



MOHAMED SATHAK A J COLLEGE OF ENGINEERING
Sponsored by Mohamed Sathak Trust
(Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai)
Siruseri IT Park, Egattur, Chennai 603 103.

M.E COMPUTER SCIENCE AND ENGINEERING

S.NO	Name of the course that include experiential learning through Project work/ Internship	Page No
1	CP5151-Advanced Data structures and Algorithms	
2	CP5152-Advanced Computer Architecture	
3	CP5154- Advanced Software Engineering	
4	CP5201 – Network Design and Technologies	
5	CP5191- Security Practices	
6	CP5293 – Big Data Analytics	
7	CP5092 – Cloud Computing Technologies	



PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.



MOHAMED SATHAK A J COLLEGE OF ENGINEERING
Sponsored by Mohamed Sathak Trust
(Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai)
Siruseri IT Park, Egattur, Chennai 603 103

M.E COMPUTER SCIENCE AND ENGINEERING

S.NO	Name of the course that include experiential learning through Project work/ Internship
1	CP5151 – Advanced Data structures and Algorithms
2	CP5152 – Advanced Computer Architecture
3	CP5154- Advanced Software Engineering
4	CP5201 – Network Design and Technologies
5	CP5191- Security Practices
6	CP5293 – Big Data Analytics
7	CP5092 – Cloud Computing Technologies



PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

**Improving Privacy and Security in Decentralizing Multi-Authority
Attribute Based Encryption in Cloud Computing**

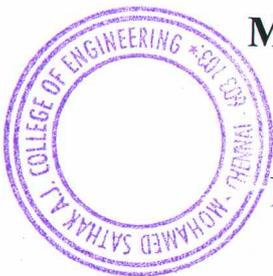
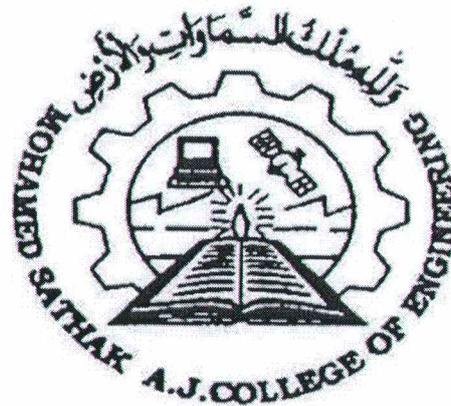
PHASE II REPORT

Submitted by

ROBERT JOHN

in partial fulfillment for the award of the degree of

**MASTER OF ENGINEERING IN
COMPUTER SCIENCE**



**MOHAMMED SATHAK A.J. COLLEGE OF
ENGINEERING**

**DEPARTMENT OF COMPUTER SCIENCE
AND ENGINEERING**

ANNA UNIVERSITY, CHENNAI

2020 - 2021


PRINCIPAL
MOHAMMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR) Siruseri, IT Park,
Chennai-603 103.



MOHAMED SATHAK A.J COLLEGE OF ENGINEERING

(Approved by AICTE, New Delhi and Affiliated to Anna University) 34, Old Mahabalipuram Road, Egattur, Chennai – 603 103.

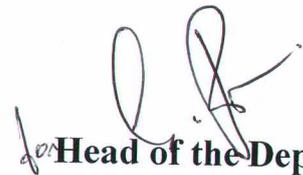
ISO 9001: 2008 Certified Institution

Sponsored By: MOHAMED SATHAK TRUST, Chennai – 600 034.

BONAFIDE CERTIFICATE

This is to certify that this Report titled “**Improving Privacy and Security in Decentralizing Multi Authority Attribute Based Encryption in Cloud Computing**” is the Bonafide Record of work done by Mr/Ms.. **.ROBERT JOHN** Register Number **311819405001**. of **2nd YEAR / IV Semester M.E-COMPUTER SCIENCE AND ENGINEERING** in the CP5411 Project Work **Phase - II** during the academic year 2020-2021.


Staff In-Charge


Head of the Department

Submitted for the Anna University M.E Project viva -voice held on

..... 05-08-2021


Internal Examiner


External Examiner




PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), S. Egattur, Chennai-603 103.

ABSTRACT:

Decentralizing multi-authority attribute-based encryption (ABE) has been adopted for solving problems arising from sharing confidential corporate data in cloud computing.

For decentralizing multiauthority ABE systems that do not rely on a central authority, collusion resistance can be achieved using a global identifier. Therefore, identity needs to be managed globally, which results in the crucial problems of privacy and security.

A scheme is developed that does not use a central authority to manage users and keys, and only simple trust relations need to be formed by sharing the public key between each attribute authority (AA).

The scheme is based on Composite order bilinear groups. A proof of security is presented that uses the dual system encryption methodology.

Cloud storage is a promising and important service paradigm in cloud computing. Benefits of using cloud storage include greater accessibility, higher reliability, rapid deployment and stronger protection, to name just a few. Despite the mentioned benefits, this paradigm also brings forth new challenges on data access control, which is a critical issue to ensure data security. Since cloud storage is operated by cloud service providers, who are usually outside the trusted domain of data owners, the traditional access control methods in the Client/Server model are not suitable in cloud storage environment. The data access control in cloud storage environment has thus become a challenging issue.

To address the issue of data access control in cloud storage, there have been quite a few schemes proposed, among which Cipher Text-Policy Attribute- Based Encryption (CP-ABE) is regarded as one of the most promising techniques. A salient feature of CP-ABE is that it grants data owners direct control power based on access policies, to provide flexible, fine-grained and secure access control for cloud storage systems.

In CP-ABE schemes, the access control is achieved by using cryptography, where an owner's data is encrypted with an access structure over attributes, and a user's secret key is labeled with his/her own attributes. Only if the attributes associated with the user's



A handwritten signature in green ink, appearing to be "Ushini".

PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMPL) S. J. S. S. S.
Chennai-003, INDIA.

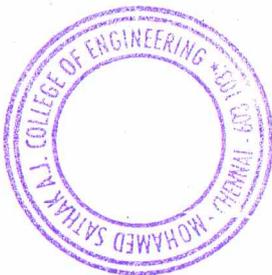
secret key satisfy the access structure, can the user decrypt the corresponding cipher text to obtain the plain text. So far, the CP-ABE based access control schemes for cloud storage have been developed into two complementary categories, namely, single-authority scenario, and multi authority scenario.

Although existing CP-ABE access control schemes have a lot of attractive features, they are neither robust nor efficient in key generation. Since there is only one authority in charge of all attributes in single-authority schemes, offline/crash of this authority makes all secret key requests unavailable during that period. The similar problem exists in mufti-authority schemes, since each of multiple authorities manages a disjoint attribute set.

In single-authority schemes, the only authority must verify the legitimacy of users' attributes before generating secret keys for them.

As the access control system is associated with data security, and the only credential a user possess is his/her secret key associated with his/her attributes, the process of key issuing must be cautious.

However, in the real world, the attributes are diverse. For example, to verify whether a user is able to drive may need an authority to give him/her a test to prove that he/she can drive. Thus he/she can get an attribute key associated with driving ability.



A handwritten signature in green ink, appearing to read "Sathak".

PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR) Siruseri, IT Park
Chennai-603 103.

7: Conclusion and Future Work:

In this paper, we proposed a new framework, named RAAC, to eliminate the single-point performance bottleneck of the existing CP-ABE schemes. By effectively reformulating CPABE cryptographic technique into our novel framework, our proposed scheme provides a fine-grained, robust and efficient access control with one-CA/multi-*AAs* for public cloud storage. Our scheme employs multiple *AAs* to share the load of the time-consuming legitimacy verification and standby for serving new arrivals of users' requests. We also proposed an auditing method to trace an attribute authority's potential misbehavior. We conducted detailed security and performance analysis to verify that our scheme is secure and efficient. The security analysis shows that our scheme could effectively resist to individual and colluded malicious users, as well as the honest-but-curious cloud servers. Besides, with the proposed auditing & tracing scheme, no *AA* could deny its misbehaved key distribution. Further performance analysis based on queuing theory showed the superiority of our scheme over the traditional CP-ABE based access control schemes for public cloudstorage.




PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.



MOHAMEDSATHAKAJCOLLEGE OFENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
1.	CP5151	ADVANCED DATA STRUCTURES AND ALGORITHMS	<ul style="list-style-type: none">• Shortest Paths and Matrix Multiplication• Shortest Paths and Matrix Multiplication

PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

CP5151

ADVANCED DATA STRUCTURES AND ALGORITHMS

L T P C
4 0 0 4

OBJECTIVES:

- To understand the usage of algorithms in computing.
- To learn and use hierarchical data structures and its operations
- To learn the usage of graphs and its applications.
- To select and design data structures and algorithms that is appropriate for problems.
- To study about NP Completeness of problems.

UNIT I **ROLE OF ALGORITHMS IN COMPUTING** 12

Algorithms - Algorithms as a Technology- Insertion Sort - Analyzing Algorithms - Designing Algorithms- Growth of Functions: Asymptotic Notation - Standard Notations and Common Functions- Recurrences: The Substitution Method - The Recursion-Tree Method

UNIT II **HIERARCHICAL DATA STRUCTURES** 12

Binary Search Trees: Basics - Querying a Binary search tree - Insertion and Deletion- Red-Black trees: Properties of Red-Black Trees - Rotations - Insertion - Deletion -B-Trees: Definition of B-trees - Basic operations on B-Trees - Deleting a key from a B-Tree- Fibonacci Heaps: structure - Mergeable-heap operations- Decreasing a key and deleting a node-Bounding the maximum degree.

UNIT III **GRAPHS** 12

Elementary Graph Algorithms: Representations of Graphs - Breadth-First Search - Depth-First Search - Topological Sort - Strongly Connected Components- Minimum Spanning Trees: Growing a Minimum Spanning Tree - Kruskal and Prim- Single-Source Shortest Paths: The Bellman-Ford algorithm - Single-Source Shortest paths in Directed Acyclic Graphs - Dijkstra's Algorithm; All-Pairs Shortest Paths: Shortest Paths and Matrix Multiplication - The Floyd-Warshall Algorithm;

UNIT IV **ALGORITHM DESIGN TECHNIQUES** 12

Dynamic Programming: Matrix-Chain Multiplication - Elements of Dynamic Programming - Longest Common Subsequence- Greedy Algorithms: An Activity-Selection Problem - Elements of the Greedy Strategy- Huffman Codes.

UNIT V **NP COMPLETE AND NP HARD** 12

NP-Completeness: Polynomial Time - Polynomial-Time Verification - NP- Completeness and Reducibility - NP-Completeness Proofs - NP-Complete Problems

TOTAL: 60 PERIODS

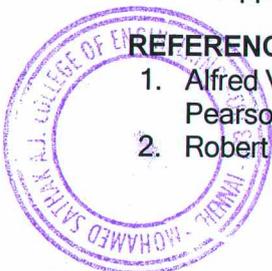
OUTCOMES:

Upon the completion of the course the students should be able to:

- Design data structures and algorithms to solve computing problems
- Design algorithms using graph structure and various string matching algorithms to solve real-life problems
- Apply suitable design strategy for problem solving

REFERENCES:

1. Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, "Data Structures and Algorithms", Pearson Education, Reprint 2006.
2. Robert Sedgewick and Kevin Wayne, "ALGORITHMS", Fourth Edition, Pearson Education.



Principal
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-605 003.



MOHAMEDSATHAKAJCOLLEGE OFENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
2.	CP5152	ADVANCED COMPUTER ARCHITECTURE	<ul style="list-style-type: none">• Cloud Computing• Graphics Processing Units

PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

CP5152

ADVANCED COMPUTER ARCHITECTURE

L T P C
3 0 0 3

OBJECTIVES:

- To introduce the students to the recent trends in the field of Computer Architecture and identify performance related parameters.
- To learn the different multiprocessor issues.
- To expose the different types of multicore architectures.
- To understand the design of the memory hierarchy.

UNIT I FUNDAMENTALS OF COMPUTER DESIGN AND ILP 9
Fundamentals of Computer Design – Measuring and Reporting Performance – Instruction Level Parallelism and its Exploitation - Concepts and Challenges -Exposing ILP - Advanced Branch Prediction - Dynamic Scheduling - Hardware-Based Speculation - Exploiting ILP - Instruction Delivery and Speculation - Limitations of ILP - Multithreading

UNIT II MEMORY HIERARCHY DESIGN 9
Introduction – Optimizations of Cache Performance – Memory Technology and Optimizations - Protection: Virtual Memory and Virtual Machines - Design of Memory Hierarchies – Case Studies.

UNIT III MULTIPROCESSOR ISSUES 9
Introduction- Centralized, Symmetric and Distributed Shared Memory Architectures -Cache Coherence Issues – Performance Issues – Synchronization – Models of Memory Consistency - Case Study-Interconnection Networks - Buses, Crossbar and Multi-stage Interconnection Networks

UNIT IV MULTICORE ARCHITECTURES 9
Homogeneous and Heterogeneous Multi-core Architectures - Intel Multicore Architectures - SUN CMP architecture - IBM Cell Architecture. Introduction to Warehouse-scale computers- Architectures- Physical Infrastructure and Costs- Cloud Computing -Case Study- Google Warehouse-Scale Computer.

UNIT V VECTOR, SIMD AND GPU ARCHITECTURES 9
Introduction-Vector Architecture - SIMD Extensions for Multimedia - Graphics Processing Units – Case Studies – GPGPU Computing – Detecting and Enhancing Loop Level Parallelism-Case Studies.

TOTAL : 45 PERIODS

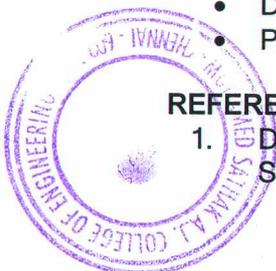
OUTCOMES:

Upon completion of this course, the students should be able to:

- Identify the limitations of ILP.
- Discuss the issues related to multiprocessing and suggest solutions
- Point out the salient features of different multicore architectures and how they exploit parallelism.
- Discuss the various techniques used for optimising the cache performance
- Design hierarchical memory system
- Point out how data level parallelism is exploited in architectures

REFERENCES:

1. Darryl Gove, Multicore Application Programming: For Windows, Linux, and Oracle Solaris, Pearson, 2011




PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.



MOHAMEDSATHAKAJCOLLEGE OFENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
3.	CP5154	ADVANCED SOFTWARE ENGINEERING	<ul style="list-style-type: none">• Interaction diagrams• Unit testing• White box testing

PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

OBJECTIVES:

- To understand Software Engineering Lifecycle Models
- To do project management and cost estimation
- To gain knowledge of the System Analysis and Design concepts.
- To understand software testing approaches
- To be familiar with DevOps practices

UNIT I INTRODUCTION

9

Software engineering concepts - Development activities - Software lifecycle models - Classical waterfall - Iterative waterfall - Prototyping - Evolutionary - Spiral - Software project management - Project planning - Estimation - Scheduling - Risk management - Software configuration management.

UNIT II SOFTWARE REQUIREMENT SPECIFICATION

9

Requirement analysis and specification - Requirements gathering and analysis - Software Requirement Specification - Formal system specification - Finite State Machines - Petrinets - Object modelling using UML - Use case Model - Class diagrams - Interaction diagrams - Activity diagrams - State chart diagrams - Functional modelling - Data Flow Diagram.

UNIT III ARCHITECTURE AND DESIGN

9

Software design - Design process - Design concepts - Coupling - Cohesion - Functional independence - Design patterns - Model-view-controller - Publish-subscribe - Adapter - Command - Strategy - Observer - Proxy - Facade - Architectural styles - Layered - Client-server - Tiered - Pipe and filter.- User interface design

UNIT IV TESTING

9

Testing - Unit testing - Black box testing- White box testing - Integration and System testing- Regression testing - Debugging - Program analysis - Symbolic execution - Model Checking

UNIT V DEVOPS

9

DevOps:Motivation-Cloud as a platform-Operations- Deployment Pipeline:Overall Architecture-Building and Testing-Deployment- Case study: Migrating to Microservices.

TOTAL: 45 PERIODS**OUTCOMES:**

At the end of this course, the students will be able to:

- Understand the advantages of various Software Development Lifecycle Models
- Gain knowledge on project management approaches as well as cost and schedule estimation strategies
- Perform formal analysis on specifications
- Use UML diagrams for analysis and design
- Architect and design using architectural styles and design patterns
- Understand software testing approaches
- Understand the advantages of DevOps practices

REFERENCES:

1. Bernd Bruegge, Alan H Dutoit, Object-Oriented Software Engineering, 2nd edition, Pearson



PRINCIPAL
MOHAMED SATHAR A.S. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Sikkari, IT Park
Chennai-603 103.



MOHAMEDSATHAKAJCOLLEGE OFENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
4.	CP5201	NETWORK DESIGN AND TECHNOLOGIES	<ul style="list-style-type: none">• Bluetooth• Protocol STACK• Mobility Management and Power Optimization• 4G Networks• Green Wireless Networks


PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

OBJECTIVES:

- To understand the principles required for network design
- To explore various technologies in the wireless domain
- To study about 3G and 4G cellular networks
- To understand the paradigm of Software defined networks

UNIT I NETWORK DESIGN**10**

Advanced multiplexing - Code Division Multiplexing, DWDM and OFDM - Shared media networks - Switched networks - End to end semantics - Connectionless, Connection oriented, Wireless Scenarios -Applications, Quality of Service - End to end level and network level solutions. LAN cabling topologies - Ethernet Switches, Routers, Firewalls and L3 switches - Remote Access Technologies and Devices - Modems and DSLs - SLIP and PPP - Core networks, and distribution networks.

UNIT II WIRELESS NETWORKS**9**

IEEE802.16 and WiMAX - Security - Advanced 802.16 Functionalities - Mobile WiMAX - 802.16e - Network Infrastructure - WLAN - Configuration - Management Operation - Security - IEEE 802.11e and WMM - QoS - Comparison of WLAN and UMTS - Bluetooth - Protocol Stack - Security - Profiles

UNIT III CELLULAR NETWORKS**9**

GSM - Mobility Management and call control - GPRS - Network Elements - Radio Resource Management - Mobility Management and Session Management - Small Screen Web Browsing over GPRS and EDGE - MMS over GPRS - UMTS - Channel Structure on the Air Interface - UTRAN -Core and Radio Network Mobility Management - UMTS Security

UNIT IV 4G NETWORKS**9**

LTE - Network Architecture and Interfaces - FDD Air Interface and Radio Networks - Scheduling - Mobility Management and Power Optimization - LTE Security Architecture - Interconnection with UMTS and GSM - LTE Advanced (3GPP Release 10) - 4G Networks and Composite Radio Environment - Protocol Boosters - Hybrid 4G Wireless Networks Protocols - Green Wireless Networks - Physical Layer and Multiple Access - Channel Modelling for 4G - Introduction to 5G

UNIT V SOFTWARE DEFINED NETWORKS**9**

Introduction - Centralized and Distributed Control and Data Planes - Open Flow - SDN Controllers - General Concepts - VLANs - NVGRE - Open Flow - Network Overlays - Types - Virtualization - Data Plane - I/O - Design of SDN Framework

TOTAL : 45 PERIODS**OUTCOMES:**

Upon completion of this course, the students should be able to

- Identify the components required for designing a network
- Design a network at a high-level using different networking technologies
- Analyze the various protocols of wireless and cellular networks
- Discuss the features of 4G and 5G networks
- Experiment with software defined networks



Mohamed Sathak A.J.
PRINCIPAL
 MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
 34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
 Chennai-603 103.



MOHAMEDSATHAKAJCOLLEGE OF ENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
5.	CP5291	SECURITY PRACTICES	<ul style="list-style-type: none">• Cellular Network Security• Key establishment Protocols• Password based authenticated• Privacy Enhancing Technologies

PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

CP5291

SECURITY PRACTICES

L T P C
3 0 0 3

OBJECTIVES:

- To learn the core fundamentals of system and web security concepts
- To have through understanding in the security concepts related to networks
- To deploy the security essentials in IT Sector
- To be exposed to the concepts of Cyber Security and encryption Concepts
- To perform a detailed study of Privacy and Storage security and related Issues.

UNIT I SYSTEM SECURITY 9
Building a secure organization- A Cryptography primer- detecting system Intrusion- Preventing system Intrusion- Fault tolerance and Resilience in cloud computing environments- Security web applications, services and servers.

UNIT II NETWORK SECURITY 9
Internet Security - Botnet Problem- Intranet security- Local Area Network Security - Wireless Network Security - Wireless Sensor Network Security- Cellular Network Security- Optical Network Security- Optical wireless Security.

UNIT III SECURITY MANEGEMENT 9
Information security essentials for IT Managers- Security Management System - Policy Driven System Management- IT Security - Online Identity and User Management System - Intrusion and Detection and Prevention System.

UNIT IV CYBER SECURITY AND CRYPTOGRAPHY 9
Cyber Forensics- Cyber Forensics and Incidence Response - Security e-Discovery - Network Forensics - Data Encryption- Satellite Encryption - Password based authenticated Key establishment Protocols.

UNIT V PRIVACY AND STORAGE SECURITY 9
Privacy on the Internet - Privacy Enhancing Technologies - Personal privacy Policies - Detection of Conflicts in security policies- privacy and security in environment monitoring systems. Storage Area Network Security - Storage Area Network Security Devices - Risk management - Physical Security Essentials.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course the students should be able to

- Understand the core fundamentals of system security
- Apply the security concepts related to networks in wired and wireless scenario
- Implement and Manage the security essentials in IT Sector
- Able to explain the concepts of Cyber Security and encryption Concepts
- Able to attain a through knowledge in the area of Privacy and Storage security and related Issues.

REFERENCES:

1. John R.Vacca, Computer and Information Security Handbook, Second Edition, Elsevier 2013.
2. Michael E. Whitman, Herbert J. Mattord, Principal of Information Security, Fourth Edition, Cengage Learning, 2012.




PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.



MOHAMEDSATHAKAJCOLLEGE OF ENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
6.	CP5293	BIG DATA ANALYTICS	<ul style="list-style-type: none">• HDFS concepts• Real Time Analytics Platform (RTAP) Applications• Data Types and File Formats

PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

CP5293

BIG DATA ANALYTICS

L T P C
3 0 0 3

OBJECTIVES:

- To understand the competitive advantages of big data analytics
- To understand the big data frameworks
- To learn data analysis methods
- To learn stream computing
- To gain knowledge on Hadoop related tools such as HBase, Cassandra, Pig, and Hive for big data analytics

UNIT I INTRODUCTION TO BIG DATA

7

Big Data – Definition, Characteristic Features – Big Data Applications - Big Data vs Traditional Data - Risks of Big Data - Structure of Big Data - Challenges of Conventional Systems - Web Data – Evolution of Analytic Scalability - Evolution of Analytic Processes, Tools and methods - Analysis vs Reporting - Modern Data Analytic Tools.

UNIT II HADOOP FRAMEWORK

9

Distributed File Systems - Large-Scale FileSystem Organization - HDFS concepts - MapReduce Execution, Algorithms using MapReduce, Matrix-Vector Multiplication - Hadoop YARN

UNIT III DATA ANALYSIS

13

Statistical Methods:Regression modelling, Multivariate Analysis - Classification: SVM & Kernel Methods - Rule Mining - Cluster Analysis, Types of Data in Cluster Analysis, Partitioning Methods, Hierarchical Methods, Density Based Methods, Grid Based Methods, Model Based Clustering Methods, Clustering High Dimensional Data - Predictive Analytics – Data analysis using R.

UNIT IV MINING DATA STREAMS

7

Streams: Concepts – Stream Data Model and Architecture - Sampling data in a stream - Mining Data Streams and Mining Time-series data - Real Time Analytics Platform (RTAP) Applications - Case Studies - Real Time Sentiment Analysis, Stock Market Predictions.

UNIT V BIG DATA FRAMEWORKS

9

Introduction to NoSQL - Aggregate Data Models - Hbase: Data Model and Implementations - Hbase Clients - Examples - .Cassandra: Data Model - Examples - Cassandra Clients - Hadoop Integration. Pig - Grunt - Pig Data Model - Pig Latin - developing and testing Pig Latin scripts. Hive - Data Types and File Formats - HiveQL Data Definition - HiveQL Data Manipulation - HiveQL Queries

TOTAL: 45 PERIODS

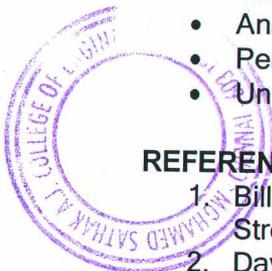
OUTCOMES:

At the end of this course, the students will be able to:

- Understand how to leverage the insights from big data analytics
- Analyze data by utilizing various statistical and data mining approaches
- Perform analytics on real-time streaming data
- Understand the various NoSql alternative database models

REFERENCES:

1. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", Wiley and SAS Business Series, 2012.
2. David Loshin, "Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph", 2013.
3. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, Second Edition,



Handwritten signature
MOHAMMED SAIFULLAH A.J. COLLEGE OF ENGINEERING
34, 1st/14th Ganga Road, Siruseri, IT
Chennai-600 103.



MOHAMEDSATHAKAJCOLLEGEENGINEERING

(Approved by AICTE, New Delhi and affiliated to Anna University, Chennai)

Siruseri IT Park, Egattur, Chennai - 603 103

S.No.	Subject Code	Subject Name	Topics Covered
7	CP5092	CLOUD COMPUTING TECHNOLOGIES	<ul style="list-style-type: none">• virtual clusters and Resource Management• Layered cloud Architectural Development• Setting up Hadoop Cluster• Cloud Security• and Trust Management


PRINCIPAL
MOHAMED SATHAK A.J.COLLEGE OF ENGINEERING
34, Rajiv Gandhi Road (OMR), Siruseri, IT Park
Chennai-603 103.

CP5092

CLOUD COMPUTING TECHNOLOGIES

L	T	P	C
3	0	0	3

OBJECTIVES:

- To understand the concepts of virtualization and virtual machines
- To gain expertise in server, network and storage virtualization.
- To understand and deploy practical virtualization solutions and enterprise solutions
- To gain knowledge on the concept of virtualization that is fundamental to cloud computing
- To understand the various issues in cloud computing
- To be able to set up a private cloud
- To understand the security issues in the grid and the cloud environment

UNIT I VIRTUALIZATION

9

Basics of Virtual Machines - Process Virtual Machines - System Virtual Machines -Emulation - Interpretation - Binary Translation - Taxonomy of Virtual Machines. Virtualization -Management Virtualization – Hardware Maximization - Architectures - Virtualization Management - Storage Virtualization - Network Virtualization

UNIT II VIRTUALIZATION INFRASTRUCTURE

9

Comprehensive Analysis - Resource Pool - Testing Environment -Server Virtualization - Virtual Workloads - Provision Virtual Machines - Desktop Virtualization - Application Virtualization - Implementation levels of virtualization - virtualization structure - virtualization of CPU, Memory and I/O devices - virtual clusters and Resource Management - Virtualization for data center automation.

UNIT III CLOUD PLATFORM ARCHITECTURE

9

Cloud deployment models: public, private, hybrid, community - Categories of cloud computing: Everything as a service: Infrastructure, platform, software- A Generic Cloud Architecture Design - Layered cloud Architectural Development - Virtualization Support and Disaster Recovery - Architectural Design Challenges - Public Cloud Platforms : GAE,AWS – Inter-cloud Resource Management

UNIT IV PROGRAMMING MODEL

9

Introduction to Hadoop Framework - Mapreduce, Input splitting, map and reduce functions, specifying input and output parameters, configuring and running a job -Developing Map Reduce Applications - Design of Hadoop file system -Setting up Hadoop Cluster - Cloud Software Environments -Eucalyptus, Open Nebula, Open Stack, Nimbus

UNIT V CLOUD SECURITY

9

Cloud Infrastructure security: network, host and application level – aspects of data security, provider data and its security, Identity and access management architecture, IAM practices in the cloud, SaaS, PaaS, IaaS availability in the cloud - Key privacy issues in the cloud -Cloud Security and Trust Management

TOTAL : 45 PERIODS




PRINCIPAL
MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING
34 Rajiv Gandhi Road (OMR), Siruseri, Chennai
Chennai-603 103.