POWER CENTRIC NODES IN CLUSTERED KNOWLEDGE NETWORKS
ENSURING OPTIMAL SECURITY

PhD Thesis

Submitted By:
Qaisar Javaid
Reg. No.1094-212001

Supervised By:
Prof. Merit. (CS) Dr. Mohammad Daud Awan
Dean Computer Science, Preston University

Co-Supervised By:
Dr. Syed Hussnain Abbas Naqvi
Chairman
Department of Computer Science & Software Engineering
International Islamic University, Islamabad.

DEPARTMENT OF COMPUTER SCIENCE
FACULTY OF COMPUTER SCIENCE
PRESTON UNIVERSITY, KOHAT
ISLAMABAD CAMPUS
2015
Dedication

I dedicate this thesis to my loving
Parents and family
Whose prayers and attention enabled me to reach this destination.
Acknowledgements

All praise be to Allah Almighty, the most merciful, the most gracious, without whose help and blessings, I would have been unable to complete the research thesis.

And, many thanks to my parents and supporters who were always there to pray for my success and kept my morale high. Then I am also most obliged and grateful to my tutors and supervisors for their active and inspiring guidance regarding the successful completion of my research work.
DECLARATION

I Solemnly declare that the dissertation “Power Centric Node in Clustered Knowledge Networks Ensuring Optimal Security” contained in the following pages is the result of my own work except where due acknowledgement or reference is made and that I have not used any material either published or otherwise to which legitimate claim by other exists.

QAISAR JAVAID

1094-212001

August, 2015
Abstract

Knowledge Networks are gaining momentum within cyber world. Knowledge leads to innovation and for this reason organizations focus on research and information gathering in order to gain and improve existing knowledge. In this age of information, which is primarily based on internet and world wide web technologies, enable significantly expanded networks of people to communicate and collaborate ‘virtually’ across teams of the entire organization, across organizations and the world, anytime and anyplace. Therefore, wonderful innovations in computing and telecommunications have transformed the corporations from structured and manageable type to interwoven network of blurred boundaries such as; ad hoc networks and mobile wireless networks, semantic web technology etc. This study explores the means of knowledge networks in Information Technology and security leaks that are found and measures that are taken to counter this menace are coming up with optimal Secure Clustered Power Centric node network. The study estimates these measures, evaluating and integrating them to come up with a secured network design.

This study seeks to establish these threats and attacks and counter security measures in Wireless Sensor Networks, Mobile Ad-hoc Networks and Semantic Web technology and propose architecture for the purpose of establishing secure routing in knowledge networks called Optimized Network Security Solution for Knowledge Networks – ONSS-KN.

Design and development of power-aware, scalable and performance efficient routing protocols for Knowledge Networks (KNs) is an active area of research. In this dissertation, we show that insect colonies based intelligence – commonly referred to as Swarm Intelligence (SI) – provides an ideal metaphor for developing routing protocols for KNs because they consist of minimalists, autonomous individuals that through local interactions self-organize to produce system-level behaviors that show lifelong adaptability to changes and perturbations in an external environment. In this context, we propose a clustering and secure data routing protocol for KNs (iRoutCluster) - inspired by the foraging principles of honey bees.
To conclude this study, we perform simulation studies to analyze and compare the performance of the final *iRoutCluster* design with existing protocols. The simulation results demonstrate that *iRoutCluster* outperforms its competitors in all assumed scenarios and metrics. We then implement the *iRoutCluster* protocol in MATLAB® simulator to further investigate its performance in mobile networks and large-scale static sensor networks. The results clearly show that *iRoutCluster* not only performs well in large-scale networks, but is equally good in MANETs as well. Therefore, *iRoutCluster* is a viable protocol for hybrid ad hoc networks.
# Table of Contents

**Chapter 1 Thesis Introduction** ................................................................. 1

1.1 Knowledge Networks ................................................................. 1

1.1.1. Sensor Networks ................................................................. 3

1.1.2. Mobile Ad-hoc Networks ....................................................... 3

1.1.3. Wireless Networks ............................................................... 4

1.1.4. Semantic Web ........................................................................ 5

1.1.5. Similarities and Differences of Knowledge Networks ............... 5

1.2 Research Problem and Research Questions ..................................... 7

1.3 Thesis Motivation, Objectives, Methodology and Significance ............ 8

1.3.1. Motivation ............................................................................. 9

1.3.2. Research Objectives .............................................................. 11

1.3.3. Research Methodology .......................................................... 11

1.3.3.1 Research Assumptions ....................................................... 12

1.3.3.2 Research Techniques .......................................................... 12

1.3.3.3 Tools of Research ............................................................... 13

1.3.3.4 Data Collection and Analysis .............................................. 13

1.3.4. Significance of Study ............................................................. 13

1.4 Major Contributions of This Study ................................................... 13


1.4.2. Intrusion Detection Based Security Scheme for Cluster-Based Wireless Networks ............................................................. 14

1.4.3. Verification of Proposed Network Design .................................. 14

1.4.4. Simulation Using MATLAB ...................................................... 15

1.4.5. A Bee-Inspired Routing Protocol for Knowledge Networks to Attain the Clustering Optimization- *iRoutCluster* ................................................. 16

1.5 Organization of Thesis ...................................................................... 17

References............................................................................................... 20
Chapter 2 Problem Evaluation, Security Optimization and Routing Methodology

2.1 Literature Review and Evaluating Problem Domain

2.1.1. Intelligent Routing

2.1.2. Group Authentication

2.1.3. Data Integrity and Confidentiality

2.1.4. Intrusive Behavior, Detection and Prevention

2.1.5. Botnet Boost-up Detection System

2.1.6. Creating a Robust Network Society in Sweden

2.1.7. Security Framework

2.1.8. Multistage Attack Recognition System (MARS)

2.1.9. Centralized Host-Based Security Scanning Architecture

2.1.10. Clustered Networks

2.1.11. Simulation Model

2.1.12. Turing Assessor

2.1.13. Supervisory Control and Data Acquisition Systems (SCADA) – A test bed

2.1.14. IPSec Network Protocol

2.1.15. IPSec Policy Architecture

2.2 Optimized Network Security Solution for Knowledge Networks

2.2.1. Monitoring for Intrusive Activities

2.2.2. IPSec Security Protocol

2.3 System Model of Optimized Security Solution

2.3.1. Cluster of Nodes

2.3.2. Clustered Networks

2.3.3. Power Centric Nodes

2.4 Clustering and Secure Routing Algorithm for Knowledge Network

- iRoutcluster

2.4.1. Complying with Knowledge Networks

2.4.2. Deployment is Not Pre-Determined

2.4.3. Pre-Processing Data
Chapter 3 Sensor Networks Security and Routing Technique ................. 47

3.1 Introduction ............................................................................. 47
3.2 Wireless Sensor Networks ..................................................... 48
3.3 Communication Architecture of Sensor Node ............................. 49
   3.3.1. Hardware Framework of a Sensor Node ............................ 50
   3.3.2. Attribute of Sensor Networks ........................................ 52
3.4 Network Topology of Sensor Networks ..................................... 52
3.5 Wireless Sensor Network Security .......................................... 54
   3.5.1. Security Requirements of WSN ................................. 54
   3.5.2. Wireless Sensor Network Communication Protocols ........... 55
   3.5.3. Attacks on WSN and Their Mitigation ........................... 57
3.6 Routing Challenges and Design Issues in WSNs ................. 61
   3.6.1. Node Deployment ......................................................... 62
   3.6.2. Energy Consumption without Losing Accuracy ............... 62
   3.6.3. Link/Node Heterogeneity ........................................... 62
   3.6.4. Data Reporting Model ............................................... 62
   3.6.5. Fault Tolerance .......................................................... 63
   3.6.6. Scalability ................................................................. 63
   3.6.7. Network Dynamics ..................................................... 63
   3.6.8. Transmission Media .................................................... 63
   3.6.9. Data Aggregation ......................................................... 64
   3.6.10. Coverage ................................................................. 64
   3.6.11. Quality of Service ..................................................... 64
3.7 Clustering Routing Protocol in Wireless Sensor Network .......... 64
3.8 Classification of Clustering Elements in WSNs ..................... 65
   3.8.1. Cluster Attributes ...................................................... 65
   3.8.2. Cluster-Head Characteristics ....................................... 66
   3.8.3. Clustering Process ..................................................... 66
3.8.4. Entire Proceeding of Algorithm ........................................... 67
3.9 Range-Based and Range-Free Algorithms .................................. 69
3.10 Pairwise Key Distribution Scheme ......................................... 70
3.11 Coverage in a Wireless Sensor Network ................................... 71
3.12 TDMA (Time Division Multiple Access) Transmissions .............. 72
3.13 Wireless Body Area Networks (WBANs) ................................. 73
3.14 Smart Sensors and Actuator for Power Management in Intelligent Buildings ......................................................... 74
3.15 Chapter Conclusion ................................................................. 75
References ...................................................................................... 75

Chapter 4 Mobile Ad-hoc Network Security and Routing Protocol .... 78

4.1 Introduction ............................................................................. 78
4.2 Mobile Ad-hoc Networks ......................................................... 79
4.3 Topology Management ............................................................... 81
  4.3.1. 4G and Ad-hoc networking .................................................. 82
  4.3.2. Bluetooth Networking ........................................................ 83
4.4 Mobile Ad-hoc Network (MANET) Architecture ...................... 83
4.5 Mobile Ad-hoc Network Security ............................................. 85
  4.5.1 Vulnerabilities of the MANETs ............................................ 87
  4.5.2 Security Attacks ................................................................. 88
4.6 Security Scheme in MANET ...................................................... 89
  4.6.1 Intrusion Detection Techniques .......................................... 90
  4.6.2 Secure Routing Techniques in Mobile Ad Hoc Network ....... 90
4.7 Security Solutions to the Mobile Ad-hoc Networks .................... 92
  4.7.1 Preventing External Attacks : Key and Trust Management ...... 93
    4.7.1.1 Trust model ................................................................. 93
    4.7.1.2 Cryptosystems .......................................................... 93
    4.7.1.3 Key Creation and Key Storage ..................................... 93
    4.7.1.4 Key Distribution ........................................................ 94
  4.7.2 Preventing Internal Attacks: Secure Routing Protocols ......... 94
    4.7.2.1 Dynamic Source Routing (DSR) .................................... 95
    4.7.2.2 Advanced On-demand Distance Vector (AODV) .......... 95
4.7.2.3 Secures Advanced On-demand Distance Vector (SAODV) ..... 95
4.7.2.4 Optimized Link State Routing (OLSR) .......................... 96

4.8 Routing Protocol in MANETs ........................................ 96

4.9 Clustering in Mobile Ad-hoc Network ................................. 98

4.9.1 Algorithms for Cluster Heads Election in MANETs .............. 99

4.9.2 Clustering Schemes in Mobile Ad-hoc Network ................. 99

4.9.2.1 Identifier Neighbor Based Clustering ............................ 99

4.9.2.2 Topology Based Clustering ................................... 99

4.9.2.3 Mobility Based Clustering .................................. 100

4.9.2.4 Energy Based Clustering .................................. 100

4.9.2.5 Weight Based Clustering .................................. 101

4.10 Chapter Conclusion .................................................. 101

References ......................................................................... 102

Chapter 5 Semantic Web Model and Security Protocol ................. 106

5.1 Introduction .................................................................... 106

5.2 Semantic Web Technology ........................................... 107

5.3 Languages Representing Knowledge on the Web .................. 107

5.3.1 Web Semantic Protocol ...................................... 108

5.3.2 Software Component .......................................... 109

5.4 Security and Encryption .............................................. 110

5.4.1 Security Standards for the Semantic Web ...................... 111

5.4.2 Semantic Web Security Foundations ............................ 112

5.4.3 Injection Attack in Semantic Query Language ................. 113

5.5 Key Issues in Recognizing the Semantic Web Vision ............. 113

5.6 Web Semantic for Decision Support Systems (DSS) ............... 114

5.7 Online Encyclopedia Wikipedia .................................... 117

5.8 Web 2.0 and Semantic Web Technologies .......................... 119

5.9 Knowledge Extraction on the Semantic Web ...................... 120

5.10 Data and Information Integration ................................... 121

5.11 Watson System .................................................. 122

5.12 NetKAT .................................................. 124
5.13 Chapter Conclusion ................................................................. 125
References .................................................................................. 126

Chapter 6 Cluster Based Security Architecture and Communication Protocols for
Ad-hoc Wireless Networks ............................................................. 130
6.1 Introduction ........................................................................... 130
6.2 Related Work ........................................................................ 133
6.3 Cluster Formations ................................................................. 136
6.4 Classification of Clustering Algorithm ................................. 138
   6.4.1 Probabilistic Algorithms .................................................. 139
   6.4.2 Non-Probabilistic Algorithms ........................................... 143
6.5 A Cluster-Based Security Architecture for Ad-hoc Wireless Networks ..... 145
   6.5.1 Authentication Using Distributed Certification .................. 145
   6.5.2 Integrity and Confidentiality : Secure Data Routing ............ 149
   6.5.2.1 Secure Data Routing Model ....................................... 149
   6.5.2.2 Routing Algorithm .................................................... 150
6.6 Intrusion Detection Based Security Scheme for Cluster-Based Wireless
   Sensor Networks ........................................................................ 153
   6.6.1 Intrusion Detection Based Security Scheme ....................... 154
   6.6.2 Design of the Wireless Sensor Network ......................... 154
   6.6.3 Expected Intrusion Scenarios ......................................... 155
6.7 Security of Cluster Based Communication Protocols ................. 159
   6.7.1 Secure Routing in Cluster Based WSNs Using Symmetric
       Cryptography with Session Keys ........................................... 159
   6.7.2 Comparison of Low Energy Adaptive Clustering Hierarchy and Secure
       Routing in Clustered Based WSNs ....................................... 159
6.8 Chapter Conclusion ............................................................... 162
6.9 References ............................................................................. 163

Chapter 7 BCO – A Simple, Scalable and Energy Efficient Routing Protocol
For WSNs & MANETs ................................................................. 165
7.1 Introduction ............................................................................ 165
7.2 The Clustering Problem ............................................................... 168
7.3 Artificial Bee Colony Algorithm .................................................. 169
  7.3.1 Honey Bee Modeling .............................................................. 169
    7.3.1.1. Mathematical Model ......................................................... 171
    7.3.1.2 Steps of Algorithm ............................................................ 171
  7.3.2 The ABC Algorithm for Clustering Problems ................................. 173
7.4 WSN Routing Protocols Using ABC Algorithm and Hierarchal WSN
    Clustering Approach (LEACH) ..................................................... 177
  7.4.1 Theoretical and Mathematical Model ........................................... 177
  7.4.2 Analysis and Result ............................................................... 183
7.5 Fitness Function for Artificial Bee Colony ........................................ 185
7.6 Chapter Conclusion .................................................................... 187
References ....................................................................................... 188

Chapter 8 Design Review, Simulation and Performance Analysis of Clustered
    Network Design – (ONSS-KN) .............................................................. 190
8.1 Introduction .................................................................................. 190
8.2 ABC Algorithm in Deployment for Wireless Networks – iRoutCluster-I ... 194
  8.2.1 Modified ABC Algorithms and Its Application in Deployment ............ 195
    8.2.1.1 Modification in Onlooker Bee Phase ........................................ 195
    8.2.1.2 Modification in Scout Bee Phase .............................................. 196
  8.2.2 Simulation Results ..................................................................... 197
    8.2.2.1 Parameter Selection for the Proposed Algorithm ...................... 197
    8.2.2.2 Simulations in An Ideal Room ............................................... 198
8.3 Design Review and Performance Analysis of Clustering-iRoutCluster-II ..... 200
  8.3.1 Data Transmission ..................................................................... 200
  8.3.2 Cluster Network Design and Simulation ....................................... 201
  8.3.3 Experimental Setup ................................................................... 206
    8.3.3.1 WIFI- WEP Authentication .................................................. 206
    8.3.3.2 Cellular and Wireless Path - Wireless Network ....................... 207
    8.3.3.3 Simulation of IPsec- Security Protocol .................................... 208
  8.3.4 Proposed Solution - Artificial Bee Colony Algorithm (ABC) ............ 208
Chapter 8 Simulation and Analysis

8.3.5. Network Model ................................................................. 209
8.3.6. Simulation and Analysis ...................................................... 211
  8.3.6.1. Thresholds for Clustering ........................................... 211
  8.3.6.2. Optimum Number of Clusters .................................... 213
  8.3.6.3. Parameters to Evaluate the Performance of the Network...... 215
  8.3.6.4. Energy Consumption .................................................. 217
  8.3.6.5. Findings ................................................................... 218

8.4 Comparative Analysis of Swarm Intelligence Algorithms .................. 219
8.5 Chapter Conclusion .................................................................. 220
References .................................................................................. 221

Chapter 9 Conclusion and Future Work .............................................. 223
9.1 Summary of Important Findings ............................................... 223
9.2 Future Work: Suggestion and Directions .................................... 225
Appendixes .................................................................................. 227
# List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>Artificial Bee Colony</td>
</tr>
<tr>
<td>ACO</td>
<td>Ant Colony Optimization</td>
</tr>
<tr>
<td>ATA</td>
<td>Asynchronous Turing Assessor</td>
</tr>
<tr>
<td>BCO</td>
<td>Bee Colony Optimization</td>
</tr>
<tr>
<td>BDD</td>
<td>Bad Data Detector</td>
</tr>
<tr>
<td>CSMA</td>
<td>Carrier Sense Multiple Acess</td>
</tr>
<tr>
<td>DNS</td>
<td>Domain Name System</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision Support System</td>
</tr>
<tr>
<td>EAR</td>
<td>Energy-Aware Routing</td>
</tr>
<tr>
<td>ECC</td>
<td>Elliptic Curve Cryptosystems</td>
</tr>
<tr>
<td>EP</td>
<td>Evolutionary Programming</td>
</tr>
<tr>
<td>ESARP</td>
<td>An Efficient Security Aware Routing Protocol</td>
</tr>
<tr>
<td>GPSR</td>
<td>Greedy Perimeter Stateless Routing Protocol</td>
</tr>
<tr>
<td>GSR</td>
<td>Gigabit Switch Router</td>
</tr>
<tr>
<td>IBPSO</td>
<td>Ionic Bond Directed Particle Swarm Optimization</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technologies</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>KE</td>
<td>Knowledge Extraction</td>
</tr>
<tr>
<td>KNs</td>
<td>Knowledge Networks</td>
</tr>
<tr>
<td>LD</td>
<td>Linked Data</td>
</tr>
<tr>
<td>LEACH</td>
<td>Low Energy Adaptive Clustering Hierarchy</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MANETs</td>
<td>Mobile Ad-hoc Networks</td>
</tr>
<tr>
<td>MARS</td>
<td>Multistage Attack Recognition System</td>
</tr>
<tr>
<td>MCG</td>
<td>Maximum Cycle Number</td>
</tr>
<tr>
<td>MPR</td>
<td>Multi-Point Relays</td>
</tr>
<tr>
<td>NCS</td>
<td>Network Controlled Systems</td>
</tr>
<tr>
<td>NER</td>
<td>Named Entity Recognition</td>
</tr>
<tr>
<td>NetKAT</td>
<td>Network Kleene Algebra with Tests</td>
</tr>
<tr>
<td>NIDS</td>
<td>Network Intrusion Detection Systems</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>NIPS</td>
<td>Network Intrusion Prevention Systems</td>
</tr>
<tr>
<td>NLP</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>ONSS-KN</td>
<td>Optimized Network Security Solution for Knowledge Networks</td>
</tr>
<tr>
<td>OPF</td>
<td>Optimal Power Flow</td>
</tr>
<tr>
<td>OWL</td>
<td>Web Ontology Language</td>
</tr>
<tr>
<td>PARP</td>
<td>Privacy Aware Routing Protocol</td>
</tr>
<tr>
<td>PDA</td>
<td>Personal Digital Assistant</td>
</tr>
<tr>
<td>PEGASIS</td>
<td>Power-Efficient Gathering in Sensor Information Systems</td>
</tr>
<tr>
<td>PKI</td>
<td>Public Key Infrastructure</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RIF</td>
<td>Rule Interchange Format</td>
</tr>
<tr>
<td>SAR</td>
<td>Sequential Assignment Routing</td>
</tr>
<tr>
<td>SCADA</td>
<td>Supervisory Control and Data Acquisition Systems</td>
</tr>
<tr>
<td>SE</td>
<td>State Estimator</td>
</tr>
<tr>
<td>SI</td>
<td>Swarm Intelligence</td>
</tr>
<tr>
<td>SPINE</td>
<td>Secure Positioning for Sensor Networks</td>
</tr>
<tr>
<td>SRCWSNS</td>
<td>Secure Routing in Cluster Based WSNs Using Symmetric Cryptography with Session Keys</td>
</tr>
<tr>
<td>SSH</td>
<td>Secure Shell</td>
</tr>
<tr>
<td>TBF</td>
<td>Trajectory Based Forwarding</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>TLS</td>
<td>Transport Layer Security</td>
</tr>
<tr>
<td>URI</td>
<td>Uniform Resource Identifiers</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
<tr>
<td>WAP</td>
<td>Wireless Application Protocol</td>
</tr>
<tr>
<td>WOL</td>
<td>Web Ontology Language</td>
</tr>
<tr>
<td>WSD</td>
<td>Web Services for Devices</td>
</tr>
<tr>
<td>WSN</td>
<td>Wireless Sensor Networks</td>
</tr>
<tr>
<td>XML</td>
<td>eXtensible Markup Language</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1 Distribution of node degree in different topologies with different bandwidths ................................................................. 3
Figure 1.2 Thesis route map ................................................................. 19
Figure 2.1 Optimized Security Solution Process Model and Architecture .................................................................................. 32
Figure 2.2 Botnet Detection System ........................................................................................................................................ 34
Figure 2.3 Clustered Network Design with Cluster Head ........................................................................................................ 38
Figure 2.4 Communication of Sensor to Cluster Node ..................................................................................................................... 39
Figure 2.5 Power Centric Node of Cluster Network Design ..................................................................................................... 40
Figure 2.6 COPE increasing throughput by limiting transmission of a pair of packets to 3 transmissions rather 4 transmissions ................. 44
Figure 3.1 Sensor node spotted in a sensor area ......................................................................................................................... 50
Figure 3.2 The component of a sensor Unit .................................................................................................................................. 50
Figure 3.3 Sensor node hardware framework ................................................................................................................................... 51
Figure 3.4 Primary topologies of wireless sensor networks ...................................................................................................... 53
Figure 3.5 Layered Architecture of WSN ...................................................................................................................................... 56
Figure 3.6 Sybil Attack ................................................................................................................................................................. 59
Figure 3.7 Blackhole/ Sinkhole Attack .......................................................................................................................................... 59
Figure 3.8 Wormhole Attack .......................................................................................................................................................... 60
Figure 3.9 Clustering Methods in WSNs .......................................................................................................................................... 68
Figure 3.10 Coverage Dilemma ..................................................................................................................................................... 72
Figure 3.11 Intra-body and extra-body communication ............................................................................................................. 73
Figure 3.12 Place of a WBAN in the area of wireless networks .................................................................................................. 74
Figure 3.1-a Wireless LAN, Ad-hoc network versus MANET ........................................................................................................ 80
Figure 3.1-b MANET - At Airport ................................................................................................................................................... 81
Figure 4.1 4G Network ................................................................................................................................................................. 82
Figure 4.2 A simple MANET architecture ....................................................................................................................................... 84
Figure 4.3 Wireless networks mapping ........................................................................................................................................ 86
Figure 4.4 Emerging protocol and approaches for authentication and access control - MANET ........................................................................ 86
Figure 4.5 Establish public key infrastructure authentication and access control protocols Fixed Networks ........................................ 86
Figure 4.6 Rush Attack in Mobile ad-hoc network ......................................................................................................................... 91
Figure 4.7 Combined Mechanisms to Secure MANET against Rushing Attacks ......................................................................................... 91
Figure 4.8 Classification of MANET routing protocols ................................................................................................................. 97
Figure 4.9 Semantic Web Language and Protocol Standards ...................................................................................................... 108
Figure 5.1 Semantic Web Components Architecture .................................................................................................................. 109
Figure 5.2 Layers for Secure Semantic Web ........................................................................................................................................ 111
Figure 5.3 Fraction of the publications each year treating different types of DSS ........................................................................... 115
Figure 5.4 Type of study reported .................................................................................................................................................... 116
Figure 5.5 Number of publications reporting the use of specific Semantic Web technologies ......................................................... 117
Figure 5.6 Wikipedia Miner toolkit Architecture .......................................................................................................................... 118
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8-a</td>
<td>Impact of thresholds for disambiguation and detection</td>
</tr>
<tr>
<td>5.8-b</td>
<td>Impact of probability thresholds for performance of the detection</td>
</tr>
<tr>
<td>5.9</td>
<td>Hybridization of formal and linguistic knowledge in semantic technologies</td>
</tr>
<tr>
<td>5.10</td>
<td>Overview of the Watson architecture</td>
</tr>
<tr>
<td>5.11</td>
<td>Overview of the Watson Web interface</td>
</tr>
<tr>
<td>6.1</td>
<td>One hop model</td>
</tr>
<tr>
<td>6.2</td>
<td>Multi-hop model</td>
</tr>
<tr>
<td>6.3</td>
<td>Cluster based model</td>
</tr>
<tr>
<td>6.4</td>
<td>CH-5 shifted into cluster 2 it take off its role as CH according to its higher ID. Nodes A and B which lost their CH</td>
</tr>
<tr>
<td>6.5</td>
<td>Simple Architecture of clustered Senor Network</td>
</tr>
<tr>
<td>6.6</td>
<td>Authentication Process</td>
</tr>
<tr>
<td>6.7</td>
<td>States of a New Node during Log-On</td>
</tr>
<tr>
<td>6.8</td>
<td>Secure Data Transmission for ONSS-KN</td>
</tr>
<tr>
<td>6.9</td>
<td>Design of Wireless Senor Network</td>
</tr>
<tr>
<td>6.10</td>
<td>Data communication</td>
</tr>
<tr>
<td>6.11</td>
<td>Data communication between CHs and BS</td>
</tr>
<tr>
<td>6.12</td>
<td>Intruder introduced in the network</td>
</tr>
<tr>
<td>6.13</td>
<td>Intruder communicating with its nearest node</td>
</tr>
<tr>
<td>6.14</td>
<td>Intruder trying to communicate with node N9 and established in listen mode</td>
</tr>
<tr>
<td>6.15</td>
<td>Intruder tries to create duplicate of N9 node</td>
</tr>
<tr>
<td>6.16</td>
<td>Un-known MAC address identified by Cluster Head</td>
</tr>
<tr>
<td>6.17</td>
<td>Intruder communicating with Cluster Head CH-1</td>
</tr>
<tr>
<td>6.18</td>
<td>Intruder communicating with Cluster Head</td>
</tr>
<tr>
<td>6.19</td>
<td>Intruder creates a copy of CH-1</td>
</tr>
<tr>
<td>6.20</td>
<td>Un-known MAC address identified by Base Station</td>
</tr>
<tr>
<td>6.21</td>
<td>Base Station identifies intruder and publicize alert</td>
</tr>
<tr>
<td>6.22</td>
<td>Mitigation of Sinkhole Attack</td>
</tr>
<tr>
<td>7.1</td>
<td>Hive and the kind of bees involved in the process</td>
</tr>
<tr>
<td>7.2</td>
<td>Major steps of ABC Algorithms</td>
</tr>
<tr>
<td>7.3</td>
<td>The flow chart of ABC algorithm used for clustering</td>
</tr>
<tr>
<td>7.4</td>
<td>Main steps of Protocol</td>
</tr>
<tr>
<td>7.5</td>
<td>Gathering of advertisement messages and calculation of distances</td>
</tr>
<tr>
<td>7.6</td>
<td>Data gathering – direct communication</td>
</tr>
<tr>
<td>7.7</td>
<td>Data gathering using cluster method</td>
</tr>
<tr>
<td>7.8</td>
<td>Distribution of Sensor using LEACH Algorithm</td>
</tr>
<tr>
<td>7.9</td>
<td>Distribution of sensor node using ABC Algorithm</td>
</tr>
<tr>
<td>7.10</td>
<td>Number of Received Items versus Residual Network Energy</td>
</tr>
<tr>
<td>7.11</td>
<td>Residual Network Energy versus Round</td>
</tr>
<tr>
<td>8.1</td>
<td>Architecture of ONSS-KN – Optimized Network Security Solution for Knowledge Networks</td>
</tr>
<tr>
<td>8.2</td>
<td>Nodes Distribution: (a) initial, (b) 100 iterations, (c) 500 iterations, (d) 1000 iterations, and (e) 10000 iterations</td>
</tr>
<tr>
<td>8.3</td>
<td>Coverage rate comparisons between FNF-BL-ABC, original ABC</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>8.4-a</td>
<td>Cisco network design model.</td>
</tr>
<tr>
<td>8.4-b</td>
<td>Power Centric Node Resides on Core Layer and Distribution Layer security of the Network Infrastructure.</td>
</tr>
<tr>
<td>8.5</td>
<td>Design and Simulation of A cluster – Ad-hoc Mobile and Wireless Network.</td>
</tr>
<tr>
<td>8.6</td>
<td>Cluster member – In a Cluster.</td>
</tr>
<tr>
<td>8.7</td>
<td>WIFI- WAP Authentication.</td>
</tr>
<tr>
<td>8.8</td>
<td>Cellular and Wireless Path for PDA device.</td>
</tr>
<tr>
<td>8.9</td>
<td>IPSec Protocol Simulation – Network Design Model.</td>
</tr>
<tr>
<td>8.10-a</td>
<td>Total Energy Dissipation for various $E_{elec}$ values.</td>
</tr>
<tr>
<td>8.10-b</td>
<td>Total Energy Dissipation for various $E_{amp}$ values.</td>
</tr>
<tr>
<td>8.11</td>
<td>Hop Count against Network Size of each cluster.</td>
</tr>
<tr>
<td>8.12</td>
<td>Calculated Response time between two clusters.</td>
</tr>
<tr>
<td>8.13</td>
<td>Success rate achieved keeping limited to a considerable network size within a cluster.</td>
</tr>
<tr>
<td>8.14</td>
<td>Optimum number of clusters with respect to cluster headset size.</td>
</tr>
<tr>
<td>8.15</td>
<td>Number of Clusters with Respect to Distance from the Base Station for various Number of Cluster Head-set sizes.</td>
</tr>
<tr>
<td>8.16</td>
<td>Throughput in ABC as per number of Round.</td>
</tr>
<tr>
<td>8.17</td>
<td>Delay factor in ABC as per number of Round.</td>
</tr>
<tr>
<td>8.18</td>
<td>Packet Lost in ABC as per number of Round.</td>
</tr>
<tr>
<td>8.19</td>
<td>Packet Delivery Ratio.</td>
</tr>
<tr>
<td>8.20-a</td>
<td>Optimum Number of Clusters size 1.</td>
</tr>
<tr>
<td>8.20-b</td>
<td>Optimum Number of Clusters size -3.</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1: Similarities and differences among four types of knowledge networks.... 5
Table 3.1: Attributes of Sensor Networks......................................................... 52
Table 4.1: MANETs security issues on different layers........................................ 87
Table 4.2: Solutions to Secure Routing Protocols............................................. 94
Table 4.3: Properties of proactive and Reactive Routing.................................... 98
Table 6.1: Pros and Cons of Different Secure Routing Protocols......................... 160
Table 7.1: Mapping of Biological Terminology with Network Terminology........... 169
Table 7.2: Field of distance vector ................................................................. 178
Table 7.3: A Solution Array.............................................................................. 184
Table 8.1: Corresponding parameters in ABC algorithm and wireless node
deployment...................................................................................................... 194
Table 8.2: Deployment results comparison between 3 algorithms at some
iterations........................................................................................................... 197
Table 8.3: Parameters of the ABC algorithm as used in.................................... 208
Table 8.4: Parameter values used for experiment............................................ 208
Table 8.5: Parameters of wireless sensor network........................................... 214
Table 8.6: Comparison between LEACH, PSO and ABC................................. 218
Research List of Publications


Chapter 1

Thesis Introduction

This study presents a Secure Energy-efficient Cluster-based Data Routing Protocol for KNs based on Symmetric Cryptography with Session Keys. The network is divided into a number of clusters with selected cluster head sets within each cluster. The cluster head-set members are responsible for control and management of the network. On rotation basis, a cluster head-set member receives data from its cluster nodes and transmits the aggregated results to the distant base station. Communication between node and base station takes place at four levels; node → cluster head → gateway → base station. Encryption of the sensed data is transmitted to the cluster head, which aggregates the data received from the network nodes of its cluster before forwarding to the next cluster head on the path or through its gateway to the base station.

1.1 Knowledge Networks

Network is made up of connecting different nodes or vertices that are correlated to each other. Internet or World Wide Web is combination of millions even billions of vertices. Virtual Knowledge Networks are good to form interactive learning mechanism, promoting innovation and bringing new advantages with it. Virtual or Cyber Knowledge Networks consist of different mediums; social networking (blogs, twitter, 2015 etc.), mobile networks (such as iPhone) that converts the collected information to semantic network through interlinking agents [1]. Knowledge networks are highly complex networks in terms of transportation of data and mobility [2]. There are various questions that arise [3] like an analyst can ask: which node would prove most crucial to network connectivity if it is removed? This probability would be minimal when talking about such a large network. To draw a meaningful picture and form an understanding from all these vertices is unbelievable. Thus analysts try to convert these connections to statistical data that would tell path lengths and degree distributions to help in measuring network properties and structure. Network models are created to understand structure and how the
interactions are made in such a huge network. Then it determines the behavior of this
network, for example, how the network structure would affect the traffic on the Internet?
In the seven layers OSI model the third layer is network layer that is responsible for
packet forwarding including routing through intermediate routers.

When studying the behavior of these networks, there are threats of security
breaches that endanger the secure transmission of information. Topology of the Internet
has to be first understood to address the criticalities of this complex infrastructure. First-
principles theory [4] reflects the constraints and tradeoffs in network topology. It asserts a
plain model that links bandwidth and connectivity into the hard technological constraints
jointly with conceptual models defining network performance and user demand. Cisco
Gigabit Switch Router - 12416 is limited by its bandwidth and number of available m-
line-card connections has 15 for which it is configured and throughput as per degree is
constrained by maximum line-card speed of 10Gbps. Number of connections increases
as total bandwidth increases but when the numbers of connections exceed 15 it starts to
degrade. There can be maximum possible connections up-to 120 for this router but
connections exceeding 15 are not recommended. There are different types of bandwidths
requests as per user ranging from 56 Kbps to very low number of users requesting 10
GBPS. Thus, considering these requirements the network has to be designed and built.
For the network solution providers the economic constraints also pose the limitations of
the type of network deployment. They are influenced both by the number of users on the
network and their increasing bandwidth requests with time.
It has been clarified in our problem statement that knowledge networks may be defined in
four categories that are:

1. Sensor networks,
2. Mobile ad-hoc Networks (MANET),
3. Wireless networks, and
4. Semantic web

Moving forward in our research it is necessary to understand the context of these
network types and security considerations that are prevalent in these networks and
measures are being taken so far to add security.
Figure 2.1 Distribution of node degree in different topologies with different bandwidths.
   a. Regularly high bandwidth users
   b. Highly irregular bandwidth
   c. Consistently low bandwidth (Li et al., 2004)

1.1.1. Sensor Networks

Sensor networks have been realized with the increasing use of wireless communications through electronic devices. The sensor networks are used within many scenarios (health, military, home, etc.). What is a sensor network? It is a collection of several sensor nodes and its deployment is not predetermined. Thus, requesting the algorithms and protocols engineered to tackle these sensor nodes must be intelligent enough to enable the nodes to manage themselves. The processors of sensor nodes are intelligent enough to preprocess the raw data and send only the required and partially processed data. Sensor networks do require wireless ad-hoc networking techniques. But there is a difference between ad-hoc networks and sensor networks. Main difference that would help distinguish the two is that ad-hoc networks is a point to point communication while sensor networks use broadcast communication and are densely deployed at higher orders of magnitude than the ad-hoc networks.

1.1.2. Mobile Ad-hoc Networks (MANET)

As discussed earlier Knowledge networks are inclusive of Mobile Ad hoc Network (MANET). This is inclusive of autonomous self-organized nodes using wireless medium for connecting two ends for communication that are within each other’s
transmission radius in a multi hop infrastructure. There has been progressive research conducted in the past that proposed various routing algorithms for MANETs but the recent study shows that the emerging trend in routing research is towards Swarm Intelligence (SI).

Swarm Intelligence (SI) refers to an artificial intelligence technique implemented on a complex system that is meant to be decentralized and self-organized. It could result in observing intelligent and complex behavior emerging from simple unsupervised interaction of all swarm members part of a system or any network. There are various routing techniques based on Swarm Intelligence (SI) two of these are (i) Ant colony optimization (ACO) algorithm based routing protocols, and (ii) Bee colony optimization (BCO) based routing protocols. These algorithms are inspired by ants and bees behavior in their colonies. ACO based on ant’s characteristics of searching food is applied as such on MANETs for routing packets. The shortest path is found for routing packets. If it is found dead then an alternative path is utilized and likewise the routing packets adapt intelligently to the dynamically changing environment. Forwarded packets leave their track for following packets along the same path. ACO and BCO differ in their principles based on the different nature of ants and bees as ants walk and bees fly. As with ants, ACO lets the forwarding packets leave their trail behind them while in BCO, visual communication plays the same role. BCO based on bee characteristics works as a well-knitted team work, coordination, and simultaneous task performance. Such nature inspired routing protocols would be capable of removing at least one of several problems like; survivability, scalability, adaptability, maintainability, battery life, and further more.

1.1.3. Wireless Networks

Wireless communication has gained momentum in the twenty first century. And, with the growth in cellular use, and other networks within businesses and homes the wireless communication has replaced wired networks. Every future appliance is coming with the wireless technology and being smart every day whether it is wireless sensor networks, automated highways and businesses, laptops and palmtops, and home
appliances have all emerged in the recent researches. Now, the wireless communications network follow a highly successful IEEE standard i.e. IEEE 802.11 family of standards.

1.1.4. Semantic Web

NetKAT is a mathematical networking programming language for establishing semantic foundation. NetKAT design incorporates in itself the techniques for filtering, transmitting, and modifying packets. It uses a mathematically structured Kleene Algebra with tests (KAT) that is proved in its capacity for sound and complete equational theory catering substantial semantics. It has been practically implemented with syntactic techniques for reachability, non-interference properties implementation that isolates programs, and correct algorithm compilation is thus proved. It is thus established that the equational theory applied in several diverse domains provides reasoning while reaching, isolating traffic, and correct compilation.

1.1.5. Similarities and Differences of Knowledge Network

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Sensor Networks</th>
<th>MANET</th>
<th>Wireless Networks</th>
<th>Semantic web</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deployment is not predetermined</strong></td>
<td>Yes. Intelligently deployed</td>
<td>Self-organized</td>
<td>Yes. Nodes are always on move</td>
<td>Servers are frequently being changed. IPs are not fixed</td>
</tr>
<tr>
<td><strong>Preprocessing data</strong></td>
<td>Yes. Partially processes data before sending to the desired destination</td>
<td>No preprocessing of data being transferred</td>
<td>No preprocessing</td>
<td>Extracting meaningful information</td>
</tr>
<tr>
<td><strong>Communication channels</strong></td>
<td>Broadcast</td>
<td>Point-to-point</td>
<td>COPE architecture, Relay network</td>
<td>NetKAT algorithm. Forwarding packets through node to node</td>
</tr>
</tbody>
</table>

Before the proposed theory there have been numerous measures proposed to counter these threats different security measures were implemented. One such architecture proposed [5] was Resilient Overlay Network (RON). It has provided
distributed Internet Applications with architecture that would detect and recover, within several seconds, from path outages and periods of degraded performance. RON is an overlay on the top of Internet routing stake of application layer that monitors the quality of functioning of Internet paths amongst them. This information is used to decide either to route the packet straight over the Internet or through the RON nodes using optimum application specific routing metrics.

With the emergence of wireless communications low cost sensor networks are developed. The sensor networks are composed of different sensor nodes which are densely placed as their positions are not predetermined. Sensor network has an onboard processor that processes and computes raw data and transmits only the required and partly processed data [6]. Such applications are deployed in healthcare, military, and homes. Characteristics like self-organization, fault tolerance, and rapid deployment make it useful to the military in tasks like command, control, surveillance, communication, targeting, etc. Such applications need to employ characteristics of wireless ad-hoc networking techniques.

There are so many of differences in between typical wireless ad-hoc and sensor networks, few are illustrated here:

- Number of nodes found in sensor networks is several magnitudes more than the nodes in ad-hoc networks.
- Sensor nodes are compactly deployed
- Sensor nodes have risk of failures
- As they are self-organized the topology structure of sensor networks transform very often
- Sensor networks use broadcasting for communication while ad hoc networks use p-2-p contact
- Sensor nodes have restrictive power, memory and computational capabilities
- Sensor nodes may not have worldwide ID because of that increase overhead

Researchers want to diminish these differences and are trying to cater these loop holes. In ad-hoc networks throughput is increased by using techniques like watchdog and path rater that identify misbehaving nodes and help routing protocols to avoid these nodes respectively [7]. Then there is Greedy Perimeter Stateless Routing protocol -
GPSR [8] for wireless datagram networks that use locations of router and a destination address of packet’s to make decisions for packet forwarding. It is a greedy approach to transmit information knowing only about the immediate neighbors of the routers. GPSR scales are better than ad-hoc and shortest path algorithms as the number of network destination increases keeping in per-router state in a local topology. The two dominant factors in scaling of routing protocol are:

- The number of routers in a domain
- The frequency with which a topology changes

Freenet [9] is another example of peer-to-peer network application. It permits retrieval, publication, and replication of data keeping anonymity of both authors and readers and taking measures to enhance the security of information storage. Freenet operates as a network of similar nodes that make a pool of storage space to store data and collaboratively rout requests for data to its most likely physical place. There is no centralized location or broadcast search to locate data.

It is designed keeping up to five goals:

- Protecting the anonymity of readers and authors
- Data is stored and routed dynamically at location independent distributed file system pooled by the nodes
- Resistance to denial of service attacks by third party intruders
- Deniability of keepers of information
- Decentralization of all network systems

Likewise, many such developments are in the pipeline that would be discussed in this study. This research would further highlight the various types of knowledge networks that are working with different levels of security precautions protecting the information being transmitted within these networks.

1.2 Research Problem and Research Questions

Knowledge networks as complex as it is, is harder to secure against intruding parties that can be malicious users, hackers, crackers, unauthorized access, etc., spreading destructive viruses and Trojans, making dissemination of important information a risky
endeavor. There have been considerable researches done to embed security in knowledge networks that are often termed as sensor networks, wireless networks, mobile ad-hoc networks (MANET), and semantic web. Still there are gaps left to be taken care of.

*Back Track Software:* There is software available that helps the intruder to hack the systems over the open networks whether it is an enterprise or a home based network. Back track (backtrack-linux.org) application has been distributed by Linux as a penetration testing tool that is used for security tests of LANs, Wi-Fi, Bluetooth and the list goes on. The same application is also the favorite of skilled hackers who get automated access to the open network.

*Massive Knowledge Network:* This network channel is used for different forms of communication that are required to be safe from eavesdropping. The high magnitude of nodes present in knowledge networks is itself a problem. It is hard to manage and monitor for external threats that are vast in number and type.

The research questions or concerns posed by the research at hand are:

1. Determining the optimum solution to knowledge networks security, integrating the security measures formed till now;
2. To come up with the best network design that enables the solution to work securely;
3. How can the sender and receiver host communication channel between them be isolated from network attacks?;
4. What mechanism is adopted to monitor such a huge network?
5. What would be the optimum security standards followed?; And
6. How these cyber security standards are implemented?

### 1.3 Thesis Motivation, Objectives, Methodology and Significance

Ideally, a WSN should be scalable and energy-efficient, smart and configurable, reliable and long-lived, highly responsive and incur small installation and maintenance cost. However, the goals are optimistic taking into view the available limited resources.
WSNs are distinguished from other types of ad hoc networks primarily due to the limited capabilities of a sensor node. On the other hand, ad hoc networks form a distinct category of networks. Nodes are wirelessly connected to each other and may be in constant random motion (e.g. Mobile Ad Hoc Networks (MANETs)). Now, in such a network, wireless medium and mobility poses another set of constraints. A MANET routing protocol’s prime function is to maintain the connectivity between a pair of nodes in the face of a frequently changing network topology.

1.3.1. Motivation

The remarkable increase in usage and complexity of new communication infrastructure systems associated to the Inter-net set demands security management to protect organizations’ critical information and assets from malicious users. Malicious attacks by hackers and intruders take advantage of weak points and flaws in deployed systems via number of sophisticated techniques that cannot be disallowed by traditional counter measures, such as access controls, firewalls and user authentication. Resultantly, timely responsive and automated detection systems are immediately required to find abnormal activities by monitoring the system events and network traffic. Network Intrusion Detection Systems and Network Intrusion Prevention Systems are technologies that look for the traffic and analyze system behavior to provide more attack protection.

A methodology to evaluate the generated alerts using a framework [10] for alert relationship has established the security which forms a global view of the security. The alerts that are missed could be recovered absolutely by using the contextual system techniques to detect multistage attack set-up. The alerts are combined, verified, demonstrated using vulnerability modeling, and interrelated to create multi-stage attacks [10].

There are various algorithms projected in the research [10] to give base to the operations of this framework which includes: alert aggregation, alert correlation and graph reduction. These techniques have been deployed in the tool entitled Multi-stage Attack Recognition System which consists of an integrated and collected mechanism.
With the increase in the number of events of detected attacks and rapid advancement in the communication networks, NIDSs became a key element of the security systems.

Swarm intelligence is a comparatively new subject that was basically described as a challenge to design problem-solving devices and algorithms or reflecting by the collective behavior of social insects and from other animal societies [11]. However, nowadays it refers more commonly to the research of the collective behavior of systems composed of a number of components which communicate using de-centralized controls and self-organization. But from an engineering aspect, SI lays emphasis to ‘bottom-up’ design of independent distributed systems which can show robust, scalable and adaptive behaviors. The SI includes other popular frameworks such as Ant Colony Optimization [12, 13], Bee Colony Optimization and Particle Swarm Optimization [11]. Most of the work in the area of SI design has been inspired by the adaptation of collective behaviors examine in natural phenomenon.

The similarity with the requirements and the characteristics of routing in MANETs, WSNs and a bee colony is remarkable. A honey bee colony collectively solves the routing problems to discover and setup optimized paths to foraging the places and then route back and forth the individuals from colony hive to the sources of food. This has been observed that the colony behavior is robust to losses of individuals, adaptive to environmental variations, and fully scalable and distributed. The bee colony could be considered as a ‘distributed adaptive system’ of small controls which uses energy efficient resources and little computing to discover the network, and that proficiently is cooperated by sharing routing information with their social group about the revealed paths and its quality.

These characteristics have already been applied productively for the development of two high-techs routing protocols: (1) BeeHive [14] which is particularly designed for ‘packet switched networks, and (2) BeeAdHoc [15] for ‘mobile ad-hoc networks’. BeeHive produces improved performance in fixed networks with a simple agent model as compared to other associated algorithms. On the other hand, BeeAdHoc carries alike or better performance and is more energy-efficient than other ad-hoc routing algorithms.
The success of BeeHive and BeeAdHoc and similarity of a bee colony food collection with a WSN and MANETs and NIDS, NIPS components are the two prime motivations of this work.

1.3.2. Research Objectives

The main objective of this study is to propose a framework to establish Secure Energy-efficient Cluster-based Data Routing in knowledge network. The specific objectives of the study are to:

- Find out the security and data routing issues in Wireless Sensor Networks (WSNs), Mobile Ad hoc Networks (MANETs) and Semantic Web Technology.
- Find out the existing data aggregation approaches and protocols in WSNs, MANETs and Semantic Web.
- Establish secure energy-efficient cluster-based data routing from source to base station so that data can be transmitted in a more secured manner with less energy consumption.
- Propose a secure energy-efficient protocol that performs secure data routing using hierarchical clustering and a cryptographic algorithm, with resource rich dynamic cluster head.

1.3.3. Research Methodology

To achieve the above objectives, the methodology taken in this thesis is followed to develop the design of an optimized security solution and discussion on Artificial Bee Colony clustering routing protocol for Knowledge Networks.

1. We first investigate the security threat, counter measures and architecture of each element (i.e WSN, MANET, Semantic Web) of the knowledge network. This helps in understanding the environment in which the newly proposed optimized security solution and routing protocol is expected to operate. We call it the requirement engineering phase.
2. In the second stage of dissertation, we study the principles of Cluster Based Network Design, Routing Protocol and Security and comparison of different clustering protocols in-order to take inspiration from the relevant concepts for our new security and routing protocol for knowledge network.

3. We then map the principles of a bee colony to a real world problem - routing in WSN and MANETs. Steps 2 and 3 are a part of the natural engineering phase.

4. Having done the above steps, we enter the security optimization and protocol engineering phase. We transform the natural concepts into the algorithmic details of deployment and clustering the network design.

5. We analyze the performance of the proposed solution, through empirical and formal techniques to gain insights into the parameters governing its behavior in WSN and MANET and Semantic Web.

6. In the evaluation phase, we extensively evaluate the performance of the final network design of our optimized security solution and routing protocol through simulation studies to show that it meets the design objectives.

1.3.3.1 Research Assumptions

Assumption 1: Clusters are administered by multiple base stations. Parameters are set and they are configured as a bubble effect. From one cluster other clusters get created adjacent to it.

Assumption 2: For power centric routers, the energy level and specifications for system compatibility is similar.

Assumption 3: Base stations have already been set up and running throughout the network.

Assumption 4: All nodes are covered by the clusters.

1.3.3.2 Research Techniques

Network design is proposed that would be simulated using:

- Wireless network simulator in MATLAB: would define the parameters set as threshold for clustering network


- **Swarm Intelligence Methodology**: For the deployment of heterogeneous nodes in a network.

1.3.3.3 **Tools of Research**

Tools of this research would be:

1. Evaluating and integrating proposed solutions by previous research papers and thesis,
2. Simulations, and
3. Validating through mapping on to the already established standards.

1.3.3.4 **Data collection and analysis**

Data would be collected for analysis through previous research and web pages of organizations that are famous for developing network security standards. The previous research would be evaluated and integrated to form the optimal solution that would be mapped on to the well-defined standards.

1.3.4. **Significance of Study**

Finding an optimum solution to implementing security in knowledge networks is very crucial for future development. Any future development towards knowledge networks would require a strong foundation that could not be outperformed by security threats. It would be thoroughly tested against current penetration techniques employed by intruders. Security within knowledge networks would add value to knowledge management and business intelligence techniques for further improvements and enhancement. Consumers of knowledge networks need to be assured of secure transmission and storing of their highly valuable data.

1.4 **Major Contributions of This Study**

The primary contributions of this thesis are summarized as follows:
1.4.1. Cluster Based Security Architecture - An Optimized Security Solution

Secure communication is an essential element in computer networking and authentication is one of the most renowned prerequisite. However, general authentication mechanisms are not feasible to be deployed in ad-hoc networks due to centralized certification authority (public key infrastructures) being difficult to be implemented there. Researchers proposed a distributed certification facility based on security concept in which a network is segregated into clusters having a special node CH. These CHs hold the network key of nodes used for certification and perform administrative tasks. New nodes can only become a full member of the network once if it has signed certificate received on successful ‘authenticity’. In the design every node has a session key at the time of deployments. Primarily the node encrypts the captured data by applying the Blowfish Algorithm, which makes the data transmission more secured. It sends encrypted data to the CH and from CH to the gateway and lastly to the base station.

1.4.2. Intrusion Detection Based Security Scheme for Cluster-Based Wireless Networks

An intrusion is a set of activities that can control an authoritative access or modifications of the wireless network system. These detection methods can identify malicious user on the bases of some anomalies; normally, the adjacent of a malicious node are the first points observed or influenced by those abnormal behaviors. Intrusion detection system-IDS deployed to supervise and monitor the systems and computer networks, detects the probable intruders in the network, and notifying users once if intruders are detected, if essential and possible then reconfiguring the network.

1.4.3. Verification of Proposed Network Design

A simulation for producing real-time network traffic (using CISCO Packet Tracer) is adopted, to generate test data in the presence of security intrusion detection systems. The security protocol being used is IPSec on each router in the networks. A
network design is thus proposed to hold power centric intelligent nodes (routers) within clustered knowledge networks. In real network system, the clusters would be formed such that each town or city would have one intelligent cluster.

**Logical Verification of Design:** Proposed network design involves the conditional implementation of statements that specify the attribute or determine its value. The basic feature includes Parent-child connection, subnet implementation, services stored on device and access rights list.

**Visually Verified Features:** The visual shell of the modeling process to verify the design features includes the graphical representation of design model and the attribute values. The design model features that have been verified visually are Graphic connections, IP addressing, and Network model saving process.

### 1.4.4. Simulation Using MATLAB

This research focuses on forming a clustered network combining towns and cities and thus, catering to a massive network of nodes. And, the routers are the nodes that would be power centric and behave intelligently. Thus, IPSec is configured on routers to implement gateway-to-gateway security as node-to-node security becomes too annoying. A simulated network design is thus proposed to hold power centric intelligent nodes (routers) within clustered knowledge networks. In real network system, the clusters would be formed such that each town or city would have one intelligent cluster.

- Simulation-based investigations are conducted to determine the optimum number of clusters, energy dissipation and time for message transfer of proposed network. Other two simulations for wireless sensor networks are done using MATLAB® to relate ‘hop count’ with ‘network size’ estimate. This is the way each cluster nodes would relate itself to the other cluster in terms of ‘hop count’ and within a certain ‘response time’. Then we validated the network against success rate that is achieved keeping a considerable network size in a cluster:
  - After exhaustive testing experimental results of multiple simulations the conclusion has been reached to create clusters of up to 10 to 20 nodes.
  - Simulation-based investigations are carried out to test and validate the success rate
of response and hop time for various networks within clusters of 10 to 20 nodes has been found enough for getting the desired result.

Corresponding with intelligent routers in place for getting the desired success in monitoring the activities of nodes within a cluster. The average success rate after evaluating the response time and hop count of a considerable network size viewed that nodes between 10 to 20 give the optimal result or otherwise the cluster would grow too large to manage and keep track of.

1.4.5. A bee-Inspired Routing Protocol for Knowledge Networks to Attain the Clustering Optimization- iRoutCluster.

This research caters to attending the knowledge networks inclusive of; (i) Sensor Networks, (ii) Mobile Ad-hoc Networks (MANETs), (iii) Wireless Networks, (iv) Semantic Web. The Intelligent Routing Cluster protocol designed to be implemented in this research is such that it would be covering all three domains of knowledge networks discussed in this research. Mainly it is influenced by Bee colony optimization (BCO) based routing protocols that is itself a Swarm Intelligence (SI) based technique for networking. As BCO itself keeps a nature of bees in real life scenario, so it could be implemented in all types of knowledge networks. There have already been many protocols been implemented for MANETs but it is specifically good to utilize for wireless, sensor and semantic web networks.

The enhancement is made in adapting to a clustered network. Where a center point (cluster head) has already been established that in turn is responsive to the base station executing it. These clusters would be having all three domains of network connections and protocol is intelligent enough to respond towards each one. The proposed protocol would be made responsible for all communications taking place within a respective cluster and it would intelligently keep track of end to end communication. It would be supported by IPSec protocol being embedded with it into the network that would be responsible for gateway-to-node communication. And, IPSec security protocol is also applicable at all levels of knowledge networks.
Utilizing the BCO visual monitoring feature, the designed protocol enables the cluster head to take snapshots of the operating system within the nodes that are included in its range. These snapshots are simultaneously being visualized by base stations at the very detail level. Thus, the network design captures every minute detail of malfunctioning in the knowledge networks that is composed of sensor networks, mobile ad-hoc networks, wireless networks, and semantic web. Moreover, it is a mix of protocols leveraged from these networks. It is inclusive of COPE and NetKAT algorithms to enhance its managing capabilities among broad spectrums of knowledge networks.

1.5 Organization of Thesis

The rest of this thesis is arranged into eight chapters. Figure 1.2 shows the route map which presents the interconnections among these chapters in structuring this thesis as follows:

Chapter 1: Thesis Introduction: Introduces the motivation, objectives, methodology, and the primary and significant contributions of this thesis

Chapter 2: Problem Evaluation, Security Optimization and Routing Methodology: The purpose of this survey is to understand the existing intelligent routing techniques, Security frameworks and technologies which help to optimize the security solution. We have described our design and methodologies concisely followed by the problem of Intrusive Behavior, Detection and Prevention for knowledge Networks. Finally, propose security architecture for the purpose of establishing Cluster based security architecture and secure routing that is discussed in chapter 6 and 7.

Chapter 3: Sensor Networks Security and Routing Techniques: Wireless and Sensor network are the core element of Knowledge networks and purpose of this section is to study the communication structure of sensor node, attacks on WSN and their mitigation, routing challenge and design issue of WSN, and clustering routing protocols.

Chapter 4: Mobile Ad-Hoc Network Security and Routing Protocol: In this section, deliberately discussed are mobile ad-hoc network architecture, its security concerns i.e vulnerabilities, attacks, and detection and prevention schemes. Furthermore, Secure
Routing Protocols and MANETs clustering model are also studied with its feature and implications.

**Chapter 5: Semantic Web Model and Security Protocol:** Languages for Representing Knowledge on the Web, semantic web technology, key security issues, protocol and security standards, some explanatory applications and supporting instances are researched and discussed.

**Chapter 6: Cluster Based Security Architecture and Communication Protocol for Ad-hoc Wireless Networks:** We continue our analysis and evaluation of optimal network security solution, routing and associated algorithms based on Hieratical clustering technique and energy efficient routing in WSN and MAENTs. In addition to this, we have also defined the intrusion detection based security structure and respective solution in a systematic layout.

**Chapter 7: BCO – A Simple, Scalable and Energy Efficient Routing Protocol for WSNs & MANETs:** BCO inspired by the bees’ behavior in the nature, the main theme of the BCO is to define multi-agent system – a colony of artificial bees, which is competent enough to solve the complex combinational optimization problems successfully. In this chapter we study the theoretical, mathematical and algorithmic representation of routing and clustering model with some analysis and comparative result, are also discussed.

**Chapter 8: Design Review, Simulation and Performance Analysis of Clustered Network Design:** Formal modeling process and optimum solutions have been described in the last two chapters which unleashed several important insights. These insights are extremely beneficial and can be instrumental in the design of any ad-hoc and sensor network routing algorithm. To this end, we utilize these insights and review the design of network protocol to rectify its shortcomings. After that, the results of simulation to illustrate the performance of proposed protocol in static and mobile networks are presented and compared with different adaptive and swarm intelligent algorithms.

**Chapter 9: Conclusions and Future Works:** This chapter outlines the key findings of this study and makes suggestions for potential further work.
Figure 1.2  Thesis route map
References


Chapter 2
Problem Evaluation, Security Optimization and Routing Methodology

This chapter presents a comprehensive introduction on problem evaluation and respective proposed security framework for knowledge networks in the light of literature survey. The purpose of this survey is to understand the design principles of existing security mechanisms, simulation model and routing protocols and study their strengths and weaknesses to discuss individually (WSNs, MANETs and Semantic Web) in next three chapters 3, 4 and 5, respectively.

To this end, we first describe a set of challenging issues in the design of a Optimized Network Security Solutions. We then introduce a comprehensive taxonomy that helps us in identifying the key features of existing security measures and models in the support of proposed framework. After discussing the basics of security inductions, we propose a framework that represents a unified view of the different components of system model and SI-based routing algorithms. It not only facilitates the mutual comparison of different SI-implementations but can also guide the design of future SI routing protocols.

In section 2.1 we have presented the comprehensive Literature review to enlighten the problem domain on the bases of problem statement and research methodology illustrated in chapter 1. This section discusses the relating recent work on intelligent routing, group authentication, data Integrity and confidentiality, intrusive behavior: detection and prevention, security framework, centralized host security scanning architecture, cluster networks, IPSec network security protocol and standard based on this study. We have proposed an optimized network security solution for knowledge network (ONSS-KN) in section 2.2.

In section 2.3 based on literature review, it is learnt that there have been many mechanism already proposed but we have selected the most suitable features of these and incorporate in to clustered network design model that is integrated with security mechanism defined in section 2.2. The system model is designed in such a way that power centric node would hold and control the security implementation and execution of
the entire network. The cluster based network design and security architecture is presented in chapter 6 of this research thesis.

In section 2.4, we have proposed a distributed, scalable and energy-efficient bee inspired routing protocol for wireless networks – iRoutCluster. Like other SI algorithms, iRoutCluster is designed with the so called “bottom-up approach” in which the behavior of individual nodes is defined keeping in view the desired network level behavior. This routing algorithm applied on cluster based network model is described in section 2.3 to facilitate the ONSS-KN. In detail BCO – a simple, scalable and energy efficient routing protocol for wireless sensor networks and mobile ad hoc network is presented in chapter 7.

At the end in section 2.5 the most recent wireless communication channel, IEEE standards and COPE architecture are featured.

2.1 Literature Review and Evaluating Problem Domain

Knowledge networks are as complex as it is harder to secure against intruding parties such as malicious users, hackers, crackers, unauthorized access, etc., spreading destructive viruses and Trojans, making dissemination of important information a risky endeavor. There have been considerable researches done to embed security in knowledge networks that are often termed as sensor networks, wireless networks, mobile ad-hoc networks (MANET), semantic web, etc. Still there are gaps left to be taken care of.

Tasks like authentication, data integrity, intrusion detection and prevention and firewall systems are all concerns of security in cyber knowledge world. Authenticating the parties on both ends in knowledge networks, then the data being transmitted needs to safely reach on to the other end without being hijacked in the middle, intrusive behavior within the networks has to be discouraged through early detection and prevention mechanisms and implementing firewalls are all tasks to be considered in this research for implementing optimal security in cyber networks. There are different types of routing attacks, man-in-the-middle, and lack of privacy concerns. As the networks have advanced considerably including cloud computing, the mobile internet, voice over IP (VoIP),
intelligent systems, smart phones as well as home environments give way to countless attacks from malicious users [1] find the need for security measures as strong as the attackers and intruders to these systems thus delivering the Internet World with the novel algorithms, frameworks, and theories to improve the ever increasing threat to Internet security and leverages. The work recognizes the foundation of the current security measures and architectures to be topology parameters, operations under isolation, reactive reconfiguring mechanisms, and human advisory. Intermittently connected networks thus also have problem of not having a connected path between desired nodes where communication is required.

2.1.1. Intelligent Routing

Knowledge networks are inclusive of Mobile Ad hoc Network. This is inclusive of self-organized nodes using wireless setup for connecting two ends for communication that are within each other’s transmission radius in a multi-hop infrastructure. There has been conducted progressive research in the past that proposed various routing algorithms for MANET, but the recent study shows that the emerging trend in routing research is towards Swarm Intelligence (SI) [2].

Swarm Intelligence refers to an artificial intelligence technique [2] implemented on to a complex system that is meant to be decentralized and self-organized. It could result in observing intelligent and complex behavior emerging from simple unsupervised interaction of all swarm members part of a system or any network. There are various routing techniques based on Swarm Intelligence that are (i) Ant Colony Optimization (ACO) algorithm based routing protocols, and (ii) Bee Colony Optimization (BCO) based routing protocols. These algorithms are inspired by ants and bees behavior in their colonies. Such nature inspired routing protocols would be capable of removing at least one of several problems like; survivability, scalability, adaptability, maintainability, battery life, and further more.
2.1.2. Group Authentication

Now, there are measures taken to enable secure group communications within the network. Commonly, there are two types of authentications; knowledge based (e.g. passwords) and key based authentication (e.g. public/private key encryption). Knowledge based authentication has some flaws due to which passwords could be hacked. Key based authentication concerns with the computational time involved using large integers. So, a new type of authentication is being proposed called group authentication [3]. All users in a group are authenticated at once. There is a group manager who manages this communication using group authentication protocol. Manager distributes a special token to each user in a group while registering that determines if the user belongs to the same group. These tokens can be reused without compromising the security of tokens. The identity of each user remains protected and hidden from each other in a group.

2.1.3. Data Integrity and Confidentiality

In networked society governmental decision making and public services are enhanced through increased use of ICT [4]. But with increased network communication information became invasive and complicated, interdependencies started to occur that gave way to various types of vulnerabilities. These vulnerabilities formed the basis for serious failures in critical infrastructures and introduced highly intrusive cyber-attacks. The ICT enabled countries are therefore giving high priority to counteract these threats. Botnets provide platform for serious threats including distributed denial of service, information stolen and spamming. There have been observed two complex scenarios; there are stealthy Botnet attacks that are hard to be identified, and legitimate P2P applications (e.g. skype, and bittorrent, etc.) are running on bot-infected hosts. Lastly, traffic analysis framework[5] is provided boosting scalability of Botnet detectors being used. This framework can identify number of hosts that are Botnet infected. But, the counter attack that would diminish these Botnet infected attacks has not yet been realized.

Network Controlled Systems (NCS) [6] became vulnerable to attack in the presence of Internet and wireless communications delivering pervasive and non-proprietary information. Traditionally, when developing security mechanisms, the
interdependencies between the physical systems and cyber connections in IT get ignored. Security in power systems should also be focused when creating security and reliability mechanisms for NCS.

2.1.4. Intrusive Behavior, Detection and Prevention

It was recognized that automated intrusion detection and prevention [7] should be in place against intrusion attacks. Network traffic has to be continuously monitored through diagnostic systems for abnormal activity. Security scanning [8] should be performed at all sub-systems whether they are private enterprises or public cloud. Vulnerabilities found on enterprise computer systems can also be exploited by intruders reaching through open networks. When networking, effective security architectures should assess the vulnerabilities found on the host. Cyber-crime is present to exploit network traffic to collect useful and private information. Intrusion detection tools are lacking to give full protection against these malicious attacks. Thus, developing effective tools is trivial job but to do this a test bed is required to represent a network and intruders attacks. But, monitoring or analyzing the network is not just sufficient unless the passive behavior being recognized is not blocked or discouraged. In this research, the importance is given to blockage and discouragement of intrusive occurrences.

2.1.5. Botnet Boost-up Detection System

Botnets provide platform for serious threats including distributed denial of service, information stolen and spamming. As the networks are growing and high scalability is required by Botnet detection systems, the study [9] presents three contributions towards enhancing the functionality already provided by current Botnet detection systems to detect hard to observe Botnet attacks through sampling Botnet generated network packets. Firstly, drive-by download attacks are detected that were missed before by existing detection systems; secondly, a new system is built adopting peer-to-peer (P2P) command and control (C&C) structures to detect serious Botnets that are disrupting. There are two complex scenarios; there are stealthy Botnet attacks that are hard to be identified, and legitimate P2P applications (e.g. skype, and bittorrent, etc.) are
running on bot-infected hosts. Lastly, traffic analysis framework is provided boosting scalability of Botnet detectors being used. This framework can identify number of hosts that are Botnet infected. Thus, sending the network traffic is associated with these hosts to boost-up existing detection systems for fine grained analysis. It includes the algorithms for novel Botnet-aware and adaptive packet sampling with scalable flow-correlation technique.

2.1.6. Creating a Robust Network Society in Sweden:

Network community faced lots of digital threats [10], a strategy derived for collaboration and information sharing between military and civilians for highly robust society. This societal structure is based on abundant access to Information and Communication Technologies (ICT). In networked society governmental decision making and public services are enhanced through increased use of ICT. But with increased network communication, information became invasive and complicated, interdependencies started to occur that gave way to various types of vulnerabilities. These vulnerabilities formed the basis for serious failures in critical infrastructures and introduced highly intrusive cyber-attacks. To counteract on these loopholes in the highly technological dependent society then became the high priority task of government bodies and agencies in the ICT enabled countries. The Swedish government took this task seriously in 2010 and decided to come up with a strategy for the safety of their public services and societal infrastructure[10]. This is an extremely demanding task with huge complexity to counter attack an intruder whose origin and even the way of attack was hard to be identified in the massive network. To come up with a well formed strategy a controlled analysis is required so that risks could be identified. US came a long distance to set up the Cyber Command that creates defense against cyber threats.

2.1.7. Security Framework:

Network Controlled Systems (NCS) became vulnerable to attack in the presence of Internet and wireless communications delivering pervasive and non-proprietary
information. Traditionally, when developing security mechanisms, the interdependencies between the physical systems and cyber connections in IT get ignored. Security in power systems should also be focused when creating security and reliability mechanisms for NCS [11]. Thus, the State Estimator (SE) in power networks is analyzed in presence of malicious sensor data corruption attacks. The Bad Data Detector (BDD) schemes are analyzed for getting bypassed. Using the results the security framework is proposed highlighting minimum-effort attack policy as a constrained optimization problem. The optimal attack cost is evaluated as a security metric to provide the basis for developing a strong protective scheme. Simulations and experiments demonstrate the features of the proposed framework. Secondly, the Optimal Power Flow (OPF) algorithm is analyzed during stealthy data corruption and consequences that occur to power network operation. Set of attacks are characterized for which OPF is made available and analytical expression is proposed as an optimal solution to simplified OPF problem.

2.1.8. Multistage Attack Recognition System (MARS):

User authentication firewalls, and access control measures were insufficient to prevent exploitation of flaws and weaknesses found in the architectural security in modern communications and networks [12]. It was recognized that automated intrusion detection and prevention should be in place against intrusion attacks and network traffic has to be continuously monitored through diagnostic systems for abnormal activity. Such systems are: Network Intrusion Detection Systems and Network Intrusion Prevention Systems. Current protection mechanisms lack scalability and do not keep up with the increasing network speed, new services and protocols. With installation of multi-giga networks the resources get exhausted in the presence of various attacks, thus making NIDS susceptible. Focus was put on two areas lacking in NIDS: (i) due to NIDS performance limitations packets were lost and there were missing alerts, (ii) event observation had become undoubtedly difficult for security analyst due to overwhelming alerts generated. Therefore, a methodological framework was proposed to correlate security alerts and provide security analyst with a global perspective. This framework implemented number of algorithms collection in a tool called Multistage Attack
Recognition System that included a combination of integrated modules: alert correlation, graph reduction, and alert aggregation.

2.1.9. Centralized Host-Based Security Scanning Architecture

Security scanning [13] should be performed at all sub-systems whether they are private enterprises or public cloud. Vulnerabilities found on enterprise computer systems can also be exploited by intruders reaching through open networks. When networking effective security architectures should assess the vulnerabilities found on the host. Scanning the host needs the scanner to have administrative privileges that create a legitimate concern and threat from the scanner itself if it has some vulnerabilities of its own. Installing the security analysis is not easy as every host would have to carry it on an enterprise level. The architecture is presented having minimal coding so that its correctness can be verified through thorough inspection. This architecture allows third party security tools to incorporate with it to leverage security knowledge more conveniently.

2.1.10. Clustered Networks

Today large data centers consisting of thousands of expensive units such as routers and switches catering to varied aggregate bandwidth requirements may support fifty percent of the total bandwidth at the frame of the network in a topology. And, the resulting cost would also be maximum depending on the topology. Varied bandwidth among the networks complicates the arrangement and design of networks and overall system performance.

Thus, study [14] is done to develop clusters for getting full aggregate bandwidth out of network topology in large data centers having tens of thousands of nodes. It is argued that proper installation and design of network switches/routers would give high performance with minimal cost. The proposed approach needs no modification on the end host network interface and is entirely compatible with TCP, IP, and Ethernet.
The routing algorithm [14] used in the proposed architecture is built with first two levels of switches acting as filters for traffic distribution, the upper and lower end switches in a given shell implements prefixes to terminate subnets. So if a host sends a packet to the destination within same pod but having different subnet the terminating prefix would be directing to the destination subnet’s switch. The core switches have the terminating first-level prefixes for all network IDs pointed to the right shell containing that specific network. Thus, a single path is recognized between the core switch and the destination shell. To keep it simple it is assumed that a central element or unit of the network has complete knowledge of the interconnected cluster devices or node.

2.1.11. Simulation Model

Cyber-crime is present to exploit network traffic to collect useful and private information. Intrusion detection tools are lacking to give full protection against these malicious attacks. Thus, developing effective tools is trivial job but to do this a test bed is required to represent a network and intruders attacks. A simulation [15] was designed that could generate test data that models a reality based network activity and cyber-attacks in presence of security intrusion detection systems. The simulation model developed was object-oriented in nature using Java providing a customized interface representing real-time network structure and attack scenarios. This simulation model provided user the capability to detect intrusion in a virtual network. User could load a network topology with vulnerabilities that would be attacked in a given scenario and thus sensor alerts would be generated.

2.1.12. Turing Assessor

Creating security metrics is a challenging task as network communicates within an external environment which makes it very vulnerable to outside threats that can be passive like; eavesdropping, etc. or active attacks like; inserting, modifying, or deleting packets over the communication channel. A network is divided into subsystems to evaluate its associated functionalities. To evaluate a security for functionality within
every individual module, information learning is enabled through oracle queries. Each query in oracle can be effective and ineffective based on the attack. As the type of environment is very integral for creating security parameter, an environment-based security model is provided that is robust and applicable to any environment. A security measure called Turing Assessor [16] is realized for a network decomposed into subdomains having related functionalities. Turing Assessor caters to synchronous as well as asynchronous communication channels. A platform for Synchronous Turing Assessor (STA) is built and then it is reduced to Asynchronous Turing Assessor (ATA) by removing synchronous communication channel restriction.

2.1.13. Supervisory Control and Data Acquisition Systems (SCADA) – A test bed:

Sridharan [17] discussed the vulnerabilities found in Smart Grid concept in 2012 during his research. These deficiencies gave way to intrusion in cyber space. Power system devices that are for managing and supplying electrical energy can be threatening to information sharing on the mobile networks. Till now there are no metrics formed to measure the threats and attacks on cyber networks. Sridharan proposed the test to measure the multiple-threat methods for monitoring cyber security on a multi-laboratory test bed that would assist in developing SCADA – a test bed devised at Georgia University, USA.

2.1.14. IPSec Network Protocol:

IPSec [18] is a standard security protocol that ‘encapsulates’ an encrypted network layer packet within a typical network by keeping its decoding transparent to mediator device or nodes which must have to process the packet headers for routing, etc. authentication, encryption, encapsulation is done on outgoing packets. Therefore, incoming data traffic is decrypted, de-capsulated and verified on recipient side. Key management in this system is simpler.
The encapsulation technique or design adopted for authentication and confidentiality is not such as hard to implement and IPSec is standardized by IETF and implemented by commercial vendors.

The problem with standardized network layer security protocols is that they do not address packet handling management in way of a host running the encapsulation protocol. Security protocol protects the packets from external altering and eavesdropping but does not enforce security policies for hosts on transferring particular kinds of traffic. If these policies are implemented within firewalls and VPNs it can be very complex.

2.1.15. IPSec Policy Architecture

In any protocol architecture, here specifically IPSec is discussed [18] – whenever packets leaving a network security endpoint and when reaching a destination, policies must be enforced on routers, endpoints, gateways, firewalls, etc.

The packet is traditionally not secured and only packet filtering can be done for acceptance. If a packet is encapsulated on security protocol, (i) the key substances should be controlled in a security association (SA) that is used to de-capsulate the packet, and (ii) the packet after de-capsulation is accepted through packet filtering of another type and there is a chance that packet may be rejected if it is illegal.

2.2 Optimized Network Security Solution for Knowledge Networks – (ONSS-KN)

Research is carried out by intensive literature review and analyzing the data and experiments previously been implemented. The security concerns in knowledge networks would have to be tackled by adopting various means that would comprise of cryptographic algorithms and secure communication protocols. These mechanisms would enforce security parameters [19]:

1. Confidentiality: data transmitted between two endpoints remains private
2. Integrity: the data does not get tampered during transmission
3. Availability: the endpoints are accessible whenever required
4. Authenticity: data sender has to authenticate himself and data receiver should not be spoofed

By this study we bring to close that secure routing scheme based on session keys by implementing symmetric cryptography using Blowfish Algorithm (symmetric key algorithm) with session keys is the Optimized Network Security Solution for Knowledge Networks (ONSS-KN). This model conveys optimal security as compared to previous secure routing algorithm adopted for cluster wireless networks. ONSS-KN add security to cluster based communication protocols - a distributed, scalable and energy-efficient bee-inspired routing protocol for wireless sensor networks and mobile ad-hoc networks is illustrated in chapter 7.

Our proposed cluster based security architecture ONSS-KN consists of three modules. Each module has specific task which has detailed discussion in chapter 6, 7 and 8;

*Figure 2.1 Optimized Security Solution Process Model and Architecture.*

**Module I:** Deploying process is an efficient energy utilization base which extends the life time of network and provides optimal node distribution using modified artificial bee colony algorithm iRoutCluster -I (a modified version of ABC, chapter 8). Afterward, it places the node into cluster network with symmetric key.
**Module II:** Optimized security model add security to cluster based communication protocols (Bee inspired, swam intelligence base routing algorithm will be studied in detail in chapter 7) in homogeneous wireless and mobile ad-hoc networks with resource-constrained nodes. Proposed network clustering protocol (**iRoutCluster – II**) is based on a centralized control algorithm that is implemented at the base station. The base station is a node with unlimited energy supply.

**Module III:** There are two major tasks under this module: the authentication using distributed certification which is controlling access to services and resources by authorized certificates and integrity and confidentiality using secure data routing. The techniques adopted are discussed in chapter 6 and 8 and benefit of this system is that, it improves communication security and requires very less energy as compared to other cryptography algorithms.

IETF has greatly contributed in standardizing different protocols (i.e. TCP/IP, HTTPS, UDP, IPSec etc) for providing open network security. With protocol compression where overhead is increased, the size of transmitted data has increased. Security duplication is avoided by providing cross-layer interaction between protocols. At network layer, Internet Protocol Security (IPSec) is implemented to keep the data exchange secure at different configuration parameters; (i) Gateway-to-gateway, (ii) Gateway-to-host, and (iii) host-to-host communication [19]. IPSec is known for providing data-origin authentication, confidentiality, integrity and prevention against replay attacks. IPSec uses two kinds of security protocols such that authentication header and Encapsulated Security Payload used in combination with Internet key exchange (IKE).

This research focuses on forming a clustered network combining towns and cities and thus, catering to a massive network of nodes. And, the routers are the nodes that would be power centric and behave intelligently. Thus, IPSec is configured on routers to implement gateway-to-gateway security as node-to-node security becomes too annoying.

### 2.2.1. Monitoring for Intrusive Activities

Boost-up Botnet detection system [5] is there contributing three novel strategies for tackling the cyber-attacks like; drive-by download attacks that were missed before by
existing detectors, then there is P2P Command & Control (C&C) structures adopted to identify attacks that could disrupt the network. Finally, a framework is provided for traffic analysis to boost the effectiveness of Botnet detection system. Algorithms for adaptive packet sampling and novel Botnet aware system are there with scalable flow-correlation technique. Network Intrusion Detection Systems and Network Intrusion Prevention Systems are inbuilt in central network connecting devices such as routers (nodes) integrating them with the tool called Multistage Attack Recognition System [7] that includes integrated modules: alert correlation, graph reduction, and alert aggregation.

**(a) Centralized Botnet**

**(b) P2P Botnet**

*Figure 2.2 - Botnet Detection System [5]*
Network Intrusion Detection Systems and Network Intrusion Prevention Systems would be inbuilt in central network connecting devices such as routers (nodes) integrating them with the tool called Multistage Attack Recognition System [20] that included compiled integrated modules: alert correlation, graph reduction, and alert aggregation. Then there is centralized host-based security scanning architecture [21] that would also be made part of the routers within a cluster. Creating clusters would make the knowledge networks easy to manage and monitor for faults and intrusive malicious activities that endangers this massive network. The network design is thus proposed to hold power centric nodes (routers) within clustered knowledge networks.

2.2.2. IPSec Security Protocol

IPSec [20] is a standard security protocol that ‘encapsulates’ a decoded network layer packet within a network packet by keeping the decoding transparent to middle network nodes or device that must have to process the packet headers for routing, etc. authentication, encryption, encapsulation is done on outgoing packets being sent to network. And, thus incoming packets are decrypted, de-capsulated and verified upon receipt. Key management in this system is simpler.

The design of encoding techniques for simple authentication and confidentiality is not that hard to implement and IPSec is standardized by IETF and implemented by commercial vendors. IPsec is an end-to-end security scheme operating in the Internet Layer of the Internet Protocol Suite, whereas some other Internet security systems in general use, such as Secure Shell (SSH) and Transport Layer Security (TLS), operate in the upper layers at Application layer. Hence, only IPsec protects any application traffic over an IP network. Applications can be automatically secured by IPsec at the IP layer.

2.3 System Model of Optimized Security Solution

In this section we are presenting the system model of cluster of nodes, clustered networks mechanism, and features of power centric node.
2.3.1. Cluster of Nodes

This research study proposes a security mechanism forming a network design in a clustered manner. Based on literature review, it is learnt that there have been many mechanisms already proposed. This study gathers the most appropriate features of these and incorporates them into a clustered network that enables monitoring in a dense network of nodes that would be discussed in detail in chapter 6. Then:

- Best and most appropriate security mechanisms are filtered through literature review and refined to integrate within the clustered network.
- The knowledge networks are broken into dense clusters given some threshold parameters.
- The routers in these clusters are made intelligent being power centric with the ability to diagnose sender and receiver hosts for any type of malicious activities.
- It is being assumed that the base stations are already established to execute and monitor these clusters in a group.
- Within each cluster there would be a cluster head that would be regulating the whole network traffic becoming part of the particular cluster.

2.3.2. Clustered Networks

Today large data centers consisting hundred and thousands of expensive entities such as routers and switches catering to varied aggregate bandwidth requirements may support fifty percent of the total bandwidth and the resulting cost would also be maximum depending on the topology. Varied bandwidth among the networks complicates the arrangement and design of networks and overall system performance [22].

Thus, study [22] was done to develop clusters for getting full aggregate bandwidth out of network topology in large data centers having tens of thousands of nodes. It is argued that proper installation and design of network switches/routers would give high performance with minimal cost. The proposed approach needs no modification on the end host network interface and is fully compatible with IP, Ethernet, and TCP.
The routing algorithm [20] used is taken from the previously proposed architecture built with first two levels of switches acting as filters for traffic circulator, the upper and lower end switches in a given shell implements prefixes to terminate subnets of that shell. So if a host sends a packet to the destination within same shell but having different subnet the terminating prefix would be indicating to the destination subnet’s switch. The core switches have to terminate the 1st level prefixes for all network IDs pointed to the right shell containing that specific network. Thus, a single path is founded between the core switch and the destination shell. To keep it simple it is assumed that a central unit has complete knowledge of the interconnected cluster network.

In this research, a new design is proposed (figure 2.3 & 2.4) and implemented through simulating in a large network split into multiple clusters of relatively same size and area. Each cluster would be regulated and its traffic tracked by the cluster head appointed at the center of that cluster. The base station would be responsible for managing group of clusters and would mainly be in contact with the cluster head within each cluster coming in its range. This cluster head would be recognizing every node that sleeps or incurs in the specific cluster.
Figure 2.3 Clustered Network Design with Cluster Head
2.3.3. Power Centric Nodes

Systems running on a power supply are also known to be threatening to knowledge networks as they are known to be sharing informative data along the network channel. As this research already clarifies Sridharan’s proposed test bed – SCADA, that monitors the cyber security in a multi-laboratory configured in Georgia. Routers in a clustered network termed as central nodes in this research study, hold such monitoring system and act as power centric for the traffic coming from different gateways. Security in power systems themselves being maintains as well when creating a robust network [6]. As there were mechanisms formed for security of NCS such optimal solution is designed to protect routers and gateways too, sensing the malicious data corruption attacks in information channels connecting routers and within the routers.
Then there is centralized host-based security scanning architecture [8] that is also made part of the routers within a cluster.

**Figure 2.5** Power Centric Node of Cluster Network Design [8]
2.4 Clustering and Secure Routing Algorithm for Knowledge Network – \textit{iRoutCluster}

This research caters to attending the knowledge networks inclusive of; (i) Sensor Networks, (ii) Mobile Ad-hoc Networks (MANETs), (iii) Wireless Networks, and (iv) Semantic Web.

The Intelligent Routing Cluster protocol designed to be implemented in this research is such that it would be covering all four domains of knowledge networks discussed in this research. Mainly it is influenced by Bee colony optimization (BCO) based routing protocols that itself is a Swarm Intelligence (SI) based technique for networking [21]. As BCO itself keeps a nature of bees in real life scenario so it could be implemented in all types of knowledge networks. There are already many protocols, being implemented for MANETs but it is specifically good to utilize for wireless, sensor and semantic web networks.

The enhancement is made in adapting to a clustered network. Where a center point (cluster head) has already been established that in turn is responsive to the base station executing it. These clusters would be having all four domains of network connections and suggested protocol is intelligent enough to respond towards each one. It would be made responsible for all communications taking place within a respective cluster and it would intelligently keep track of end to end communication. It would be supported by IPSec protocol [20] being embedded with it into the network that would be responsible for gateway-to-node communication. And, IPSec security protocol is also applicable at all levels of knowledge networks.

Utilizing the BCO visual monitoring feature, the protocol enables the \textit{cluster head} to take snapshots of the operating system within the nodes that are included in its range. These snapshots are simultaneously being visualized by base stations at very detail level. Thus, the network design captures every minute detail of malfunctioning in the knowledge networks that is composed of sensor networks, mobile ad-hoc networks, wireless networks, and semantic web.
2.4.1 Complying with Knowledge Networks

In the previous section, we have already studied the similarities and differences between four types of knowledge networks that are being considered in this study. Now, there is a need to realize the mechanism of proposed protocol which has to be submissive with all four knowledge networks domains spread between sensor networks, MANETs, wireless networks and semantic web.

2.4.2 Deployment is not Pre-Determined

We have already realized that all four types of knowledge networks are intelligently deployed and are self-organized. Proposed protocol as understood follows the swarm intelligence technique using BCO routing protocols that are themselves quite intelligent in organizing the nodes through visual sensing and feeling the movement of other nodes in the same network. Plus, this protocol is a mix of protocols leveraged from these networks. It is inclusive of COPE and NetKAT algorithms to enhance its capabilities in managing among broad spectrums of knowledge networks.

2.4.3 Pre-processing Data

It is known that sensor networks and semantic web use preprocessing techniques for partially or somewhat extracting meaningful interpretation over the web. This feature is provided in proposed protocol as it is inclusive of NetKAT and the sensor nodes in sensor networks are already been controlled via processors having inbuilt capability of preprocessing data.

2.5 Communication Channels

Wireless communications [23] have gained momentum in the twenty first century. And, with the growth in cellular use, and other networks within businesses and homes the wireless communication has replaced wired networks. Every future appliance is coming with the wireless technology and being smart every day whether it is wireless sensor
networks, automated highways and businesses, laptops and palmtops, and home appliances have all emerged in the recent researches. Now, the wireless communications network follow a highly successful IEEE standard i.e. IEEE 802.11 family of standards.

Communication whether it is broadband or point-to-point is managed by authorizing through tokens of various types. If a node or server is about to broadcast, it would circulate the broadcast token to its member nodes that has to be reached and other nodes would be kept isolated from this broadcast. Likewise if there is point-to-point communication taking place the token keys would be shared among the two nodes that need to communicate and be understood by only these two nodes. If there is a relay channel then the packets would be transmitted using COPE algorithm.

COPE [24] presents architecture for wireless networks which have significantly increased the ‘throughput’ of wireless networks. Moreover, routers intelligently merge the packets that have been received from multiple nodes to increase the throughput of network and content of each transmission.

COPE’s architecture is based on the following key points.

- The COPE design using the coding technique which sets sides of the routers to merge the information residing in the packets before forwarding them whereas the previous network coding methods are more or less theoretical and emphasize on multi-cast transmission.
- The proposed architecture [24] supports the broadcast kind of the wireless communication channel and sets out of the point-to-point abstraction for the same channel. Afterward it uses the routing techniques designed for wired networks.
- COPE encapsulates its own coding style between IP and MAC layers [24]. In an example (figure 2.5) it is simplified by showing the transmission between Alice and Bob through a router passing a pair of packets. In old scenario, Alice transmits a packet and router forwards it to Bob and vice versa. Thus, this flow of transmission requires 4 channels. In the coding approach of COPE, Alice and Bob both send their packets to the router and it in turn XORs the packets and broadcast the ‘XOR-ed’ version of the packet. Now, Bob and Alice, both obtain their packets by ‘XOR-ing’ again with their own packets. This mode of transmission limits the channels of communication to three instead of four.
2.6 Chapter Conclusion

It is a common fact that security is always of great importance in all human endeavor and due to the wireless nature of communication in sensor and ad hoc networks applications, various security threats may occur. These threats and attacks could pose serious problems to people using the wireless devices. This study finds to establish these threats and attacks and counter security measures in WSNs, MANETS and Semantic
Web Technology (In chapter 3, 4 and 5 respectively), and proposes a security architecture for the purpose of establishing Cluster based security architecture and secure routing that is discussed in chapter 6 and 7.

References


Chapter 3.
Sensor Networks Security and Routing Techniques

3.1 Introduction

This chapter presents some fundamentals about sensor communication system and architecture, hardware framework, network topology and communication protocols for Wireless Sensor Networks (WSNs). The main focus of this chapter is on the constraints of sensor security, security requirements in WSNs, attacks on WSNs, their defensive measures, and the core challenges to the design of wireless sensor network routing protocols. The clustering routing scheme, standards for industry and the latest research related to wireless sensor networks and respective application studies are also inclusive in this chapter.

The study begins with the brief introduction on wireless sensor networks, respective major applications, constraints and requirements for different applications presented in section 3.2. In Section 3.3, the communication architecture of sensor node is presented; the major components of hardware framework are also identified along with sensor network attribute and the function of each component is also explained.

In section 3.4, the network topology of sensor node is presented with concise explanatory notes. The core section of this chapter 3.5 illustrating; wireless sensor network security presenting the conventional security mechanism for WSNs are symmetric key encryption & decryption and Public Key Infrastructure (PKI) cryptography. This section further explains the layered architecture wireless sensor networks communication protocol; the functions of each layer and their requirements are also elaborated. Moreover, types of attacks and their defensive mechanisms are also discussed.

The challenges to the designing of wireless sensor network communication algorithms (routings) are presented in section 3.6; the key challenges are identified and possible solutions are proposed.
In the regards of proposed framework ONSS-KN in chapter 2, the section 3.7 is a study of clustering routing protocol for wireless sensor network, which further explains the classification of clustering, its attributes, process and characteristics of cluster head in section 3.8. Cluster based security architecture is discussed in chapter 6 which is proposed in chapter 2 as a part of security framework. The next four sections 3.9 - range-based and range-free algorithms, 3.10 - pair-wise key distribution scheme, 3.11 - coverage in wireless sensor network and 3.12 - TDMA transmission are elaborated in terms of some solution of issues listed in section 3.6.

At the end in section 3.13 and 3.14 the most recent applications of wireless sensor network are elaborated i.e. the wireless body area network and the smart sensor and actuator for power management in intelligent buildings.

### 3.2 Wireless Sensor Networks

Sensor networks have been comprehended with the growing use of wireless communications through electronic devices. The sensor networks are used within many scenarios and some of the wireless sensor networks applications are: military applications, such as battlefield surveillance, enemy tracking and battle damage assessment; civil applications, such as environment-friendly buildings and logistics, disaster detection and rescue, environment observation and forecast system; health applications such as biodiversity and habitat monitoring; these are all in addition to other commercial applications. With all these applications, there are constraints on wireless sensor networks i.e., the limited battery power of sensor nodes. Therefore, energy efficient communication is the primary requirement in wireless sensor networks.

What is a network? It is collection of several sensor nodes and its deployment is not predetermined [1]. Though, requesting the algorithms and protocols engineered to deal with these sensor nodes must be adequate intelligent to facilitate the nodes to manage themselves. The processors of sensor nodes are intelligent enough to preprocess the raw data and forward to communication channel only the essential data. Sensor networks do require wireless ad-hoc networking techniques. But the main difference between the sensor and ad-hoc network would help distinguish the two is that sensor
networks use broadcast communication and are densely deployed at higher orders of magnitude than the ad-hoc networks where as ad-hoc networks is a point to point communication.

Time delays might occur during the transmission in networked control systems. This problem has been tried to be solved by introducing fuzzy neural networks [2] for forecasting the anticipated time delay and by using this predicted time delay, sampling duration of the networked control systems is determined.

Ionic bond directed particle swarm optimization (IBPSO) [3] is proposed enhancing the Particle Swarm Optimization (PSO) where ionic bond builds a close bond between sensor nodes while PSO solves a multidimensional function optimization in moving space. Through simulation it has been seen that IBPSO functions perform better than PSO for global searching and regional convergence and implements dynamic WSN deployment in most efficient manner.

### 3.3 Communication Architecture of Sensor Node

The sensor nodes are mostly scattered in a sensor space as shown in figure 3.1. A sensor node could gather the data and send back to the sink, data routed to sink by multi-hop infrastructure-less architecture. After receiving the data sink communicates with the ‘task-manager-node’ through the satellite or may be internet. The design architecture of the sensor network is subjected by many aspects, i.e hardware constraints production cost, operating environment, fault tolerance, sensor network topology, scalability, transmission media, and power consumption.
3.3.1. Hardware Framework of a Sensor Node

The basic requisite of wireless sensor networks is to create node hardware using simple and cheap electronic components in such a way to decrease the size and cost of a sensor node. There are five major components of a wireless sensor node [4, 5, 6]. These components are as shown in figure 3.2:

![Diagram of sensor node components](image)

**Figure 3.2.** The component of a sensor Unit [1]
- Power supply
- Wireless communication device
- Computation device
- Memory
- Sensor

The architecture of a sensor node and the interconnection between the components are presented in [4, 5, 7] and shown in Figure 3.3.

![Sensor node hardware framework](image)

**Figure 3.3.** Sensor node hardware framework [4]

The component of the sensor node framework take over the data from the physical environment then transforms it into the signals and passes it on to the device for computation. The computation device executes the essential computing for the received signal and passes it to the communication device or to the memory. The performance of electronic components can be enhanced by using the high performance computing equipments. But, high performance equipments such as the transceivers in cell-phones, and the processors in modern personal computers, are very costly and consume so much of power. However, in wireless sensor networks, a battery is the only power supply for all other onboard parts of a sensor node. Due to these limitations in term of cost and power, microcontrollers are used for computing in wireless networks. The main advantage of these microcontrollers is flexible programmable/reprogrammable capability and low power consumption at a reasonable expenditure [4].
3.3.2. Attribute of Sensor Networks:

**Table 3.1**: Attributes of Sensor Networks

<table>
<thead>
<tr>
<th>Sensros</th>
<th>Size: small (e.g., micro-electro mechanical systems (MEMS)), large (e.g., radars, satellites)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number: small, large</td>
</tr>
<tr>
<td></td>
<td>Type: Passive (e.g., acoustic, seismic, video, IR, magnetic), active (e.g., radar, lidar)</td>
</tr>
<tr>
<td></td>
<td>Composition or mix: homogeneous heterogenous</td>
</tr>
<tr>
<td></td>
<td>Spatial Coverage: dense, parse</td>
</tr>
<tr>
<td></td>
<td>Deployment: fixed and planned (e.g., factory networks), ad hoc (e.g., air-dropped)</td>
</tr>
<tr>
<td></td>
<td>Dynamic: stationary (e.g., seismic sensors), mobile (e.g., on robot vehicles)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensing entities of interest</th>
<th>Extent: distributed (e.g., environmental monitoring), localized (e.g., target tracking)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mobility: static, dynamic</td>
</tr>
<tr>
<td></td>
<td>Nature: cooperative (e.g., air traffic control), non-cooperative (e.g., military targets)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operating environment</th>
<th>Benign (factory floor), adverse (battlefield)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>Networking: wired, wireless</td>
</tr>
<tr>
<td></td>
<td>Bandwidth: high, low</td>
</tr>
<tr>
<td>Processing architecture</td>
<td>Centralized (all data sent to central site), distributed (located at sensor or other sites), hybrid</td>
</tr>
<tr>
<td>Energy availability</td>
<td>Constrained (e.g., in small sensors), unconstrained (e.g., in large sensors)</td>
</tr>
</tbody>
</table>

Security: Since the sensor network may operate in a restrictive environment, security should be built into its design. Network system techniques are required to provide survivable, low-latency and secure networks. Low probability of detection communication is looked-for networks because sensors are being visualized for use across the enemy lines. For the same reasons, the network should be protected against spoofing and intrusion [8].

3.4 Network Topology of Sensor Networks

The network layer is responsible for packet forwarding through intermediate nodes to the receiver. In sensor networks, data is sent from the source node to the destination using the intermediate nodes, also known as multi-hops communication where as considerable energy can be saved using Multi-hop routing algorithms [4, 9]. The routing algorithm forms the wireless sensor network topologies on the basis that data
streams from the source to the destination node. There are three important network topologies for wireless sensor networks, as shown in figure 3.4.

These topologies are:

- Tree topology
- Cluster topology
- Mesh topology

In the tree topology, the destination resides in the centre and other nodes are around the receiver node. Data from sensor nodes are communicated to the receiver through intermediate nodes. Therefore data traffic converges in the direction of the receiver and shapes a tree-like network topology.

![Diagram showing primary topologies of wireless sensor networks](image)

**Figure 3.4**. Primary topologies of wireless sensor networks[4]
Cluster topology is also known as star topology in which a set of sensor nodes gather themselves to transmit data. This collection of sensor nodes is called a cluster. In a cluster, the topology set of sensor nodes send the data to one node within the cluster; this is called as the Cluster Head (CH) node. The CH node then forwards the data to the destination. Generally, CH nodes located in the centre of the cluster. All the nodes within the cluster transmit the data to the CH and then the CH forwards the data towards the destination. The cluster based topology is commonly used in many wireless sensor networks, such as intelligent agriculture and environment monitoring. In cluster based topology, the data flow pattern of network is described as being “many-to-one” [1, 9,10].

In mesh topology, wireless links can be established between any two sensor nodes. Mesh topology is very effective to achieve the high data transmission reliability. In mesh topology, some redundant links for sensor nodes are established to transmit the data which ensure the reliable data transmission when there is congestion or path failure in the network.

3.5 Wireless Sensor Network Security

The wireless sensor networks generally function in uncontrolled public area, therefore the security is a key challenge in sensor applications. The conventional security mechanisms are symmetric key encryption & decryption, authentication and Public Key Infrastructure (PKI) cryptography [11,12,13]. The main challenge is to set up the encryption techniques and their counterparts in a sensor network which is characterized with constrained processing capability, power supply and memory. For WSN several security schemes were proposed but they have restricted attributes for example they are only concerned to attack on a particular layer. Some others have also proposed a theoretical framework that is not suitable at deployment time [14].

3.5.1. Security Requirements of WSN

Data Confidentiality: In many applications, nodes communicate highly sensitive data, (e.g., key distribution) therefore it is extremely important to build a secure
communication channel in a wireless sensor network. The adversary can change the data, so as to send the sensor network into disarray. For example, a malicious node may add some fragments or manipulate the data within a packet.

*Data Freshness:* Data freshness suggests that the data is current, and it ensures that no old messages have been replayed. This requirement is especially important when there are shared-key strategies employed in the design.

*Self-Organization:* A wireless sensor network is typically an ad-hoc network, which requires every sensor node to be independent of the others and flexible enough to be self-organizing and self-healing according to different situations.

*Time Synchronization:* Sensors may wish to compute the end-to-end delay of a packet as it travels between two pairs of wise sensors. A more collaborative sensor network may require group synchronization for tracking applications, etc.

*Secure Localization:* A sensor network will rely on its ability to accurately and automatically locate each sensor in the network. A sensor network designed to identify faults will need accurate location information in order to pinpoint the location of a fault. For large sensor networks, the SPINE (Secure Positioning for Sensor Networks) algorithm is used. It is a three-phase algorithm based on verifiable multi-alteration.

*Authentication:* Data authentication allows a receiver to verify that the data really is sent by the claimed sender. In the case of two-party communication, data authentication can be achieved through a purely symmetric mechanism: the sender and the receiver share a secret key to compute the message authentication code (MAC) of all communicated data.

### 3.5.2. Wireless Sensor Network Communication Protocols

Because of the presence of extensive applications of wireless sensor networks, there is deep interest in the design of wireless sensor network communication protocols equally in industry and in academia. Several solutions are also proposed to handle the challenges of design in wireless sensor networks. In this section some of the more significant solutions are presented. The solutions for every issue related to each layer of wireless sensor network communication protocols are developed. For practical
application, for the design of wireless sensor networks, the Zigbee protocol [15] is adopted, in addition to the IEEE 802.15.4 standard [15].

Wireless sensor networks use layered architecture same as wired network architecture. Characteristics and purpose of each layer are given below.

![Layered Architecture of WSN](image)

**Figure 3.5** Layered Architecture of WSN[15]

**Physical Layer:** The objective of physical layer is to reducing the path loss effect and shadowing and increasing the reliability. This layer is responsible for establishing the connection, modulation, data encryption, frequency generation, data rate, and signal detection.

**Data Link Layer:** The objective of Data link layer is to ensure interoperability between node to node communications. This layer is also responsible for error detection, prevention of collision of packets, multiplexing, repeated transmission etc. To secure the data link layer, Karlof has proposed link layer security architecture “TinySec” for wireless sensor networks. Some researchers also proposed the use of public key cryptography [16,17], secure code distribution [18] to create secure key during network
deployment and maintenance. Naveen Sastry proposed Zigbee or the 802.15.4 standard for hardware based symmetric key encryption.

**Network Layer:** The objective of Network layer is to find the best path for efficient routing mechanism. This layer is responsible for routing the data from node to sink, node to node, node to cluster head, node to base station, and vice versa. The PEGASIS and LEACH are the protocols which explain the techniques to save the energy from consumption so as to improve the life of the sensors. LEACH presents the cluster based transmission where as PEGASIS is a chain protocol [19]. In WSN, a node in the network acts as a router because of the broadcast mechanism, so that to create a secure routing protocol encryption and decryption techniques are used for secure routing [20, 21, 22].

**Transport Layer:** The objective of Transport Layer is to establish communication for external networks i.e. connection to the internet is the most confronting problem in wireless sensor networks.

**Application Layer:** The objective of Application Layer is to present ultimate outputs which makes sure the smooth information flow to lower layers. This layer is responsible for the collection of data, its management and processing through the application software.

### 3.5.3. Attacks on WSN and Their Mitigation

The security breaches appear mainly in the form of Interception- an unauthorized access of WSN, Interruption-breakdown of communication links, Modification-Change of data by unauthorized access and the fabrication-addition of false data by unauthorized accesses [20,23].

**Denial of service:** “Any event that diminishes or eliminates a network’s capacity to perform its expected function”. This type of attack might occur by occupying the resources to their respective intended users. At physical layer it may occur by blocking the fabrication of data packet. In Data Link Layer it is by producing unfairness in use of networks, exhaustion of resources and collision data. In network layer, it occurs by neglecting packets resulting into path failure and the way of greediness. DOS attack occurs due to flooding and de-synchronization in transport layer. The significant
identification and authentication mechanisms could be useful to prevent denial of service attacks. To handle jamming at the MAC layer (Medium Access Control Layer), nodes might utilize a MAC admission control that is rate limiting. This would allow the network to ignore those requests designed to exhaust the power reserved for a node.

**Attack of information in transit:** Typically for WSNs each node notifies the modification to the cluster head or to a base station only if the data exceeding form some threshold value. Information in transit in such communication patterns may be replayed again, spoofed, vanished or altered. This type of attack may be prevented by data aggregation and authentication techniques.

**Sybil attack:** Sybil attack is defined as a “malicious device illegitimately taking on multiple identities” [23]. It was originally described as an attack able to defeat the redundancy mechanisms of distributed data storage systems in peer-to-peer networks. In addition to defeating distributed data storage systems, the Sybil attack is also effective against routing algorithms, data aggregation, voting, fair resource allocation and foiling misbehavior detection. For instance, in a sensor network voting scheme, the Sybil attack might utilize multiple identities to generate additional “votes.” In this kind of attack the attacker acquire illegally some identities on one node as shown in figure 3.6 attacker node gets ID1 and ID2. Sybil attacks are typically banned by the ‘validation’ methods that have influenced the routing mechanism. In direct validation a trusted node directly tests whether the joining identity is valid. Direct validation techniques include a radio resource test. Indirect Validation: -In the indirect validation, another trusted node is allowed to vouch for (or against) the validity of a joining node [23].

**Blackhole/ Sinkhole Attack:** In this kind of attack, the attacker exists in a network with high bandwidth and high processing power so it will be possible to create the shortest path. Resultantly, all data passes through the attacker’s node. According to Liu and Ning, the Trust Routing for Location Aware Sensor Networks (TRANS) routing protocol is designed to be used in data centric networks. It also makes use of a loose time synchronization asymmetric cryptographic scheme to ensure message confidentiality. In their implementation, µTESLA is used to ensure message authentication and confidentiality. Using µTESLA, TRANS is able to ensure that a message is sent along a path of trusted nodes while also using location aware routing.
Figure 3.6 Sybil Attack [23]

Figure 3.7 Blackhole/ Sinkhole Attack [23]
'Hello flood' Attack: This is one of the most effort-less attack in sensor networks in which sender or receiver broadcasts the ‘HELLO’ message packets with high transmission power. The receiver nodes suppose that the sender node sending the message is the closest to them and send the packet. By this attack, congestion may occur in the network. 

Wormhole Attack: In this type of attack, tunneling mechanism is used by the attacker to establish the communication link between them by confusing the routing protocol. As shown in figure 3.8, suppose the node ‘Y’ broadcasts the data to find the path. Whereas the ‘X’ presents himself as a node ‘X’ and sends back the acknowledgement to ‘Y’. ‘Y’ sends data to ‘X’ that is received by ‘X’ and ‘X’ sends that data to ‘X’ by tunneling, hiding its own identity. In this case ‘Y’ and ‘X’ are not in a one hop but they suppose they are in a single hop range. And the attacker ‘X’ may destroy the security mechanism by interruption, interception, fabrication and modification. Additional hardware, such as a directional antenna is used to defend against wormhole attacks. If a wormhole exists within the network, the shape of the virtual network will bend and curve towards the offending nodes. Using this strategy the nodes that participate in the wormhole can be identified and removed from the network. If a network does not contain a wormhole, the virtual network will appear flat [23].

![Figure 3.8 Wormhole Attack](image)

Secure Data Aggregation: An aggregator is responsible for collecting the raw data from a subset of nodes and processing/aggregating the raw data from the nodes into more usable data. Aggregate-commit-prove technique comprises three phases. Aggregate: The aggregator collects data from the sensors and computes the aggregation result according
to a specific aggregate function. Each sensor shares a key with the aggregator. This allows the aggregator to verify that the sensor reading is authentic.

**Defensive Measures - Key Establishment:** Traditionally, key establishment is done using one of the many public-key protocols. One of the more common public-key protocols is the Diffie-Hellman public key protocol, but the protocols are incompatible in low power devices such as wireless sensor networks. This is due largely to the fact that typical key exchange techniques use asymmetric cryptography, also called public key cryptography. In this case, it is necessary to maintain two mathematically related keys, one of which is made public while the other is kept private. This allows data to be encrypted with the public key and decrypted only with the private key. The problem with asymmetric cryptography, in a wireless sensor network, that it is typically too computationally intensive for the individual nodes in a sensor network. This is true in the general case. [23,24] show that it is feasible with the right selection of algorithms.

Symmetric schemes utilize a single shared key known only between the two communicating hosts. This shared key is used for both encrypting and decrypting data. The traditional example of symmetric cryptography is DES (Data Encryption Standard). The use of DES, however, is quite limited due to the fact that it can be broken relatively easily. Other symmetric cryptography systems have been proposed including 3DES (Triple DES), RC6, AES, and so on.

### 3.6 Routing Design Issues and Challenges in WSNs

In spite of the several applications of WSNs, sensor networks have numerous confines such as limited bandwidth, limited computing power and, limited energy supply for the communication links. The major objectives regarding the design of the WSN is to bring out the data communication, however, prevent the connectivity deprivation by utilizing aggressive ‘energy management’ techniques and to increase the lifetime of the network.
3.6.1. Node Deployment

Node deployment in WSNs could affect the performance of the routing protocol and is application dependent. The deployment can be either randomized or deterministic. In random deployment mechanism, the nodes are spread randomly in ad-hoc manners which creates an infrastructure whereas in deterministic, the sensor nodes are deployed manually and data is transmitted via predestined paths. However, an optimized clustering developed to facilitate the energy efficient network function and connectivity if the distribution of nodes is not consistent.

3.6.2. Energy Consumption without Losing Accuracy

Sensor nodes have very restrictive power energy to perform the computations and Data transmission in a wireless infra-structure. The energy exhaustion of some nodes may originate the serious connectivity problems and result comes out as in a network breakdown [24]. In a multi-hop wireless network, each node acts a double role as data router and data sender. The out of order sensor nodes due to power cut down can be the reason of basic topological alteration and requiring re-routing of data packets and re-organization of the network.

3.6.3. Link/Node Heterogeneity

The existence of heterogeneous sensor network has many technical data routing issues. Even though the data collection, its analysis and results produced from these nodes are subject to distinct constraints of quality of service and at different rates and can use several reporting models. By clustering the network, the cluster head can be chosen from the deployed nodes which are more powerful to other sensor in-terms of subject to distinct quality of service constraints such as memory, energy, and bandwidth. Consequently, the load of transmission is managed by the group of cluster heads.

3.6.4. Data Reporting Model

Data collection and analysis in WSNs could be independent of the application requirements and time factor which criticality matters for data reporting. It can be
grouped as event driven, hybrid, query driven or time-driven. The routing protocol particularly concerned by the ‘reporting model’ with respect to the route stability and energy consumption.

### 3.6.5. Fault Tolerance

Few sensor nodes might be blocked or failed for the reason of physical damage, environmental interference or lack of power. The main point noticed is that makes the network fault tolerance as the failure of sensor nodes should not disturb the general operations of the sensor network. If numerous nodes get failed, routing protocols and MAC must have to be compensated for new or duplicate routes and links to the base stations for data collection.

### 3.6.6. Scalability

Communication protocol should produce a high performance for a huge number of nodes as a sensor network having big number of sensor nodes and is deployed in wide area, [1]. Any routing adjustment must have to be worked with this large number of sensors.

### 3.6.7. Network Dynamics

Most of the sensor network design supposed that the nodes are fixed. Though, in various applications the mobility of either sensor nodes or BS’s is someway necessary [22]. It could be more challenging to rout the messages from or to non-stationary nodes as route consistency becomes a critical task, in addition to energy, bandwidth, etc.

### 3.6.8. Transmission Media

The typical issues related with a wireless communication channel may also disturb the operations of the sensor networks. In a multi-hop sensor network, nodes are linked by a wireless medium for communication [25]. TDMA based protocols can be adopted to design the MAC of sensor network that might save the energy in contrast of ‘contention’ based protocols such as CSMA, Bluetooth technology [26] can also be used.
3.6.9. Data Aggregation

Data aggregation is the collection and emergence of data received from various sources as have to apply aggregation function, such as average, duplicate suppression, maxima, and minima. These techniques have been used to attain the data-transfer optimization in routing protocols and efficiency of energy.

3.6.10. Coverage

A specified sensor’s characteristics are bounded in accuracy and in range; it can cover a definite physical area. So, the coverage of area is also a key design constraint in WSNs.

3.6.11. Quality of Service

In some applications, it is a basic and critical requirement that the data should be transferred within a definite period of time from once it is sensed; otherwise that data will be insignificant. As the energy gets exhausted, it may be required by the network to reduce the ‘results quality’ in the nodes and therefore extend the total lifetime of network. So the, energy aware routing protocols have major role to reflect this constraints.

3.7 Clustering Routing Protocol in Wireless Sensor Network

Routing is one of the significant technologies in sensor networks as divergent to the classical ad-hoc networks; it is more challenging due to their inherent features [27] as discussed in last section.

Routing protocols based on network structure in sensor networks can be distinguished into two groups: the hierarchical and flat routing. In a flat WSNs structure, all nodes in the network have the similar functionalities and perform the similar tasks. A form of flooding into transmission is act upon hop to hop channel. The distinctive flat routings in WSNs include Sensor Protocols for Information through Negotiation, Flooding and Gossiping, Rumor, Greedy Perimeter Stateless Routing (GPSR), Directed
Diffusion (DD), Energy-Aware Routing (EAR), Trajectory Based Forwarding (TBF), Sequential Assignment Routing (SAR), Gradient-Based Routing (GBR), etc. [27]

In contrast to flat, nodes of hierarchical topology carry out special designated activities in WSNs and characteristically are grouped into clusters as per defined metrics or requirements. However, each cluster has a monitor called as cluster head and other ordinary or member nodes. Typically, nodes having higher most energy play a role of CH and perform the job of information transmission and data-processing, whereas nodes having least energy considered as a member node carry out information sensing activity. The clustering routings protocols in wireless networks are Low-energy Adaptive Clustering Hierarchy, Hybrid Energy-Efficient Distributed clustering, Two-Level Hierarchy LEACH, Distributed Weight-based Energy-efficient Hierarchical Clustering protocol, Energy-Efficient Uneven Clustering algorithm, Position-based Aggregator Node Election protocol, Energy Efficient Clustering Scheme, Unequal Clustering Size model, Base-Station Controlled Dynamic Clustering Protocol, Algorithm for Cluster Establishment, Power-Efficient Gathering in Sensor Information Systems, Threshold sensitive Energy Efficient sensor Network protocol, Two-Tier Data Dissemination, Hierarchical Geographic Multicast Routing, Adaptive Threshold sensitive Energy Efficient sensor Network protocol, and Concentric Clustering Scheme, etc. Clustering routing in WSNs have advantages, few of them are such as scalability, data aggregation, less energy utilization, less load and robustness.

3.8 Classification of Clustering Elements in WSNs

The following are the clustering elements in Wireless Sensor Networks which includes cluster attributes, cluster head properties, clustering process and categories of clustering algorithms.

3.8.1. Cluster Attributes

Variability of Cluster Count: On the basis of variability cluster count, clustering methods can be categorized into variable form of clustering and fixed form. In the fixed method,
the number of clusters is defined as it could not be changed and its sets are predestined. In spite of this, in the latter scheme the number of clusters is changeable and selection accomplished on the bases of predetermined rules or randomly [27].

**Uniformity of Cluster Sizes:** The clustering schemes have adopted to avoid energy hole and attain the more uniform energy consumption in which the network is divided into cluster by using different sizes. Its routing protocols can be categorized into two forms on the base of uniformity of cluster size, even and uneven, respectively.

**Inter-Cluster Routing:** On the basis of communication and connectivity inter-cluster routing can be distinguished into further two types, one is single hop and other is multiple hopes inter-clustering routing mechanism. In a Single hop routing between clusters, all routers directly communicate to the base station whereas for multiple hop inter clustering routing scheme CHs are used for data transmission.

### 3.8.2. Cluster-Head Characteristics

**Difference of Capabilities:** On the basis of uniformity of energy distribution activity, clustering mechanism in sensor networks can be categorized as into two: heterogeneous and homogeneous. Homogeneous clustering allocates an equal energy; computation power and resource for communication to each node of the cluster. In contrast to it, in heterogeneous scheme CH’s are elected already and designate the energy, computation and resources in an unequal capacity.

**Mobility:** The mobility feature of Cluster heads defines that its membership toward leader of the cluster could be changed dynamically and consequently this cluster required to be maintained on regular bases. Despite that Stationary scheme maintains the cluster stability and it would be simple to manage the network operations [26].

**Role:** A CH performs aggregation of data that has been gathered from member nodes of the cluster. At the same time, CHs works as a BS that performs operations on the bases of data analysis and detection [26].

### 3.8.3. Clustering Process

**Control Approaches:** The routing schemes in WSNs on the basis of some control of clustering have been classified into distributed, centralized routing scheme, and hybrid. In
the distributed method, a sensor node could become a CH without having any world wide information about the network. Whereas in centralized approach, CH functioning is totally based on worldwide information about the network cluster. Hybrid is a combination of both the distributed and centralized approaches. In this situation, centralized mechanism is adopted for CHs to build exclusive clusters and distributed manners are used for co-ordination between the CHs.

*Convergence Time:* Clustering schemes in sensor networks are classified as constant and variable by considering the convergence time, after a completion of defined iterations, an algorithm must have to be converged without considering the scaling constrains of the network. The convergence time in variable; convergence algorithms depending upon the quantity of node present into the network at that time, which adjusted finely to small scale networks.

*Parameters for CH Selection:* For the election of CH, there have been three approaches defined such as adaptive, deterministic approach, and randomly election. The selection criteria might be on the parameters i.e. weights, communication cost, residual energy, number of neighbors and the identifier etc.

### 3.8.4. The Proceeding of Algorithm

Generally, all clustering routing algorithm consists of two fundamental steps,

- Cluster construction
- Data transmission.

We can say that the complete clustering routing protocol and algorithm procedure could be organized into cluster construction and data transmission.
Figure 3.9 Clustering Methods in WSNs [27]
3.9 Range-Based and Range-Free Algorithms

Wireless Sensor Networks [28] acquired lot of research interest in the recent time as there is continued technological uplift. Gathering information from its surroundings WSNs consist of thousands of built-in sensor nodes. When collecting the desired information there is need to know the exact location of its occurrence and here researchers face problem in localizing the position of information. Localization of position would be then required in applications built for surveillance, intrusion detection, fire alarms, traffic monitoring, and lot more. The location is estimated by recording the position of other few sensors in the area and finding out their distance and bearing measurements. GPS is not a possible solution for WSNs for cost limitations thus recognition is done for a need for specific localization algorithm for WSNs.

The importance is being given to self-localization algorithms that are defined as range-based and range-free. Range-based algorithms [28] use point-to-point and angle information of their neighboring nodes to calculate distance and position. They have inbuilt hardware that makes range measurements. Most of the range-based algorithms are Angle of Arrival, Time of Arrival, Received Signal Strength Indicator, and Time Difference of Arrival etc.

Range-free algorithms [28] work without absolute range information. Accuracy gets affected but it is still helpful to many applications. There is little requirement for built-in hardware. These algorithms are cost effective and feasible for large scale WSNs. Examples of range-free algorithms include; DV-hop, centric algorithm, MDS-MAP, amorphous, and APIT, etc. Range-free algorithms are preferred because WSN devices have hardware limitation and range-free algorithm is cost-effective and handy in this perspective.

**DV-Hop Localization Algorithm:**

*Step 1: Finding the Minimum HopCount:*

At first, all Beacon-Nodes broadcast beacon-messages in the format of the beacon message is \{id, x_i, y_i, HopCount\} to other nodes; which set the initial value of HopCount is 0. Each receiving node keeps the minimum HopCount value from other nodes and maintains a Beacon-Node information table. Once a process is completed after a specific
period of time, the entire nodes network will have the minimum HopCount of other nodes.

**Step 2: Calculation of the Distance between Beacon-Nodes and Unknown-Nodes:**
Secondly, the Beacon-Nodes get minimum HopCount value of other Beacon-Nodes as defined in previous step by using it can estimate the average HopSize to the other Beacon-Nodes on the basis of the minimum HopCount and the distance between Beacon-Nodes. The i’th Beacon-Node’s average HopSize can be achieved by

\[ \text{HopSize}_i = \frac{\sum_{j=1, j \neq i}^{N} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{\sum_{j=1, j \neq i}^{N} h_{ij}} \]  

(3.1)

Where \((x_i, y_i)\) and \((x_j, y_j)\) are coordinates of Beacon-Nodes \(i\) and \(j\), \(N\) presents the quantity of Beacon-Nodes, \(h_{ij}\) is the minimum HopCount value between Beacon-Nodes \(i\) and \(j\). Every Beacon-Node broadcasts its average HopSize to the whole network. Each Unknown-Node receives all Beacon-Node’s average HopSize and selects the HopSize of a Beacon-Node, which has the minimum HopCount value to this Unknown-Node, as its average HopSize. So at the end, we can compute the distance of every Unknown-Node to Beacon-Node by:

\[ d_{ij} = \text{HopSize}_i \times \text{HopCount}_{ij} \]  

(3.2)

Where \(d_{ij}\) is distance between the \(i^{th}\) Unknown-Node and \(j^{th}\) Beacon-Node. \(\text{HopCount}_{ij}\) indicates the minimum HopCount between the \(i^{th}\) Unknown-Node and \(j^{th}\) BeaconNode.

**Step 3: Calculation of Estimated Location:**
In the last step, the Unknown-Node can calculate their locations by maximum likelihood estimation.

3.10 Pairwise Key Distribution Scheme

A novel pairwise key distribution scheme is proposed based on the deployment knowledge within the self-organized WSNs and irreversibility of some hash chains [20].
Before deploying the nodes, the offline server constructs a number of hash chains and using values from pairs of reverse hash chains to establish their pairwise keys among the nodes in the same region. While among the neighboring nodes in different regions, few pairs of hash chains based on the deployment knowledge are established for pairwise keys. Such procedures employed give the attackers tough time to break the network ensuring the probability of establishing pairwise keys is close to 1. It has been compared with q-composite’s scheme and Dai scheme and hence discovered that DKH-KD scheme efficiently improves the probability of pairwise key establishment and invulnerability.

3.11 Coverage in a Wireless Sensor Network

Recently, there have been lots of researches committed to the sensor networks, which highlights its design challenges associated to routing, transport protocols [3,5,7] and media access layers (physical layer) [12,16,29]. Since sensors may be scattered in a random manner, the primary concern is the coverage of sensors.

Solution of Coverage problem: The coverage problem look like to be difficult and its immature resolution yet has to be discovered by sensing the all divided region borders of all \( n \) (n circle) and then verify that if each part or region is absolutely covered by \( k \), as shown in figure 3.10. It would be computationally expensive task \( O(n^2) \) in geometry to manage all sub-regions.

Solution to the k-UC problem: A cost of \( O(nd \log d) \), where \( d \) is the maximum number of sensors whose sensing ranges may intersect a sensor’s sensing range. Instead of determining the coverage of each sub-region, our approach tries to look at how the perimeter of each sensor’s sensing range is covered.

Definition 1. Consider any two sensors \( s_i \) and \( s_j \). A point on the perimeter of \( s_i \) is perimeter-covered by \( s_j \) if this point is within the sensing range of \( s_j \).

Definition 2. Consider any sensor \( s_i \). We say that \( s_i \) is \( k \)-perimeter-covered if all points on the perimeter of \( s_i \) are perimeter-covered by at least \( k \) sensors other than \( s_i \) itself. Similarly, a segment of \( s_i \)’s perimeter is \( k \)-perimeter-covered if all points on the segment are perimeter-covered by at least \( k \) sensors other than \( s_i \) itself.
Figure 3.10. Coverage Dilemma [29]

(a) The sensing ranges are unit disks, and
(b) The sensing ranges are non-unit disks

3.12 TDMA transmissions

Multi-hop networks provide channel for TDMA transmissions for which algorithms are defined for scheduling to determine the smallest lengths of slots that are conflict-free where each link or node is activated for once at least [26]. Self-organized point-to-point flows form the bases of this assumption. Sensor networks are the source of data transmission through sensor nodes to few central data collectors. To schedule it requires finding the smallest distance comprising of conflict-free availability of slots for the packets to travel from source to destination node. Also the interference graph comes out of it where conflicting nodes transmissions can be visualized and it is different from the connectivity graph due to broadcast nature of these transmissions. This problem is proved as NP-complete [26]. Two centralized heuristic algorithms are proposed; one is based on direct scheduling of nodes and this approach is adapted from classical multi-hop scheduling algorithms for general ad hoc networks, while the other approach levels the routing tree before scheduling the nodes and it is novel approach for many-to-one communication in sensor networks. These algorithms perform best depending on distribution of nodes across the levels. Then a distributed algorithm is proposed that depends on distributed coloring of nodes increasing the delay by factor of 10-70 over
centralized algorithms where there are 1000 nodes. Upper bound is also determined for these schedules as a function of total number of packets generated in the network [26].

Multi-hop TDMA scheduling is more challenging than one-hop scheduling because spatial reuse of a time slot may be possible: More than one node can transmit at the same time slot if their receivers are in non-conflicting parts of the network.

### 3.13 Wireless Body Area Networks

Wireless Body Area Networks (WBANs) [30] has emerged with the increasing use of wireless networks and the constantly diminishing electrical devices. A Wireless Body Area Network attached on or implanted in the body consists of small, intelligent devices which are capable of establishing a wireless communication link. These devices provide continuous health monitoring and real-time feedback to the user or medical personnel. These networks involve implantation of various sensors on clothing or on the body or even implanted under the skin. This is a breakthrough in order to improve quality of life as in healthcare. The patient becomes more mobile rather staying in hospital when on WBAN. Variety of applications on WBAN is surveyed used for patient monitoring. WBAN communication and its positioning within different technologies have also been focused.

Generally, sensors and actuators are well-known two types of devices. These sensors can compute body temperature, heartbeat, used to record a lengthened electrocardiogram being in WBAN, see figure 3.11.

![Figure 3.11 Intra-body and extra-body communication](image)

Figure 3.11 Intra-body and extra-body communication [30]
Figure 3.11 shows that the control of the data between personal device and sensor [31], which later make sure the connection between the PDA and external network [32].

Figure 3.12 shows a WBAN associated Wide Area Networks (WAN), Wireless Metropolitan (WMAN), Wireless Local (WLAN) and Wireless Personal (WPAN) [30].

![Diagram of WBAN in the area of wireless networks]

**Figure 3.12.** Place of a WBAN in the area of wireless networks [30]

### 3.14 Smart Sensors and Actuator for Power Management in Intelligent Buildings

A power management system [33] is a technical solution for monitoring and easy control of household appliances. WSNs are increasingly being used in the home for energy controlling services. Regular household appliances are monitored and controlled by WSNs installed in the home. New technologies include cutting-edge advancements in information technology, sensors, metering, transmission, distribution, and electricity storage technology, as well as providing new information and flexibility to both consumers and providers of electricity.

*Control of Home Appliances:* System has been designed for measurement of electrical parameters of household appliances. Important functions to the system are the ease of modeling, setup, and use. From the consumer point of view, electrical power consumption of various appliances in a house along with supply voltage and current is the key parameter. Thus, the real-time monitoring of the electrical appliances can be viewed...
through a website. The sensor networks are programmed with various user interfaces suitable for users of varying ability and for expert users such that the system can be maintained easily and interacted.

3.15 Chapter Conclusion

Wireless sensor networks attained substantial interests over the last few years, and can be engaged in a vast scale of applications in both military and civilian scenarios. The design of robust, effective, and scalable routing protocols for WSNs is a challenging assignment. It is observed, that the clustering routing algorithms, generally is more suitable to tackle the constraints and the challenges of WSNs. Resultantly, considerable hard works have been done in attending the techniques to design efficient and effective and clustering routing protocols for WSNs in last year.

This chapter presents study of the fundamentals of the wireless sensor networks. The importance of wireless sensor networks and their real time application is presented. The sensor node hardware framework and wireless sensor network protocols are also presented. Wireless sensor network protocol layer are elaborated as; the challenges of the requirements for the design of efficient protocols are identified and presented. The literature review relates to the state-of-the-art research on wireless sensor networks communication protocol; the security threats, its mitigations and design requirements are conducted and elaborated.

References


Chapter 4

4.1 Introduction

Ad-hoc networks are a wireless networking prototype for mobile devices. In contrast to conventional wireless networks, MANETS do not depend on any predetermined infrastructure. In fact, these networks are self-directed and self-configurable systems which are intelligent enough to hold up movability. These exclusive attributes of MANETS causes several challenges for the implementation of secure communication in the wireless network system design. In this chapter, we study some basics about the MANETs system, its architecture, topology management, understanding the vulnerabilities and security goals. Moreover, we present the different security attacks and investigate approaches to secure the mobile ad hoc communication.

First we concisely introduce the basic characteristics and major applications of Mobile Ad hoc networks then the constraints and requirements for different networks presented in section 4.2. In Section 4.3, the network topology of mobile ad hoc devices is discussed with briefing on 4G and Bluetooth technology features. In section 4.4 a simple architecture of MANETS is presented; the major components and cross layer issues such as enabling technology, Networking layers and Middleware applications versus QoS, simulation, energy conservation and security respectively.

In section 4.5 we discuss some distinctive and hazardous vulnerabilities and security attacks in the MANETs and most of which are posed by the basic features of the MANETs such as restrictive power supply, scalability, lack of secure boundaries, threats from compromised node, and lack of centralized management facility. The presence of these vulnerabilities and threats has made it mandatory to discover some proficient security solutions and secure the MANETs from all types of security hazards.

In the core section 4.6 of this chapter we introduce the most recent development in security schemes for the MANETs which are intrusion detection and secure routing techniques. In each of security schemes, various methods are pointed out and compared
with each other and observed that some methods have deficiencies. The section 4.7 start with the discussion on the security mechanism in MANETs, which is supposed as a direction to the security-related research works in this field; Key Management as for preventing external attacks and secure routing protocols as for preventing internal attacks are the two suggested solutions for aforementioned security mechanism.

In the regards of the proposed security framework ONSS-KN in chapter 2, in section 4.9, we discuss the procedure that breakup the network into interlined substructure (Clustering), respective schemes and algorithms. Cluster based security architecture is discussed in chapter 6 which is proposed in chapter 2 as a part of security framework.

### 4.2 Mobile Ad-hoc Networks

In the recent years, we have seen a prompt expansion in the area of mobile computing due to the production of inexpensive, extensively available wireless devices. However, current devices, protocols and applications only focused on wireless local area networks (WLANs) or cellular, not taking into account the great potential offered by mobile ad-hoc networking. A mobile ad hoc network is an autonomous collection of mobile devices (laptops, sensors, smart phones, etc.) that communicate with each other on wireless links and co-operate in a distributed way in order to offer the essential network functionality in the nonappearance of a fixed infrastructure. This type of network, operating as a stand-alone network or with one or multiple points of attachment to cellular networks or the Internet, open the way for several exciting and new applications. Application scenarios include, but are not limited to: conference or campus settings, emergency and rescue operations, personal networking, car networks, etc.

Self-administering wireless network [1], contrasting to infrastructure wireless networks, where each user directly communicates with a base station or access point, MANET, does not base on a fixed infrastructure for its operations as shown in Figure 4.1-a The network is an independent transient connection of mobile nodes that communicates with each other over the wireless links.
In a MANET, user’s mobile machines are the actually network, and devices co-operatively provide the infrastructure that basically provided by the network device (e.g., routers, servers, switches). MANET itself is an infrastructure required to facilitate the data transmission within the user’s mobile devices [2]. MANETs are growing fast because it doesn’t require communication infrastructures in the areas to get the network services [3, 4]. MANET establishes a proper framework by providing a multi-hop wireless network.

Figure 4.1-a Wireless LAN, Ad-hoc network versus MANET [1]

Figure 4.1-b showing the implementation of MANETs, it’s on an airport, where crowed can access LAN and WAN, notebook computers and wireless-CDMA mobile phones, Bluetooth technology are used to interconnect carried devices on ad-hoc, such as PDAs,. For example, a client can access e-mail via a HiperLAN interface to a laptop but peruse and reply to these messages via his PDA.
4.3 Topology Management

The mobile devices in a mobile ad-hoc network (MANET) [5, 6] play a very important role as compared to a conventional wireless LAN (WLAN). In a traditional WLAN, interactions are centered on the access point or base station; the communication infrastructure up to the base station is typically fixed, so the topology is consistent whereas in MANET the mobile devices act as routing device as end systems. Ad-hoc networking model is not an innovative one, indeed it’s been in variety of mode around us for more than 20 years. Conventionally, there have been few networking application that go along with the ad-hoc standards. A decade of, the presenting the advanced technologies which are the IEEE 802.11, Bluetooth, and Hyperlan make it possible to deploy the commercial MANET outside the military realm. The Mobile ad-hoc topology that is potentially extremely dynamic is dependent on the connections of nodes and relative locations within the network. This affects in all aspects of an ad-hoc network, as well as the routing protocols and medium access control (MAC) layer. In
respect to achieve the adequate performance, the MANET must find an efficient ways to manage the side affected by the changing topology as a whole.

4.3.1. 4G and ad-hoc networking

The aims for the evolution of 4G Wireless is given that of persistent computing world that can ubiquitously and seamlessly help the users in accessing information, performing their activities, or making a communication with other users from any device at anywhere, [7]. This view profoundly based on 4G mobile and wireless communications. Basically, 4G technology is a global network of an integrated system. Main focus of 4G is to integrate the different type of wireless network by the support of wired back-bone infrastructure for data traffic on an IP based core network. As shown in Figure 4.2. A 4G networks architecture and its components.

![Figure 4.2 4G Network][1]

In such an environment, computer devices get a step further into background; network connectivity and computing power are embedded virtually in all devices regardless of

---

[1]: https://example.com/image.png
where they are placing and in what environment to bring computation to users. These computing machines personalized themselves to discover new software or information.

The intersecting of different network borders correspondence to the mixing of different networks in 4G. The theory on the bases of 4G has been launched that upcoming networks will completely be based on packet-switching core technology. This has also concluded that 4G-IP systems will be more efficient and cheaper than 3G.

4.3.2. Bluetooth Networking

The major function of Bluetooth [8] is to provide a substitute of the cables between electronic devices by using a low-cost radio chip, such as laptop, computers, telephones, PDAs, printers, digital cameras, and fax machines. Such short range connectivity fits greatly into the broader context, so it can extend IP based networking into the personal network area domain.

4.4 MANET Architecture

As we can see in figure 4.3 the MANET architecture, there are three main areas:

- Middleware and applications;
- Networking and;
- Enabling technologies.

Moreover, there are various issues listed as cross layers issues across all areas which will be discussed later in this section.

Figure 4.4 represents the different dimension related to ad-hoc networking in which several wireless networks plotted to two self-governing expressions of ad-hoc networking: horizontally: the level of centralized control, and vertical: the use of data lines- multi-hopping.
Figure 4.3.  A simple MANET architecture [6]

Figure 4.4 Wireless networks mapping [6]
4.5 Mobile Ad-hoc Network Security

Wireless system functions with a centralized sustaining structure such as access points. Recent development of wireless technologies IEEE 802.11[9], Bluetooth [10] have introduced another kind of wireless system known as (MANETs) [11, 12, 13, 14], which operates without the support of central access point. It provides high mobility and device portability that makes it possible for the devices to communicate and establish connections in the network. It allows the devices to maintain the connections to the network and simply add and remove device or in the network. User has huge flexibility to design a network in least time and at low cost.

In most of MANETs specifications it was supposed that all the nodes in the network were friendly so the routing protocols for Mobile Ad-hoc Networks were originally designed without having any security in mind. The security issue was deferred and altered on it would be feasible to make these routing protocols secure by pre-existing retro-fitting cryptosystems. However, it is confirmed in this case that securing the network transmissions without securing the routing protocols is not adequate. Furthermore, by retrofitting cryptosystems (i.e. IPSec) security is not essentially achieved. Therefore, in MANET networks with security requirements, there must be two security systems: one to make the routing protocol secure and one to protect the data transmission. There has already been enough consideration on point to point security systems that would be used for protecting network transmissions. But there was no much work done on how to make the routing protocols discover the routes in a secure manner.

MANETs is not free from different active and passive attacks [15]. Because of the lack of resource constraints and central authority it is much more vulnerable. The attacks are classified into two different types, called internal attacks and external attacks depending upon the malicious node location and depending upon the operations it is also categorized into two types, which are active attacks and passive attacks [16].

MANET resources are mostly used in military and rescue operations even though mobile phones are using Bluetooth. As it is a topology, less kind of network security has become one of its main issues and different challenges faced by MANETs are [17] such as Confidentiality, Integrity, Authentication, Non-Repudiation.
Figure 4.5 Emerging protocol and approaches for authentication and access control. MANET [6]

Figure 4.6 Establish public key infrastructure authentication and access control protocols Fixed Networks [6]
Table 4.1: MANETs security issues on different layers

<table>
<thead>
<tr>
<th>Layer</th>
<th>Security issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application layer</td>
<td>Detecting and preventing viruses, worms, malicious codes and application abuses</td>
</tr>
<tr>
<td>Transport layer</td>
<td>Authenticating and securing end-to-end communications through data encryption</td>
</tr>
<tr>
<td>Network layer</td>
<td>Protecting the ad hoc routing and forwarding protocols</td>
</tr>
<tr>
<td>Link layer</td>
<td>Protecting the wireless MAC protocol and providing link-layer security support</td>
</tr>
<tr>
<td>Physical layer</td>
<td>Preventing signal jamming denial-of-service attacks.</td>
</tr>
</tbody>
</table>

4.5.1 Vulnerabilities of the MANETs

As nodes in mobile network can move without any restriction; therefore, the network topology is adjusted very often. This nature of the mobile nodes network might create security and other issues in MANETs [18] such as:

Lack of Centralized Management Facility: It would be not easy to monitor the network traffic and make the detection of attacks in a highly dynamic and large scale ad hoc network in the absence of centralized management architecture [3].

Restricted Power Supply: A mobile node in the ad-hoc network acts in a self-interested way as it has very limited energy resources, and the self-interest can create few issues specifically regarding the co-operation behaviors is expected with other nodes to facilitate the core functions of network. Just take an example of cluster-based intrusion detection technique [19], a cluster of adjacent MANET nodes can fairly and randomly select the monitoring node that will watch the malicious behaviors in the network traffic for the entire cluster so there would be no need that every node in the network performs as monitoring node.

Scalability: It is difficult to forecast the number of devices that will be part of the network in future, because of the mobility of the nodes in the mobile ad-hoc network. Therefore, the services and protocol that are applied to the network such as key management service
and routing protocol should be compatible to the continuously changing scale of the ad hoc network.

Threats from Compromised Nodes: Since mobile nodes are self directed unit of network that can leave or join the network anytime, so it is difficult for the mobile nodes to manage effective policies to prevent the malicious attacks from all the devices it connected with.

Lack of Secure Boundaries: The mobile ad-hoc network go through from all kind of weather attacks, which can come at any time, from any node that is in the radio range, and can target to any other node in the network.

### 4.5.2 Security Attacks

There are various types of attacks on ad-hoc network which are as follows [20]:

**Location Disclosure:** Location disclosure attack aims at the privacy requirements of an ad-hoc network by using the traffic analysis techniques [21], by using simple inquiring and monitoring techniques, an attacker is able to realize the structure of the entire network or even the location of a node.

**Black Hole:** In a black hole attack a malicious node advertising itself as having the shortest path to a destination and infuse false route respond to the route requests it receives [22]. These bogus responses can be fabricated to switch the network traffic via malicious node for eavesdropping.

**Replay:** This attack normally targets the novelty of routes, but can be used to damage the poorly adopted security solutions by replaying, attack injects into the network routing traffic that has been encapsulated recently.

**Wormhole:** The wormhole attack is the most influential offered as it implies the co-operation between two malicious mobile nodes that are involved in the network [23]. The wormhole link connectivity of the nodes that have established routes is completely under the direction of the two scheme attackers and the solution to this attack is packet leashes.

**Blackmail:** The routing protocol that used the method to identify the malicious mobile nodes and broadcast messages is targeted by this attack to blacklist the offender [24]. An attacker may formulate such messages and make an attempt to isolate valid nodes from
the network. The security feature of ‘non-repudiation’ can be helpful as it attaches a
device to the messages it created.

Denied of Service: Denial of service attacks aims at the whole disruption of the routing
function and consequently the entire operation of the ad hoc network [25]. Specific case
of denial of service attacks, node floods the network with fake route packets in order to
utilize the resources of the connecting nodes and interrupt the formation of legitimate
routes.

Routing Table Poisoning: Routing protocols maintain record that keep the information
related to the routes of the network. In these attacks, the malicious mobile nodes create
and send fictitious traffic, or modify valid data from other nodes, for the purpose to make
wrong recording in the routing tables of the connecting nodes [25].

Rushing Attack: Rushing attack comes in results of denial-of-service when using against
all on-demand routing protocols [26] such as AODV, DSR, and secure protocols such as
SAODV, ARAN, and Ariadne, are not be able to discover the routes that are longer than
two hops when affected by this attack. A basic defense against the rushing attack Rushing
Attack Prevention (RAP) has developed for on-demand protocols that can be applied to
any on-demand routing protocol.

Breaking the neighbor relationship: An intelligent filter or sorter is positioned on a
communication link between two information system by an intruder that could change or
modify the information in the routing updates or may interrupt traffic belonging to any
data session.

Passive Listening and traffic analysis: In such an attack, the operation of routing protocol
cannot be effected, but it breaks the trust of user to routing the protocol. Intruder could
passively collect the exposed routing information though sensitive routing information
should be protected.

4.6 Security scheme in MANET

There are two types of well-known security techniques in the mobile ad hoc
network, which are intrusion detection techniques and secure routing techniques. In each
of the security schemes, numerous specific methods are pointed out and compared with
each other. There are some points that some of the methods have deficiencies based on our observations.

4.6.1 Intrusion Detection Techniques

Due to the frequently changing topology and limited battery power, the intrusion detection system in the mobile ad hoc networks should be energy-efficient and cooperative, which are shown in the two papers written by Huang et al., and Zhang et al. respectively [19,27]. Due to the mobility of the nodes and the continuously changing topology in the ad hoc network, it is sometimes probably difficult to gather the sufficient evidences for a node if it is based on the single layer detection method. As a result, the concept of cross-layer or multi-layer detection mechanism is developed in [27] and [28].

4.6.2 Secure Routing Techniques in Mobile Ad Hoc Network

Several secure routing techniques have been designed that can ensure the ad-hoc routing security. There are various kinds of attacks against the routing layer in the mobile ad hoc networks, some of which are more complicated and harder to detect, such as Rush attacks and Wormhole attacks [29,30]. Watchdog and Pathrater, are two main mechanism in a system that aim to reduce the routing mis-behaviors which effects in mobile ad hoc networks [31]. Finally we move to a secure ad-hoc routing approach which is using localized self-healing communities [32].

Defense Mechanism against Rushing Attacks in Mobile Ad Hoc Networks

Rushing attack is a comparatively new attack that appears in denial-of service which is used against all previous on-demand routing protocols [33]. Protocol limits the number of packets that a node will transmit in reply to a single Route Discovery. The implementation details of rushing attacks are shown in the Figure 4.7.
A set of general mechanisms that makes a joint defense against the rushing attack are: secure route delegation, secure Neighbor Detection, and randomized ROUTE REQUEST forwarding. The relations among these security mechanisms are shown in Figure 4.8 below.

Secure Neighbor Detection permits each neighbor to verify first that the other node is within a defined transmission range. Suppose a node A forwarding a ‘ROUTE REQUEST’ decided that node B is a neighbor, it marks a ‘Route Delegation’ message, permitting node B to forward the ROUTE REQUEST. When node B clarifies that node A is within the permissible range, it marks an Accept Delegation message. By this procedure, the neighborhood relationships between nodes verified has been guaranteed and verified to be original.
An Efficient Security Aware Routing Protocol for Mobile Ad Hoc Networks: Mobile Ad-Hoc Networks (MANETs) uses security routing system for setting up secure route and determinedly data transmission. The major security hazard in MANET is non-repudiation, integrity, and privacy. In the contest of these security threats, several secure routing protocols have been designed to overcome the security threats in MANET. The research [34] has proposed ESARP - Privacy Aware Routing Protocol to improve security in the routing protocol to prevent the network attacks. The proposed protocol consists of three parts.

- At first, each node performs a key exchange operation with its neighbors residing in one and two hop distance. And this key exchange operation further takes two steps, in the first step, source node (S) exchanges public key(e) with its one hop distance nodes and source node exchanges public key with its two hop distance nodes and establishes a secret key(SK) in the second step.,

- In the second stage, establishes secure route and

- In the last step, secure data communication is performed.

### 4.7 Security Solutions to the Mobile Ad Hoc Networks

Several vulnerabilities that potentially make the mobile ad-hoc networks unprotected have been illustrated in the last section. Attacks encounters, such as encryption and authentication, would used as defense to decrease the risk of attacks. Major security researches have been done in MANET to date, e.g., [35, 36-39], are based on attack prevention techniques. For instance, the session secret key mechanism has been adopted to encrypt messages towards making sure the confidentiality, and to some extents to the authenticity of data, group membership, and routing information. The public key mechanism can be used to encrypt and sign the messages to make sure confidentiality, authenticity of nodes, and non-repudiation communications. These prevention schemes suggested so far vary in many ways, depending on the requirements on the MANET applications.
4.7.1 Preventing External Attacks : Key and Trust Management

Authentication, encryption, and key management are generally used to prevent external attacks. However in case of ad-hoc networks they face many challenges [40]. Firstly, it is necessary to manage the dynamic network topologies, in trust relationship and in communications as well. Secondly, consider the infrastructure-less network of MANET; it would be difficult to deploy the centralized security scheme. Key management system consists of various services which provide solutions must be able to reply on the following queries:

4.7.1.1 Trust model

The trust relationships between the components of network really affect the way the key management system has built in network. It is mandatory to resolute that how much different components can trust each other in the network. The area of application and environment of the network significantly affects the requested trust model.

4.7.1.2 Cryptosystems

Cryptosystem is offered for key management: according to the requirements, in some situations symmetric key (public key) mechanisms can be useful, while in some other scenarios Elliptic Curve Cryptosystems (ECC) are useful. Although symmetric-key cryptography presents more convenience as using the digital signature schemes but public key cryptosystems are considerably gradual at their secret key counterparts when same level of security is required. Comparatively, a newer field of cryptography in term of implementation - ECC cryptosystems is already in use extensively, for case in point smart card systems.

4.7.1.3 Key Creation and Key Storage

At first it is necessary to determine that which parties will have to generate keys and what kind of key structure is required from other parties. Secondly there may not be possible to replicated key storage for fault tolerance in ad-hoc networks neither centralized storage for keys. In mobile ad-hoc networks each element of the network is
responsible to store its own key and might be the keys of other elements as well. Moreover, the concept of shared secrets key is proposed [41] to distribute the parts of key to several of nodes. In such case the compromising of a mobile node does not compromise with the whole secret keys.

4.7.1.4 Key Distribution

The key management service should make sure that the generated keys are securely distributed to respective node so that authenticity, confidentiality, and integrity are not spoiled. In public key cryptography scheme the key distribution mechanism is obliged to assure that the private keys are delivered to only authorized users. The distribution of public keys needs to preserve the integrity and authenticity of the keys but not the confidentiality.

4.7.2 Preventing Internal Attacks: Secure Routing Protocols

An appropriate routing protocol in Mobile Ad-Hoc networks is essential to create a secure route accurately and manage it for transportation of data. Generally, if, in a protocol, the manipulated data signal cannot be injected into the network, the points such that routing signals do not fabricate, routing messages do not change while transporting, the shortest routes do not change by adverse nodes, routing loops do not create during hostile activities, and so on are considered, it can be called a secure protocol [42]. To observe these points, we are reviewing several protocols.

**Table 4.2: Solutions to Secure Routing Protocols**

<table>
<thead>
<tr>
<th>Solutions</th>
<th>Attacks Prevented</th>
<th>Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentication during all phases</td>
<td>All external attacks, and the following internal attacks, Spoofing, Redirect by modifying route sequence number</td>
<td>Requires certificate authority or key sharing mechanism</td>
</tr>
<tr>
<td>Trust-level metric</td>
<td>All attacks prevented by authentication, All attacks on higher trust-level nodes</td>
<td>Requires certificate authority or key sharing mechanism, Difficulty to define trust level</td>
</tr>
<tr>
<td>Secure neighbor verification</td>
<td>All attacks prevented by authentication, Rushing</td>
<td>Requires certificate authority or key sharing mechanism, Important overheads when mobility increases</td>
</tr>
<tr>
<td>Randomize message forwarding</td>
<td>Rushing</td>
<td>Latency</td>
</tr>
<tr>
<td>Onion encryption</td>
<td>All external attacks, and the following internal attacks, Spoofing, DoS by modifying source route</td>
<td>Requires certificate authority or key sharing mechanism, High computational cost</td>
</tr>
</tbody>
</table>
4.7.2.1 DSR (Dynamic Source Routing)

In DSR protocol, the source node constructs a package identified as RREQ in which it is defined the source and destination node and these packages via flooding [24], node in the network receives RREQ package, if the node doesn't know about destination route, then, it includes its name to the package list and broadcast it again. So, as the package reaches to the destination, it includes the data of all route nodes and its planning will be presented for the destination node. The destination node creates RREP and returns it back via list available in RREQ package header. The middle nodes know the destination and follow it according to the prepared list. So, the package travels on the route opposite to reach the source node. Even though, DSR is certainly applicable method but uses high bandwidth and adds to the network load. As growing the rate of header volumes consequently greater the distance between links this approach may not work appropriately.

4.7.2.2 AODV (Advanced On-demand Distance Vector)

As compared to DSR protocol, AODV protocol does not add the route in to the package header. However, each node manages it on receiving PREQ according to tables it already had. If final node in the rout has received its table, RREP will be sent back. Else, it will broadcast the RREQ message.

4.7.2.3 SAODV (Secures AODV)

SAODV provided to create more security in AODV [43]. It used Hash functions as it is shown in equation (4.1)

\[ h^{n-1} = H(h^n) \]  \hspace{1cm} (4.1)

H is the function of Hash in the equation (1) and h referring to the hop. In this protocol, it is used to measure the number of hops – hop count in which the packages will travel. If the hop count gets increase than the number of Max Count, the package will be discarded. Because of this hash function equation, each node can be definite about the
authenticity of message. In the equation \( n \) defines the maximum hop from that a package can travel.

4.7.2.4 OLSR

*Optimized Link State Routing protocol (OLSR)* [40, 41], is a table driven and proactive protocol that applies *multi-point relays* (MPR) and multi-tiered approach. MPRs allow the network to apply scoped flooding; resultantly the amount of control data exchanged can significantly be minimized. Since the MPR approach and OLSR is most suitable for dense ad-hoc and large networks kind of environment in which the network traffic is sporadic and random.

4.8 Routing Protocol in MANETs

In recent years many routing protocols have been introduced for MANETs. These protocols are divided into three major categories – Reactive – Proactive – Hybrid. Several protocols and algorithmic approaches have been developed and implemented so far to avoid and remove the issues associated.

- Flat Routing Protocols
  - Proactive Routing - Table-Driven
  - Reactive Routing - On-Demand

- Hierarchical - Zone/Cluster-Based Routing Protocols
- Geographic Position Assisted Routing Protocols
**Figure 4.9 Classification of MANET routing protocols [44]**

*Proactive Approach:* This technique is used by the internet routing algorithms such as OSPF, RIP, and IS-IS is also called a *table-driven approach*. In these algorithms, the routers maintain the routing information to every node in the network constantly and up-to-date. These routing tables are modified or upgraded every time as the network topology changes.

The following are examples of proactive ad-hoc routing protocols [44]:

- Cluster head gateway switch routing;
- Destination-sequenced distance vector routing;
- Wireless routing.

*Reactive techniques:* These techniques maintaining the alterations of topology structure is on-demand information in every router. When there is not a suitable route to take out this delivery on new packet requires to be delivered, a new route is discovered.

The following are the examples of reactive techniques.

- Flooding;
- Dynamic source routing (DSR);
- Ad-hoc on-demand distance vector routing (AODV);
- Associativity-based routing;
- Signal stability routing.

On the other hand, the cost of maintaining it up-to-date all the time may be lower than the cost of route discovery as it is needed. This depends on the topology change rates and the traffic generation. For advance wireless ad-hoc network applications, reactive techniques such as DSR and AODV are preferred.

Table 4.3: Properties of proactive and Reactive Routing

<table>
<thead>
<tr>
<th>Proactive Protocols</th>
<th>Reactive Protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete knowledge of the topology</td>
<td>Incomplete knowledge of network topology</td>
</tr>
<tr>
<td>Routing table with a route to all destinations</td>
<td>Routes maintained only between nodes that need to communicate</td>
</tr>
<tr>
<td>Exchange of full routing table for convergence</td>
<td>Store routes already traversed</td>
</tr>
<tr>
<td>No on-demand</td>
<td>On-Demand basis</td>
</tr>
</tbody>
</table>

4.9 Clustering in Mobile Ad hoc Network

The procedure that breaks up the network into interlinked sub-structures is called clusters. Each cluster has a specific node selected as cluster head (CH) based on a certain metric or a group of metrics such as degree, identity, weight, mobility, density, etc. The CH plays the role of co-coordinator within its sub-structure. Each CH represents as a transient base station within its cluster and exchanges with other CHs [17, 18]. A cluster is therefore composed of member’s node, cluster head, and gateways.

*Gateway:* is a common node between two or more clusters.

*Cluster Head (CH):* it is the coordinator of the cluster.

*Member Node:* Each node exclusively belongs to single cluster or might reside in a different cluster independently of its neighbors.
4.9.1. Algorithms for Cluster Heads Election in MANETs

There is a number of algorithms in the literature for cluster heads election in mobile ad hoc networks: Weighted Clustering Algorithm [30], Highest-Degree [28], Lowest-ID [27], Distributed Clustering Algorithm [29] and Distributed Weighted Clustering Algorithm [32]. We categorized the clustering algorithms based on their objectives, the cluster heads selection criteria and literature review [12, 45-47] is based on as follows:

4.9.2. Clustering Schemes in Mobile Ad hoc Network

The following are the major schemes for clustering the Mobile Ad-hoc networks

4.9.2.1. Identifier Neighbor Based Clustering

A unique ID is assigned to each node in identifier neighbor based clustering. Each node in the network keep the ID information of its neighbors. The cluster head is selected based on some measuring criteria relating these IDs such as the lowest ID, highest ID...etc.

*Linked Cluster Algorithm-LCA*: Ephremides et al [27] proposed a clustering algorithm called Linked Cluster Algorithm (LCA).

*Least Cluster Change (LCC)*: Chiang [20] proposed Least Cluster Change (LCC), an enhanced version of LCA algorithm which includes a maintenance activity to reduce the cost of re-clustering.

*Adaptive Clustering Algorithm*: Lin and Gerla [21] proposed this protocol called Adaptive Clustering Algorithm (ACA). In which, once the clusters are created, the idea of cluster head vanishes and all nodes play the same role in the net-work.

*Max-Min D-cluster*: A heuristic based algorithm [10] called Max-Min D-cluster builds non-overlapping D-clusters. The node ID is used for CH election.

4.9.2.2. Topology Based Clustering

In the topology based clustering, the cluster head is selected based on a metric (i.e node connectivity) calculated from the network topology.
**High-Connectivity Clustering:** Gerla and Tsai proposed a protocol called High-Connectivity Clustering (HCC) [28] based on the degree of connectivity to the formulated clusters.

**(3hBAC):** Yu and Chong proposed [24] 3-hop between closest CHs (3hBAC) which constructs three hops between Adjacent CHs by the initiation of a new node status, called guest with a 1-hop non-overlapping clusters structure.

\[\alpha\text{-SSCA}\] Guizani et al [25] proposed a new clustering algorithm named \(\alpha\)-Stability Structure Clustering (\(\alpha\)-SSCA). This algorithm increases moderately the number of clusters in the aim of reducing the overheads and improving clusters stability.

**Associativity-based Cluster:** Associativity based Cluster Formation and Cluster Management [26] use the relative stability of nodes in their neighborhood metric. This algorithm creates overlapping that stays stable over a long period of time.

### 4.9.2.3. Mobility Based Clustering

**MOBIC:** Lowest Relative Mobility Clustering Algorithm (MOBIC) [42] is based on the LCA algorithm using the relative mobility of nodes as a selection criteria for CH.

**MPBC:** Ni et al planned a mobility prediction based clustering (MPBC) scheme [41] for MANETs with high mobility nodes.

**MobDHop:** The purpose of creating D-hop clusters is to keep up more than one-hop radius clusters which decrease the number of CHs. Mobility based D-hop clustering algorithm (MobDHop) [35] distributes the network into D-hop clusters based on relative mobility metric.

### 4.9.2.4. Energy based Clustering

The battery power of node is a control that directly affects the life span of the network, therefore the energy constraint cause a serious issue on behalf of network performance. Cluster head execute some particular activities such as routing which causing extreme energy utilization.

A _multicast power greedy clustering_ (MPGC) [13] is based on heuristic to reduce the energy consumption.
A Flexible Weighted Clustering Algorithm based on Battery Power: (FWCABP) for MANETs [48] is projected to keep clusters stable by preventing nodes having low power from being elected as a CH.


4.9.2.5. Weight based Clustering

Weight based clustering techniques use a multiple or group of weighted metrics such as: node degree, transmission power, mobility, distance difference and battery power of mobile nodes… etc. Some of these algorithms are:

Flexible Weight Based Clustering Algorithm: A Flexible Weight Based Clustering Algorithm (FWCA) uses a combination of metrics to build clusters.

Score based clustering algorithm: Adabi et al proposed Score based clustering algorithm (Sbca) [39] for MANETs which intended to maximize lifespan of mobile nodes and minimize the number of clusters.

efficient weight-based clustering algorithm: An efficient weight-based clustering algorithm (EW-BCA) for MANETs is proposed [49] to enhance the usage of limited resources such as energy and bandwidth by minimizing routing overhead, producing stable clusters, and escalating end-to-end throughput.

4.10 Chapter Conclusion

The fast progression in the subject of mobile computing is motivating a new substitute for mobile ad-hoc communication, in which devices form a self-organizing, creating and administering wireless network, called a MANETs. Its inherit characteristics such as infrastructure-less network, simplicity of deployment, self-configuration, low cost and promising applications make it a necessary element of the evolution with regard to future generation communication networks. The short arguments in this chapter explain that, regardless the huge work of the MANET researcher and the swift development made
all through the last years, a number of challenging technical questions remains unresolved.

We have presented that the nature of MANETs has inherent vulnerabilities which cannot be resolved. Seemingly, several security attacks that take advantage of these vulnerabilities have been discussed. New attacks will appear in the future, particularly when MANETs will be used widely. Mitigation against these threats can be attained by implementing key management and secure routing protocols. This is an imperative and yet has large number of research opportunities with several open questions for technical developments.

References


[2] Imrich Chlamtac, Marco Conti, Jennifer J.-N. Liu, “Mobile ad hoc networking: imperatives and challenges”, School of Engineering, University of Texas at Dallas, Dallas, TX, USA, Istituto IIT, Consiglio Nazionale delle Ricerche, Pisa, Italy, Department of Computer Science, University of Texas at Dallas, Dallas, TX, USA, 2003.


[6] Brent Ishibashi, Raouf Boutaba, “Topology and mobility considerations in mobile ad hoc networks”, School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada, 2004


[21] A REVIEW PAPER ON AD HOC NETWORK SECURITY


[40] Security in Ad Hoc Networks, Vesa Kärpijoki, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, Vesa.Karpjoki@hut.fi


[44] Haas and Liang, 1999; Royer and Toh, 1999


[47] Haas and Liang, 1999; Royer and Toh, 1999


Chapter 5
Semantic Web Model and Security Protocol

5.1 Introduction

The Semantic Web technologies have been growing at a prompt step since their origins. The evolution of information from human to machine using up is composed to be as reflective as the preliminary impetus of the commercial and cultural acceptance of the Internet. This chapter presents an overview of semantic web technology, key security issues, encryption, protocol and security standards, some explanatory applications and supporting instances such as Decision Support System, online encyclopedia, Watson system, data and information integration, knowledge extraction on semantic web, knowledge soup and boundary, NetKAT.

The study begins with section 5.2 an introductory note on semantic web and its techniques. In section 5.3 we present the component architecture of semantic web which are standard languages and protocols sets by W3C and could maintain the semantic of the information representation i.e. RIF - Rule Interchange Format, Web Ontology Language OWL, Resource Description Framework RDF and XML.

In section 5.4 we discuss some security standards and schemes adopted for secure transportation of data such as HTTPS, identification mechanisms (OpenID, FOAF with SSL). Further, a five layer secure semantic web model ad semantic technologies presents the security status, respective roles and challenges with their shortcomings. Web technology standards can be used for data integration, sharing, and reusing on the bases of knowledge. Though, in respect to direct this idea for practical implementation, many challenges or hazards require to be reduced which are discussed in section 5.5.

Decision Support Systems undertake many goals of semantic web are presented in section 5.6. DSS takes semantic web towards realization of applications evolving out of business intelligence, as it uses collaboration support through social networking, retrieving information towards knowledge management, to situation awareness, managing emergencies and simulations. In section 5.7 we present the architecture and procedure of interlinking the document online encyclopedia Wikipedia. A concise note
on web 2.0 and semantic web technology has been provided in section 5.8. Moreover to this interlinking section 5.9 represents the relation between linguistic knowledge and formal language which holds for knowledge extraction on the semantic web and once information is extracted next step to data and information integration is discussed in section 5.10.

Watson is a semantic web engine (search tool) used to locate ontologies and semantic data as well as explores the content in these searched documents. Section 5.11 represents an overview of Watson architecture; it is not only a tool but a unique resource for researchers to create the data and formulize the knowledge.

At the end in section 5.12 the most recent mathematical programming language NetKAT for establishing semantic foundation is elaborated.

5.2 Semantic Web Technology

The Semantic Web, as an inter-disciplinary study ground, come out of the aspiration to expand the WWW in such a manner that combination of multi-authored, multi-perspective information, multi-thematic, services and interoperability could be picked up on-the-fly and flawlessly [1, 2]. Therefore, the Semantic Web is related with the techniques that have developed for information annotation, reusing, extraction, formations, and mixing. It is also illustrated in order which includes reasoning in knowledge management, databases, software engineering, artificial intelligence, knowledge representation, machine learning, and natural language processing and information visualization [3, 4].

Semantic Web Technologies are being researched now-a-days in many areas where information integration and reuses on internet added important value. On the other hand, Semantic Technologies also support in the field of application development solutions which are not originally intended to the WWW, but have to face interrelated issues, such as data analysis, intelligence information incorporation and expert and decision support systems.
5.3 Languages Represent Knowledge on the Web

Ontology languages are the recommended standard sets by the WWW Consortium (W3C) and are formal languages which could maintain the semantic of information representation; explicitly the Rule Interchange Format RIF [5, 6], the Web Ontology Language OWL [7, 8] and the Resource Description Framework RDF [9]. Although syntactically, these standards relies on XML [10] but the primary data models are extensively different.

5.3.1 Web Semantic Protocol

The major focuses for the development of the Semantic Web architecture are protocols and languages, whereas in order to be compatible with the web technology HTTP was the most proficient as known for the Semantic protocol, but where the language part is being considered, the key point is the flexibility of developing and creating a Web of data [11].

Uniform Resource Identifiers – URI is defined to identify the internet resources [12] and the data publishers use some methods to provide explicit reference due to decentralized nature of the Semantic Web. URIs on the Semantic Web generally use HTTP based identification, which provides facilitation on the DNS that would make sure its uniqueness world widely.

Figure 5.1: Semantic Web Language and Protocol Standards [11]
5.3.2 Software Component

Semantic Web architecture could be analyzed in different ways and one of them is related with standards used by languages. And another possible aspect of analysis could be; the various components of software applications that implement its functionality based on these languages and standards. Figure 5.2 shows the classical Semantic Components Web model while applying the functionality in applications.

![Semantic Web Components Architecture](image)

**Figure 5.2:** Semantic Web Components Architecture [12]

The fundamental components of infrastructure are related to de-referencing content and publishing RDF. As per convention, URIs provide unique identifiers and probably applied to recover the content and web server assists the communication for provision of contents by HTTP and its clients facilitate to search the functionality. The Semantic Web acquires the fundamental infrastructure for providing the reference and supporting search from the internet. Material replied or returned on HTTP query is processed either using RDF parsers and APIs or XML parsers and APIs. There is no
existing standardized design for proceeding RDF content and at the same time W3C construct a requirement to manage the XML documents.

A key point to communicate sensitive and critical data is to make sure that data transmissions would have not been interrupted, modified or read. Crypto implementation enforces encryption and authentication techniques to secure the data transmission such as SSL processors offer cryptographic authentication and validate digital certificates.

An integration and alignment feature of architecture merges and provides tighter integration of data contents sourced from multiple sites and use several identifiers to point to the same real-world object. At last, the user Interfaces permit to communicate with the data of Semantic Web depending on its functionality; these user interfaces are common such as graph structure of the data or might be customizable on requirements.

5.4 Security and Encryption

In open systems for instance Internet, the methods for secure transportation of data should have to be acknowledged, where data is transmitted on un-secured links connecting infrastructure managed and upheld by a group of corporation. Moreover, to verify the identity of users, authentication system is required along with the technology which makes sure the authenticity of documents by digital signatures. These schemes are described in the crypto module of the Semantic Web architecture.

A secure version of Hyper Text Transfer Protocol HTTPS [13] has been developed to ensure that the data is not changed during its transmission, a secure version of HTTP. Data is not changed, could be achieved by using encryption protocol to deal with the man-in-the-middle attacks and by using various ports on server measuring up to HTTP and eavesdropping of links. Digital signing of RDF graphs ensures that the relevant material has been formed by a recognized or authenticated source [14].

There are two similar mechanisms used to establish the identification of a user while logging into a web service or web site.

- With OpenID [15], a user is authenticated by the identity provider which has been forwarded by the service provider and identity provider returns the credentials of the user.
FOAF and SSL [16] with the contrast of OpenID is a distributed method on the support of digital certificates, in which RDF permits for authentication by signed FOAF information and published profiles.

5.4.1 Security Standards for the Semantic Web

Tim B. Lee [17] has specified different layers for the semantic web see figure 5.3.

1. The lowest layer security protocol requires secure sockets, secure TCP/IP, and secure HTTP. Secure TCP/IP based on un-trusted communication tire and required entire network security.

2. Subsequent layer is XML security and XML secure schemas. Access controlled must have to be applied at different components of the document for reading, browsing, and modifications.

3. The next layer RDF security which requires security for the interpretations and semantics.

4. So that XML and RDF have been secured once, further is to investigate the security for interoperation and ontologies. The challenge at this stage is as to how to secure this secure information integration using ontologies?

5. Security is also closely associated with privacy as some part of the document may be public or private whereas some parts might be semi-private.

Figure 5.3 Layers for Secure Semantic Web [17]
5.4.2 Semantic Web Security Foundations

The Semantic Web basically provides foundation and assists the sharing of information but most recent Semantic Web technologies such as SPARQL, OWL, RDF are not capable enough to provide satisfactory security to the information being shared [18]. The present status of security related to XML and semantic technologies is as:

eXtensible Markup Language (XML): XML security has been explored by various including [19] and the W3C has defined suggestions and standards for XML Security. XML is considered as a main building block for the Semantic Web and enforce sufficient security in order to protect other upper layers.

Resource Description Framework (RDF): Various researches have addressed RDF security [20, 21] but some issues remain unresolved.

Web Ontology Language (OWL): As with RDF, few research works relevant to security and ontologies have been performed [21] but some questions remain open. OWL is basically a semantic model consists of rules and vocabulary about relationships among objects.

Security policies: Policies are written in a natural language but can be defined using RDF, XML or OWL. Policies related Research work in Semantic Web have been carried out by Bonatti [22]. In Semantic Web perspective policies are rules relating to access control, privacy, and trust.

Query Processing: Previous research [23] has extended the RDF and SPARQL to integrate the trust by introducing new query engine. In a recent research [24] an idea proposed that using ‘views’ in SPARQL is same as access control used in SQL.

Secure Semantic Web Services: The OASIS and W3C have proposed standards related to Secure Semantic Web Services.

Access Control: Access control is a standard security control which provides protection of information by restricting access to authorize users only. Access control applies to XML documents, RDF stores, rules, queries, and ontologies; it can also be used to restrict access to a resource that has been researched by Gabillon [24].

Inference: Reduce information by aggregating other information. Preventing inference has been verified to be impossible ([15]) so reducing the likelihood of inference is the
best we can hope to accomplish. Previous research related to the inference problem has
been conducted by Farkas [21].

Trust: Research related to trust has been performed by Bertino [25]. In which relations to
the Semantic Web can apply to data, agents, or individuals.

Provenance: Provenance is related to trust and represents the history related with
information such as who created the information, who has accessed the information, who
has modified the information, etc. because information may become more or less
trustworthy depends upon who created or modified it.

5.4.3 Injection Attack in Semantic Query Language

Code injection [26] attacks are known security threats. Although other query
languages also get affected from this exposure perhaps this is the most well known attack
based on code injection in SQL injection, but other query languages also get effected
from this exposure.

In the semantic web the former vulnerabilities re-appear with new semantic
mechanism. This semantic mechanism arrives with new risky prospective to malicious
user. It is important for developer to keep in knowledge and practice and prevent the
vulnerabilities in their developments. The semantic query language (SPARUL, SPARQL)
has been found vulnerable to various attacks on the bases of invalid user inputs [27].
These vulnerabilities come out when the inputs are directly embedded into query;
allowing attackers to get the control over query, therefore unexpected results and output
are produced by the application.

5.5 Key Issues in Recognizing the Semantic Web Vision

Web technology standards can be used for data integration, sharing, and reusing on the
bases of knowledge. Though, in respect to direct this idea for practical implementation,
many challenges or hazards require to be reduced. Here is the list of some most important
ones:
Different Web sites may use different URIs for an article to provide its identification on the Semantic Web.

Different sites may use different ontologies in background. These ontologies need to be formally related to each other use by fitting appropriate formats, recommended by ontologies.

The data on the internet may be semantically diverse. For instance, different sources may report different dates for historical time duration of an event or may not agree with the fact.

Who is providing information on the internet with meta-data. Can it be automated totally or partially by using the data mining or machine learning processes? What are the tools recommended for domain experts and developers to create the metadata?

The creation of ontologies normally demanding expertise in both application domain and ontology modeling. However, their repairs can be very costly and the formation of workflows and methods to base ontology modeling is of the most importance for the development of their applications.

5.6 Web Semantic for Decision Support Systems (DSS)

Decision Support Systems undertake many goals of semantic web [28] to achieve the DSS, as it precisely interprets information to deliver reliable, relevant and right information to the peer at the right place at right time. DSS further uses this ability to making particular decisions based on the information received like, planning to react to a certain situation. Thus, DSS takes semantic web towards realization of applications evolving out of business intelligence, as it uses collaboration support through social networking, retrieving information towards knowledge management, to situation awareness, managing emergencies and simulations. Information management is the key to unify all these systems and not just limiting to specific technologies adopted to reach these goals. Semantic web technologies are adopted by DSS [28] in the past to find solution to number of problems regarding information integration and sharing,
representing knowledge for reasoning purpose, and web service annotation and discovery.

A hazard to the acceptance of Semantic Web technologies is its scalability and immaturity with respect to optimization and efficiency, is in contrast to typical data management solutions, but not all DSS require efficiency. However, Semantic Web technologies helps to provide basic DSS solution such as information linking, integration, and interoperability.

Once these obstructions have been removed, Semantic Web can exclusively include to BI. As noted by [29], BI is an area focused by many IT based companies where so much of money is spent, made this an interesting focus also for Semantic Web-related research.

As seen in Figure 5.4 the application domains of Semantic Web-supported DSS where Intelligent DSS has been one of the main categories of SemanticWeb-supported DSS throughout the years. An explanation for this is also the fact that DSS and Semantic Web share a common ancestry in early AI technologies, which are still acknowledged as a part of DSS legacy for data analysis.
In Figure 5.5 we note that most articles describe applied research, rather than theoretical work, while there is also a small number of survey articles and position papers in our collection.

![Study Report]

**Figure. 5.5.** Type of study reported [28]

When studying the contribution of Semantic Web technologies to the DSS field, as illustrated by Figure 5.6, we note that the two high-impact areas have been ontologies and semantics and Semantic Web data, with particular emphasis on the former. Many DSS applications use ontologies and rules as a means for making the DSS “intelligent” in some data analytics sense.
Semantic Web data can be used for DSS in different formats such as OWL and RDF which allows access and integrates the data from existing data sources [30-35], e.g., Data Warehousing sort of methodologies to integrate the database, but with modern and creative formats, whereas other key point is utilizing Semantic Web Data in RDF, or extracting Web Data [36] as an entirely advance data source e.g., by adding Linked Data (LD) in their DSS application, or advising to shift from existing data broadcasting principles to LD [37].

![Figure. 5.6. Number of publications reporting the use of specific Semantic Web technologies [28]](image)

**5.7 Online Encyclopedia Wikipedia**

Interlinking articles is constantly evolving within online encyclopedia Wikipedia [38] it forms a massive knowledge base for researchers and developers through multiple language databases of ideas, themes, and relations, forming a powerful source for researcher and natural language processing. Wikipedia miner toolkit is an open-source software system allowing developers and researchers to make the wikipedia’s rich semantics more relevant. Databases are created using this toolkit [38] having summarized
content in Wikipedia that includes java api to provide access. Articles within Wikipedia are categorized and directed representing classes that could be searched and iterated efficiently. Wikipedia dumps are parallelized processed, XML-based web services, and machine-based semantics relatedness measures and annotation measures. Wikipedia miner can be used further for sharing data mining measures.

![Figure 5.7 Wikipedia Miner toolkit Architecture](image)

**Figure 5.7** Wikipedia Miner toolkit Architecture [38]

Figure 5.8-a plots the recall, measure, and precision of the disambiguation organized as the minimum sense probability threshold that has been adjusted between 0% and 1%.

In figure 5.8-b plots the performance of the detection grouped as the minimum link probability threshold is settled between 0% and 1%.
We reach a logical conclusion through argumentation over human expressed views and reasoning over them. There many models have been proposed since 1950’s to come up with different scenarios of informal argumentations. Web emergence brought with it semantic web technology that shifted towards ontologies and in development context, it is witnessed that essential increase in web 2.0 human-centered collaborative
deliberation tools is seen. 150 scholastic studies [39] have been reviewed to come up with comprehensive and comparative view of modeling argumentation approaches. Social web tools have been investigated and semantic web argumentation models are studied in detail to come up with the conclusion of online applications should combine web 2.0 and semantic web technologies to generate future global worldwide argument web.

5.9 Knowledge Extraction on the Semantic Web

Few years back NLP tasks have been configured [40] including; NER, WSD, relation extraction, etc. for tasks related to semantic web that are; ontology learning, entity resolution, linked data population, NL querying to linked data, etc. existing state of the art knowledge extraction (KE) tools have been assessed as desired for semantic web. Landscape analysis is done of various tools either conceived particularly for KE on the semantic web or adapted to it or conceived as aggregators of data extractors. Currently available capabilities compared to a rich palette of ontology design constructed are visualized to come up with an overview of actual semantic reusability for KE output. Resultantly, many tools developed have provided that useful, scalable, application, and brief learning of basic semantic data structures, such as tagged named entities, factual relations, topics for document classification, and its integration with software languages are increasing fast.

Figure 5.9, represents the relations between linguistic knowledge and formal language, linguistic knowledge uses formal as background knowledge, but can access the formal knowledge as well.
5.10 Data and Information Integration

Data and information integration [41] becomes easy in semantic web within an infrastructure based on RDF and ontologies. Spatial data web is developed as detailed on the collective collaboration of Open Street Map data is interactively transformed and shown adhering to RDF data model. Information integration and aggregation is simplified through transforming as well as requiring comprehensive background knowledge related to spatial features like ways, landscapes and structures. It is described that how the data is interrelated to other spatial data sets and is accessible to machines according to linked data paradigm and to humans through various applications that include the faceted geo-browser. LinkedGeoData[41] project openly offers applications, interlinks, spatial data and vocabularies.

OpenStreetMap project is provided by Wikipedia to create a free amendable map of the whole world. It also provides wiki features such as full version history of the edits and an edit-tab. The objective of the LinkedGeoData project is to add an open, rich, and integrated geographical information to the Semantic Web by OpenStreetMap as its basic support. This is similar to the well-known DBpedia project, which follows an identical method based on Wikipedia.
5.11 Watson System

Watson System is reviewed in a tool report as being a semantic web engine with various capabilities not just limited to find and locate ontologies and semantic data online, but it also explores the content in these semantic documents. It is shown [42] that it is used to develop semantic applications such that it exploits the content within the semantic web. Watson has set of APIs containing high level functions for locating, exploring, and querying semantic data and ontologies published online. From these APIs new applications are developed that connect activities like ontology construction, sense disambiguation, matching, and question answering to the semantic web. Watson is also described as an unprecedented research platform to study semantic web and formalized knowledge in general.

Watson performs three main activities:

- Firstly, collects obtainable semantic content on the internet,
- Secondly, performs the analysis of collected data to extract indexes and useful metadata.
- Thirdly, it applies proficient query services to access the data.
Semantic search engine Watson is not only a development tool but also represents a unique resource for researchers, particularly, how data are created and formalized, utilized and online [43].

Semantic web can be seen as an empirical science with the web of data [44]. There are two problems that have to be dealt with; (i) Knowledge soup, is about semantic heterogeneity considered to be a difficult technical issue needing appropriate transformation and inferential pipelines to help make sense of the various knowledge contexts, (ii) knowledge boundary problem at the core of empirical investigation of semantic web searching for meaningful units that set up the research objects for semantic web. This issue touches many areas of semantic web studies: schemata, data, representation and reasoning, linguistic grounding, and interaction, etc.

Figure 5.10. Overview of the Watson architecture [42].
5.12 NetKAT

NetKAT [45] is a mathematical networking programming language for establishing semantic foundation. NetKAT design incorporates in itself the techniques for filtering, transmitting, and modifying packets. It uses a mathematically structured Kleene Algebra with tests (KAT) that is proved in its capacity for sound and complete equational theory catering substantial semantics. It has been practically implemented with syntactic techniques for reachability, non-interference properties implementation that isolates programs, and correct algorithm compilation is thus proved. It is thus established that the equational theory applied in several diverse domains provides reasoning while reaching, isolating traffic, and correct compilation.
Forwarding technique used in NetKAT is studied and packet details are known that includes fields for standard headers that have source address (src), destination address (dst), and protocol type (typ) and other two fields for switch (sw), and port (pt) that would identify the accurate location of the packet in the network.

Other atomic NetKAT techniques include filtering and modifying packets. A filter \( f = n \), would take any input packet \( pk \) yielding a singleton set \( \{ pk \} \), on condition if \( f \) (field) of \( pk \) (input packet) equals \( n \) and \( \{ \} \) otherwise. Modifying \( (f \leftarrow n) \) gets input packet \( pk \) and yields singleton set \( \{ pk' \} \) where \( pk' \) is the modified packet setting \( f \) to \( n \).

Reach-ability properties are formulated based on questions like:

1. Does the intrusion detection system implied reaches all un-trusted traffic circulating the network?
2. Either all hosts communicate among themselves?, and
3. Whether managed hosts are kept isolated from unmanaged hosts?

To cater to these problems an approach has been designed into NetKAT that show the encoding of two important classes of reach-ability properties within NetKAT equations. These equations are hence proved to be sound and complete being intuitive, semantically designed for reach-ability using its own language model.

For compilation [45] NetKAT program is executed on an OpenFlow switch and compiled using flow table, low level programming abstraction for which OpenFlow gives support. Prioritization is done in flow table based on list of rules where each rule has a pattern for matching packet headers and actions to other packets. Packet reaching a switch is catered as per highest priority matching rule.

5.13 Chapter Conclusion

Researcher and Institutions are regularly producing data using information assets and aforementioned semantic web technologies, however pursuing the extensive business and cultural acceptance of the Internet, the destructive impression of threats and the security closure on the individuals and organizations demands proficient security services. This chapter discussed semantic security topics and its techniques.
The major focuses for the development of the Semantic Web architecture are protocols and languages, whereas in order of compatibility with the web technology HTTPS was the most proficient as known for the Semantic protocol, but where the language part is being considered, the key point is the flexibility of developing and creating a Web of data.

This study presents an overview of semantic web technology, key security issues, protocol and security standards, some explanatory applications and supporting instances. It has observed that the secure version of HTTP has been developed to ensure that the data is not changed during its transmission and encryption protocol could handle the man-in-the-middle attacks, however several issues remain unresolved.

References


[18] Blake Middleton, James Halbert, and Frank P. Coyle, “Security Impacts on Semantic Technologies in the Coming Decade”, Southern Methodist University, Dallas TX 75205, USA


Chapter 6
Cluster Based Security Architecture and Communication Protocols for Ad-hoc Wireless Networks

6.1 Introduction

From the study review presented in chapter 3 and 4, it is revealed that the secure communication is an essential and critical element in networks and data communication which meets on most renowned authentication preconditions. However, in case of ad-hoc networks general authentication methods are not pertinent as public key infrastructures with a centralized certification authority are difficult to set up there. We propose and assess a security model based on a distributed certification feature and for this firstly we divide the node network into clusters keeping a distinctive node called head node. These nodes are designated a role of administrator of a cluster and responsible to share and store the network key used for certification. This chapter highlights several challenges in scalable clustered network architectures for MANETs and WSNs. This also comprising the classification and of cluster and cluster based security architectures and communication protocol. A major focus of clustering algorithms is cluster head selection.

The forthcoming generation of wireless systems, will consist of menageries technologies which organized both in multi-hop, self organizing networks called ad hoc networks distinguished by the lack of any fixed mechanism and in infrastructure-based architectures such as cellular networks, wireless LAN. This innovation of substantial heterogeneous networks is made possible by very large numbers of nodes into multi-hop, ad hoc network are able to self-organize them-selves using technological advancements. This network may then be inter-connected with classical wired networks such as the general Internet or the conventional telephone backbone.

A typical example of large multi-hop networks is given by wireless sensor networks (WSNs) [1]. The well-known prototype of ad-hoc networking focuses to
consider a higher number of nodes that are highly resource constrained. Whereas general ad hoc network or mobile or a large sensor network with static, small nodes, the major challenge that multi-hop networks protocol designers have to face concerns the dynamic nature of network topology and its growth. Therefore, network clustering is the most inspected solution proposed for scaling-down the networks having large number of nodes.

Clustering is observed as an initial step which provides the flat network with a hierarchal formation, the subsequent backbone structure implies the cluster-base hierarchy to design an infrastructure that is purposeful in providing required characteristics such as choosing data aggregation points, minimizing the overall power consumption, increasing the probability of aggregating redundant data, and minimizing communication overhead.

As a multi-hop, high-density, large-scale, network is pleasing of many applications; it increases the demand for extendable MANET architecture. The major challenges are because of increased route length in multi-hop MANET is to improve spatial variety in a large network area within the limited scalability. Routing protocols operates the primary cluster structure to up-date the location information and routing in an efficient mode.

In chapter 3 and 4, a review of literature pertaining to wireless sensor network, mobile ad hoc networks and respective security challenges and threats and solution has been presented. In the section 6.2, an extensive analysis of the related research achievements is presented. The literature reveals that there has been substantial work exertion for the progress and expansion of routing protocols in wireless networks which stands on the specific architecture and application requirements of the network and there quite valuable effort has been made to make it more robust, scalable and energy efficient technology.

Section 6.3 present the process of cluster formation that can be accomplished using connectivity or identifier mechanism. Further it defines the architecture of cluster based sensor network, different states of the nodes existing in a cluster, and respective functions. Contrary to distributed probabilistic clustering techniques, non-probabilistic
are mostly more consistent as concerning the direction of taking out well-balanced and robust clusters. A brief comparison of clustering algorithms is reported in the section 6.4.

In the similar class we are also discussing a relatively new and challenging category of clustering algorithms for wireless network and it is found that the biologically inspired techniques designed on swarm intelligence are certainly the most assured substitute for clustering in MANETs and WSNs in these days.

In section 6.5, a cluster based security architecture for ad-hoc wireless networks is proposed and evaluated for secure communication and access control. It supports an open network which allows new node to join the network without having any information about the network on basic access control for resources and services in the network. The model adopts distributed certification scheme to deploy authentication mechanism and secure data routing mechanism to achieve the integrity and confidentiality for secure network communication.

In section 6.6, intrusion detection and prevention based security solution for cluster based WSNs is presented which is constructed on MAC addresses of the nodes. In detail the design of wireless network and expected intrusion scenarios are illustrated with graphic explanation. Through simulation results and analysis it has been observed that number of issues can be prevented by early detection and prevention of intruder using efficient security system. Base station maintains security track of the network and ensuring high security without any substantial overheads.

In the last section 6.7 of this chapter, the cluster based communication model for homogenous mobile ad-hoc and wireless networks is discussed. Its cryptographic function provides the safe-shield for outside attacks and is the most suitable for this kind of networks because of energy efficiency and scalability characteristics of the network. It is observed that implemented symmetric key algorithm with session keys carry better security as compared to previous secure routing algorithm in cluster wireless sensor networks.

In summary, the contribution of this chapter is three-fold:

1. Proposed the cluster based security architecture and evaluated for secure communication and access control. It is a base of proposed optimized network
security solution for which CHs holds the network key of nodes used for certification and performs administrative tasks.

2. Simulation based analysis to conclude that early detection and prevention can secure the system more efficiently which is constructed on MAC addresses of the nodes.

3. Defense of the secure routing model adopted for cluster based WSNs using symmetric cryptography with session keys, is the optimized security solution.

### 6.2 Related Work

There has been significant work exertion for the progress and expansion of routing protocols in wireless networks which stands on the specific architecture and application requirements of the network. Although, there are many issues requesting to have significant attention while designing and mounting the routing protocols and from all the most critical factor is efficient energy utilization, as it directly influences to prolong the network. There have been a few efforts in the literature working on energy efficiency in WSNs.

In [2] a hierarchical protocol - Low Energy Adaptive Clustering Hierarchy - LEACH, presents that all of the member nodes communicate to CHs directly rather to base station. The process of LEACH protocol comprises of two stages:

- **Setup Stage**: In this phase, the clusters are formed and the CHs are elected by adopting the stochastic techniques.
- **Steady State Stage**: In this phase, the data is sent directly to the base station in such a way to minimize the overheads.

LEACH attended to decrease the utilization of energy for wireless network on one-hop routing, but this protocol is not suggested for the network covering larger area.

**Power-Efficient Gathering in Sensor Information Systems (PEGASIS)** [3], which provides positive developments over LEACH as each node, establishes a connection to the nearest node for data transmission to the BS. It has concluded that PEGASIS performance is better than LEACH [4].
Threshold Sensitive Energy Efficient, The Shortest Hop Routing Tree protocol [5], Extending Lifetime of Cluster Head [6] Energy Efficient Cluster Formation Protocol, Base-Station Controlled Dynamic Clustering Protocol [7], all of these protocols contribute in a way that helps to decrease the utilization of energy by adopting various algorithms.

Quite valuable effort has been made to make it more robust, scalable and energy efficient technology. In [8] explains three different routing models used for WSN as follows:

- One Hop Model,
- Multi-Hop and
- Cluster Based Model

**One Hop Model:** The simplest network model which establishes direct data transmission to the base station is one hop model. The most significant and widely used routing algorithms are cluster structure which select the optimum cluster size, CH selection and its re-selection, and cluster maintenance are the key points that must have to be considered while designing and developing the clustering algorithms.

![Figure 6.1 One hop model [8]](image-url)
**Multi-Hop Model:** It’s an energy efficient routing model in which nodes select respective neighbors to make the transmission link to the base station.

![Multi-hop model](image)

*Figure 6.2 Multi-hop model [8]*

**Cluster Based Model:** In this model network is assembled into clusters. Each cluster containing one head is called cluster head and remaining are its members. In this model, CH collects the data from the members and sends its aggregate to the base station after some processing.

![Cluster based model](image)

*Figure 6.3 Cluster based model [8]*
6.3 Cluster Formations

Clustering is a process of dividing the network into inter-linked components called clusters. Each cluster has a cluster-head which acts as a coordinator or medium for data transfer between nodes within the sub-structure. The Gateway node has two or more cluster heads as its neighbors or at least one cluster head and another gateway node. Cluster heads also communicate with each other by using gateway nodes. Cluster formation can be accomplished using these two mechanisms:

- Connectivity - based Clustering (Highest Degree Heuristic): The node, which has the most neighbor nodes, is elected as the cluster head.
- Identifier- based Clustering (Lowest-Id Heuristic): A node selected as the cluster head if it has the highest or lowest ID in its neighbor nodes.

The following status explains the clustering process depending on the current node state [9], these states are:

**Cluster head**: If a CH finds that it has a two way communication link established to another CH for the time being and have lower ID, it will change its state to ‘member’. Else the case, it will be in the same state, see figure 6.4.

**Member**: If a member node drops its CH, it will be looking for another communication link to the member node. If it has found one of them it will change its status to CH having lowest ID, else case it will replace the status to ‘undecided’.

**Undecided**: If the node does not belong to any cluster in ‘undecided’ state and this usually happens if a new node comes into the network. Thus, if it receives a HELLO message from a CH it will replace its status as the member node of the cluster on the indication of CH. Otherwise it looks up in its neighbor table if it has any bi-directional links.

To achieve the goal to minimum cluster re-structuring, the formation of the clusters should change as hardly ever as possible and it can be possible if a ‘non-cluster head’ does not challenge the status of present CH even though it has a lower ID.

In WSN, each node is presented with a limited energy source and once exhausted, it would be difficult to substitute them, and so it is essential to minimize the utilization of power in WSNs for to maximize their lifetime.
Figure 6.4. CH-5 shifted into cluster 2 it take off its role as CH according to its higher ID. Nodes A and B which lost their CH.

Clustering approach is energy efficient in a way to slow down the ‘quick energy consumption’ process from the network. Figure 6.2 shows Architecture of Clustered Sensor Network. Distributed clustering techniques are much more functional than centralized clustering methodologies because of that the centralized monitoring and control system for a large group of node is not proficient or possible to manage [10].

In WSNs large number of nodes is deployed for data collecting, but if all the nodes transmit and communicate data in network then there will be data collision and congestion, in fast energy reduction of sensor networks. Hierarchical clustering technique solves this issue by efficient resource utilization [8]. In a hierarchical network structure each cluster has a cluster head and normally carry out the particular activity such as aggregation, fusion, and have many sensor nodes as members [11].
In WSN base station (BS – a data processing point) is usually placed and fixed far away from the sensor nodes. All the data from cluster head is transmitted to base station for processing, and then information is accessible to the user. Actually cluster head plays a role of ‘gateways’ between the nodes and the base station [12].

6.4 Classification of Clustering Algorithm

There have been several different ways to organize the algorithms used for the clustering of WSNs. One of the most conventional are:

- Centralized Clustering Algorithms
- Distributed Clustering Algorithms
Distributed clustering scheme is in our consideration because centralized schemes are not appropriate [12]. All of the well recognized clustering algorithms depending on its formation parameters can be divided into two major groups:

- Probabilistic (hybrid or random) Clustering Algorithms
- Non Probabilistic Clustering Algorithms

### 6.4.1 Probabilistic Algorithms

In the probabilistic clustering algorithms, each sensor node is designated with some probability value or a node ID used to decide the initial CH. The probabilities at first allocate to each node considered as the basic condition, for each node existing in network decides itself the election as CHs; though different criteria could be used for the cluster formation process i.e., the communication cost or during Selection process, the residual energy in order to attain the better network lifetime and energy consumption. Probabilistic Clustering algorithms proof fast convergence addition to the efficient load balancing, energy efficient network utilization, and reduced volume of exchanged messages and low message overheads [8,12].

**LEACH:** LEACH is a Probabilistic Clustering Algorithm, which uses distributed scheme for one-hop cluster formation without any centralized control. The main purpose of LEACH algorithm is to reduced energy utilization by sensor node, extend network lifetime and reduce the communication using data aggregation so that congestion on link can be controlled [12].

In LEACH, the CHs squeeze the data receiving from member nodes and transmit an collected package to the BS to reduce the amount of data that must have to send to the BS. [2, 10, 12]

This protocol is more appropriate if BS is not much away from the nodes. In MTE, nodes transmit data to BS via transit nodes. These intermediate nodes provide a routing path for other nodes and also sense the network. For instance, If there are three nodes A, B and C in network, the transmission is done by node A to C via B if and only if this condition fulfills [13].
\[ E_{Tx}(k,d = d_{AB}) + E_{Tx}(k,d = d_{BC}) < E_{Tx}(k,d = d_{AC}) \]  \hspace{1cm} (6.1)

If the transmission energy as defined in equation 1 used from A to B and from B to C is less than the transmission energy of nodes A to C then the node B will be used as middle node.

**Hybrid Energy-Efficient Distributed Clustering Algorithm:** The main objective of HEED is extending the lifetime of network by distributing energy consumption among cluster nodes. It is a Probabilistic Clustering Algorithm. In contrast to LEACH, in HEED cluster head nodes are not elected randomly. Indeed, distributed scheme is used for cluster formation; it regularly elects cluster heads on the bases of a nodes residual energy as a first parameter and node degree or node proximity to its neighbors as a secondary parameter. By reducing the clustering process in a pre-determined range of iterations, minimizing control overhead and producing well distributed cluster heads. In HEED, single-hop communication model is maintained within each cluster for nodes communication, whereas multi-hop communication model is permissible between CHs and BS [10,14].

In HEED every CH contains the information of residual energy of each sensor nodes as ‘primary’ parameter, but awareness of the whole network might be difficult in the realistic world as it required to-settle-on intra cluster communication configuration and cost of these parameters. [14]

**DEEC:** DEEC is designed for heterogeneous network in which some nodes are higher with more energy than normal nodes. All the nodes are operational with diverse energy intensity to enhance the network lifetime. CH selection is based on the average energy of the network and residual energy of node. The data is collected by CHs from its members and dispatch to the BS [15]. The nodes having high initial energy and high residual energy will be CHs.

DEEC is multi-level heterogeneous network; there are two types of nodes such as normal nodes and advanced nodes. The total initial energy of this network is represented by:

\[ E_{total} = \sum_{i=1}^{N} E_0(1+a_i) = E_0(N+\sum_{i=1}^{N} a_i) \]  \hspace{1cm} (6.2)
In DEEC, there are some nodes having extra energy represented by $a$ and $m$ is the probability of extra nodes. CH selection is based on the proportion between average energy of the network and residual energy of each node. $e_{\text{pochs}}$ of being cluster head is set according to their residual and initial energy. The nodes having high initial and residual energy would be having more probability to become cluster head (4). CH selection is based on the threshold equation (5):

$$p_i = p_{\text{opt}} \left[ 1 - \frac{E(r) - E_i(r)}{E(r)} \right] = p_{\text{opt}} \frac{E_i(r)}{E(r)}$$

(6.3)

Where $p_{\text{opt}}$ is the average probability of $p_i$, $E_i(r)$ is the residual energy of the network at round $r$ and $E(r)$ is the average energy of the network. Total number of CHs per round per $e_{\text{pochs}}$ is:

$$\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} p_{\text{opt}} \frac{E_i(r)}{E(r)} = p_{\text{opt}} \sum_{i=1}^{N} \frac{E_i(r)}{E(r)} = N p_{\text{opt}}$$

(6.4)

$$T(s_i) = \begin{cases} 
\frac{p_i}{1 - p_i \left( \frac{1}{r \mod \frac{1}{p_i}} \right)} & \text{if } s_i \in G \\
0 & \text{otherwise} 
\end{cases}$$

(6.5)

Threshold value is given in equation 5 on which CH selection criteria is based. In two level heterogeneous network the value of $p_{\text{opt}}$ is given by:

$$p_{\text{adv}} = \frac{p_{\text{opt}}}{1 + am}, \quad p_{\text{nrm}} = \frac{p_{\text{opt}} (1+a)}{1 + am}$$

(6.6)

Then use the above $p_{\text{adv}}$ and $p_{\text{nrm}}$ instead of in above equation for two level heterogeneous network as supposed in equation (7):

$$p_i = \begin{cases} 
\frac{p_{\text{opt}} E_i(r)}{(1+am)E(r)} & \text{if } s_i \text{ is the normal node} \\
\frac{p_{\text{opt}} (1+a)E_i(r)}{(1+am)E(r)} & \text{if } s_i \text{ is the advanced node} 
\end{cases}$$

(6.7)
It can be extended to multi-level heterogeneous network which is given as:

\[
P_{multi} = \frac{p_{opt} N(1+a_i)}{(N+\sum_{i=1}^{N} a_i)} \tag{6.8}
\]

Above \( p_{multi} \) is used instead of \( p_{opt} \) to get \( p_i \) for heterogeneous node \( p_i \) for the multilevel heterogeneous network is given as:

\[
P_i = \frac{p_{opt} N(1+a)E_i(r)}{(N+\sum_{i=1}^{N} a_i)E(r)} \tag{6.9}
\]

In DEEC average energy \( E(r) \) of the network for any round \( r \) as in:

\[
E(r) = \frac{1}{N} E_{total} \left( 1 - \frac{r}{R} \right) \tag{6.10}
\]

\( R \) denotes total number of rounds and is expected as follows:

\[
R = \frac{E_{total}}{E_{round}} \tag{6.11}
\]

There would be more chances for a node to be selected as CH if the residual energy of the nodes is higher than average energy of network.

**H-TEEN**: H-TEEN [15] is a variation of TEEN protocol, a hierarchy of clustering to better manage with huge network area. H-TEEN performs better because of less consumption of energy in large network; in contrast to TEEN where small number of layers in hierarchy consumes lot of energy because of larger distance. H-TEEN is a four layer hierarchal clustering model where sensors are self-organized and construct a tree for communication and transmit data to the CH. CH selection is same as in LEACH and TEEN.

**Two Level Low Energy Adaptive Clustering Hierarchy (TLEACH)**: TLEACH algorithm is an expansion LEACH and utilizes primary and secondary levels of CHs. The two-level formation of TLEACH reduces the data transmitting to the base station by efficiently reducing the energy usage. T-LEACH uses random alternation of local cluster with base stations by this; it forms of primary cluster-heads and secondary cluster-heads. Resultantly, improves the distribution of the energy load along with the sensors in the network. T-LEACH uses localized co-ordination to facilitate robustness and scalability.
The T-LEACH uses the below mentioned techniques to attain the latency and energy efficiency [16]:

- Adaptive, Randomized, and self-configuring cluster formation
- Localized control for data transfers

**Energy Efficient Clustering Algorithm (EECS):** EECS is a Distributed clustering scheme, in this cluster head and BS communicate in one-hop method. Mainly it uses data gathering applications in wireless sensor networks. EECS is same as to LEACH with some advancement in cluster head selection and cluster formation process.

In the cluster head election phase, a persistent range of nominee nodes are elected and compete for cluster heads without iteration, according to the node residual energy. Moreover, in cluster formation phase clusters are formed by resizing of clusters build on distance of cluster from the base station. [10,17]

**Conclusion:** Energy efficiency and network lifetime are attentive issues in WSNs. Clustering technique have been proposed to resolve these issues by using different clustering schemes. We conclude from our analysis that H-TEEN is more energy efficient because of hierarchical clustering and threshold value; DEEC performs better in sending packets among all protocols. HTEEN outperforms in case of mobile-sink and all other protocols perform better in case of static-sink. [15]

### 6.4.2 Non-Probabilistic Algorithms

In Non Probabilistic Clustering Algorithms, more specific criteria is adopted for cluster formation and CH election is primarily set on the proximity of node i.e. degree, distance, connectivity, etc. and on the information gathered from the adjacent nodes. Whereas some algorithms jointly use the group of metrics like the transmission power, residual energy, mobility, etc. to accomplish more generic targets [12].

The cluster formation process generally depends on the communication model adopted for nodes with their neighbours’ i.e. in one-hop or multi-hop neighbours and usually expects more exhaustive exchange of messages.

On the contrary to these techniques are mostly more consistent as concerning the direction of take outing well-balanced and robust clusters. In the similar class we are also
discussing a reasonably new and challenging category of clustering algorithms for Wireless Network and it has found that the biologically inspired techniques designed on swarm intelligence are certainly the most assured substitute for clustering in MANETs and WSNs in these days [16].

**Highest-Connectivity Cluster Algorithm (HCC):** HCC is a distributed multi-hop non-probabilistic hierarchical clustering algorithm, which could also proficiently form a multi-level cluster hierarchy. HCC is processed in two phases, one is “Cluster Formation” and other one is “Tree Discovery” [12].

Any node in the network can start or initiate the cluster formation process. Each node transmits its number of neighbors’ on the network, connectivity of node is measuring metric and the node with highest connectivity is elected as CH. A Node that already has been chosen as a head of cluster might get withdraws to becoming a cluster head. The creation of clock synchronization and one-hop cluster conditions bound the algorithm more functional [10].

**Weight-Based Clustering Algorithm:** WCS is a distributed non-probabilistic algorithm for cluster formation in single-hop communication scheme. For power saving objective, it executes a new selection procedure whenever node loses its connection with cluster head and performs configuration of the network’s structure [12]. WCA is a metrics based that is considering multiple parameters i.e. mobility, transmission power, node degree, and the residual energy. Form these any or all of these parameters, can be used as a metric to select the cluster head depending on the specification of an application. This algorithm performs well against load balancing because of less number of nodes in a cluster. [4,12]

**Bio-inspired Clustering Algorithm:** In last few years there have been introduced new clustering techniques based on swarm intelligence model of social insects behavior like that ants, bees [12]. In colonial model, two objects gather or assemble as they identify either belong to similar cluster or not. For wireless network, at the start it has recognized that nodes having a lot of residual energy become CHs several time. Then, every chosen node met another to exchange data after that by using these information clusters are formed, combined, and removed. All remaining node keeping small amount of residual energy chooses a cluster the bases of some criteria such as its distance to the CH, the
residual energy of the CH, and calculation of expected cluster size. Ultimately, energy efficient clusters are created.

In general, bio-inspired clustering technique can dynamically manage the selection of cluster head and attain uniform and energy efficient distribution of cluster heads.

6.5 A Cluster-Based Security Architecture for Ad-hoc Wireless Networks

Secure communication is an essential element in computer networking and authentication is one of the most renowned prerequisite. However, general authentication mechanisms are not feasible to be deployed in ad-hoc networks due to centralized certification authority (public key infrastructures) is difficult to implement there. Researchers propose [18] a distributed certification facility based security concept in which a network is segregate into clusters having a special node CH. These CHs hold the network key of nodes used for certification and perform administrative tasks. New nodes can only become a full member of the network once if it has signed certificate received on successful ‘authenticity’.

Ad-hoc networks are subject to several kinds of attacks. In contrast to wire-line networks, known attacks such as man-in-the-middle, masquerading and replaying of messages can be taken out easily. Furthermore, implementing security mechanisms is not easy because of inherent characteristics of ad-hoc networks i.e. limited resources of end systems, the high dynamics topology and perhaps asymmetrical communication links.

Also, gateways (GW) regularly send out GW signals to inform their particular clusters about adjacent clusters. Routing is typically categorized into two classes: intra-cluster - routing within a cluster and inter-cluster - routing between different clusters [3].

6.5.1. Authentication Using Distributed Certification

A network-wide end-to-end secure communication can be established by using public key cryptography to form a distributed certification infrastructure. For controlling access to services and resources ‘authorized certificates’ are used. Furthermore
‘symmetric encryption’ could be applied to secure the communication links within the clusters.

**Network-Wide Certification Infrastructure:** Public keys are distributed in the ad-hoc network using certificates issued by an authority-CA. On the other hand for PKIs in fixed networks; CH of the network assigned the role of the distributed CA. It is created by a sub-set of all network nodes. The private key of the CA is distributed over the CHs, i.e., every CH holds a part of the complete key. Furthermore, the protocols used for common certification key generation, its management and usage for organizing the whole ad-hoc network, the CHs therefore form a logical network.

![Authentication Process](image)

**Figure 6.6 Authentication Process** [18]
**Intra-Cluster Security:** Independence to end-to-end security measures ‘symmetric key’ mechanism can be used for cluster-wide that is recognizable to the cluster’s nodes. This key can be used to secure intra-cluster’s traffic that cannot secure by any other methods, and also to hide source and destination addresses from eavesdroppers not existing into the cluster.

**Node Status and Authorization:** A new node that comes into and joins a cluster has designated an initial status of a ‘guest’ without having any access rights. Once an authentication has completed successfully and public key has been assigned by the CH network, then it will become a ‘member’ of cluster and can obtain further access rights by using that authorization certificates issued by the CA. In contrast to identity-based key certificates, authorization certificates can be issued by any network node managing an individual resource or service, such as Internet access, printer.

Figure 6.7. States of a New Node during Log-On [18]

In order to log-on to the network, at first new node has to find a cluster. If it receives CH signal, it sends its log-on request to the CH. The new node and the CH negotiate some parameters (like how the symmetric cluster key is to be used later on and the number of warranty certificates required), and then new node becomes a guest.
Otherwise, if nodes do not receive any CH signal, it will setup its own cluster and acts as a CH of this cluster.

When CHs are being requested for a certificate shares by a new node, it will first ensure that the issuers of the given ‘WarrantCerts’ are authorized enough to guarantee for a ‘guest’. Its validity can be confirmed using warranty authorization certificates ‘WarrantAutCert’. Each warrant $S$ sends a copy of its certificate to guest node $A$:

$$WarrantAutCert(S) := \text{Node}(S), \text{PubKey}(S), \text{Fct}(“S \text{ may warrant”), Sign((CH-Network)$$

The CHs verify the WarrantAutCert and the WarrantCert presented by node $A$ and send their shares of an ‘identity certificate’, once the certificate validity has been confirmed. After $A$ received certificate shares, it can complete its identity certificate:

$$IdCert (A) := \text{Node} (A), \text{PubKey} (A), \text{Validity} (t), \text{Sign} (CH-Network)$$

Now new node $A$ is a full member of cluster and has a signed key. Finally, the CH sends the symmetric cluster key which is encrypted with $A$’s public key to $A$.

**Gateways:** Each node of cluster that gets in contact to outside the home cluster can probably act as a gateway. An optional scenario, to take authorization to act as a gateway can be managed by using gateway authorization certificates - GwAutCert approved by the CH of the network:

$$GwAutCert = \text{Node}(N), \text{PubKey}(N), \text{Fct (“Gateway”), Sign} (CH-Network)$$

**Access Control:** Access to resources and services can be managed by using authorization certificates. Individuals that are accountable for managing the access to specific resources or services, or owner of the resources itself or the service provider can provide authorization certificates to the users on request. These certificates comprise of some authorization information and the public key of the subject.

Various types of keys (asymmetric public key, symmetric cluster key) and certificates can be used to make the communication secure. Each node of the network itself decides and defines the requirement and level of security and accordingly applies an encryption scheme. Complexity levels are describes as:

- Public node keys – issued by the cluster head
- Public node keys – for direct exchange
- Secret cluster key – for intra-cluster
- No encryption
Permitting a flexible complexity is an advantage for nodes having low resources, which can adopt an appropriate security level. But, if node does not allow or agree upon an ordinary level, communication would not be possible.

6.5.2. Integrity and Confidentiality: Secure Data Routing

In the design every node has a session key at the time of deployments. Primarily the node encrypts the captured data by applying the Blowfish Algorithm, which makes the data transmission more secured. It sends encrypted data to the CH and from CH to the gateway and lastly to the base station.

The benefit of this system is that, it improves communication security and requires very less energy as compared to other cryptography algorithms. After initial session successfully, the base station will generate a new session key using a pseudorandom function \( f() \) and old session key and sends it to the respective gateway. The new session key is sent to CH for data encryption of the new session. Every time BS changes the session key dynamically for communication process.

6.5.2.1 Secure Data Routing Model

*Notation uses in Secure Data Routing Algorithm (SECDRA):*

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_t )</td>
<td>Packet</td>
</tr>
<tr>
<td>( E() )</td>
<td>Encryption function</td>
</tr>
<tr>
<td>( SK )</td>
<td>Session Key</td>
</tr>
<tr>
<td>( S )</td>
<td>Base Station</td>
</tr>
<tr>
<td>( CH_1, CH_2, CH_3 )</td>
<td>Cluster Head IDs</td>
</tr>
<tr>
<td>( G )</td>
<td>Gateway</td>
</tr>
<tr>
<td>( D() )</td>
<td>Decryption function</td>
</tr>
<tr>
<td>( f() )</td>
<td>Pseudorandom function</td>
</tr>
<tr>
<td>( TGS )</td>
<td>Logical time stamp of the gateways</td>
</tr>
<tr>
<td>( TSG )</td>
<td>Logical time stamp of the Base Station</td>
</tr>
<tr>
<td>( SK_n )</td>
<td>New session key</td>
</tr>
<tr>
<td>( SK_c )</td>
<td>Current session key</td>
</tr>
<tr>
<td>( X )</td>
<td>Random number</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are four types of communications placing in the proposed model:-

- Node → Cluster Head
- Cluster Head → Gateway
- Gateway → Gateway
- Gateway → Base Station
**Node to Cluster Head:**

A node $S_i$ encrypts the packet $P_i$ using current session key $SK$, which is assigned at the time of deployments and sends it to its Cluster Head (CH).

$$S_i \rightarrow CH$$

$ESK(P_i)$

**Cluster Head to Gateway**

The following actions are performed at the CH:

i. CH concatenates the encrypted packets received from the nodes.

ii. Increments the value of logical time stamps (TGS) by one and appends it to the concatenated packets.

iii. Concatenates its own ID and sends it to the next cluster head via its gateway on the path to the Base Station.

$$CH \rightarrow G$$

$$\{ \text{ESK}(P_n) \} \| \text{TGS} \| \text{CH_ID} \| \text{CH}_1 \} \| \ldots \| \{(\text{ES}_1K_1(P_d)) \| \text{TGS} \| \text{CH_ID} \| \text{CH}_2\} \| \ldots \| \{(\text{ES}_2K_2(P_x)) \| \text{TGS} \| \text{CH_ID} \| \text{CH}_3\}$$

Where,

$SK(P_n)$ is the encrypted packet from node $n$ belonging to Cluster head $CH_1$ of a cluster.

$(ES1K1(P_d) \| TGS \| CH_ID \| CH2)$ is encrypted packet received from cluster head $CH_2$.

$(ES2K2(P_x) \| TGS \| CH_ID \| CH3)$ is encrypted packet received from cluster head $CH_3$.

TGS: - Time stamp belonging to each cluster head.

$CH\_ID$: - Cluster Head ID
**Gateway to Gateway**

All encrypted packets from the different cluster heads are sent to their respective gateway nodes to forward to the base station. The gateways also allow communication between different cluster nodes.

\[ G_{n1} \rightarrow G_{n2} \rightarrow G_{n3} \]

**Gateway to Base Station**

The Gateway sends the concatenated encrypted packets to the sink.

\[ G \rightarrow \text{Base Station} \]

\[ \{ \text{ESK}(P_n) \} | | TGS | | CH\_ID | | G_n | | . . . | | \{ \{ \text{ES}_1 \text{K}_1(P_d) \} | | TGS | | CH\_ID | | G_d \} | | . . . \}

The base station processes the received packets before it forward to the destination node.

**Activities Performed by Base Station**

i. To check the credibility of time stamp, base station decrypts the received packets using the current session key

\[ \text{DSK} \{ \{ \text{ESK}(P_n) \} | | TGS | | CH\_ID | | G_n | | . . . | | \{ \{ \text{ES}_1 \text{K}_1(P_d) \} | | TGS | | CH\_ID | | G_d \} | | . . . \}

If \( TGS \geq TSG \),
then time stamp is credible and data is authentic and decrypt the \( P_x \) to get the original message, which is \( \text{DSK ESK}(P_x) \rightarrow P_x \),

if \( TGS < TSG \),
then the base station discards the packet or sends a re-transmit request to CH.

ii. Verify CH IDs in the packets.
Base Station Generate new Session Key On Expiry

On expiry of current session, base station increments the value of TSG by 1, and generates the new session key using the pseudorandom function \( f \). The new session key is a function of current session and a random number.

- New Session \( SK_{n1} = f(SK_c.x) \) where \( x \) is a random number.
- Another new session key \( SK_{n2} = f(SK_{n1}) \)

Session key is modified for the next session as following:

i. Base station encrypts the new session key \( (SKn2) \) using the new session key \( (SKn1) \) and sends it to the corresponding gateway, \( ESKn1 (SKn2) \).

ii. Gateway sends the new session key \( (SK n2) \) to the Cluster head to broadcast the new session key \( (SK n2) \) to the sensor nodes in its cluster.

iii. Sensor nodes update its session keys, with the new session key.

6.6 Intrusion Detection Based Security Scheme for Cluster-Based Wireless Networks.

An intrusion is a set of activities that can control an authoritative access or modifications of the wireless network system. These detection methods can identify malicious user on the bases of some anomalies; normally, the adjacent of a malicious node are the first points observed or influenced by those abnormal behaviors. Intrusion detection system-IDS deployed to supervise and monitor the systems and computer networks, detect the probable intruders in the network, and notifying users once if intruders detected, if essential and possible then reconfiguring the network [1, 13].

In case of clustered wireless network where, network topology depends on location information, communication range of the nodes, remaining battery power and distance between the nodes. A malicious intruder can influence these attributes and
attract the network in the direction to create a sinkhole which may turn into black hole. Constant observation can be an energy taking activity, that may not be required in wireless network. However, a cluster-based intrusion identification scheme is developed for wireless network system.

### 6.6.1 Intrusion Detection Based Security Scheme

The proposed scheme [19] applying MAC addressing based intruder tracking system to detect and prevent the intruder in the network.

### 6.6.2 Design of the Wireless Sensor Network

Suppose that a WSN consisting of ten Cluster Heads (CH-01 to CH-10) with respective node is forming the clusters. As shown in Figure 6.10, the CH collects data from each node, compacts in, to pack together and sends it to the BS. CHs maintain the status and information of each node in the cluster and send this information to the BS regularly.

![Design of Wireless Sensor Network](image1.png)

![Data communication](image2.png)

Figure 6.9. Design of Wireless Sensor Network

Figure 6.10. Data communication

Base station establishes connection to the nodes via CHs and also maintains the status and track of health of each node in all clusters by using the MAC address. As shown in Figure 6.11, any change in the topology of cluster, and re-organization of the
clusters due to change in CH is monitored and controlled by the BS. For instance, suppose that an intruder appears as illustrated in figure 6.12 in the network and it will try to connect with one of the closest to become a part of the network.

\[
\text{Figure 6.11. Data communication between CHs and BS}
\]

\[
\text{Figure 6.12. Intruder introduced in the network}
\]

### 6.6.3 Expected Intrusion Scenarios

There are following two ways in which an intruder could be a part of the network

**Scenario I**

Suppose that the intruder finds out N-6, N-9 node and N-15 in the 1st cluster as closest potential nodes. Shown in figure 6.13, the intruder attempted to connect with N-16, N-9 or N-15 node in listen mode with unknown MAC address. As shown in Figure. 6.14, it successfully connects to communicate with node N-9 by keeping its identity unknown in listen mode. The intruder has the ability to decode the data packets of node N9 being sent to and received from the network.
The attempt to join the network has succeeded and intruder has translated the MAC address of the CH and N-9, as can be seen in Figure 6.14.
As shown in figure 6.15, the intruder attempted to make a copy of N-9 node and a copy cluster head as CH1 by using MAC address. The main target of intruder is to shutdown the N-9 node and to route the data via intruder. The time in which intruder sends data on the copied node; the cluster head recognize its as unknown MAC address, consequently this information is transmitted urgently to the base station to take the mandatory remedial action, can be seen in Figure 6.16. The BS transmits a publicize alert to network on the presence of such intruder identity and the cluster head CH-1 is commanded to prevent the sending and receiving of data to make sure that the intruder cannot contaminate the operations of the network further.

**Scenario II**

Suppose that the intruder has found the CH-1 as the closest promising node. As shown in Figure. 6.17 the intruder attempted to connect with CH-1 in listen mode only with hidden MAC address.

![Diagram 6.17](image)

**Figure 6.17 a, b** Intruder communicating with Cluster Head CH-1

Attempt to join the network succeeds and intruder has attached with CH-1 in listen mode. As shown in Figure. 6.18, intruder has translated the MAC address of the CH-1
By using the MAC address information of CH-1 the intruder successfully forms a replica of CH, as can be seen in Figure 6.19. As intruder route on the data via replicated ID and CH finds un-known MAC address of copied CH-1 and instantly communicates to the base station for further remedial actions.

As a remedial action the base station publicizes an alert to the network related to the existence of intruder with its identity data. As illustrated in Figure 6.21, the CH is
commanded to stop the sending and receiving of data to make sure that the intruder cannot impact on the operations of the network further.

6.7 Security of Cluster-Based Communication Protocols

Ad-hoc and WSNs containing generally of small nodes having restricted resources swiftly promising technology for low cost, scaled, automatic monitoring of different environments. The proposed communication model for this kind of network is Cluster based communication because of energy efficiency and scalability of the network. Research [20] emphasized on security solution LEACH for cluster based homogeneous wireless network having limited resources.

SLEACH is a modified form of LEACH adding cryptographic safe-shield for outside attacks. It protects the network from becoming an intruder to Cluster Head or inserting false data.

6.7.1 Secure Routing in Cluster Based WSNs using Symmetric Cryptography with Session Keys (SRCWSNS)

Sensors which are acquiring huge attacks are not suitable for complex cryptographic algorithms due to its low computational power and limited resources. An implemented symmetric key algorithm with session keys carry better security as compared to previous secure routing algorithm in cluster wireless sensor networks [21].

6.7.2 Comparison of S-LEACH and SRCWSNS

SLEACH is one of the well-known protocols used in cluster based MANETs and WSNs. SLEACH mitigates almost all attacks excluding ‘sink hole’ and ‘wormhole’ attacks. The main aim of sinkhole attack is to steal the information which will lead to wormhole attack.

Consider the following scenario will give explanation for the mitigation of sinkhole attack. Here node M acts as valid node, which compromises on the location of neighbors cluster head node. The forwarding cluster head checks the identity of M node.
and checks the symmetric key. Even M got symmetric security key the forwarding cluster head checks loading time of the node or bootstrap time, it ensures more security. These keys are encrypted by using ‘blowfish symmetric algorithm’. So to find out these keys is very difficult. It takes more time to crack the keys, in meanwhile the base station will change the session key. So to steal the information from nodes intruder again try to decode the content to get the session key.

![Figure 6.22. Mitigation of Sinkhole Attack](image)

Even if the malicious node will compromise on one node there is no risk to compromise other nodes through this one. So the effect of node compromising will be controlled to only that part of the networks. Similarly forwarding node will find out the malicious node and also mitigates wormhole attack.
Table 6.1. Pros and Cons of Different Secure Routing Protocols [21]

<table>
<thead>
<tr>
<th>Protocol Name</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secure Routing Based on Distributed Key Sharing in Large-Scale Sensor Networks[22]</td>
<td>proposed two protocols for secure routing: a secure geographic forwarding protocol (SGFP) and a temporal-key establishment Protocol (TKEP).</td>
<td>It’s light weight protocol and mitigates Sybil and selective forwarding attacks.</td>
<td>To implement this we need more equipment like servers for distribution of keys. But this is not possible in hostile environment like border area security system.</td>
<td>Still it possesses attacks like worm hole; flooding and lap tap class attackers.</td>
</tr>
<tr>
<td>SNEP &amp; μTESLA[11]</td>
<td>These two protocols are using for DoS attacks and broadcast authentication</td>
<td>It provides broadcast authentication based on symmetric cryptography by delayed disclosure of authentication keys.</td>
<td>A major limitation of μTESLA and its variations is the Authentication delay. In other words, a receiver cannot authenticate a broadcast packet immediately after receiving it.</td>
<td>Still facing DoS and authentication of broadcasting messages.</td>
</tr>
<tr>
<td>Access control in wireless sensor networks[8]</td>
<td>It is developed on Elliptic Curve Cryptography (ECC).</td>
<td>it provides great security compare to RSA and takes less memory size</td>
<td>For digital signature verification it takes more time, nearly 1.6sec</td>
<td>Cannot mitigate the malicious new in the vicinity of compromised node.</td>
</tr>
<tr>
<td>A Deterministic Key Management Scheme for Securing Cluster-Based Sensors Networks[3]</td>
<td>Author introduced DK-LEACH for secure communication.</td>
<td>The secure communication overhead is very low compare with LEACH protocol.</td>
<td>Managing of keys between base station and cluster head, cluster head and its member nodes takes time.</td>
<td>DoS and lap tap class attacks</td>
</tr>
<tr>
<td>SLEACH Secure Low. Energy Adaptive Clustering Hierarchy Protocol for Wireless Sensor Networks[4]</td>
<td>It is an extension to LEACH protocol and uses symmetric efficient one-way hash chains and inexpensive symmetric operations rather than expensive asymmetric cryptography operations.</td>
<td>This scheme has given extensive practical scenarios of attacks and analysis of those attacks.</td>
<td>Cluster could not change dynamically and it does not provide more nodes, which are nearer to base station to distribute high overhead.</td>
<td>Because of symmetric mechanism if node compromised, then whole system is able to compromise by hacker</td>
</tr>
<tr>
<td>Dynamic End-route Filtering Scheme for Data Reporting in Wireless Sensor Networks[23]</td>
<td>It has developed on diffie-hellman key algorithm. Author proposed architecture</td>
<td>It mitigates DoS attacks and false reports injection.</td>
<td>But proposed scheme applicable for only fixed size networks, they are not discussed about new node deployment</td>
<td>Wireless sensors have limited capabilities. But takes large memory to save keys and leads</td>
</tr>
</tbody>
</table>
6.8 Chapter Conclusion

Wireless sensors are limited computation and resource devices that are not recommended for complicated asymmetric key algorithms. Whereas ad hoc networks characteristics are varying link qualities and highly dynamic topology, therefore it is not recommended to deploy central authentication and key management system as network is vulnerable to definite points of attack and failure. Alternatively, the task of certification authority (distributed public key infrastructure) jointly performed by the cluster heads of the network that was divided into clusters and new nodes can only become a full member of the network once if it has signed certificate received on successful ‘authenticity’.

It is observed that symmetric cryptograph methods adopted by some algorithm are not mitigating all attacks whereas comparing it to clustering based network architecture which has applied distributed public key infrastructure is highly recommended to the distinctiveness of ad hoc networks. So the detection of session key and the whole scenario cannot detect it and mitigates more or less all attacks.

By this study we conclude that secure routing scheme based on session keys by implementing symmetric cryptography using Blowfish Algorithm (symmetric key algorithm) with session keys is the Optimized Network Security Solution for Knowledge Networks (ONSS-KN). This model conveys better security as compared to previous secure routing algorithm adopted for cluster wireless networks. Furthermore it will provide more security by checking one more parameter that is loading time of node along will check serious global and session key.

Optimized security model add security to cluster based communication protocols (Bee inspired, swam intelligence base routing algorithm will study in detail in chapter 7) in homogeneous wireless and mobile ad-hoc networks with resource-constrained nodes.
References


Chapter 7
BCO – A Simple, Scalable and Energy Efficient Routing Protocol for WSNs & MANETs

7.1 Introduction

From the cluster based security architecture and communication protocol presented in chapter 6, it is revealed that secure routing scheme based on session keys by implementing symmetric cryptography using symmetric key algorithm with session keys is the Optimized Network Security Solution for Knowledge Networks (ONSS-KN). This model conveys better security and adds to cluster based communication protocols for homogeneous wireless sensor and mobile ad-hoc networks.

Clustering performs a vital role and often adopts to solve the complex problem of science and engineering. Clustering process divides a set of objects into clusters according to some predefined criteria such that objects in the same cluster are more alike to each other. The existing clustering algorithms have been discussed in chapter 6 and the aim of this chapter is to discuss an optimization algorithm called the Artificial Bees Colony Algorithm, to find the optimal solution.

Swarm Intelligence - SI is categorized into a combined problem resolving capability of social animal and as a part of Artificial Intelligence based on analysis of actions of individuals in different de-centralized systems. SI comes out on result of self-organization in which the small elements interconnections build a dynamic global formation that may be considered as intelligence. However, lower level interactions between the components are intended on a basic set of rules that entity of the colony may have to pursue without having the know-how of its overall effects. Individual in the colony can communicate directly or indirectly on the bases of only local information which may have an effect on the global organization of the colony [1]. The basic idea behind the BCO is to construct a multi-agent system - colony of artificial bees motivated
by the bees’ behavior in the nature, which is proficient enough to solve the complex ‘combinational and optimization’ problems successfully.

There is a number of progressive bio inspired methods including Neural Network genetic programming, Evolutionary programming (EP) developed for routing optimization in WSN and MANETs. The Swarm Intelligence based algorithmic techniques; Particle Swarm Optimization (PSO), Bee Colony Optimization (BCO) and Ant Colony Optimization (ACO) are more carried out for provision of energy-aware, loop free and multi-path routing in MANETs and WSN. As it is discussed in detail, there are number of parametric limitations regarding the mobility of node and non-infrastructure kind of systems in sensor and ad-hoc networks protocols [2]. We study in this research the probabilistic performance assessment frameworks and Swarm Intelligence methods (PSO, BCO, ACO) for routing protocols to measure the (i) energy consumption, (ii) routing overhead, and (iii) route optimality, as performance assessment metrics that has been used for wireless and ad hoc routing algorithms. PSO is very effective for load balancing and energy optimization in WSNs whereas, ACO approaches assure for route optimization in MANETs [2,3].

MANET’S is a type of networks that is infrastructure less, self-configured, network of mobile devices connected via wireless links, is identified as Mobile Ad-hoc Network (MANET) [4]. The MANETs are distinct with other wireless networks because of their characteristics such as mobility of nodes, dynamic topology configuration, frequently node failure, distrusted multi-hop forwarding, infrastructure less and limited energy power. The more recommended, efficient, promising and effective routing protocol for MANETs is that requiring comparatively less control overhead, less packet delays, produce high throughput, efficiently adjustable to dynamic topological changes, and optimized power use. These considerable factors of protocol in MANETs improve the productivity and efficiency of the wireless system; although, very hard to keep balance in these conflicting objective. For the optimization of the listed objectives, Swarm Intelligence based meta-heuristics methodology - PSO, BCO, ACO are more proficient than other algorithms in MANETs [2].

Wireless Sensor Networks [5] is a standard architecture of the networks that controls the computation, sensing, and communications potential with small devices. Due
to the topology nature, architecture, and functionality of ad-hoc and wireless networks, Swarm Intelligence approaches are most appropriate for the energy resources optimization and routing related issues in MANETs and WSNs.

Bio inspired, Swarm Intelligence approaches are more competent for ad-hoc and wireless sensor networks because of the following dominated features i.e.[6]:

- Locality of interactions
- Failure backup,
- Scalable performance robustness to failures,
- Losses internal to the protocol,
- Availability of multiple paths,
- Self-organizing behaviors
- Losses internal to the protocol,
- Ability to adopt in a quick and robust manners to topological and traffic changes and component failures,
- Easiness of design and tuning.

The clustering problem is to discover the partitions and in section 7.2 a concise clustering problem is presented.

The forage selection model that tends to come out of social intelligence of honey bee swarms consists of three basic modules described by Karaboga which are presented in section honey bee modeling of section 7.3. Food sources, status of forage and mode of behavior are discussed and defined in mathematical model and algorithmic steps. On bases of clustering problem elaborated in section 7.2 and honey bee general model inspects for it forages behavior we solve the clustering problem using ABC algorithm. Its mathematical representation and algorithm are presented in section 7.3.2.

In the core section of this chapter 7.4 we present the WSNs routing protocols scenarios using ABC and LEACH algorithms that are meant for the networks are not based on global positioning system. Its theoretical and mathematical model shows that LEACH protocol increases the network life-time by managing and handing over diverse roles to the nodes with some drawbacks whereas objective of this comparative simulation was to reduce the energy utilization of the network and this is achieved by applying ABC algorithm an optimized clustering, on least energy usage for the transportation in a network.
The Fitness function is the improved version of cluster formation and cluster head selection part of ABC algorithm and its mathematic representation is presented and discussed in section 7.5 of this chapter. This fitness function will be used in chapter 8, for simulation of iRoutCluster algorithm using MATLAB®.

In summary, the contribution of this chapter is three fold:

1. The development of mathematical model for clustering problem and honey bee model; by contrast of these two representations is derived the solution of clustering problem using Artificial Bee Colony algorithm.
2. Simulation based comparative investigation to conclude that the bee colony model of swarm intelligence based engineering solution is the most suitable for clustering problem as compare to LEACH algorithms.
3. The derivation of theoretical expression for fitness function (Cluster formation) of ABC algorithm which will be simulated and analyzed in chapter 8.

7.2 The Clustering Problem

Clustering is the process of identifying groups or clustering in multi-dimensional data found on certain resembled measures [7]. Usually distance dimension is used to estimate similar pattern and specifically the clustering problem is as following:

Given N objects, assign each object to a K cluster and reduce the ‘Euclidean distances’ between the center of cluster and respective assigned object.

Let \( O = \{O_1, O_2, \ldots, O_n\} \) is a set of \( n \) objects and \( X_{nxp} \) is the profile data matrix, with \( n \) rows and \( p \) columns. Each \( i \)th object is featured by a real-value \( p \)-dimensional profile vector \( x_i \) ( \( i=1, \ldots, n \)), where each element \( x_{ij} \) related- to the \( i \)th real-value feature (\( j = 1, \ldots, p \)) of the \( i \)th object (\( i = 1, \ldots, n \)). Given \( X_{nxp} \), the purpose of a dividing the clustering algorithm is to find a partition \( G = \{C_1; C_2; \ldots; C_k\} \) such that the objects that belong to a cluster are similar to each other, whereas the objects that belongs to different clusters having different pattern. A well-known function used to measure the integrity of a distribution is the total mean-square quantization error (MSE) or the total within-cluster variance as defined as:

\[
\text{Perf}(O,G) = \sum_{i=1}^{n} \min \{ ||Oi - Cl||^2 | l = 1, \ldots, k \} \tag{7.1}
\]
Where \(\|O_i - C_l\|\) pass on the similarity between object \(O_i\) and center \(C_l\). Euclidean distance is the most used similarity metric in clustering process which is derived from the Minkowski metric.

\[
d(x,y) = \left(\sum_{i=1}^{m} |x_i - y_i|^p\right)^{\frac{1}{p}} 
\Rightarrow d(x,y) = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2} \tag{7.2}
\]

The clustering problem is to discover the partition \(G\) that is optimized w.r.t all other possible solutions \(G = \{G^1, G^2, \ldots, G^{N(n,k)}\}\) (i.e., \(G^i \neq G^j, i \neq j\)) where

\[
N(n,k) = \frac{1}{k!} \sum_{g=0}^{k} (-1)^g \binom{k}{g} (k-g)^n \tag{7.3}
\]

With all possible distribution, clustering problem is NP-hard if it goes above to number three.

### 7.3 Artificial Bee Colony Algorithm

The Honey Bee algorithm is the developing technique, which was proposed in 2005, for many engineering applications. Michael has a detailed study of honeybee about the foraging behavior and in the biological features and observed many pattern from honey bee colony such as waggling dance, aggression syndrome that are used for solving optimization problems.

#### 7.3.1 Honey Bee Modeling

The forage selection model that tends to come out of social intelligence of honey bee swarms consists of three basic modules described by Karaboga:

**Food Sources:** A food source worth relies on different factors, such as its energy awareness, closeness to nest, and the simplicity of energy extraction.

**Employed foragers and unemployed foragers:** The employed foragers are related with specific food sources (node), which they are presently “employed”. These carry information with them about these food sources and share.
**Modes of the behavior:** Abandonment of a source and recruitment to a nectar source are the two leading mode of behavior.

Figure 7.1 Hive and the kind of bees involved in the process [7]

Scouts and onlookers are types of unemployed foragers. Scouts explore the nature world near by the nest in search of new food sources, and onlookers stay in nest and search the food on the information provided by the employed foragers.

**Table 7.1** Mapping of Biological Terminology with Network Terminology [7]

<table>
<thead>
<tr>
<th>In Biological Terms</th>
<th>Network Routing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bee</td>
<td>Hello Message</td>
</tr>
<tr>
<td>Food Source (or Flower)</td>
<td>Node</td>
</tr>
<tr>
<td>Nectar</td>
<td>Energy / Power</td>
</tr>
<tr>
<td>Nectar Table</td>
<td>Routing Table</td>
</tr>
<tr>
<td>Waggling dance</td>
<td>Waiting Time</td>
</tr>
<tr>
<td>Elite Site</td>
<td>Cluster Head</td>
</tr>
<tr>
<td>Hive</td>
<td>Control Station (Real / Imaginary Node)</td>
</tr>
</tbody>
</table>
7.3.1.1 Mathematical Model

As other social foragers, bees look for the food sources in such a way that increases the ratio of $E/T$ where $E$ is the energy achieved and $T$ is the time used for foraging. If it is artificial bee swarms, $E$ is comparative to the food sources nectar amount revealed by bees. In a maximization problem, the target is to find the ‘maximum’ of the objective function $F(\theta)$, where $\theta \in R^p$. Let’s assume that $\theta_i$ is the position of the $i$th food source; $F(\theta_i)$ represents the nectar amount of the food source searched at $\theta_i$ and is proportional to the energy $E(\theta_i)$. Let $P(c) = \{\theta_i (c) | i \text{ 1, 2, . . . , } S\}$ where C: cycle, S: number of food sources, represent the population of food sources being visited by bees.

As discussed, the first choice of a food source by an onlooker relies on the nectar amount $F(\theta)$ of the food source. As the $F(\theta)$ raise, the probability of source selected by an onlooker bee will also raise relatively. However, the probability with the food source searched at $\theta_i$ will be elected by a bee can be computed as

$$P_i = \frac{F(\theta_i)}{\sum_{k=1}^{S} F(\theta_k)} \quad (7.4)$$

After inspecting the dances of ‘onlooker’ bees move towards the area of food source placed at $\theta_i$ which determines a neighbor food source. It is understandable from the previous clarification that there are four parameters used in the Bee Colony algorithm: the value of ‘limit’ and the maximum cycle number, the number of food sources and the number of employed bees -S.

7.3.1.2 Steps of Algorithm

The major steps of the algorithm can be described as following:

**Step 1:** Initialize the population of solution. $\theta_i$, $i = 1, 2, \ldots, S$ and evaluate them.

**Step 2:** Produce new solutions for the employed bees, evaluate them and apply the greedy selection process.

**Step 3:** Calculate the probabilities of the current sources with which they are preferred by the onlookers.
**Step 4:** Assign onlooker bees to employed bees according to probabilities, produce new solutions and apply the greedy selection process.

**Step 5:** Stop the exploitation process of the sources abandoned by bees and send the scouts in the search area for discovering new food sources, randomly.

**Step 6:** Memorize the best food source found so far.

**Step 7:** If the termination condition is not satisfied, go to step 2, otherwise stop the algorithm.

**Figure 7.2.** Major steps of ABC Algorithms
7.3.2 The ABC Algorithm for Clustering Problems

As discussed earlier, there is a population of bees in the Bee colony algorithm. Each bee has instructions for a contestant solution (food source) - and respective quality is defined as its ‘fitness’. For the purpose to solve clustering problem bee colony algorithm has been applied which used floating point arrays to code the instruction into cluster centers.

Therefore, if \( X_{nxp} \) is the matrix and \( k \) is number of clusters and \( C = \{ C_1, C_2, \ldots, C_k \} \) of the set of \( n \) objects \( O = \{ O_1; O_2, \ldots, O_n \} \), each contestant solution in the population contains \( p \) times \( k \) cells \( m_{ij} (i \in \{1,\ldots,k\}, j \in \{1,\ldots,p\} ) \) Each group of \( p \) cells that is related to the vector \( m_i \) represents the coordinates of the \( i \)th cluster center. The \( k \) groups of \( p \) cells that construct the vector \( m \) represent the \( k \) cluster centers. A set of \( k \) cluster heads identify an objects division by mapping the cluster space to the division space \( G = \{ G^1, G^2, \ldots, G^{N(n,k)} \} \).

In order to construct a source food position \( \theta^q \) from the current remember \( q \)th source position \( C^q \), the modified ABC algorithm uses the following representation:

\[
\theta_{ij}^q = \begin{cases} 
C_{ij}^q + \Phi_{ij} (C_{ij}^q - C_{ij}^r), & \text{if } R_j < MR \\
C_{ij}^q, & \text{otherwise}
\end{cases}
\]

(7.5)

where \( r \in \{1; \ldots; S_N\} \) is a randomly chosen index, and \( i \in \{1,\ldots,k\}, j \in \{1,\ldots,p\} \). \( \Phi_{ij} \) is a random number between \([-1, 1]\). It controls the production of neighbor food sources around \( C_q \) and represents the comparison of two food positions visually by a bee. Although \( r \) is defined randomly, it has to be different from \( q \). \( R_j \) is randomly chosen real number in the range \([0, 1]\). A control parameter \( MR \) controls whether the element \( C_{ij}^q \) will be modified or not.

In the actual bee colony, the bee who has consumed the food source change its state to scout from employed bee similarly if a scout found a wealthy food, it will change into employed from scout. In line to pretend this activity of real bees, the following approach is adopted. The use of predefined number of cycles (“limit” for abandonment)
is a key control parameter of the ABC algorithm. As the colony’s discoverer the scouts
do not have any assistances while looking for food. ABC algorithm pretended by
producing a position $\theta$ randomly as following:

$$\theta_{ij} = \theta_{j}^{min} + r \text{and}(0,1)(\theta_{j}^{max} - \theta_{j}^{min}); \quad i \in \{1,\ldots,p\},$$
$$j \in \{1,\ldots,k\},$$

By rule any point in $\mathbb{R}^p$ can be taken as a possible option for a cluster center.
Typically it will select the profile matrix $\theta_{min}, \theta_{max}$, where $\theta_{min}$ and $\theta_{max}$ are two vectors
characterizing the minimum and maximum object values found in the data set. Where
$\theta_{j}^{min}$ and $\theta_{j}^{max}$ are the minimum and maximum values of the $j$th object feature.

On the bases of the aforementioned narrative, the steps of the ABC algorithm
used in this research are shown as Figure 7.3.
Figure 7.3. The flow chart of ABC algorithm used for clustering.
Step 1: Initialization
Set the control parameter values. Make the first half of the colony consists of the employed bees and the second half includes the onlookers. Then randomly generate a position for each candidate and evaluate it. Set the current scout number \( s = 0 \).

Step 2: Introduce new food sources discovered by scouts
If \( s > UB \), order the first half of colony, make the bees with worst solution quality as scouts and others as employed bees. Update the scout number \( s \).

Step 3: Employed bees exploitation
Produce new solution for each employed bee by using (7.5) and evaluate it. Then the selection process by using Deb’s method is applied. If the “limit” for abandonment is reached, the employed bee forgets its memory and becomes a scout for exploration. The scouts number \( s = s + 1 \).

Step 4: Scouts exploration
Send each scout into the search area for discovering new food sources randomly by using (7.6). When a new food source is found, evaluate it and the selection process of Deb’s method is applied.

Step 5: Preferences computation for the current food sources
Calculate the probability values of the current food sources with which they are preferred by the onlookers according to Eq. (7.5).

Step 6: Onlookers exploitation
For the onlookers, produce new solutions from the current food sources selected depending on the computed probabilities and evaluate them. Then the selection process by using Deb’s method is applied to update the corresponding employed bee’s memory or the current food sources.

Step 7: Memorize the best position
For each employed bee and scout, if its memorized position is better than the previous achieved best position, then the best position is replaced by it.

Step 8: Check the termination criteria
If the termination condition is not satisfied, go to step 2, otherwise stop the algorithm.
7.4 WSN Routing Protocols using ABC Algorithm and Hierarchal WSN Clustering Approach (LEACH)

In this section we present the WSNs routing protocols scenarios using ABC and LEACH algorithms that are meant for the networks not based on GPS global positioning system [8]. The main operational objective of these algorithms is to extend the life-time of the network by increasing the transmitted data packets by using clustering. The clustering technique of this protocol [8] is based on the LEACH protocol where CHs performs data aggregation activity for respective clusters. The basic theme of data collection and combination is to merge the data collected from different nodes, remove un-necessary items and reduce transferring data to the base station [9]. Cluster heads use CDMA MAC communication with the base station and TDMA MAC in intra-cluster communication.

The main functional difference between the LEACH and the proposed protocol is the election of the cluster heads CH; ABC algorithm is used to perform the selection of cluster head for proposed solution whereas LEACH protocol uses a random selection method.

7.4.1 Theoretical and Mathematical Model

LEACH protocol increases the network life-time by managing and handing over diverse roles to the nodes [10]. LEACH protocol split the network into clusters and chooses the respective cluster heads arbitrarily in distinct time period. CHs are liable for assembling information within the clusters and transfer this information to the base station. However, the nodes utilize low energy as they communicate in a cluster region and cluster heads spend more energy because of their communication with base station. As the base station distributes the role of CH randomly to cluster in each round, consequently, energy consumption equalizes for each node. But there are some drawbacks are in the query.
Firstly, all the nodes of network are capable enough to communicate with the base station.

Following that, as CH elected randomly some partitions may have greater number of nodes than the others. As seen in Figure 8.4.

Thirdly, the distances from cluster heads to base station and to the member nodes are not taken into account.

The last two drawbacks come out by utilizing supplementary energy in the network. The architect of LEACH algorithm has solved these issues by uniformly distribution of the CHs as elaborated in [8, 10, 11].

In LEACH, CHs are elected arbitrarily and each node could be elected as CH by the probability of \( P \) using the equation (7.7).

\[
T(n) = \begin{cases} 
\frac{P}{1-P|\{r \mod(\frac{1}{P})|} & \text{if } n \in G, \\
0 & \text{otherwise}
\end{cases}
\] (7.7)

Where, \( T(n) \) is the threshold value, \( n \) is node, \( r \) is the current round index and \( G \) is a set of unelected CHs of the nodes in the last rounds. There are two operating phases in LEACH [10]. In the first phase of the process, a member node of a cluster directly communicates with its CH and in the second phase, the CH transmits the combined collected data to the base station [8].

**Setup Phase:** In the setup phase, nodes make a decision on either it will be a cluster head or not by using (7.7). After that, cluster heads broadcast themselves with classified messages. Once the nodes receive these messages, they will connect to the closest clusters by sending ‘request messages’ to the CHs. At this stage distances from node to the cluster heads are calculated using strength of classified messages sent to join the cluster to the member nodes. The selection process of CHs requires distance information between the nodes to the base station. The information are recorded in a table as its format is shown in Table 7.2.

This approach adopted a centralized control system (Figure 7.4) deployed at the base station. At the first stage, initialization of the network is configured and at second stage, energy status and distance information between all nodes are gathered.
In the artificial bee colony CH selection, used total of the ‘squares of the distances’ to the base station from cluster heads and to cluster heads from their member nodes for the selection of CHs in such a way that the dissolute energy is proportional to the square of the distance for a transmission process.

Therefore, the distance table (table 7.2) containing the distance values from node to node. This information that is collected at once at initialization phase, is kept safe by the base station; latterly BS use these distance values for the computation of the fitness function.

Table 7.2 Field of distance vector

<table>
<thead>
<tr>
<th>Receiver identity (2-byte/16-bit)</th>
<th>Transmitter identity (2-byte/16-bit)</th>
<th>Distance value (2-byte/16-bit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The objective achieved from this selection is to decrease the energy utilization in a transmission system; research [8] has used the fitness function as in (7.8) and (7.9). The residual energy of each node is computed and forecasted using (7.9) in algorithm at the base station.

\[
F = \sum_{j}^{n} \left[ \sum_{i}^{m_j} d(N_i, CH_j)^2 + d(CH_j, B)^2 \right] \tag{7.8}
\]
Constraints:

\[ E_j \geq (m_j \cdot E_{rx} + E_{tx}) \]

\[ E_{rx} = E_{rx} \cdot k, \quad (7.9) \]

\[ E_{tx} = E_{elec} \cdot k + E_{amp} \cdot k \cdot B^2 \]

In the hypothesis,

- \( d(x,y) \) is the distance between \( x \) and \( y \), where \( i \) is the index of joining node to \( j^{th} \) cluster and \( j \) is the cluster index,
- \( n \) is number of cluster heads, and \( k \) is the size of packet,
- \( m_j \) is number of members of \( j^{th} \) cluster and \( d \) is distance,
- \( B \) is the base station and \( CH \) is cluster head,
- \( E_{rx} \) is required energy to collect a packet, \( E_{elec} \) is transmitter/receiver electronics, \( E_{tx} \) is required energy to send a packet, \( E_{amp} \) is transmitter amplifier.

The process of gathering classified messages is shown in flow chart Figure 7.5.

The selection process of CHs using ABC is different from the LEACH Algorithm. For this, a distance table has been used which contains the calculated distance values between all nodes. Once a deployment procedure of the network has completed, at first each node sends an introductory message to other nodes and finds the distances to other nodes using (7.10) and sends the values to the base station only once after the deployment.
Figure 7.5 Gathering of advertisement messages and calculation of distances [8]

\[ d(m,n) = \alpha \cdot [P_{rx}(m,n)]^{1/2}, \text{ where } \alpha = c \cdot (P_{tx})^{1/2} / 4\pi f \]  \hspace{1cm} (7.10)

Where \( m \) and \( n \) are connecting node index, \( P_{rx} \) is received signal strength, \( P_{tx} \) is transmitted signal power, \( c \) is the speed of light, and \( f \) is the frequency of the transmission route [12].

**Data Gathering Phase:** Each node of the network sends the computed distances to the base station; using these distance values and artificial bee colony algorithm BS selects the CHs at each round. After publications of CHs to the network by the BS, each node communicates and connects to closest CH by sending a membership request. This process performs basic configuration of the network. After that, CHs collect and monitor
the data using TDMA MAC protocol from their member nodes, combine them and send to the base station.

In Figure 7.6, communicating unit (nodes) and a base station have direct communication is point-up for distances of vector $d$.

![Figure 7.6 Data gathering – direct communication [8]](image)

Another example is shown in Figure 7.7 with data collecting by clustering technique, where $n$ is the number of clusters, $m_j$ is the number of links of the $j$th cluster. Vectors $d$ and $b$ represent the distance between CHs and nodes, and the distance between the base and CHs, respectively.

![Figure 7.7. Data gathering using cluster method [8]](image)
7.4.2 Analysis and Result

To assess the functioning of Wireless Sensor Network Clustering using Artificial Bee Algorithm through simulations and comparison with a well-known hierarchical protocol LEACH. In the simulation the similar assumptions and settings have been used for WSNCABC and LEACH algorithms to make reliable comparisons. There has been used uniform distribution to deployed 100 nodes arbitrarily in the area of 500m x 500m. And set the value 5 as a dimension of parameters for ABC Algorithm, it stands for the number of CHs and sets the probability of a node to be a CH to 0.05 for the LEACH algorithm.

In this [10] work, a radio model specification as in [7] is used which is given in Table 7.3. The value of $E_{tx}$ and $E_{rx}$ is calculated by using $d$ distance and the number of messages ($k$ bit) in (11) and (12), respectively.

\begin{align}
E_{rx} &= E_{elec} \cdot k, \\
E_{tx} &= E_{elec} \cdot k + E_{amp} \cdot k \cdot B^2
\end{align}

(7.11) \hspace{2cm} (7.12)

In Figure 7.8, randomly 100 nodes are deployed in the network at area of 500 x 500 m is randomly deployed, whereas BS is sited at the position of X=250 m, Y=575 m. CHs (represents in squares and members are displayed with different marks) chosen using LEACH algorithm. In Figure 7.9 with the similar parameter CHs are chosen using ABC algorithm [8]. As can be seen from the Figure 7.9, cluster heads are more uniformly selected providing that clusters have approximately equal size of regions and energy levels of the nodes are also considered for the selection process. As the aim of this selection is to minimize energy dissipation in the network system, we use the fitness function as in (7.8) and (7.9). The remaining energy levels of each node are calculated and predicted by the algorithm using (7.9) at the base station.

The objective of this comparative simulation was to reduce the energy utilization of the network. This is also achieved by applying ABC Algorithm an optimized clustering, requiring minimum energy consumption for the communication, in a given network.
Figure 7.8 Distribution of Sensor using LEACH Algorithm [7]

Figure 7.9 Distribution of sensor node using ABC Algorithm [7]

Figure 7.10 and 7.11 show the total received items versus remaining energy of the network and the total of rounds. It shows that WSNCABC algorithm evidently expands the life-time of the network over to LEACH algorithms by consuming less energy and that the network is in more communication.
7.5 Fitness function for Artificial Bee Colony

ABC is used to find or settle the CHs where each solution represents an array shown in Table 7.3. Having $k$ item ($k$ represents the number of cluster in the network) in which each item is containing a set of node. The ABC algorithm uses a population of bees to determine the CH with $k$ dimensions in search space.

<table>
<thead>
<tr>
<th>Table 7.3. A Solution Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1$</td>
</tr>
<tr>
<td>$CH_1$</td>
</tr>
</tbody>
</table>

The fitness value stated as the fitness of CHs selection, which is inversely proportional to the total of energy utilization for single round. An equation (7.13) and (7.14) provides the least energy required for a cluster, derivates from the equation (7.13). If it is declared that a definite transfer time is essential for a data package, the calculated energy utilization is the multiplication of the time ($t$) and transmitting power ($P_s$). In the equations,
- \(i\) is the node index and \(m\) is the number of nodes.
- \(d_i\) is the distance between CH and \(i\)th node,
- \(b\) is the distance between the base station and CH,
- \(E\) is the transfer energy of the cluster.

Assume that for different clusters, the computation of least energy utilization emphasis will effect of distances as in equation (7.15) stated total of clusters energy utilization, \(d_{ij}\) is the distance between \(j\)th CH and \(i\)th node, \(j\) is the cluster index, and \(b_j\) is the distance between base station and \(j\)th CHs.

\[
E = \sum_i^m (P_i^s \cdot t) \geq \alpha \cdot \left( \sum_i^m d_i^2 + b^2 \right) \cdot t
\]

(7.13)

If \(w\) is assumed as the multiplication of \(t\) and \(a\), Equation. (7.14) is as

\[
E \geq w \cdot \sum_i^m d_i^2 + b^2
\]

(7.14)

\[
\sum_{j=1}^n E \geq w \cdot \sum_j^n \left( \sum_i^m d_{ij}^2 + b_j^2 \right)
\]

(7.15)

To achieve the aim of least energy utilization, the distances between CH and nodes, and the distances between base station and CHs must have to be considered in the selection process. Since a CH should have sufficient energy level for communication in the recent round as a managing unit, as well the energy level of the contestant nodes is also important for selection.

A contestant CH should have sufficient energy of transmitting the aggregate message (\(E^{TX}\)) to the BS and receiving messages (\(E^{RX}\)) from the nodes. Considering these, the fitness function (\(f^{dist}\)) is stated by Equation (7.16) and its respective constraints are illustrated in Equation (7.9) where energy utilization of \(E^{RX}\) and \(E^{TX}\) are formed as in [13].

The short-forms in the equations are \(i, j, k, m_j, E_j, E^{RX}, E^{TX}, E^{amp}, E^{elec}, \) and \(b\) are node index, cluster index, number of bits of the transmitting message, number of nodes in the \(j\)th cluster, energy level of the \(j\)th node, receive energy, transmit energy, parameter of transmit amplifier, parameter of radio electronics, and distance between base station and the \(j\)th cluster-head, respectively.
\[ f^{\text{dist}} = \left[ w \cdot \sum_{j}^{n} \left( \sum_{i}^{m} d_{ij}^2 + b_i^2 \right) \right]^{-1} \] (7.16)

The ABC protocol [10] using ‘fitness function’ expressed in equations (7.16) and (7.9) illustrates the CHs defined as gateways. For scheduled data collection by CHs, TDMA technique is more suitable for intra-cluster communication. Whereas the CDMA communication technique is best between base station and cluster heads [14, 15, 16].

### 7.6 Chapter Conclusion

The major goal of optimization algorithm is to discover an optimal resolution to a problem, in such a way to full-fill one or more set of constraints. Study of social insects and animals concludes on an introduction of several swarm intelligence based computational model. Researchers have designed the biologically inspired computational optimization methodologies such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Ant Bee Colony. The aim of this chapter is to discuss an optimization algorithm called the Artificial Bees Colony Algorithm which leads for cluster formation module of proposed architecture to find the optimal solution for such specific problem discussed in chapter 1 and 2. After an explanation of the natural foraging behavior of honey bees, the basic Bees Algorithm is described and implemented in order to optimize several benchmark functions. The simulation base outcomes show that the Bees Algorithm is contributing improvements on other optimization methods as per the nature of the problem. Which is a distributed, scalable and energy-efficient bee-inspired routing protocol for wireless sensor networks and mobile ad-hoc networks. Like other SI algorithms, Bee colony is designed with the so called “bottom-up approach” in which the behavior of individual nodes is defined keeping in view the desired network level behavior.

The experimental results show that artificial bee colony delivers superior performance in terms of node deployment, energy utilization and cluster formation (fitness function of ABC algorithm) as compared to most popular cluster based routing protocol LEACH. The important reasons for this behavior of bee colony are: (1) a simple routing agent model, (2) agent-agent communication to discover optimal paths, (3) fixed
size of route discovery agents that not only saves significant amount of energy during their transmission but also makes the algorithm scale to large networks, (4) distributed and decentralized control, and (5) self-organization to make it resilient to external failures.

In next chapter we will perform the design review of optimized security model by simulation and analysis of (1) optimal node distribution using ABC algorithm for the solution of distribution of wireless node deployment, (2) by clustering formation and cluster head selection mechanism using Intelligent Bee Colony Algorithm - \textit{iRoutCluster} and (3) Session Key establishment process.

References


Chapter 8.  
Design Review, Simulation and Performance Analysis of Clustered Network Design – (ONSS-KN)

8.1 Introduction

Formal modeling process and optimum solutions have been described in the last two chapters which unleashed several important insights. These insights are extremely beneficial and can be instrumental in the design of any ad-hoc and sensor network routing algorithm. To this end, we utilize these insights and review the design of proposed protocol to rectify its shortcomings.

Wireless sensors are limited computation and resources devices that are not recommended for complicated asymmetric key algorithms. Whereas ad-hoc networks characteristics are varying link qualities and highly dynamic topology, therefore it is not recommended to deploy central authentication and key management system as network is vulnerable to definite points of attack and failure. Alternatively, the task of certification authority (distributed public key infrastructure) jointly performs by the cluster heads of the network that was divided into clusters as discussed in chapter 6. By this study we conclude that secure routing scheme based on session keys by implementing symmetric cryptography using symmetric key algorithm with session keys is the Optimized Network Security Solution for Knowledge Networks (ONSS-KN). This model conveys better security as compared to previous secure routing algorithm adopted for cluster wireless networks. ONSS-KN add security to cluster based communication protocols - a distributed, scalable and energy-efficient bee-inspired routing protocol for wireless sensor networks and mobile ad-hoc networks illustrated in chapter 7.

Our proposed scheme ONSS-KN architecture contains three modules as presented in chapter 2, 6 and 7 and shown in figure 8.1. Each module have particular task;

I. Wireless Node Deployment and Pre-Defined Key Distribution
II. Cluster Formation/ Cluster Head Selection
III. Session Key Establishment
I - **Wireless Node Deployment and Pre-defined key distribution:** Firstly we deploy the optimal node distribution using modified artificial bee colony algorithm *iRoutCluster* - part I as presented in section 8.2 with sets parameters description and simulation result analysis. This process might be one time or may continue as in our study case, continue deploying process is an efficient energy utilization base would extend the life time of network. These wireless devices should be capable enough of self managing power as deploying in antagonistic infrastructure.

Once the deployment process has completed, put the devices into cluster network with symmetric key that is defined already. We opt the blowfish symmetric algorithm for encryption which is the best for wireless devices, it uses larger S-box, depends on key of 16-round Feistel cipher. It provides 64-bit block size and 32 bits to 448 bits key length.

II - **Cluster Formation / Cluster Head Selection:** Like any swarm intelligence mechanism, ABC is companionable to any phase of a wireless network design and implementation. Deployed devices of network are now ready to form the clusters.

---

**Figure 8.1** Architecture of ONSS-KN – Optimized Network Security Solution for Knowledge Networks
The proposed network clustering protocol (*iRoutCluster* - II) is based on a centralized control algorithm that is implemented at the base station. The base station is a node with unlimited energy supply.

*iRoutCluster* part –II algorithm for cluster formation and cluster head selection is implemented in MATLAB® as following:

This iterative process will form all clusters till covered all the regions of the network.

**III - Session Key Establishment/Routing**: Base station select generate a prime number of points in range that is defined on a stored elliptic curve. The number of point in range defined on elliptic curve is prime field, and it is very hard near to impossible to crack a prime number field that is defined on elliptic curve. It provides a strong security on prime number randomized on prime field mechanism.

Basically base store provides the elliptic curve and unicast the random prime number (session key) with time stamp to cluster head. The cluster head validates or authenticates time stamp value allocated by base station. If the cluster head has confirmed the validation of time stamp value then it will broadcast the session key to all its neighbors.

Algorithm for session key establishment/routing
Section 8.2 presents the deployment of wireless network using Artificial Bee Colony algorithm based *iRoutCluster* – I algorithm and purpose to attain optimized deployment some modification has accomplished on requirement of application. Here are two major alterations made into as in Onlooker Bee phase and Scout Bee phase. Simulation and comparison show that modified covered entire area efficiently and gives faster convergence than others.

In chapter 2 we present the cluster based network model and design of power centric node of cluster network design, in section 8.3.2 we design our purposed network in CISCO Packet Tracer tool to perform its logical and visual verification on IPSec routing protocol. We have deployed 8 power centric intelligent routers which are configured on IPSec routing. It has successfully simulated the network traffic.

In second section of 8.3 *iRoutCluster* part two we present the simulation of *iRouterCluster* – II in MATLAB®. Simulation-based investigations are conducted to determine the optimum number of clusters, energy dissipation and time for message transfer of proposed network. Other two simulations for wireless sensor networks are done using MATLAB® to relate “hop count” with “network size” estimate. This is the way each cluster nodes would relate itself to the other cluster in terms of “hop count” and within a certain “response time”. Then we had validated the network against success rate that is achieved keeping a considerable network size in a cluster.

In the end a comparative analysis between swarm intelligence algorithms (LACH, ABC and PSO) is presented in section 8.4. This is done on the selection of cluster head, location of cluster head, distance between cluster head and non-cluster head node, number of clusters, intra cluster distance and number of slave node.

- BS selects the random number from prime field over curve in the range of prime number.
- BS unicast the information of session key to CH.
- CH verifies BS with time stamp value assigned by BS previously. (except initial round)
- CH broadcast the session key to all its neighbors.
- This procedure repeats for all dusters in the network.
- After expiring session CH has to send request for new session key.
In Summary, the contribution of this chapter is three-fold as follows:

1. Simulation and Analysis on nodes deployment algorithm based on the modified ABC algorithm is investigated. Some new parameters such as forgetting and neighbor factor for accelerating the convergence speed and PM-Probability of Mutant for maximizing the coverage rate are introduced.

2. Followed by the deployment methodology, we utilize the insights of the formal modeling process and make design changes to the initial Artificial Bee Colony design presented in Chapter 8. We also answer the third question raised in Chapter 1 and that is: To come up with the best network design (clustering formation using fitness function discussed in last chapter) that enables the solution to work efficiently and securely.

3. Research analyze and compare the results of simulations to illustrate the performance of protocol both in static as well as mobile networks. We compare it with several adaptive and swarm intelligent algorithms. The results of our simulations clearly show that adopted algorithm meets it intended objectives defined in chapter 1.

8.2 ABC Algorithm in Deployment for Wireless Networks – iRoutCluster-I

The purpose of wireless-node deployment algorithm is to decide an optimal node distribution in an area of interest. So the ABC algorithm could be more appropriate and helpful into the solution of the wireless node deployment problem [1]. For each node in wireless network there is a range which specifies its sensing power and a communication range. These ranges help to control the performance of a deployment mechanism.

In the research there are three assumptions sets for the deployment of wireless node:

1. The Sensing radius of the wireless node represents the sensing capability;
2. All of the nodes are interconnected;
3. All of the nodes can move.
Table 8.1 shows the connection between the problem of wireless node deployment and the respective parameters in ABC algorithm. Results [2] have demonstrated that the Artificial Colony algorithm is more flourishing and performed better than another swarm intelligence algorithm - PSO for all of 30 independent rounds ‘dynamic deployment’ of wireless network.

Table 8.1. Corresponding parameters in ABC algorithm and wireless node deployment.

<table>
<thead>
<tr>
<th>ABC Algorithm</th>
<th>Wireless Node deployment problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution of a food source</td>
<td>Wireless nodes distribution</td>
</tr>
<tr>
<td>M dimension in each solution</td>
<td>M node coordinate</td>
</tr>
<tr>
<td>Fitness of the solution</td>
<td>Coverage rate of interest area</td>
</tr>
<tr>
<td>Lowest fitness</td>
<td>Optimum node distribution</td>
</tr>
</tbody>
</table>

8.2.1 Modified ABC Algorithms and Its Application in Deployment

In wireless nodes deployment, coverage and convergence are two very critical questions. In order to attain the faster convergence speed and higher coverage rate, two alterations are implemented in ‘FNF-BL-ABC’ based approach.

8.2.1.1. Modification in Onlooker Bee Phase.

The initial food source informing mechanism is computed by Xiangyu [1] that the fitness of the neighbor food source may be higher because it is chosen randomly, which means the distribution is not as good as the present one. Moreover, the formulation equation remains the same in entire iterations. Consequently, the bee cannot find the optimum food sources efficiently using original updating mechanism.

For the purpose to speed up whole process, neighbor factor and forgetting factor; two new parameters [3] are defined into equation 8.1.

\[
\nu_{il} = \eta \times x_{il} + \tau \times \text{rand} (-1,1) \times (x_{il} - x_{kl}) \tag{8.1}
\]
Where $\eta$ is the neighbor factor and $\tau$ is the forgetting factor. The neighbor factor $\eta$ can weaken or strengthen the co-relation between current and neighbor food source. The forgetting factor represents the food source memory when looking for the subsequent food source. While using the full information of neighbor $\eta$ and discovering Excellency, the forgetting factor gets decreased dynamically as:

$$\tau = \lambda \times \omega_{\eta} \quad (8.2)$$

The value of neighbor factor $\eta$ is as:

$$\eta = \lambda \times \omega_{\eta} \quad (8.3)$$

The equation 8.2 will be used when the fitness value of neighbor food source is greater than the current, $\lambda < 1$. An equation 8.3 in case when the fitness value of neighbor food source is less than the current one, $\lambda > 1$;

Consider

$$\omega_{\eta} = \omega_4 - \left(\frac{\text{maxcycle} - \text{iter}}{\text{maxcycle}}\right)^{\alpha} \times \omega_4 - \omega_3 \quad (8.4)$$

$$\omega_{\tau} = \omega_1 - \left(\frac{\text{maxcycle} - \text{iter}}{\text{maxcycle}}\right)^{\beta} \times \omega_2 - \omega_1 \quad (8.5)$$

In equation (8.4) and (8.5), $\omega_1$, $\omega_2$, $\omega_3$, and $\omega_4$ are constants. The interval $\alpha$ is $[0.8, 1]$ and $\beta$ is $[1, 1.2]$. The parameter $\omega_{\tau}$ become smaller from $\omega_2$ to $\omega_1$ and $\omega_{\eta}$ upgraded from $\omega_3$ to $\omega_4$, as the iteration increase.

### 8.2.1.2. Modification in Scout Bee Phase.

In the original ABC, the function of scout bee is to discard the remaining unaffected food source by the “threshold” time and randomly generating a food source. In order to achieve the performance of deployment algorithm and high coverage rate the value of “threshold” is more significant. Therefore, alterations have been introduced into ABC algorithm labeled by backward learning [4]. This approach can be described as shown in Pseudocode.

The following equation is used to compute the value of $\text{PM} – \text{‘probability of mutant’}$:

$$\text{PM} = 0.01 + 0.1 \times \left(2 - e^{iter \times \ln 2 / \text{maxcycle}} \right) \quad (8.6)$$
The ‘backward learning’ approach can be defined as

\[ v_{il} = \eta \times x_{w_l} + x_{b_l} + \text{rand} \ (0,1) \times x_{il} \]  \hspace{1cm} (8.7)

Where \( x_w \) and \( x_b \) are the ‘worst’ and ‘best’ classes of food source. It will remain memorized in \( v_i \) if it is better than \( x_i \).

Pseudocode of backward learning method

```
For i = 1:n (n stands for the number of food source)
    If rand < probability of mutant (PM)
        backward learning.
    end
end
```

**8.2.2 Simulation Results**

**8.2.2.1 Parameter Selection for the Proposed Algorithm.**

As for the modified algorithm [1], there are three different sets of parameters:

- \( \omega_i (i = 1, \ldots, 4) \),
- \( \lambda \),
- \( \alpha, \beta \).

For the reason described in [5], research [1] sets \( \alpha = 0.8 \) and \( \beta = 1.2 \). Among all the tests for \( \omega_i \), the best five combinations are

1. \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 1.2 \);
2. \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 1.6 \);
3. \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 2.0 \);
4. \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 2.4 \);
5. \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 2.8 \).

Among these five, the results of combinations 2 and 3 give better convergence speed than the first \( \omega_1 = \omega_3 = 0.2 \) and \( \omega_2 = \omega_4 = 1.2 \) [5], whereas the combination 4 and 5 performed similar [5]. And for the parameter \( \lambda \), from best six sets 1.4 or 0.6; is chosen for a better performance in convergence speed after many experiments.
8.2.2.2. Simulations in an Ideal Room.

In line to get more consistent result, the parameters (number of node, sensing radius) are the same as in [2]. There are 100 mobile wireless nodes in the experiment and the radius of each node \( r \) is 7. The targeted area is \( 100 \times 100 \), colony size is 20 and the max cycle is 10000. Figure 8.2 shows the distribution of wireless node at initial, then after 100 iterations, 500 iterations, 1000 iterations, and 10000 iterations. It can be found from these drawing that the modified ABC mechanism covered the whole area efficiently.

The performance of ABC, FNF-BL-ABC and PSO based deployment algorithm is compared in Figure 8.3 For now, we can find that modification gives faster convergence speed than the other two in Figure 8.3 and Table 8.2.

| Table 8.2 : Deployment results comparison between 3 algorithms at some iterations. |
|-----------------------------|-----------------|-----------------|------------------|-----------------|-----------------|
| Numbe of Iterations         | 100             | 500             | 2000            | 5000            | 10000           |
| FNF-BL-ABC                  | 0.9095          | 0.9740          | 0.9897          | 0.9945          | 0.9971          |
| Original ABC                | 0.9184          | 0.9713          | 0.9826          | 0.9848          | 0.9863          |
| PSO                         | 0.7993          | 0.8097          | 0.8403          | 0.9215          | 0.9431          |
Figure 8.2 Nodes Distribution: (a) initial, (b) 100 iterations, (c) 500 iterations, (d) 1000 iterations, and (e) 10000 iterations.
### 8.3 Design Review and Performance Analysis of Clustering – iRoutCluster-II

The innovative hierarchical clustering mechanism for wireless networks is to retain energy exhaustion of the network at least level using ABC Algorithm which is a latest swarm based heuristic algorithm [6]. In this approach, a protocol using ABC Algorithm attempts to present an ‘optimum cluster organization’ (Selection of CH and respective members) in such an efficient way to make a least energy utilization.

As discussed in previous chapter that more secure, effective and efficient routing protocol for MANETs is one that require comparatively low packet delays, low control overhead, provides high throughput, effectively adjusted to dynamic topological changes, and optimized power utilization.

#### 8.3.1 Data Transmission

Once all nodes obtain respective TDMA slot information, the data transmission process begins. Each node communicates to respective cluster head only in designated
time slot. The nodes positioned close to the cluster head pass on low-energy data signal and reverse to it; transmission energy of each node increases as distance increases between CH and the node. When a cluster head gathers data from nodes, it compresses the data by applying few essential signal processing practices and transmits it to the base station. Considering the fact that the CH has lot of energy load due to high energy information transmission to the BS; that is the key reason of revolving the CHs during the entire network operation.

8.3.2 Cluster Network Design and Simulation

Cisco presents a three-layered hierarchy as the favor approach to network design. A hierarchical network design model offers the complex problem of network design into minor, more manageable problems. Each level in the hierarchy addresses a different set of problems. This helps the designer optimize network hardware and software to perform specific roles.

In the three-layer network design model as shown in figure 8.4, network devices and links are grouped according to three layers Core, Distribution and Access. Each layer of the three-tiered design model may include the following: router, switch, link or combination of these.

**The Core Layer:** The core layer provides an optimized and reliable transport structure by forwarding traffic at very high speeds. In other words, the core layer switches packets as fast as possible. Devices at the core layer should not be burdened with any processes that stand in the way of switching packets at top speed. This includes the following:

- Access-list checking
- Data encryption
- Address translation

**The Distribution Layer:** The distribution layer is located between the access and core layers and helps differentiate the core from the rest of the network. The purpose of this layer is to provide boundary definition using access lists and other filters to limit what gets into the core. Therefore, this layer defines policy for the network. A policy is an approach to handling certain kinds of traffic, including the following:
Routing updates
Route summaries
VLAN traffic
Address aggregation

Use these policies to secure networks and to preserve resources by preventing unnecessary traffic.

**The Access Layer:** The access layer supplies traffic to the network and performs network entry control. End users access network resources by way of the access layer. Acting as the front door to a network, the access layer employs access lists designed to prevent unauthorized users from gaining entry. The access layer can also give remote sites access to the network by way of a wide-area technology, such as Frame Relay, ISDN, or leased lines.

![Cisco network design model](image)

**Figure 8.4-a** Cisco network design model
Figure 8.4 -b Power Centric Node Resides on Core Layer and Distribution Layer security of the Network Infrastructure

A simulation for producing real-time network traffic, Cisco Packet Tracer has adopted that generated test data and cyber-attacks in presence of security intrusion detection systems. The security protocol being used is IPSec on each router in the network. A simulated network design (figure 8.5) is proposed to hold power centric intelligent nodes (routers) within clustered knowledge networks. In real network system, the clusters would be formed such that each town or city would have one intelligent cluster.

Simulation represents in figure 8.5 is one of the clusters of a Network design, which consists of eight intelligent routers with IPsec security protocol configuration. It has successfully simulated the network traffic.
Figure 8.5 Design and Simulation of A cluster – Ad-hoc Mobile and Wireless Network
Figure 8.6 Cluster member – In a Cluster

Figure 8.6 Shows the member of a cluster which has been simulated and these cluster nodes (routers) are connected to other device to receive the data from wireless network devices.

These routers are considered to be placed on core layer of Network infrastructure and a power centric node of the network topology as seen in figure 8.4-b.
8.3.3 Experimental Setup

8.3.3.1 WIFI- WEP Authentication

The following commands are configured on the router:

```
! 
ip dhcp excluded-address 1.1.1.1
! 
ip dhcp pool dhcp_pool
network 1.1.1.0 255.255.255.0
default-router 1.1.1.1
! 
dot11 ssid test
authentication open
guest-mode
! 
interface Dot11Radio0/3/0
ip address 1.1.1.1 255.0.0.0
encryption key 1 size 40bit 0 1234567890
encryption mode wep mandatory
speed 12.0 18.0 24.0 36.0 48.0 54.0 6.0 9.0 basic-1.0 basic-11.0
basic-2.0 basic-5.5
ssid test
station-role root
! 
```

It has been observed that the wireless connections are established between the router and PCs and PC receives an IP address. Ping between each other to confirm the connection is established.
8.3.3.2 Cellular and Wireless Path - Wireless Network

When PDA is connected to both Cellular and Wireless the default gateway of the device is wireless and sends packet via its wireless port, otherwise via Cellular.

Figure 8.8 Cellular and Wireless Path for PDA device
8.3.3.3 Simulation of IPsec- Security Protocol

In this simulation, sent pings or http requests from PC0 to Server0 (10.4.0.2) and Server1 (10.3.0.2) to observe how Router0 and Router2 encrypt and decrypt packets as well as change source and destination addresses on Layer 3.

![Figure 8.9 IPSec Protocol Simulation – Network Design Model](image)

8.3.4 Proposed Solution - Artificial Bee Colony Algorithm (ABC)

In an original bee colony, few activities are performed by dedicated bees which try to increase the stored nectar amount in hive by using the proficient distribution of work task and self-organization."Economic Load dispatch with value-point effect using artificial bee colony algorithm" [7].

Like any swarm intelligence mechanism, ABC is companionable to any phase of a wireless network design and implementation. Although the original version doesn’t perform as per requirement or to achieve high target goals, but a few alterations to fit into formulated problem generate vibrant outputs. This research has adopted such a model,
which has been molded to optimize the total deployment scenario as demonstrated in previous section and later modifying the same algorithm with small parameter changes for clustering and the routing scheme. Therefore it can construct the backbone of a complete wireless network layer protocol suite. The analysis part can be done in MATLAB simulator.

Table 8.3: Parameters of the ABC algorithm as used in.

<table>
<thead>
<tr>
<th>Swarm Size</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of scouts</td>
<td>5%-10% of number of bees</td>
</tr>
<tr>
<td>Number of onlookers</td>
<td>50% of the swarm</td>
</tr>
<tr>
<td>Number of employed bees</td>
<td>50% of the swarm</td>
</tr>
<tr>
<td>Limit</td>
<td>#onlookers. Dimension</td>
</tr>
</tbody>
</table>

8.3.5 Network Model

To conduct experiment, a corresponding discrete model given and data transferred was developed in MATLAB.

In the experiment, we use a model as specified in [8] where radio $E_{elec}$ for transmitting and receiving, and $E_{amp}$ for the amplifier power as defined in Table 8.4. $E_{tx}$ and $E_{rx}$, which has been computed by using the number of messages $-k$ bit and distance $d$ in equation (8.8) and (8.9) shown in Figure 8.10 (a) and 8.10 (b).

\[
E_{rx} = E_{elec} \cdot k \tag{8.8}
\]

\[
E_{tx} = E_{elec} \cdot k + E_{amp} \cdot k \cdot B^2 \tag{8.9}
\]

Table 8.4. Parameter values used for experiment

<table>
<thead>
<tr>
<th>$E_{elec}$</th>
<th>$E_{amp}$</th>
<th>$K$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50nJ/bit</td>
<td>100pJ/bit/m^2</td>
<td>1024 bit</td>
<td>1024 bit/s</td>
</tr>
</tbody>
</table>
In Figure 8.2, a network of 100 nodes is deployed and CHs are elected by using ABC algorithm [9] on the configuration discussed before. We can also see from the Figure 8.2 that clusters are almost equal in size of regions providing that CHs are selected uniformly.
8.3.6 Simulation and Analysis

In this section the results are analyzed for the proposed routing protocol (iRoutCluster II) using the radio communication model for the quantitative analysis.

8.3.6.1 Thresholds for Clustering

Other two simulations for wireless networks are done using MATLAB to relate hop count with network size estimate (figure 8.11). This is the way each cluster of nodes would relate itself to the other cluster in terms of hop count and within a certain response time (figure 8.12). Then figure 8.13 is there to validate the network against success rate that is achieved keeping a considerable network size in a cluster.

![Figure 8.11 Hop Count against Network Size of each cluster](image)
This set of multiple experiments and simulations to test and validate the success rate of response and hop time for various networks within clusters of 15 to 20 nodes has been found enough for getting the desired result that is shown in above figures.

**Figure 8.12** Calculated Response time between two clusters

**Figure 8.13** Success rate achieved keeping limited to a considerable network size within a cluster
8.3.6.2 Optimum Number of Clusters

Figure 8.14 Optimum number of clusters with respect to the cluster head-set size

Figure 8.14 shows the optimum number of clusters with respect to the cluster head-set size. From the graph, the maximum number of clusters is approximately 20 and its corresponding cluster head-set size is 6. As the graph shows, the number of clusters decreases as the number of cluster head-set size exceeds 6. The minimum number of clusters is approximately 7 and its corresponding cluster head-set size is 1. Therefore as the number of clusters increases the cluster head-set size also increases. As a result, bigger number of clusters can manage bigger cluster head-set size whereas smaller number of clusters can manage smaller cluster head-set size. The reason for this observation is that, in the wireless sensor network, if a cluster head-set size is not carefully chosen for its respective number of clusters, during data transmission much burden is put on the cluster head-set nodes. This causes fast exhaustion of the energy of the cluster head set nodes since the nodes depend on their respective batteries power.
Figure 8.15 shows the change in number of clusters with respect to distance from the base station for different number of cluster head-set sizes. In the graph the number of clusters decreases as the distance from the base station increases. This change is the same for the different cluster head set sizes (that is from cluster head set size 1 to cluster head set size 5). This implies that the cluster head set size does not vary with respect to distance but rather the number of clusters. So, as the number of clusters increases, the distance from the base station decreases. This is because, the sensor node is a tiny device and because of its tiny nature it has challenges with its power control unit and communication unit. Due to these limitations, energy consumption should be as efficient as possible to extend the network lifetime. Apparently, when the transmission range from the base station is long, nodes would require high transmission power to reach the base station and also increases the chance of eavesdropping.
8.3.6.3 Parameters to Evaluate the Performance of the Network

Table 8.5: Parameters of wireless sensor network.

<table>
<thead>
<tr>
<th>No of Iterations</th>
<th>Total Packet</th>
<th>Delay</th>
<th>Loss</th>
<th>Throughput</th>
<th>Packet Delivery Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>760</td>
<td>3.2544</td>
<td>7.5</td>
<td>209</td>
<td>0.9901</td>
</tr>
<tr>
<td>500</td>
<td>760</td>
<td>3.2444</td>
<td>7.5</td>
<td>232</td>
<td>0.9901</td>
</tr>
<tr>
<td>1000</td>
<td>430</td>
<td>3.2435</td>
<td>7.5</td>
<td>218</td>
<td>0.9901</td>
</tr>
<tr>
<td>1500</td>
<td>430</td>
<td>3.2391</td>
<td>2.4</td>
<td>233</td>
<td>0.9966</td>
</tr>
<tr>
<td>2000</td>
<td>490</td>
<td>3.2387</td>
<td>3.3</td>
<td>210</td>
<td>0.9973</td>
</tr>
</tbody>
</table>

Figure 8.16 Throughput in ABC as per number of Round
**Figure 8.17** Delay factor in ABC as per number of Round

**Figure 8.18** Packet Lost in ABC as per number of Round
8.3.6.4 Energy Consumption

Figure 8.20-a and 8.20-b illustrates the energy consumption with respect to the number of clusters for various cluster head-set sizes.

From Figure 8.20-a the energy consumption reduces as the number of clusters increases. The optimum variation in the energy consumption ranges between 0 (Joules) and 6 (Joules) when the cluster head size is 1.

From Figure 8.20-b the energy consumption reduces as the number of clusters increases. The optimum variation in the energy consumption ranges between 0 (Joules) and 1.8 (Joules) when the cluster head size is 3. Therefore comparing the two graphs, the energy consumption in Figure 8.20-b is comparatively lower when cluster head size is 3 as compared to Figure 8.20-a when cluster head size is 1. As a result, the bigger the cluster head size the lower the energy consumption during transmission and vice versa. The reason for this is that, if the energy consumption is reduced in the network, the lifetime of the network would be prolonged and more transmissions can take place in the network.
8.3.6.5 Findings

- The cluster size up to 10 to 20 nodes corresponding with intelligent routers in place for getting the desired success rate and response time in monitoring the activities of nodes within a cluster.
- Bigger cluster size can manage bigger cluster head-set size while smaller cluster size can manage smaller cluster head-set size.
- Distance from the base station reduces when there are more clusters; hence less energy is spent during data transmission.
- Energy consumption reduces as the number of clusters increases.
- For these reasons, the number of clusters and cluster head-set size should be carefully chosen to extend the network life time and also control and manage the sensor nodes efficiently within a network.
8.4 Comparative Analysis of Swarm Intelligence Algorithms

Table 8.6. Comparison between LEACH, PSO and ABC

<table>
<thead>
<tr>
<th>Parameters</th>
<th>LEACH</th>
<th>ABC</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection of Cluster Head</td>
<td>CH selection is done randomly. It does not depend upon energy consumption</td>
<td>CH selection is based on the fitness value calculated by each node</td>
<td>CH selection is based on energy consumed by the sensor nodes.</td>
</tr>
<tr>
<td>Location of Cluster Head</td>
<td>CH are non-uniformly distributed i.e. at the edge of cluster</td>
<td>CH are optimally distributed</td>
<td>CH are optimally distributed across the network</td>
</tr>
<tr>
<td>Distance between CH and Non-CH Nodes</td>
<td>Distance between CH and non-CH nodes is not minimized</td>
<td>Minimum distance between CH and non-Ch nodes</td>
<td>Distance between CH and non-CH nodes is minimized</td>
</tr>
<tr>
<td>No. of Clusters</td>
<td>Numbers of clusters does not ensure the expected number of clusters</td>
<td>Ensures the expected number of clusters</td>
<td>It gives the expected number of clusters</td>
</tr>
<tr>
<td>Intra Cluster Distance</td>
<td>It does not give guarantee for minimum intra-cluster distance</td>
<td>Intra-cluster distance is minimum</td>
<td>Achieve minimum intra-cluster distance</td>
</tr>
<tr>
<td>Number of Nodes Alive</td>
<td>15000 – 25000</td>
<td>More than 40000</td>
<td>30000 – 40000</td>
</tr>
</tbody>
</table>

From above comparison it is concluded that LEACH presents significant energy savings and expands the lifetime of network but does not assure that the required or expected quantity of cluster heads is selected and cluster heads are not uniformly placed or deployed across the network space. Also it does not give minimum distance between CH and non-CH nodes. Whereas, in PSO and ABC algorithm Cluster-Heads are optimally distributed. These two algorithms give minimum distance between CH and non-CH nodes. Artificial Bee Colony gives more alive nodes as compared to LEACH and PSO algorithm. So, ABC algorithm is used to increase the Lifetime of WSN. The ABC protocol using ‘fitness function’ expressed in equations (7.16) and (7.9) illustrates the CHs defined as gateways.
8.5 Chapter Conclusion

Finding an optimum solution to implementing security in knowledge networks is very crucial for future developments. Clustering technique leads the knowledge networks towards easy to manage and monitor for faults and intrusive malicious activities that endangers this massive network, standardized measures have been experimented and adopted for better results and future considerations to implement them in real scenarios. IPSec protocol is implemented as it is standardized by Internet Engineering Task Force (IETF) already 70% implemented. Any future development towards knowledge networks would require a strong foundation that could not be outperformed by security threats. It would be thoroughly tested against current penetration techniques employed by intruders. Therefore, security within knowledge networks would add value to knowledge management and business intelligence techniques for further improvements and enhancement. Consumers of knowledge networks need to be assured of secure transmission and storing of their highly valuable data.

Also, after exhaustive testing experimental results of multiple simulations the conclusion has been reached to create clusters of up to 10 to 20 nodes corresponding with intelligent routers in place for getting the desired success in monitoring the activities of nodes within a cluster. In summary when optimum number of clusters with respect to the cluster head-set size was compared, it was observed that bigger cluster size can manage bigger cluster head-set size while smaller cluster size can manage smaller cluster head-set size. Comparing the cluster size with respect to distance from base station, the maximum number of clusters decreases when the distance from the base station increases. As a result, when the distance from the base station is increased, more energy is spent for a distant transmission.
References


Chapter 9. Conclusion and Future Work

This thesis tackles several problems and challenges that are related to embed security in knowledge networks that are often termed as wireless sensor networks, mobile ad-hoc networks (MANET), and semantic web. The ultimate target was to design a secure routing scheme based on session keys by implementing symmetric cryptography using symmetric key algorithm with session keys which is an optimal resolution in such a way to full-fill one or more set of constraints. The studies are dedicated to developing a security mechanism forming a network design in clustered method and power centric node in clustered knowledge networks ensuring the optimal security of the network.

The developed Optimized Network Security Solution for Knowledge Networks provides better security as compared to previous secure routing algorithm adopted for cluster wireless networks and emerging the loading time of node along secrete global and session key cross check criteria.

This chapter summarizes the important findings in the development of these algorithms, the performance evaluation of the algorithms and the performance evaluation of the networks based on these algorithms. Also, suggestions for future work are presented in the reminder of this chapter.

9.1 Summary of Important Findings

The first step of this study was to understand the feasibility of optimal security models and mechanism i.e security challenges, threat, attacks, mitigation plans for wireless sensor networks, Mobile Ad-hoc networks and semantic web technology. The design of robust, effective, and scalable routing protocols for WSNs was a challenging task. It is observed, that the clustering routing algorithms are generally, more suitable to tackle the constraints and the challenges of wireless sensor networks and Mobile Ad-hoc Networks.

From the study of several routing algorithms for WSNs and MANETS, It was found that clustering routing protocols is the only technique to design an efficient and effective
network design. We have proposed a clustered based security model, theoretically analyzed and simulated on several threshold and parameters in MATLAB code, graphically represents the results and draws findings on it.

We have developed a distributed, scalable, energy and performance efficient routing protocol for ad hoc and sensor networks, *iRoutCluster*. The major objective of this dissertation was to engineer a simple – both in terms of processing complexity and communication overhead – routing solution keeping in view the strict constraints posed by WSNs and MANETs. Such a routing algorithm is expected to utilize the available network resources efficiently and hence achieve competitive performance compared with the existing protocols. The major guideline in developing such a solution was to combine the power of resource constrained wireless sensor and Mobile ad-hoc nodes so that they collectively produce the desired network level behavior.

ONSS-KN uses positive features of symmetric key cryptography and cluster based routing. In development phase, Firstly we deploy the optimal node distribution using modified artificial bee colony algorithm *iRoutCluster I*. The ABC algorithm is found more appropriate and helpful into the solution of the wireless node deployment problem and modified version covering entire area efficiently, gives faster convergence than original and PSO swarm intelligence methods. After that blowfish symmetric algorithm is used to provide encryption. It was found that, in encryption and decryption Blowfish is better than the other mentioned algorithms in throughput and power consumption.

Finding an optimum solution to implementing security in knowledge networks is very crucial for future developments. Clustering technique leads the knowledge networks towards easy to manage and monitor for faults and intrusive malicious activities that endanger this massive network, standardized measures have been experimented and adopted for better results and future considerations to implement them in real scenarios. Like any swarm intelligence mechanism, ABC is companionable to any phase of a wireless network design and implementation. We proposed and developed network clustering protocol (*iRoutCluster - II*) is based on a centralized control algorithm that is implemented at the base station. Multiple simulation analysis concludes the range of corresponding nodes with power centric intelligent that is most preferred design for monitoring the activities of devices within a cluster.
IPSec protocol is implemented as it is standardized by Internet Engineering Task Force (IETF) already 70% implemented. Any future development towards knowledge networks would require a strong foundation that could not be outperformed by security threats. It would be thoroughly tested against current penetration techniques employed by intruders Therefore. Security within knowledge networks would add value to knowledge management and business intelligence techniques for further improvements and enhancement. Consumers of knowledge networks need to be assured of secure transmission and storing of their highly valuable data.

An accurate and detailed IDS and IPS are strongly desired to be implemented in the cyber security area. The functionality of network design is verified and validated through some of different modeling techniques and tools. SCADA monitoring system deployed on central nodes in a clustered network, acts as power centric for the traffic coming from different gateways integrating with the tool MARS that includes components of alert correlation, graph reduction, and alert aggregation.

Standard Network layer IPSec security protocol secured the centric node and cluster network communication, as it provided the security scheme of confidentiality, origin authentication, integrity and prevention against replay attacks in addition to error detection and diagnostic capability. Whenever communication data packets are about to depart a cluster network centric node and when getting a destination node, security policies are enforced on endpoints, routers, gateways, firewalls.

In response to the Problem statement, Clustered network design provide power centric node successfully securing the gateway by implementation of the studies and illustrated standards, algorithms, techniques.

9.2 Future work: Suggestion and Directions

Protocol engineering is an iterative process. The design goes through several intermediate stages before it gets finalized. The testing and evaluation process is generally a designer’s choice. Therefore, some network design and secure routing are evaluated through simulations while others are tested on a real test-bed. We believe that observed analysis is not sufficient to get a true image about an optimized security
solution and protocol’s performance in the real world. In this perspective, we recommend following guiding principle that should be followed in the design cycle of any protocol designed for ad hoc networks.

It is concluded that insect colonies provide an ideal metaphor for developing a WSN routing protocol. The resulting protocol – like its natural counterpart – is adaptive, scalable and robust for the loss of individual units. We also identify areas that may be targeted in future. The first and the foremost is an extension of our formal evaluation framework. We believe that the framework – in its existing form – has tremendous potential to evolve into a de-facto formal evaluation framework which can be used in future research. We also conclude that the theory of reliability can be further investigated for evaluation of ad hoc and sensor networks. Last but not the least, Bee Colony Optimization (BCO) metaheuristic is also emerging as a new field in swarm intelligence that can be a focus of future research.

Future work will focus on investigation of the variation in the cluster head-set size for different network conditions. After the simulation model, targets will be on the implementation of ONSS-KN in NS-2 as a separate module so that it could be tested more accurately.
Appendixes

Appendix A.1

Blowfish Algorithm

It is a symmetric (i.e. uses the same secret key for both encryption and decryption) block cipher (encrypts data in 8-byte blocks) that uses a variable-length key, from 32 bits (4 bytes) bits to 448 bits (56 bytes). The algorithm consists of two parts: a key expansion part and a data-encryption part. Key expansion converts a variable key of at least 4 and at most 56 bytes into several sub-key arrays totaling 4168 bytes. Blowfish has 16 rounds. Each round consists of a key-dependent permutation, and a key and data-dependent substitution. All operations are XORs and additions on 32-bit words. The only additional operations are four indexed array data lookups per round. Blowfish uses a large number of sub-keys. These keys must be pre-computed before any data encryption or decryption (Schneier, 1993).

Sharma (2009) evaluated the performance of different symmetric cryptographic algorithms and found out that AES algorithm (Tamimi, 2008) is a very fast algorithm but requires at least 800-byte memory space for lookup tables, DES also uses large lookup tables and its throughput is very less hence weak DES, which makes it an insecure block cipher. RC6 (El-Fishawy, 2007) is a small algorithm, but it is slower than blowfish. It was found that, in encryption and decryption Blowfish is better than the other mentioned algorithms in throughput and power consumption. Thus, analysis concluded that Blowfish has better performance than other general encryption algorithms in term of the battery and time consumption.

Algorithm: Blowfish Encryption

x is a plain text

Pi is an array of string containing hexadecimal values – 32 bits subkeys

XOR is an exclusive OR function
Blowfish Encryption

Divide x into two 32-bit halves: xL, xR
For i = 1 to 16:
  xL = XL XOR Pi
  xR = F(XL) XOR xR
  Swap XL and xR
  Swap XL and xR (Undo the last swap.)
  xR = xR XOR P17
  xL = xL XOR P18
  Recombine xL and xR
Figure A-1  Blowfish Encryption

[Source: http://i.cmpnet.com/embedded/gifs/2003/0308/0308feat2fig1.gif]
Appendix A.2

Elliptic curve

Elliptic curves appear in many diverse areas of mathematics, ranging from number theory to complex analysis, and from cryptography to mathematical physics. An elliptic curve is the set of points that satisfy a specific mathematical equation. The equation for an elliptic curve looks something like this: $y^2 = x^3 + ax + b$ to generate a random curve over a prime field, choose $a, b$ at random from the prime field. Elliptic curves are applicable for encryption, digital signatures, pseudo-random generators and other tasks. They are also used in several integer factorization algorithms that have applications in cryptography, such as Lenstra elliptic curve factorization. We can use geometry to make the points of an elliptic curve into a group.

**Elliptic curves over the rational number field- finite field:**

We first consider polynomial equations of degree 3 with integer coefficients. For example, take

$$y^2 = x^3 - x$$

Here, we consider an easier problem: For a prime number $l$, count the number of solutions in $\mathbb{F}_l = \mathbb{Z}/l\mathbb{Z}$ of the equation modulo $l$. We regard this problem as easy because for a given prime number, finding solutions can be done, in principle, by a finite number of steps of computations. Let us actually find the solutions. We have the following list.

- $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$
  
  $(0, 0), (1, 0)$

- $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$
  
  $(0, 0), (1, 0), (2, 0)$

- $\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$
  
  $(0, 0), (1, 0), (2, 1), (2, 4), (3, 2), (3, 3), (4, 0)$

- $\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$
  
  $(0, 0), (1, 0), (4, 2), (4, 5), (5, 1), (5, 6), (6, 0)$

The reader should pretend as if it were Fermat or Gauss to find the pattern of the number of solutions. As a matter of fact, we have the following.

**Theorem (Gauss).** Let $l$ be an odd prime number

1) If $l \equiv 3 \mod 4$, then the number of solutions in $\mathbb{F}_l$ to the equation $y^2 = x^3 - x$ equals $l$. 

2) If \( l \equiv 1 \mod 4 \), then, as we stated in Proposition 0.2 in Number Theory 1, \( l \) can be written in the form

\[
\ell = a^2 + b^2, \quad a, b \in \mathbb{Z}.
\]

We choose \( a \) as an odd number, \( b \) an even number, and we choose the sign of \( a \) so that we have

\[
a \equiv 1 \mod 4 \text{ if } b \equiv 0 \mod 4,
\]
\[
a \equiv 3 \mod 4 \text{ if } b \equiv 2 \mod 4.
\]

Then, the number of solutions in \( \mathbb{F}_l \) to the equation \( y^2 = x^3 - x \) equals \( l - 2a \).

What we want to note here is that there exists such a law for the number of solutions. The number of solutions could be at random for each prime number.

Let us consider a slightly different equation

\[
y^2 + y = x^3 - x^2.
\]

Similar calculations show the following.

\[
\begin{align*}
\mathbb{F}_2 &= \mathbb{Z}/2\mathbb{Z} \quad (0, 0), (0, 1), (1, 0), (1, 1) \\
\mathbb{F}_3 &= \mathbb{Z}/3\mathbb{Z} \quad (0, 0), (0, 2), (1, 0), (1, 2) \\
\mathbb{F}_5 &= \mathbb{Z}/5\mathbb{Z} \quad (0, 0), (0, 4), (1, 0), (1, 4) \\
\mathbb{F}_7 &= \mathbb{Z}/7\mathbb{Z} \quad (0, 0), (0, 6), (1, 0), (1, 6), (4, 2), (4, 4), (5, 1), (5, 5), (6, 3)
\end{align*}
\]

**Theorem (Eichler).** For a positive integer \( n \), let \( a_n \) be the integer coefficient of \( q^n \) in the power series

\[
q \prod_{n=1}^{\infty} (1 - q^n)^2 (1 - q^{11n})^2 = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 + \cdots = \sum_{n=1}^{\infty} a_n q^n.
\]

Then, for \( \ell \neq 11 \), the number of solutions to the equation \( y^2 + y = x^3 - x^2 \) in \( F_{\ell} \) equals \( \ell - a_{\ell} \).

Note that the above power series is a modular form of weight 2 and level 11. If we write

\[
q \prod_{n=1}^{\infty} (1 - q^{4n})^2 (1 - q^{8n})^2 = \sum_{n=1}^{\infty} a_n q^n
\]

We can prove that the number of solutions to the equation \( y^2 = x^3 - x \) in \( F_{\ell} \) equals \( \ell - b_{\ell} \).

**Appendix A.3**

**Pseudocode for Proposed iRoutCluster II Clustering Algorithm**

\( \alpha \) is time scale factor and \( \beta \) volume scale factor

\begin{verbatim}
Procedure_EnergyEfficient_iRoutCluster_Clustering
\end{verbatim}
Initialization
Generate the initial population of the bees

Selection of the best bee as the queen
Selection of the maximum number of mating flights (n)

Main Phase
\( \text{do while } i \leq n \)
\( \quad \text{Initialize queen spermatheca, energy and speed.} \)
\( \quad \text{Select } \alpha \)
\( \quad \text{do while } \text{energy} > \text{threshold and spermatheca is not full} \)
\( \quad \quad \text{Select a drone} \)
\( \quad \quad \text{If the drone passes the probabilistic condition then} \)
\( \quad \quad \quad \text{Add sperm of the drone in the spermatheca} \)
\( \quad \quad \text{endif} \)
\( \quad \text{Update Speed} \)
\( \quad \text{Update Energy} \)
\( \text{enddo} \)
\( \text{do } j = 1, \text{Size of Spermatheca} \)
\( \quad \text{Select a sperm from the spermatheca} \)
\( \quad \text{Generate a brood by applying a crossover operator between} \)
\( \quad \text{the queen, the selected drones and the adaptive memory} \)
\( \quad \text{Select, randomly, a worker} \)
\( \quad \text{Use the selected worker to improve the brood’s fitness} \)
\( \quad \text{if the brood’s fitness is better than the queen’s fitness then} \)
\( \quad \quad \text{Replace the queen with the brood} \)
\( \quad \text{else} \)
\( \quad \quad \text{if the brood’s fitness is better than one of the drone’s} \)
\( \quad \quad \text{fitness then} \)
\( \quad \quad \quad \text{Replace the drone with the brood} \)
\( \quad \text{endif} \)
\( \text{endif} \)
\( \text{enddo} \)
\( \text{Return The Queen (Best Solution Found)} \)
\( \end Procedure \)

Appendix A.4

Bee Colony Optimization – iRouteCluster - II Algorithm in MATLAB

\%/* Reference Papers from chapter 8 * [13] [14] [15] [16]/
NP=20;
FoodNumber=NP/2
limit=100;
maxCycle=2500;
objfun='Sphere';
D=100;
ub=ones(1,D)*100;
lb=ones(1,D)*(-100);
runtime=1;

GlobalMin=zeros(1,runtime);
Range = repmat((ub-lb),[FoodNumber 1]);
Lower = repmat(lb, [FoodNumber 1]);
Foods = rand(FoodNumber,D).*Range + Lower;
ObjVal=feval(objfun,Foods);
Fitness=calculateFitness(ObjVal);
BestInd=find(ObjVal==min(ObjVal));
BestInd=BestInd(end);
GlobalMin=ObjVal(BestInd);
GlobalParams=Foods(BestInd,:);

for iter=1:
while ((iter <= maxCycle)),
%%%%%%%%%%%%%%%%%%%%%%%% EMPLOYED BEE PHASE %%%%%%%%%%%%%%%%%%%%%%%%
for i=1:(FoodNumber)
    Param2Change=fix(rand*D)+1;
    neighbour=fix(rand*(FoodNumber))+1;
    while(neighbour==i)
        neighbour=fix(rand*(FoodNumber))+1;
    end;
    sol=Foods(i,:);
    sol(Param2Change)=Foods(i,Param2Change)+(Foods(i,Param2Change)-
        Foods(neighbour,Param2Change))*(rand-0.5)*2;
    ind=find(sol<lb);
    sol(ind)=lb(ind);
    ind=find(sol>ub);
    sol(ind)=ub(ind);
    ObjValSol=feval(objfun,sol);
    FitnessSol=calculateFitness(ObjValSol);
    if (FitnessSol>Fitness(i))
        Foods(i,:)=sol;
        Fitness(i)=FitnessSol;
        ObjVal(i)=ObjValSol;
        trial(i)=0;
    else
        trial(i)=trial(i)+1;
    end;
end;
%%%%%%%%%%%%%%%%%%%% Calculate Probabilities %%%%%%%%%%%%%%%%%%%%%
prob=(0.9.*Fitness./max(Fitness))+0.1;
%%%%%%%%%%%%%%%%%%%  ONLOOKER BEE PHASE  %%%%%%%%%%%%%%%%%%%%%%%%

i=1;
t=0;
while(t<FoodNumber)
    if(rand<prob(i))
        t=t+1;
        Param2Change=fix(rand*D)+1;
        neighbour=fix(rand*(FoodNumber))+1;
        while(neighbour==i)
            neighbour=fix(rand*(FoodNumber))+1;
        end;
        sol=Foods(i,:);
        sol(Param2Change)=Foods(i,Param2Change)+(Foods(i,Param2Change)-
        Foods(neighbour,Param2Change))*(rand-0.5)*2;
        ind=find(sol<lb);
        sol(ind)=lb(ind);
        ind=find(sol>ub);
        sol(ind)=ub(ind);
        ObjValSol=feval(objfun,sol);
        FitnessSol=calculateFitness(ObjValSol);
        if (FitnessSol>Fitness(i))
            Foods(i,:)=sol;
            ObjVal(i)=ObjValSol;
            trial(i)=0;
        else
            trial(i)=trial(i)+1;
        end;
    end;
    i=i+1;
    if (i==(FoodNumber)+1)
        i=1;
    end;
end;

ind=find(ObjVal==min(ObjVal));
ind=ind(end);
if (ObjVal(ind)<GlobalMin)
    GlobalMin=ObjVal(ind);
    GlobalParams=Foods(ind,:);
end;

%%%%%%%%%%%%%%%%%%% SCOUT BEE PHASE  %%%%%%%%%%%%%%%%%%%%%%%%

ind=find(trial==max(trial));
ind=ind(end);
if (trial(ind)>limit)
    trial(ind)=0;
    sol=(ub-lb).*rand(1,D)+lb;
    ObjValSol=feval(objfun,sol);
    FitnessSol=calculateFitness(ObjValSol);
    Foods(ind,:)=sol;
    ObjVal(ind)=ObjValSol;
end;

fprintf('Ýter=%d ObjVal=%g\n',iter,GlobalMin);
iter = iter + 1;
end
GlobalMins(r) = GlobalMin;
end;
save all

function ackley

function ObjVal = ackley(Chrom, switch1);
Dim = size(Chrom, 2);
[Nind, NVar] = size(Chrom);
A = 1/Dim;
Omega = 2 * pi;
sum1 = A * sum((Chrom .* Chrom)')';
sum2 = A * sum((cos(Omega * Chrom))')';
ObjVal = -20 * exp(-0.2 * sqrt(sum1)) - exp(sum2) + 20 + exp(1);

function calculate fitness

function fFitness = calculateFitness(fObjV);
fFitness = zeros(size(fObjV));
ind = find(fObjV >= 0);
fFitness(ind) = 1 ./ (fObjV(ind) + 1);
ind = find(fObjV < 0);
fFitness(ind) = 1 + abs(fObjV(ind));

function griewank

function ObjVal = griewank(Chrom, switch1);
[Nind, NVar] = size(Chrom);
nummer = repmat(1:Nvar, [Nind 1]);
ObjVal = sum(((Chrom.^2) / 4000)')' - prod(cos(Chrom ./ sqrt(nummer)))' + 1;

function rastrigin

function ObjVal = rastrigin(Chrom, switch1);
Dim = size(Chrom, 2);
[Nind, NVar] = size(Chrom);
A = 10;
Omega = 2 * pi;
ObjVal = Dim * A + sum(((Chrom .* Chrom) - A * cos(Omega * Chrom))')';

function rosenbrock

function ObjVal = rosenbrock(Chrom, switch1);
Dim = size(Chrom, 2);
\[ [\text{Nind}, \text{Nvar}] = \text{size(Chrom)}; \]
\[
\text{Mat1} = \text{Chrom(:,1:Nvar-1)}; \\
\text{Mat2} = \text{Chrom(:,2:Nvar)}; \\
\text{if Dim == 2} \\
\quad \text{ObjVal} = 100*(\text{Mat2-Mat1}.*\text{Mat1}).^2+(1-\text{Mat1}).^2; \\
\text{else} \\
\quad \text{ObjVal} = \text{sum}((100*(\text{Mat2-Mat1}.*\text{Mat1}).^2+(1-\text{Mat1}).^2)').'; \\
\text{end} \\
\]

\% function schwefel \% function schwefel
\text{function ObjVal} = \text{schwefel(Chrom,switch1)}; \text{function ObjVal} = \text{schwefel(Chrom,switch1)};
\text{Dim}=\text{size(Chrom,2)}; \text{Dim}=\text{size(Chrom,2)};
\text{[Nind,Nvar]} = \text{size(Chrom)}; \text{[Nind,Nvar]} = \text{size(Chrom)};

\% function 7, sum of \(-\text{xi}\sin(\sqrt{\text{abs(xi)}}))\) for \(i = 1:\text{Dim}\)
\(\text{(Dim}=10)\)
\% \(n = \text{Dim}, -500 \leq \text{xi} \leq 500\)
\% \text{global minimum at} \((\text{xi})=(420.9687)\); \text{fmin}=?
\text{ObjVal} = \text{sum}((\text{-Chrom} .* \sin(\sqrt{\text{abs(Chrom))})))'); \text{ObjVal} = \text{sum}((\text{-Chrom} .* \sin(\sqrt{\text{abs(Chrom))})))');

\% function Sphere \% function Sphere
\text{function ObjVal=Sphere(Colony,xd)} \text{function ObjVal=Sphere(Colony,xd)}
\text{S=Colony.*Colony;} \text{S=Colony.*Colony;}
\text{ObjVal=sum(S')}; \text{ObjVal=sum(S')};

**Appendix A.5**

**MATLAB Source Codes thee several graphs**

**MATLAB source code for the graph of Figure 8.14**

\(\text{el}=0.0013*10^{-12}; \text{es}=10*10^{-12}; \text{Ee}=50*10^{-9}; \text{Ebf}=5*10^{-9};\)
\(\text{n}=2000; \text{l}=4000; \text{Nf}=10000; \text{M}=100;\)
\([\text{s}]=[(1:8)];\)
\(\text{d}=150;\)
\(\text{f1}=\text{el}*(\text{d}^4)-\text{Ebf}.*((2*\text{s}-1)*\text{Ee};\)
\(\text{c}=((n/(2*3.14))^2*(0.5))^{((\text{es}/\text{f1}).^2*(0.5)))^{\text{M}};\)
\text{plot(s,c);} \text{plot(s,c);}
\text{xlabel('Cluster head-Set Size', 'FontSize',18)} \text{xlabel('Cluster head-Set Size', 'FontSize',18)}
\text{ylabel('Number of Clusters', 'FontSize',18)} \text{ylabel('Number of Clusters', 'FontSize',18)}

**MATLAB source code for the graph of Figure 8.15**

\(\text{el}=0.0013*10^{-12}; \text{es}=10*10^{-12}; \text{Ee}=50*10^{-9}; \text{Ebf}=5*10^{-9};\)
\(\text{n}=2000; \text{d}=110; \text{l}=4000; \text{M}=100;\)
\([\text{s,d}]=\text{meshgrid}([1:1:5],[150:10:250]);\)
MATLAB source code for the graph of Figure 8.20-a and 8.20-b

```matlab
el=0.0013*10^-12; es=10*10^-12; Ee=50*10^-9; Ebf=5*10^-9;
n=2000; M=100; d=100; l=4000; Nf=10000;
[c]=[1:1:20];
s=1;
Ech_elect=l.*(Ee.*((n./c)-s+1)+es*d^2);
Enon_ch_elect=l.*(Ee.*(1+c)+es*d^2);
Ech_frame=l.*(((n./c)-s+1)*Ebf+el*d^4+((n./c)-s+1)*Ee);
Enon_ch_frame=l.*(Ee+es*((M^2)/((2*3.14159).*c)));
Ech_data=(1./((n./c)-s+1)).*(Nf./c).*Ech_frame;
Enon_ch_data=((n./c)-s)./((n./c)-s+1)).*(Nf./c).*Enon_ch_frame;
Ech_iter=Ech_elect+Ech_data;
Enon_ch_iter=Enon_ch_elect+Enon_ch_data;
Eround=(Ech_iter/s)+(((n./(c*s))-1).*Enon_ch_iter)./((n./c)-s);
subplot(1,2,1);plot(c,Eround);
xlabel('Number of Clusters', 'FontSize',14)
ylabel('Energy Consumption(J)', 'FontSize',14)
```

```matlab
el=0.0013*10^-12; es=10*10^-12; Ee=50*10^-9; Ebf=5*10^-9;
n=2000; M=100; d=100; l=4000; Nf=10000;
[c]=[1:1:20];
s=3;
Ech_elect=l.*(Ee.*((n./c)-s+1)+es*d^2);
Enon_ch_elect=l.*(Ee.*(1+c)+es*d^2);
Ech_frame=l.*(((n./c)-s+1)*Ebf+el*d^4+((n./c)-s+1)*Ee);
Enon_ch_frame=l.*(Ee+es*((M^2)/((2*3.14159).*c)));
Ech_data=((n./c)-s)./((n./c)-s+1)).*(Nf./c).*Ech_frame;
Enon_ch_data=((n./c)-s).*((Nf./c).*Enon_ch_frame;
Ech_iter=Ech_elect+Ech_data;
Enon_ch_iter=Enon_ch_elect+Enon_ch_data;
Eround=(Ech_iter/s)+(((n./(c*s))-1).*Enon_ch_iter)./((n./c)-s);
subplot(1,2,2);plot(c,Eround);
xlabel('Number of Clusters', 'FontSize',14)
ylabel('Energy Consumption(J)', 'FontSize',14)
```