Design of Resilient Heterogeneous Wireless Networks

by

Ozgur Kabadurmus

A dissertation submitted to the Graduate Faculty of Auburn University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Auburn, Alabama
August 3, 2013

Keywords: Survivability, Reliability, Heterogeneous wireless networks, Network design, Optimization

Copyright 2013 by Ozgur Kabadurmus

Approved by

Alice E. Smith, Chair, Professor of Industrial and Systems Engineering
Chase C. Murray, Assistant Professor of Industrial and Systems Engineering
Alvin S. Lim, Associate Professor of Computer Science and Software Engineering
Abdullah Konak, Associate Professor of Information Sciences and Technology
Abstract

With the growing use of new telecommunications technologies such as 4G and wireless hotspots, heterogeneous wireless networks (HetNets) are gaining more attention. The source of heterogeneity of a HetNet can either be the differences in nodes (such as transmission ranges, failure rates and energy levels) or the differences in services offered in the network (such as GSM and WiFi). Quality of service (QoS) is an issue for users while the cost of the infrastructure is an issue for the network provider. In telecommunications network design problems, survivability and reliability are well known QoS metrics. Most previous studies considered survivability and reliability as constraints (vertex connectivity or edge-disjoint paths), while other papers used traditional reliability metrics (such as two-terminal reliability or all-terminal reliability). In this dissertation, a new metric that combines network reliability with network resilience in capacitated networks is devised. Exact and approximate methods to evaluate this capacitated resilience metric are formulated and solved. Capacitated resilience is used to solve HetNets network designs ranging from 10 to 150 users. Results are compared to the popular reliability and survivability metrics in the literature. It is shown that networks designed by this new measure are significantly different than other network designs. This metric is the first to consider rerouting under capacity constraints in the instance of failure and thus reflects more realistic practice. It is also computationally tractable for use during network design.
Acknowledgments

I owe my deepest gratitude to my advisor, Dr. Alice E. Smith, for her encouragement, guidance, and support throughout my PhD process. Also, I would like to thank my committee members, Dr. Abdullah Konak, Dr. Chase C. Murray and Dr. Alvin Lim for their support. They gave me great suggestions and ideas throughout the dissertation process. I would also like to thank Dr. Jeffrey S. Smith for providing me research opportunities. I have learned a lot. I would also like to thank Dr. Bülent M. Durmuşoğlu, my former academic adviser from Istanbul Technical University, for encouraging me to start M.S. degree and seek career in academia.

Finally, but perhaps most importantly, this dissertation would have not been possible without the unquestioning support, sacrifice, and patience of my family and friends. I’d like to thank my dad (İsmail), my mom (Mürvet) and my brother (Tolga) for their love and encouragement. Graduate school is a long process, and it would have been even longer and lonelier without the support of Fatmanur Karaman. She was always there to motivate me whenever I needed it the most. I also would like to thank my friends Ramiz Boy, Gökhan Özden, Özgür Özmen, Eren Sakınç and Burcu Özkülhancı for their invaluable friendship.
# Table of Contents

Abstract ................................................................. ii  
Acknowledgments ......................................................... iii  
List of Figures ............................................................ viii  
List of Tables .............................................................. xv  
List of Abbreviations ................................................... xvii  
List of Notation .......................................................... xix  
1 Introduction ............................................................ 1  
  1.1 Wireless network types ............................................ 2  
  1.2 HetNets ............................................................. 3  
  1.3 Problem definition ................................................ 6  
  1.4 Primary contributions ............................................ 11  
2 Literature Review ..................................................... 14  
  2.1 Heterogeneous wireless networks ................................. 14  
    2.1.1 Nonsensor heterogeneous wireless networks ................ 14  
    2.1.2 Heterogeneous wireless sensor networks .................... 19  
  2.2 Survivability/reliability measures ............................... 26  
    2.2.1 Two-terminal reliability ..................................... 26  
    2.2.2 Probability that two nodes can communicate ............... 32  
    2.2.3 Expected loss of traffic (ELT) ............................. 33  
    2.2.4 Traffic efficiency ........................................... 34  
    2.2.5 k-connectivity and edge-disjoint paths .................... 36  
    2.2.6 Summary and discussion ..................................... 36  
2.3 Shortest path problem variations: Constraint shortest path and k shortest path 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.1</td>
<td>Summary and discussion</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>Some issues of wireless networks</td>
<td>44</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Reliability of a wireless link and network density</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>Capacity of a wireless network and interference issues</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>Proposed metric: Capacitated resilience</td>
<td>51</td>
</tr>
<tr>
<td>3.1</td>
<td>The need for a new metric</td>
<td>51</td>
</tr>
<tr>
<td>3.2</td>
<td>Proposed metric: Capacitated resilience</td>
<td>52</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Finding the most reliable path between the user to an access point</td>
<td>53</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Identifying alternative paths between the user and any access point</td>
<td>54</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Determining independent subgroups of the alternative paths</td>
<td>57</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Reliability calculation of independent subgroups</td>
<td>58</td>
</tr>
<tr>
<td>3.2.5</td>
<td>Calculating reliability of the alternative paths (resilience factor)</td>
<td>60</td>
</tr>
<tr>
<td>3.3</td>
<td>An example network</td>
<td>60</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Capacitated resilience calculation</td>
<td>60</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Connectivity measures</td>
<td>66</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Two-terminal and all-terminal reliability</td>
<td>66</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Traffic efficiency</td>
<td>66</td>
</tr>
<tr>
<td>3.3.5</td>
<td>ELT</td>
<td>66</td>
</tr>
<tr>
<td>3.3.6</td>
<td>Summary</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>Proposed Method and Solution Approach</td>
<td>68</td>
</tr>
<tr>
<td>4.1</td>
<td>Single objective Evolutionary Strategies (ES) model</td>
<td>69</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Pseudocode of ES model</td>
<td>70</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Encoding</td>
<td>71</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Population initialization procedure</td>
<td>73</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Edge selection algorithm</td>
<td>73</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Calculation of cost, reliability and capacitated resilience</td>
<td>74</td>
</tr>
</tbody>
</table>
4.1.6 Mutation operators ......................................................... 75
4.1.7 Replacement of population with children .......................... 77
4.1.8 Stopping criteria .......................................................... 77
4.1.9 Penalty functions ......................................................... 77
4.2 Bi-objective Evolutionary Strategies (ES) model ....................... 78
4.2.1 Definitions: “Non-dominated rank”, “Crowding distance” and “Pareto optimality” .......................................................... 78
4.2.2 Pseudocode of ES model ................................................ 81
4.2.3 Global Pareto front operations ....................................... 82
4.2.4 Multiobjective sorting .................................................. 84
4.3 Calculation of other survivability and reliability metrics ............ 86
4.3.1 \(k\) and all-terminal reliabilities .................................... 86
4.3.2 Traffic efficiency (TE) ................................................... 87
4.3.3 \(k\)-connectivity ......................................................... 89
5 Preliminary Results ............................................................. 90
5.1 Single objective vs. bi-objective ES .................................... 91
5.1.1 Pareto front diversity .................................................... 91
5.1.2 Comparison of non-dominated solutions for bi-objective and single objective Pareto fronts .................................................. 92
5.1.3 Single and bi-objective Pareto Fronts ............................... 96
5.2 Equalizing computational effort ......................................... 97
5.2.1 Pareto front diversity .................................................... 97
5.2.2 Comparison of non-dominated solutions for bi-objective and single objective Pareto fronts .................................................. 101
5.3 Variation of single objective ES ........................................... 105
5.4 Summary .......................................................... 105
6 Results .......................................................... 110
6.1 Comparison of single (capacitated resilience) and bi-objective (cost and capacitated resilience) ES .......................... 110
6.2 Correlation among capacitated resilience and other metrics .................. 111
6.3 Network structure: Comparison of capacitated resilience and other metrics . 118
   6.3.1 A problem instance of the 10 user scenario .......................... 119
   6.3.2 Another problem instance of the 10 user scenario ................. 132
   6.3.3 A problem instance of the 25 user scenario ....................... 138
   6.3.4 Summary of the differences of the network designs .............. 146
6.4 Bi-objective (TE and capacitated resilience) ES .......................... 161
6.5 Solution time comparison of capacitated resilience and other metrics .... 162
   6.5.1 Solution time comparison of capacitated resilience: Effect of number
        of paths and cut-set size ..................................... 162
   6.5.2 Solution time comparison of capacitated resilience: Problem size ... 176
   6.5.3 Solution time comparison of all metrics ............................ 177
7 Conclusions and Further Research ........................................... 180
   7.1 Conclusions .......................................................... 180
   7.2 Further research .................................................... 182
Appendices ................................................................. 185
A Input data of problems ..................................................... 186
   A.1 The 10 user scenario ................................................. 186
   A.2 The 25 user scenario ................................................. 188
   A.3 The 50 user scenario ................................................. 191
   A.4 The 75 user scenario ................................................. 196
   A.5 The 100 user scenario ............................................... 204
   A.6 The 150 user scenario ............................................... 214
List of Figures

1.1 A wireless mesh network design (from Amaldi et al. (2008)) ............................. 3
1.2 A HetNet design (from Yang et al. (2007)) ...................................................... 5
1.3 A HetNet design (from Capone et al. (2010)) .................................................... 6
3.1 An example network .............................................................................................. 61
3.2 Independent subgroup 1 of the network in Figure 3.1 ........................................... 62
3.3 Independent subgroup 2 of the network in Figure 3.1 ........................................... 62
3.4 Independent subgroup 3 of the network in Figure 3.1 ........................................... 63
3.5 Capacitated version of the example in Figure 3.1 (c and f denote device capacity and user flow requirement, respectively) ................................................................. 65
4.1 Crowding distance calculation ................................................................................ 81
5.1 Comparison of Pareto front solutions (Problem instance 1) ................................. 93
5.2 Comparison of Pareto front solutions (Problem instance 2) ................................. 94
5.3 Comparison of Pareto front solutions (Problem instance 3) ................................. 95
5.4 Pareto fronts for problem instance 1: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds) .................. 98
5.5 Pareto fronts for problem instance 2: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds) ....... 99

5.6 Pareto fronts for problem instance 3: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds) ....... 100

5.7 Comparison of Pareto front solutions (Problem instance 1) ................. 101

5.8 Comparison of Pareto front solutions (Problem instance 2) ................. 102

5.9 Comparison of Pareto front solutions (Problem instance 3) ................. 103

5.10 Variation of the single objective ES for different budget values ............ 106

5.11 Variation of the single objective ES for different budget values (5000 generations, 10 random number seeds) ............................................... 107

6.1 Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 10 user scenario ............ 112

6.2 Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 25 user scenario ............ 113

6.3 Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 50 user scenario ............ 114

6.4 Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 75 user scenario ............ 115
6.5 Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 100 user scenario

6.6 Inputs of 10 users scenario (problem instance 1): User locations ($x$ and $y$ are the coordinates, $f$ is the traffic flow requirement)

6.7 Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for Capacitated Resilience with unconstrained capacities

6.8 Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for Capacitated Resilience with unconstrained capacities

6.9 Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for Capacitated Resilience with constrained capacities

6.10 Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for Capacitated Resilience with constrained capacities

6.11 Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for TE

6.12 Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for TE

6.13 Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for two-terminal reliability

6.14 Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for two-terminal reliability

6.15 Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for all-terminal reliability
6.16 Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for all-terminal reliability .......................... 134

6.17 Summary of the designs of the problem instance 1 of the 10 user scenario, budget=500 ......................................................... 135

6.18 Summary of the designs of the problem instance 1 of the 10 user scenario, budget=600 ......................................................... 136

6.19 Inputs of the 10 user scenario (problem instance 2): User locations ($x$ and $y$ are the coordinates, $f$ is the traffic flow requirement) ........................................ 137

6.20 Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for Capacitated Resilience with unconstrained capacities 139

6.21 Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for Capacitated Resilience with constrained capacities 140

6.22 Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for TE ............................................ 141

6.23 Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for two-terminal reliability ......................... 142

6.24 Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for all-terminal reliability ......................... 143

6.25 Summary of the designs of the problem instance 2 of the 10 user scenario, budget=600 ......................................................... 144

6.26 Inputs of the 25 user scenario (problem instance 1): User locations ($f$ is the traffic flow requirement) ........................................ 145
6.27 Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for Capacitated Resilience with unconstrained capacities

6.28 Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for Capacitated Resilience with unconstrained capacities

6.29 Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for Capacitated Resilience with constrained capacities

6.30 Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for Capacitated Resilience with constrained capacities

6.31 Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for TE

6.32 Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for TE

6.33 Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for two-terminal reliability

6.34 Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for two-terminal reliability

6.35 Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for all-terminal reliability

6.36 Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for all-terminal reliability

6.37 Summary of the designs of the problem instance 1 of the 25 user scenario, budget=1000
6.38 Summary of the designs of the problem instance 1 of the 25 user scenario, budget = 1200 .......................................................... 158

6.39 Pareto fronts for the 10 user scenario problem instance 1 with budget constraint of 500: Bi-objective optimization of TE and capacitated resilience .................. 163

6.40 Pareto fronts for the 10 user scenario problem instance 1 with budget constraint of 600: Bi-objective optimization of TE and capacitated resilience .................. 164

6.41 Network structure of the 10 user problem (problem instance 1, budget = 500) found by bi-objective optimization for capacitated resilience and TE (high TE and low capacitated resilience case) ............................................ 165

6.42 Network structure of the 10 user problem (problem instance 1, budget = 500) found by bi-objective optimization for capacitated resilience and TE (low TE and high capacitated resilience case) ............................................. 166

6.43 Network structure of the 10 user problem (problem instance 1, budget = 500) found by bi-objective optimization for capacitated resilience and TE (medium TE and medium capacitated resilience case) ................................. 167

6.44 Comparison of the designs presented in Figures 6.41 through 6.43 .................. 168

6.45 Solution time per iteration (one ES generation) comparison of single objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) .................................................. 170

6.46 Solution time per iteration (one ES generation) comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) .................................................. 171
6.47 Capacitated resilience comparison of single objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) ................................................................. 172

6.48 Capacitated resilience comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) ................................................................. 173

6.49 Number of evaluated alternative paths per iteration (one ES generation) comparison of single-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) ................................................................. 174

6.50 Number of evaluated alternative paths per iteration (one ES generation) comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each) ................................................................. 175
List of Tables

2.1 Summary of objective functions used in the nonsensor heterogeneous wireless network literature .................................................. 15
2.2 Summary of decision variables used in the nonsensor heterogeneous wireless network literature ............................................... 16
2.3 Summary of parameters used in the nonsensor heterogeneous wireless network literature ...................................................... 17
2.4 Summary of optimization methods used in the nonsensor heterogeneous wireless network literature ........................................ 17
2.5 Summary of survivability/reliability measures used in the nonsensor heterogeneous wireless network literature .................. 18
2.6 Summary of objective functions used in the heterogeneous wireless sensor network literature .................................................. 20
2.7 Summary of decision variables used in the heterogeneous wireless sensor network literature ..................................................... 21
2.8 Summary of parameters used in the heterogeneous wireless sensor network literature ......................................................... 22
2.9 Summary of constraints used in the heterogeneous wireless sensor network literature ......................................................... 23
2.10 Summary of optimization methods used in the heterogeneous wireless sensor network literature ........................................ 23
2.11 Summary of survivability/reliability measures used in the heterogeneous wireless sensor network literature .................. 24
2.12 Comparison of reliability/survivability metrics ................................................................. 37
4.1 Input parameters used in the ES model .................................................. 69
4.2 Decision variables used in the ES model .................................................. 70
4.3 Model parameters used in the ES model .................................................. 71
5.1 Performance of single and bi-objective ES for three problem instances \dotfill 96

5.2 Comparison of Pareto solutions of single (1000 generations) and bi-objective (16,000 generations) ES for three problem instances \dotfill 104

5.3 Mean and standard deviation of capacitated resilience for single and bi-objective runs (10 problem instances of 10 user scenario with 10 random number seeds for each problem instance. Capacitated resilience is calculated for number of alternative paths=10 and cut set size=4.) \dotfill 109

6.1 Correlations calculated using the best solutions of 10 problem instances of the 10 user scenario \dotfill 117

6.2 Summary of all metrics for the 10 user scenario, problem instance 1 \dotfill 132

6.3 Summary of all metrics for the 10 user scenario, problem instance 2 \dotfill 138

6.4 Summary of all metrics for the 25 user scenario, problem instance 1 \dotfill 159

6.5 Solution time comparison of single (capacitated resilience) and bi-objective (cost and capacitated resilience) ES for different problem sizes \dotfill 176

6.6 Solution time comparison of all metrics for the 10 and the 25 user scenarios \dotfill 178

6.7 The largest solvable scenarios in 24 hours, given as number of users for a medium budget \dotfill 179
List of Abbreviations

AP    Access Point
BS    Base station
CPU   Central processing unit
CR    Capacitated resilience
ELT   Expected loss of traffic
ES    Evolutionary strategies
GA    Genetic algorithms
GPU   Graphics processing unit
HetNet Heterogeneous wireless network
LAN   Local area network
MANET Mobile ad hoc network
MAP   Mesh access point
MC    Mesh client
MG    Mesh gateway
MH    Mobile host, mobile handset
MR    Mesh relays
QoS   Quality of service
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>RN</td>
<td>Relay node</td>
</tr>
<tr>
<td>RP</td>
<td>Relay point</td>
</tr>
<tr>
<td>SN</td>
<td>Sensor node</td>
</tr>
<tr>
<td>TCP</td>
<td>Transport control protocol</td>
</tr>
<tr>
<td>TE</td>
<td>Traffic efficiency</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless local area network</td>
</tr>
<tr>
<td>WMAN</td>
<td>Wireless metropolitan area network</td>
</tr>
<tr>
<td>WMN</td>
<td>Wireless mesh network</td>
</tr>
<tr>
<td>WSN</td>
<td>Wireless sensor network</td>
</tr>
</tbody>
</table>
### List of Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G$</td>
<td>Graph</td>
</tr>
<tr>
<td>$V$</td>
<td>Set of vertices (nodes)</td>
</tr>
<tr>
<td>$E$</td>
<td>Set of edges</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of nodes</td>
</tr>
<tr>
<td>$m$</td>
<td>Number of edges</td>
</tr>
<tr>
<td>$U$</td>
<td>Set of users</td>
</tr>
<tr>
<td>$r_j$</td>
<td>Range of device $j$</td>
</tr>
<tr>
<td>$c_j$</td>
<td>Capacity of device $j$</td>
</tr>
<tr>
<td>$f_i$</td>
<td>Traffic requirement of user $i$</td>
</tr>
<tr>
<td>gridX and gridY</td>
<td>The limits of the area in which nodes are deployed</td>
</tr>
<tr>
<td>budgetLimit</td>
<td>Budget limit for installing devices</td>
</tr>
<tr>
<td>maxDevice</td>
<td>Maximum number of devices</td>
</tr>
<tr>
<td>$N_p$</td>
<td>Population size</td>
</tr>
<tr>
<td>$N_c$</td>
<td>Number of children</td>
</tr>
<tr>
<td>$P_m$</td>
<td>Probability to mutate</td>
</tr>
<tr>
<td>maxGen</td>
<td>Maximum number of generations</td>
</tr>
<tr>
<td>maxNonImprovingGen</td>
<td>Maximum number of non-improving generations</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>The mutation rate for coordinates</td>
</tr>
<tr>
<td>successRate</td>
<td>A parameter to adjust $\sigma$ in ES</td>
</tr>
<tr>
<td>sigmaAdjuster</td>
<td>Adjustment rate of $\sigma$ in ES</td>
</tr>
<tr>
<td>penaltyUnassignedUser</td>
<td>Penalty value if a user is not assigned to a device</td>
</tr>
<tr>
<td>penaltyInfeasibleRouting</td>
<td>Penalty value if a device cannot connect to the wired backbone</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Reliability and survivability of telecommunication networks is a popular area of optimization. With the growing use of wireless services, e.g. 3G/4G and wireless hotspots, the topic has assumed more importance. The needs of users are increasing as the Internet is becoming the new source of entertainment. Bandwidth requirements of users have increased dramatically with the available multimedia services and even more bandwidth will be demanded in the future due to the increasing number of communication devices (such as cell phones, laptops and tablets). Connection speed, coverage, session continuity and reliability are the main concerns of the users. The competitive structure of the telecommunication industry forces service providers to invest more in their infrastructure to satisfy user demands. With the emergence of new technologies, service providers try to combine different solutions to serve customers better. This creates the network design problem.

In the literature, most of the efforts were to evaluate the reliability or the survivability of a network. However, design of the network is very important for service quality. In general, a telecommunication network design problem is to minimize the cost of a network while ensuring some quality of service constraints. The network can be any type, such as wireless or fiber optic. In this dissertation, wireless telecommunication networks are considered. In a typical wireless communication network, users connect to an intermediate node; then intermediate nodes connect to an end node (e.g., a base station - usually connected to a wired backbone). While “session continuity” is a concern for users, the cost of the infrastructure is an issue for the network provider. Although networks vary, the requirement is always the same: coverage and reliability/survivability at a low cost. Network type does not affect this requirement but the related properties, constraints and assumptions vary.
This dissertation defines and investigates “Capacitated Resilience” (CR) in heterogeneous wireless networks (HetNets). The structure is mesh type networks, but the methodology presented in this dissertation can be easily extended to other wireless networks (such as sensor networks). Capacitated resilience, a new metric proposed in this research, is related to reliability and it will be explained in detail in Section 3. Before defining capacitated resilience, the basic wireless network types (mesh, sensor and ad hoc networks) and HetNets are summarized.

1.1 Wireless network types

A wireless mesh network (WMN) consists of “interconnected mesh access points (MAPs), relays (MRs) and gateways (MGs) in which mesh clients (MCs) connect to MAPs to access the Internet and MGs act as bridges between the wireless infrastructure and the Internet while MRs relay the traffic” (Benyamina et al., 2011). According to Amaldi et al. (2008), MRs and MAPs are often fixed and electrically powered. Also, MAPs are connected to a wired link, such as local area network (LAN), DSL or fiber. Wireless network devices in a mesh network have an organizational hierarchy (Amaldi et al., 2008). An illustration of the structure of a mesh network is given in Figure 1.1. In this structure, users connect to a MR and MRs connect to a MAP which is connected to the wired backbone.

A wireless sensor network (WSN) consists of “low-cost, low-power and energy-constrained sensors to monitor a physical phenomenon” (Guidoni et al., 2010). Sensors are aimed to monitor an area and gather information, however, they have limited ranges and they are usually battery operated with limited energy levels. Generally, relay nodes (RNs) are used to transmit information from sensor nodes (SNs) to a server node (or an end node).

Another type of wireless networks is “ad hoc networks”. According to Moraes et al. (2009), an ad hoc network has transceivers that can receive and transmit packets (information) by using multiple consecutive wireless links. Mobile ad hoc networks (MANETs) are wireless networks in which the nodes (transceivers) are mobile. Unlike mesh type networks,
where the network has a hierarchy, MANETs are self configuring networks without a fixed infrastructure.

### 1.2 HetNets

HetNets are an integration of different kinds of wireless networks. Chen et al. (2007) state that telecommunication service providers, such as Verizon, Sprint and T-Mobile, are integrating or planning to integrate multiple wireless technologies with partially overlapped coverage areas. One example is to offer wireless LAN access for their 3G/4G customers. However, to benefit from these services, the users (mobile hosts) must be equipped with one or more wireless access technologies. In this case, a mobile host can choose WiFi in one location and 3G/4G in another location because of different cost rates, bandwidth or coverage properties. Because customer satisfaction is closely related to the quality of service (QoS) level and better service usually means more investment, a service provider should optimize cost by employing the best combination of available heterogeneous wireless technologies (Chen et al., 2007). Niyato and Hossain (2009) claim that the future of wireless networks is to provide seamless mobility to users with the integration of different wireless access technologies. For example, they foresee the integration of 802.16 based metropolitan area
networks (WMANs) and 802.11 based wireless local area networks (WLANs) in the near future. This integration is an example of a HetNet. They cite load balancing and preventing network congestion as important issues of this integration.

Pei et al. (2010) state that research on integration of third-generation (3G) mobile systems (and, currently, 4G service) and WLAN systems are gaining popularity. According to the authors, the main reason for this popularity is to satisfy more diversified QoS needs of users. Although 3G systems can provide more bandwidth and economic revenues, WLAN technologies are widely used and demanded by customers. The important issue is to select an appropriate wireless access technology to have service continuity (Pei et al., 2010).

Yang et al. (2007) give an example of HetNets in Figure 1.2. In this example, the authors introduce ad hoc network technology on top of a cellular system to increase system performance. If a request from mobile handset 2 (MH2) is blocked due to limited bandwidth on base station 1 (BS1), traffic diversion station 2 (TDS2) reroutes the flow to BS2. The new path becomes MH2-TDS2-TDS1-BS2. In this type of system, the traffic diversion station can connect either base stations or mobile handset devices. In other words, the traffic diversion stations and mobile handset devices can use both ad hoc technologies and cellular network technologies. Yang et al. (2007) refer to the IEEE 802.11 protocol for their ad hoc interface. They also mention that all base stations in their example (Figure 1.2) operate at a cellular network frequency.

Yang et al. (2007) assert that destination nodes are unknown and this is the main difference between HetNets and wireless mobile ad hoc networks in which the destination nodes are known. In HetNets the destination can be any node, as long as the bandwidth requirements (or other QoS requirements) are met. Hence, this involves a unique issue for HetNets - the destination selection problem. In this problem, the best suitable base station must be selected to satisfy QoS requirements. However, many important issues must be addressed to solve it. In HetNets, users can choose from variety of wireless access technologies (such as, a cellular network or a WLAN) with different channel characteristics (such as bandwidth,
loss or delay). Another issue is that requirements differ due to the different service providers that users interact with, because each service provider offers different services. Yang et al. (2007) state that this routing problem is NP-hard due to the heterogeneity of the system.

Another example of HetNets is given in Figure 1.3. In this design, a sink node gathers all of the information sent from sensor nodes. It is basically a heterogeneous wireless sensor network. Gateways are added to the system to balance load and increase network life time (because sensors are battery powered). Since gateway devices are more expensive than sensors, the objective is to minimize the number of gateway devices without decreasing the QoS level (different QoS metrics are covered 2.2). In this example, the sink node is connected to a direct geographical link, such as a wired backbone or a satellite connection. However, instead of connecting the gateways with these expensive technologies, they are connected to each other with wireless links such as ZigBee, WiFi or WiMax. The authors state that the connection of gateways constitutes a wireless mesh network and sensors are connected to this mesh backbone. Gateways can be fixed or mobile and finding the best location of gateways is an issue. In this example, sensor to gateway and gateway to gateway connections may use different wireless technologies. The structure of this example by Capone et al. (2010) is very similar to the one of Yang et al. (2007).
1.3 Problem definition

A telecommunications network can be defined as a graph $G$ consisting of vertices (nodes) of users and network devices. Edges are the connections between users (or devices) and network devices. To be connected with a network device $v$, a user or network device must be within the communication range of $v$. The communication range is determined by the technical specification of the network device. Throughout this dissertation, the sets of vertices (devices or users) and edges (wireless or wired links) of $G$ will be denoted as $V$ and $E$, respectively.

Design of survivable/reliable heterogeneous wireless networks is a new area of optimization which has applications in mesh and sensor networks. With the growing use of new telecommunications technologies such as 4G and wireless hotspots, this subject is gaining more attention. The source of heterogeneity of a HetNet is the difference in services offered in the network (such as 3G/4G and WiFi). Some authors in the literature also use the term heterogeneity as the differences in nodes (such as transmission ranges, failure rates and energy levels). In this dissertation, the research problem includes both different nodes and different services in the wireless network. The network structure will be similar to the ones in Figures 1.2 and 1.3. In this type of network, users connect to a gateway (relay node) and gateways connect to a sink node (e.g., base station) using potentially different wireless
technologies. In a study for homogeneous wireless networks (Amaldi et al., 2008), a similar structure for mesh networks was presented and the network design problem was solved to minimize the installation cost of the network under full coverage constraints. However, the heterogeneous wireless network presented by Capone et al. (2010) integrates different wireless technologies and different nodes (with different properties), and it will be used in this dissertation as a base structure of the proposed model.

The main focus of this dissertation is the resilience of a telecommunications network. Capacitated resilience is related to reliability and survivability; however, it has not yet been considered in survivable networks. Survivability/reliability of heterogeneous wireless networks has gained attention recently and the two important network designs are mesh and sensor type networks. Most studies consider survivability/reliability as a constraint (vertex connectivity or edge-disjoint paths). For example, Kashyap et al. (2010) minimized the number of relay nodes in a sensor network under $k$-connectivity constraints. Benyamina et al. (2011) worked on a reliable mesh network design problem with $k$-connectivity constraints. The remaining papers on optimization of heterogeneous wireless networks do not consider survivability/reliability. Among them, Amaldi et al. (2008) proposed a mathematical model to minimize the cost of a mesh network without considering survivability/reliability. Benyamina et al. (2009b) worked on the same problem proposed by Amaldi et al. (2008); however, they used bi-objective optimization (minimizing cost and maximizing network throughput) without survivability/reliability constraints. In another similar work, Benyamina et al. (2009a) proposed a bi-objective algorithm for the minimum gateway placement problem in mesh networks (without survivability/reliability) to minimize cost and congestion of gateways. Benyamina et al. (2009b) and Benyamina et al. (2009a) only considered different nodes in the networks as the source of heterogeneity. In this dissertation, the integration of different wireless technologies are also considered. More importantly, link capacity for rerouting is included in the context of survivability/reliability.
In studies using traditional reliability calculations, cost is minimized under a minimum reliability constraint. For example, Dengiz et al. (2002) used Formulation 1.1 where $c_{ij}$ is the cost to use edge $x_{ij}$. In this formulation, “all-terminal reliability” of a solution is required to exceed a predefined threshold.

$$\min z = \sum_{i=1}^{N-1} \sum_{i=1+1}^{N} c_{ij}x_{ij}$$

s.t.

$$R(x) \geq R_0$$

The calculation of all terminal reliability is simply the multiplication of the reliable components. To demonstrate, a simple formulation from Grover (2004) is shown in Equation 1.2. As two-terminal reliability includes only two nodes in the network, all terminal reliability includes all node pairs in the network.

$$R(G, s, t, p) = \sum_{i=0}^{m} N_i(G, s, t)p^i(1 - p)^{m-i}$$

where $\{s, t\} \in E$, $p$ = fixed failure rate of edges,

$$N_i(G, s, t)$$ is the number of operating states

This reliability calculation is nonlinear due to the multiplication of decision variables $x_{ij}$ (1 if edge $(i, j)$ is selected). However, it can be linearized by using the logarithm of reliability as explained in Equation 3.4 in Section 3.2.1. In this dissertation, “capacitated resilience” metric is used instead of reliability because capacitated resilience includes reliability and also considers rerouting of flows when a failure occurs. Rerouting is very important to ensure survivability of the network. The motivation to use capacitated resilience instead of reliability is explained in detail in Section 1.4.

In this dissertation, a telecommunication network consists of users and devices; and capacitated resilience is defined at the user level. “User” can represent an individual or
total traffic requirements in an area. Equation 1.3 shows the capacitated resilience of User \( i \) \((U_i)\). \( R(p_i) \) denotes the reliability of the assigned path of User \( i \) where the user is assigned to a device whose path has maximum reliability (from the user to an AP). The resilience factor scales the user reliability. To calculate it, all alternative paths from User \( i \) to any available access point are identified first. Then, reliabilities of the alternative paths are calculated. These steps are explained in detail in Section 4.1.4. Capacity of devices and links are considered for finding initially assigned and alternative paths.

\[
\text{Capacitated resilience} = R(p_i) \times \text{Resilience Factor} \quad (1.3)
\]

After finding the user capacitated resistances (which is computationally the most expensive part of the capacitated resilience calculation), network level capacitated resilience is calculated easily according to Equation 1.4. Network capacitated resilience calculation is simply the weighted average of user level capacitated resistances in terms of user traffic requirements. In this equation, \( w_i \) is the weight of User \( i \), which is calculated as flow of User \( i \) / total flow of users. Both network and user level capacitated resistances are between 0 and 1 because network level capacitated resilience is the weighted average of user level capacitated resistances which are found by scaling the reliability of the assigned paths \( R(p_i) \), where \( 0 \leq R(p_i) \leq 1 \).

\[
\text{Capacitated resilience} = w_i \times \text{Capacitated resilience} \ (U_i) \quad (1.4)
\]

Similar to a reliability calculation, the capacitated resilience calculation is nonlinear. Nevertheless, this calculation can be handled with metaheuristic search methods, such as Evolutionary Strategies (ES) and genetic algorithms (GA), because of the flexibility of those methods. For heuristic search methods, an important issue is to use penalty functions to accommodate the capacitated resilience constraint in the objective function to eliminate
infeasibility issues. In this dissertation, new aspects of survivability/reliability (with capacitated resilience) are investigated. A practical yet effective solution methodology has been developed to minimize the cost (maximize capacitated resilience) of a network for a required capacitated resilience level (budget). The details of the solution method are covered in Chapter 4.

Typical objective functions used in previous studies are cost (Shahnaz and Erlebach, 2010), energy consumption (Moraes et al., 2009) and lifetime of the network (Yang et al., 2009). In general, the objective functions of survivable/reliable networks are either cost or reliability. In models where the objective function is cost, reliability/survivability is considered as a constraint (either edge-disjoint paths or $k$-connectivity). If reliability is the objective, then cost is usually constrained to an upper bound.

In addition to single objective optimization, two objective functions are optimized simultaneously in this dissertation: cost and capacitated resilience. Pareto optimality is used as the bi-objective optimization method. Benyamina et al. (2009a) and Benyamina et al. (2009b) used bi-objective optimization for mesh networks, but they did not consider survivability/reliability. Capone et al. (2010) minimized cost and energy consumption, but they used a simple weighted additive objective function and they did not consider survivability/reliability. Therefore, this research fills a gap by using cost and a survivability/reliability metric that considers edge and node capacities in bi-objective optimization of heterogeneous wireless networks.

In this dissertation, the decision variables are the number, types and locations of the devices, user to device assignments, and routing decisions from users to APs (see Section 4.1 for more details). In addition, to distinguish this research from others, realistic features are included in the optimization model. These include variable transmission ranges of devices and the presence of relays in the network. Although the problem becomes more complicated with the addition of different components, the solution becomes more meaningful for real life applications.
1.4 Primary contributions

The primary objectives of this dissertation are to present a new and valuable survivability metric, capacitated resilience, for wireless heterogeneous networks and to develop effective methods for solving realistic heterogeneous wireless network design problems where the objective functions are cost and capacitated resilience. Models with one objective function (cost or capacitated resilience) have been developed to present a solution approach. As there are two metrics to optimize, Pareto optimality is also used in a biobjective model. Using penalty functions for capacitated resilience helps the heuristic optimization method move from the infeasible region to the feasible region easily (see Section 4.1.9) during search. Realistic size problems instances with many (more than 100) users and devices are solved to assess the proposed survivability metric.

The main hypothesis of this research is that a network design with better allocation of redundancies for rerouting options considering node capacities can be obtained by using capacitated resilience instead of using traditional reliability metrics or $k$-connectivity/edge disjoint paths constraints. Network designs using capacitated resilience, traditional reliability constraints and connectivity constraints are compared to verify this hypothesis in this dissertation (Section 6.3). Capacitated resilience enables more realistic operation because it considers rerouting possibilities in case of a failure. Traditional reliability measures (such as two-terminal/all-terminal reliability, traffic efficiency, probability that two nodes can communicate and expected loss of traffic), do not directly consider rerouting options. Survivability constraints, such as $k$-connectivity and edge disjoint paths, consider rerouting options; however, they allocate redundant paths to the network without consideration of reliability or capacity. They assume the components (nodes or edges) are perfectly reliable. The redundancy provided by $k$-connectivity and edge disjoint paths increases the chance of a user remaining connected to the network but it causes more slack capacity and higher costs. On the other hand, capacitated resilience also includes redundancy by considering rerouting
and leads to designs with less, but more effectively distributed, slack capacities. More on this discussion is presented in Section 2.2.6.

This research provides contributions to the survivable/reliable wireless telecommunication networks field in several ways:

- First, capacitated resilience as a survivability/reliability measure is the main contribution of this dissertation. As previously explained, in the survivable/reliable wireless networks literature, survivability/reliability was often considered as a constraint of edge-disjoint paths (or $k$-connectivity). Some studies used traditional reliability measures, e.g. two-terminal reliability or all-terminal reliability, as objective functions. Survivability/reliability has been similarly studied in the heterogeneous wireless networks literature. However, wireless networks (heterogeneous or homogeneous) require a better design approach considering rerouting of flows (in case of a failure). Therefore, a different metric other than reliability is needed. Capacitated resilience is devised because it includes both reliability and rerouting alternatives. The capacitated resilience metric is applicable not only to telecommunications networks, but also to other network types, such as transportation networks. It can also be used in military applications where instantaneous changes can affect the success of an operation. Both exact calculation and estimation methods are presented for capacitated resilience calculation.

- A second contribution is to present a bi-objective method to optimize HetNets for cost and capacitated resilience. In most real life applications, cost is the main limitation to achieve higher QoS levels that are demanded by users. Bi-objective optimization with Pareto optimality allows decision makers to minimize cost and maximize capacitated resilience by selecting the most appropriate solution from a set of non-dominated solutions.

- A third contribution is the comparison of the network designs obtained by optimization of different reliability/survivability metrics. Specifically, network designs optimized for
capacitated resilience are compared with designs optimized for other metrics. The specific network designs of capacitated resilience and the other metrics are identified and their effectiveness are discussed.
Chapter 2
Literature Review

In this chapter, the previous work in literature is summarized. The main focus is on network types, survivability/reliability measures and wireless communication issues.

2.1 Heterogeneous wireless networks

Throughout this dissertation, the term “heterogeneous wireless network” refers to a heterogeneous wireless telecommunications network. As explained before, the source of heterogeneity can be different services used in a wireless network or differences among the properties of nodes.

In this section, heterogeneous wireless networks, including sensor and nonsensor networks, are summarized. In this dissertation, nonsensor HetNets are termed as heterogeneous wireless networks and do not have sensor nodes. The reason for this classification is the large and specific literature of sensor networks.

2.1.1 Nonsensor heterogeneous wireless networks

In nonsensor HetNets various objective functions have been used to optimize a given network. Among them, cost and throughput are the widely used ones. The total number of nodes is a variation of cost. Overhead is a well known metric that defines the extra data that is sent with the actual communication data to route it over a network. It is an important metric for routing protocols. Availability is defined by Choi et al. (2011) as “the probability that a system or service is available at a specific time”. Another objective function, redundancy ratio, is the ratio of number of backup systems to working systems. Connected dominated sets are defined in detail in Tiwari et al. (2007) and it can be briefly
defined as a set of vertices $V' \in V$ in which a vertex $i \in V'$ can connect to another vertex $j \in V'$ by using only vertices of $V'$ and every other vertex of $V - V'$ is adjacent to a vertex $j \in V'$. Therefore, this metric is related to the connectivity of a network. Energy consumption of nodes has also been an objective. Relays help connect users with an end node, for example a base station. They are generally used to increase coverage by providing multi-hop connections with a lower cost. Table 2.1 summarizes the objective functions used in the literature.

Table 2.1: Summary of objective functions used in the nonsensor heterogeneous wireless network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overhead</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Throughput</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redundancy ratio</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of connected dominated set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Energy consumption</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td># of relay nodes (RNs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Many different decision variables have been used in nonsensor HetNets literature as shown in Table 2.2. Shahnaz and Erlebach (2010) used Steiner subgraphs, where a Steiner tree spans all the nodes in a graph $G$. Similarly, a Steiner subgraph spans a given set of nodes. As another decision variable, transmission power of a node affects its range, with higher power yielding a larger range. To find the best routing from users to a base station, “edges to select” was used as a decision variable. Number of backup systems is simply the number of additional systems other than the working ones. The location of nodes has a
direct impact on reliability of a network. Nonoverlapping wireless channel assignment is used to prevent interference (see Section 2.4.1 for more information on interference).

Table 2.2: Summary of decision variables used in the nonsensor heterogeneous wireless network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission power of each node</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Edges to connect nodes</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of backup systems for control systems</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relay locations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Wireless channel assignments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

In the optimization process, many variables have been fixed. For example, number of nodes, transmission ranges of nodes, network density, and number of users are common parameters. Network density is one such parameter and is defined by Tiwari et al. (2007) as the number of nodes per unit area. Choi et al. (2011) used node weight as a parameter. It determines the priority of a node to be secured against failures. Similar to this dissertation, traffic requirements of nodes are fixed in So and Liang (2009). The parameters used in heterogeneous wireless networks are given in Table 2.3.

Various methods have been used to optimize a nonsensor HetNet, including mixed integer programming, approximation algorithms and custom algorithms (Table 2.4).

As discussed in Chapter 1, instead of using survivability/reliability measures, edge-disjoint paths and \( k \)-connectivity constraints are usually used in the nonsensor HetNets literature (Table 2.5).

Some studies focused on the architecture of HetNets. For example, Birkos et al. (2012) proposed ROMEO (Remote collaborative real-time multimedia experience over the future Internet) architecture for live 3D video delivery networks. This architecture aims for seamless
Table 2.3: Summary of parameters used in the nonsensor heterogeneous wireless network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different transmission ranges of nodes</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network density</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Unidirectional/bidirectional edges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Weight of the nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic requirements of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 2.4: Summary of optimization methods used in the nonsensor heterogeneous wireless network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximation algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Custom algorithm</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIP</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Queuing model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

video session continuity over heterogeneous networks consisting of WiFi, LTE (Long-term evolution, usually referred to 4G LTE) and DVB (Digital video broadcasting) for mobile and roaming users. This dissertation will not cover the technical aspects of these technologies.
Table 2.5: Summary of survivability/reliability measures used in the nonsensor heterogeneous wireless network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two edge-disjoint paths</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k-connectivity</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No survivability/reliability (only connectivity)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Summary and discussion

In the above listed studies either $k$-connectivity or edge-disjoint path constraints are used for survivability/reliability. In this research, “capacitated resilience” is maximized while cost is minimized.

Another main difference of these nonsensor HetNet studies with this dissertation is the level of comprehensiveness of the models. Each of these studies captured a specific aspect of the problem. For example, Dilmaghani and Rao (2007) worked on a possible disaster scenario. In their problem, they deployed a mesh network to increase throughput and reduce overhead. They used Transport Control Protocol (TCP) for their mesh network and observe the bottlenecks in a real application. However, they did not use any optimization approach. The main focus of Choi et al. (2011) was the redundancy of devices. They solved the problem as a queuing system ($M/M/c$), which restricts their model with unrealistic assumptions, for example, exponentially distributed failures and repairs. Also they did not consider the path of communication links from nodes $i$ to $j$, which is considered in this dissertation as a decision variable. Shahnaz and Erlebach (2010) considered a “node-weighted graph” having different nodes with different costs. However, their problem was to find a Steiner tree and connect all nodes together. Similarly, Tiwari et al. (2007) worked on connected subgraphs. Although
their idea of a node-weighted graph is similar to the one studied in this dissertation, their connectivity assumption does not apply here because connectivity of a user to a base station (or an end node) is required in this dissertation. In other words, users are not required to be connected to each other in this dissertation, but they must be connected to an AP. Ad hoc networks were studied by Moraes et al. (2009) with a focus to find optimum power assignments of nodes. Instead of locating the nodes, they had device positions as an input. It is a different problem because positions of nodes are optimized in this dissertation.

2.1.2 Heterogeneous wireless sensor networks

Although this dissertation does not include sensor networks, some of the studies are related to the HetNets and therefore this section summarizes the relevant wireless sensor network literature. There have been many studies aimed at optimizing heterogeneous wireless sensor networks. In the literature, an important objective function is cost while others are total energy consumption or lifetime of the network. The number of additional nodes (either additional relay or sensor nodes) is minimized in some studies. In addition, transmitted communication packets and latency are two important performance metrics for sensor networks in which realtime information is crucial. Latency is the time delay in data communications (Guidoni et al., 2010). In one study (Wang et al., 2008), the number of high transmission power state nodes was minimized to reduce total power usage, which is an important issue for battery powered sensors. Also, Machado et al. (2010) minimized vacancy which is defined as the area outside of the sensing region. Al-Turjman et al. (2013) used $k$-connectivity as the performance metric. The objective functions of the selected literature are summarized in Table 2.6.

The decision variables used in heterogeneous wireless sensor networks have been various. The number (and/or location) of relay nodes has been used in many studies. Also, number of additional sensor nodes is another decision variable that is used in some studies. Similarly, the number (and/or location) of base stations has been used as a decision variable. The
Table 2.6: Summary of objective functions used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of relay nodes</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Lifetime of network</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy consumption</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmitted communication packets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of a node can securely deliver info.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of high transmission power state nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifetime of a node (related to energy consumption)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. and Coef. Var. of capacity utilization of RNs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of linear programming models have been applied in the algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of additional sensor nodes (SNs)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU time</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k-connectivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

decision variable “edges to select” was used to route information from sensors to a base station with flow on each edge (and edges to select) as a decision variable. The decision variables of the selected literature are given in Table 2.7.

Many parameters have been used in the heterogeneous wireless sensor networks literature. Number of nodes, especially sensor nodes, is one of the most frequently used parameters in optimization models. Transmission ranges of sensor and relay nodes are also used in many studies. Grid (or region) size is used as a parameter for initial deployment of sensors or other nodes. As an indicator of network activity of sensor nodes, number of broadcast messages was used by Machado et al. (2010). The network activity parameter is important because it
Table 2.7: Summary of decision variables used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># and (or) locations of RNs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td># and (or) locations of base station (BS) nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># and locations of (additional) SNs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edges to use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binary if SN connected to BS and RN</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow from SNs to BS and RN</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmission power of each node</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow of a commodity from node $i$ to $j$</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

directly affects the power usage of the sensor nodes. Table 2.8 summarizes the parameters used in the heterogeneous wireless sensor networks literature.

Many of the constraints in heterogeneous wireless sensor networks are related to either link or device capacity. Qian et al. (2007) used a different connectivity constraint, key connectivity, which affects only less powerful sensor nodes. They defined key connectivity as the probability that less powerful (limited ranged) sensors can connect to network. Lifetime of a network was defined in Al-Turjman et al. (2013) as the time period that the sensor network operates. The constraints previously used are summarized in Table 2.9.

To optimize heterogeneous wireless sensor networks different methods have been used. Among them, traditional optimization methods (e.g., mixed integer programming), custom algorithms and approximation methods are the common ones. Also, a metaheuristic search method, genetic algorithms, has been used. One study used Markov chains to solve the heterogeneous wireless sensor network design problem (Machado et al., 2010). Simulation is another common tool to verify solutions. Custom mixed integer programming methods
Table 2.8: Summary of parameters used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of SNs</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of other nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of edges</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmission ranges of RNs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmission ranges of SNs</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic load</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial energy levels of SNs</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grid/region size</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average distance between nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction of wireless communication (one-way, two-way)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network activity level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr. of a SN to be active</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bidirectional/unidirectional links</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(such as Lagrangian relaxation) have also been also used. Table 2.10 summarizes methods that have been used to optimize heterogeneous wireless sensor networks.

Two/$k$-edge disjoint paths and $k$-connectivity are the most commonly used constraints to ensure survivability/reliability in heterogeneous wireless sensor networks. As a variation of connectivity constraints, Guidoni et al. (2010) worked on small world concept, in which there are some shortcuts that reach from a node $i$ to a farther node $j$. These shortcuts are obtained by using more powerful devices (yielding a larger range). As a variation of all-terminal reliability, Qian et al. (2007) used probability that a node can securely deliver information to ensure survivability in wireless sensor networks. To ensure secure communication between nodes, they proposed assigning different communication keys to sensor nodes (for sending/receiving encrypted information). In their model, a set of edges $E' \in E$ fails after a key is compromised by a successful attack on a node which uses the same key as $E'$. The survivability/reliability measures used in sensor networks are given in Table 2.11.
Table 2.9: Summary of constraints used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Restriction on RN placement (only certain areas)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survivable path from nodes to RNs</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survivable path from RN to BS</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key connectivity</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>Capacity of relay/BS nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>Link capacity (bandwidth)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>Max transmission range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>Candidate/restricted locations for RNs</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifetime of network</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>No constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>usable</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 2.10: Summary of optimization methods used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Custom algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>Approximation method (algorithm)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>usable</td>
<td>usable</td>
</tr>
<tr>
<td>Mixed integer programming</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>usable</td>
<td>usable</td>
</tr>
<tr>
<td>Markov chain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>usable</td>
<td>X</td>
</tr>
<tr>
<td>Lagrangian relaxation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>usable</td>
<td></td>
</tr>
<tr>
<td>New routing scheme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>usable</td>
<td></td>
</tr>
</tbody>
</table>

Summary and discussion

Similar to this dissertation, the number and locations of RNs were optimized by Misra et al. (2010). However, instead of considering cost, they minimized the number of deployed
Table 2.11: Summary of survivability/reliability measures used in the heterogeneous wireless sensor network literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two/k-edge disjoint paths</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k-connectivity (vertex connectivity)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of a node can securely deliver information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probabilistic failures on edges/components</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small world concept</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Routing based survivability</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No survivability/reliability (only connectivity)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

relay nodes under connectivity constraints. Yang et al. (2010) extended the same problem to have two edge-disjoint paths for both sensor to relay nodes and relay to base station nodes whereas the original problem only considered relay to base station nodes. Han et al. (2010) also worked on the same problem. Differing from these studies, devices with different capabilities (such as range or capacity) and costs are considered in this dissertation. In addition to a mixed integer programming model, these studies proposed approximation methods for special cases of the problems. Yang et al. (2010) reported that their approximation algorithm never exceeds an optimality gap of 100%. Although this gap is large, the method could be useful for large problem instances. In this dissertation, optimum solutions of maximum capacititated resilience problem are not known. Therefore, an optimality gap cannot be calculated for the method in this dissertation. In another similar study, Yang et al. (2009) proposed an approximation algorithm for optimum base station placement. They were the first to present a polynomial time approximation algorithm to solve this problem. Although they did not compare their algorithm with an exact method, their algorithm was reported to be better than other approximations. Major difference compared to this dissertation is
that none of these studies used capacitated edges or considered rerouting of flows in case of a failure. Al-Turjman et al. (2013) considered three dimensional node placement, whereas this dissertation and other studies mentioned in this section place nodes in a two dimensional area.

Wang et al. (2007) minimized cost under connectivity and energy constraints. Their problem is different than the research problem presented in this dissertation, because they did not assign paths from sensors to base stations. Also, they only studied the problem as a minimum set covering problem and flows were not considered. They did not model a mixed integer programming formulation, but they provided an exact algorithm to solve the problem. They also presented approximation methods. Machado et al. (2010) worked on the lifetime of a sensor network. The unique thing about their model is the Lagrangian relaxation, because it is a powerful method to solve large sized problems.

Similar to nonsensor networks, edge-disjoint paths and $k$-connectivity are the two most common survivability measures in the heterogeneous wireless sensor network literature. Wang et al. (2008) presented an approximation algorithm to minimize the number of high transmission power state nodes under $k$-connectivity constraints and it outperformed competing approaches in the literature. Kashyap et al. (2010) minimized the additional number of relay nodes in a given network while ensuring connectivity constraints. They gave an exact method and also presented an approximation algorithm with a $k$-connectivity constraint. Their approximation method works for larger values of $k$ as well as small ones. Similar to Kashyap et al. (2010), Bredin et al. (2010) minimized the additional number of sensor nodes to satisfy $k$-connectivity constraints for any value of $k$. It is a different problem than the one in this dissertation, because in their model the network is already deployed. They also provided an approximation algorithm for their problem. Al-Turjman et al. (2013) defined connectivity in a different way by using Laplacian Matrix of nodes. Matrix element $(i,j)$ is $-1$ if nodes $i$ and $j$ are connected, 0 otherwise. The element $(i,i)$ is the number of edges connected to node $i$. They explained that $\lambda_2$, the second smallest Eigen value of the matrix,
defines the minimum number of nodes and links to disconnect the network. Therefore, they maximized $\lambda_2$ to maximize connectivity. Guidoni et al. (2010) worked on a small world concept. Similar to this dissertation, they decided on paths in their model (“edges to use” as a decision variable). It is basically a connectivity problem without considering link reliabilities of the links. Qian et al. (2007) considered network attacks and connectivity issues. They proposed a multiobjective GA to minimize cost and maximize the probability that a node can securely deliver information. As explained earlier, the latter metric is a variation of all-terminal reliability. In their multiobjective optimization they used a simple weighted additive objective function, whereas Pareto optimality is used in this dissertation. They did not consider flow assignment or capacity issues, but they chose the edges to build paths. Also, there were no relay nodes in their model. Therefore, this dissertation fills an important gap by proposing capacitated resilience as a survivability/reliability metric and using it in a bi-objective optimization model.

2.2 Survivability/reliability measures

In this section, the most common reliability measures are summarized and compared with capacitated resilience, which is proposed in this dissertation as a new survivability/reliability measure.

2.2.1 Two-terminal reliability

According to Grover (2004), there are two probabilistic failure models used to analyze the survivability of a network: (1) Given occurrence of failure models and (2) Random occurrence of failure models. In the first one the typical question is: “If failure $x$ occurs, how well are network services protected from it?” The second type of models deals with: “How likely is it that a path between nodes has total outage of over $x$ minutes a year?”

Grover (2004) also asked the question “How likely is it that at least one path between nodes exists?” These questions are related to “two-terminal reliability” which is defined as
the likelihood that at least one distinct path between \((s,t)\) works. Grover (2004) stated that calculating two-terminal reliability is \(NP\)-complete when link failure probabilities are known. The formula of two-terminal reliability is given in Equation 1.2.

A more general form of two-terminal reliability is \(k\)-terminal reliability. Ball (1986) defined \(k\)-terminal reliability as the probability that there exists an operating path from a source node \(s\) to each node in \(K\), where \(s \in K\) and \(k \equiv |K|\). If \(k = 2\), then the reliability is called two-terminal reliability. All-terminal reliability is calculated when \(k = n\) (in other words, \(K\) is \(V\)) where a network can be defined as a graph \(G = (V,E)\) having \(n\) nodes and \(m\) edges. Harms (1995) defined all-terminal reliability as “the probability that there is a path of operating edges from node \(s\) to all other nodes”. A different explanation would be to define unreliability as the likelihood that every distinct path between \((s,t)\) contains at least one failed (blocked) edge. Two-terminal reliability is a special case of all-terminal reliability where it is calculated for only node pairs \((i,j)\).

According to Dengiz et al. (2002), all-terminal network reliability (also called uniform or overall network reliability) can also be defined as “the probability that every pair of nodes can communicate with each other”. The primary design problem is to choose a set of links for a given set of nodes to either maximize reliability given a cost constraint or to minimize cost given a minimum network reliability constraint, as given in Equation 1.1 (in Section 1.3). Dengiz et al. (2002) also used the upper bound estimation technique of Jan (1993) for their all-terminal reliability measure.

These measures consider only edge failures, and nodes are assumed to perfectly reliable. Another assumption is that the failures are independent than each other. This is important for computational ease. Also, in most of the studies, repairs are assumed to take a very short time and are therefore ignored in the models. And, links are either working or failed - degradation is not considered.
These reliability measures can be calculated exactly or approximately. For larger networks, exact calculation of reliability requires extensive computational work, therefore approximation approaches are required.

Ball (1986) stated that a cutset (minimal subset of components whose failure implies the failure of network) and a pathset (minimal subset of components required to operate the network) can be defined by using a “stochastic binary system” (which was first presented by Ball and Nemhauser (1979)) as described below ($S \subseteq T$, where $T$ is the set of all components):

$$\phi(S) = \begin{cases} 
1 & \text{if when } S \text{ operates and } T - S \text{ fails then the system operates} \\
0 & \text{if when } S \text{ operates and } T - S \text{ fails then the system fails} 
\end{cases}$$

Pathsets and cutsets were enumerated to evaluate reliability in early studies (Ball, 1986). To use this enumeration method for reliability calculation, the number of pathsets and cutsets must be small.

In an early paper, deMercado et al. (1976) used a partitioning approach to calculate 2-terminal network reliability. Ball (1979) provided a partition based approach to calculate 2, $k$, and all terminal reliability and compared the proposed algorithm with the enumeration methods. The proposed method is more effective than enumeration approaches in terms of CPU time. Provan and Ball (1984) proposed an algorithm to compute two-terminal reliability between nodes $s$ and $t$ in polynomial time based on the number of $(s,t)$-cuts in the network. The complexity of their algorithm is $O(|E| + |N|\mu^2)$, where $\mu$ is the number of $(s,t)$-cuts. This algorithm is intractable for large values of $\mu$. Abraham (1979) used an algorithm based on boolean algebra to find the 2-terminal reliability of a network. Disjoint (mutually exclusive) paths between nodes $i$ and $j$ are identified first and then reliability is calculated. Beichelt and Spross (1987) improved the algorithm proposed by Abraham (1979) and calculated 2-terminal reliability more effectively because they reduced the number of disjoint terms and therefore reduced the computational time.
Despite these efforts, exact calculation of reliability becomes intractable for large problems. Therefore, alternative methods are needed to solve the problem.

Van Slyke and Frank (1971) developed an effective approach to calculate the probability of a network being connected and the fraction of communicating node pairs. Their approach was based on simulation. Deeter and Smith (1998) worked on all-terminal reliability (actually they used source-sink reliability, which is a variation of all-terminal reliability) and they used a GA. In their GA model, they used reliability in a penalty function (to ensure a minimum reliability constraint). The calculation method of reliability in their model is taken from Ball and Van Slyke (1977). It is a backtracking algorithm and is a reasonable choice for small sized networks. For larger networks, they suggested Monte Carlo simulation to estimate reliability.

Monte Carlo simulation is a popular tool that has been used in reliability analysis; however, it is an approximation. In early works, Kumamoto et al. (1977) used Monte Carlo simulation estimate system reliability and Karp and Luby (1985) proposed Monte Carlo simulation to estimate $k$-terminal network reliability. Lomonosov (1994) estimated reliability using Monte Carlo simulation combining various states of the network. Nel and Colbourn (1990) used Monte Carlo simulation and improved a well-known estimation of reliability bounds (Ball and Provan, 1982) by adding additional constraints to the original estimation. Colbourn and Harms (1988) used linear programming to obtain tighter bounds than (Ball and Provan, 1982) for estimation of all-terminal network reliability.

Dengiz et al. (1997) also worked on all-terminal reliability. They minimized cost under a minimum reliability constraint. They calculated reliability using Equation 2.1:

$$\sum_{\Omega} \left( \prod_{l \in L'} p_l \right) \left( \prod_{l \in L \setminus L'} q_l \right),$$

where $\Omega = \text{all operational states}$, $L' = \text{set of operational links}$, $L' \in L$  

$p_l = \text{reliability of link } l$, $q_l = 1 - p_l$
However, Dengiz et al. (1997) reported that this formula is not tractable for large values of Ω. Therefore, they proposed an estimation method based on Monte Carlo simulation. In their GA methodology, they included reliability as a penalty in the objective function.

Hui et al. (2003) proposed a Monte Carlo approach based on simulation and partitioning edge sets to estimate network reliability. They only considered edge failures (nodes are assumed to be perfectly reliable) and they calculated k-terminal reliability. They used a structure function \( \phi(x) \) which is originally proposed by Ball and Nemhauser (1979). They reported that the new method improved the performance dramatically in comparison to the Crude Monte Carlo and Permutation Monte Carlo approaches. This approach could have been adopted in this dissertation to estimate reliability because the problem herein is focused on heterogeneous wireless networks, but an exact method has been developed to calculate capacitated resilience.

Hui et al. (2005) combined a Cross-Entropy method and a Monte Carlo simulation approach to estimate k-terminal network reliability. They used cross-entropy with Crude Monte Carlo, Permutation Monte Carlo and Merge Monte Carlo simulation based approaches. In their estimation approach, they sampled only up time of each edge and then calculated the probability of the network functioning at a specific time. By iteratively estimating the reference parameter of the cross-entropy method and using likelihood formulations, reliability of the network is estimated. Kroese et al. (2007) also approached reliability estimation by using a cross-entropy method. In their problem, they designed the most reliable network under a budget constraint. Hui et al. (2005) reduced the variation in the reliability estimation dramatically and increased the speed of the calculation. Also, they concluded that Permutation Monte Carlo and Merge Monte Carlo yield better solutions than Crude Monte Carlo for estimating reliability. This proposed approach might be used for extensions of this dissertation, because the authors reported promising performance for large sized networks.

Hui (2007) proposed a new method (Synchronous Construction Ranking) for Monte Carlo
simulation to rank different network designs according to their $k$-terminal reliability values. Their main concern was to find the most reliable design among many candidates.

Ramirez-Marquez and Coit (2005) estimated two-terminal reliability in multi-state networks by using Monte Carlo simulation. Multi-state networks are consisted of edges with different available capacities where the current state (capacity) of the edge can take values of $b_i \geq 0$. Thus, multistate two-terminal reliability ($M2TR_d$) is defined as the probability that a flow requirement ($d$ units of flow) between nodes $i$ and $j$ can be served by the network of multi-state edges.

Cancela and El Khadiri (1995) proposed a variance reduction technique for Monte Carlo simulation to estimate $k$-terminal reliability. They used a recursive method to reduce variation of Monte Carlo simulation by evaluating the unreliability of the network. Their method yielded faster results than existing methods, namely Dagger Sampling, Sequential Construction, Bounds, Failure Set and Merge Process approaches. This method was further improved by Cancela and El Khadiri (1998) by applying series-parallel reductions. A series reduction is to replace two serially connected links (with reliabilities $r_{ij}$ and $r_{jl}$) by one link where the new reliability is calculated as $r_{ij} \ast r_{jl}$. A parallel reduction is to replace two parallel connected links by one link where the reliability of the new link is $r_{ij} + r_{kl} - r_{ij} \ast r_{kl}$. These reduction techniques were proposed by Rushdi (1984) in an earlier study to calculate $k$-terminal reliability. Similar to Cancela and El Khadiri (1995), Cancela and El Khadiri (2003) proposed a new formulation to reduce variance of Monte Carlo simulation to estimate $k$-terminal unreliability. They also combined the new formulation with the series-parallel reduction approaches of Cancela and El Khadiri (1998). Their formulation yielded better solutions in terms of CPU time.

In a recent study by Grosan et al. (2009), the objective functions of cost and resilience were used. They defined resilience by assigning a backup path in addition to a primary path between commodity pairs. Similar to $k$-connectivity and edge-disjoint paths, the backup paths of Grosan et al. (2009) ensure survivability but the paths are not required to be
disjoint (they can have more than one common edge or node). Also, they only consider link failures. Their cost function is simply the edge costs. They minimized the number of common edges with primary and backup paths to ensure resilience. Inversely, they also maximized the number of non-common edges between backup paths. They used Pareto optimality for these three objective functions. In summary, their approach only assigns primary and backup paths and it minimizes common edges between the primary and backup paths. This can be accomplished using edge-disjoint path constraints more effectively. They did not compare their method with other methods, and only discussed the network designs generated by their algorithm.

2.2.2 Probability that two nodes can communicate

Wilkov (1972) presented the “probability of two nodes communicating” measure assuming that all communication-link failures and computer-center breakdowns are statistically independent and that each communication link fails with probability \( p \) and each computer center goes down with probability \( q \). This metric is similar to two-terminal reliability; however, it includes both link and node failures. Calculation of \( P_c(a, b) \), the probability of successful communication of any operating nodes \( a \) and \( b \), is approximated by Equation 2.2.

\[
P_c(a, b) = \sum_{i=0}^{b} A_{a,b}^c(i)(1-p)^i p^{b-i}, \quad p \gg q
\]  

(2.2)

\( A_{a,b}^c(i) \) is the number of combinations of \( i \) nodes such that if they are operative and the remaining \( (n - 2 - i) \) nodes fail, there is at least one communication path between nodes \( a \) and \( b \). Wilkov (1972) also gave an approximation formula for the \( q \gg p \) case. However, a comparison with an exact method is not given.
This approximation calculation could be used in our problem, but a downside is the requirement of homogeneous failures (with failure probability $p$). Also the calculation of $A_{a,b}^n(i)$ is computationally expensive.

### 2.2.3 Expected loss of traffic (ELT)

Another measure for survivability/reliability is the expected loss of traffic (ELT). According to Grover (2004), it is basically the expected number of lost demand-minutes over a year. In this calculation, the total demand (flow) between $i$ and $j$ is not required to be along a single path – multiple paths are allowed. The demand $d_{ij}$ may be realized by routing over $P$ different routes, but total demand must be satisfied. Grover (2004) gives the formula of ELT in Equation 2.3.

$$ELT_{i,j} = \sum_{p=1}^{P} \left( \sum_{k \in (S,N)} d_{ij}^p \delta_{ij}^p(k) U_k \right) M_0$$ (2.3)

Thus, ELT is the “sum of the demand-weighted unavailability of each distinct (not disjoint) path employed to satisfy total demand $d_{ij}$” (Grover, 2004). Disjoint paths can be assured by additional constraints.

Similar to capacitated resilience, ELT can be calculated if flow from $i$ to $j$ is routed over different paths. ELT can also be interpreted as unreliability with split flows, however, it returns the lost demand-minutes over a year. Therefore, unlike capacitated resilience, ELT is not scaled within 0 and 1. In the capacitated resilience calculation, if a flow cannot be rerouted, capacitated resilience is equal to zero. On the other hand, ELT does not capture rerouting possibilities. It only calculates the unreliability for given splits of a flow and use them in the lost-demand calculation. As another difference, ELT calculation assumes that all distinct paths (the given splits) are independent. This assumption may lead to a higher ELT value than it should be if the splits are routed over two paths having a common network.
component. In capacitated resilience calculation, independent subgroups of alternative paths (see Section 3.2.3 for details) are used to overcome the path dependency issues.

2.2.4 Traffic efficiency

Konak and Bartolacci (2007) used the traffic efficiency (TE) measure which was originally proposed by Kubat (1989). The main reasons for them to use TE instead of reliability are to include node failures and to consider traffic flows. They also included a 2-node connectivity constraint in their model.

They defined the state of edges and nodes as either 0 or 1. If a component is operational, then its state is 1, otherwise it is 0. They assumed that reliability (and unreliability) of nodes and edges are given as inputs.

They defined TE as the expected value of $\Omega(x)$, which is the fraction of the traffic that the network delivers in state $x$ (Equation 2.4). The definition of $\Omega(x)$ is given in Equation 2.5. They defined $\tau_{ij}$ to be 1 if there is a path between nodes $i$ and $j$. $\gamma$ is defined as the total traffic demand of the network. This formulation needs to consider $2^{m+n}$ states for calculation of TE.

$$TE = E[\Omega(x)] = \sum_{x \in S} \Omega(x) \ast P\{x\}$$ \hspace{1cm} (2.4)

$$\Omega(x) = \frac{1}{\gamma} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \tau_{ij}(x)t_{ij}$$ \hspace{1cm} (2.5)

There are four main differences between TE and the capacitated resilience metric proposed in this dissertation. First, TE does not consider capacity of edges (or nodes), however, capacitated resilience considers capacity in finding the initial path (the most reliable path)
and in rerouting. Second, TE considers node failures as well as link failures, whereas capacitated resilience considers only link failures. This is an important difference of TE with capacitated resilience and the other connectivity based reliability measures, such as all-terminal reliability. Third, paths from $i$ to $j$ are found for each state of network to calculate TE, whereas all distinct paths are found to calculate capacitated resilience for a given state of network. In other words, capacitated resilience calculates all possible paths from originating node $i$ to all operating APs. And for capacitated resilience, rerouting is only enacted if the initially assigned path fails (any failure in the path). Therefore, distinct paths are calculated without using the failed component and all components other than the failed one(s) are assumed operational. Nevertheless, TE considers all possible states of all components in the network where network components are either operational or failed. Last, split flows are allowed in routing process of capacitated resilience when maximum capacity of an alternative path is reached. On the other hand, as seen from Equation 2.5, TE does not split the traffic flow between nodes $i$ and $j$. These are the main differences between TE and capacitated resilience.

An additional difference is the calculation method of TE. It is calculated by an efficient simulation technique based on “sequential construction” (summarized in Section 4.3.2) by Konak and Bartolacci (2007). In the original article proposing the TE measure, Kubat (1989) combined a Monte Carlo based simulation and analytical approach to calculate TE. The motivation of the calculation approaches of these studies is the intractability of the calculation of TE for large sized networks. However, in this dissertation, capacitated resilience is calculated by enumeration of paths. This enumeration may be intractable for large networks. Therefore, limits on the number of paths (see Section 3.2.2) and cut sets (see Section 3.2.4) are used to estimate capacitated resilience in larger networks.

Similar to capacitated resilience metric, TE is already normalized because it is the expected fraction of the traffic that the network delivers. Therefore, two different network
designs can be readily compared using TE. Both TE and capacitated resilience prefer larger values, and for both metrics the upper limit is 1.

2.2.5 \(k\)-connectivity and edge-disjoint paths

A two-connected (or two-node connected) graph has at least two node-disjoint paths between every pair of vertices (Bondy and Murty, 2008). If two paths are edge-disjoint, they do not have a common edge (but can have a common node). Similarly, if two paths are node-disjoint, they do not have a common node (Kashyap et al., 2010). Therefore, if a graph is two-connected, it is also a two-edge connected graph. However, the reverse of this statement is not necessarily true. These connectivity constraints are widely used to ensure survivability/reliability of a network.

2.2.6 Summary and discussion

To summarize, some survivability/reliability metrics have been used in the HetNets literature although it is most common to add \(k\)-connectivity or edge-disjoint path constraints. Among them, two-terminal and all-terminal reliability measures are the most frequently used ones. These metrics consider link failures. Similar to these, “probability that two nodes can communicate” also captures link failures, but it includes node failures as well. However, flow between nodes is not taken into consideration in these metrics. It is important to have flows and capacities of the devices because rerouting of flow is possible in case of a failure. The proposed metric in this dissertation, capacitated resilience, uses reliability and rerouting options at the same time. It assumes capacitated links due to capacity of wireless devices. Capacitated links have been considered in some studies (such as Benyamina et al. (2009a) and Capone et al. (2010)), whereas some studies (for example Shahnaz and Erlebach (2010)) did not consider capacity of links. Using capacitated wireless links and devices for resilient heterogeneous wireless network design problem is more realistic.
Table 2.12 compares capacitated resilience with the other reliability/survivability metrics. The most important difference of capacitated resilience is the consideration of capacity. Another difference is that capacitated resilience prioritizes the availability of rerouting options. Also, it allows split flows in rerouting. On the other hand, capacitated resilience assumes that nodes are perfectly reliable. In conclusion, capacitated resilience is a more comprehensive metric as it includes both reliability and rerouting under capacity constraints.

Table 2.12: Comparison of reliability/survivability metrics

<table>
<thead>
<tr>
<th></th>
<th>Two-terminal rel.</th>
<th>All-terminal rel.</th>
<th>Pr. that nodes can com.</th>
<th>ELT</th>
<th>Traffic efficiency</th>
<th>k-connectivity</th>
<th>Edge-disjoint paths</th>
<th>Capacitated resilience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rerouting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Capacity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Node Failures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Link Failures</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Split flows</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

In Section 3.3, the reliability/survivability metrics presented in this chapter are calculated for an example network and the differences are discussed.

2.3 Shortest path problem variations: Constraint shortest path and \( k \) shortest path

The rerouting calculation of capacitated resilience requires enumeration of all possible alternative paths under capacity constraints. This problem is known as the constrained \( k \) shortest path problem. The problem reduces to the constrained shortest path problem if \( k \) equals to one. Constrained shortest path and \( k \) shortest path problems are well known and they are summarized in this section.

In one of the earliest works on the shortest path problem Yen (1971) proposed an exact method to find \( k \) shortest paths between two nodes. However, this algorithm does not include
constraints. Since that time, the shortest path problem has been solved under constraints and the problem was termed the “constrained shortest path problem”. According to Dumitrescu and Boland (2001), the weight constrained shortest path problem is to minimize the cost of a route between two nodes while ensuring the total edge weight is less than a predefined value. Another variation of this problem uses node weights instead of edge weights. The basic formulation of this problem is taken from Dumitrescu and Boland (2001) and given in Equation 2.6.

\[
\begin{align*}
\min & \sum_{e \in E} c_e x_e \\
\text{s.t.} \quad & \sum_{e \in \delta^+(i)} x_e - \sum_{e \in \delta^-(i)} x_e = \begin{cases} 
1, & \text{if } i = s \\
-1, & \text{if } i = t \\
0, & \text{if } i \in V \setminus \{s,t\} \end{cases} \quad \forall i \in V \\
& \sum_{e \in E} w_e x_e \leq W \\
x_e & \in \{0, 1\}, \forall e \in E, \delta^+(i) = \{(i,j) \in E\}, \delta^-(i) = \{(j,i) \in E\}
\end{align*}
\]

Dumitrescu and Boland (2001) reported that the weight constrained shortest path problem is closely related to the shortest path problem with time windows and the resource constrained shortest path problem. Also, the bi-objective (cost and weight minimization) shortest path problem is another variation of this problem.

To solve this problem, mixed integer programming formulations were developed in many studies. However, other methods were devised to solve the problem more effectively. Among
them, Lagrangian relaxation, column generation and dynamic programming are the most common ones.

Desrochers et al. (1992) worked on vehicle routing with time windows problem, which is a variation of the weight constrained shortest path problem. In this problem, time windows (allowable delivery times) are added to the vehicle routing problem in which the minimum cost route is sought between two nodes. They used distance between nodes as cost and the weight of edge \((i, j)\) is defined as the time duration (including the service time at customer \(i\)). They defined the problem as a set partitioning problem to select a set of minimal cost routes (Equation 2.7):

\[
\min \sum_{r \in R} c_r x_r \\
\text{s.t.} \\
\sum_{r \in R} \delta_{ir} x_r = 1, \ i \in N \setminus \{d\} \\
x_r \in \{0, 1\}, \ r \in R, \ \delta_{ir} = 1 \text{ if route } r \in R \text{ visits customer } i \in N \setminus \{d\}, \ 0 \text{ otherwise}
\]

where \(d\) is the central depot that routes are originating and terminating at.

In this formulation, \(R\) is the set of feasible routes for the vehicle routing problem with time windows. \(x_r\) is a binary variable (1 if route \(r\) is used). Since the number of columns (i.e., feasible routes) are extremely large, Desrochers et al. (1992) presented a column generation approach.

Barnhart et al. (1998) also used column generation to solve an aircraft routing problem which is similar to vehicle routing with time windows. Differing from the problem solved by Desrochers et al. (1992), they simultaneously solved a fleet assignment problem and an aircraft routing problem (The problem reduces to the aircraft routing problem when there is only one fleet to assign.) Holmberg and Yuan (2003) used column generation to find the
shortest path under time-delay and reliability constraints for capacitated multicommodity network flow problems.

In an early influential paper, Beasley and Christofides (1989) solved the resource constrained shortest path problem by a Lagrangian relaxation formulation. In their problem, the resource constraint was the budget of the traveler. Again in an early study, Handler and Zang (1980) solved the constrained shortest problem by a Lagrangian relaxation algorithm. They used a knapsack constraint in their problem. Their algorithm terminates at the $k$th shortest path which is the first path satisfying the constraints.

Irnich and Desaulniers (2005) summarized solution methodologies as dynamic programming, Lagrangian relaxation, constraint programming and heuristics. Dynamic programming is used to build new paths. A trivial path $P$ is extended for all feasible combinations and new paths are checked whether they are useful or not. Paths are often encoded by labels in dynamic programming algorithms to solve the shortest path with resource constraints problem (Irnich and Desaulniers, 2005). They explained the usage of Lagrangian relaxation in detail. They also defined how to use constraint programming using a “domain reduction algorithm” for each constraint. Heuristic methods, such as “preprocessing” techniques, can be used to eliminate edges and reduce the network. “Dynamic programming heuristics” stop if a predefined number of negative (that is, cannot enter the basis) columns are found.

Carlyle et al. (2008) improved the Lagrangian relaxation approach by closing the optimality gap using an enumeration of near-shortest paths. In this dissertation, this approach would be useful because rerouting of flows (for the capacitated resilience calculation) requires enumeration of all possible paths. However, this approach yields an approximate solution.

Avella et al. (2002) proposed a heuristic to solve large resource constrained shortest path problems. Their heuristic depends on penalty functions and provides approximate solutions. In many cases, their method yielded better solutions with lower number of iterations than Lagrangian relaxation. They eliminate resource constraints by using an exponential penalty
function. This approach could be used in this dissertation to find the most resilient paths under a capacity constraint, but it, too, is an approximate method.

Ribeiro and Minoux (1985) proposed a heuristic method to get good feasible solutions for double constrained shortest path problems. Their algorithm generates partial solutions and combines them to a final solution using enumeration. They showed that the method produces good, feasible solutions even for hard constraints. This method might be useful as a future work of this dissertation by finding the most resilient path for instances with capacity and minimum hop constraints.

Mehlhorn and Ziegelmann (2000) worked on the resource constrained shortest path problem where there are \( k \) resource constraints. They proposed an algorithm, "hull approach", to solve this problem and showed that it provides good upper and lower bounds for the problem. It is a polynomial algorithm when \( k \) equals 1. This algorithm could be useful in this dissertation if more than one constraint (such as capacity and minimum number of hops) are applied. However, only capacity is considered herein.

van der Zijpp and Catalano (2005) proposed an algorithm to enumerate \( k \)-shortest paths with resource constraints. Their problem is equivalent to a shortest path problem with resource constraints if \( k \) is equal to 1. Their algorithm finds the constrained shortest paths directly instead of selecting feasible ones from a large set, which reduces CPU times dramatically.

Feillet et al. (2004) presented an exact algorithm to find the optimum path for the constrained shortest path problem. Their algorithm is a label-correcting algorithm and the problem is a restricted case where the paths are constrained to use a node only once. In their problem, capacity is defined for nodes, which is similar to the problem presented in this dissertation. Righini and Salani (2008) worked on the same problem and proposed a dynamic programming model. They also used a “state-space relaxation” based dynamic programming method in which an infeasible region can be projected onto a feasible region without guaranteeing optimality.
As another variant, Villeneuve and Desaulniers (2005) worked on the shortest path problem with forbidden paths. In this problem some paths (edges) are restricted in building shortest paths. They proposed a polynomial time algorithm and stated that their method can be used to eliminate cycles in $k$ different shortest paths between nodes $i$ and $j$. In this dissertation, this problem may have been useful to find distinct paths (for rerouting) if forbidding some edges was necessary.

Similar to Desrochers et al. (1992), Irnich and Desaulniers (2005) defined the same problem and they stated that this problem is very close to a multiobjective function where time and cost are considered. They also noted that paths can be incomparable because one path may be better than another for one criterion (for example, cost) but can be worse for the other one (e.g., time). They also classified different variants of this problem. They distinguished problems according to formulation of the resource constraints, existence of path-structural constraints, objective and underlying network.

Irnich and Villeneuve (2006) proposed an algorithm to eliminate cycles with length $k$ from resource constrained shortest paths. In this study, they proposed a pseudo-polynomial labeling algorithm. They found that this method solves hard problem instances (with wide time windows) faster than the other methods in literature. They also showed that their algorithm can handle Pareto optimal resource constrained shortest paths. In an early work, Aneja and Nair (1978) solved the constrained shortest path problem with a bicriteria model. Warburton (1987) presented a method to approximate Pareto optimal paths for multiobjective shortest path problems.

2.3.1 Summary and discussion

In relation to constrained shortest path problems, the problem in this dissertation is somewhat different. In this dissertation, the route with maximum reliability is sought under capacity constraints. Edge reliabilities have been transformed with a logarithm to linearize the model and the $k$ shortest path problem has been solved under a capacity constraint.
However, similar to previous studies (Desrochers et al., 1992; Barnhart et al., 1998) there are a large number of paths to enumerate. The $k$ shortest path problem helps to reduce the size of the problem for larger networks by limiting the number of rerouting options ($k$). These distinct paths are important for the rerouting calculation of capacitated resilience. Reliability of a path is maximized, instead of minimizing the cost. Also, the weight constraint is replaced with capacity of nodes. In this dissertation, the $k$ shortest path problem is solved using Yen’s algorithm (Yen, 1971), presented in Section 3.2.2, because of its ease for implementation. Nevertheless, the solution approaches presented in Section 2.3 might be adopted to solve the problem. For example, as a future study, column generation could be applied to calculate capacitated resilience. All possible paths could be generated as columns and the restricted problem solved.

In this dissertation, capacity constraints are based on node (device) constraints. Hence, a transformation from node capacities to edge capacities is required. As demonstrated in Equation 2.8, device capacities can easily be converted to edge capacities. The capacity of edge $(i, j)$ is the minimum of the capacities of devices $i$ and $j$ that the edge connects. Therefore, edge capacity constraints ensure that the capacity of nodes are not exceeded.

$$l_{ij} = \min\{l_i, l_j\}, \text{ where } l \text{ denotes capacity} \quad (2.8)$$

In this dissertation, enumeration of all possible paths is an appropriate solution for small sized networks; however, for realistic sized problems more computationally effective solution approaches must be used. A limit on the number of alternative paths ($k$) helps to approximate for larger networks (see Section 3.2.2 for more details). Cut set size is another parameter in the capacitated resilience calculation (see Section 3.2.4 for more details).
2.4 Some issues of wireless networks

2.4.1 Reliability of a wireless link and network density

In one of the earliest and the most influential works, Takagi and Kleinrock (1984) worked on multihop packet radio networks and defined “probability of successful transmission” by Equation 2.9a. In this formula, $S$ is one-hop throughput, which is defined as the average number of successful transmissions per slot from the terminal (Equation 2.9b), and $N$ is the expected number of nodes within transmission range. $Pr.(P \rightarrow Q)$ is the probability that the transmission from $P$ to $Q$ is successful. Transmission is only possible if none of the terminals (including $Q$) within the range ($R$) of $Q$ transmits when $P$ is transmitting to $Q$. $p$ is the probability that a terminal transmits a packet and it depends on $N$ (number of nodes within range). $p$ is calculated according to Equation 2.9c. Takagi and Kleinrock (1984) found the optimal number of nodes in a transmission range as 7.72 to maximize throughput. The key of this formulation is that the probability of successful transmission is calculated considering interference in the network due to density. Similarly, several other studies defined probability of successful transmission for wireless networks by taking interference concept into consideration (Hunter et al., 2008; Huang et al., 2008; Hira et al., 2007).

\[
\text{Pr. of Successful Transmission} = \frac{S}{p}, \text{ where} \tag{2.9a}
\]

\[
S = Pr.(\text{there is at least one terminal within } R) \tag{2.9b}
\]

\[
\times Pr.(P \text{ transmits}) \times Pr.(P \rightarrow Q)
\]

\[
p = \frac{2}{N + 2 + \sqrt{N^2 + 4}}, \tag{2.9c}
\]

The reliability of a link defined in this dissertation and “probability of successful transmission” (Takagi and Kleinrock, 1984) are related to each other in terms of their definitions.
Both definitions consider the probability if nodes $i$ and $j$ can communicate with each other. However, they have a major difference in terms of their calculation method. The probability of successful transmission takes interference (due to number of nodes in transmission range) into account, however, the reliability of a link only considers the distance (which affects signal quality) between the nodes. Reliability also depends on signal quality which is affected by distance, and is assumed to be negligible in this dissertation.

In another study, Camp et al. (2006) defined the probability that there is a route between nodes $i$ and $j$ (denoted as $R_{ij}$) for wireless mesh networks. Similar to the definition of path reliability given in this dissertation, they calculated the product of the probabilities of all links in a path between $i$ and $j$ (Equation 2.10). In this formulation, $S_l$ is the signal strength of link $l \in R_{ij}$ and $T_{\text{min}}$ is the minimum required signal strength.

\[
Pr(R_{ij} \text{ exists}) = \prod_{\forall l \in R_{ij}} Pr(S_l > T_{\text{min}}) \tag{2.10}
\]

Xue and Kumar (2004) stated that the number of connected neighbors affect the capacity of an ad hoc network due to interference of nodes. Node $i$ interferes with node $j$ when it broadcasts within the range of $j$. Transmission range ($r$) determines the number of nodes in the neighborhood and that interference increases on order of $r^2$ (Xue and Kumar, 2004; Gupta and Kumar, 2000). In another study, Raniwala and Chiueh (2005) stated that the bandwidth of a wireless network using the IEEE 802.11 protocol can be decreased by interference due to the relays in the same path or neighboring paths.

Hekmat and Van Mieghem (2004) used a path-loss law model for radio propagation to calculate interference of mobile ad hoc networks. In their model, the mean value of received signal power ($p_a$) decreases as distance ($d$) between devices increases. The calculation of $p_a$ is given in Equation 2.11 where constant $c$ is affected by transmission power and some other properties of the devices, and $\eta$ is the path loss exponent between 2 and 6. $\eta$ is 2 for free
space, and 5 for a building environment with obstacles. Thus, they modeled interference with this assumption. Their assumption of degradation of signal quality has not been used in this dissertation to keep the model more tractable. Also, the model in this dissertation is more general and the specific environment is not known.

\[ p_a = c d^{-\eta} \tag{2.11} \]

Hekmat and Van Mieghem (2004) also showed that delays in communication increase and network throughput reduces as the number of nodes increases. They found that when network density increases, the average number of hops to communicate decreases but fewer nodes transmit simultaneously. Therefore, capacity per node reduces. In this dissertation, density of a network is adjusted according to the number of users in the network to address this issue.

Takagi and Kleinrock (1984) stated that if the average number of terminals in a transmission range is \( N \), then the probability of successful transmission is proportional to \( 1/N \) in multihop packet radio networks. A shorter transmission range includes a smaller number of terminals in a neighborhood and therefore means a lower possibility of collision. However, a long transmission range increases the chance of finding an appropriate receiver in a desired direction. Takagi and Kleinrock (1984) worked on this trade-off and investigated the optimum number of average terminals in a transmission range. In another study, Grossglauser and Tse (2002) assumed that all nodes can communicate with each other in a mobile ad hoc network. Their analysis showed that throughput for each communication pair in the network decreases proportional to \( 1/\sqrt{n} \) (where \( n \) is the number of nodes per unit area) as \( n \) increases.

In a real life application, reliability of a link is affected by the number of nodes (links) in the transmission range. Probability of successful transmission is proportional to \( 1/N \) (where
is the average number of nodes in transmission range) due to the interference caused by
density (Takagi and Kleinrock, 1984).

2.5 Capacity of a wireless network and interference issues

As a simplistic view, the number of APs definitely affects the capacity in this disser-
tation. All users must connect to an AP; however, possible routes to reach an AP can be
extremely difficult to find if a limited number of APs exist in a network. In other words,
more APs increase the chances to balance the flow of the network. The main bottleneck of
a given network can be investigated by a graph theoretical approach. To determine whether
a specific node is a bottleneck, the cut-sets of the network should be checked. If the node
is repeatedly included in many of the node cut-sets of a network, that device is a bottle-
neck. On the other hand, another definition of a bottleneck would be a device that has no
or limited capacity to route flows. Therefore, the number of devices affects the capacity of
a network. In this dissertation, both device capacities and cut-sets are considered for the
capacitated resilience calculation.

Another issue with the number of devices in a network (or density) is the interference in
the wireless channels. Nevertheless, the effects of interference caused by neighbors have not
been included in the dissertation because this complicates the mathematical model which
is already hard to solve. It is assumed that nonoverlapping channels are used to avoid
interference issues.

The issues on interference and capacity have been studied in the wireless networks
literature and therefore they are summarized in this section, even though they are not
included in this dissertation.

Xue and Kumar (2004) showed that an ad hoc wireless network (having \( n \) nodes) be-
comes (asymptotically) connected if each node is connected to more than \( 5.1774 \log n \) nearest
neighbors \( n \) goes to \( \infty \). By definition, the number of connected neighbors affect the capacity
of the network due to interference of nodes. They also mentioned that device \( i \) interferes
with device $j$ when it broadcasts within the range of $j$. Therefore, more links cause more interference and, eventually, a lower capacity of the network. However, there is a trade-off between the number of links (number of neighbors) and reliability of a network. As the expected number of neighbors (number of links) increases, the chance of having a one hop connection increases (“relaying burden” decreases) but overall interference of network grows. Relaying burden means the requirement of a node to relay the packets from other nodes. Although this is a natural result of multi-hop communication, a higher level of relaying burden is not desired because it increases the loads on RPs by requiring them to transceive a larger number of packets from other nodes (users or RPs). Transmission range ($r$) affects not only the number of neighbors but also the relaying burden and interference (Xue and Kumar, 2004). Relaying burden grows with order of $1/r$, whereas interference grows with the order of $r^2$ (also discussed in Gupta and Kumar (2000)). Therefore, less range and fewer neighbors are better in terms of interference, but the network may be disconnected if the range is too small.

In the problem solved in this dissertation, the network has a different structure. Its structure is closer to the one presented by Gastpar and Vetterli (2002) in which there is only one active link between $i$ and $j$ and all remaining nodes are simply relay nodes for conveying information between $i$ and $j$. This is similar to the user to AP connection of this dissertation. However, in this dissertation instead of the one active link of Gastpar and Vetterli (2002) there are as many connections as the number of users. By using the min-cut maximum flow theorem, Gastpar and Vetterli (2002) showed that the capacity of the network goes to $O(\log n)$ bits per second when $n$ (number of nodes in the network) goes to $\infty$.

Li et al. (2001) used simulation to analyze the capacity of a wireless ad hoc network. One important result related to this dissertation was the effect of multi-hop communication in an ad hoc network. They showed that throughput significantly reduces as the number of hops increases, but the reduction becomes very small after some number of hops. If $n$ is the total number of nodes in the network, they showed that a wireless network (using
IEEE 802.11) approaches its theoretical maximum capacity per node \(O(1/\sqrt{n})\), which was originally proposed by Gupta and Kumar (2000). Therefore, capacity of a network reduces as the network becomes denser.

Reliability of a wireless mesh network increases when there are more redundant paths for communication pairs to alleviate catastrophic effects due to failure of bottleneck links (Bruno et al., 2005). Raniwala and Chiueh (2005) explained the term “path capacity” which is the minimum residual bandwidth of the path that connects a WMN node to the wired network. Another approach to analyze wireless network capacity is “gateway capacity” which takes into account the capacity of the gateway (AP) link with the assumption that a bottleneck is caused by the gateway (AP) connection. A gateway is simply an AP, which is explained in this dissertation. Raniwala and Chiueh (2005) also stated that a bottleneck is generally located on links around an AP due to heavy collision and interference. Therefore, the number of APs affects the capacity of a network. However, Raniwala and Chiueh (2005) added that bottlenecks can occur even in intermediate wireless links due to other radio sources and therefore path load balancing can be a more effective strategy to improve network capacity instead of AP (gateway) load balancing. Similarly, Jun and Sichitiu (2003) showed that gateways (APs) in a WMN are the main bottlenecks. If \(n\) is the number of users served by an AP (gateway), they claimed that the available capacity of each node is affected by order of \(1/n\). Adding more APs increases the capacity and reliability of the network (Jun and Sichitiu, 2003). The number of APs directly affects the capacity of both individual nodes and the network. Jun and Sichitiu (2003) computed throughput through identification of bottleneck collision domain which was defined as the area in the network limiting the amount of data that can be transmitted in the network.

Wang and Liu (2006) proposed a linear programming model to find maximum capacity (in terms of throughput) for a given network. They investigated ad hoc networks and did not consider mesh type networks, but their methodology could be adapted to this dissertation as a future work to calculate throughput.
Wu et al. (2006) assumed that mesh routers use different channel assignments to prevent interference in their model. The same assumption has been included in this dissertation. A non-overlapped channel is used for one of the routers when two of the routers are close enough to interfere with each other. As summarized in Wu et al. (2012), IEEE 802.11a has 12 non-overlapping channels, whereas IEEE 802.11b/g has only three. In their proposed queuing model, a bottleneck is defined as the average delivery delay of data requests at the gateway nodes (APs). Also bottleneck throughput is defined as the “maximum feasible data requesting frequency from all users attached to a mesh router”. They showed that the bottleneck delay decreases when the number of APs (gateways) increases. They assumed the location of gateways do not affect bottlenecks because they ignored the delay caused by routers. However, this is an issue in real life problems. By using their $M/D/1$ queuing model, they found that the number of APs must be greater than $\lambda N s$ to ensure a steady network where $N$ is the number of mesh routers, $\lambda$ is the mean arrival rate of data to the mesh routers and $s$ is the server time at an AP node.
3.1 The need for a new metric

In network design literature, requiring $k$-connectivity or edge-disjoint paths are the most common ways to ensure survivability. Other reliability metrics, namely $k$-terminal reliability (Grover, 2004), all-terminal reliability (Harms, 1995), “traffic efficiency” (Konak and Bartolacci, 2007), “probability that two nodes can communicate” (Wilkov, 1972) and “expected loss of traffic” (Grover, 2004), have also been used. Among them, two-terminal and all-terminal reliability measures are the most frequently used ones. These well known reliability/survivability measures, and the differences between them and capacitated resilience are summarized in Sections 2.2 and 2.2.6, respectively.

The main motivation for developing a new metric is to include rerouting options as well as reliability when considering capacitated links and devices. Capacity is one of the main differences. If a device (or a link) does not have enough capacity then traffic cannot be routed on that device (or link). Connectivity based survivability metrics and most reliability/survivability metrics do not consider capacity.

Rerouting is an important issue for session continuity in telecommunication networks and it is the essential part of the capacitated resilience metric. For example, capacitated resilience becomes zero if there is no alternative path available. Capacitated resilience emphasizes session continuity in network design. Unlike connectivity constraints, it also allows comparison of different network designs. In other words, two different $k$-connected network designs can have different reliability and capacitated resilience values.

The capacitated resilience metric proposed in this dissertation can be calculated with an exact method. In the network reliability/survivability literature, metrics were mostly
calculated by approximate methods, such as Monte Carlo simulation. The studies using exact methods made many simplifying assumptions to keep problems tractable and do not consider alternative paths. Connectivity based metrics, such as $k$-connectivity, have only $k$ different alternative paths and they do not consider reliabilities of the paths. This is another difference between capacitated resilience and the others.

### 3.2 Proposed metric: Capacitated resilience

Network level capacitated resilience is the weighted average of capacitated user resistances in terms of their traffic requirements. Equation 3.1 presents the calculation of capacitated network resilience. In this equation, $w_i$ is the weight of User $i$, which is the proportion of traffic flow of User $i$ to the total traffic flow of all users. Note that the traffic flow of a user is assumed to be constant over time. The capacitated resilience calculation requires user level capacitated resilience values to be known. If the network consists of one user, then capacitated user resilience is equal to network resilience.

$$\text{Capacitated Resilience} = \sum_{i \in \text{Users}} w_i \ast \text{Capacitated Resilience} (U_i) \quad (3.1)$$

User level capacitated resilience is defined in Equation 3.2. $R(p_i)$ denotes the reliability of the assigned path of User $i$ where the user $(U_i)$ is assigned to a device whose path has maximum reliability (from the user to an AP). The resilience factor scales the user reliability. To calculate it, all alternative paths (those having available capacity) from User $i$ to any available access point are identified first. Then, reliabilities of the alternative paths are calculated. These steps are explained in Sections 3.2.1 through 3.2.5. Capacity of devices and links are considered for finding assigned and alternative paths, but split flows are not allowed due to complexity of calculations.
Capacitated Resilience \( (U_i) = R(p_i) \times \text{Resilience Factor} (U_i) \), \( i \in Users \) \hspace{1cm} (3.2)

The next sections explain the calculation steps of capacitated resilience (user level) components: (1) Finding the most reliable path between the user to an access point, (2) Identifying alternative paths between the user and any access point, (3) Determining independent subgroups of the alternative paths, (4) The reliability calculation of independent subgroups, and (5) Calculating reliability of the alternative paths (Resilience factor) using subgroup reliabilities.

A subgroup consists of a subset of alternative paths of a user. Obviously, a subgroup is a subgraph of the network. Therefore, subgroup and subgraph are used interchangeably in this dissertation.

3.2.1 Finding the most reliable path between the user to an access point

There might be more than one AP in a HetNet. For a given user, finding the most reliable path to connect an AP is important for service quality. The most reliable path from a user to an access point is found by Dijsktra’s shortest path algorithm (Algorithm 3.1). It is a well known algorithm which labels every vertex \( v \) with its predecessor \( p(v) \). In this algorithm, the distance from vertex \( r \) to \( v \) is defined as \( l(v) \). More information can be found in Bondy and Murty (2008). As given in Equation 3.3, the algorithm works by minimizing the logarithm of link reliabilities which is equivalent to maximizing reliability of the path.

Note that the reliability of the wireless link between user \( i \) and device \( j \) is calculated by \( \max\{0, (r_j - d_{ij})/r_j\} \), where \( d_{ij} \) is the distance between user \( i \) and device \( j \) and \( r_j \) is the range of device \( j \). Similarly, the reliability of the wireless link between device \( j \) and device \( k \) is calculated by \( \max\{0, (r - d_{jk})/r\} \), where \( r = \min\{r_j, r_k\} \).
Algorithm 3.1 Pseudocode of Dijkstra’s shortest path algorithm (Bondy and Murty, 2008)

1: set \( p(v) \leftarrow \emptyset, v \in V, \ l(r) \leftarrow 0, \ d(v) \leftarrow \infty, v \in V \setminus r \)
2: while there is an uncolored vertex \( u \) with \( l(u) < \infty \) do
3: choose a vertex \( u \) with minimum \( l(u) \)
4: color \( u \) black
5: for \( \forall \) uncolored neighbor \( v \) of \( u \) with \( l(v) > l(u) + w(u,v) \) do
6: replace \( p(v) \) by \( u \) and \( l(v) \) by \( l(u) + w(u,v) \)
7: end for
8: end while
9: return \( (p,l) \)

\[
\log R(P) = \log \prod_{(i,j) \in P} p_{ij}
\]
\[
\log R(P) = \sum_{(i,j) \in P} \log p_{ij}
\]

(3.3)

\( R(P) \), reliability of a path of the user to an AP, is maximized when \( \sum_{(i,j) \in P} \log(p_{ij}) \) is maximized (equivalently, \(- \sum_{(i,j) \in P} \log(p_{ij}) \) is minimized), because \( 0 \leq p_{ij} \leq 1 \), as given in Equation 3.4.

\[
\max\{R(P) = \prod_{(i,j) \in P} p_{ij}\} \equiv \max\{\sum_{(i,j) \in P} \log(p_{ij})\} \equiv \min\{- \sum_{(i,j) \in P} \log(p_{ij})\}
\]

(3.4)

After evaluating reliabilities from the user to all available APs (having enough capacity), the AP with the most reliable path is assigned to the user.

3.2.2 Identifying alternative paths between the user and any access point

Upon assigning the most reliable AP to a user, the next step is to identify all alternative paths from the user to all APs. This subproblem is computationally the most expensive part of the capacitated resilience calculation.
For each available AP, a $k$ shortest path problem is solved to find the $k$ most reliable paths from the user to that AP. Identifying all available paths might be intractable for large networks, therefore limiting the number of paths ($k$) reduces the complexity of this subproblem. Obviously, the solution becomes exact if $k$ is sufficiently large.

The $k$ shortest path problem is not new and there are many algorithms to solve it. Section 2.3 summarizes the $k$ shortest path problem and its solution methods. Among them Yen (1971), Eppstein (1998), Katoh et al. (1978), and Katoh et al. (1982) are the most important ones. Katoh et al. (1982) is a generalization of Yen (1971). Eppstein (1998) is faster than Yen (1971) but it allows repeated vertices (which makes the search space larger). In this paper, Yen’s (Yen, 1971) $k$ shortest path algorithm has been used, because it provides an effective and easy implementation and allows only simple paths (no loops or repeated vertices).

Details of Yen’s algorithm

The pseudo code of Yen’s Algorithm to find $k$ shortest paths is given in Algorithm 3.2 (the interested reader may refer to the original article (Yen, 1971) for more information):

Yen’s algorithm starts with finding the most reliable path using any shortest path algorithm. In here, Dijkstra’s shortest path algorithm (Algorithm 3.1) has been used because it is an exact algorithm and easy to implement. Note that solving the shortest path problem is equivalent to maximizing reliability of the path according to Equation 3.4. Upon identifying the shortest path between the user and the AP, the second shortest path is found by modifying the shortest path obtained in the first step. The main idea in this algorithm is to use previously obtained shortest paths to generate the remaining shortest paths. Therefore, the algorithm runs Dijkstra’s shortest path algorithm from an intermediate node in the shortest path (called a spur) to the AP without using the rest of the edges in the shortest path. After finding the second shortest path the algorithm finds the rest of the $k$ shortest paths, $(k - 2)$
Algorithm 3.2 Pseudocode of Yen’s $k$ shortest path algorithm

1: for $k$=1 do
2:   Step 1: Use Dijkstra to find shortest path from a fixed node to other nodes. The result will be $A_1$. This step assumes there are no negative loops in the shortest path.
3:   Step 1a: Store $A_1$ into List $A$.
4:   end for
5: for $k$=2 to $K$ do
6:   Step 2: Check if a node sequence (1) − ... − $i$ of $A^{k-1}$ is same with first $i$ nodes of previously generated path $j$ = 1, 2, ..., $k$ − 1. Go to Step 3.
7:   Step 3: Find the shortest path from $i$ to $N$ without including any nodes from (1) − ... − $i$ of $A^j$ (which is called as $R_i^k$). Therefore, we are finding the shortest path of spur of $A^k_i$, which is $S_i^k$.
8:   Step 3a: Add $A^k_i$ (joins $R_i^k$ and $S_i^k$) to candidate List $B$. Note that we need only $K$ − ($k$ − 1) many items in List $B$.
9: if Number of paths found at Step 3 + number of paths in List $A$ > $K$ then
10:      $K$ Shortest paths found, save the paths to List $A$
11:      Break
12: else
13:      Step 4: Move path $A^k$ from List $B$ to List $A$.
14:      Step 4a: Leave remaining items in $B$ and $k$++. Repeat steps 2−4 until having $K$ shortest paths.
15: end if
16: end for
17: return List $A$ ($k$ shortest paths)
paths in a similar fashion. To do this, new candidate shortest paths are found by using previously identified shortest paths. The algorithm ends when $k$ shortest paths are found.

### 3.2.3 Determining independent subgroups of the alternative paths

Upon identifying all alternative paths from the user to all available APs, the next step is to group the alternative paths into independent subgroups. For any user having at least one alternative path (except the assigned path), there must be at least one independent subgroup of alternative paths. Each independent subgroup consists of paths having one or more common edges. Any path of an independent subgroup cannot have a common edge with a path of another independent subgroup. Independent subgroups are determined by a simple algorithm, which is summarized in Algorithm 3.3.

In this algorithm, all alternative paths are compared with each other to check for common edges. If so, one of the paths and all other paths in the same group are labeled as the other path’s group. Therefore, the number of subgroups is dynamic in this algorithm.

**Algorithm 3.3** Pseudocode of clustering alternative paths into independent groups

1: Save all alternative paths to List $P$
2: groupNo $\leftarrow 1$
3: for each path $i \in P$ do
4:     if path $i$ is not in a group then
5:         if $i = 1$ then
6:             Assign $groupNo_i \leftarrow groupNo$
7:             $groupNo + +$
8:         end if
9:     for $j = 1$ to $P$ do
10:        if $i$ and $j$ has a common edge then
11:            Assign $groupNo_i$ to all paths of the group that $j$ belongs to
12:        end if
13:     end for
14: end if
15: end for
16: return Independent subgroups of alternative paths
3.2.4 Reliability calculation of independent subgroups

To find the capacitated resilience, reliabilities of the independent subgroups (that are identified in the previous step) must be calculated. However, minimum (or minimal) independent cuts must be found first to get reliability of each subgroup.

This problem is similar to s-t cut set problem, which is basically finding the minimum cut set between source and sink nodes. However, as there can be more than one AP in a subgroup, the problem is not the same as the s-t cut set problem. Another version of this problem is the multiterminal cut problem which is finding a set of edges that consist of non-terminal nodes to disconnect terminal nodes. Xiao (2010) and Hartvigsen (1998) stated that the multiterminal cut problem is NP-hard for \( n \geq 3 \), where \( n \) is the number of nodes in a graph. Again, this problem is not the same problem herein, because the terminal vertices, i.e. the user and APs, do not necessarily communicate with each other. In other words, APs do not communicate with each other because they only serve as a connection to the wired backbone. Also, unlike MANETs, users are not required to communicate with each other.

Thus, algorithms from the literature are not suitable for the capacitated resilience calculation. Therefore, an algorithm has been developed to find all minimum independent cut sets to disconnect the user from any AP with which user can communicate directly or indirectly (via RPs). One way to solve this problem is to enumerate all possible cut sets and then find the independent ones. The main drawback of this approach is the large computational effort for realistic sized networks. Although estimation methods can be used, such as Monte Carlo simulation (for example, Dengiz et al. (1997) and Deeter and Smith (1998) use Monte Carlo for all-terminal reliability estimation), an exact approach has been presented in this paper.

The pseudo code for finding minimum independent cut sets is given in Algorithm 3.4. This algorithm checks all combinations of edges beginning with one edge, two edges, three edges and so on. If any combination of the edges disconnects the user from all APs in the subgroup, then that combination of edges is a cut. If any combination of edges uses an edge
from a previously found cut set, that combination is not considered. The algorithm terminates when there are not enough edges left to form a unique combination or all combinations have been examined.

**Algorithm 3.4** Pseudocode of finding minimum independent cutsets of a subgroup

1: Save all unique edges of alternative paths in the subgroup to set \( E \)
2: Cutsets \( \leftarrow \emptyset \)
3: for \( k = 1 \) to \( m \) (number of edges in \( E \)) do
4: \[ \text{if } |E| \text{ - number of unique edges in Cutsets } < k \text{ then} \]
5: \[ \text{break} \]
6: end if
7: for each unique combination of edges \( (C(m,k)) \in E \) do
8: \[ \text{if } C(m,k) \cap \text{Cutsets } = \emptyset \text{ then} \]
9: \[ \text{if removal of edges in } C(m,k) \text{ disconnects user from all APs} \text{ then} \]
10: Cutsets \( \leftarrow \) Edges in \( C(m,k) \)
11: end if
12: end if
13: end for
14: end for
15: return Cutsets

Upon identifying all minimum independent cut sets of a subgroup, reliability of a subgroup is calculated. Failure of any cut set disconnects the user from APs. Therefore, all cut sets must be reliable to make the network reliable. Equation 3.5 presents the reliability calculation of a subgroup \( (S_i) \). In this formulation, \( C \) denotes the total number of combinations of edges that form cut sets. Similarly, \( R(C_{ij}) \) denotes the reliability of cut set combination \( C_{ij} \). \( R(C_{ij}) \) is calculated by parallel reliability calculations (Equation 3.6). The reliability of cut \( C_{ij} \) requires that at least one edge (or serial edges from a RP to an AP), denoted as \( e \) in the Equation 3.6, to be operational.

\[
\text{Reliability}(S_i) = \prod_{j=1}^{C} R(C_{ij}) \quad (3.5)
\]
\[ R(C_{ij}) = 1 - \prod_{e \in C_{ij}} (1 - r_e) \], where \( r_e \) is the reliability of edge \( e \in C_{ij} \)  

### 3.2.5 Calculating reliability of the alternative paths (resilience factor) using subgroup reliabilities

After calculating reliabilities of the independent subgroups, the last step is to calculate the reliability of the alternative paths (union of subgroups). This calculation is simply the parallel reliability calculation, where the system is reliable if at least one of the subgroups is reliable. Equation 3.7 shows this calculation where \( S \) denotes the number of independent subgroups and \( R(S_i) \) denotes the reliability of subgroup \( i \).

\[
\text{Resilience Factor} = 1 - \prod_{i=1}^{S} [1 - R(S_i)]
\]  

### 3.3 An example network

In this section, calculations of capacitated resilience and the other reliability/survivability metrics are demonstrated on an example. The network presented in Figure 3.1 is used throughout this section. In this network, there are four APs (having enough capacities) and many rerouting options for a single user.

#### 3.3.1 Capacitated resilience calculation

Capacitated resilience is calculated for both uncapacitated and capacitated cases. Although capacitated resilience considers capacities, uncapacitated case is important to show the effect of the capacity constraint because the other metrics neglect capacities.
Uncapacitated case

Uncapacitated case considers that device capacities are larger than the user flow requirement and therefore the problem becomes uncapacitated. To calculate the capacititated resilience of this network, all possible alternative paths from the user to all available APs are found (where capacity is available) by Yen’s $k$ shortest path algorithm. In this dissertation, $k$ is sufficiently large ($k = 7$) to find all shortest paths. Then, alternative paths from User to APs are grouped into independent groups. Lastly, reliability of the independent subgroups are calculated.

Independent subgroups 1, 2, and 3 are identified using the procedure given in Section 3.2.3.

The cut set of subgroup 1 (Figure 3.2) is $U - AP1$ and the reliability of the subgroup is 0.7.
Figure 3.2: Independent subgroup 1 of the network in Figure 3.1

Figure 3.3: Independent subgroup 2 of the network in Figure 3.1
Figure 3.4: Independent subgroup 3 of the network in Figure 3.1

The cut set of subgroup 2 (Figure 3.3) consists of \((U - R P 2)\) and \((R P 2 - A P 1, R P 2 - A P 3)\). The reliability of the subgroup is 0.644. The calculation steps are presented in Equation 3.8.

\[
R(S2) = R(U - R P 2) (R(R P 2 - A P 1, R P 2 - A P 3)) \\
= 0.7(1 - (1 - 0.8)(1 - 0.6)) = 0.7(0.92) = 0.644
\]

(3.8)

The cut set of subgroup 3 (Figure 3.4) consists of \((U - R P 4)\), and \((R P 4 - A P 6, R P 4 - A P 7, R P 4 - R P 5\) or \(R P 5 - A P 6)\). The reliability of the subgroup is 0.4275 (Equation 3.9).

\[
R(S3) = R(U - R P 4) (R(R P 4 - A P 6, R P 4 - A P 7, R P 4 - R P 5\) or \(R P 4 - A P 6)) \\
= 0.5 ((1 - (1 - 0.5)(1 - 0.5))[1 - (0.6 * 0.7)]) = 0.5(0.855) \\
= 0.4275
\]

(3.9)

After calculating reliabilities of the subgroups, the reliability of the alternative path system (resilience factor) can be calculated by:
\[ 1 - (1 - 0.7)(1 - 0.644)(1 - 0.4275) = 0.938857 \]

The resilience factor is the reliability of alternative path system (rerouting options). The alternative path system remains reliable if at least one subgroup is reliable. The capacitated resilience of the network has been defined in Equation 1.3. Therefore, the capacitated resilience (for the single user) can be calculated as:

\[ \text{Capacitated resilience} = 0.8(0.938857) = 0.751086 \]

Capacitated resilience for the entire network is calculated by the weighted average (in terms of flows) of user level capacitated resiliences. In this example there is only one user.

**Capacitated case**

For the capacitated case (Figure 3.5), the calculations of subgroups 1 and 2 remain unchanged. However, the reliability calculation of subgroup 3 is changed due to lack of capacity of AP6 and AP7. The flow of the user must be split between AP6 (traffic flow of 10) and AP7 (traffic flow of 10), therefore the subgroup can only be operational if the two split paths, U-RP4-AP7 and U-RP4-AP6 (or U-RP4-RP5-AP6), are both reliable. The calculation is given in Equation 3.10. The reliability of the subgroup 3 reduced from 0.4275 to 0.1775. The system reliability becomes 0.912157 (Equation 3.11). The capacitated resilience is calculated as 0.72973 (Equation 3.12). The reduction in the capacitated resilience is not very high (from 0.7511 to 0.7297) because there are three alternative path systems in this example. If the subgroup 3 was the only alternative path system, the reduction in the capacitated resilience due to limited capacity would be more significant, i.e., from \(0.8 \times 0.4275 = 0.342\) to \(0.8 \times 0.1775 = 0.142\).
Figure 3.5: Capacitated version of the example in Figure 3.1 (\(c\) and \(f\) denote device capacity and user flow requirement, respectively)

\[
R(S3) = R(U - RP4) \cdot (R(RP4 - AP6, RP4 - AP7, RP4 - RP5orRP4 - AP6))
\]

\[
= 0.5 \cdot 0.5 \cdot (1 - 0.6)(0.7) \cdot (1 - 0.5)) = 0.5 \cdot 0.355
\]  \hspace{1cm} (3.10)

\[
= 0.1775
\]

\[
1 - (1 - 0.7)(1 - 0.644)(1 - 0.1775) = 0.912157 \hspace{1cm} (3.11)
\]

\textit{Capacitated resilience} = 0.8 \cdot 0.912157 = 0.72973 \hspace{1cm} (3.12)

The next sections demonstrate the calculation of the other reliability/survivability metrics. As explained in Section 2.2, capacity is not considered for those metrics.
3.3.2 Connectivity measures

The connectivity measures are also investigated for the same example (Figure 3.1). Two-connectivity (two node-disjoint paths) is assured by the paths U-AP3 and U-RP2-AP3. Similarly, the same set of paths (U-AP3 and U-RP2-AP3) ensures two edge-disjoint paths. Therefore, this network is both two node-connected and two edge-connected.

3.3.3 Two-terminal and all-terminal reliability

Similar to the connectivity measures, two-terminal reliability is calculated for U-AP3. The paths leading to AP3 from the user are U-AP3 and U-RP2-AP3. Therefore, at least one of the paths must be operational to reach AP3. Two-terminal reliability becomes \(1 - (1 - 0.8)(1 - 0.6(0.7)) = 0.844\).

All-terminal reliability considers all paths from the user to all available APs. Therefore, the parallel reliability calculation of the paths from user to AP1, AP3, AP6 and AP7 yields all-terminal reliability: \(1 - (1 - 0.8)(1 - 0.42)(1 - 0.7)(1 - 0.56)(1 - 0.25)(1 - 0.25)(1 - 0.21) = 0.993196\). Not surprisingly, all-terminal reliability is larger than two-terminal reliability because it considers all possible APs whereas two-terminal reliability considers only AP3.

3.3.4 Traffic efficiency

Traffic Efficiency/Resilience (Konak and Bartolacci, 2007; Kubat, 1989) is calculated by both exact method and simulation. The exact method yields 0.9878, whereas simulation (using 5000 iterations, according to Algorithm 4.6 in Section 4.3.2) yields 0.9902.

3.3.5 ELT

The ELT calculation requires the input of the splits of the flow. Assuming the user traffic is split equally to U-AP3 and U-AP1, the ELT is calculated by (according to Section 2.2.3):
\[
ELT = [(1 - R(U - AP1))(0.5) + (1 - R(U - AP3))(0.5)] (M_0) \\
= [(1 - 0.7)(0.5) + (1 - 0.8)(0.5)] 5.26(10^5) \\
= 0.25(5.26(10^5)) = 131,500 \text{ demand-minutes/year}
\]

The value of the constant \(M_0\) is \(5.26(10^5)\) to convert the fraction of lost flow to demand-minutes per year (Grover, 2004). However, the resulting number is not scaled, it is the expected number of lost demand-minutes over a year. Considering the fraction of the flow that cannot be routed is 0.25, then a conversion of ELT can be done to obtain a scaled measure (like reliability), i.e., ELT becomes 0.75 \(1 - [(1 - 0.7)(0.5) + (1 - 0.8)(0.5)]\).

### 3.3.6 Summary

The procedures for finding the above listed reliability/survivability measures are summarized in detail in Section 4.3. Each metric presents a different way to evaluate the network. Capacitated resilience emphasizes the importance of rerouting options in case of failure of the assigned path. Therefore, it considers the redundancy level of the network under capacity constraints. Connectivity metrics focus on redundant paths, but they do not consider their reliability. Two terminal and all terminal reliabilities are calculated by parallel reliability calculations and they focus on the redundancy in the network. In the example given in Figure 3.1, terminal reliability values are high, however, these values can be misleading because these metrics do not consider the common links that are used by many paths. In case of a failure of these links, the redundancy of the network can be reduced or lost. Capacitated resilience considers common links and their impact on reliability of redundant paths in its subgroup and cut set calculations. Traffic efficiency also indirectly considers common links in the reliability calculation. However, it does not consider capacity. As explained in Section 6.3, all these differences result in different network designs.
Chapter 4
Proposed Method and Solution Approach

The network design problem, even without survivability/reliability constraints (or objectives), is a well known $NP$-hard problem. In this dissertation, an Evolutionary Strategies (ES) algorithm has been developed to solve this problem. Both single objective and bi-objective models are considered in the proposed ES.

The main reasons to use ES for solving the resilient heterogeneous wireless network design problem are summarized below:

- The capacitated resilience calculation is nonlinear and traditional optimization methods (such as mixed integer programming) cannot be applied effectively. And for computational reasons, for large problem instances, an alternative methodology other than traditional exact optimization methods is needed to solve the problem. This is the main motivation for using a metaheuristic such as ES.

- ES is known for its success on continuous problems. The problem solved in this dissertation has device coordinates as continuous decision variables as well as some discrete variables (e.g., device types). Also, ES is a flexible method that can handle many restrictions and assumptions.

- Bi-objective optimization is one of the main contributions of this research. ES can easily be extended to bi-objective optimization using “Pareto optimality” (defined in Section 4.2.1). Here, bi-objective optimization is ensured by “non-domination ranking” (see Section 4.2.1) as used in NSGA-II (Deb et al., 2002). NSGA-II is a multiobjective Genetic Algorithm (GA) technique, however, its main idea has been adapted to ES in this dissertation.
There are some drawbacks of using ES. First of all, it does not guarantee optimality. Second, choosing a good representation (encoding) of problem and finding efficient mutation operators are generally challenging issues. Third, parameter tuning is important for better solution quality; for example, premature convergence can be observed due to certain parameter settings. These issues have been addressed in this dissertation.

### 4.1 Single objective Evolutionary Strategies (ES) model

Cost and capacitated resilience are two objective functions that are considered in this dissertation. The single objective ES uses cost (capacitated resilience) as the objective function and capacitated resilience (budget) as a constraint. The parameters of the ES model are given in Table 4.1. Budget is important because it determines the maximum number of devices in the network. As there are two types of devices (AP and RP), the number of devices and their types change the cost. As explained in Section 4.1.9, the cost and the capacitated resilience of a solution are penalized if the solution is infeasible. A solution becomes infeasible if any user is not assigned to a device or a device does not have a feasible route.

#### Table 4.1: Input parameters used in the ES model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td># of users and their locations</td>
<td>Locations as ((x, y)) coordinates</td>
</tr>
<tr>
<td>flow of the users</td>
<td>Traffic requirement of each user</td>
</tr>
<tr>
<td>gridX and gridY</td>
<td>The limits of the area in which nodes are deployed</td>
</tr>
<tr>
<td>budgetLimit</td>
<td>Budget limit for installing devices</td>
</tr>
<tr>
<td>maxDevice</td>
<td>Maximum number of devices</td>
</tr>
<tr>
<td>Cost of devices</td>
<td>Cost to deploy a relay or an access point</td>
</tr>
<tr>
<td>Ranges of devices</td>
<td>Ranges of relay or access points</td>
</tr>
<tr>
<td>Capacity of devices</td>
<td>Capacities of relay or access points</td>
</tr>
</tbody>
</table>

The decision variables of the ES model are listed in Table 4.2. The objective function of the ES model is defined in detail in Section 4.1.5.
Table 4.2: Decision variables used in the ES model

<table>
<thead>
<tr>
<th>Decision Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td># of devices</td>
<td>Number of devices used in the solution ($\leq$ MaxDevice)</td>
</tr>
<tr>
<td>Locations of devices</td>
<td>Defined as $(x, y)$ coordinates</td>
</tr>
<tr>
<td>Mixture of type of devices</td>
<td>Either router or access point</td>
</tr>
<tr>
<td>User to device assignments</td>
<td>Each user $i$ is assigned to a device $j$</td>
</tr>
<tr>
<td>Routing of the flow</td>
<td>From a router to an access point the route of user traffic</td>
</tr>
<tr>
<td>Reliability among devices</td>
<td>It depends on the locations of devices and routing</td>
</tr>
<tr>
<td>Reliability among users and devices</td>
<td>It depends on the locations of users and devices</td>
</tr>
</tbody>
</table>

4.1.1 Pseudocode of ES model

The pseudocode of the single objective ES model is given in Algorithm 4.1. This ES model maximizes capacitated resilience (minimizes cost) under a budget (capacitated resilience) constraint.

At each generation, a parent is selected from the population randomly and saved as a child solution. Then, mutation and swap operators are applied to the child. Mutation and swap operators are explained in Section 4.1.6. In the proposed ES, the device types of a solution are changed every ten generations because finding good coordinates for a given set of devices (mixture of APs and RPs) is very hard. Therefore, 10 generations give ES enough time to adjust device coordinates for a given set of devices. Children are replaced with population members according to $(\mu + \sigma)$ rule (see Section 4.1.7 for more details). The best solution is saved throughout the generations.

In the proposed ES algorithm, there are ES specific operations. First of all, mutation success is calculated for a predetermined number of generations ($g'$). A successful mutation means that the child solution (mutated solution) has a better objective function values than its parent (lower cost or higher resilience, according to the objective function). Note that the success of a mutation is defined differently in bi-objective optimization (see Section 4.2 for details). According to the “rate of successful mutations”, $\sigma$ is dynamically adjusted at $n$ generations. If the percentage of successful mutations is less than the rate of successful mutations, the search space is narrowed by 0.85. In this dissertation, rate of successful
mutations is selected as 0.2, which is very common in ES literature and known as the “one fifth rule” (Dréo et al., 2006). $\sigma$ is the mutation rate for coordinates and is explained in Section 4.1.6 in detail. At the beginning of the ES, $\sigma_x$ and $\sigma_y$ are initialized to one third of the grid dimensions (gridX and gridY of Table 4.1). The algorithm terminates if the maximum number of generations is reached or the best solution has not updated for a given number of generations (Section 4.1.8). This ES algorithm yields the best network design (does not guarantee optimality), the lowest cost or the highest resilience, according to the objective function.

Parameters for running the ES model are given in Table 4.3.

Table 4.3: Model parameters used in the ES model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_p$</td>
<td>Population size</td>
</tr>
<tr>
<td>$N_c$</td>
<td>Number of children</td>
</tr>
<tr>
<td>$P_m$</td>
<td>Probability to mutate</td>
</tr>
<tr>
<td>maxGen</td>
<td>Maximum number of generations</td>
</tr>
<tr>
<td>maxNonImprovingGen</td>
<td>Maximum number of non-improving generations</td>
</tr>
<tr>
<td>numberOfAlternativePaths</td>
<td>Number of alternative paths in capacitated resilience calculation</td>
</tr>
<tr>
<td>cutSetSize</td>
<td>The maximum number of edges in a cut set</td>
</tr>
<tr>
<td>successRate</td>
<td>Also known as one fifth rule. It is a parameter to adjust $\sigma$ in ES</td>
</tr>
<tr>
<td>sigmaAdjuster</td>
<td>Adjustment rate of $\sigma$ in ES</td>
</tr>
<tr>
<td>numberOfMutationsToAdjustSigma</td>
<td>It is used to estimate successful mutations in ES</td>
</tr>
<tr>
<td>penaltyUnassignedUser</td>
<td>Penalty value if a user is not assigned to a device</td>
</tr>
<tr>
<td>penaltyInfeasibleRouting</td>
<td>Penalty value if a device cannot connect to the wired backbone</td>
</tr>
</tbody>
</table>

4.1.2 Encoding

An adjacency matrix is used to represent the network structure. This matrix is $n \times n$, where $n$ is the maximum number of devices. Different from a traditional adjacency matrix, reliabilities between devices are used instead of 0/1 values in the matrix. In other words, if devices $i$ and $j$ can communicate with each other (both devices are within the communication ranges of each other), then the entry $ij$ of the adjacency matrix is the reliability between devices $i$ and $j$. Also, device types and ($x, y$) coordinates of the devices are stored in an
Algorithm 4.1 Pseudocode of the single objective ES model

Ensure: Maximum budget (minimum capacitated resilience) constraint, assigning all users to a device, feasible (defined in Section 4.1.9) routings

1: Random initialization of population (Section 4.1.3)
2: Sort population (from the best solution to the worst)
3: bestSolutionSoFar ← Population(0)
4: \( g \leftarrow 0 \) //g is the generation counter
5: while (\( g < \text{maxGen} \)) do
6: \( i = 0 \) to \( N_c-1 \) do
7: Randomly select a parent \( (i) \) from population to mutate, \( i \in \{0,1,\ldots,(N_p-1)\} \)
8: Mutate device coordinates of Child\([i]\)
9: Select two random devices \( j \) and \( k \) of Child\([i]\)
10: Swap device types of \( j \) and \( k \) within Child\([i]\)
11: if (\( g\%10 = 0 \)) then
12: Mutate device types of Child\([i]\)
13: end if
14: end for
15: Sort population and children (from the best solution to the worst)
16: Replace worst population members with new generated children
17: Sort population (from the best solution to the worst)
18: if (Population[0] is better than bestSolutionSoFar) then
19: bestSolutionSoFar ← Population[0]
20: end if
21: if (\( g\%g' = 0 \)) then
22: if (Mutation success rate > 0.20) then
23: Increase \( \sigma_x \) and \( \sigma_y \) to \( 1/0.85 \) of their values
24: else
25: Decrease \( \sigma_x \) and \( \sigma_y \) to 0.85 of their values
26: end if
27: end if
28: if (bestSolutionSoFar has not been updated for maxNonImprovingGen generations) then
29: Terminate ES
30: end if
31: \( g \leftarrow g+1 \)
32: end while
33: return Network design with maximum capacitated resilience (or minimum cost)
array with the size of the maximum number of devices. Users to device assignments are stored in a separate array with the size of the number of users.

4.1.3 Population initialization procedure

The population is randomly initialized using a different random number seed than the one used in the ES. In the initialization procedure, only device types and device coordinates (in continuous space) are created. Users are assigned to the closest devices and a routing algorithm (Section 4.1.4) is applied to route user flows. Cost, reliability and capacitated resilience values are calculated according to the equations of Section 4.1.5.

4.1.4 Edge selection algorithm

The edge selection algorithm is basically the routing of flow from device (or user) $i$ to $j$. Device capacities are checked by using the adjacency matrix, and the best path from user $i$ to $AP_j$ is determined by using Dijkstra’s shortest path algorithm. Dijkstra’s algorithm minimizes the total cost of the links of a path, however, in this routing problem cost is meaningless because wireless links do not incur costs. Instead, the logarithms of link reliabilities are used as “cost” in Dijkstra’s algorithm. The algorithm maximizes the reliability between user $i$ and an access point. This procedure was explained in Equation 3.4. Thus, an access point with the most reliable path is selected using Dijkstra’s algorithm. If the device is already an access point, Dijkstra’s algorithm is not applied, because it is already connected to the wired backbone.

The main issue of this algorithm is the capacity of devices. Users are sequentially assigned to the devices and routed over the devices. In other words, Dijkstra’s algorithm is applied to each user in a randomized order. Therefore, depending on the order of assigning flows to devices a better path may become infeasible later in the algorithm (due to lack of capacity of a device). It is possible that a user’s flow cannot be routed due to insufficient capacity of devices. Also, the ES may yield non-optimal results due to this ordering issue.
On the other hand, this algorithm works satisfactorily for problem instances in which the capacities of devices are not highly constrained. Although the effect of ordering on the solution quality has not been observed in the preliminary testing of the ES, this issue will be investigated as a future work of this dissertation.

4.1.5 Calculation of cost, reliability and capacitated resilience

The cost calculation is the summation of the deployment costs of devices (Equation 4.1). Also, penalty values are added to the cost function. As given in Table 4.3, two different penalties are defined in this ES model. A penalty is added to the cost if a user is not assigned to a device and another penalty is added if a device does not have a feasible path to an AP. Section 4.1.9 discusses these penalties in detail.

\[
\text{Cost} = \sum_{i \in \text{Devices}} c_i \times x_i
\]

where \(c_i\) is the cost of device \(i\), and

\(x_i = 1\), if the device \(i\) is included in the solution

(4.1)

The capacitated resilience calculation is given in Equation 1.4 (in Section 1.3). According to this equation, the weighted average (in terms of flows) of all resilience values for each device is calculated to find capacitated network resilience. A penalty is applied to capacitated resilience if the budget is exceeded. Capacitated resilience is also penalized if there are unassigned users or infeasible device paths. These penalties are discussed in Section 4.1.9 in detail.

The reliability of a network is calculated similarly to the capacitated resilience calculation. The reliabilities of assigned paths of all users are combined using the weighted average of user flows. Therefore, the network reliability is the weighted average of the user reliabilities.
4.1.6 Mutation operators

The first mutation is to alter device coordinates. In this mutation, the coordinates of a device are changed by $N(0, \sigma_x)$ and $N(0, \sigma_y)$ for $x$ and $y$ axes, respectively. The $(x, y)$ coordinates of devices are restricted to be inside the predefined grid (given in Table 4.3). This is also known as “self-adaptive Gaussian” mutation (Dréo et al., 2006) because the mutation rates of coordinates, $\sigma_x$ and $\sigma_y$, are dynamically adjusted during the ES. The search is widened by increasing $\sigma$ parameters by $1/0.85$ if the percentage of successful mutations is larger than $0.2$ (“one fifth rule”). Otherwise, the search is focused on a smaller region by decreasing $\sigma$ values by $0.85$. This decision is made every $n$ generations.

A swap of device types is the second type of mutation. In this mutation, also known as “2-opt swap”, the two devices to swap are randomly selected and device coordinates are not changed. To perform the swap operator, the devices are not required to be in the solution. This encourages the optimizer to search more diverse candidate solutions. Also, it works faster since the devices are not checked whether they are in the solution or not.

The last mutation is to change a device type of a child using a variation of bit-flip mutation (Dréo et al., 2006). In this mutation, an RP (AP) device changed to an AP (RP) if a random number is less than the mutation probability (given in Table 4.3). Similarly, the status of the device is changed by this mutation operator. For example, a device is included in the solution if another random number is less than the mutation probability and the device was not included in the solution prior to the mutation. If a previously not included device is included in the solution, the coordinates of the device are assigned using an algorithm instead of random assignment. In this algorithm, the number of devices in each quadrant are calculated and the device coordinates are randomly assigned within the quadrant having the least amount of devices. There are four quadrants (I, II, III and IV) in the $xy$ plane, the first quadrant is the area where $x > 0$ and $y > 0$, the second is the area where $x < 0$ and $y > 0$, the quadrant is the area where $x < 0$ and $y < 0$ and the last is the area where $x > 0$.
and $y < 0$. A slight increase in the solution quality has been observed with using this device coordinate assignment instead of random coordinate assignment.

The coordinate change and swap mutations are done at each generation. However, device type change mutation is done at every 10 generations to give the ES enough time to optimize device coordinates as explained in Section 4.1.1. The swap mutation is done at each iteration instead of doing it infrequently because the probability of changing a device type (from AP to RP or RP to AP) is not very high and therefore the effect of swap on the solution quality is not very large. Since only two devices are swapped and there are two device types (AP and RP), the probability of getting a different device type after the swap operator is 0.5 if there are equal amount of APs and RPs in the solution. Also, only two devices in a solution are swapped, therefore the net effect of the change (if there is any) on the solution quality is very limited. In fact, the initial experimentation of the model showed that swapping at each iteration outperformed swapping infrequently. Obviously, the main mutation operator is the coordinate change. Note that the original solution, which was mutated and saved as the child solution, stays in the population as the parent solution.

After mutations, users are assigned to the devices having the most reliable path to an AP using the routing algorithm. Cost, reliability and capacitated resilience values are calculated.

Budget is not considered in mutation operators. However, a penalty of exceeding the budget is included in the capacitated resilience calculation (see Section 4.1.9 for details), which is applied after mutation. The reason of not including budget control in mutations is to encourage the optimizer to find more diversified solutions even though this may temporarily create some infeasible solutions.
4.1.7 Replacement of population with children

After generating children, population and children solutions are combined and the best solutions are selected for the next generation of population. This is also known as $\lambda + \mu$ replacement in ES literature.

As an elitist approach, the “best so far” solution is stored and updated throughout generations. At each generation, the worst individual of the population is replaced with the “best so far” solution if the “best so far” is better than the current best individual of the population. The reason of this approach is to maintain the “best so far” solution in the population at all times.

4.1.8 Stopping criteria

The stopping criteria are the total number of generations and the number of non-improved generations. The number of non-improved generations criterion prevents the ES running longer without improving solution quality. The key is to select the number of non-improved generations is to balance early termination prevention and avoid running excessively.

4.1.9 Penalty functions

Three main penalties are used in the proposed ES: (1) Unassigned users, (2) Infeasible device routes, and (3) Budget overruns. These penalties change cost (penalties 1 and 2) and capacitated resilience (penalties 1, 2 and 3) to ensure that infeasible solutions do not proliferate in the population.

The cost of a solution is penalized if a user is not assigned to a device. A user may not be assigned to a device due to lack of capacity of the device, being beyond the ranges of devices, or lack of connection of device to an AP. Another penalty is applied to the cost when the routing of a device is infeasible. Infeasible routing occurs when an RP cannot connect to an AP. There is no routing needed for APs, therefore this penalty does not apply to APs.
Each penalty is a “death penalty” (Dréo et al., 2006), in other words, these penalty values are very large to prevent reproduction of infeasible solutions. These penalties are defined in Table 4.3 for each user and device. An additional cost penalty is not applied for exceeding the budget limit, but capacitated resilience is penalized.

Capacitated resilience is scaled by percentage of assigned users. For example, if one out of ten users is not assigned to a device, the capacitated resilience value is scaled by 0.9 (= 9/10). The number of devices with infeasible paths is incorporated in the capacitated resilience calculation in a similar way. The percentage of devices with feasible paths scales capacitated resilience. In addition to these two penalties, capacitated resilience is further penalized by the cost of the solution. If the cost of the devices (not including penalties) is greater than the budget, capacitated resilience is scaled by “budget/cost of devices”. This dynamically penalizes capacitated resilience for exceeding the number of devices and forces the ES to reduce the total number of devices or number of APs (either by removing them or replacing them with RPs).

The details of the penalty functions are given in Algorithm 4.2.

4.2 Bi-objective Evolutionary Strategies (ES) model

The bi-objective ES simultaneously optimizes capacitated resilience and cost using Pareto optimality. The parameters and decision variables of the bi-objective ES model are same as in the single objective model (Tables 4.1, 4.2 and 4.3).

4.2.1 Definitions: “Non-dominated rank”, “Crowding distance” and “Pareto optimality”

In bi-objective optimization of resilient heterogeneous wireless networks, solution \( i \) dominates solution \( j \) only if its cost (capacitated resilience) is lower (higher) than \( j \) and capacitated resilience (cost) is not lower (higher) than \( j \) (Equation 4.2).
Algorithm 4.2 Pseudocode of penalty functions

Ensure: Penalize cost and capacitated resilience of solution[i] (if the conditions are met)
1: Define $n_{UU}$ and $n_{ID}$ as the number of unassigned users and the number of infeasible devices, respectively
2: Define $Cost_{Devices}$ as the deployment cost of devices in solution[i]
3: Set $n_{UU} \leftarrow 0$, $n_{ID} \leftarrow 0$ and $Cost_{Devices} \leftarrow 0$
4: Calculate solution[i].cost and solution[i].CR // CR denotes Capacitated Resilience
5: for (∀ user[j] in solution[i]) do
6: if (user[j] is not assigned to a device) then
7: $n_{UU} \leftarrow n_{UU} + 1$
8: end if
9: end for
10: if ($n_{UU} > 0$) then
11: solution[i].cost $\leftarrow$ solution[i].cost $+ n_{UU} \times penalty_{UnassignedUser}$
12: solution[i].CR $\leftarrow$ solution[i].CR $\times \frac{\text{Total number of users in solution[i]} - n_{UU}}{\text{Total number of users in solution[i]}}$
13: end if
14: for (∀ device[j] in solution[i]) do
15: if (device[j] is not connected to an AP (has an infeasible route)) then
16: $n_{ID} \leftarrow n_{ID} + 1$
17: end if
18: end for
19: if ($n_{ID} > 0$) then
20: solution[i].cost $\leftarrow$ solution[i].cost $+ n_{ID} \times penalty_{InfeasibleRouting}$
21: solution[i].CR $\leftarrow$ solution[i].CR $\times \frac{\text{Total number of device in solution[i]} - n_{ID}}{\text{Total number of device in solution[i]}}$
22: end if
23: for (∀ device[j] in solution[i]) do
24: $Cost_{Devices} \leftarrow Cost_{Devices} + device[j].cost$
25: end for
26: if ($Cost_{Devices} > budgetLimit$) then
27: solution[i].CR $\leftarrow$ solution[i].CR $\times \frac{\text{budgetLimit}}{Cost_{Devices}}$
28: end if
Solution $i$ dominates $j$ if

\[ Cost_i \leq Cost_j \text{ and } \text{Cap. Resilience}_i > \text{Cap. Resilience}_j , \text{ or} \]

\[ Cost_i < Cost_j \text{ and } \text{Cap. Resilience}_i \geq \text{Cap. Resilience}_j \]

(4.2)

Deb et al. (2002) defined “non-dominated ranking” and “crowding distance” to ensure population diversity for their multiobjective GA. A “non-dominated rank” of a population is a set of non-dominated solutions. In other words, a solution does not dominate another within a rank. For a population having more than one solution, there is at least one non-dominated rank which is called the first rank. There can also be other ranks in a population; i.e., second, third and so on. A solution of a lower non-dominated rank (e.g., first rank) dominates the solutions of higher ranks (e.g., second rank). The solutions in the first rank are “Pareto optimal” because they are not dominated by any solution in the population.

“Crowding distance”, as first defined by Deb et al. (2002), measures “uniqueness” of a solution in objective function space. It is defined for solutions within the same non-dominated rank. In this dissertation the crowding distance is calculated according to the algorithm proposed by Deb et al. (2002). This algorithm (Algorithm 4.3) deals with each non-dominated rank separately. Within each rank, the crowding distances of solutions are calculated for each objective function. For any objective function, the crowding distance of a solution becomes infinity (or a very large number) if its objective function value is the minimum (or maximum) of its rank. For all other solutions crowding distance is calculated by adding the difference of objective function values of two neighboring solutions ($i + 1$ and $i - 1$). Therefore, a smaller crowding distance means that the objective function value of solution $i$ is very close to its neighbor solutions ($i + 1$) and ($i - 1$). Note that, the neighboring solutions of a solution are defined for each objective function and they can be different for each objective. In Figure 4.1, the crowding distance values of solutions 1 and $n$ (where $n = |r|$) are infinity. As given in Algorithm 4.3, the crowding distance is normalized for each.
objective function $m$ by $(f_{m}^{\text{max}} - f_{m}^{\text{min}})$. Although the closeness of the objective function values of two solutions does not guarantee their similarity in solution space, initial experimentation showed that two solutions are very similar in solution space if their objective function values (cost and capacitated resilience) are close to each other.

**Algorithm 4.3** Pseudocode of crowding distance calculation (Deb et al., 2002)

1: for $(r = 1$ to total number of non-dominated ranks $)$ do
2: for (each objective $m$ (cost or capacitated resilience)) do
3: Solution$(i).\text{distance} \leftarrow 0$, $\forall i \in \text{Rank } r$
4: Sort solutions within the rank according to objective $m$
5: $n \leftarrow$ number of solutions in Rank $r$, Solution$(1) \leftarrow \infty$, Solution$(n) \leftarrow \infty$
6: $f_{m}^{\text{max}} \leftarrow$ Maximum value of objective function $m$ within rank $r$
7: $f_{m}^{\text{min}} \leftarrow$ Minimum value of objective function $m$ within rank $r$
8: for $(i = 2$ to $(n - 1))$ do
9: Solution$(i).\text{distance} = \text{Solution}(i).\text{distance} + \frac{\text{Solution}(i+1).m - \text{Solution}(i-1).m}{f_{m}^{\text{max}} - f_{m}^{\text{min}}}$
10: end for
11: end for
12: end for

4.2.2 Pseudocode of ES model

The pseudocode of the bi-objective ES model is given in Algorithm 4.4. The algorithm is exactly the same as the single objective ES except that the sorting is different and Pareto
optimality is used. The solutions within a population are sorted according to non-dominated ranks and then the crowding distances within the rank (see Section 4.2.4 for details). This idea of sorting was first introduced by Deb et al. (2002) for a multiobjective Genetic Algorithm (GA). Each rank has non-dominated solutions. However, by definition, the solutions of the second rank are dominated by the solutions of the first rank. The first rank is also called the “Pareto Front” which includes the best solutions of the population. Non-dominated rank and crowding distance are used to keep population diversified for bi-objective optimization.

4.2.3 Global Pareto front operations

The proposed bi-objective ES keeps a “global” Pareto front throughout the search. The global Pareto front is different than the current Pareto front of the population. The current Pareto front consists of solutions of the current population, whereas the global Pareto front includes solutions from any population generated during the ES search. In other words, the global Pareto front keeps the best solutions that have been found so far. Therefore, the global Pareto front solutions may be different than the Pareto front of the current population. A solution may be included in the current Pareto front but it may also be dominated by some solutions of the global Pareto front. At the first generation the global Pareto front is empty and at the end of the last generation there is at least one non-dominated solution in the Pareto front. At each generation the population members are modified by mutation operators. Therefore, new solutions are generated and they are compared with existing global Pareto front solutions. If a new solution is better than an existing Pareto front solution, the solution becomes a member of the Pareto front and the dominated solutions of the Pareto front are discarded. Thus, the size of the global Pareto front is dynamic.

Pareto front diversity

Global Pareto front members do not necessarily exist in the population. They can be lost during the iterations due to mutation operators. To increase variability within the
Algorithm 4.4 Pseudocode of the bi-objective ES model

Ensure: Maximum budget (minimum capacitated resilience) constraint, assigning all users to a device, feasible routings

1: Initialization of population
2: Calculate non-domination rankings and crowding distances of solutions
3: Sort population (according to non-domination rankings and crowding distances, explained in Section 4.2.4)
4: bestSolutionSoFar ← Population(0)
5: $g ← 0$
6: while ($g < \text{maxGen}$) do
7:     for ($i = 0$ to $N_c - 1$) do
8:         Randomly select a parent ($i$) from population to mutate, $i \in \{0, 1, \ldots, (N_p - 1)\}$
9:         Mutate device coordinates of Children[$i$]
10:        Select two random devices $j$ and $k$ of Children[$i$]
11:        Swap device types of $j$ and $k$ within Children[$i$]
12:        if ($g \% 10 = 0$) then
13:            Mutate device types of Children[$i$]
14:        end if
15:     end for
16:     Sort population and children (according to non-domination rankings and crowding distances)
17:     Replace worst population members with new generated children
18:     if (If the solution Population[$i$] is better than bestSolutionSoFar, $i \in \{0, 1, \ldots, N_p\}$) then
19:         bestSolutionSoFar ← Population[$i$]
20:     end if
21:     Update Pareto front with new solutions (Add new non-dominated solutions from Population and remove dominated solutions)
22:     Add some Pareto front members to the population (see Algorithm 4.5)
23:     if ($g \% n = 0$) then
24:         if (Mutation success rate $> 0.20$) then
25:             Increase $\sigma_x$ and $\sigma_y$ to $1/0.85$ of their values
26:         else
27:             Decrease $\sigma_x$ and $\sigma_y$ to $0.85$ of their values
28:         end if
29:     end if
30:     if (bestSolutionSoFar has not been updated for maxNonImprovingGen generations) then
31:         Terminate ES
32:     end if
33:     $g ← g + 1$
34: end while
35: return Network design with maximum capacitated resilience (minimum cost)
population and increase the solution quality of future generations some inferior population members are replaced with some non-dominated solutions from the global Pareto front. Details of this process are given in Algorithm 4.5. In this algorithm, the population is sorted for increasing crowding distance. A larger crowding distance is preferred because it shows the “uniqueness” of a solution. Therefore, sorting the population according to the crowding distance is equivalent to sorting it from the least unique solution to the most unique. To have a more diversified Pareto front and a better solution quality during the search, the population is preferred to have a larger number of “unique” solutions. The proposed algorithm compares each solution in the Pareto front with all members of the population and the number of similar solutions in the population is saved for each Pareto front solution. A population solution is determined to be similar to the Pareto front solution if its cost and capacitated resilience values are within a predetermined percentage of the Pareto front solution’s cost and capacitated resilience. If a Pareto front solution has a smaller number of similar solutions than a given threshold, that solution is replaced with the population member having the smallest crowding distance value. Other Pareto front solutions are added to the population in a similar way. This process ends if there is no Pareto front solution qualifying for the given conditions or a predetermined number (\(\text{countMax}\) parameter in the algorithm) of Pareto front members are added to the population. The value of \(\text{countMax}\) has been selected as four after initial testing of the ES. Thus, at most four least unique solutions in the population (size of the population is selected as 30) are replaced with the most unique global Pareto front solutions.

4.2.4 Multiobjective sorting

Deb et al. (2002) defined “partial order” \((\prec_n)\) where Solution\((i)\) \(\prec_n\) Solution\((j)\) means that Solution\((i)\) is better (or more preferable) than Solution\((j)\). A solution having a lower non-dominated rank is better than the other. Within the same rank, solutions with larger
Algorithm 4.5 Pseudocode of adding Pareto front solutions to the population

1: \( n \) ← number of solutions in Pareto front
2: \( \text{minNumber} \leftarrow N_p/10, \text{countMax} \leftarrow N_p/5, \text{counter} \leftarrow 0 \)
3: Initialize array \( \text{numberOfSolutionsWithinRange} \) with size of \( n \)
4: Sort Population for increasing order of crowding distance
5: \textbf{for} \( (i = 0 \text{ to } n - 1) \) \textbf{do}
6: \quad \textbf{for} \( (j = 0 \text{ to } N_p - 1) \) \textbf{do}
7: \quad \quad \textbf{if} \ (\text{Cost}_i \ast (1 - \text{rangePercentage}) < \text{Cost}_j < \text{Cost}_i \ast (1 + \text{rangePercentage}) \text{ and} \)
8: \quad \quad \quad \quad \text{CR}_i \ast (1 - \text{rangePercentage}) < \text{CR}_j < \text{CR}_i \ast (1 + \text{rangePercentage})) \textbf{then}
9: \quad \quad \quad \quad \text{numberOfSolutionsWithinRange}[i] \leftarrow \text{numberOfSolutionsWithinRange}[i] + 1
10: \quad \textbf{end if}
11: \textbf{end for}
12: \textbf{for} \( (i = 0 \text{ to } n - 1) \) \textbf{do}
13: \quad \textbf{if} \ (\text{numberOfSolutionsWithinRange}[i] < \text{minNumber}) \textbf{then}
14: \quad \quad \text{Population[counter] \leftarrow ParetoFrontSolution[i]}
15: \quad \quad \text{counter} \leftarrow \text{counter} + 1
16: \quad \textbf{end if}
17: \textbf{if} \ (\text{counter} \geq \text{countMax}) \textbf{then}
18: \quad \textbf{break}
19: \textbf{end if}
20: \textbf{end for}
21: \text{Sort Population (multiobjective sort)}
“crowding distance” values are more preferable than ones with lower crowding distance values. Equation 4.3 formally defines this preference rule between solutions.

\[
\text{Solution}(i) \prec_n \text{Solution}(j) \quad \text{if} \\
\text{Solution}(i)_{\text{rank}} < \text{Solution}(j)_{\text{rank}}
\]

or

\[
\text{Solution}(i)_{\text{rank}} = \text{Solution}(j)_{\text{rank}} \quad \text{and} \quad \text{Solution}(i)_{\text{Crowd. Dist.}} > \text{Solution}(j)_{\text{Crowd. Dist.}}
\]

This partial ordering process is used for the multiobjective sorting in this dissertation. The solutions are first sorted according to increasing order of non-dominated ranks. Then, the solutions within the same rank are sorted according to the decreasing order of crowding distance values.

4.3 Calculation of other survivability and reliability metrics

In this dissertation the proposed metric, capacitated resilience, is compared with the metrics explained in Section 2.2. The calculation steps of these metrics are explained in the following sections.

4.3.1 \textit{k} and all-terminal reliabilities

\textit{k}-terminal reliability is a special case of all-terminal reliability. In the literature, the parameter \(k\) is usually selected as two. Similar to capacitated resilience, \(k\)-terminal reliability is first calculated at the user level and then the network level \(k\)-terminal reliability is obtained by a weighted average in terms of user traffic requirements (Equation 4.4).

\[
k\text{-terminal reliability} = \sum_{U_i \in Users} w_i \times k\text{-terminal reliability}(U_i)
\]
In this dissertation, two-terminal reliability is calculated at the user level by calculating the parallel reliability of the assigned path of the user to an AP and a disjoint path from the user to the same AP (Equation 4.5). All terminal reliability is the parallel reliability of the assigned user path and all other disjoint paths from the user to all available APs.

Two-terminal reliability\( (U_i) = 1 - \prod_{p_i \in \text{Paths from } U_i \text{ to the AP}} [1 - R(p_i)] \) \hspace{1cm} (4.5)

### 4.3.2 Traffic efficiency (TE)

Due to intractability of exact calculation of TE, Konak and Bartolacci (2007) proposed a simulation based method, Sequential Construction. One replication of their simulation algorithm is summarized in Algorithm 4.6.

In this algorithm, both node and edge failures are considered. The algorithm randomly starts network components (node or link) from either operative or failed states according to reliabilities \((x_{ij} = 1 \text{ if the component is operational})\) and improves connectivity incrementally by repair of each component (changing its state from failed to operative), where component \((i, j)\) is an edge if \(i \neq j\), a node otherwise \((i = j)\). The permutation of the component states is randomly generated (according to node and link reliabilities) for each replication of the algorithm. The algorithm calculates the weighted average of the successful traffic flow to estimate TE using the total traffic demand on the network \((\gamma)\) and the traffic demand between nodes \(i\) and \(j\) \((t_{ij})\).

In this dissertation, the Sequential Construction algorithm of Konak and Bartolacci (2007) has been used. They used 300 iterations for all solutions and 5000 iterations for promising solutions to estimate TE. In here, the number of replications is selected as 2000 to estimate TE because the initial experimentation showed that the difference between the estimations by 2000 and by 5000 replications is negligible.
Algorithm 4.6 Pseudocode of Sequential Construction method to estimate TE (Konak and Bartolacci, 2007)

1: Sample network state $x$ by assigning a random number $U$ for each component $(i, j)$, if $U > p_{ij}$ then $x_{ij} = 0$ ($x_{ij} = 1$, otherwise).
2: Generate random permutation $\Pi$ of components in operative and failed states
3: Set $T \leftarrow 0$, $h \leftarrow 1$, $c \leftarrow 0$, $a \leftarrow 0$, $\text{label}[i] \leftarrow i$ for nodes $i = 1, 2, ..., n$ and $x_{ij} \leftarrow 0$ for each component $(i, j)$.
4: // $T$, $h$, $a$ and $c$ are the temporary variables used in the simulation.
5: for ($r \in 1, 2, ..., (n + m)$) do
6: 
7: $c = c(n+m-r+1)(\frac{p_{[r]}}{1-p_{[r]}})$ where $p_{[r]}$ is the reliability of the $r$th component in permutation $\Pi$ that is sampled in the previous step.
8: Repair the $r$th component in permutation $\Pi$.
9: if ($r$th component is a link $(i_{[r]}, j_{[r]})$, and both nodes $i_{[r]}$ and $j_{[r]}$ are in the operative state and $\text{label}[i_{[r]}] \neq \text{label}[j_{[r]}]$) then
10: Set the label of each node whose label is equal to the label of node $i_{[r]}$ to the label of node $i_{[r]}$
11: end if
12: if ($r$th component is a node $(i_{[r]}, i_{[r]})$) then
13: for (each link $(i_{[j]}, r)$ incident to node $i_{[r]}$) do
14: if ($i_{[j]}$, $r$) and node $j$ are operative and $\text{label}[i_{[r]}] \neq \text{label}[j_{[r]}]$) then
15: Set the label of each node whose label is equal to the label of node $j$ to the label of node $i_{[j]}$
16: end if
17: end for
18: end if
19: for (each node pair $i$ and $j$) do
20: if ($\text{label}[i] = \text{label}[j]$) then
21: $T = T + t_{ij}$
22: end if
23: end for
24: $h = h + c$, $a = a + c \times T/\gamma$
25: end for
26: return Estimation of $TE$ as $a/h$. 

88
4.3.3 $k$-connectivity

$k$-connectivity is calculated similar to $k$-terminal reliability. The user level $k$-connectivity is checked and the network is said to be $k$-connected if all users are $k$-connected. If any user is not $k$-connected then network $k$-connectivity fails.

In this dissertation, $k$ is selected as two. Vertex (node) and edge disjoint connectivity are both examined. A user is two vertex connected if two different paths connecting the user to its assigned AP have no common vertex. For two edge connectivity, the two paths must not have a common edge.
Chapter 5  
Preliminary Results

This chapter summarizes the performance of the proposed single and bi-objective ES models. It shows the variation of the ES in terms of solution quality and Pareto front diversity. It also compares single and bi-objective ES performance to validate that the ES is working satisfactorily prior to the work presented in Chapter 6. This chapter also explains the key concepts of Pareto front diversity and non-dominated solutions which were introduced in the previous chapter.

The problem size is selected as 10 users and the budget is constrained to 600 (17 devices) because it is a small sized scenario with short solution times. The budget is selected as 600 to make problem quite constrained and therefore harder than a scenario with a larger budget. Capacitated resilience, cost, and Pareto front diversity are compared for different problem instances. The bi-objective ES ran for 2000 generations with an early termination criterion of 500 non-improved generations, whereas the single objective optimizer ran for 1000 generations with an early termination criterion of 250 non-improved generations. Both single objective and bi-objective ES use 30 population members and 30 children. These parameters were specified during the experimentation of the model building phase with consideration of the trade-off between solution time and solution quality. The ES is not very sensitive to the parameter values of size of population and children, however, a larger size of population (and children) helps improve solution quality with an increase of the solution time. Also, it improves population diversity for bi-objective optimization. After extensive experimentation with different parameter values, these values were selected. For example, bi-objective ES runs longer (2000 generations) than single objective ES (1000 generations) because Pareto optimality requires more time to find good solutions. Running ES longer increases solution
time but provides a slight improvement on the solution quality (more on this in Sections 5.2 and 5.3). The selected parameter values were tested on a wide range of test problems and they are valid for different problem sizes.

The bi-objective Pareto front is updated at each generation. Therefore, the size of the Pareto front is dynamic. On the other hand, the single objective ES is solved for maximum capacitated resilience under a budget constraint. The single objective “Pareto front” was generated using different budget values, i.e., 26 problem instances with different budget constraints (from 350 to 600 in increments of 10) are solved for maximum resilience. The best solutions of these 26 runs were combined and the non-dominated solutions are extracted to form the “Pareto front” of the single objective ES. For the bi-objective Pareto front, eight replications (different random number seeds) were used for each problem instance, however, only one random number seed was used to generate the single objective Pareto front.

The Pareto fronts of the single and bi-objective ES are compared in Section 5.1. In Section 5.2, the bi-objective ES is run longer to match the time of the single objective Pareto front generation process (the total time of 26 single objective runs) and its performance is compared with the single objective version. The variation of the single objective ES is investigated in Section 5.3. Note that variation in the bi-objective ES has been investigated with eight random number seeds in Sections 5.1 and 5.2.

5.1 Single objective vs. bi-objective ES

For each problem instance, eight random number seeds were used for the bi-objective optimization, and they were compared with the solutions that were generated from the single objective ES.

5.1.1 Pareto front diversity

The ranges of capacitated resilience and the cost of single and bi-objective Pareto front solutions are given in Figures 5.1, 5.2 and 5.3. These are the non-dominated solutions. To
find the non-dominated solutions, single and bi-objective Pareto fronts are compared with each other. A solution of the single objective Pareto front is non-dominated if it is not dominated by any solution of the bi-objective Pareto front. Non-dominated solutions of bi-objective ES are found similarly. The total number of solutions in each Pareto front and the number of non-dominated solutions are given in the next section.

The performance of the bi-objective ES varies with random number seed. Only one random number seed is used to generate the single objective Pareto front due to the long solution time (requires 26 runs) of the Pareto front generation process. The variation of the single objective ES is investigated in Section 5.3. For example, random number seed five of problem instance two has a better Pareto front diversity and a better capacitated resilience value than the single objective Pareto front. However, random number seed seven of problem instance one has a worse diversity and lower capacitated resilience value. For problem instances two and three, bi-objective ES outperforms single objective in terms of capacitated resilience. However, single objective performs better for problem instance one. Therefore, the performance comparison for capacitated resilience is not conclusive. However, the bi-objective Pareto front has a better diversity in terms of cost.

5.1.2 Comparison of non-dominated solutions for bi-objective and single objective Pareto fronts

The Pareto front solutions of bi-objective ES were compared with the Pareto front solutions of single objective ES, and vice versa.

In Table 5.1, the non-dominated solutions are reported for each Pareto Front. In the table, the non-dominated solutions of bi-objective ES were found by calculating the number of bi-objective Pareto front solutions that are not dominated by single objective Pareto front solutions. The single objective Pareto front solutions were compared with the bi-objective Pareto front solutions in the same way. Bi-Objective outperforms single objective for some scenarios (e.g., random number seed four of problem instance three) but single objective
Figure 5.1: Comparison of Pareto front solutions (Problem instance 1)
Figure 5.2: Comparison of Pareto front solutions (Problem instance 2)
Figure 5.3: Comparison of Pareto front solutions (Problem instance 3)
dominates bi-objective for most scenarios. However, these results can be misleading because
the solution times of bi-objective and single objective are not equal (single objective combines
26 different runs to generate the Pareto front). Therefore, Section 5.2 repeats the same
analysis presented in this section where the bi-objective optimizer runs longer.

Table 5.1: Performance of single and bi-objective ES for three problem instances

<table>
<thead>
<tr>
<th>Method</th>
<th>Bi-Objective</th>
<th>Single Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rnd Number Seed 1</td>
<td>4/14</td>
<td>6/6</td>
</tr>
<tr>
<td>Rnd Number Seed 2</td>
<td>8/14</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 3</td>
<td>1/15</td>
<td>6/6</td>
</tr>
<tr>
<td>Rnd Number Seed 4</td>
<td>2/10</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 5</td>
<td>5/15</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 6</td>
<td>6/11</td>
<td>3/6</td>
</tr>
<tr>
<td>Rnd Number Seed 7</td>
<td>3/11</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 8</td>
<td>2/15</td>
<td>6/6</td>
</tr>
</tbody>
</table>

* Number of nondominated solutions (single and bi-objective Pareto fronts are
compared with each other) / total solutions in the Pareto front

5.1.3 Single and bi-objective Pareto Fronts

This section shows the Pareto fronts of both single and bi-objective ES. Specifically, it presents the diversity of the Pareto front solutions. Also, single and bi-objective Pareto front solutions are graphically compared. Figures 5.4, 5.5 and 5.6 compare the population and Pareto front members of single and bi-objective ES. Population members of the single
objective are the 26 solutions that are generated using different budget constraints. Non-dominated solutions of the single objective population members form the single objective Pareto front.

Except from some instances (e.g., random number seed four of problem instance three), the single objective outperforms bi-objective. However, the bi-objective Pareto front has better diversity than the single objective. There is also a gap between the bi-objective population and the bi-objective Pareto front in terms of solution quality and therefore these results confirm the need of keeping a global Pareto front in an external repository for bi-objective ES (Section 4.2.3).

5.2 Equalizing computational effort

In this section, different from Section 5.1, the bi-objective ES was run longer to match the time of the single objective Pareto front generation process. To match the time of generating 26 different budget values, the bi-objective optimizer was run for 16,000 generations with an early termination of 4000 non-improving generations. The rationale of using 16,000 generations instead of 26,000 (26 runs of 1000 generations each) is that the total time of 26 single objective runs for budget values from 350 to 600 is equivalent to the running time of 16,000 bi-objective generations. For each problem instance, the single objective was run for only one random number seed (due to the long process of generating single objective Pareto front) and the bi-objective optimizer was run for 10 random number seeds. Variation in the single objective ES is discussed in Section 5.3.

5.2.1 Pareto front diversity

Figures 5.7, 5.8 and 5.9 summarize the diversity of the bi-objective and single objective Pareto fronts. For three problem instances, bi-objective ES has a better diversity (in cost and capacitated resilience) than the single objective version. Also, the capacitated resilience
Figure 5.4: Pareto fronts for problem instance 1: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds)
Figure 5.5: Pareto fronts for problem instance 2: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds)
Figure 5.6: Pareto fronts for problem instance 3: Single objective (obtained by 26 different budget values) and bi-objective (for different random number seeds)
values have been improved over the shorter bi-objective runs with the longer solution times (in comparison to Section 5.1).

5.2.2 Comparison of non-dominated solutions for bi-objective and single objective Pareto fronts

The Pareto front solutions of bi-objective ES were compared with the Pareto front solutions of single objective ES, and vice versa. Table 5.2 summarizes the number of non-dominated solutions in each combined Pareto front. As expected, running the bi-objective
Figure 5.8: Comparison of Pareto front solutions (Problem instance 2)
Figure 5.9: Comparison of Pareto front solutions (Problem instance 3)
Table 5.2: Comparison of Pareto solutions of single (1000 generations) and bi-objective (16,000 generations) ES for three problem instances

<table>
<thead>
<tr>
<th>Method</th>
<th>Bi-Objective</th>
<th>Single Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rnd Number Seed 1</td>
<td>6/13</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 2</td>
<td>9/12</td>
<td>2/6</td>
</tr>
<tr>
<td>Rnd Number Seed 3</td>
<td>3/10</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 4</td>
<td>5/12</td>
<td>6/6</td>
</tr>
<tr>
<td>Rnd Number Seed 5</td>
<td>8/16</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 6</td>
<td>9/14</td>
<td>4/6</td>
</tr>
<tr>
<td>Rnd Number Seed 7</td>
<td>11/18</td>
<td>5/6</td>
</tr>
<tr>
<td>Rnd Number Seed 8</td>
<td>9/18</td>
<td>6/6</td>
</tr>
</tbody>
</table>

Problem Instance 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Bi-Objective</th>
<th>Single Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rnd Number Seed 1</td>
<td>6/14</td>
<td>7/8</td>
</tr>
<tr>
<td>Rnd Number Seed 2</td>
<td>3/13</td>
<td>8/8</td>
</tr>
<tr>
<td>Rnd Number Seed 3</td>
<td>5/13</td>
<td>7/8</td>
</tr>
<tr>
<td>Rnd Number Seed 4</td>
<td>8/18</td>
<td>6/8</td>
</tr>
<tr>
<td>Rnd Number Seed 5</td>
<td>7/19</td>
<td>8/8</td>
</tr>
<tr>
<td>Rnd Number Seed 6</td>
<td>6/18</td>
<td>7/8</td>
</tr>
<tr>
<td>Rnd Number Seed 7</td>
<td>7/13</td>
<td>7/8</td>
</tr>
<tr>
<td>Rnd Number Seed 8</td>
<td>7/13</td>
<td>7/8</td>
</tr>
</tbody>
</table>

Problem Instance 3

<table>
<thead>
<tr>
<th>Method</th>
<th>Bi-Objective</th>
<th>Single Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rnd Number Seed 1</td>
<td>9/15</td>
<td>5/9</td>
</tr>
<tr>
<td>Rnd Number Seed 2</td>
<td>10/15</td>
<td>4/9</td>
</tr>
<tr>
<td>Rnd Number Seed 3</td>
<td>12/17</td>
<td>3/9</td>
</tr>
<tr>
<td>Rnd Number Seed 4</td>
<td>11/16</td>
<td>3/9</td>
</tr>
<tr>
<td>Rnd Number Seed 5</td>
<td>14/17</td>
<td>3/9</td>
</tr>
<tr>
<td>Rnd Number Seed 6</td>
<td>5/12</td>
<td>8/9</td>
</tr>
<tr>
<td>Rnd Number Seed 7</td>
<td>12/14</td>
<td>4/9</td>
</tr>
<tr>
<td>Rnd Number Seed 8</td>
<td>7/13</td>
<td>5/9</td>
</tr>
</tbody>
</table>

* Number of nondominated solutions (single and bi-objective Pareto fronts are compared with each other) / total solutions in the Pareto front

version longer helped improve the Pareto front; the bi-objective Pareto front has better solutions (in terms of number of non-dominated solutions) than the ones in Section 5.1. And, the number of solutions in the bi-objective Pareto front is increased. Comparing to the single objective Pareto front, the bi-objective Pareto front performs better for problem instances one and three (for most random number seeds), however, the single objective Pareto front has more non-dominated solutions for problem instance two (except for random number seed four).
5.3 Variation of single objective ES

Variation of the single objective ES is summarized in this section. To show the variation in capacitated resilience and cost, three problem instances of the single objective ES was run for 10 random number seeds under different budget constraints (for 1000 generations).

Figure 5.10 summarizes the results. Interestingly, there is variation in cost, which means that the optimizer can not always fully utilize the budget. This also causes variation in the capacitated resilience. To see the relationship between the variation and the number of generations, the same experiment was performed with 5000 generations (Figure 5.11). A similar variation has been observed for 5000 generations. In summary, the single objective ES is sensitive to random number seed.

A comparison of the variability of single and bi-objective ES is provided in Section 5.4.

5.4 Summary

According to the preliminary results presented in this chapter, the performance of single and bi-objective ES are comparable in terms of Pareto front diversity, capacitated resilience and cost. The Pareto front diversity of the bi-objective version was compared with the single objective Pareto front and it was found satisfactory (in solution quality and diversity) given that the single objective Pareto front requires a longer time to construct. According to the two-sample t-tests, the single objective ES found better capacitated resilience values than bi-objective only for problem instance 4 (p-value=0.0202) of the budget=500 case and problem instances 3 (p-value=0.0220) and 10 (p-value=0.1002) of the budget=600 case. For the remaining problem instances, the difference in capacitated resilience between single and bi-objective ES was not statistically significant. Also, the difference in cost was not statistically significant in any of the problem instances. The advantage of using bi-objective over single objective ES is that the bi-objective provides a non-dominated set of solutions instead of a single solution. A decision maker can decide on the cost and capacitated resilience trade-off
Figure 5.10: Variation of the single objective ES for different budget values
and select the best non-dominated solution. Also, as seen in Sections 5.1 and 5.2, single run of bi-objective ES provides a better diversity than multiple runs of single objective ES and runs faster.

Both single and bi-objective ES have some variation in solution quality. Table 5.3 shows the variation in capacitated resilience of the single and bi-objective ES for the 10 user scenario with two budget levels (10 problem instances with 10 random number seeds each). The standard deviation of the runs does not exceed 0.105 for the bi-objective (problem instance 2/budget=500) and 0.166 for single objective (problem instance 9/budget=600). Also, the standard deviations of the single and bi-objective ES are very similar. According to the F-tests for equal variance of capacitated resilience, the variances of single and bi-objective ES are not statistically different except in four problem instances: problem instances 4 (p-value=0.0351), 7 (p-value=0.0343) and 8 (p-value=0.0435) of the budget=500 case and problem instance 9 (p-value=0.0105) of the budget=600 case. It can be concluded that the variation to seed in capacitated resilience is acceptable (standard deviation is 0.05 on average) for both single and bi-objective ES.
This chapter analyzed the performance of the proposed model on a small sized problem. Succinctly, the performance of the bi-objective ES is comparable to the single objective. The next chapter compares capacitated resilience with other reliability/survivability metrics.
Table 5.3: Mean and standard deviation of capacitated resilience for single and bi-objective runs (10 problem instances of 10 user scenario with 10 random number seeds for each problem instance. Capacitated resilience is calculated for number of alternative paths=10 and cut set size=4.)

<table>
<thead>
<tr>
<th>Problem Instance</th>
<th>Budget = 500</th>
<th></th>
<th></th>
<th>Budget = 600</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
<td>Bi-Objective</td>
<td>Single</td>
<td>Bi-Objective</td>
<td>Single</td>
<td>Bi-Objective</td>
</tr>
<tr>
<td></td>
<td>Objective</td>
<td></td>
<td>Std. Deviation</td>
<td></td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td>1</td>
<td>0.53649</td>
<td>0.02250</td>
<td>0.55387</td>
<td>0.03767</td>
<td>0.68558</td>
<td>0.04548</td>
</tr>
<tr>
<td>2</td>
<td>0.65178</td>
<td>0.06636</td>
<td>0.68746</td>
<td>0.10526</td>
<td>0.75458</td>
<td>0.06276</td>
</tr>
<tr>
<td>3</td>
<td>0.72454</td>
<td>0.05161</td>
<td>0.69514</td>
<td>0.06401</td>
<td>0.79778&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.02430</td>
</tr>
<tr>
<td>4</td>
<td>0.84978&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.02366&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.81169</td>
<td>0.00691</td>
<td>0.84355</td>
<td>0.01812</td>
</tr>
<tr>
<td>5</td>
<td>0.75027</td>
<td>0.04256</td>
<td>0.71340</td>
<td>0.02207</td>
<td>0.84719</td>
<td>0.04205</td>
</tr>
<tr>
<td>6</td>
<td>0.64184</td>
<td>0.05897</td>
<td>0.66774</td>
<td>0.05940</td>
<td>0.82046</td>
<td>0.04186</td>
</tr>
<tr>
<td>7</td>
<td>0.75881</td>
<td>0.05361&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.74815</td>
<td>0.01556</td>
<td>0.81342</td>
<td>0.02375</td>
</tr>
<tr>
<td>8</td>
<td>0.70869</td>
<td>0.01769</td>
<td>0.70261</td>
<td>0.05704&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.82065</td>
<td>0.03138</td>
</tr>
<tr>
<td>9</td>
<td>0.82082</td>
<td>0.04539</td>
<td>0.80234</td>
<td>0.02511</td>
<td>0.79605</td>
<td>0.16603&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>10</td>
<td>0.59992</td>
<td>0.05781</td>
<td>0.63547</td>
<td>0.02829</td>
<td>0.77596&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.04754</td>
</tr>
</tbody>
</table>

<sup>a</sup> Significantly larger mean (p-value < 0.05)

<sup>b</sup> Significantly larger standard deviation (p-value < 0.05)
Chapter 6

Results

6.1 Comparison of single (capacitated resilience) and bi-objective (cost and capacitated resilience) ES

In this section, single and bi-objective ES are compared for capacitated resilience. The Pareto front comparison of single and bi-objective ES was presented in Section 5.1, therefore Pareto fronts are not analyzed here.

Figures 6.1 through 6.5 summarize the results of the 10, 25, 50, 75 and 100 users scenarios, respectively. These figures compare the best capacitated resilience values of the single and bi-objective runs. Ten problem instances with five replications are analyzed for each scenario. The capacitated resilience in these figures is exactly calculated by setting the number of alternative paths to 10 and the cut set size to 4.

As seen from the figures, the single objective ES outperforms bi-objective in most problem instances. This is expected because the bi-objective ES maintains a diversified Pareto front to optimize the two objective functions simultaneously. For some problem instances (e.g., problem instance 1 of the 50 user scenario) the bi-objective ES finds better capacitated resilience values than the single objective. However, according to the two-sample t-tests, single objective is significantly better for only some problem instances. The single objective was significantly better for the 10 user scenario problem instance 3 with budget 600 (p-value=0.0220), problem instance 4 with budget 500 (p-value=0.0202), problem instance 10 with budget 600 (p-value=0.0102). The single objective ES was also statistically better than bi-objective ES for the 25 user scenario problem instance 2 with budget 1000 (p-value=0.0345) and budget 1200 (p-value=0.0053), problem instance 5 with budget 1200 (p-value=0.0111), problem instance 7 with budget 1000 (p-value=0.0011), problem instance
9 with budget 1000 (p-value=0.0049) and budget 1200 (p-value=0.0073), problem instance 10 with budget 1000 (p-value=0.0009) and budget 1200 (p-value=0.0257). For the 75 user scenario, the single objective was better than the bi-objective only for problem instance 2 with budget 2700 (p-value=0.0057) and problem instance 8 with budget 2700 (p-value=0.00185). For the 50 and 100 user scenarios the single objective ES was not significantly better for any problem instance. On the other hand, there was no problem instance that the bi-objective ES was statistically better than the single objective ES. To sum up, according to the results of the t-tests, the performance of the bi-objective ES is comparable to the single objective ES.

The performance gap between single and bi-objective optimizers reduces as the problem size increases (from 50 users to 100 users). For the 100 user scenario (Figure 6.5) bi-objective optimization was able to find better solutions than the single objective for most problem instances. For example, bi-objective was statistically better than the single objective for problem instance 8 at 10% significance level (p-value=0.0622). This may suggest that the bi-objective performs well for larger sized problems. Its diversified population due to the Pareto front helps to search the solution space more effectively than the single objective.

6.2 Correlation among capacitated resilience and other metrics

In this section, correlations among the capacitated resilience and the other metrics are investigated. The main reason to check correlations is to identify possible differences or similarities of the network structures.

Table 6.1 summarizes the correlations using the best solutions over five random seeds of each problem instance (total of 10). In this table, each row shows correlations among the objective function and the other metrics. For example, the first row presents the correlations among capacitated resilience and the other metrics for the network designs found by optimization for capacitated resilience. The correlation between capacitated resilience
Figure 6.1: Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 10 user scenario.
Figure 6.2: Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 25 user scenario
Figure 6.3: Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 50 user scenario.
Figure 6.4: Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 75 user scenario
Figure 6.5: Comparison of single and bi-objective ES performance for capacitated resilience (using number of alternative paths and size of cut-sets as 10 and 4, respectively), 10 problem instances with 5 replications of the 100 user scenario.
and all-terminal reliability is high because both capacitated resilience and all-terminal reliability are calculated by considering all paths from a user to all available APs. There is a similarity between their network designs such that both favor redundancy in their designs. Also, the correlation of two-terminal reliability and all-terminal reliability is high when optimizing for two-terminal reliability because two-terminal reliability is simply a subset of all-terminal reliability where only one AP is considered. However, the correlation between two-terminal and all-terminal reliabilities is low when optimizing for all-terminal reliability. Therefore, there is a difference between the designs of two-terminal and all-terminal reliabilities. Similarly, the correlation between capacitated resilience and two-terminal reliability is high when optimizing for capacitated resilience but the correlation is low when optimizing for two-terminal reliability. Hence, there are differences in their network designs. Also, the correlation between capacitated resilience and TE is low and there are significant differences in their network designs. The correlations among traffic efficiency, two-terminal and all-terminal reliabilities are high when optimizing for traffic efficiency, however, the correlations are low when optimizing for two-terminal or all-terminal reliabilities. Therefore, the network designs of traffic efficiency, two-terminal reliability and all-terminal reliability are different. The similarities and the differences in the network designs are discussed in detail in the Section 6.3.

Table 6.1: Correlations calculated using the best solutions of 10 problem instances of the 10 user scenario

<table>
<thead>
<tr>
<th>Optimized for</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
</tr>
<tr>
<td>CR</td>
<td>-</td>
</tr>
<tr>
<td>TE</td>
<td>0.3075</td>
</tr>
<tr>
<td>Two-term.</td>
<td>0.2991</td>
</tr>
<tr>
<td>All-term.</td>
<td>0.7318</td>
</tr>
</tbody>
</table>

* Correlation between capacitated resilience and traffic efficiency (when the network is designed for capacitated resilience)
6.3 Network structure: Comparison of capacitated resilience and other metrics

In this section, the differences between the network structures obtained by optimization for capacitated resilience and the other metrics are compared. For this comparison, different problem instances are solved for each metric. Input data of a problem instance consists of user locations and traffic requirements. In the proposed ES model of this dissertation, a user represents an area which may consist of some users. If the number of users in the optimization model were equal to the number in the physical world, then that model would have a real representation of users. However, this increases the number of users dramatically and therefore makes the problem intractable for real life applications. Hence, the traffic requirements of users represent the total traffic requirements of the users in an area. Also, because of this aggregation of data, the user locations are combined to a single location in the model for simplification.

The main goal of this section is to identify the key differences in the network structures in terms of number, types and locations of the devices. For this comparison 10 user and 25 user scenarios are used due to their small problem size (which helps to visually detect the differences in the network structures). The single objective ES was run for each metric and the resulting network structures are presented.

The network designs for capacitated resilience are obtained by two cases: unconstrained and constrained capacity cases. For the first case, capacities of APs and RPs are set to 150 and 80 (large enough to assume uncapacitated operation), respectively. For the constrained capacity case, capacities are set to 30 and 20, respectively. There are two cases for capacitated resilience for two reasons. The first is to make a fair comparison with other metrics, one basically uncapacitated and the other quite capacitated. Recall that the other metrics do not consider capacity. The second is to assess the effects of different levels of capacity constraint on the network designs.
6.3.1 A problem instance of the 10 user scenario

Figure 6.6 presents the input data of problem instance 1 of the 10 user scenario. The devices are deployed to optimize capacitated resilience, TE, two-terminal reliability and all-terminal reliability for the given user coordinates and traffic requirements.

Network structures obtained by optimization for capacitated resilience

Problem instance 1 of the 10 user scenario is optimized for capacitated resilience for two capacity cases. First, the unconstrained capacity case is solved for budget constraints of 500 and 600. Figure 6.7 shows the network structure obtained by optimization of problem instance 1 for capacitated resilience under a budget constraint of 500. The network structure shows the allocation of device redundancy in the high traffic requirements areas. U0, U1 and U8 are the low traffic flow users (which have smaller weights in the capacitated resilience calculation), therefore they are not prioritized by the ES. In other words, APs (or RPs) are not located near these users because their effect on the capacitated resilience is limited due to their low traffic requirements (see Equation 3.1 for the definition of user weight). For example, the two APs (AP5 and AP8) are not located close to U0 due to its low traffic requirement. Also, AP5 is located far from U5 but close to U9. The reason is to increase the reliability of the alternative path (U9 - AP5) without losing connectivity of U5 and U0. The traffic requirements of U9 and U5 are very high (14.803 and 15.753, respectively) and it might be argued that the location of AP8 nearer U5 would provide a slightly higher capacitated resilience because of its higher traffic requirement. However, this improvement would be very limited due to the small difference between the traffic requirements of U9 and U5. If the budget was larger (e.g., 600) capacitated resilience would be increased by adding a second AP close to U5 to improve its assigned path and use AP5 as its alternative path. Similarly the ES preferred to allocate redundancy for U2 and U3 instead of U7 and U8. This decision is reasonable because U8 is a low traffic node (3.737) and U3 and U7 are comparable in terms of their traffic requirements (14.729 and 14.269). Also, locating redundancies around
Figure 6.6: Inputs of 10 users scenario (problem instance 1): User locations ($x$ and $y$ are the coordinates, $f$ is the traffic flow requirement)
U2 and U3 helps to further improve capacitated resiliences of those nodes because the nodes are located close to each other and this improves the network capacitated resilience. The capacitated resilience of this network is 0.5054 and the TE is 0.3043.

Figure 6.8 shows the network structure of the same problem with a larger budget (600 instead of 500). As the budget increases, redundancies are added to U0, U4, U5, U6, U7 and U8. On the other hand, U9 has lost its redundant path due to this new arrangement of the devices. However, since the traffic flow requirements of U7 and U9 are comparable (14.269 and 14.803, respectively) this new arrangement does not have an adverse effect on the network level capacitated resilience. Also, locating devices in the more crowded areas (in this case, the areas of U7, U8, U3 and U2) helps improving the overall capacitated resilience. Note that U1 still has no redundancy, which is similar to the budget=500 scenario, due to its low traffic requirement (which is 5.11).

Figures 6.9 and 6.10 show the solutions for constrained capacity case. For the same problem instance, the capacity of the APs and RPs are set to 30 and 20, respectively. Both network designs, with budget=500 (Figure 6.9) and budget=600 (Figure 6.10) have lower capacitated resilience values than the previous designs (Figures 6.7 and 6.8) due to the limited capacity. However, the designs are similar because the users are connected with an AP and then the redundancies are provided for high traffic users. Isolated users (such as U0, U1, U4 and U5 of Figure 6.9) are connected with distant APs that are primarily serving the high traffic users. For the budget constraint of 600 case (Figure 6.10), similar design rules have been observed. High traffic user (e.g., U5 and U7) have strong connections with a high level of redundancy ensured by nearby APs. Isolated users (such as, U1) have no or limited redundancy. The capacity constraint does not change the design rules.

One of the common design rules from the network structure obtained by optimizing for capacitated resilience is that the redundancy is allocated to the most promising area (crowded or having larger traffic nodes) instead of to isolated nodes or low traffic nodes. As the capacitated resilience of a user becomes zero if there are no alternative paths available,
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
490, 0.5054, 0.3043, 0.604, 0.6114

# of APs = 8, # of RPs = 1

Figure 6.7: Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for Capacitated Resilience with unconstrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
570, 0.6429, 0.642, 0.7675, 0.7875
# of APs = 9, # of RPs = 3

Figure 6.8: Network structure of the 10 user problem (problem instance 1, budget=600)
found by optimization for Capacitated Resilience with unconstrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
500, 0.4975, 0.4502, 0.6556, 0.7047
# of APs = 8, # of RPs = 2

Figure 6.9: Network structure of the 10 user problem (problem instance 1, budget=500)
found by optimization for Capacitated Resilience with constrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
570, 0.5839, 0.5826, 0.7408, 0.7651

# of APs = 9, # of RPs = 3

Figure 6.10: Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for Capacitated Resilience with constrained capacities
the location of the assigned devices are strategically made to improve reliabilities of the alternative paths of other users (see U1 and AP0 in Figure 6.8).

**Capacitated resilience vs. TE**

Figure 6.11 shows the network structure generated by traffic efficiency optimization under budget constraint of 500. The major difference of this network between the one generated by capacitated resilience is the redundancy allocation. In the one generated by TE optimization, there is no or very limited redundancy in the network. Instead an AP is located for most users. For the users that cannot have a “dedicated” access point, they share an access point which is close to another user. For example, U0 and U8 are the low traffic nodes and they are connected to the network by a distant AP. However, this decision of assigning the dedicated APs to high traffic nodes is needed to maximize TE.

In terms of redundancy, the network has very limited redundancy but high reliability for the node connections. Only U2, U3, U4, U6, U7 and U8 have alternative paths. The other users have no alternative paths, therefore their capacitated resilience values are zero. The capacitated resilience and TE of this network are 0.2056 and 0.8577, respectively. In comparison to the network optimized by capacitated resilience (Figure 6.7), TE improved from 0.3043 to 0.8577, but the capacitated resilience reduced dramatically from 0.5054 to 0.2056 due to lack of alternative paths in the network. This difference is caused by the structural differences between the two networks.

Figure 6.12 has a larger budget (600 instead of 500), but the structure of the resulting network is similar to the one of smaller budget (500). The devices are located very close to users so that the reliabilities between users and APs are very high. Only U0 does not have an AP located nearby, but U0 has a very small traffic requirement (0.023) therefore it does not have a significant effect on the TE. The capacitated resilience of the network is 0.2039, which is similar to one given in Figure 6.11, but the TE has improved from 0.8577 to 0.9854 because of the larger number of RPs and their locations.
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
490, 0.2056, 0.8577, 0.9759, 0.9795

# of APs = 8, # of RPs = 1

Figure 6.11: Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for TE
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
510, 0.2039, 0.9854, 0.977, 0.9787
# of APs = 8, # of RPs = 3

Figure 6.12: Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for TE

128
In summary, the typical network structure obtained by TE optimization is to have high reliability for high traffic nodes by locating APs near them, and to have some redundancy for high traffic nodes within the given budget and allow low reliability connections for low traffic nodes (if the budget is tightly constrained).

**Capacitated resilience vs. two-terminal reliability**

As seen in Figure 6.13, which is similar to the structure obtained by TE optimization, the APs (or RPs) are located very close to the users to get maximum reliability. U0 is assigned to an AP which is not very close to its location, (which was also seen in Figure 6.12), but it does not affect two-terminal reliability significantly due to its low weight in the network level two-terminal reliability calculation. The two-terminal reliability of this network is 0.9802 and capacitated resilience is 0.1054. Similar to TE optimization, the main reason of the low capacitated resilience value for two-terminal optimization is the lack of redundancy in the network.

As the budget increases from 500 to 600 (Figure 6.14), an AP has been assigned to the U0 user to increase overall two-terminal reliability. Also, an additional device has been located near U6, which is a high traffic node with traffic requirement of 19.662. This improved capacitated resilience from 0.1054 to 0.3695, but two-terminal reliability improved slightly (from 0.9802 to 0.9831). The network structure remained similar.

Thus, the structure obtained by two-terminal reliability optimization is very similar to the one obtained by TE optimization. It allocates APs very close to the high traffic nodes. If the budget is not available, the low traffic nodes are connected to a distant AP which serves a nearby high traffic user.

**Capacitated resilience vs. all-terminal reliability**

The network structure found by all-terminal reliability optimization is very similar to the one of two-terminal reliability. With a budget constraint of 500, APs are located very
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
490, 0.1910, 0.7134, 0.9855, 0.9869

# of APs = 8, # of RPs = 1

Figure 6.13: Network structure of the 10 user problem (problem instance 1, budget=500)
found by optimization for two-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
560, 0.3695, 0.8565, 0.9861, 0.9862

# of APs = 9, # of RPs = 2

Figure 6.14: Network structure of the 10 user problem (problem instance 1, budget=600)
found by optimization for two-terminal reliability
close to high traffic nodes (Figure 6.15). When the budget increases to 600 (Figure 6.16), some redundancies are added to the high traffic nodes and all-terminal reliability increases from 0.9831 to 0.9877. Not surprisingly, capacitated resilience has increased from 0.2836 to 0.3298 because of the new alternative paths.

It can be said that the network structure obtained by optimization for all-terminal reliability is very similar to the one obtained from two-terminal reliability optimization. However, the redundancies for all-terminal reliability are more decentralized. For example, in Figure 6.16, AP2 and AP12 provide redundancies for all users on the left side (i.e., \( x < 0 \)) of the network. This is different than Figure 6.14 in which the redundancies are located near the users for two-terminal maximization.

Table 6.2 summarizes the values of all metrics for all network designs presented in this section. Figure 6.17 and 6.18 summarize the comparison of the network designs presented in this section.

Table 6.2: Summary of all metrics for the 10 user scenario, problem instance 1

<table>
<thead>
<tr>
<th>Budget</th>
<th>Optimized by</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CR</td>
</tr>
<tr>
<td>CR (unconstrained capacities)</td>
<td>0.5054</td>
<td>0.3043</td>
</tr>
<tr>
<td>CR</td>
<td>0.4975</td>
<td>0.4502</td>
</tr>
<tr>
<td>TE</td>
<td>0.2056</td>
<td>0.8577</td>
</tr>
<tr>
<td>Two-terminal rel. (2-term)</td>
<td>0.1910</td>
<td>0.7134</td>
</tr>
<tr>
<td>All-terminal rel. (all-term)</td>
<td>0.3284</td>
<td>0.7238</td>
</tr>
<tr>
<td>CR (unconstrained capacities)</td>
<td>0.6429</td>
<td>0.6420</td>
</tr>
<tr>
<td>CR</td>
<td>0.5839</td>
<td>0.5826</td>
</tr>
<tr>
<td>TE</td>
<td>0.2039</td>
<td>0.9854</td>
</tr>
<tr>
<td>Two-terminal rel. (2-term)</td>
<td>0.3695</td>
<td>0.8565</td>
</tr>
<tr>
<td>All-terminal rel. (all-term)</td>
<td>0.3298</td>
<td>0.7115</td>
</tr>
</tbody>
</table>

6.3.2 Another problem instance of the 10 user scenario

In this section the network structures obtained by optimization for different metrics are compared using another problem instance of the 10 user scenario. The location and flows of the users are given in the Figure 6.19.
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
500, 0.3284, 0.7238, 0.9852, 0.9871
# of APs = 8, # of RPs = 2

Figure 6.15: Network structure of the 10 user problem (problem instance 1, budget=500) found by optimization for all-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
600, 0.3298, 0.7115, 0.972, 0.9877
# of APs = 10, # of RPs = 0

Figure 6.16: Network structure of the 10 user problem (problem instance 1, budget=600) found by optimization for all-terminal reliability
Figure 6.17: Summary of the designs of the problem instance 1 of the 10 user scenario, budget=500
Figure 6.18: Summary of the designs of the problem instance 1 of the 10 user scenario, budget=600
Figure 6.19: Inputs of the 10 user scenario (problem instance 2): User locations (x and y are the coordinates, f is the traffic flow requirement)
Figure 6.20 shows the network optimized for maximum capacitated resilience. Similar to the network design found for problem instance 1 (Section 6.3.1) this network emphasises strong connections for high flow nodes and redundancies around those nodes. Similarly, the constrained capacity case of the same problem instance (solved for capacitated resilience) yields a similar network design (Figure 6.21) with slightly less capacitated resilience value (0.6753 instead of 0.7461) due to limited capacities. On the other hand, the network structure of TE (Figure 6.22) focuses on strong connections for most of the nodes. Redundancy is not as important as for capacitated resilience.

The networks found for two-terminal and all-terminal reliabilities (Figures 6.23 and 6.24) provide strong connections for high flow nodes, but also some level of redundancy are allocated by RPs. In capacitated resilience, the redundancies for high flow nodes are mostly ensured by APs but the two-terminal and all-terminal reliability designs use RPs to provide redundancies in this problem instance because APs are mainly used to create strong connections (between users and APs) for two and all-terminal reliability designs and RPs provide alternative paths at a lower cost.

The summary of the network designs is given in Table 6.3. The designs presented in this section are compared in Figure 6.25.

Table 6.3: Summary of all metrics for the 10 user scenario, problem instance 2

<table>
<thead>
<tr>
<th>Budget</th>
<th>Optimized by</th>
<th></th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CR</td>
<td>TE</td>
</tr>
<tr>
<td>CR (unconstrained capacities)</td>
<td>0.7461</td>
<td>0.6308</td>
<td>0.8169</td>
</tr>
<tr>
<td>CR</td>
<td>0.6753</td>
<td>0.4523</td>
<td>0.7830</td>
</tr>
<tr>
<td>600 TE</td>
<td>0.3655</td>
<td>0.9101</td>
<td>0.9036</td>
</tr>
<tr>
<td>Two-terminal rel. (2-term)</td>
<td>0.5248</td>
<td>0.5361</td>
<td>0.9787</td>
</tr>
<tr>
<td>All-terminal rel. (all-term)</td>
<td>0.5210</td>
<td>0.8247</td>
<td>0.9764</td>
</tr>
</tbody>
</table>

6.3.3 A problem instance of the 25 user scenario

As another example of differences the network design, a larger problem is presented in this section. Figure 6.26 shows the user locations and traffic requirements of the 25 user
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
580, 0.7461, 0.6308, 0.8169, 0.8363

# of APs = 9, # of RPs = 4

Figure 6.20: Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for Capacitated Resilience with unconstrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
600, 0.6753, 0.4523, 0.7830, 0.8306

# of APs = 10, # of RPs = 0

Figure 6.21: Network structure of the 10 user problem (problem instance 2, budget=600)
found by optimization for Capacitated Resilience with constrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:

570, 0.3655, 0.9101, 0.9036, 0.9412

# of APs = 9, # of RPs = 3

Figure 6.22: Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for TE
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
580, 0.5248, 0.5361, 0.9787, 0.9844
# of APs = 9, # of RPs = 4

Figure 6.23: Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for two-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
590, 0.521, 0.8247, 0.9764, 0.9845
# of APs = 9, # of RPs = 5

Figure 6.24: Network structure of the 10 user problem (problem instance 2, budget=600) found by optimization for all-terminal reliability
Figure 6.25: Summary of the designs of the problem instance 2 of the 10 user scenario, budget=600
Figure 6.26: Inputs of the 25 user scenario (problem instance 1): User locations \((f\) is the traffic flow requirement)

scenario. Similar to the previous examples, this problem is solved for maximum capacitated resilience, TE, two-terminal reliability and all-terminal reliability and differences in the network designs are discussed.

Solving the 25 user problem for maximum capacitated resilience (Figures 6.27 and 6.28) yields a similar network design found for the previous examples (Sections 6.3.1 and 6.3.2). The high traffic nodes are connected with the devices located close to them and also some redundancy is allocated as the budget allows. Low traffic nodes, for example U15 and U16, are connected with devices which are located further from these users to connect high
traffic nodes or provide redundancy for high traffic nodes. This trend was observed in the previous examples as well. The design with a larger budget (Figure 6.28) has a higher level of redundancy in the network. Similarly, constrained capacity case (Figures 6.29 and 6.30) has similar designs.

Solving the same problem for maximum TE yields a network design with high reliability connections for high traffic nodes. Redundancies are allocated as the budget constraint allows. Unlike with resilience the primary objective is to connect as many high traffic nodes with high reliability devices as possible. Both the 1000 (Figure 6.31) and the 1200 (Figure 6.32) budget scenarios result in similar structures.

For two-terminal (Figures 6.33 and 6.34) and all-terminal (Figures 6.35 and 6.36) reliability design, the network is different than for capacitated resilience. As discussed in the previous sections, these yield a design which is similar to the one obtained by TE maximization. Network designs obtained by maximization of two-terminal and all-terminal reliabilities emphasize high reliability connections for high traffic nodes and add redundancies as the budget permits. Redundancy allocation for two-terminal and all-terminal reliability is different than for capacitated resilience because the redundancies are located near the nodes that do not have a high reliability connection. This increases the chance for those nodes to keep connected to the network.

Table 6.4 summarizes the values of all metrics for all network designs presented in this section. Figures 6.37 and 6.38 compare the designs that are presented in this section.

6.3.4 Summary of the differences of the network designs

According to the examples given in the previous sections, there are some differences in the network designs obtained by optimization for different reliability/survivability metrics.

The most distinct network design is obtained by capacitated resilience maximization. Users are connected to the network with a nearby device to ensure a high reliability connection in capacitated resilience optimization. As seen in the provided examples of the previous
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
990, 0.4667, 0.3904, 0.5942, 0.6305
# of APs = 16, # of RPs = 3

Figure 6.27: Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for Capacitated Resilience with unconstrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
1200, 0.5559, 0.4717, 0.6992, 0.7434

# of APs = 19, # of RPs = 6

Figure 6.28: Network structure of the 25 user problem (problem instance 1, budget=1200)
found by optimization for Capacitated Resilience with unconstrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
990, 0.3958, 0.4795, 0.5751, 0.6199

# of APs = 16, # of RPs = 3

Figure 6.29: Network structure of the 25 user problem (problem instance 1, budget=1000)
found by optimization for Capacitated Resilience with constrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
1170, 0.5165, 0.4823, 0.6909, 0.7549

# of APs = 19, # of RPs = 3

Figure 6.30: Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for Capacitated Resilience with constrained capacities
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
990, 0.0501, 0.8422, 0.8515, 0.8538
# of APs = 16, # of RPs = 3

Figure 6.31: Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for TE
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
1160, 0.0712, 0.9084, 0.8805, 0.8848
# of APs = 18, # of RPs = 8

Figure 6.32: Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for TE
Range of access point
Range of relay point

Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
980, 0.0959, 0.6682, 0.9157, 0.9180

# of APs = 16, # of RPs = 2

Figure 6.33: Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for two-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
1180, 0.2349, 0.6378, 0.9488, 0.9554
# of APs = 19, # of RPs = 4

Figure 6.34: Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for two-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
990, 0.1072, 0.7007, 0.9143, 0.9192
# of APs = 16, # of RPs = 3

Figure 6.35: Network structure of the 25 user problem (problem instance 1, budget=1000) found by optimization for all-terminal reliability
Cost, Capacitated Resilience, TE, Two-terminal Reliability, All-terminal Reliability:
1190, 0.3301, 0.7674, 0.9425, 0.9619
# of APs = 19, # of RPs = 5

Figure 6.36: Network structure of the 25 user problem (problem instance 1, budget=1200) found by optimization for all-terminal reliability
Figure 6.37: Summary of the designs of the problem instance 1 of the 25 user scenario, budget=1000
Figure 6.38: Summary of the designs of the problem instance 1 of the 25 user scenario, budget=1200
Table 6.4: Summary of all metrics for the 25 user scenario, problem instance 1

<table>
<thead>
<tr>
<th>Budget</th>
<th>Optimized by</th>
<th>CR</th>
<th>TE</th>
<th>2-term</th>
<th>All-term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>CR (unconstrained capacities)</td>
<td>0.4667</td>
<td>0.3904</td>
<td>0.5942</td>
<td>0.6305</td>
</tr>
<tr>
<td></td>
<td>CR</td>
<td>0.3958</td>
<td>0.4795</td>
<td>0.5751</td>
<td>0.6199</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.0501</td>
<td>0.8422</td>
<td>0.8515</td>
<td>0.8538</td>
</tr>
<tr>
<td></td>
<td>Two-terminal rel. (2-term)</td>
<td>0.0959</td>
<td>0.6682</td>
<td>0.9157</td>
<td>0.9180</td>
</tr>
<tr>
<td></td>
<td>All-terminal rel. (all-term)</td>
<td>0.1072</td>
<td>0.7007</td>
<td>0.9143</td>
<td>0.9192</td>
</tr>
<tr>
<td>1200</td>
<td>CR (unconstrained capacities)</td>
<td>0.5559</td>
<td>0.4717</td>
<td>0.6992</td>
<td>0.7434</td>
</tr>
<tr>
<td></td>
<td>CR</td>
<td>0.5165</td>
<td>0.4823</td>
<td>0.6909</td>
<td>0.7549</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.0712</td>
<td>0.9084</td>
<td>0.8805</td>
<td>0.8848</td>
</tr>
<tr>
<td></td>
<td>Two-terminal rel. (2-term)</td>
<td>0.2349</td>
<td>0.6378</td>
<td>0.9488</td>
<td>0.9554</td>
</tr>
<tr>
<td></td>
<td>All-terminal rel. (all-term)</td>
<td>0.3301</td>
<td>0.7674</td>
<td>0.9425</td>
<td>0.9619</td>
</tr>
</tbody>
</table>

sections, this is also a very common design rule for other metrics. The main difference of capacitated resilience is with the redundancy allocation. Redundancy is achieved by locating an additional device nearby a user. Capacitated resilience optimization locates the additional device(s) near a user to increase its “resilience”. This approach, providing alternative paths, is at the core of the capacitated resilience calculation (Equation 3.2 in Section 3.2). However, budget constraint limits the number of additional devices. Therefore, redundancies are allocated near high traffic nodes or crowded regions to maximize capacitated resilience as calculated by the weighted average of user level capacitated resiliences in terms of user traffic requirements (Equation 3.1 in Section 3.2). Because of this prioritization some low traffic nodes are connected to the network without redundancies. This is a reasonable strategy because satisfying a larger number of users (or larger traffic requirements) maximizes the overall satisfaction of the network. Also, low traffic users are still connected to the network and some level of redundancy is provided to them within the given budget constraint. The ideal location for a low traffic user would be near a high traffic node so that it can benefit from redundancy of the high traffic node. If the low traffic node is isolated from other users, that user is assigned with a distant device with limited or no redundancy.

TE, two-terminal and all-terminal reliabilities have some common design properties. For example, they try to assign users a high reliability device to connect the network. Although
it seems to be similar to the design of capacitated resilience, this is actually their main
difference from the capacitated resilience. Capacitated resilience adds redundancies near
high traffic nodes, whereas the other metrics try to connect more nodes to a high reliability
device then redundancies are allocated if budget allows. In most cases, the redundancies
in the network generated by TE, two-terminal and all-terminal reliabilities are very limited.
There is a slight difference between the designs of two-terminal and all-terminal reliability.
The design for two-terminal reliability has more emphasis on strong connections, however,
the all-terminal design uses more redundancy in network design. For example, in Figure
6.16 the redundancies are allocated in a decentralized way to serve many users instead of a
limited number of users. On the other hand, two-terminal reliability allocates redundancies
nearby users as observed in Figure 6.14.

The network design of capacitated resilience can provide better connectivity if there are
catastrophic failures or attacks in the network because of its higher level of redundancy in the
network. Any potential attack in the network most likely aims for high traffic areas. These
high traffic areas can be more crowded or consist of users with high traffic requirements.
This dissertation solves network design problem for wireless networks. For other networks,
disconnection may mean transportation breakdowns, telephone breakdowns or interdiction in
military supply service. From the rich literature of the network interdiction area, an earlier
study shows that an attacker’s goal might be removing $n$ edges to reduce the maximum
amount of flow (Wollmer, 1964). This is based on max-flow min-cut network problems.
Since the calculation of capacitated resilience actually takes the “min-cut”s into account, it
provides a more resilient design than the others. The main goal is to keep users connected
to the network. Obviously, a planned network interdiction causes more harm than random
failures or attacks, therefore capacitated resilience provides even more resilient network for
random interdiction (or failures). As further study, network designs obtained by these metrics
will be compared for planned interdictions on the network.
6.4 Bi-objective (TE and capacitated resilience) ES

In previous sections, the bi-objective ES was solved for cost and capacitated resilience. In this section, the bi-objective ES is solved for traffic efficiency and capacitated resilience. As seen from Section 6.3, different network designs are obtained by optimization for TE and for capacitated resilience.

As Section 6.2 summarizes, the correlation between TE and capacitated resilience is weak when optimizing for TE. Therefore, high TE values may not result in high capacitated resiliences. Similarly, high capacitated resilience values may not guarantee high TE values. Figures 6.39 and 6.40 present the Pareto fronts of the bi-objective ES optimized for traffic efficiency and capacitated resilience for the 10 user scenario with budget constraints of 500 and 600, respectively. The search space of the bi-objective ES of capacitated resilience and TE is larger than the one of capacitated resilience and cost due to the continuous values of TE. Therefore, instead of 3000 generations with population size of 30 and children size of 30, ES is run for 5000 generations with population size of 40 and children size of 40. Five random number seeds are used for each scenario. Traffic efficiency is calculated by the simulation method presented in Section 4.3.2 and capacitated resilience is calculated by setting the values of number of alternative paths and cut-set size to 10 and 4, respectively.

The first thing to notice is the number of solutions in the Pareto fronts. Compared to the Pareto front solutions of cost and capacitated resilience (Section 5.1.3), there are more solutions for traffic efficiency and capacitated resilience optimization. The reason is that the search space is larger as both metrics (TE and capacitated resilience) are continuous, whereas cost is discrete.

As expected, a budget increase (from 500 to 600) increases the values of TE and capacitated resilience of most solutions. Specifically, with a budget increase, the highest capacitated resilience increased from 0.5889 (random number seed 2) to 0.6537 (random number seed 4). Similarly, the highest TE increased dramatically (0.8670 for random number seed 4 of
budget=500 case and 0.9665 for random number seed 4 of budget=600). The lowest capacitated resilience remained almost same (0.0345 for random number seed 4 of budget=500 case and 0.0483 for random number seed 2 of budget=600). The lowest TE decreased from 0.5089 (random number seed 5 of budget=500 case) to 0.4238 (random number seed 1 of budget=600 case), however, this lowest value seems to be an extreme case. Without considering the random number seed 1 of budget=600 case, the lowest TE increases to 0.5561.

The two ends of the Pareto front (low TE/high capacitated resilience and high TE/low capacitated resilience) suggest that the structure of the networks are different in these areas. Specifically, the network structures of the solutions in the high TE/low capacitated resilience section have limited redundancies but more reliable connections for most nodes (Figure 6.41). The solutions in the other extreme section (low TE/high capacitated resilience) have more redundancies but some of the low or moderate traffic nodes have less reliable connections (Figure 6.42). However, the solutions in the middle sections of the Pareto fronts have more balanced network designs with reliable connections and some level of redundancy (Figure 6.43). Figure 6.44 provides a comparison of these three different designs.

6.5 Solution time comparison of capacitated resilience and other metrics

In this section, computational experience is presented. The solution methods were coded in Java and the problems were solved on a Linux computer with 2.66 Ghz Intel Quad Core Xeon W3520 CPU and 8 GB memory.

6.5.1 Solution time comparison of capacitated resilience: Effect of number of paths and cut-set size

This section summarizes the solution times of the 10 user scenario (10 problem instances with 10 random number seed for each problem instance) with budget of 500 and 600 for single (capacitated resilience) and bi-objective (cost and capacitated resilience) optimization. The effects of number of alternative paths and cut set sizes are investigated.
Figure 6.39: Pareto fronts for the 10 user scenario problem instance 1 with budget constraint of 500: Bi-objective optimization of TE and capacitated resilience
Figure 6.40: Pareto fronts for the 10 user scenario problem instance 1 with budget constraint of 600: Bi-objective optimization of TE and capacitated resilience
Cost, Capacitated Resilience, TE:
500.0, 0.2278, 0.7953
# of APs = 8, # of RPs = 2

Figure 6.41: Network structure of the 10 user problem (problem instance 1, budget=500) found by bi-objective optimization for capacitated resilience and TE (high TE and low capacitated resilience case)
Cost, Capacitated Resilience, TE:
500.0, 0.5633, 0.5959
# of APs = 8, # of RPs = 2

Figure 6.42: Network structure of the 10 user problem (problem instance 1, budget=500) found by bi-objective optimization for capacitated resilience and TE (low TE and high capacitated resilience case)
Cost, Capacitated Resilience, TE:
500.0, 0.3769, 0.6818
# of APs = 8, # of RPs = 2

Figure 6.43: Network structure of the 10 user problem (problem instance 1, budget=500) found by bi-objective optimization for capacitated resilience and TE (medium TE and medium capacitated resilience case)
Figures 6.45 and 6.46 summarize the solution time per iteration (one ES generation). The reason of using time per iteration instead of total time is that the early termination criterion (250 and 500 non-improving generations for single and bi-objective ES) affects the total solution time. Therefore, solution time per iteration provides a more accurate comparison. As seen from these figures, the bi-objective ES takes significantly more time than the single objective because it runs for 2000 generations with early termination criterion of 500 non-improved generations whereas the single objective runs for 1000 generations with early termination criterion of 250 non-improved generations. The solution times of the bi-objective are about the double of the time required for single objective ES due to the extra operations to maintain Pareto front. As also expected, the solution times of the problems with a budget constraint of 600 are higher than the ones with a budget constraint of 500 due to larger number of available devices.

More importantly, the time per iteration increases as the number of alternative paths increases from 1 to ten (Figures 6.45 and 6.46). Cut set size does not seem to affect the solution time per iteration significantly for the 10 user scenario because the cut-set size does not have a significant effect on the capacitated resilience calculation due to the small number of users (and devices). To support this, Figures 6.47 and 6.48 show that capacitated resilience does not change significantly for different cut-set sizes. Interestingly, the difference
in capacitated resilience between the number of paths of two and 10 is insignificant and it validates that the estimation of capacitated resilience using a smaller number of alternative paths and cut-set size performs well, i.e., finds comparable capacitated resilience values (near-exact values) faster.

At each ES generation, some number of alternative paths are evaluated. This number varies for each generation and finding alternative paths requires solving a $k$-shortest path problem, therefore it affects the solution time directly. Figures 6.49 and 6.50 compares the number of evaluated alternative paths. Clearly, these figures have the same trend with the solution time (per generation) figures (Figures 6.45 and 6.46). The number of evaluated paths increases as the parameter value of number of alternative paths increases from one to 10. As expected, cut-set size does not have any effect on the number of evaluated alternative paths.

The effect of the problem size on the total solution time is investigated in detail in Section 6.5.2.
Figure 6.45: Solution time per iteration (one ES generation) comparison of single objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
Figure 6.46: Solution time per iteration (one ES generation) comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
Figure 6.47: Capacitated resilience comparison of single objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
Figure 6.48: Capacitated resilience comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
Figure 6.49: Number of evaluated alternative paths per iteration (one ES generation) comparison of single-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
Figure 6.50: Number of evaluated alternative paths per iteration (one ES generation) comparison of bi-objective ES optimized according to different budgets, number of alternative paths and cut set sizes for capacitated resilience for the 10 user problem (10 problem instances, 10 random number seeds each)
6.5.2 Solution time comparison of capacitated resilience: Problem size

Table 6.5 summarizes the solution times of single (capacitated resilience) and bi-objective (cost and capacitated resilience) ES. The smallest problem size is the 10 user and the largest test problem is the 150 user scenario. The solution time difference between single and bi-objective ES increases as the problem size increases. For the 10 user scenario the bi-objective takes 2.89 times longer than the single objective, whereas it takes 4.45 times longer for the 150 user scenario. The reason is that the single objective terminates earlier (1000 generations with early termination of 250 non-improved generations) as the problem size increases, whereas the bi-objective does not terminate (2000 generations with early termination of 500 non-improved generations) as early as the single objective because finding an improved non-dominated solution is easier. According to these results, solving problem instances having more than 200 users is not practical.

Table 6.5: Solution time comparison of single (capacitated resilience) and bi-objective (cost and capacitated resilience) ES for different problem sizes

<table>
<thead>
<tr>
<th>Objective</th>
<th>Scenario</th>
<th>Solution Time (s)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average</td>
<td>Std. Dev.</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Single obj</td>
<td>u10_600_10_4</td>
<td>170.009</td>
<td>58.589</td>
<td>68.252</td>
<td>480.021</td>
</tr>
<tr>
<td></td>
<td>u25_1200_10_4</td>
<td>858.131</td>
<td>135.665</td>
<td>499.680</td>
<td>1,503.129</td>
</tr>
<tr>
<td></td>
<td>u50_1700_10_4</td>
<td>2,157.812</td>
<td>136.623</td>
<td>1,992.201</td>
<td>2,656.511</td>
</tr>
<tr>
<td></td>
<td>u75_2700_10_4</td>
<td>7,885.627</td>
<td>462.430</td>
<td>7,137.655</td>
<td>8,810.718</td>
</tr>
<tr>
<td></td>
<td>u100_3000_5_4</td>
<td>10,978.277</td>
<td>412.083</td>
<td>10,350.390</td>
<td>12,057.203</td>
</tr>
<tr>
<td></td>
<td>u150_4700_5_4</td>
<td>45,656.592</td>
<td>5,853.616</td>
<td>25,814.137</td>
<td>53,271.754</td>
</tr>
<tr>
<td>Bi-obj</td>
<td>u10_600_10_4</td>
<td>492.209</td>
<td>127.283</td>
<td>190.815</td>
<td>667.567</td>
</tr>
<tr>
<td></td>
<td>u25_1200_10_4</td>
<td>2,688.267</td>
<td>454.826</td>
<td>1,589.506</td>
<td>3,309.185</td>
</tr>
<tr>
<td></td>
<td>u50_1700_10_4</td>
<td>7,174.958</td>
<td>1,519.856</td>
<td>3,754.245</td>
<td>9,124.685</td>
</tr>
<tr>
<td></td>
<td>u75_2700_10_4</td>
<td>30,748.788</td>
<td>3,836.392</td>
<td>19,615.756</td>
<td>34,733.680</td>
</tr>
<tr>
<td></td>
<td>u100_3000_5_4</td>
<td>46,205.246</td>
<td>3,294.315</td>
<td>30,808.370</td>
<td>49,085.168</td>
</tr>
<tr>
<td></td>
<td>u150_4700_5_4</td>
<td>203,254.431</td>
<td>13,536.241</td>
<td>154,795.730</td>
<td>220,643.620</td>
</tr>
</tbody>
</table>

*a, b, c, d: a=number of users, b=budget, c=number of alternative paths and d=cut set size

In summary, the solution time of bi-objective is higher than single objective. Solution times of both single and bi-objective ES are sensitive to the parameter value of the number of alternative paths due to total number of evaluated paths. Cut-set size seems to have an
insignificant effect on the solution time for the 10 user scenario. An important result is that the estimation of capacitated resilience with lower number or alternative paths and cut-set sizes (e.g., 2 and 2 instead of 10 and 4, respectively) performs comparably with the exact capacitated resilience calculation and runs faster.

### 6.5.3 Solution time comparison of all metrics

In this section, the solution times for optimization of all metrics for different problem sizes are summarized. Table 6.6 presents the solution times of the 10 and the 25 user scenarios for single objective (capacitated resilience, TE, two-terminal and all-terminal reliability) and bi-objective (cost and capacitated resilience, and TE and capacitated resilience) ES.

According to these results, two-terminal and all-terminal reliability have the lowest solution times, however, capacitated resilience optimization performs very closely. On the other hand, traffic efficiency is the worst in terms of solution time. Similarly, the bi-objective ES of cost and capacitated resilience is significantly faster than the bi-objective ES of TE and capacitated resilience. Solving bi-objective ES of TE and capacitated resilience is not practical beyond problem sizes larger than 30 users, whereas bi-objective of cost and capacitated resilience can solve for the 100 user scenario within the same time of solving the 25 user scenario for TE and capacitated resilience.

The largest solvable problem sizes within 24 hours for all metrics are given in Table 6.7. All scenarios are for the medium budget for their size (number of users). According to these results, using capacitated resilience can solve larger problems than using TE and its largest solvable problem size is very close to using all-terminal or two-terminal reliabilities.
### Table 6.6: Solution time comparison of all metrics for the 10 and the 25 user scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Budget</th>
<th>Objective</th>
<th>Method</th>
<th>Average</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 users</td>
<td>500</td>
<td>Single obj</td>
<td>Capacitated Resilience</td>
<td>108.330</td>
<td>22.588</td>
<td>64.379</td>
<td>188.263</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Traffic Efficiency</td>
<td>16,151.041</td>
<td>3,360.070</td>
<td>9,058.007</td>
<td>22,496.754</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Two-terminal Reliability</td>
<td>96.009</td>
<td>20.574</td>
<td>53.801</td>
<td>145.500</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All-terminal Reliability</td>
<td>96.096</td>
<td>18.203</td>
<td>61.903</td>
<td>153.425</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>Bi-obj</td>
<td>Cost and Capacitated Res.</td>
<td>305.470</td>
<td>76.779</td>
<td>141.138</td>
<td>454.692</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TE and Capacitated Res.</td>
<td>30,680.334</td>
<td>8,955.493</td>
<td>14,522.584</td>
<td>48,832.880</td>
</tr>
<tr>
<td>25 users</td>
<td>1000</td>
<td>Single obj</td>
<td>Capacitated Resilience</td>
<td>571.159</td>
<td>70.104</td>
<td>388.844</td>
<td>826.490</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Traffic Efficiency</td>
<td>42,517.725</td>
<td>6,287.600</td>
<td>32,546.073</td>
<td>52,489.889</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Two-terminal Reliability</td>
<td>538.086</td>
<td>82.644</td>
<td>374.393</td>
<td>682.823</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All-terminal reliability</td>
<td>534.854</td>
<td>59.740</td>
<td>431.624</td>
<td>728.873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bi-obj</td>
<td>Cost and Capacitated Res.</td>
<td>1,747.886</td>
<td>345.188</td>
<td>943.294</td>
<td>2,141.489</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TE and Capacitated Res.</td>
<td>141,753.159</td>
<td>13,036.430</td>
<td>119,028.622</td>
<td>150,931.612</td>
</tr>
</tbody>
</table>
Table 6.7: The largest solvable scenarios in 24 hours, given as number of users for a medium budget

<table>
<thead>
<tr>
<th>Method</th>
<th># of users in problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacitated resilience*</td>
<td>200</td>
</tr>
<tr>
<td>Traffic efficiency</td>
<td>50</td>
</tr>
<tr>
<td>Two-terminal reliability</td>
<td>240</td>
</tr>
<tr>
<td>All-terminal reliability</td>
<td>220</td>
</tr>
<tr>
<td>Bi-objective cost vs. capacitated resilience</td>
<td>125</td>
</tr>
<tr>
<td>Bi-objective TE vs. capacitated resilience</td>
<td>20</td>
</tr>
</tbody>
</table>

* Number of alternative paths and cut set size are 10 and 4, respectively.
Chapter 7
Conclusions and Further Research

7.1 Conclusions

In network design, survivability is a challenging but important problem. In the network reliability/survivability literature, many metrics have been proposed. Among them all-terminal reliability, $k$-terminal reliability, traffic efficiency, $k$ edge-disjoint paths and $k$ node-disjoint paths are commonly used. However, a new survivability metric has been proposed in this dissertation to overcome limitations of the previous metrics and provide a new aspect to survivability.

Capacitated resilience, the proposed metric in this dissertation, considers capacity, reliability and rerouting options simultaneously. This is the main difference of capacitated resilience compared with the other metrics. If a node or an edge does not have enough capacity then a path using that node or edge is not feasible. The availability of rerouting options in case of a failure helps to maintain connectivity and session continuity. Capacitated resilience becomes zero when rerouting is not available. In a wireless network, redundant devices create alternative paths which are the rerouting options. Capacitated resilience uses reliability and scales it with rerouting options to find the true resilience of the network under capacity constraints. Therefore, it allows a comparison of different network designs whereas connectivity based metrics do not allow comparison of different designs. The connectivity based metrics, $k$ edge-disjoint or node-disjoint paths, consider neither reliability nor capacity in the network. They focus on the redundant paths in the network design assuming perfectly reliable nodes and edges. Terminal reliability ($k$-terminal or all-terminal) focuses on reliability but does not consider capacity. Traffic efficiency considers rerouting options but it does not consider capacity.
In the literature, exact methods to calculate the reliability/survivability metrics have been presented. However, due to intractability of exact methods for even moderately sized problems, many approximation methods or simulation (mostly Monte Carlo simulation) approaches have also been developed. For example, Konak and Bartolacci (2007) proposed a simulation based method to estimate TE. In this dissertation, an exact method to calculate capacitated resilience is given as well as an approximation method. The exact calculation is based on the $k$-shortest path calculation and cut-set identification. However, a fast and effective estimation of capacitated resilience is obtained by changing the values of $k$ and the size of the cut-set.

In this dissertation both single and bi-objective optimization are considered. The Evolutionary Strategies method has been used to solve the network design problem because of its success on nonlinear problems for which traditional optimization methods (e.g., mixed integer programming) do not perform satisfactorily. Also, ES works best with continuous problems and the network design problem in this dissertation has device coordinates as continuous decision variables. Besides, ES can easily be extended to multiobjective optimization, which is one of the contributions of this dissertation. Single objective optimization solves for maximum capacitated resilience, traffic efficiency, two-terminal reliability and all-terminal reliability. Bi-objective optimization simultaneously optimizes cost and capacitated resilience because a main constraint of real life projects is the budget. Pareto optimality is used for bi-objective optimization. A well known multiobjective method, NSGA-II (Deb et al., 2002), has been adapted for the proposed bi-objective ES model in this dissertation.

Different problem instances were solved for capacitated resilience as well as the other metrics and the resulting network designs were compared. The network designs obtained by optimization for capacitated resilience have structures such that the high traffic users are prioritized. Redundancies are allocated as the budget allows but the redundant devices are mostly allocated around the high traffic users to ensure maximum capacitated resilience. The other metrics prioritized reliable connections for all users (starting from the high traffic...
ones) and redundancies are considered as a secondary objective. Therefore, it can be concluded that capacitated resilience prioritizes the allocation of redundancies (survivability) considering the capacities.

From the network survivability perspective, it is interesting to consider the performance of the different designs in case of failures and attacks. Obviously, a planned attack has a more severe effect on the network than a random attack (or failure). An attack on the network may aim for removal (or elimination) of some edges (or nodes) to disconnect users (the max-flow min-cut problem). Capacitated resilience considers cut-sets in its calculation and its primary design goal is to create redundancies to maintain connectivity in case of a failure. Since capacitated resilience maximizes survivability by taking cut-sets into account, it provides a resilient design that minimizes the adverse effects of planned or unplanned attack.

In this dissertation, heterogeneous wireless networks are considered. Design of survivable heterogeneous wireless networks is a new area in optimization which has gained popularity because of the growing use of new telecommunication technologies such as 4G and wireless hotspots. Heterogeneity is defined as the differences in both offered services and device properties in a wireless network. However, capacitated resilience and its design rules can be applied to any network including homogeneous ones (such as military networks, transportation networks, communication networks or electrical networks) to ensure resilience and survivability.

7.2 Further research

For future research, node failures can be added to capacitated resilience calculation. In the current model presented in this dissertation the nodes are assumed to be perfectly reliable. Including node failures requires the independent subgroup identification process (Section 3.2.3) to be redesigned because common nodes among different independent subgroups create dependencies. To solve this, two independent subgroups having one (or more)
common node should be merged. This may increase computational effort by reducing the
number but increasing the size of independent subgroups.

As an interesting extension, the effect of interference can be included in the calculation
of capacitated resilience and other metrics. Currently, wireless interference is not considered
as a parameter. Different wireless channels are assumed to be assigned for each wireless
link within the same communication range to eliminate interference. Although interference
has been mostly considered negligible in the network reliability/survivability literature, as
discussed in Section 2.5, it may affect the reliabilities of wireless links in dense networks.
Addition of the wireless channel assignment as a decision variable (this also requires the
routing algorithm to be changed accordingly) and including interference in the wireless link
reliability calculation would make the model presented in this dissertation more realistic.

As discussed in Section 4.1.4, the current assignment of users to devices is done in a
randomized order. Although this is not an issue for most cases, for a few instances it can
lead to suboptimal user-device assignments if the capacity of the network is restricted. The
obvious, but not ideal, solution for this user-device assignment issue would be considering
all possible combinations of user to device assignments (enumeration) and choosing the best
one. However, this approach would make the proposed model intractable for moderate to
large size problem instances. An alternative heuristic method could be devised as future
work.

As another change in the proposed method, the use of Monte Carlo simulation will
be investigated for calculation of the reliability of independent subgroups. In the proposed
model, the reliability is calculated by identification of minimal cut-sets which provides a
lower bound of reliability (worst case). Simulation may provide the expected value of the
reliability of the subgroups, however, it is computationally more expensive.

Network failures and attacks are two important topics in the survivable network design
literature. As future work, the network designs obtained by optimization for capacitated
resilience and the other metrics could be compared in terms of network attacks (that is,
planned interdictions) or different failure scenarios. The proposed metric, capacitated resilience, could be applied to other networks, such as transportation networks or military supply networks.

In this dissertation, the locations of users are fixed. However, in real life the users are mobile for wireless network applications. Therefore, to be more realistic, mobility of the users could be included in the proposed model.

One of the most popular topics in today’s computational optimization area is to use parallel computing. It gained popularity as devices with multiple core processors became widespread after 2000. The solution approach presented in this dissertation and calculation methods of the other metrics are computationally expensive. Effective use of parallel computing could reduce solution times dramatically. Parallel computing will be even more important in the future as the parallel use of the GPUs (Graphics Processing Unit) and the CPUs (Central Processing Unit) are becoming popular. Therefore, adaptation of the solution method of this dissertation to parallel computing would be an interesting extension.
Appendix A

Input data of problems

All problem data (10, 25, 50, 75, 100, 150 user scenarios) are given. Each scenario consists of 10 problem instances. Coordinates and traffic requirement of each user are given as an array of size three, \([x, y, \text{traffic requirement}]\).

A.1 The 10 user scenario

**Problem instance 1:** \([0.88, 2.696, 0.013], [3.467, 0.078, 5.11], [-3.432, -2.914, 19.153], [-1.689, -2.242, 14.729], [3.571, -3.636, 14.45], [1.098, 3.948, 15.753], [1.333, -3.518, 19.662], [-3.518, -0.173, 14.269], [-3.227, -0.516, 3.737], [-3.227, -0.516, 3.737], [-1.684, 2.159, 14.803]\)

**Problem instance 2:** \([0.425, 2.529, 6.56], [1.242, -3.97, 9.247], [2.913, 1.513, 17.002], [-2.775, -1.519, 12.444], [0.649, 0.403, 4.59], [3.559, -3.492, 0.998], [-1.307, 2.951, 5.34], [-2.467, 3.871, 9.607], [1.285, 0.367, 17.313], [-1.424, -0.027, 3.466]\)

**Problem instance 3:** \([-3.447, 2.797, 0.317], [-2.521, 2.604, 8.453], [1.481, -3.948, 9.813], [-0.72, 0.644, 13.425], [0.864, 3.667, 3.374], [0.85, 2.552, 3.359], [3.91, 0.577, 14.871], [-3.499, 2.782, 18.653], [-3.675, 0.586, 2.583], [-0.111, -2.855, 18.295]\)

**Problem instance 4:** \([2.065, -3.872, 4.379], [-1.016, -0.351, 0.317], [1.165, 3.918, 5.042], [0.292, 0.695, 13.085], [3.838, -3.714, 1.632], [2.008, -0.472, 3.663], [0.525, -2.181, 0.488], [2.914, 2.585, 12.139], [1.778, -2.903, 18.476], [0.906, -2.608, 13.341]\)
Problem instance 5: \([-2.095, -1.775, 5.265], [2.815, -0.196, 16.87], [1.497, -0.779, 15.504], [-1.886, -1.92, 2.493], [2.775, -2.977, 9.111], [-1.996, 3.514, 7.032], [1.765, 2.797, 0.603], [3.243, -1.44, 11.039], [-2.868, -1.304, 2.82], [2.098, 1.215, 5.066]\]

Problem instance 6: \([-2.764, -3.462, 1.966], [1.331, 2.051, 4.234], [2.643, 3.312, 12.872], [-0.611, -2.715, 19.712], [3.23, 1.423, 14.972], [-3.822, -1.639, 5.411], [-3.382, 3.901, 16.055], [0.416, 1.244, 14.181], [-1.236, -3.566, 9.328], [0.882, 1.563, 2.772]\]

Problem instance 7: \([-0.941, -3.71, 4.635], [1.681, 0.351, 18.569], [3.508, 2.687, 19.517], [3.035, 0.02, 19.447], [-2.124, -1.193, 2.273], [3.853, -2.038, 8.534], [2.469, 2.661, 3.19], [-3.177, -3.079, 15.768], [-3.33, -2.657, 18.658], [-0.477, -2.643, 8.68]\]

Problem instance 8: \([-3.041, 2.415, 12.086], [-1.889, 3.811, 11.329], [1.091, 2.325, 18.516], [2.073, -0.832, 4.929], [1.031, -0.61, 14.808], [0.086, 1.011, 0.342], [-0.998, 0.798, 10.204], [1.432, -2.797, 1.415], [0.207, -3.128, 12.247], [-0.151, -0.136, 17.919]\]

Problem instance 9: \([-1.034, -3.91, 0.69], [-1.066, 1.384, 19.162], [-2.36, -3.127, 4.057], [0.787, -1.283, 5.559], [2.27, -3.082, 0.4], [-2.979, -1.388, 7.635], [-1.676, -2.389, 0.179], [-0.238, 1.017, 9.586], [-2.458, -3.617, 19.424], [-2.329, 3.876, 4.893]\]

Problem instance 10: \([0.931, 2.503, 2.432], [-0.405, -0.237, 8.911], [2.67, 2.788, 4.89], [1.548, -1.768, 19.765], [0.077, 0.535, 11.362], [1.731, 0.163, 13.669], [-3.519, 2.104, 16.373], [1.195, 2.416, 11.378], [-1.397, -3.387, 8.271], [3.657, 1.635, 19.93]\]
A.2 The 25 user scenario

Problem instance 1: [1.391, 4.263, 0.013], [5.481, 0.123, 5.11], [-5.426, -4.608, 19.153], [-2.671, -3.544, 14.729], [5.647, -5.749, 14.45], [1.737, 6.242, 15.753], [2.107, -5.562, 19.662], [-5.562, -0.273, 14.269], [-5.102, -0.816, 3.737], [-2.663, 3.414, 14.803], [-1.794, 0.083, 8.585], [-1.765, -6.3, 12.228], [-6.562, 0.425, 1.925], [0.107, -3.39, 10.052], [5.823, -2.713, 12.703], [0.624, -3.538, 1.513], [4.668, 3.128, 4.606], [3.592, -1.387, 14.304], [-5.913, 0.436, 10.943], [-6.245, -3.151, 10.304], [6.009, -5.457, 5.83], [1.177, 2.556, 0.042], [2.594, 4.961, 0.755], [-6.183, 5.84, 7.282], [0.603, 3.298, 6.955]

Problem instance 2: [-5.566, 5.313, 16.264], [-3.916, -3.112, 1.571], [-2.464, 2.949, 16.276], [4.19, -1.131, 16.322], [0.502, 4.679, 11.183], [3.039, -1.857, 10.679], [-0.306, 3.989, 9.016], [0.757, 6.284, 13.796], [-4.933, -1.147, 5.947], [5.54, 0.711, 12.313], [-2.606, 1.93, 9.247], [-4.703, 1.202, 10.161], [2.93, -4.495, 2.67], [4.17, 2.936, 0.845], [2.111, 5.758, 18.317], [4.327, 0.98, 10.013], [-0.483, 5.198, 16.846], [-3.782, 3.959, 8.267], [5.978, 0.944, 0.845], [-5.419, 3.253, 11.699], [4.47, -0.273, 2.33], [-4.352, -2.492, 19.088], [-5.81, 4.619, 9.598], [2.951, 4.91, 9.552], [4.872, 4.065, 8.333]

Problem instance 3: [-5.45, 4.422, 0.317], [-3.986, 4.118, 8.453], [2.342, -6.243, 9.813], [-1.139, 1.018, 13.425], [1.365, 5.798, 3.374], [1.344, 4.035, 3.359], [6.183, 0.912, 14.871], [-5.533, 4.399, 18.653], [-5.811, 0.926, 2.583], [-0.175, -4.515, 18.295], [-0.375, -1.563, 6.404], [2.462, 1.265, 19.06], [5.013, 3.641, 7.024], [0.755, -3.928, 18.024], [1.791, 3.54, 5.174], [0.11, -2.224, 17.889], [-3.362, -2.509, 17.724], [-0.892, -1.127, 10.177], [-3.769, 5.502, 0.987], [1.736, 0.437, 15.58], [-2.44, -0.885, 17.143], [1.644, -2.805, 9.306], [4.615, -3.753, 18.278], [2.696, 5.115, 5.954], [-2.939, -0.918, 14.029]

Problem instance 4: [3.265, -6.121, 4.379], [-1.606, -0.556, 0.317], [1.841, 6.195, 5.042], [0.461, 1.099, 13.085], [6.068, -5.872, 1.632], [3.174, -0.746, 3.663], [0.83, -3.448, 0.488], [4.608,
Problem instance 5: [-3.312, -2.807, 5.265], [4.451, -0.309, 16.87], [2.367, -1.232, 15.504], [-2.982, -3.035, 2.493], [4.387, -4.707, 9.111], [-3.157, 5.556, 7.032], [2.791, 4.422, 0.603], [5.128, -2.277, 11.039], [-4.535, -2.062, 2.82], [3.317, 1.922, 5.066], [-1.397, 3.963, 13.456], [-2.088, -3.56, 0.458], [-4.108, -6.32, 9.832], [-3.632, 5.126, 12.081], [-0.112, -4.312, 13.064], [4.601, -1.721, 17.17], [-3.599, 4.186, 6.903], [2.646, 5.441, 2.404], [4.423, 6.196, 9.041], [4.859, 2.084, 6.417], [0.037, -0.233, 18.423], [-4.411, 2.11, 15.584], [-0.403, 2.886, 10.464], [3.847, 1.455, 5.727], [4.88, -3.506, 7.859]

Problem instance 6: [-4.371, -5.474, 1.966], [2.105, 3.242, 4.234], [4.179, 5.236, 12.872], [-0.967, -4.293, 19.712], [5.107, 2.25, 14.972], [-6.043, -2.592, 5.411], [-5.348, 6.168, 16.055], [0.659, 1.967, 14.181], [-1.955, -5.638, 9.328], [1.394, 2.471, 2.772], [5.51, 4.801, 10.97], [-6.127, 5.125, 3.239], [4.912, 5.017, 19.512], [-2.181, -0.594, 9.299], [2.377, 5.962, 3.182], [-3.123, -2.471, 13.902], [5.171, 0.651, 10.818], [5.858, -2.852, 5.917], [0.691, 6.203, 6.802], [-1.6, 5.693, 4.677], [3.335, -6.136, 16.245], [-6.044, 5.252, 9.539], [-1.857, -6.269, 12.376], [-0.159, -1.381, 6.3], [2.656, -1.048, 3.789]

Problem instance 7: [-1.488, -5.866, 4.635], [2.657, 0.555, 18.569], [5.547, 4.248, 19.517], [4.799, 0.032, 19.447], [-3.358, -1.886, 2.273], [6.092, -3.222, 8.534], [3.904, 4.208, 3.19], [-5.023, -4.869, 15.768], [-5.265, -4.2, 18.658], [-0.755, -4.179, 8.68], [4.321, -4.504, 5.384], [-0.63, -3.094, 14.327], [-4.734, 5.981, 15.098], [1.274, -4.947, 19.668], [-5.521, 5.805, 17.34], [5.828, 0.364, 0.627], [-0.755, 0.014, 9.008], [0.303, -4.944, 17.036], [-5.534, 3.857,
Problem instance 8: [-4.809, 3.818, 12.086], [-2.987, 6.026, 11.329], [1.725, 3.676, 18.516], [3.277, -1.315, 4.929], [1.631, -0.965, 14.808], [0.136, 1.598, 0.342], [-1.578, 1.261, 10.204], [2.265, -4.423, 1.415], [0.327, -4.946, 12.247], [-0.239, -0.214, 17.919], [-4.47, -1.628, 17.803], [-1.982, -3.592, 9.652], [-2.152, 6.195, 3.955], [-0.285, 4.425, 0.322], [1.308, 4.117, 2.191], [1.24, 2.583, 18.785], [-4.832, 5.997, 2.62], [5.004, -2.876, 4.381], [-1.251, 1.814, 14.071], [-0.995, -2.462, 13.04], [2.772, 2.814, 7.315], [1.45, -1.239, 9.32], [2.168, 4.116, 19.273], [4.811, 5.321, 19.853], [-4.593, 5.806, 0.627]

Problem instance 9: [-1.635, -6.182, 0.69], [-1.686, 2.188, 19.162], [-3.731, -4.945, 4.057], [1.244, -2.029, 5.559], [3.59, -4.873, 0.4], [-4.711, -2.195, 7.635], [-2.651, -3.778, 0.179], [-0.377, 1.607, 9.586], [-3.887, -5.719, 19.424], [-3.682, 6.128, 4.893], [-2.285, -2.343, 5.256], [1.849, -1.127, 8.19], [1.761, 1.058, 4.807], [2.575, 2.063, 4.96], [-5.249, -5.235, 14.013], [6.193, 2.75, 13.133], [-5.673, 5.592, 1.832], [5.16, 1.674, 15.116], [-6.005, 3.459, 0.088], [-0.004, -3.038, 1.035], [-1.479, -0.532, 1.893], [-4.764, 3.13, 19.831], [-0.283, -2.142, 15.42], [0.246, -2.758, 18.417], [2.146, -3.154, 15.359]

Problem instance 10: [1.472, 3.958, 2.432], [-0.641, -0.375, 8.911], [4.222, 4.409, 4.89], [2.448, -2.796, 19.765], [0.122, 0.846, 11.362], [2.736, 0.258, 13.669], [-5.563, 3.326, 16.373], [1.889, 3.821, 11.378], [-2.208, -5.355, 8.271], [5.783, 2.585, 19.93], [-3.586, 5.174, 0.389], [-0.618, -2.202, 15.35], [0.968, -4.059, 7.338], [-0.584, 5.487, 7.232], [-1.206, 3.216, 18.935], [-5.609, -6.124, 16.754], [-2.32, 1.538, 13.13], [-2.393, -2.823, 4.36], [3.066, -0.899, 16.088], [-5.305, -4.712, 16.385], [-5.085, -2.181, 8.741], [1.482, -0.739, 5.719], [-3.308, 4.405, 17.658], [0.803, 0.9, 6.896], [5.483, -0.384, 6.717]
A.3 The 50 user scenario

**Problem instance 1:** [1.967, 6.028, 0.013], [7.752, 0.173, 5.11], [-7.673, -6.517, 19.153], [-3.777, -5.012, 14.729], [7.986, -8.131, 14.45], [2.456, 8.827, 15.753], [2.98, -7.866, 19.662], [-7.866, -0.386, 14.269], [-7.215, -1.155, 3.737], [-3.766, 4.828, 14.803], [-2.536, 0.117, 8.585], [-2.496, -8.909, 12.228], [-6.58, 0.601, 1.925], [0.151, -4.794, 10.052], [8.235, -3.837, 12.703], [0.882, -5.004, 1.513], [6.601, 4.424, 4.606], [5.08, -1.961, 14.304], [-8.362, 0.616, 10.943], [-8.832, -4.456, 10.304], [8.498, -7.717, 5.83], [1.665, 3.615, 0.042], [3.669, 7.016, 0.755], [-8.744, 8.259, 7.282], [0.853, 4.664, 6.955], [-5.683, 7.133, 17.293], [5.72, 2.821, 9.802], [-8.359, -3.888, 7.194], [6.145, -8.099, 18.491], [0.071, -8.477, 4.166], [-6.133, -8.925, 9.775], [-8.537, 6.21, 14.335], [-0.172, -2.864, 13.144], [-2.63, 0.086, 17.368], [-7.151, 1.85, 1.009], [0.984, -6.537, 19.711], [-4.739, 7.82, 7.03], [5.741, -6.644, 19.099], [3.277, -4.065, 8.193], [8.564, 7.878, 0.284], [3.931, -0.811, 10.279], [-0.602, -4.814, 19.315], [-3.122, -7.569, 13.917], [-5.749, -2.906, 12.623], [5.612, -1.037, 18.098], [-3.058, -7.471, 16.108], [4.954, 2.657, 18.198], [-2.154, 7.527, 12.932], [-3.58, -4.181, 1.301], [-2.822, 6.504, 12.931]

**Problem instance 2:** [-7.872, 7.514, 16.264], [-5.538, -4.401, 1.571], [-3.485, 4.17, 16.276], [5.925, -1.599, 16.322], [0.71, 6.618, 11.183], [4.297, -2.626, 10.679], [-0.433, 5.641, 9.016], [1.071, 8.887, 13.796], [-6.976, -1.622, 5.947], [7.835, 1.006, 12.313], [-3.685, 2.729, 9.247], [-6.651, 1.7, 10.161], [4.144, -6.356, 2.67], [5.898, 4.152, 0.845], [2.985, 8.143, 18.317], [6.119, 1.386, 10.013], [-0.683, 7.351, 16.846], [-5.349, 5.599, 8.267], [8.454, 1.336, 0.845], [-7.664, 4.6, 11.699], [6.321, -0.386, 2.33], [-6.154, -3.524, 19.088], [-8.216, 6.532, 9.598], [4.173, 6.943, 9.552], [6.89, 5.748, 8.333], [4.707, 1.239, 19.653], [-5.624, 1.501, 9.7], [-1.807, 8.596, 0.65], [-0.064, -0.291, 16.267], [-2.304, 2.736, 1.652], [-0.458, 7.4, 6.83], [6.017, 6.694, 12.227], [3.208, 3.439, 11.02], [4.516, -3.939, 1.113], [0.853, 1.728, 9.313], [-2.23, -6.862, 11.642], [-2.95, 2.502, 0.817], [4.774, -0.301, 0.667], [4.099, -8.193, 7.482], [7.399, 7.657, 2.596], [-5.54, 3.527, 8.587], [-5.854, 1.03, 5.198], [-5.356, 7.693, 17.39], [-1.275, -7.606, 7.763], [8.217, -2.825, 13.669], [8.022, -7.272, 18.338], [2.967, 7.497, 7.284], [3.305, -4.26, 18.8], [-7.35, 3.19, 11.776], 191
Problem instance 3: [-7.708, 6.254, 0.317], [-5.637, 5.823, 8.453], [3.313, -8.828, 9.813], [-1.611, 1.439, 13.425], [1.931, 8.199, 3.374], [1.9, 5.706, 3.359], [8.744, 1.29, 14.871], [-7.825, 6.221, 18.653], [-8.218, 1.31, 2.583], [-0.247, -6.385, 18.295], [3.481, 1.789, 19.06], [7.089, 5.149, 7.024], [1.068, -5.555, 18.024], [2.533, 5.006, 5.174], [0.156, -3.145, 17.889], [-4.754, -3.549, 17.724], [-1.262, -1.594, 10.177], [-5.311, 7.781, 0.987], [2.455, 0.618, 15.58], [-3.45, -1.252, 17.143], [2.325, -3.967, 9.306], [6.527, -5.307, 18.278], [3.813, 7.233, 5.954], [-4.156, -1.298, 14.029], [-4.834, 5.622, 0.981], [-0.569, -6.238, 3.029], [5.733, -8.013, 14.799], [5.123, -7.712, 1.984], [1.087, -5.509, 2.361], [-7.378, 6.138, 13.116], [-5.294, -3.467, 2.613], [-0.41, 4.894, 6.345], [-2.528, -4.958, 7.044], [5.87, -8.054, 8.951], [-6.493, -7.845, 18.027], [-1.895, -8.677, 11.516], [6.33, 0.578, 18.171], [5.075, -1.556, 7.65], [3.887, -7.349, 7.521], [3.671, 3.462, 19.24], [8.672, 8.32, 8.083], [-1.701, 5.357, 15.402], [-7.216, -0.743, 11.301], [4.885, 8.199, 4.504], [-1.903, 3.86, 18.638], [-4.729, -2.005, 11.642], [6.873, 3.528, 17.364], [-6.35, 2.27, 1.108], [-7.854, -5.837, 13.617]

Problem instance 4: [4.618, -8.657, 4.379], [-2.271, -0.786, 0.317], [2.604, 8.761, 5.042], [0.653, 1.555, 13.085], [8.582, -8.305, 1.632], [4.489, -1.055, 3.663], [1.174, -4.876, 0.488], [6.517, 5.78, 12.139], [3.977, -6.492, 18.476], [2.027, -5.832, 13.341], [1.97, 4.661, 16.015], [-0.02, -2.383, 4.304], [-3.234, 7.808, 0.392], [-4.084, 7.894, 0.617], [-6.579, 5.553, 18.274], [4.359, -4.853, 17.171], [7.34, 8.876, 7.79], [-8.514, 6.877, 11.239], [-5.48, 2.625, 18.628], [-3.993, -7.816, 8.739], [-7.047, 4.187, 3.949], [4.555, 5.054, 8.156], [-4.336, 2.287, 4.196], [-1.476, -3.535, 4.303], [-0.75, 2.194, 2.976], [3.053, -8.532, 14.047], [2.013, 8.305, 7.218], [5.133, 1.917, 14.707], [2.096, -3.251, 11.635], [0.841, -7.685, 12.252], [4.627, -0.525, 16.254], [6.964, -1.883, 15.11], [-8.234, -1.653, 2.702], [-6.237, -1.709, 2.029], [-1.878, 6.714, 19.324], [-0.986, -3.293, 11.953], [7.185, -5.307, 7.675], [-2.363, -1.602, 12.774], [2.46, -6.027, 7.67], [0.885, -2.879, 0.047], [-8.478, 4.651, 10.272], [-7.095, -6.568, 7.34], [-3.179, 8.696, 16.591], [-1.958, 0.993, 192]
Problem instance 5: [-4.684, -3.969, 5.265], [6.295, -0.437, 16.87], [3.348, -1.743, 15.504], [-4.217, -4.292, 2.493], [6.205, -6.656, 9.111], [-4.464, 7.857, 7.032], [3.947, 6.253, 0.603], [7.252, -3.221, 11.039], [-6.413, -2.917, 2.82], [4.691, 2.718, 5.066], [-1.975, 5.605, 13.456], [-2.953, -5.034, 0.458], [-5.81, -8.937, 9.832], [-5.137, 7.249, 12.081], [-0.158, -6.099, 13.064], [6.507, -2.434, 17.17], [-5.089, 5.919, 6.903], [3.742, 7.695, 2.404], [6.256, 8.762, 9.041], [6.872, 2.947, 6.417], [0.052, -0.33, 18.423], [-6.238, 2.984, 15.584], [-0.57, 4.081, 10.464], [5.44, 2.058, 5.727], [6.901, -4.958, 7.859], [-1.092, -4.252, 11.065], [-5.405, 7.442, 12.474], [0.478, -1.797, 2.861], [-2.722, 8.343, 2.661], [4.818, 6.089, 11.968], [1.895, -4.747, 8.615], [-6.239, 7.166, 10.234], [-1.863, -6.647, 12.609], [2.76, 8.864, 4.975], [-2.446, -6.652, 5.043], [0.851, -0.323, 1.589], [-2.698, -8.921, 7.352], [4.437, 0.886, 1.715], [-1.365, 1.375, 7.809], [-5.697, 1.62, 10.048], [6.568, -4.313, 5.129], [6.562, 8.646, 5.473], [-0.543, 4.53, 2.858], [3.465, -8.344, 19.529], [-5.885, 3.2, 2.279], [-2.542, -7.456, 6.143], [8.518, 1.808, 11.395], [7.08, 7.618, 0.966], [-6.254, 6.745, 1.138], [-7.571, 4.247, 5.406]

Problem instance 6: [-6.181, -7.741, 1.966], [2.977, 4.585, 4.234], [5.91, 7.405, 12.872], [-1.367, -6.071, 19.712], [7.222, 3.181, 14.972], [-8.546, -3.665, 5.411], [-7.563, 8.722, 16.055], [0.931, 2.781, 14.181], [-2.764, -7.973, 9.328], [1.972, 3.495, 2.772], [7.792, 6.789, 10.97], [-8.665, 7.248, 3.239], [6.946, 7.095, 19.512], [-3.084, -0.84, 9.299], [3.362, 8.431, 3.182], [-4.417, -3.495, 13.902], [7.313, 0.92, 10.818], [8.284, -4.033, 5.917], [0.978, 8.772, 6.802], [-2.263, 8.051, 4.677], [4.716, -8.678, 16.245], [-8.547, 7.428, 9.539], [-2.627, -8.866, 12.376], [-0.225, -1.953, 6.3], [3.756, -1.482, 3.789], [-6.064, -5.289, 17.436], [-5.991, 4.627, 10.374], [-4.721, 1.921, 0.395], [7.207, -6.194, 2.688], [6.116, -2.253, 0.645], [-2.659, 5.401, 16.618], [0.295, 4.274, 10.454], [8.064, 4.843, 3.913], [-4.847, -0.224, 19.209], [7.963, 2.433, 3.14], [6.766, 4.78, 5.8], [-0.813, 7.542, 7.531], [-8.842, -8.467, 5.591], [-0.487, 6.725, 16.624], [1.122, -4.89, 18.796],
Problem instance 7: [-2.105, -8.295, 4.635], [3.758, 0.785, 18.569], [7.844, 6.008, 19.517], [6.787, 0.046, 19.447], [-4.749, -2.667, 2.273], [8.615, -4.557, 8.534], [5.521, 5.951, 3.19], [-7.103, -6.885, 15.768], [-7.445, -5.94, 18.658], [-1.068, -5.91, 8.68], [6.111, -6.369, 5.384], [-0.89, -4.376, 14.327], [-6.94, 8.458, 15.098], [1.802, -6.997, 19.668], [-7.808, 8.209, 17.34], [8.243, 0.515, 0.627], [-1.068, 0.02, 9.008], [0.428, -6.992, 17.036], [-7.826, 5.455, 14.227], [-7.557, 3.399, 4.151], [5.275, 6.138, 11.503], [1.892, -0.932, 12.834], [-1.615, -0.684, 5.835], [-3.093, -4.466, 6.496], [-8.134, -3.799, 15.109], [-1.441, 3.134, 11.053], [-7.829, -2.281, 11.29], [-7.285, 2.943, 2.679], [1.098, 2.846, 12.451], [-7.163, 0.238, 2.564], [-3.987, -6.721, 7.561], [-6.817, -6.935, 1.66], [-7.921, 2.093, 19.394], [-7.177, -2.843, 9.946], [8.025, 5.58, 9.297], [1.279, 1.027, 7.677], [-5.099, -7.356, 5.582], [-7.855, 3.011, 11.832], [6.788, 3.533, 12.035], [-3.865, 1.267, 13.452], [-7.49, 8.276, 9.623], [8.729, 7.164, 2.969], [-5.881, -1.919, 5.054], [8.094, -1.849, 1.462], [-2.728, -2.42, 13.381], [6.911, 5.327, 3.794], [3.413, -4.376, 0.614], [-0.167, -5.621, 10.23], [4.429, 7.956, 7.98], [8.064, 5.848, 11.309]

Problem instance 8: [-6.801, 5.4, 12.086], [-4.225, 8.522, 11.329], [2.439, 5.199, 18.516], [4.634, -1.86, 4.929], [2.306, -1.365, 14.808], [0.193, 2.26, 0.342], [-2.232, 1.784, 10.204], [3.203, -6.255, 1.415], [0.463, -6.995, 12.247], [-0.339, -0.303, 17.919], [-6.321, -2.303, 17.803], [-2.804, -5.08, 9.652], [-3.044, 8.761, 3.955], [-0.403, 6.258, 0.322], [1.85, 5.823, 2.191], [1.754, 3.652, 18.785], [-6.834, 8.481, 2.62], [7.077, -4.067, 4.381], [-1.77, 2.565, 14.071], [-1.407, -3.481, 13.04], [3.92, 3.979, 7.315], [2.051, -1.752, 9.32], [3.066, 5.821, 19.273], [6.803, 7.525, 19.853], [-6.495, 8.211, 0.627], [5.002, 8.653, 11.567], [2.7, 4.284, 8.756], [-6.127, -8.445, 9.424], [1.52, 3.762, 16.464], [2.542, -3.684, 5.685], [8.539, -0.981, 15.444], [1.498, -6.895, 17.785], [2.496, 4.296, 3.866], [5.432, 6.258, 5.23], [-1.226, -5.228, 0.874], [4.316, 1.29, 2.069], [2.893, -5.882, 194]
Problem instance 9: [-2.312, -8.743, 0.69], [-2.385, 3.094, 19.162], [-5.276, -6.993, 4.057], [1.76, -2.869, 5.559], [5.077, -6.891, 0.4], [-6.662, -3.104, 7.635], [-3.749, -5.343, 0.179], [-0.532, 2.273, 9.586], [-5.497, -8.087, 19.424], [-5.208, 8.667, 4.893], [-3.232, -3.314, 5.256], [2.615, -1.594, 8.19], [2.49, 1.496, 4.807], [3.642, 2.917, 4.96], [-7.423, -7.403, 14.013], [8.759, 3.89, 13.133], [-8.023, 7.908, 1.832], [7.297, 2.367, 15.116], [-8.492, 4.891, 0.088], [-0.005, -4.296, 1.035], [-2.092, -0.753, 1.893], [-6.737, 4.427, 19.831], [-0.4, -3.03, 15.42], [0.348, -3.901, 18.417], [3.034, -4.46, 15.359], [-2.934, -6.657, 18.732], [-2.953, -5.218, 16.716], [1.328, -3.79, 11.914], [-3.86, -4.733, 3.197], [-7.333, 8.117, 3.384], [-6.672, -7.051, 0.219], [-7.009, -1.396, 11.47], [-0.323, 7.877, 11.754], [3.497, 1.91, 10.66], [7.16, -4.961, 9.828], [-7.134, -4.168, 14.069], [-2.892, 6.713, 5.121], [3.746, 2.073, 13.972], [-1.605, -2.36, 17.73], [2.515, -6.354, 1.335], [-7.202, -5.957, 1.786], [8.046, 7.447, 16.882], [-8.384, 2.923, 2.787], [-4.225, 6.208, 11.437], [4.305, -1.345, 15.754], [-5.503, 3.864, 13.636], [4.059, -7.917, 3.785], [7.937, 2.549, 2.262], [5.435, -3.151, 5.49], [8.717, 3.839, 15.51]

Problem instance 10: [2.081, 5.597, 2.432], [-0.906, -0.53, 8.911], [5.971, 6.235, 4.89], [3.462, -3.954, 19.765], [0.173, 1.197, 11.362], [3.87, 0.365, 13.669], [-7.868, 4.704, 16.373], [2.671, 5.403, 11.378], [-3.123, -7.573, 8.271], [8.178, 3.656, 19.93], [-5.071, 7.318, 0.389], [-0.874, -3.114, 15.35], [1.369, -5.741, 7.338], [-0.826, 7.76, 7.232], [-1.705, 4.549, 18.935], [-7.933, -8.66, 16.754], [-3.281, 2.176, 13.13], [-3.384, -3.992, 4.36], [4.337, -1.271, 16.088], [-7.502, -6.664, 16.385], [-7.192, -3.084, 8.741], [2.096, -1.045, 5.719], [-4.678, 6.23, 17.658], [1.136, 1.272, 6.896], [7.755, -0.543, 6.717], [4.507, -1.399, 0.201], [1.51, 3.05, 6.454], [-1.575, -8.495, 3.289], [-4.321, -0.907, 1.998], [8.754, -5.077, 14.343], [7.608, 2.805, 7.084], [6.922,
A.4 The 75 user scenario

Problem instance 1: [2.409, 7.383, 0.013], [9.494, 0.212, 5.11], [-9.398, -7.981, 19.153], [-4.626, -6.139, 14.729], [9.781, -9.958, 14.45], [3.008, 10.811, 15.753], [3.65, -9.634, 19.662], [-9.634, -0.472, 14.269], [-8.836, -1.414, 3.737], [-4.613, 5.913, 14.803], [-3.107, 0.144, 8.585], [-3.057, -10.912, 12.228], [-8.059, 0.736, 1.925], [0.185, -5.872, 10.052], [10.086, -4.699, 12.703], [1.08, -6.129, 1.513], [8.085, 5.419, 4.606], [6.222, -2.402, 14.304], [-10.242, 0.755, 10.943], [-10.817, -5.457, 10.304], [10.408, -9.451, 5.83], [2.039, 4.427, 0.042], [4.493, 8.593, 0.755], [-10.709, 10.115, 7.282], [1.045, 5.712, 6.955], [-6.96, 8.736, 17.293], [7.006, 3.455, 9.802], [-10.238, -4.762, 7.194], [7.526, -9.919, 18.491], [0.088, -10.382, 4.166], [-7.511, -10.93, 9.775], [-10.455, 7.605, 14.335], [-0.211, -3.508, 13.144], [-3.221, 0.105, 17.368], [-8.758, 2.266, 1.009], [1.205, -8.006, 19.711], [-5.804, 9.578, 7.03], [7.032, -8.137, 19.099], [4.014, -4.979, 8.193], [10.489, 9.648, 0.284], [4.814, -0.993, 10.279], [-0.737, -5.896, 19.315], [-3.823, -9.269, 13.917], [-7.042, -3.559, 12.623], [6.874, -1.27, 18.098], [-3.745, -9.151, 16.108], [6.068, 3.254, 18.198], [-2.639, 9.218, 12.932], [-4.385, -5.121, 1.301], [-3.456, 7.966, 12.931], [1.629, -6.449, 1.617], [5.583, 3.279, 1.716], [-7.267, -1.026, 18.109], [-4.92, 7.412, 2.69], [3.984, -1.754, 10.246], [7.302, -0.917, 4.575], [2.738, 1.414, 5.374], [-6.525, -9.809, 2.798], [7.858, 9.542, 5.496], [-2.264, -3.468, 13.783], [-9.35, -8.791, 18.704], [2.329, 3.016, 6.007], [-5.257, 10.59, 7.005], [-10.162, -3.831, 1.601], [4.63, -0.781, 7.818], [-7, 2.155, 11.707], [-3.574, -3.025, 9.733], [-4.413,
Problem instance 2: [-9.641, 9.202, 16.264], [-6.783, -5.39, 1.571], [-4.269, 5.107, 16.276], [7.257, -1.958, 16.322], [0.87, 8.105, 11.183], [5.263, -3.216, 10.679], [-0.531, 6.909, 9.016], [1.311, 10.884, 13.796], [-8.544, -1.987, 5.947], [9.596, 1.232, 12.313], [-4.513, 3.343, 9.247], [-8.146, 2.082, 10.161], [5.076, -7.785, 2.67], [7.223, 5.085, 0.845], [3.656, 9.973, 18.317], [7.495, 1.697, 10.013], [-0.836, 9.003, 16.846], [-6.551, 6.857, 8.267], [10.354, 1.636, 0.845], [-9.386, 5.634, 11.699], [7.742, -0.472, 2.33], [-7.537, -4.316, 19.088], [-10.063, 8, 9.598], [5.111, 8.504, 9.552], [8.438, 7.04, 8.333], [5.765, 1.517, 19.653], [-6.888, 1.838, 9.7], [-2.213, 10.528, 0.65], [-0.079, -0.356, 16.267], [-2.822, 3.351, 1.652], [-0.561, 9.063, 6.83], [7.369, 8.198, 12.227], [3.929, 4.212, 11.02], [5.531, -4.825, 1.113], [1.045, 2.116, 9.313], [-2.731, -8.404, 11.642], [-3.613, 3.064, 0.817], [5.846, -0.369, 0.667], [5.02, -10.035, 7.482], [9.062, 9.378, 2.596], [-6.785, 4.32, 8.587], [-7.169, 1.262, 5.198], [-6.56, 9.422, 17.39], [-1.562, -9.315, 7.763], [10.064, -3.459, 13.669], [9.825, -8.906, 18.338], [3.634, 9.182, 7.284], [4.048, -5.218, 18.8], [-9.002, 3.906, 11.776], [-3.352, -2.775, 3.255], [-1.726, -2.732, 9.753], [-5.987, -8.049, 9.477], [7.21, -8.491, 17.368], [7.234, 3.782, 16.078], [0.87, -7.935, 4.173], [-9.613, -6.931, 15.493], [-2.815, -1.452, 7.789], [0.769, 9.295, 9.345], [-4.21, -1.665, 16.332], [6.639, 8.026, 3.99], [3.734, 7.571, 16.745], [-4.534, -8.891, 2.972], [0.533, 8.625, 2.667], [2.243, -10.532, 9.704], [-1.24, 6.822, 12.744], [-1.406, 0.307, 18.737], [1.06, -7.19, 12.95], [-7.736, -2.805, 0.231], [-9.872, -6.592, 19.123], [-3.598, -9.732, 3.749], [0.846, 4.612, 2.71], [-5.043, 0.679, 19.922], [-0.881, -9.866, 2.161], [4.817, -9.622, 0.188], [-3.079, 1.763, 3.338]

Problem instance 3: [-9.44, 7.659, 0.317], [-6.904, 7.132, 8.453], [4.057, -10.813, 9.813], [-1.973, 1.763, 13.425], [2.365, 10.042, 3.374], [2.327, 6.988, 3.359], [10.709, 1.58, 14.871], [-9.583, 7.619, 18.653], [-10.065, 1.605, 2.583], [-0.303, -7.82, 18.295], [-0.65, -2.707, 6.404], [4.264, 2.191, 19.06], [8.682, 6.306, 7.024], [1.308, -6.803, 18.024], [3.102, 6.131, 5.174], [0.19,

**Problem instance 4:** [5.656, -10.603, 4.379], [-2.782, -0.962, 0.317], [3.189, 10.73, 5.042], [0.799, 1.904, 13.085], [10.511, -10.171, 1.632], [5.498, -1.292, 3.663], [1.438, -5.972, 0.488], [7.982, 7.079, 12.139], [4.871, -7.951, 18.476], [2.482, -7.143, 13.341], [2.413, 5.708, 16.015], [-0.025, -2.919, 4.304], [-3.961, 9.562, 0.392], [-5.002, 9.668, 0.617], [-8.057, 6.802, 18.274], [5.338, -5.944, 17.171], [8.989, 10.87, 7.79], [-10.428, 8.422, 11.239], [-6.712, 3.215, 18.628], [-4.89, -9.572, 8.739], [-8.63, 5.128, 3.949], [5.579, 6.19, 8.156], [-5.31, 2.8, 4.196], [-1.808, -4.33, 4.303], [-0.918, 2.687, 2.976], [3.739, -10.449, 14.047], [2.466, 10.171, 7.218], [6.287, 2.348, 14.707], [2.567, -3.981, 11.635], [1.03, -9.413, 12.252], [5.667, -0.643, 16.254], [8.529, -2.307, 15.11], [-10.084, -2.025, 2.702], [-7.639, -2.093, 2.029], [-2.3, 8.223, 19.324], [-1.208, -4.033, 11.953], [8.799, -6.5, 7.675], [-2.895, -1.962, 12.774], [3.013, -7.381, 7.67], [1.083, -3.526, 0.047], [-10.384, 5.696, 10.272], [-8.69, -8.044, 7.34], [-3.893, 10.651, 16.591], [-2.398, 198
Problem instance 5: 

-5.737, -4.861, 5.265, 

[7.71, -0.536, 16.87, [4.1, -2.134, 15.504],

[5.165, -5.257, 2.493], [7.599, -8.152, 9.111], [-5.467, 9.622, 7.032], [4.834, 7.659, 0.603],

[8.882, -3.944, 11.039], [-7.854, -3.572, 2.82], [5.746, 3.329, 5.066], [-2.419, 6.864, 13.456],

[-3.617, -6.166, 0.458], [-7.115, -10.946, 9.832], [-6.291, 8.878, 12.081], [-0.194, -7.469, 13.046],

[7.969, -2.981, 17.17], [-6.233, 7.25, 6.903], [4.583, 9.424, 2.404], [7.661, 10.731, 9.041], [8.417, 3.61, 6.417], [0.064, -0.404, 18.423], [-7.64, 3.655, 15.584], [-0.698, 4.998, 10.464], [6.663, 2.521, 5.727], [8.452, -6.072, 7.859], [-1.337, -5.208, 11.065], [-6.619, 9.115, 12.474], [0.586, -2.201, 2.861], [-3.334, 10.218, 2.661], [5.9, 7.457, 11.968], [2.321, -5.813, 8.615], [-7.641, 8.776, 10.234], [-2.281, -8.141, 12.609], [3.38, 10.856, 4.975], [-2.996, -8.147, 5.043], [1.042, -0.395, 1.589], [3.304, -10.926, 7.352], [5.435, 1.085, 1.715], [-1.671, 1.684, 7.809], [-6.978, 1.984, 10.048], [8.045, -5.282, 5.129], [8.037, 10.589, 5.473], [-0.664, 5.548, 2.858], [4.244, -10.219, 19.529], [-7.208, 3.919, 2.279], [-3.114, -9.132, 6.143], [10.433, 2.215, 11.395], [8.671, 9.331, 0.966], [-7.66, 8.262, 1.138], [-9.273, 5.202, 5.406], [-7.633, -6.186, 6.53], [-7.133, 1.027, 9.624], [1.935, -10.133, 14.044], [1.945, 8.85, 5.136], [9.028, -10.397, 19.5], [-2.467, 9.423, 5.043], [-2.235, -1.29, 14.43], [-7.553, -9.93, 0.114], [-7.506, -10.399, 19.731], [-6.94, 3.244, 15.603], [6.289, 4.09, 7.427], [-9.168, 2.064, 11.679], [1.976, -6.51, 5.239], [4.688, 2.7, 19.531], [-0.884, 1.148, 16.939], [7.853, 2.056, 19.757], [7.458, -5.341, 3.909], [5.916, -5.196, 3.145], [0.605, 3.423, 2.629], [6.5, 5.032, 6.771], [5.974, 4.868, 16.298], [-4.026, 6.285, 7.461], [0.86, 3.335,
Problem instance 6: 

Problem instance 7: 
Problem instance 9: 

[6.262, -1.429, 6.279], [7.305, 2.7, 4.205], [5.158, 8.293, 1.349], [-1.618, 8.042, 15.003], [6.031, -0.426, 7.103], [5.4, -6.188, 17.926], [10.154, 2.544, 16.498], [6.936, -0.723, 9.684], [-1.824, -1.635, 8.728], [7.825, -10.78, 10.581], [-3.24, -7.07, 17.145], [1.045, 4.322, 0.073], [-4.116, -0.988, 6.25], [5.908, -7.644, 3.309], [0.544, 9.938, 9.972], [5.276, -8.303, 3.477], [1.49, -0.819, 8.82], [-1.279, 9.556, 16.004], [5.898, -6.787, 7.504], [-2.014, -7.761, 11.899], [-1.414, -9.013, 15.753], [-0.098, 1.262, 15.077], [8.187, -7.61, 4.088], [-9.793, -6.605, 11.06]

Problem instance 10: [2.549, 6.855, 2.432], [-1.11, -0.649, 8.911], [7.313, 7.636, 4.89], [4.24, -4.843, 19.765], [0.212, 1.466, 11.362], [4.74, 0.447, 13.669], [-9.636, 5.761, 16.373], [3.271, 6.617, 11.378], [-3.825, -9.275, 8.271], [10.016, 4.477, 19.93], [-6.211, 8.962, 0.389], [-1.07, -3.814, 15.35], [1.676, -7.031, 7.338], [-1.011, 9.504, 7.232], [-2.089, 5.571, 18.935], [-9.715, -10.606, 16.754], [-4.018, 2.664, 13.13], [-4.145, -4.889, 4.36], [5.311, -1.556, 16.088], [-9.189, -8.162, 16.385], [-8.808, -3.777, 8.741], [2.567, -1.28, 5.719], [-5.729, 7.63, 17.658], [1.391, 1.558, 6.896], [9.498, -0.665, 6.717], [5.52, -1.713, 0.201], [1.849, 3.735, 6.454], [-1.929, -10.404, 3.289], [-5.293, -1.111, 1.998], [10.721, -6.218, 14.343], [9.318, 3.436, 7.084], [8.478, 7.362, 12.915], [-4.79, -8.404, 1.272], [4.434, -1.723, 3.344], [5.33, 0.113, 7.431], [10.017, 9.26, 1.958], [-5.556, 3.952, 9.428], [-6.424, 0.147, 3.811], [-7.728, 7.359, 4.69], [-4.264, -5.179, 8.286], [3.801, -2.769, 16.521], [-10.613, -9.019, 3.542], [2.459, -4.452, 19.301], [-5.96, -8.309, 5.814], [-2.156, 4.003, 10.979], [2.074, -6.835, 15.315], [10.069, -1.065, 16.199], [4.386, 7.144, 19.338], [8.334, 0.45, 7.279], [-8.881, 3.217, 17.83], [8.04, 1.02, 2.509], [4.185, 1.749, 0.858], [-5.033, -5.806, 16.202], [-8.254, 3.797, 15.356], [5.224, 0.512, 11.41], [1.746, 0.133, 12.522], [4.267, 8.045, 4.648], [-9.781, -3.176, 2.806], [-7.465, -5.979, 4.737], [-7.934, 2.437, 3.119], [-5.556, 9.26, 18.221], [3.847, -3.524, 3.385], [1.376, -10.013, 6.345], [-2.86, -8.35, 12.14], [7.576, 4.928, 14.506], [7.705, -9.483, 0.805], [-8.82, -2.28, 8.057], [6.597, -1.586, 10.831], [6.074, 3.33, 16.897], [9.962, 6.427, 14.092], [-0.82, 9.129, 9.031], [0.601, -0.554, 3.754], [5.747, 0.879, 18.87], [4.89, -4.714, 17.664], [-4.561, -10.21, 0.424]
A.5  The 100 user scenario

Problem instance 1: [2.782, 8.525, 0.013], [10.963, 0.245, 5.11], [-10.852, -9.216, 19.153], [-5.341, -7.088, 14.729], [11.294, -11.499, 14.45], [3.473, 12.484, 15.753], [4.214, -11.125, 19.662], [-11.124, -0.546, 14.269], [-10.203, -1.633, 3.737], [-5.327, 6.828, 14.803], [-3.587, 0.166, 8.585], [-3.53, -12.6, 12.282], [-9.305, 0.85, 1.925], [0.213, -6.78, 10.052], [11.646, 0.871, 10.943], [-12.49, -6.301, 10.304], [12.018, -10.913, 5.83], [2.354, 5.112, 0.042], [5.188, 9.922, 0.755], [-12.366, 11.68, 7.282], [1.206, 6.596, 6.955], [-8.037, 10.087, 17.293], [8.09, 9.389, 9.802], [-11.821, -5.498, 7.194], [8.69, -11.454, 18.491], [0.101, -11.988, 4.166], [-8.673, -12.621, 9.775], [-12.073, 8.782, 14.335], [-0.243, -4.051, 13.144], [-3.719, 0.122, 17.368], [10.113, 2.616, 1.009], [1.392, -9.245, 19.711], [-6.702, 11.06, 7.03], [8.12, -9.396, 19.099], [4.634, -5.749, 8.193], [12.112, 11.141, 0.284], [-5.559, -1.146, 10.279], [-0.851, -6.809, 19.315], [-4.415, -10.703, 13.917], [-8.131, -4.11, 12.623], [7.937, -1.466, 18.098], [-4.325, -10.566, 16.108], [7.007, 3.758, 18.198], [-3.047, 10.644, 12.932], [-5.063, -5.913, 1.301], [-3.991, 9.198, 12.931], [1.881, -7.447, 1.617], [6.446, 3.786, 1.716], [-8.391, -1.185, 18.109], [-5.681, 8.558, 2.69], [4.6, -2.025, 10.246], [8.431, -1.059, 4.575], [3.161, 1.632, 5.374], [-7.535, -11.327, 2.798], [9.074, 11.018, 5.496], [-2.614, -4.005, 13.783], [-10.796, -10.151, 18.704], [2.69, 3.482, 6.007], [-6.07, 12.228, 7.005], [-11.734, -4.423, 1.601], [5.346, -0.902, 7.818], [-8.083, 2.489, 11.707], [-4.127, -3.492, 9.733], [-5.096, 7.213, 3.81], [-12.522, 7.439, 2.828], [9.818, -2.619, 9.626], [-7.625, -2.335, 2.721], [-3.092, -3.566, 5.35], [12.499, 2.368, 14.046], [-5.096, 9.393, 14.616], [11.337, -10.397, 13.559], [-11.523, 4.226, 19.723], [-6.444, -0.43, 11.9], [-7.277, 4.04, 11.29], [4.563, 10.959, 7.164], [11.062, 8.528, 16.289], [-1.735, -4.377, 12.526], [4.404, -10.148, 10.574], [-10.046, -9.39, 5.59], [-7.631, 4.489, 4.192], [3.708, 7.52, 17.629], [1.38, -5.396, 15.95], [10.635, 7.929, 18.686], [-0.451, 8.246, 10.804], [-3.663, -10.744, 17.785], [-10.736, -6.084, 6.505], [-10.474, -8.759, 12.94], [8.074, -10.465, 12.492], [-2.148, -4.022, 11.606], [-12.209, -7.332, 9.002], [-2.403,
Problem instance 2: 

\[-11.132, 10.626, 16.264\], 
\[-7.832, -6.224, 1.571\], 
\[-4.929, 5.897, 16.276\], 
\[8.379, -2.261, 16.322\], 
\[1.005, 9.359, 11.183\], 
\[6.077, -3.713, 10.679\], 
\[-0.613, 7.977, 9.016\], 
\[1.514, 12.568, 13.796\], 
\[-9.866, -2.294, 5.947\], 
\[11.08, 1.423, 12.313\], 
\[-5.211, 3.86, 9.247\], 
\[-9.406, 2.404, 10.161\], 
\[5.861, -8.989, 2.67\], 
\[8.341, 5.872, 0.845\], 
\[4.222, 11.515, 18.317\], 
\[8.654, 1.96, 10.013\], 
\[-0.966, 10.396, 16.846\], 
\[-7.564, 7.918, 8.267\], 
\[11.956, 1.889, 0.845\], 
\[-10.838, 6.505, 11.699\], 
\[8.939, -0.545, 2.33\], 
\[-8.703, -4.984, 19.088\], 
\[-11.62, 9.238, 9.598\], 
\[5.901, 9.819, 9.552\], 
\[9.744, 8.129, 8.333\], 
\[6.656, 1.752, 19.653\], 
\[-9.213, -6.601, 14.839\], 
\[8.022, 9.242, 2.049\], 
\[-1.624, -4.371, 5.718\], 
\[-7.869, 11.68, 0.854\], 
\[-1.113, -0.374, 12.465\], 
\[-7.829, -12.041, 14.466\], 
\[-10.681, 5.242, 0.254\], 
\[0.098, -11.909, 10.529\], 
\[2.35, 205]
Problem instance 4: [6.53, -12.243, 4.379], [-3.212, -1.111, 0.317], [3.683, 12.39, 5.042],
[0.923, 2.198, 13.085], [12.137, -11.745, 1.632], [6.349, -1.491, 3.663], [1.66, -6.896, 0.488],
[-0.029, -3.371, 4.304], [-4.573, 11.042, 0.392], [-5.776, 11.163, 0.617], [-9.304, 7.854, 18.274],
[6.164, -6.864, 17.171], [10.38, 12.552, 7.79], [-12.041, 9.725, 11.239], [-7.75, 3.712, 18.628],
[-5.647, -11.053, 8.739], [-9.965, 5.922, 3.949], [6.442, 7.147, 8.156], [-6.132, 3.234, 4.196],
[-2.088, -4.999, 4.303], [-1.06, 3.102, 2.976], [4.317, -12.066, 14.047], [2.847, 11.744, 7.218],
[7.259, 2.711, 14.707], [2.964, -4.597, 11.635], [1.189, -10.869, 12.252], [6.543, -0.742, 16.254],
[9.849, -2.663, 15.11], [-11.644, -2.338, 2.702], [-8.82, -2.416, 2.029], [-2.655, 9.495, 19.324],
[-1.395, -4.657, 11.953], [10.161, -7.505, 7.675], [-3.342, -2.266, 12.774], [3.479, -8.523, 7.67],
[1.251, -4.072, 0.047], [-11.99, 6.578, 10.272], [-10.034, -9.289, 7.34], [-4.495, 12.299, 16.591],
[-2.768, 1.405, 15.162], [2.287, 1.115, 10.452], [-5.431, -4.785, 9.934], [-8.082, -6.573, 10.825],
[8.065, 3.217, 0.946], [10.347, -7.78, 10.487], [2.028, -10.971, 8.439], [-7.135, -5.616, 18.14],
[-10.332, -1.536, 0.171], [-5.98, -6.586, 9.412], [-0.98, -6.784, 7.455], [1.99, -9.551, 17.721],
[-3.74, 9.817, 18.667], [-8.695, -9.56, 15.757], [5.616, 2.304, 16.902], [-9.467, 0.98, 10.426],
[1.751, 9.213, 18.491], [-5.405, 4.272, 18.265], [-4.49, 11.34, 18.962], [-7.527, 11.259, 18.402],
[10.936, 7.074, 5.144], [3.084, 2.609, 17.681], [-11.032, 8.865, 3.011], [-5.029, 4.197, 6.328],
[-11.376, -6.347, 8.086], [-2.51, -9.609, 2.225], [11.138, 11.994, 10.388], [2.442, -1.04, 0.032],
[6.147, -0.246, 4.465], [-1.372, -9.73, 2.666], [3.248, -12.039, 1.16], [-8.337, 11.534, 18.134],
[4.702, -10.213, 9.275], [-12.037, -0.543, 7.445], [7.867, 9.689, 19.933], [-2.154, -8.01, 17.152],
[8.393, -0.435, 4.884], [-6.337, 11.552, 1.779], [8.704, -4.951, 8.651], [-7.487, -0.082, 15.009],
[3.794, 11.801, 17.143], [-5.533, -10.183, 3.831], [-10.467, -9.185, 18.579], [2.596, 6.62, 6.828],
[-6.037, -11.141, 14.486], [6.883, -2.365, 11.66], [5.348, -7.574, 12.913], [5.862, 4.73, 4.626],
Problem instance 6:  
Problem instance 10: [2.943, 7.916, 2.432], [-1.282, -0.75, 8.911], [8.444, 8.817, 4.89],
[4.896, -5.592, 19.765], [0.244, 1.693, 11.362], [5.473, 0.517, 13.669], [-11.127, 6.653, 16.373],
[3.778, 7.641, 11.378], [-4.417, -10.709, 8.271], [11.565, 5.17, 19.93], [-7.172, 10.349, 0.389],
[-1.235, -4.404, 15.35], [1.936, -8.119, 7.338], [-1.168, 10.974, 7.232], [-2.412, 6.433, 18.935],
[-11.218, -12.247, 16.754], [-4.64, 3.077, 13.13], [-4.786, -5.645, 4.36], [6.133, -1.797, 16.088],
[-10.61, -9.424, 16.385], [-10.171, -4.361, 8.741], [2.965, -1.478, 5.719], [-6.615, 8.81, 17.658],
[1.607, 1.799, 6.896], [10.967, -0.768, 6.717], [6.374, -1.979, 0.201], [2.135, 4.313, 6.454], [-2.228, -12.013, 3.289],
[-6.111, -1.283, 1.998], [12.38, -7.179, 14.343], [10.759, 3.967, 7.084],
[9.79, 8.501, 12.915], [-5.531, -9.705, 1.272], [5.12, -1.99, 3.344], [6.154, 0.131, 7.431], [11.567, 10.693, 1.958],
[-6.415, 4.563, 9.428], [-7.418, 0.17, 3.811], [-8.924, 8.497, 4.69], [-4.924, -5.98, 8.286],
[-2.489, 4.622, 10.979], [2.395, -7.893, 15.315], [11.627, -1.229, 16.199], [5.065, 8.249, 19.338],
[9.623, 0.52, 7.279], [-10.254, 3.714, 17.83], [9.284, 1.178, 2.509], [4.833, 2.02, 0.858],
[-5.812, -6.704, 16.202], [-9.531, 4.384, 15.356], [6.032, 0.591, 11.41], [2.016, 0.154, 12.522],
[4.928, 9.289, 4.648], [-11.294, -3.668, 2.806], [-8.62, -6.904, 4.737], [-9.162, 2.814, 3.119],
[8.748, 5.691, 14.506], [8.897, -10.95, 0.805], [-10.184, -2.633, 8.057], [7.618, -1.831, 10.831],
[7.014, 3.846, 16.897], [11.503, 7.422, 14.092], [-0.947, 10.541, 9.031], [0.694, -0.64, 3.754],
[6.636, 1.014, 18.87], [5.646, -5.443, 17.664], [-5.267, -11.79, 0.424], [-10.72, -1.71, 8.619],
[-1.018, -2.937, 0.053], [6.624, 0.552, 9.238], [-10.534, 12.301, 1.187], [11.212, 2.101, 1.46], [-0.265, 8.228, 7.891],
[-10.86, -7.464, 6.031], [3.797, -4.066, 5.626], [7.644, 7.595, 4.296], [5.445, 6.519, 16.725],
[-3.541, -10.011, 15.248], [-0.77, -4.894, 5.555], [0.975, 6.139, 3.227], [8.107, 6.646, 1.64],
[8.942, 10.952, 11.063], [-2.55, 10.216, 4.098], [10.474, -6.955, 17.282], [-0.823, -5.334, 13.838],
A.6  The 150 user scenario

### Problem instance 3

Problem instance 9: [-4.004, -15.143, 0.69], [-4.13, 5.359, 19.162], [-9.139, -12.112, 4.057], [3.048, -4.969, 5.559], [8.793, -11.936, 0.4], [-11.539, -5.377, 7.635], [-6.493, -9.254, 0.179], [-0.922, 3.937, 9.586], [-9.52, -14.008, 19.424], [-9.02, 15.011, 4.893], [-5.597, -5.74, 5.256], [4.529, -2.76, 8.19], [4.313, 2.591, 4.807], [6.308, 5.053, 4.96], [-12.858, -12.823, 14.013],

Bibliography


