

ALL INDIA SHRI SHIVAJI MEMORIAL SOCIETY'S
INSTITUTE OF INFORMATION TECHNOLOGY,PUNE
Department of Computer Engineering

Technical Magazine 2017-18

Stand out from the crowd



*

Objective behind Technical Magazine

Department of Computer Engineering is very happy and proud to publish technical magazine of year 2017-18. We have gathered technical articles from industry experts working in very renowned IT industries. These articles gives guidelines to students regarding what is expected in IT industry and how various technologies are applied for the projects in IT industry. Through this technical magazine we are sharing these articles and quality of final year projects of year 2017-18.

Department has set objective to bring technical competency among the students. Department is taking efforts for the same since second year of these students. Department arranges various expert lectures,workshops,industrial visits,learning contents beyond syllabus for the students. All these activities are planned to make students aware of current need of IT industry. Outcome of these efforts is reflected through their final year projects ,placement and admission to higher studies.

We had contacted alumni who are working in various IT industries. We are sharing experience of these alumni with our students through this magazine. Our objective behind sharing this information is to motivate students and to create awareness among them about current need in IT industry.

Coordinator

S.P.Pimpalkar

HOD

S.N.Zaware

INDEX

Topic

Current Trends in IT industry :

1. Implementation of Technology in IT projects
Diksha Gulati,TechMahindra
2. I know your choices, by your choice
Omkar Gurav ,Quick Hill
3. Introduction to Blockchain Sonali Patwe,IBM
4. Need Of SQL And PL/SQL In IT Industry
Sayali Pawar- Zensar
5. Evolution of Big Data with Cloud Computing
3RI Technologies ,Pune

Final Year Projects :

6. Secure and Efficient Multi-Keyword Ranked Search Method on Cloud (Chetna Raut,Nayana Dashmukhe,Anushree Gawande,Neha Fegade)
7. Preventing Document Frauds Using Smart Centralized Qualification Card.(SCQC)
(Pratik Lodha,Parag Sanyashiv,Shubham Kadlag, Sunil Jagtap)
8. iBeacon Based Smart Attendance Monitoring System.
(Shubham Jain,Shweta Kale,Shraddha Kannurkar)
9. Stratification of Dengue Fever using SMO and NSGA-II optimization approach
(Poonam Doddamani,Preet Dalsani,Aishwarya Dixit,Swarada Deshpande)
10. Smart Agriculture System Based On Data Analytics.
(Omkar Buchade,Chandan Mehta,Shubham Ghodekar)
11. Stock Prediction Using Machine Learning Techniques
(Jenish Karia,Muskaan Khan,Satyam Anand,Tushar Mukherjee)
12. Data Compression using Wireless Distributed Computing
(Bhushan Mali,Siddhesh Pandey,Ivan Pillay,Shubham Soni)
13. Email Correlation And Analyzer Using Pattern Matching Technique
(Ashwini Kawade,Pratima Khambat,Shraddha Bhagwat,Shrirang Potale)
14. Attendance Monitoring System Using Image Processing and Machine Learning (Manasi Jadhav Tushar Indalkar Kirtiraj Kadam Shivam Mulay)

Implementation of Technology in IT projects

Diksha Gulati, TechMahindra

Software engineering – Just a theory subject, I'll read and clear it, OOPs (Object oriented Programming) concepts- I am good with coding, do I need my basics to be clear? These are a few questions that popup in a student's head, moreover in the final year , when we are just one milestone away from achieving what we have worked for since the past 22 years. And once we have crossed that milestone and got a job, that is when we realize, "I wish I had understood my concepts better". Personally, theory subjects were my weak point, but now when I have been working for 2 years, I realize how important "AGILE" is. When I write a piece of code and see my seniors suggest and discuss solutions for the same, I understand how important it is to write clean, optimal code, code that will never hamper the performance, and how redundant code is definitely not acceptable. Ever thought of testing your code? Or analyzed probable use cases for the same? Or are we just going to code for the given problem statement? Is the solution printed in our file the only way to code for that program or can we not discuss and think out of the box? I understand that college, submissions, lectures don't really leave us with a lot of time, but if a year or 4 years down the line we sit at our office desk and look at others suggest better solutions, code in a smarter way, is it really going to make us happy ? Engineers are somebody who can not only improve the efficiency of work but also help improve human lives by saving time, resources, money and the quality of the decisions taken.

In IT, we implement various projects with the underlying objective of simplifying daily business processes, improving the efficiency of work and

reduce the costs with no impact to the existing business. For instance, in a pharma company, we may have a project with IT wherein we want to standardize a process to analyze the sales of the medicines to take a decision regarding which is the drug most widely accepted by the consumers which would further impact the company decisions of where they need to invest more money for research & development. Such decisions then not only remain as a scope of a project but the correct implementation of this in turn determines the further directions such as the amount of investment to be made in the research of life-saving drugs, the talent to be employed for this research, the activity of development of the medicines along with the investment in the sales & marketing of this product. Thus we see how the application of the right technology can in turn not only impact the following activities but in a larger picture help fight critical ailments like cancer, etc.

Implementation of the projects in IT move around various depending factors such as delivering accurate algorithms, standardizing processes, implementing global protocols, improving system performance & assuring business of getting the correct data in a timely fashion. These objectives turn out to be further challenging when we have global expansion of business, offshoring resources and with budget & headcount issues. The important thing about IT is that irrespective about the kind of service the company provides, all have an IT department as all companies need infrastructure and stable hardware & software platforms.

Projects need to deliver the requirements in a fixed time period and the technology and protocols used to deliver the solution have a deep impact during the complete project lifecycle. We need to be careful of the infrastructure we use which is compatible with the platform on which we are implementing the

solution. Implementing small details like caching data or using a higher bandwidth network can finally come to making huge impacts. The efficiency in requirement gathering phase can often impact the smooth operation of all the phases that will follow. We need to take into account the most efficient testing approach where IT can fix bugs in the development before it is in the production environment and can cause huge business impacts like delayed shipments or wrong invoices which would in turn lead to dissatisfied customers and spoil the brand name. Thus accurately planning all the lifecycles taking into account the `Six Sigma` rules can often lead to huge success rates.

Another important aspect to consider is that whatever we implement has to be compliant and adhere to the legal requirements which will vary differently for each company as well as each country. The IT projects often have government restrictions binding on them and have internal company as well as external audits. Using non-compliant procedures can lead to huge penalties as well as termination of business based on the level of non-compliance. Hence, we need to ensure that the IT team is contacting the compliance team on a regular basis during the complete phase of the project.

Projects have tough deadlines and ensuring that we adhere to compliant standards and achieve the highest quality of the deliverables with satisfied customers is what truly comes to define the success of the project. In order to monitor this we need to ensure we develop several Key Performance Indicators during the various phases of the project and ensuring an efficient business operation model with good knowledge of the technology is what can truly determine the success of the project.

The right choice of resources and technology can hold the power to either take a business forward or make it decline rapidly. Today, we are living in a world where technology has advanced, the competition is strong and making important decisions in a short span of time – are some factors that can have huge business impacts. We are at a point, where we are learning new things, but we are not sure, rather unaware of how apply it. Try to update yourself with the latest technologies and that doesn't mean that you sit and read your entire "Java Reference". See what solutions the market has to offer, explore the web world, shape your thinking in a way that you will be able to think considering all the aspects and impacts. Think of the chain that will follow once you are considering a solution (What leads to what, what will be the final outcome and not just a temporary solution). Make your contribution worthwhile, come across as one platform that works for the people and delivers solutions effectively.

"Scientists study the world as it is, Engineers create the world that has never been."

Quoting Godfrey Reggio, ***"It's not that we use Technology, We live Technology."***

- Deeksha Gulati

(2016 Batch)



I know your choices, by your choice!

Omkar Gurav ,Quick Hill

Can you guess a feature over the internet which has something to do with the title?. Let me give you some hint along besides the image given above. This feature is related to the browsers we use to surf the internet. You may have seen a popup asking for a permission to remember you, your settings while surfing web pages. I am sure until now you must have given a thought about Chrome and Firefox. Yes, you were on the right track and it was the **cookies** which we would be discussing. I will the keep the technical jargons as simple as possible

What is a cookie?

A cookie is a small piece of text stored on a user's computer by their browser. Common uses for cookies are authentication, storing of site preferences, shopping cart items, and server session identification.

What is a session?

To make you understand few concepts let us consider an example. In a class, a teacher is resolving a student's query. Let us assume the following would be the steps before a conversation between the student and the teacher.

=====Session started=====

Student: Excuse me, I have a doubt teacher, could you please explain this to me again?

Teacher(Looking at the student):Yes, sure! Let me know which part exactly confused you.

And so on,

=====Session ended=====

Now technically speaking the teacher here was a server which acknowledged the request of the client student and maintained the same topic in which the student was confused. The same happens when you request for a page on the internet to the specific server. The server maintains details about who you are and in what you are interested.

How to preserve your choices? Cookies are the answer.

Now continuing the same example let us consider that the next day the student had some further doubt on the same topic. The teacher and the student had a conversation on the same topic. This time the steps would be:

=====

Student: Hello teacher, I further read about the topic you taught yesterday and have some few more questions about it.

Teacher: Ok let's discuss them.

=====

If you have noticed the difference, here the teacher(server) already had information about the student's topic(client) and hence did not need to start the conversation over again. This information that is needed to start the conversation again is stored in the form of cookies, **session cookie** especially in the above example, on the client machine via the browser.

Now unknowingly, you are already aware of what a cookie is and a special type of cookie as the session cookie.

Another use of cookies is to know more about you.

Example- when you visit YouTube and search for Bollywood songs, this gets noted in your browsing history, the next time you open YouTube on your browser, the cookies read your browsing history and you will be shown Bollywood songs on your YouTube homepage

Types of cookies

Session Cookie

This type of cookies dies when the browser is closed because they are stored in browser's memory. They're used for e-commerce websites so a user can continue browsing without losing what he put in his cart. If the user visits the website again after closing the browser these cookies will not be available. It is safer because no developer other than the browser can access them.

Persistent Cookie

These cookies do not depend on the browser session because they are stored in a file of browser computer. If the user closes the browser and then access the website again then these cookies will still be available. The lifetime of these cookies are specified in cookies itself (as expiration time). They are less secure.

Third Party Cookie

A cookie set by a domain name that is not the domain name that appears in the browser address bar these cookies are mainly used for tracking user browsing patterns and/or finding the Advertisement recommendations for the user.

Secure Cookie

A secure cookie can only be transmitted over an encrypted connection. A cookie is made secure by adding the secure flag to the cookie. Browsers which

support the secure flag will only send cookies with the secure flag when the request is going to an HTTPS page.

HTTP Only Cookie

It informs the browser that this particular cookie should only be accessed by the server. Any attempt to access the cookie from client script is strictly prohibited. This is an important security protection for session cookies.

Zombies Cookie

A zombie cookie is an HTTP cookie that is recreated after deletion. Cookies are recreated from backups stored outside the web browser's dedicated cookie storage.

That's all for this article and hope you would be excited to explore more about cookies.

Helpful links and sources:

<http://www.hackingarticles.in/beginner-guide-understand-cookies-session-management/>

<http://www.lassosoft.com/Tutorial-Understanding-Cookies-and-Sessions>

Introduction to Blockchain

Sonali Patwe,IBM

In today's digital age, it is very common to record the exchange of goods and services. As we evolve and more and more systems are being digitized, recording of the transactions become more complex whether it exchange of money between 2 parties, documenting the movement of goods across supply chain or making contractual agreements. The existing centralized systems like banks, SCM, Government are vulnerable to errors, frauds and misinterpretation. For example, diamond industry. When a diamond goes from a mine to the consumer, it travels through a complex landscape of legal, regulatory, financial, manufacturing and commercial practices. Current supply chain has to rely on intermediaries on every step of the way from government officials to lawyers, accounts, dealers and banks. This adds time and cost. There is also possibility of frauds. As multiple parties are involved in the process, it becomes difficult to track and the complete system is not transparent. This is where Blockchain technology comes into picture. It has the potential to eliminate such vulnerabilities with transparent transactions. Blockchain offers all parties involved in the transaction a secure and synchronized record of transactions. The blockchain ledger records every sequence of transaction from beginning to end whether it is hundreds of steps in supply chain or single transaction of online payment. As each transaction occurs, it is put into the block. Each block is connected to the one block before and after that block in the chain. Groups of transactions are blocked together and fingerprint of each block is added to the next, thus creating irreversible chain. That's why Blockchain is ideal for recording the transactions of valuable goods and services. There are thousands of use cases where blockchain can disrupt the existing implementations of the systems and provide more transparency, trust and security.

The blockchain, the underlying technology, is the biggest innovation in computer science—the idea of a distributed database where trust is established through mass

collaboration and clever code rather than through a powerful institution that does the authentication and the settlement.

The way it works is, if person A owes Person B \$20, we do the transaction. There's a huge community called miners, and they have a powerful computing resource. Some people have estimated that the entire computing power of Google would be 5 percent of this blockchain-computing power, for the Bitcoin blockchain. These miners go about authenticating that the transaction occurred.

Each miner is motivated to be the first one to find the truth, and once they find the truth, it's evidence to everybody else. When you find the truth and you solve a complex mathematical problem, you get paid some money, some Bitcoin. If someone wants to solve the cryptography puzzle as a miner, then he has to hack the ten-minute block. That's why it's called blockchain, and that block is linked to the previous block, and the previous block—ergo, chain. This blockchain is running across countless numbers of computers.

The unique features of Blockchain like distributed, permissioned and secure make it capable of handling regulated industries like diamond supply chain. Because the ledger is shared, it works as a shared form of record keeping instead of no one person or organization holds the ownership of the system. As a diamond cycles through its supply chain, everyone involved during the process is permissioned to have a copy of every record and no transaction can be added to the blockchain without consensus of the participants, this gives the guarantee that no one person / party can add or alter the records in the blockchain. This makes the blockchain-enabled applications tamper-proof and very secure eliminating the risk of fraud and errors. The records in the distributed ledger are auditable and indisputable information. Blockchain technology gives us the ability to transform the industries of all sorts like Supply chain, healthcare, governance, finance, etc. Blockchain technology frees up capital flows, speeds up processes, lowers transaction cost in a secure and trustworthy way. Experts say that Blockchain will do to the business

what Internet did to the communication. It will be really exciting to see how far this technology goes and what real-world problems it solves.

Need Of SQL And PL/SQL In IT Industry

Sayali Pawar- Zensar

Oracle is combination of both **SQL** and **PL/SQL**. To become an **Oracle DBA**, one probably needs to learn both of them.

The world of data is constantly changing and evolving every second. This in turn has created a completely new dimension of growth and challenges for companies around the globe. By accurately recording data, updating and tracking them on an efficient and regular basis, companies can address their challenges on one hand and make use of the immense potential offered by this sector on the other hand.

What is SQL?

The SQL language is written to comb the contents of tables in a conventional database. SQL is widely used in business and in other types of [database administration](#). It is the default tool for “operating” on the conventional database, to alter tabled data, retrieve data or otherwise manipulate an existing data set.

SQL may be considered as the source of data for our reports, web pages and screens.

What is PL/SQL?

PL/SQL can be considered as the application language similar to Java or PHP. It might be the language used to build, format and display those reports, web pages and screens.

PL-SQL is a **programming language SQL**, used to write full programs using variables, loops, operators etc. to carry out multiple selects/inserts/updates/deletes.

If you have good SQL skills and you know PL/SQL well, you can do pretty much anything with an Oracle database, including all types of web service and web development.

The value of a skill depends on the relationship between the number of people with the skill and the number of jobs requiring that skill. For PL/SQL, the amount of work is constant, but the number of people with PL/SQL skills is declining.

Oracle Applications

Oracle applications is an ERP package. So it is important to use for data integration

1. Master data is integrated
2. Financial data is integrated
3. Transactional data is integrated.

As in any IT industry maintenance of data is very much important. The data related to all the associates, their payroll need to be updated on monthly basis.

Oracle provides the database at enterprise and express editions. Oracle provides some in built schemas to practice like HR schema.

So for the betterment of the industry most of the companies work for oracle applications.

Evolution of Big Data with Cloud Computing

3RI Technologies ,Pune

Cloud computing has become the de facto platform on which enterprises are fueling digital transformations and modernizing IT portfolios. Organizations are increasingly finding business agility or cost savings by renting software through vendors such as Amazon Web Services (AWS), Microsoft, Google and IBM.

The global public cloud market will hit \$178 billion this year, up from \$146 billion in 2017, according to a Research. Public cloud adoption in enterprises will cross 50 percent for the first time. With so many large organizations offloading compute resources to focus on strategic digital initiatives, the tipping point was inevitable.

Business and IT executives are no longer looking at the Cloud solely as a tool; now the focus has shifted towards finding the right way to use it so they can accomplish their business goals. The advent of the Cloud has created significant changes to organizations in the past few years. Cloud Computing has provided Big Data with a way to store and retrieve an immense amount of information. It has evolved from personal cloud storage to entire organizations moving all of their data to the cloud.

Hadoop has become a leading platform for big data analytics today. Hadoop-based applications are used by enterprises which require real-time analytics from data such as video, audio, email, machine generated data from a multitude of sensors and data from external sources such as social media and the internet. Hadoop-as-a-service enables technical experts of organizations to perform several operations which include big data management, big data analytics and big data storage in a cloud. The HaaS platform enables organizations to use Hadoop technology in a highly cost-effective manner, along with ensuring minimal consumption of time.

Before we begin, it's important to roughly understand the three components of Apache Hadoop project:

1. Hadoop Distributed File System (HDFS) is a distributed file system based off Google File System (GFS) that splits files into blocks and stores them redundantly on several relatively cheap machines, known as DataNodes. Access to these DataNodes is coordinated by a relatively high-quality machine, known as a NameNode.
2. Hadoop MapReduce (based off Google MapReduce) is a paradigm for distributed computing that splits computations across several machines. In the Map task, each machine in the cluster calculates a user-defined function on a subset of the input data. The outputted data of the Map task is shuffled around the cluster of machines to be grouped or aggregated in the Reduce task.
3. YARN (unofficially Yet another Resource Negotiator) is a new feature in Hadoop 2.0 that manages resources and job scheduling, much like an operating system. This is an important improvement, especially in productions with multiple applications and users, but we won't focus on this for now.

Some of the ways we may see Cloud computing change:

Trend 1: Cloud as the IT norm

Scalability has always been the backbone of agile enterprise computing. Salesforce an integration partner of ours, mentions that moving to the cloud has numerous benefits, including security, collaboration, accessibility, document control, environmental friendliness, and more.

Smart file storage, audit-ready data storage, secure Cloud backup, and federated IT architecture rely on scalability in order to adapt to rapid changes. This scalability is expected to be supported by:

- Innovations in Cloud-enabling hardware and hardware virtualization
- A gradual phasing out of dedicated server storage and backup
- Dedicated virtual server environments offering nearly unlimited scalability

Trend 2: Software to Service

One of the crucial duties of IT services is the scheduled backup and retrieval of enterprise data, as it's one of the most important trends in the future of cloud computing.

While the importance of these functions to the risk management function of an enterprise cannot be underestimated, important core functions including development, implementation, deployment, and document management will move from being run on proprietary software to running entirely on a Cloud platform.

While Software as a Service (SaaS) models have been embraced by large enterprises, cheaper Cloud services are expected to make SaaS and Platform as a Service (PaaS) models accessible to small developers and companies.

Small and medium-sized businesses will be able to leverage significantly cheaper Cloud services that offer enterprise-level services at a fraction of the cost.

The cloud has been around much longer than most people give it credit for. Although it's only come into popular use within the past 10 years, it's been around since the 1960s. Additionally, some people are uncomfortable with the idea of cloud computing, but these beliefs are largely unfounded.

Trend 3: Social and Modular Capabilities: Important Trends in the Future of Cloud Computing

As more IT firms migrate to a Cloud-based platform environment, we can expect to see innovative cross-platform, pan-industry platform interfaces that allow independent companies to sell and operate their Cloud-based services as a module.

For example, an IT company building a Cloud-based software can 'write in' a module that is compatible with a remote file storage and backup service such as eFileCabinet and other Cloud-based service providers to offer a bespoke service platform solution for specific industry types.

Additionally, resource elasticity will be another definitive trend of the near future in cloud computing. One of the traits of this movement will entail internal speed collaboration only associated with that of external resource sharing.

And entire workflow processes will be allocated around the existing resources of the cloud's infrastructure. Additionally, more traditional platforms, including those of the document management variety, will be transposed to the level of the operating system. And the future of cloud interoperability will hinge on how various clouds interact with one another at the level of this operating system.

Most experts predict this is one of the most pressing trends in the future of cloud computing, as it will change the way organizations leverage their managed services, including the depth of which these managed services will have within the organizations themselves. Instead of comprising a small component of storage and collaboration, cloud environments will subsume large business processes, and with increased scale.

Project Title :”Secure and Efficient Multi-keyword Ranked Search Method on Cloud”

(Chetna Raut Anushree Gawande Nayana Dashmukhe Neha Fegade)

ABSTRACT

The extensive use of cloud services has resulted in dramatic growth in volume of data which has made information retrieval much more difficult than before. Even text documents are encrypted before being outsourced to cloud servers. This helps to protect user’s data privacy. Existing techniques to search over encrypted data are not suitable for a huge data environment. Due to the blind encryption, relationship between the documents is concealed which further leads to search accuracy performance degradation. Therefore it is necessary to adopt an approach to support more search semantic and for fast search within enormous data. A hierarchical clustering method for cipher text search is proposed in this paper. The proposed approach clusters the documents based on the minimum similarity threshold, and then partitions the resultant clusters into sub-clusters until the constraint on the maximum size of cluster is reached. The very first thing which cloud server does is that it performs search activity and then selects the k documents (previously decided by the user and sent to the cloud server) from the minimum desired sub-category. To request desired documents, instead of using single keyword search query, multi-keyword search technique is proposed.

Architecture:

Data owner: The data owner performs following tasks-

1. selects the collection documents $d_i = \{d_1, d_2 \dots d_m\}$ to be stored on cloud.
2. Builds the secure searchable index tree (I) of these documents.

3. Generate the encrypted document collection (C) for these documents.
4. At last uploads the index tree (i) and collection of encrypted documents(C) on cloud server.
5. Distributes the secret key (sk) to authorized user which is required to generate the trap door.
6. Also shares the decryption key (K) with the user.

Following things need to be considered while performing above tasks are-

1. Maintain the capability to search on encrypted documents.
2. In case the documents are updated the required information is stored locally.

Data users

The authorized users can access the documents stored on cloud by data owner. They are the Data users. With query which can be single keyword or multi-keywords (w), the authorized user can generate a trapdoor T_w . This is done according to search control mechanisms to fetch “k “encrypted documents from cloud server. The trapdoor is forwarded to cloud. The value of k as mentioned earlier is user defined. After fetching the result, the user now decrypts the documents using the Key (K) provided by data owner

Cloud server

Cloud server stores the -

1. Encrypted document collection (C)
2. Encrypted searchable index tree I form data owner.

Upon receiving the trapdoor T_w from the data user

1. The cloud server executes search over the index tree I

2. Responds the corresponding collection of top- k ranked encrypted documents to the user.

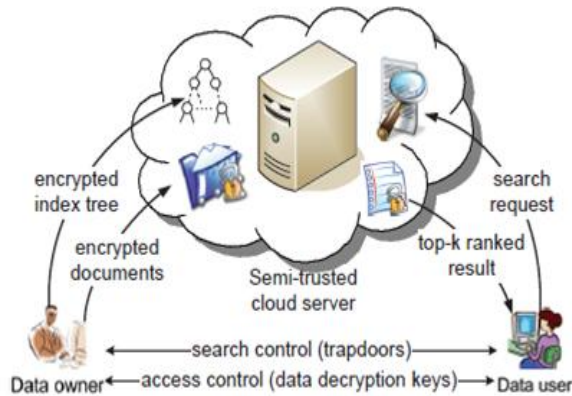


Fig: Ranked Search Over Encrypted Data

CONCLUSION

We propose a safe and efficient search scheme. It guarantees the accurate multi-keyword ranked search. We explored the problem of sustaining the relations among various plain documents over the correlated encrypted documents and give the design technique to improve the performance of the search on these documents. For data explosion and for retrieval of online information MRSE-HCI architecture is used. The documents retrieved by the user is also verified which maintains the integrity.

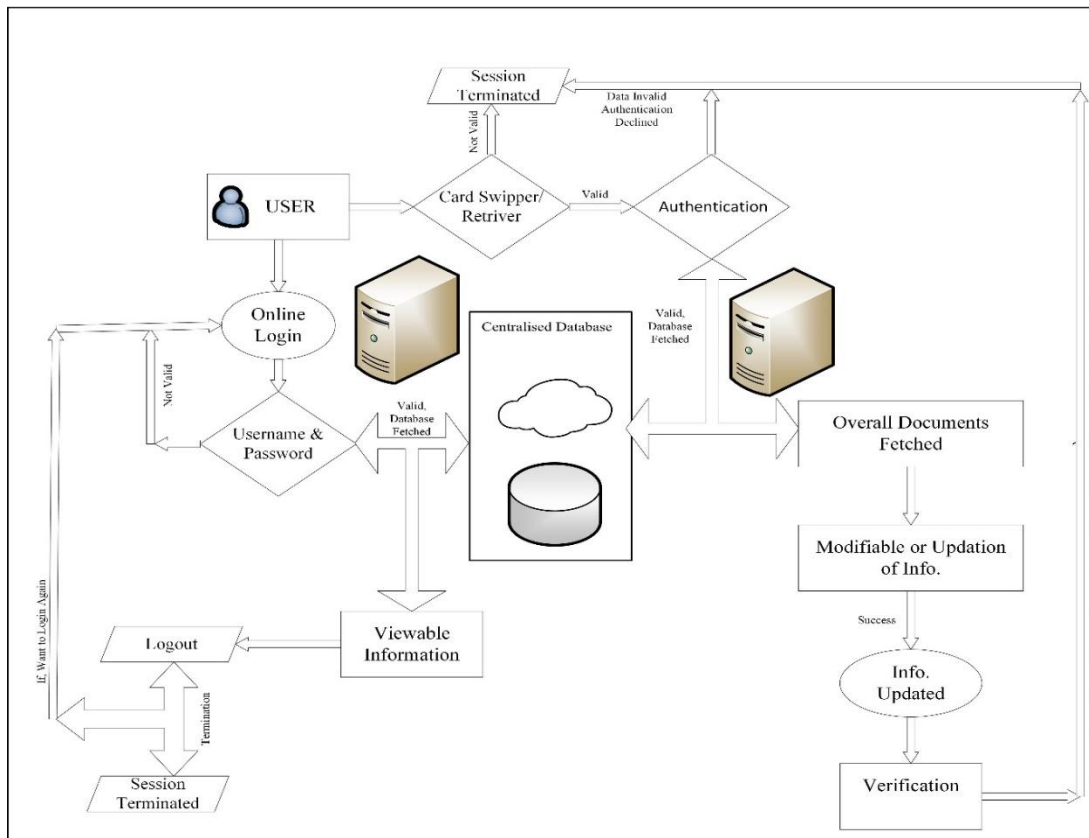
Project Title :”Preventing Educational Document Frauds using Smart Centralized Qualification Card (SCQC).”

(Pratik Lodha,Parag Sanyashiv,Shubham Kadlag, Sunil Jagtap)

Abstract:

Even though world is focusing on being digital still large amount of information or the data is represented by the papers only and that may also contain critical data. Now a days official documents are strongly secured with modern techniques such as artwork or printed pattern, but still paper documents suffer from a lack of security it means that with the high availability of cheap scanning and printing hardware, non-experts also can easily create fake documents. For the many organizations prevention of fraud has become a major concern. The industry understood the problem and is just now starting to act by innovation of new techniques for fraud detection. But as always said prevention is better than cure, we will focus on prevention of frauds, it is the best way to reduce frauds, fraudsters are adaptive and will usually find ways to break the security and circumvent such measures. So, for that purpose we have to make prevention system more and more stronger. In this paper, we present convenient and much more efficient approach for fraud prevention. In our approach, we use the smart card (can be RFID, magnetic tape based etc.). In this smart card, all the educational details can be linked so that the details can be obtained on the go, and hence there is no worries related to carrying your documents all the time with you. So, the issues related to the forgery or alteration of original documents will be reduced (or you can say fully eliminated!)

Architecture Diagram:



Conclusion:

In this way, we have proposed Smart Centralized Qualification Card (SCQC) design which will fully eliminate the document frauds as all the original documents are directly linked to the SCQC by the respective university or the board. At each retrieval time, the data is passed via encryption and decryption that makes a secured communication that the chance to change/ attack details from the database is low.

In other way, it also reduces time to verify the documents. Implementation of our proposed system takes a step towards the digitalization i.e. it makes the work paperless by providing all the original documents without tempering on just one scan of the SCQC.

Project Title :”iBeacon Based Smart Attendance Monitoring System.”

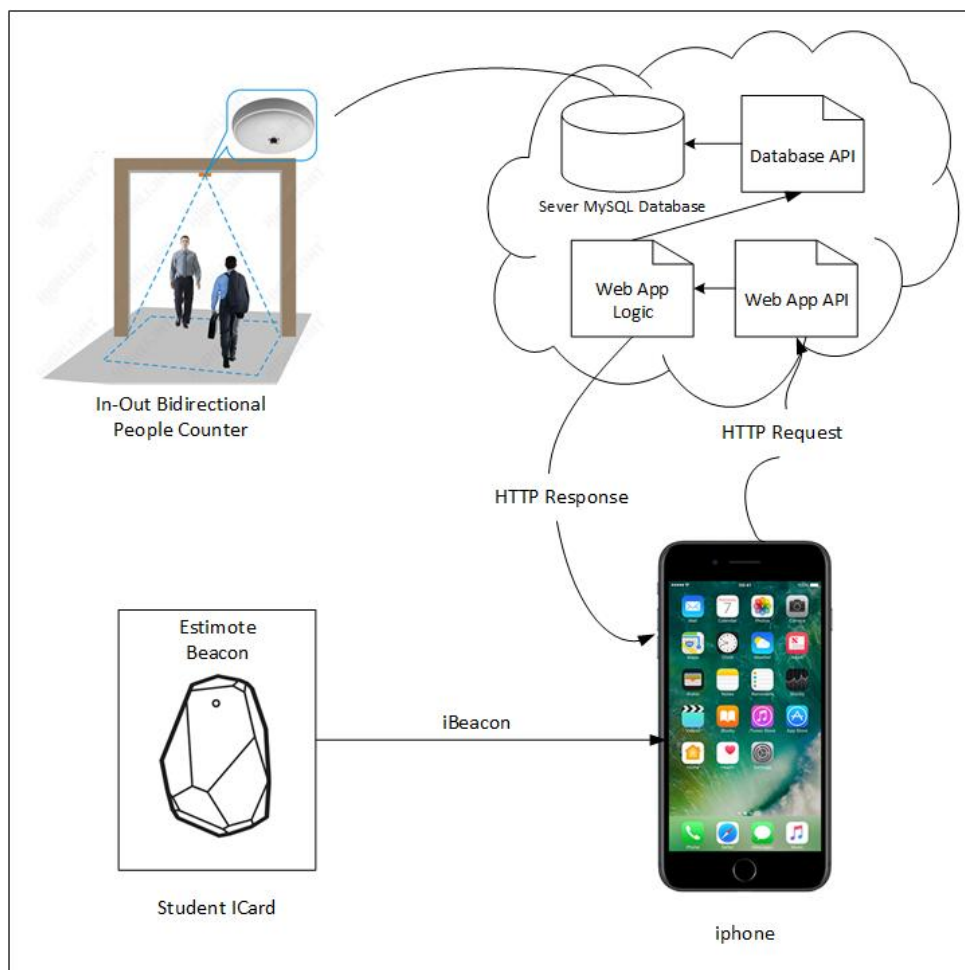
(Shubham Jain,Shweta Kale,Shraddha Kannurkar)

Abstract: The problem of fake attendance (proxy) is growing day by day. There are many technologies have been used to develop the attendance system like RFID, Biometrics, Barcode, etc. In spite of all these technologies have been used many students are able to mark the fake attendance (proxy) and some systems are very time consuming. There is a need to develop the riskless and shielded attendance monitoring and management system which is our main motto. The technology which we are going to use in the proposed system is “iBeacon Technology” which is mostly used for indoor positioning system and proximity detection and also beneficial in removing the problem of fake attendance. The newer technology which is found out to interact with hardware is iBeacon. An iBeacon works on Bluetooth Low Energy technology which is used to send a signal in a definite format. The iBeacon devices give accurate results in an indoor system in spite of the influencing factors of radio waves. In iBeacon devices, we can adjust the range of beacon up to which the device can transmit. The range of iBeacon is up to 200m, which can vary according to the type of iBeacon used. An infinite number of devices can be connected to the iBeacon. The main advantage of the proposed system is to overcome the human intervention and follow the concept of paperless smart work. The accuracy of our proposed system is very high i.e. 100% and removes human intervention to almost zero.

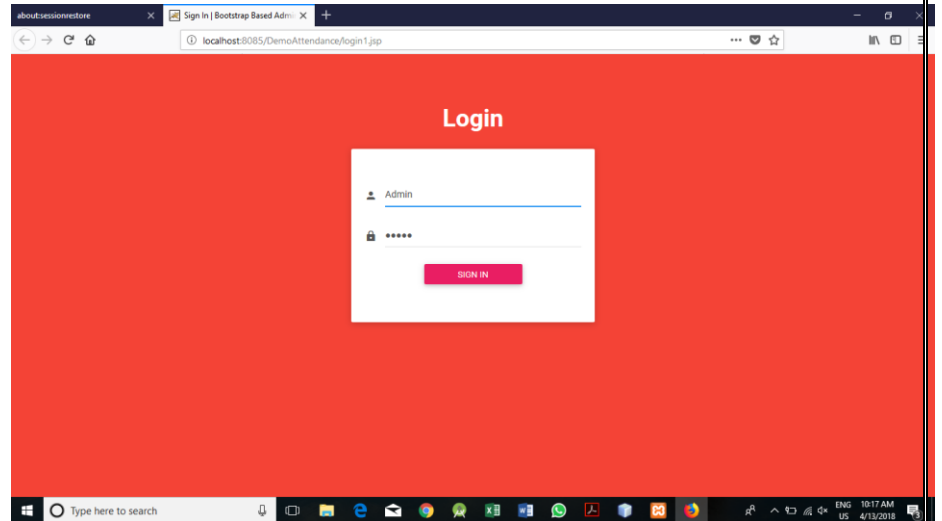
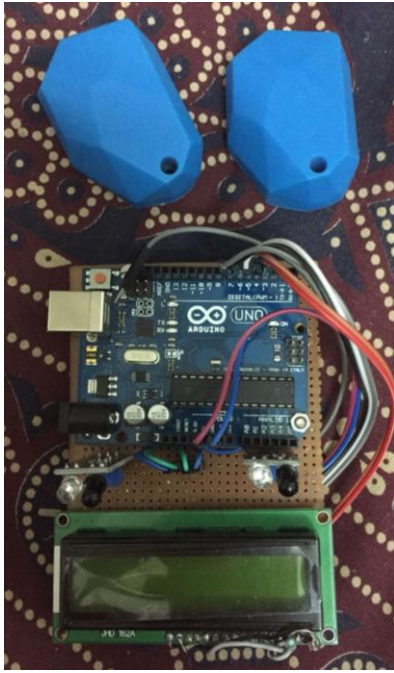
Details of the Project Work: The elements used in proposed system are bi-directional people counter, student, the central device (server), staff member, cloud, iBeacon device. Bidirectional counter will be set at the doorstep. The value of counter will be incremented by one every time student enters the classroom. Now the student’s I-Card will be consisting of iBeacons which will

get connected to the server immediately and the attendance will be marked. As the staff's device and server are connected to the cloud, the notification regarding the counter and the attendance of the student will be sent to staff member also the student will only receive the attendance notification. If any of the proxy attendance has been marked, it will be known by matching the values between the counter and the server and staff member can discard the fake attendance immediately. According to the data saved on cloud, attendance report will be generated i.e. average students' attendance is calculated and we can also generate the student's defaulter list.

Block diagram:



Snapshots:



Conclusion: In this project, we have proposed the smart attendance system which will totally overcome the problem of fake attendance (proxy) which was the main motto of our research. It will also minimize the human intervention to almost zero. Implementation of our proposed system will follow the paperless approach as there is no use of paper anywhere for attendance management leading to the green environment. The accuracy of the proposed system is very high, i.e. 100%. We can say that our proposed system is more efficient than the existing systems which are based on various technologies such as RFID, Barcode, Bio-Metric, Wireless, etc. Furthermore, based on the observations and results obtained in this work, our future work is to develop the system which will not only detect the fake attendance but it will also remove the fake attendance automatically.

Project Title “Stratification of Dengue Fever using SMO and NSGA-II Optimization Algorithms”

Preet Dalsania, Swarada Deshpande, Aishwarya Dixit, Poonam Doddamani

Abstract. In today’s world, millions of cases of dengue are reported ever year. The number of cases has increased, afflicting many individuals. For prediction of dengue clinical methodology comprises of antigens and anti-bodies tests. The tests are conducted on the blood samples collected from the patients. In our proposed paper we are stratifying dengue into Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF) and healthy patients. The dataset [GDS5093] being referred in this proposed model are of acute dengue patients. We are using optimization algorithms like Spider Monkey Optimization (SMO) and Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) and to increase the optimality of the model, we have used Probabilistic Neural Network (PNN). PNN uses feedforward technique for classification. A paper has been already published on dengue fever classification using PSO approach which achieved the accuracy of 90.91%

Keywords: SMO, NSGA-II, PNN, DF, DHF.

1 Introduction

Dengue is a mosquito borne viral disease which is mainly transmitted by the species of female mosquitoes named “Aedes aegypti”. The “Aedes aegypti” mosquito lives in urban habitats and breeds on a large scale. Symptoms of dengue include high fever, pain behind the eyes, muscle and joint pain, severe headache.[1]

The primary task is to find out whether the person is suffering from dengue or is he a healthy person. After this, the more challenging part is to find whether he is infected from Dengue fever (DF) or dengue Hemorrhagic fever (DHF)[3]. There is a need for research in this field.

Our proposed model gives an architecture of stratification of dengue disease using the PNN model along with optimization algorithms. A probabilistic neural network is widely used in classification and pattern recognition problems. This type of neural network has four layers in it the input layer, the hidden layer, pattern layer, output layer. Dataset of acute dengue patients [GDS5093][2]. These layers compute the result and compare it with the training set and produce the desired output. Probabilistic neural network sometimes converge at the local optima hence to avoid this we will apply two algorithms that are NSGA-II and SMO. These algorithms will help us to converge at the global optima rather than the local optima. These algorithms have their independent features.

The current work proposes NSGA-II and SMO trained Neural Network using greedy feedforward selection algorithm for selection of the specific prominent gene [3]. Only the useful data must be extracted from the dataset as there are 54715 genes for 56 homo-sapiens subjects [3]. We will be using PNN as we are mainly considering the probability of achieving a optimal accuracy [4] as compared with the previous PSO trained NN which also uses the greedy forward selection algorithm for gene selection. The NN-PSO model gave the accuracy of 90.91% [3]. Our proposed model is to find whether we achieve a better accuracy than NN-PSO

Rest of the paper is organized as section II discusses literature survey work, section III discusses the algorithms used, section IV discusses about result and section V discusses about conclusion followed by references.

2 Literature Survey

P. Manivannan et al [5] provided K-Mediod Clustering Algorithm for prediction of dengue fever. Research work done on predicting the people who are affected by dengue depending upon categorization of age using the K-medoid clustering algorithm. The K-mediod clustering algorithm was applied on the dengue dataset. The result obtained by using K-mediod clustering algorithm has increased the efficiency of output. This is the most effective technique to predict the dengue patients with serotypes. It lacks in scalability of large datasets. It also has high time and space complexity.

Sankhadeep Chatterjee et al [3] has provided a PSO approach for classification of dengue virus into DF (Dengue fever) and DHF (Dengue Hemorrhagic Fever). The greedy forward methodology is used to select the significant genes through the pre-processing stage from the available blood samples. Further PSO trained Neural Network (PSO-NN) was implemented for detecting and classifying the fever into DF and DHF. The accuracy obtained by this method is 90.91%. Different nature based algorithms like NSGA-II and SMO can be used to increase the accuracy of the same.

Tarmizi et al [6] has referred the weather of Thailand, Indonesia and Malaysia where the climate is more humid , and hence water borne diseases like dengue is more prone in that area. The study proposes different machine learning techniques like Data mining (DM), Artificial neural network (ANN), and rough set theory (RS). The classification algorithms which are used to predict dengue disease. The data set referred is of public health department at Selangor state. WEKA data mining tool is applied with two tests which are 10 cross fold validation and Percentage split .The accuracy obtained with 10 cross fold validation was 99.5% with DT, 99.8% with ANN, 100% accuracy with RS. Using percentage split 99.2% of accuracy was achieved with DT and ANN, whereas 99.72% of accuracy was obtained using

RS.

Fathima et al [7] The work proposed is prediction of Arbovirus- Dengue disease. The data mining algorithm that is used by them are support vector machine (SVM). The reference was taken from the King Institute of Preventive Medicine and surveys of hospitals and laboratories at Chennai, India. The data set they referred had 29 attributes and 5000 samples. T R project version tool was used for examination and achieved the accuracy of 90.42%.The only disadvantage we can talk about is accuracy achieved by using SVM and hence the accuracy achieved by rough set theory which was 100% is referred.

Ibrahim et al [8] The model was suggested which used Artificial neural network with multi-layer feedforward neural network. It is used for forecasting the defervescence fever in patients of dengue disease. The data is gathered from 252 hospitalized patients, in which 4 patients are having Dengue Fever and 252 patients had Dengue hemorrhagic fever. MATLAB neural network tool box is used and achieved the accuracy of 90%. Accuracy can be increased using different methodologies.

Project Title:” Smart Agriculture System based on Data Analytics”

*Sponsored by Bhimashankar Saharaki Karkhana, Bhimashankar, MH, India.
(Agricultural Industry)*

Omkar Buchade ,Nilesh Mehta, Shubham Ghodekar, Chandan Mehta. --BE
Computer(1st Shift)

Abstract: India being an agricultural country, its economy predominantly depends on agriculture yield and allied agro industry products. In India, agriculture is largely influenced by rainwater which is highly unpredictable. Agriculture growth also depends on diverse soil parameters, namely Nitrogen, Phosphorus, Potassium, Crop rotation, Soil moisture, pH, surface temperature and weather aspects like temperature, rainfall, etc. India now is rapidly progressing towards technical development. Thus, technology will prove to be beneficial to agriculture which will increase crop productivity resulting in better yields to the farmer. The proposed project provides a solution for Smart Agriculture by monitoring the agricultural field which can assist the farmers in increasing productivity to a great extent. Weather forecast data obtained from IMD (Indian Metrological Department) such as temperature and rainfall and soil parameters repository gives insight into crops that are suitable to be cultivated in a particular area. This work presents a system, in form of an android based application and a website, which uses Machine Learning techniques in order to predict the most profitable crop in the current weather and soil conditions. By applying Machine Learning algorithm: Multiple Linear Regression, a prediction of most suitable crops according to current environmental conditions is made. This provides a farmer with variety of options of crops that can maximize their profit.

Architecture diagrams:

There is no system existing which recommends crops based on multiple factors such as nutrients in soil (Nitrogen, Phosphorus and Potassium), pH and weather components which include temperature and rainfall. The proposed system suggests an android and a web based application, which can precisely predict the most profitable crop to the farmer. Location is the only input for the extrapolation system. The user location is identified with the help of GPS. According to user location, the feasible crops in the respective location is identified from the soil, pH and weather database. These crop are further processed along-with past year production database using Regression techniques to identify the most profitable crop in the current location. After this processing is done at server side, the result is sent to the user's android and web application. Thus depending on the numerous scenarios and additional filters according to the user requirement the most producible crop is suggested.

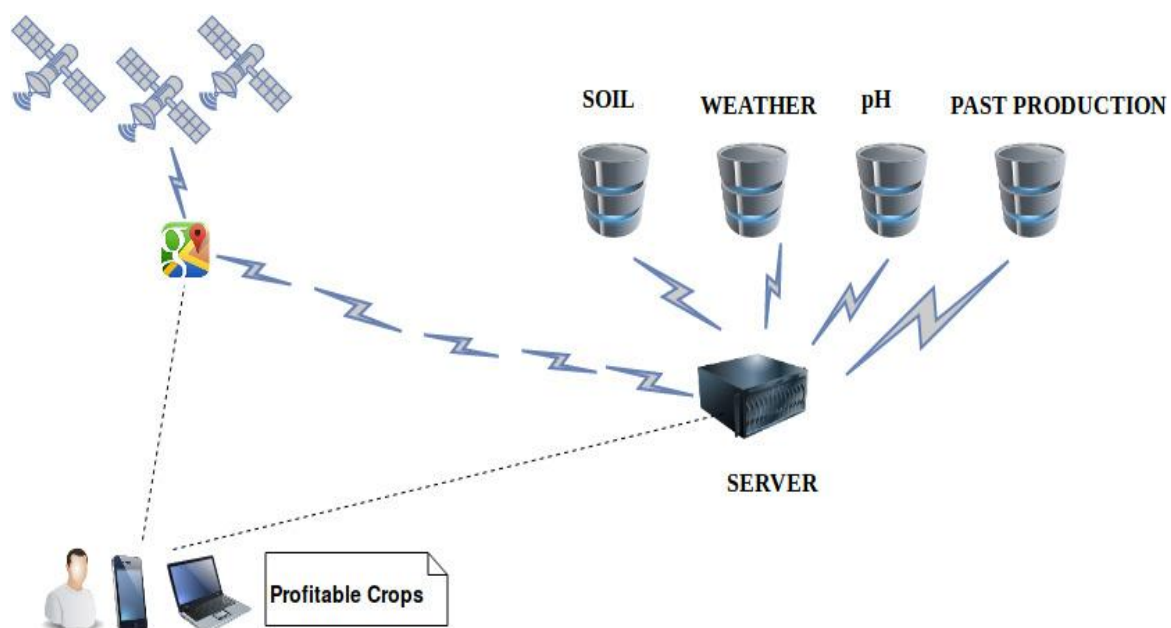


Fig 1: System Architecture

Detailed system architecture:

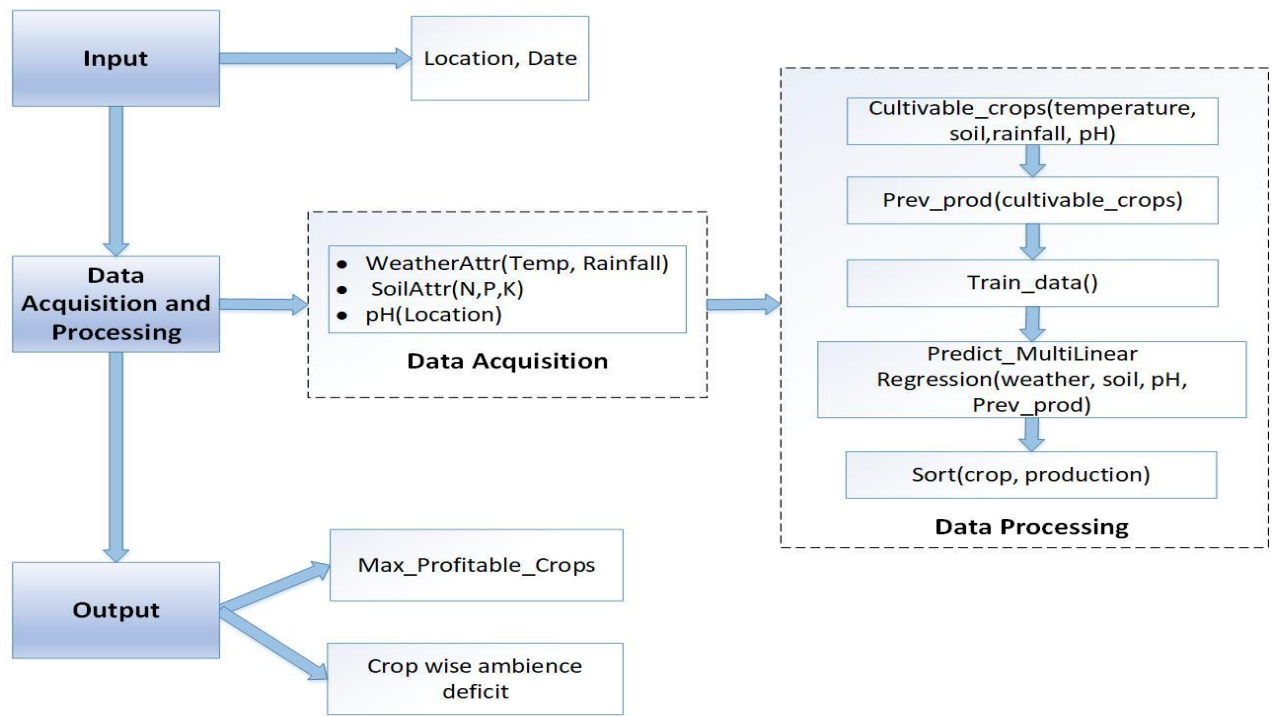


Fig 2: System Architecture

Source: [Source: https://data.gov.in](https://data.gov.in)

Farm First application Implementation

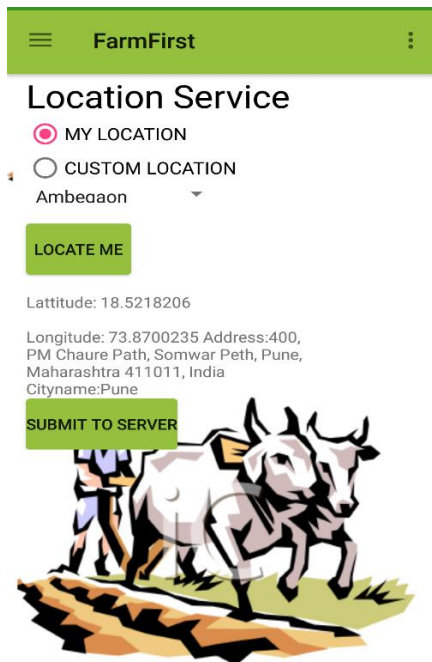


Fig. Android application home screen

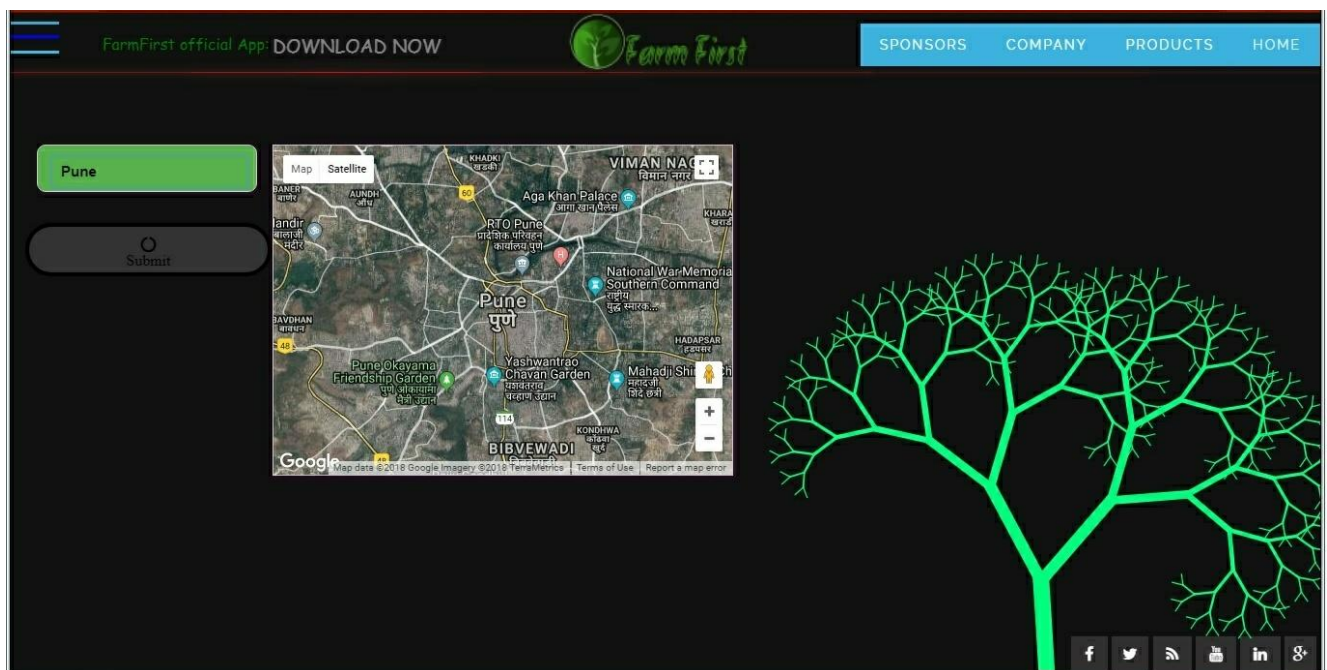


Fig. Website home page

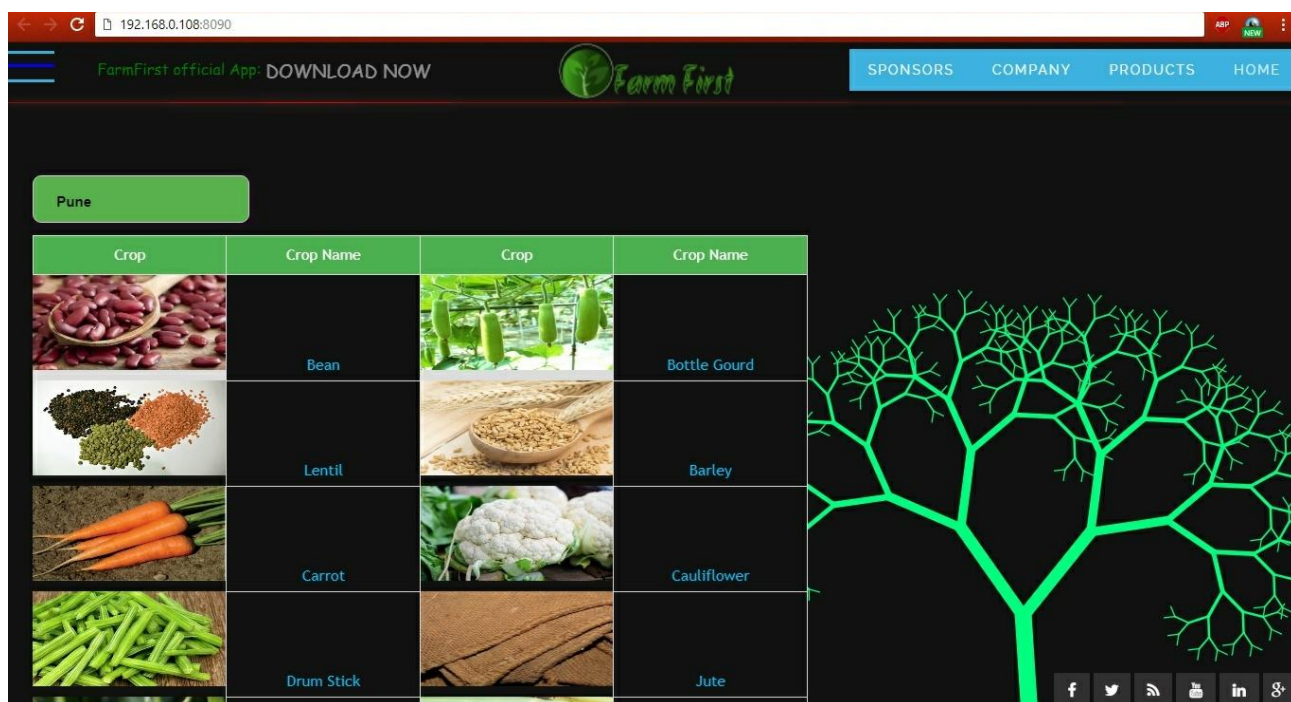


Fig. Display of crops based for location: Pune

CONCLUSION:

The proposed system lists out all crops feasible in a particular area, helping the farmer in decision making of which crop to cultivate. A careful examination of the data related to soil, weather, pH and past year production has been done by the system and suggests which are the most profitable crops which can be cultivated in the apropos environmental condition. This system examines the past production of data which will help the farmer get insight into the demand and the cost of various crops in market over multiple years. As maximum types of crops will be covered under this system, farmer will get to know about the crop which may never have been cultivated.

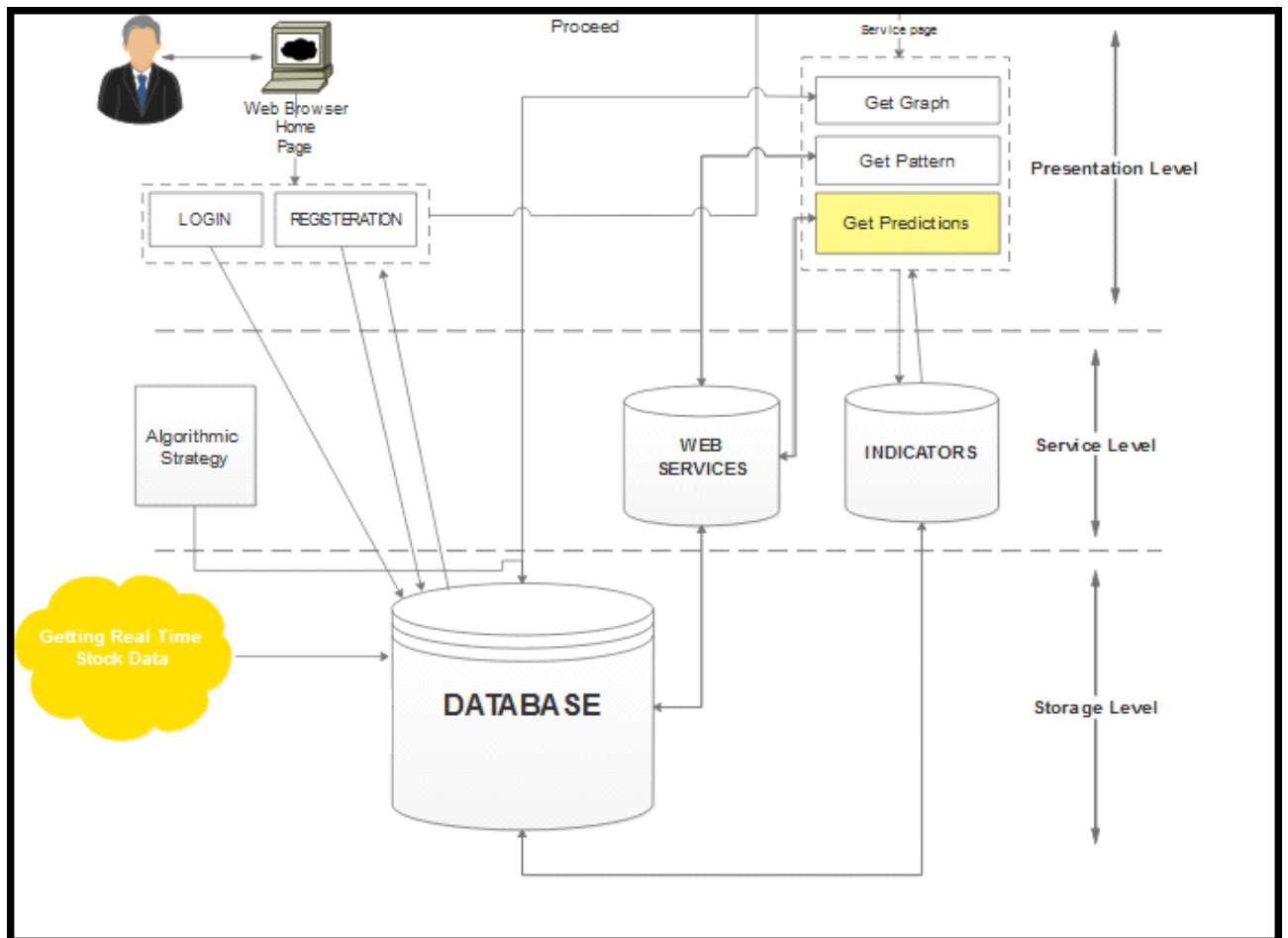
Project Title “Stock Market Prediction Using Machine Learning”

(Jenish Karia,Muskaan Khan,Satyam Anand,Tushar Mukherjee)

Abstract

The use of Neural networks has found a variegated field of applications in the present world. This has led to the development of various models for financial markets and investment. This paper represents the idea how to predict share market price using Artificial Neural Network with a given input parameters of share market. Artificial Neural Network can remember data of any number of years and it can predict the feature based on the past data. This paper makes use feed forward architecture for prediction. The network was trained using ten year data. It shows a good performance for market prediction. The network selected though was not able to predict exact value but it succeeded in predicting the trends of stock market.

Architecture Diagram



Conclusion

The Stock market prediction has at all times been tough work for analysts. Thus, we attempt to make use of vast written data to forecast the stock market indices. We use machine learning techniques analysis the accuracy in predictions can be achieved. RNN algorithm is analyzed to forecast market upcoming trends. Financial analysts, investors can use this prediction model to take trading decision by observing market behaviour.

Project Title :

“Data Compression using Wireless Distributed Computing.”

(Bhushan Mali,Siddhesh Pandey,Ivan Pillay, Shubham Soni)

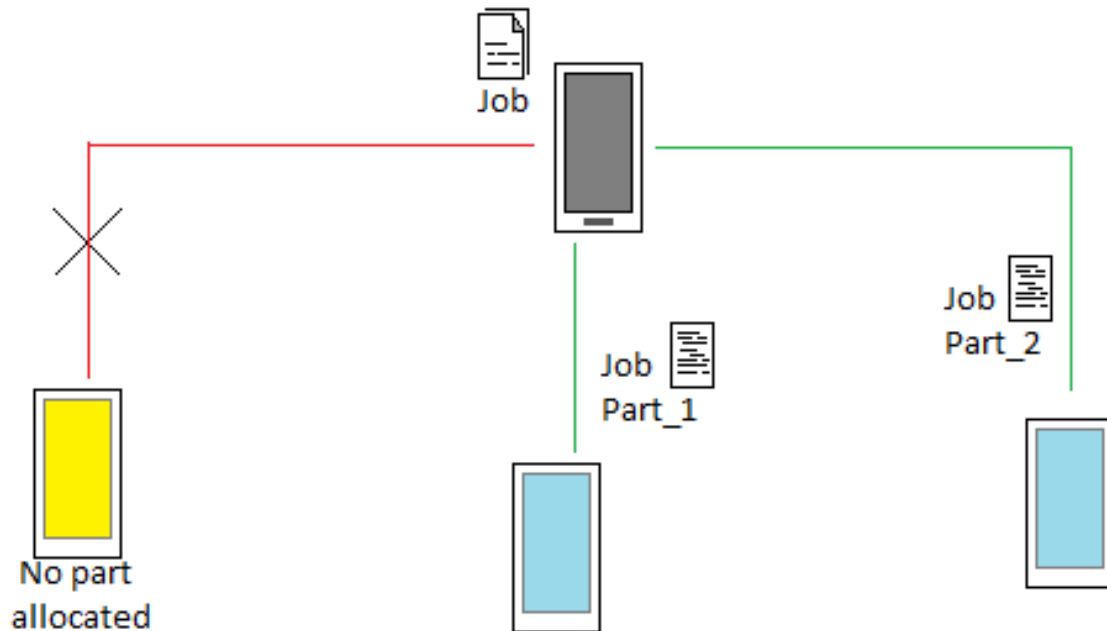
Abstract :

In this modern world, smartphone is used as a multi-purpose device. People use their smartphone for emails, calling, games, social networking, etc. which then increases the demand for smartphone. Also due to development in hardware in recent years it has become possible to perform multiple tasks simultaneously or perform complicated task just on a smartphone. On an average, people replace their smartphone every 20 months. Thus we can say there is lot of computational power available at the user end which is either idle or used infrequently. These sources include smartphones, notebooks, and also IOT devices that have sufficient capability.

We propose an application which is data compression, that utilizes these available devices to execute its own operation. Data Compression comprises of two phases encoding and decoding, out of which encoding requires more time to compress the data as it has to parse all of the data and determine which sequence can be compressed. Thus we can split this task among multiple devices to decrease resource consumption on the host device as well as utilize resources on idle devices. Offloading a task reduces resource consumption on the host device which can then be used to perform other tasks.

This application distributes the task of compression among nodes based on their energy constraints. Thus high capacity devices receives larger chunk of data as compared to others. We also use an optimal task allocation algorithm for distribution of tasks among the nodes. In this way, we can reduce the load of a device by distributing the task to its neighboring devices.

Architecture Diagram :



Conclusion :

We discussed how computational power of nearby devices can be utilized to perform high resource consuming tasks such as compression. Each of the nearby devices process on a chunk of the original data and final result is combination of all the results obtained from other devices. This implementation of utilizing peer devices for processing can be used where cloud computing is not available. In future we can extend this application by adding more tasks that can be processed in a distributed way.

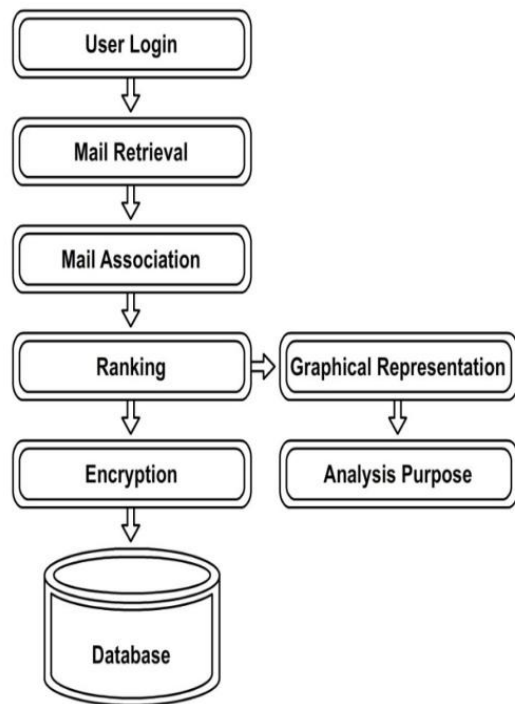
Project Title

“ Email Correlation And Analysis Using Pattern Matching Technique”

(Ashwini Kawade,Pratima Khambad,Shraddha Bhagwat,Shrirang Potale)

Abstract : Email or Electronic mail is been widely used to exchange messages between people. Electronic mail or an Email can be personal, commercial, business related messages. These messages can be used to determine association between mail holders. A report or summary can be made by retrieving mails and grouping them to get a count of total mails that has been exchanged from a particular mail id. This will help to find out an association between the two contacts and reduces the burden of reading and going through each and every mail which is being send or received. The generated report is useful in means of cyber forensics and in means of data mining, where emails are studied to find out a person, company or an organization with whom the user has been communicating using emails and how many times they have exchanged emails. Proposed system generates a report on number of mail's being exchanged with different user and provides a graphical representation of the same. In addition, to the above feature, an advanced features is added in the system which helps to monitor the workers or group members and gets feedback on progress of work assigned to them by using emails.

Architecture Diagram :



Conclusion : The proposed system would display the association between email contacts. This information is important by means of both cyber forensic and email data mining. the proposed system uses open sources software and can be installed on computers with low configuration, which makes the system more feasible. Proposed system also provides a feature of email client which would help users to monitor important mails send to colleagues, and progress on these important mail can be seen to by the user once updated by his or her colleague which reduces the effort of composing a new mail for sending updates to user. System can be made feasible and efficient if we are able to trace and count those emails which are deleted from server's mail account and making application of the proposed system which would not be limited to desktop and can be used in multiple machine any time any where can provide a bright future to cyber forensics and email data mining.

Project Title

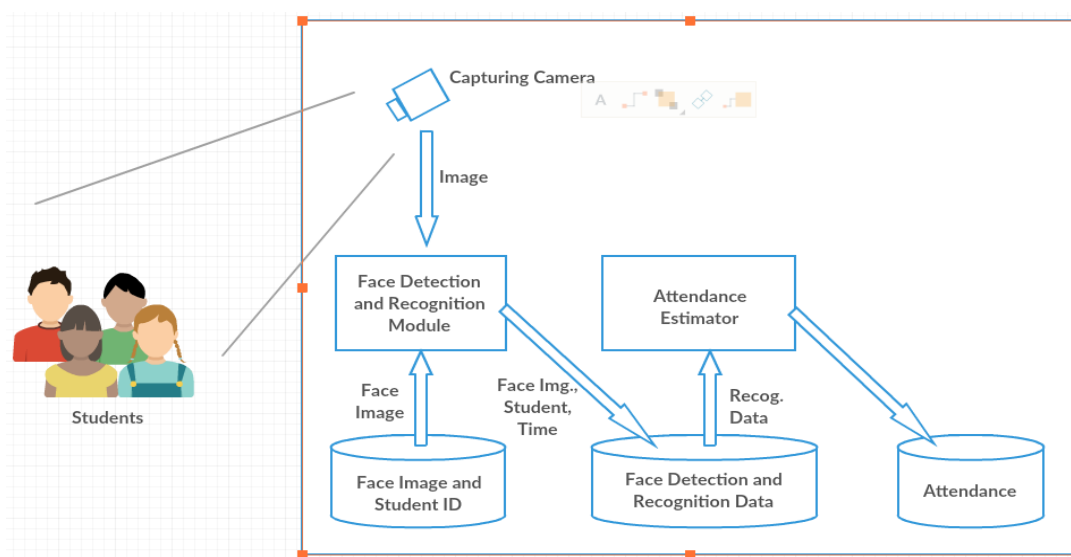
“Attendance Monitoring System Using Image Processing and Machine Learning”

(Manasi Jadhav, Tushar Indalkar, Kirtiraj Kadam, Shivam Mulay)

Abstract

“Attendance Monitoring is used in various fields like educational institutes, IT companies, Government offices etc. Educational institute require attendance for scholarships and various other important reasons, therefore attendance is marked manually or by roll call list but these methods can induce proxies/fake attendance of students who aren't present. This would prove injustice to the students who attend their colleges regularly. Hence the solution is to monitor the attendance in such a way that fake attendance is not marked and solution to this problem is by developing such a system that would mark attendance only of those students who are present in the classroom.”

System Architecture-



Conclusion-

We have proposed an Image & machine learning based attendance monitoring system for classroom. Unlike the traditional manual marking of attendance which can mark fake attendance our system will detect faces from the classroom and mark attendance accordingly by feature extraction & matching by using machine learning techniques.

We have implemented Viola-Jones algorithm and evaluated the performance. Our results showed that the detection of faces of students is accurately done in the classroom