Block chain Based Data Provenance in Supply chain

Amrita Jyoti, R. K. Chauhan

Abstract: ***Blockchain technology as an infrastructure allows an innovational platform for a new transparent and decentralized transaction mechanism for different type of industries and businesses. Different attributes of blockchain technology increase trust through traceability and transparency ability of goods, data and financial resources within any transaction. Regardless of initial uncertainty about this technology, government and many major enterprises and firms have recently examined the adoption and improvement of this technology in several areas of applications, from social, legal and finance industries to manufacturing, design and supply-chain networks. An interesting research problem in this new era is that of determining provenance. At present, goods which are produced and transported using complicated medium supply chains, in this type of supply chain it is impossible to evaluate the provenance of physical goods. We have an interest in the blockchain as there are numerous favored use cases of blockchain especially for provenance tracking. In this paper we review the basics features of the blockchain along with its type like permission less and permissioned blockchain. Then discussed the need of provenance of assets in supplychain as it increase the trust of the customer and proposed a process and architecture for providing the data provenance in supply chain with blockchain using smart contract.***

Keywords: data provenance; blockchain technology; supply chain;

I. INTRODUCTION

A supply chain has all of the links that are involved when manufacturing and distributing goods. In today's world, dozens of geographical locations and hundreds of stages can be potentially involved by a supply chain. Like for example, over 1.6 million people in 20 countries are employed by suppliers of Apple Computing supply chain (Apple 2016). This makes it very hard to track events happening in a supply chain and investigates incidents. Because in every step of the supply chain there are information losses and barriers, the further away in the chain an incident is the harder it is to obtain any information on it. The interest in the application of blockchain in supply chain is created due to the need for more transparency, traceability and provenance of assets. All

the perennial matters that compromise the effectiveness of the supply chain can be tackled by blockchain's ability to offer transparency and traceability. Currently, in order to regain consumers trust in products an improved access to information's demand is developing [1].

Revised Manuscript Received on January 05, 2020

Amrita Jyoti, P.hd. scholar, Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, India.

Dr. R. K Chauhan , Professor, Department of Computer Science and Appliaction , Kurukshetra University, Kurukshetra, India

II. PROBLEM STATEMENT

Centuries back, supply chain was not complicated as the commerce was simple, but now the businesses have widened globally. Earlier, the businesses were limited to the local areas. People used to run minor businesses and trade homemade items locally to earn money. As an outcome, the commerce was relatively fair and transparent. It is very hard to have a complete picture of all transactions within the chains in a large supply chain system [4].At present supply chain mostly depends on emails and fax machines to send and receive contracts across the globe, resulting paper work to be slower and error prone. Blockchain can solve this by providing verifiable and immutable data sources. Provenance is the history of the transmission and ownership of an object. For several reasons professionals have been interested in the provenance of an item and the utmost important of which is that the authenticity of an item can be confirmed with the help of well-documented provenance. We proposed a smart contract in etherum Blockchain to providing the provenance of assets in the Supply Chain Industry.

III. SURVEY:BASICS OF BLOCKCHAIN

A blockchain is a time-stamped series of immutable record of data that is managed by cluster of computers not owned by any single entity. A blockchain can be seen as a distributed ledger: a chronological chain of 'blocks' where every block contains a record of the valid network activity since the last block was added to the chain [2]. Cryptographic principles (namely chain) are used to secure every blocks of data (namely block) and also bound block to each other. Blockchain can be defined as a democratized system as its network has no central authority. The information stored in blockchain ledger is open to see for anyone and everyone as the ledger is immutable and shared. Hence, anything that is built on the blockchain is transparent in nature and every person involved is accountable for their every action. Blockchain technology can be described as the technology which powers the Internet of Transactions [3].

A. Three Pillars of Blockchain Technology

Blockchain Technology has three major properties that helped it drawing acclamation from widespread which are, namely:

1. Decentralization
2. Transparency
3. Immutability

Decentralization: The decentralized nature of blockchain technology means that it doesn't depend on a central point of control. The system becomes comparatively more fair and considerably more secure due to the absence of single authority. The most revolutionary quality of blockchain is its value of decentralization that can be epitomized by the way in which data is recorded onto a blockchain.

An innovative consensus protocols is utilized by blockchain across a network of nodes, for transactions validation and recording the data in a manner that is incorruptible rather than to rely on a central authority to securely transact with other users.

Transparency: The fact that the transactions and holdings of every public address are open to view defines the transparency of a blockchain. It is possible for a user to view the transactions and their holdings that they have carried out using an explorer, and equipped with a user's public address. Within the financial systems this level of transparency has never existed before, and adds a degree of accountability that has never existed till date. Previously, without anyone's knowledge large financial institutions were able to use their customers' funds as they saw fit, and not every time in the most effective or honest matter and the epitome of this same issue is the financial crisis of 2008.

Immutability: Blockchain ledger's ability to remain unchanged and a blockchain to stay indelible and unaltered can be defined as the immutability. In short, data in the blockchain can't be changed. Every single block of information, for example transaction details or facts, proceeds using a hash value or a cryptographic principle. A hash value or a cryptographic principle is used by each block of information, with facts or transactional details to proceed. That hash value has an alphanumeric string generated by each block distinctly. Every block not only carries a hash or digital signature for itself but also for the previous one. Blocks are ensured to retroactively couple together and relentless. This functionality of blockchain technology ensures that no one can intrude in the system or alter the data saved to the block. It is also essential to know that blockchains are distributed and decentralized in nature, where a consensus is made among the several nodes that store the copy of data. This consensus ensures that the originality of data must be maintained. This lets consumers to work with the utmost degree of sureness that the chain of data is unchanged and correct [2].

B. Permissioned vs. Permissionless Blockchain

In a permissionless blockchain, transaction information is validated by public. Whereas, in a permissioned systems blockchain's owner choose a particular group and only this group is responsible for validating the transaction information. Permissioned systems are a way faster and scalable as compare to permissionless but are more centralized. The basic distinction is that you need approval to use a permissioned blockchain, while anyone can participate in permissionless systems. The original Bitcoin blockchain was and is still completely open, for example, but as companies and institutions start to adopt the technology, they're willing to sacrifice trustlessness and transparency for better access controls and easier customization. Both permissioned and permissionless blockchains have some important characteristics in common:

- They're both distributed ledgers, meaning that there are multiple versions of the same data stored in different places and connected through some type of network.
- They both use some form of consensus mechanism, which means they have a way for multiple versions of the ledger to reach an agreement on what they should all actually look like.
- They are both theoretically immutable in the sense that the data they store can't be changed without having sufficient power over the network. Even then, the blocks are linked by cryptographic hashes that will change if any data is altered.

Put simply, both permissioned and permissionless blockchains use cryptography and decentralization to varying degrees to accurately store data in a format that is difficult to hack or alter.

1. Permissionless Blockchain

Most of the blockchains you've probably heard of fall into this category: Bitcoin, Ethereum, Litecoin, Dash, and Monero. Data stored on these chains is publicly available, and full copies of the ledgers are stored all around the world, which is what makes these systems very hard to hack or censor. No one runs the blockchain, no one can restrict access to it, and you can remain relatively anonymous since you don't need to identify yourself to get an address and perform transactions. Of course, this system is far from perfect. It can be slow, difficult to build for and scale up on, *too* transparent to keep sensitive data on, hard to control access to, energy-intensive, and complex. That's why permissioned blockchains are becoming a more popular solution for companies and institutions looking to use blockchains to replace more traditional systems.

2. Permissioned Blockchain

Permissioned blockchains are only open to those who are allowed access. Anyone who wants to validate transactions and/or view data on the network has to be approved by the central authority first. This is especially useful for banks, companies, and other institutions that have to comply with regulations and might not be fans of losing complete control of their data. Instead of building on a large, decentralized blockchain like Ethereum, they can instead create a custom solution run only by institutions that they approve of [10]. The big advantages of permissioned blockchains are that they have:

- Access controls
- High customizability
- An easier time changing to comply with regulations
- Better energy-efficiency
- Potentially better scalability

There are disadvantages, too. They are:

- More centralized
- Less transparent
- More vulnerable to hacks and manipulation
- More easily censored
- Less anonymous

IV. DATA PROVENANCE

Data provenance refers to the record of the systems, entities, inputs, and processes that influences data of interest, providing a historical record of the data and its origins. A historical record of the data and its origins is provided by data provenance . Complex transformations such as workflows generate the provenance of data that is of considerable value to scientists [11]. From it, one can ascertain the quality of the data based on its ancestral data and derivations, track back sources of errors, allow automated re-enactment of derivations to update a data, and provide attribution of data sources. In the business domain provenance plays an essential role where it can be used to drill down to the source of data in a data warehouse, tracking the creation of intellectual property, and audit trail can be provided for regulatory purpose.

In distributed systems the use of data provenance is proposed to trace records through a dataflow, replay the dataflow on a subset of its original inputs and debug data flows. To do so, one needs to keep track of the set of inputs to each operator which were used to derive each of its outputs. Although there are several form of provenance, such as how-provenance and copy-provenance [12], [13] the information we need is a simple form of why-provenance, as defined in [14].

A. Type of provenance documentation

Some of the following types of provenance documentation that we probably have for the object that we own are:

- **Receipt, Invoice, or Bill of Sale:** these documents help to confirm the date that an item previously changed owners, and the identity of the parties involved, such as gallery, private owner or auction house. Whether the person owns the item they are selling can also be proven with these documents and therefore has a clear title for the object that can be legally passed to the buyer upon purchase.
- **Previous appraisal:** an object might have been appraised previously, possibly for insurance purposes or as part of an estate. Because value fluctuates over time, a previous appraisal serves to document the ownership and the age of an object, rather than the present value.
- **Illustration in an exhibition catalog from a museum or gallery:** if an item has been included in a museum or gallery exhibition, it will be mentioned and usually illustrated in a catalog published along with the exhibition
- **Inclusion / illustration in an auction catalog:** the sale results are generally accessible to the public if an item has been formerly included in an auction. The item can be illustrated in the catalog for the sales if the auction house publishes catalog.
- **Inventory number indicating de-accession from a museum or corporate collection:** items held in a corporate collections or a museum are given inventory numbers, and when they leave the collection these numbers accompany the work. They serve to verify that the work or object was part of this collection during a specific time period.

V. PROPOSED ARCHITECTURE AND PROCESS FOR PROVENANCE IN SUPPLY CHAIN

Since all types of transactions can be tracked in a transparent and secured way using blockchain, it offers many possibilities across the entire supply chain. The transactions can be recorded on the blockchain to create the account of an item from the manufacturing to sale using smart contract, for every single time a product changes hands in the supply chain.Fig 1 shows the proposed architecture for the provenance of the assets in the blockchain. Raw material's information, manufacturing details, retailer and product information and inventory update information will save in blockchain using smart contract in solidity on Ethernet network which is public blockchain [7]. For testing purpose we will use any testing network like Robstan Testing Network or Rinkyby Testing Network.

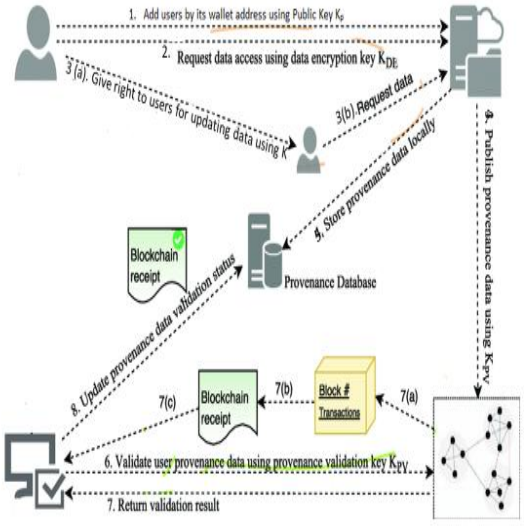

Fig.1. Proposed architecture for data Provenance in Supplychain

Figure 1 shows the proposed architecture for the provenance of the assets in the blockchain. Steps for provenance data in supplychain

1. Admin add different actors or users by its wallet address using Public Key K_p on Server
2. Admin Request Data on Server access using encryption Key K_{DE}
3. (a) Admin right to actors for updating the data for provenance as per their role
   (b) Actors update the data on server using their encryption Key K_{DE}
4. Publish Provenance data Using K_{PV}
5. Store Provence data Locally
6. Validate user Provence data using provenance validation key K_{PV}
7. (a) Create the Block Hash (# value)
   (b) Create block of transaction
   (c) Blockchain Receipt
8. Update Provence data Validation status (local server)

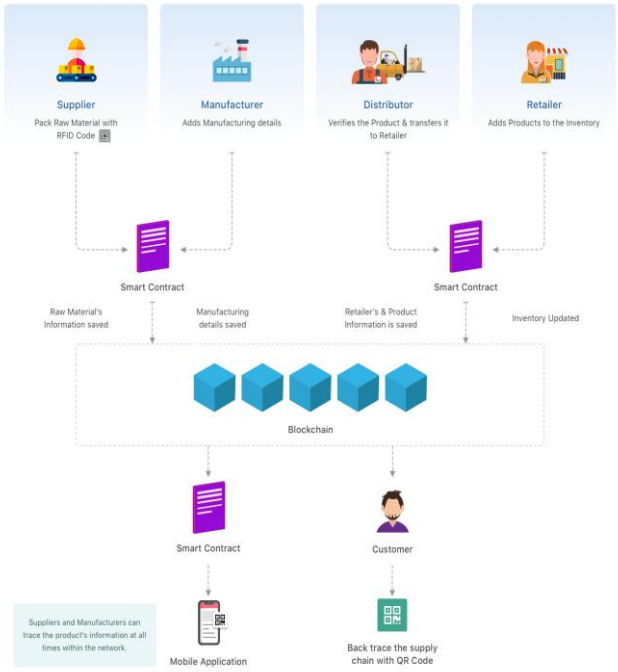

Fig 2. Usecase for Proposed architecture in Supplychain using Smart Contract

Figure 2 shows the general use case of data provenance in supply chain using smart contract. There are different actors' works as per their role. In this usecase of supplychain different actors are supplier, manufactures, retailers and distributer.

Implementing blockchain in supply chain could reduce time delays, extra costs and human errors. In the blockchain supply chain, the trust level is established by the blockchain by enabling transparency across the involved parties which has been a challenge for the last several years. Steps for creating and deploying Ethereum Smart Contracts with Solidity for provenance of data in supply chain [9].

A. Steps for Implementation of proposed approach

1) Step1: Create a Wallet at MetaMask

Install MetaMask in Chrome browser and enable it. Once it is installed, click on its icon on the top right of the browser page. Clicking on it will open it in a new tab of the browser.

2) Step2: Select a Test Network

Find the following test networks in MetaMask wallet:

- Robsten Test Network
- Goerli Test Network
- Rinkeby Test Network
- Kovan Test Network

The above networks are for testing purposes only; note that the Ethers of these networks have no real value.

3) Step 3: Add Some Dummy Ethers in Wallet

To add dummy Ethers, click on the **Deposit** and **Get Ether** button under **Test Faucet**.

4) Step 4: Use Editor Remix to Write the Smart Contract in Solidity [6]

Remix is the best option for writing smart contracts, as it comes with a handful of features. It is usually used for writing smaller sized contracts. Remix's features include:

- Warnings like Gas cost, unsafe code, checks for overlapping variable names, and whether functions can be constant or not.
- Syntax and error highlighting.
- Functions with injected Web3 objects.
- Static analysis.
- Integrated debugger.
- Integrated testing and deployment environment.
- Deploy directly to Mist or MetaMask.

5) Step 5: Create a .sol Extension File

Open Remix Browser, and click on the plus icon on the top left side next to the browser to create a .sol extension file [8].

6) Step 6: Deploy smart Contract for supply chain

Deploy the smart contract at the Ethereum test network.

B. Aim and Objectives of proposed work

Our proposed provenance works to enable businesses to build consumers trust in their goods in supply chain [5]. They leverage blockchain to generate a more clear and transparent digital record of the journey of a physical product through the supply chain. Our goal is to offer consumers improved product information, while rewarding responsible retailers and manufacturers. Proposed aim and objective can be illustrated in the following way:

- **Tracking Provenance**

To trace the history of any product from its origination to where it is delivered through the blockchain in supply chain. Such kind of traceability is helpful in detecting and resolving frauds in the supply chain process.

- **Building Trust**

Blockchain's immutable nature prevents the supply chain ecosystem from tampering while the consensus mechanism to validate every new transaction.

- **Cost Reduction**

Reduction of the extra costs, preventing frauds or counterfeits and elimination of the chances of product duplicacy can be achieved by the eliminating the intermediaries from the supply chain process. Also, rather than depending on the centralized systems, payments between the involved parties can be processed directly with cryptocurrencies

- **Reducing the complexities**

The issues related to complicated contracts can be fixed by smart contracts. For every transaction occurrence in the system, smart contracts automatically execute actions like sending alerts, payments, access to information ,transfer of ownership or other actions that do not need an intermediary.

VI. RESULT ANALYSIS

We are using a use case of rice supply chia and implemented the proposed approach for finding the provenance of data using smart contract and find the following result output.

A. Create User

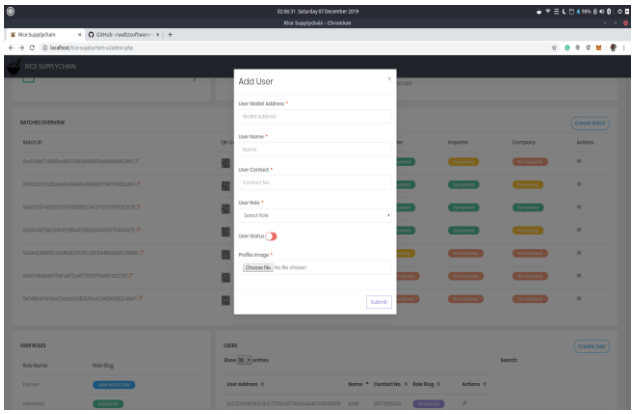

Fig 3 Create user in rice supply chain

Only admin has right to add new user in rice supply chain as show in figure 3. For add the user admin have to fill the basic information of user like user Wallet Address, username, user contact number, role of User like farmer , harvester, importer and exporter. User can play his role as per assigned right by the admin and user can update information only if he is in activated status.

B. Batch overview

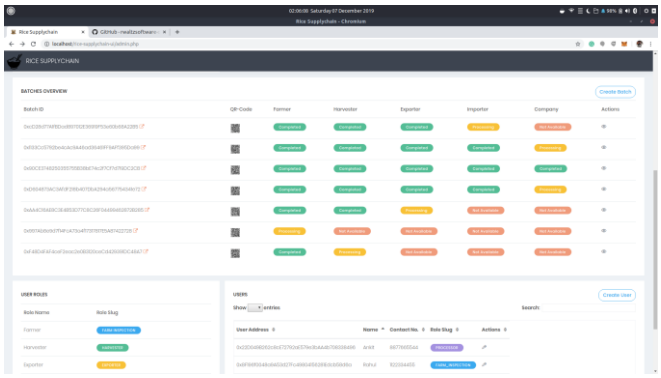

Figure 4. Batch overview

All actors like farmers, harvester, importer and exporter can fill the their respective batch detail. We can get all rice batch information in batch overview. We get the in at what stage the batch is processing as shown in the figure 4. By clicking on the read arrow at the end of batch id, admin can find out the transaction details of that particular batch transparently without any modification. Similarly admin can also scan QR-code to find out the transaction details of batch. There are different stages of batch like processing, completed or not available as depend on the current scenario of the particular batch in the rice supply chain.

C. Batch Details

In View Batch Page,/ admin will be able to see the progressive information of rice batch. Here we can get all details of each stage and also the name and address of user who updated the particular stages this is our primary objective to find the provenance of each batch of rice in rice supply chain without any modification by any actor or user because each and every data store in blockchain using the smart contract.In this way, we can track the progress of riceafter each stage in blockchain. The stages which are yet not updated in blockchain are denoted using cross sign and the stages which are completed are denoted by right tick sign. You can also find out the name, address and contact information of user who updated the particular stage in rice supply chain.

VII. CONCLUSION

We described and evaluating assets provenance as an important and ongoing issue in supplychain. Evaluating knowledge provenance has become more possible as more and more of the data required discovering the source of knowledge is recorded on the supplychain. Evaluating provenance in the supply chain has generally been more complex and hard task because so many goods are handled in international supply chains where the all the records maintain has not been possible. That is, until recently, when provenance evaluation has become more possible with the help of Blockchain technology and smartcontract.

In particular, as blockchain technology develops, as more business models conceived that hold it and as well as more researchers explore research opportunities with its use, we believe that the smart contract for data province can make a contribution to the growth of blockchain.

REFERENCES

1. G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
2. F. Mechthild and T. Ludwig, “Transparency in Supply Chains: Is Trust a Limiting Factor?,” 2006. [Online]. Available: http://ageconsearch.umn.edu/bitstream/7733/1/sp06fr01.pdf
3. S. Bogart and K. Rice, “The Blockchain Report: Welcome to the Internet of Value,” 2015.
4. M. Mainelli and A. Milne, “The Impact and Potential of Blockchain on the Securities Transaction Lifecycle,” 2016. Available http://www.zyen.com/now-and-zyen/blog/1516-the-impact-and-potential-of-blockchain-on-the-securities-transaction-lifecycle.html
5. Haq, I., Monfared, R.P., Harrison, R., Lee, L., and West, A., "A new vision for the automation systems engineering for automotive powertrain assembly," International Journal of Computer Integrated Manufacturing (IJCIM), vol. 23, pp. 308-324, 2010.
6. Allison, I. (2016). Provenance has a big year ahead delivering supply chain transparency with Bitcoin and Ethereum. [online] IBTimes. Available at: http://www.ibtimes.co.uk/.
7. Dannen C., 2017, Introducing Ethereum and Solidity: Foundations of Cryptocurrency and Blockchain Programming for Beginners, DOI 10.1007/978-1-4842-2535-6, Springer Science & Business Media New York, New York, USA.
8. Ethereum, 2017, *Solidity* [www], Tillgänglig: https://solidity.readthedocs.io/en/v0.4.21 Hämtad
9. Yann300, 2017, *Remix – Solidity IDE* [www], Tillgänglig: http://remix.readthedocs.io/en/latest/ Hämtad
10. Github.com, (2017), Ethereum White Paper. [online] Available at: http://www.citethisforme.com/harvard-referencing
11. T. T. A. Dinh, J. Wang, G. Chen, R. Liu, B. C. Ooi, and K.-L.Tan, “Blockbench: A framework for analyzing private blockchains,” in *Proceedings of the 2017 ACM International Conference on Management of Data*, ser. *SIGMOD ’17. New York, NY, USA: ACM,* 2017, pp.1085-1100[Online].

12. Yogesh L. Simmhan, Beth Plale, and Dennis Gannon. A survey of data provenance in e-science. SIGMOD Rec., 34(3):31–36, September 2005.
13. Peter Buneman, Sanjeev Khanna, and Wang-Chiew Tan. Data provenance: Some basic issues. In Proceedings of the 20th Conference on Foundations of SoftwareTechnology and Theoretical Computer Science, FST TCS 2000, pages 87–93, London, UK, UK, 2000. Springer-Verlag
14. Robert Ikeda and Jennifer Widom. Data lineage: A survey. Technical report, Stanford University, 2009.
15. Y. Cui and J. Widom. Lineage tracing for general data warehouse transformations. VLDB Journal, 12(1), 2003.

AUTHORS PROFILE

Amrita Jyoti, Ph.D. research scholar Kurukshetra University, Kurukshetra India and presently working as Associate professor in ABES Engineering College, Ghaziabad Uttar Pradesh, India. She received her B.Tech. degree in Information Technology from Kurukshetra University, India in 2003 and M.Tech. in Computer Science & Engineering from the Dr. A. P. J. Abdul Kalam Technical University, Uttar Pradesh, Lucknow, India in 2011. In 2005, she joined the Department of Computer Science & Engineering, ABES Engineering College, Uttar Pradesh, India as a lecturer and became an associate professor in 2015. She is the author of a Book ‘‘Title: Data compression’’ which has covered all the techniques of compression in text, audio and video. Her current research area include software Engineering, software testing, data compression, data structure, JAVA , cloud computing and Blockchain. She published many research paper in National and International Journals and presented many papers in various National and International Conferences.

Dr. R. K. Chauhan, is the oldest founder faculty member as well as senior most professors in the deptt of computer science & applications in Kurukshetra University, Kurukshetra. He obtained the doctor of philosophy in computer science from the Kurukshetra university. Under his supervision, fourteen scholars have been awarded Ph.D. degree on different areas of computer science and applications and three scholars are pursuing their research work. He has also guided more than fifty M.Tech desertion. He has affended and participated in numerous professional conferences meetings and has published more than hundred research papers in national and International journals.He was awarded 6 merit certificate for best research paper in Dec 1998 by Institution of engineers(India). He held the position of chairman of Deptt of Comp. Sci and Applications from Dec 2007 to Dec 2010. His research area include Advance Database, Data Mining & Warehousing, Mobile Computing, Ad-hoc Networks and Software Engineering. He has been the member of various academic and administrative bodies of Kurukshetra University and other University.

.