Users. We demonstrate the security of our protocol is better than recent jobs in visualised methods. User Accuracy - Our protocol always gives optimal results. It all depends on the basic steps for the end user Standard compliment algorithm, but implemented

Encrypted data. Maximum cloud participation - one of the primary goals increasing the use of our outsourcing cloud resources as much as possible. In our protocol, after users are outsourced C1's their databases do not belong to any of them the remaining steps in the specified protocol.

That is, all association rule mining activities. C1 and C2 work together so that it is max our sourcing advantage to the cloud environment. It is worth it note that users may be offline after our outsourcing they can also check their data and association rule mining Results when they are needed. So the scalability most functions in a specific protocol are free and well measured Supports large data applications.

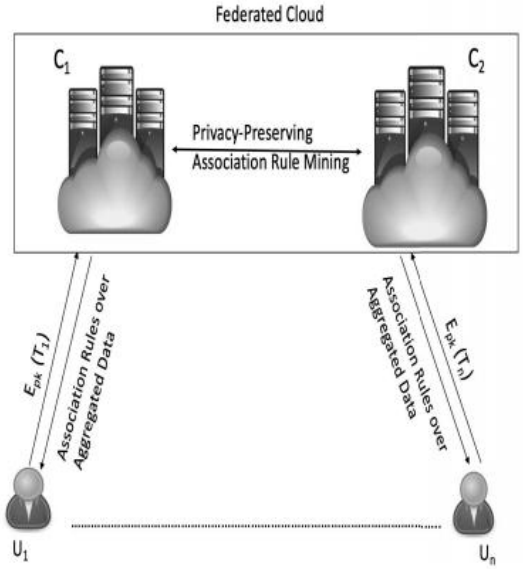


Figure 2: Proposed association rule mining for implementing Smart Grid

The above figure 1 shows that proposed framework for entities among each cloud users. 1. User u can outsource the data and set of association rule mining is applied in each C1 and C2. 2. The user group can holds all the information and set of rule mining is applied for each transactional database values. 3. The encrypted database denoted as $Encl_{IdP}$ and Privacy preserving collaborative association rule mining is applied by using below algorithm

The encryption function,

$En_{pk}(a + b) = En_{pk}(a) * En_{pk}(b) \bmod N^2$

Whereas secure modelling is applied for input values a and b, r represented as random number of each mask values

$a * b = (a + r_a) * (b + r_b) – (a * r_b) – (b * r_a) – (r_a * r_b)$

1. Secure data outsourcing Model

Our proposed model considered set of q user denoted as U1,..,Uq and apply association rule mining with aggregated data. The En_{pk} is calculated as C1 and C2. The $En_{pk}(T)$ is measured as follows

$$En_{pk}(a,b) = \prod_{i=0}^{q-1} \frac{T(ai, bi) mod\ N\ pow(2)}{N}$$

2. Secure data from frequent dataset

The Encrypted dataset has support and confidence factor denoted as α and β. Set of user U1,..,Uq and threshold values of each En_{pk} is obtained from outsource data factors.

$En_{pk}(\alpha) = \prod_{i=0}^{N-1} T(\alpha)$ and $En_{pk}(\beta) = \prod_{i=0}^{N-1} T(\beta)$

So the confidence factor is obtained from Confidence(Enpk) = α * β / N

5. Experimental Results

We measure performance of our system using Java running on Ubuntu OS with 3.3 GHz Intel i& processor, 8GB RAM. The transaction are randomly generated and following are considered for Spark evaluation. Identity Token Issue, data request is provided by the applicant Identity Properties to the identity provider, it provides an identity Token IT for each identification feature. It has a nickname to uniquely identify the name in the system Identity trait, Pederson commitment Identity attribute value. Each identification token is digitally signed

Following the identity provider and data applicant the smart contract is stored in the chain by CIdP. Signed Allows the identity provider to verify authenticity Identity token. When the data is signed by the applicant After the identity is stored in the cloud dataset as required tokens are publicly available. Without signature, another User identification tokens can be introduced in the access controls Manage and obtain unauthorized access to federal data. Identification token registration - once recognized Tokens, data can be requested from the identity provider Request access to federated data in Access Control Manager. Then, to retrieve the data decryption key, the data request must be submitted to the Access Controller Manage identification tokens that match the feature Terms of the Access Control Policy governing access Information. Then, Access Control Manager Requesting Conditional Membership Privacy (CSS) data for Name each identifying attribute of a federated organization Log in to the control system that matches the user identity Token name. User uses CSS to obtain keys decrypt shared data that satisfies access Control mechanism. The below table shows that input dataset values of different dataset values and token. The threshold values are compared with existing methods.

Dataset	Token	Parameters	Iterations	Support and Confidence	Threshold
Redshit	300	12	5	0.05	0.87
V3MD5	345	15	5	0.05	0.92
ShaF4	568	15	5	0.05	0.79
VmwT1	456	18	5	0.05	0.76
Classfish	347	21	5	0.05	0.89
Sunflash	127	24	5	0.05	0.91

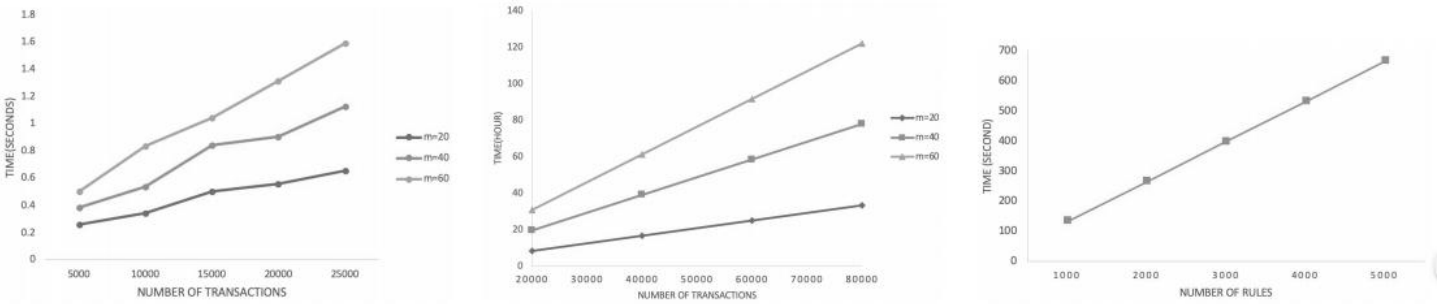
Table 1: Dataset input applied in Spark model and threshold values are obtained from our method

The same threshold values are compared with existing protocol and shown in table 2

Dataset	Support and Confidence	Iterations	Convolution Network	ToPo Specification	VSphere Method	Proposed Model
Redshit	0.05	5	0.58	0.71	0.54	0.87
V3MD5	0.05	5	0.71	0.72	0.82	0.92
ShaF4	0.05	5	0.45	0.55	0.57	0.79
VmwT1	0.05	5	0.42	0.58	0.55	0.76
Classfish	0.05	5	0.45	0.64	0.71	0.89
Sunflash	0.05	5	0.41	0.71	0.81	0.91

Table 2: Comparison result of different protocol with our proposed model.

The spark generated results for computation time, threshold values and cloud user representations. The following figure 3 shows the graphical results of each cloud user values with respect to time and transactions.



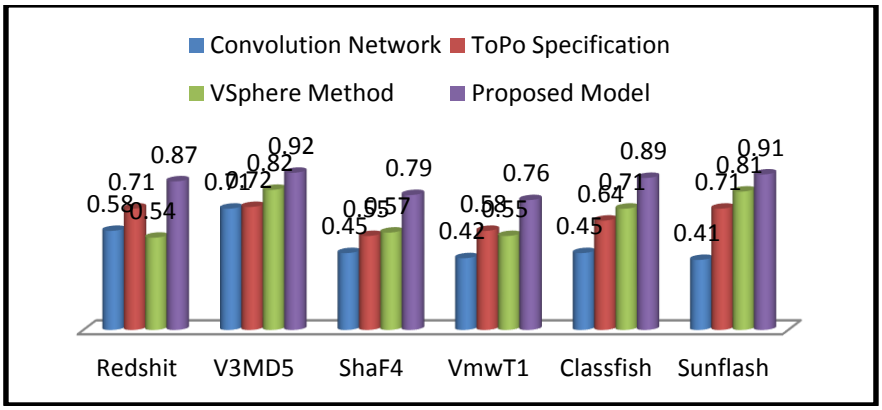


Figure 3: Computation Time, Transaction and comparison results

2. CONCLUSION

Data mining gains strength in the cloud environment companies believe that cloud computing can effectively meet their financial and performance needs processing large amounts of data. We suggest in this article Privacy-friendly framework that allows a group of users to cooperate with us on their data and association rule mining in a federated cloud environment. We have shown the proposed protocol provides a strong security guarantee compared to the current method. Also, we evaluated our protocol by demonstrating various protocols and creating low costs among end users using spark. Our solution is secure standard semi-honest model. As a future creation, we plan explore our protocol further and improve to get it Security against malicious opponents. We do extend our research to other critical data analytical task cloud encrypted data.

3. REFERENCES

[1] A. Almutairi, M. Sarfraz, S. Basalamah, W. Aref, and A. Ghafoor, “A distributed access control architecture for cloud computing,” IEEE Softw., vol. 29, no. 2, pp. 36–44, Mar. 2012.

[2] F. Zhang, E. Cecchetti, K. Croman, A. Juels, and E. Shi, “Town crier: An authenticated data feed for smart contracts,” Cryptology ePrint Archive, Report 2016/168, 2016, http://eprint.iacr.org/2016/168.

[3] M. Singhal, S. Chandrasekhar, T. Ge, R. Sandhu, R. Krishnan, G. J. Ahn, and E. Bertino, “Collaboration in multicloud computing environments: Framework and security issues,” Computer, vol. 46, no. 2, pp. 76–84, 2013

[4] Bharath K. Samanthula, Salha Albehairi and Boxiang Dong, “A Privacy-Preserving Framework for Collaborative Association Rule Mining in Cloud”, 2019 IEEE Cloud

Summit, 978-1-7281-3101-6/19/$31.00 ©2019 IEEE DOI 10.1109/CloudSummit47114.2019.00025

[5] X. Yi, F.-Y. Rao, E. Bertino, and A. Bouguettaya, “Privacy-preserving association rule mining in cloud computing,” in Proceedings of the 10th ACM symposium on information, computer and communications security. ACM, 2015, pp. 439–450

[6] C.-H. Tai, P. S. Yu, and M.-S. Chen, “k-support anonymity based on pseudo taxonomy for outsourcing of frequent itemset mining,” in Proceedings of the 16th ACM SIGKDD international conference on discovery and data mining. ACM, 2010, pp. 473–482

[7] A. Machanavajjhala and J. P. Reiter, “Big privacy: Protecting confidentiality in big data,” XRDS: Crossroads, The ACM Magazine for Students - Big Data, vol. 19, no. 1, pp. 20–23, Sep. 2012.

[8] B. K. Samanthula, Y. Elmehdwi, and W. Jiang, “k-nearest neighbor classification over semantically secure encrypted relational data,” IEEE Transactions on Knowledge and Data Engineering (TKDE), vol. 27, no. 5, pp. 1261–1273, 2015.

[9] S Manikandan, K Raju, R Lavanya, R.G Gokila, "Web Enabled Data Warehouse Answer With Application", Applied Science Reports, Progressive Science Publications, E-ISSN: 2310-9440 / P-ISSN: 2311-0139, DOI: 10.15192/PSCP.ASR.2018.21.3.8487, Volume 21, Issue 3, pp. 84-87, 2018

[10] M. Hogan, F. Liu, A. Sokol, and J. Tong, ”NIST Cloud Computing Standards Roadmap”, 2011, https://www.nist.gov/sites/default/files/ documents/itl/cloud/SP 500 293 volumeII.pdf.

[11] Shorouq Alansari, Federica Paci, Andrea Margheri, Vladimiro Sassone, “Privacy-Preserving Access Control in Cloud Federations”, 2017 IEEE 10th International Conference on Cloud Computing, 2159-6190/17 $31.00 © 2017 IEEE DOI 10.1109/CLOUD.2017.108