CLUSTERING SENSORS IN WIRELESS *AD HOC* NETWORKS USING DYNAMIC EXPECTED COVERAGE MODEL

Dipesh Patel, Rajan Batta, and Rakesh Nagi

Center for Multisource Information Fusion
and
Department of Industrial Engineering
University at Buffalo (SUNY), Buffalo, NY 14260, USA

November 18, 2002

Outline

- Introduction
- Literature Review
- Problem Description and Formulation
- Solution Approaches
  - I. Relocation Heuristic
  - II. No Relocation Heuristic
  - III. Column Generation
- Numerical Results
- Conclusion and Future Work

INTRODUCTION

Data Fusion: An Introduction

- Data Fusion can be formally defined as
  "a multilevel, multifaceted process dealing with detection, association, correlation, estimation and combination of data and information from multiple sources to achieve refined state and identity estimation, and complete and timely assessments of situation and threat."
- The combination of data from multiple sensors is termed as MultiSensor Data Fusion Process
- Multisensor Data Fusion can be classified into:
  - Centralized Data Fusion and
  - Decentralized Data Fusion
- Logistics of fusion process

- Entities
  - Sensors (e.g. Radars, Battletanks, UAV's, Soldiers, Sonars etc.)
  - Clusterheads (AWACS and UAV's)
- Issues
  - Sensor Mobility
  - Link Jamming or Failure
  - Limited number of clusterheads
  - Transmission Range
- Objective
  - To facilitate collection of data from multiple mobile sensors, we suggest an optimal clusterhead location in time
  - The optimal clusterhead location strategy, is modeled as a covering location problem

LITERATURE REVIEW

- Covering Location Problem
  - Deterministic Covering Models
  - Stochastic Covering Models
- Wireless *Ad Hoc* Networks
  - Network Topology
  - Location Management
  - Routing Protocols
  - Clustering Algorithms

Covering Location Problem

- Deterministic Covering Location Problem
  - Toregas et. al. (1971) location of emergency service facilities as Set Covering Location Problem (SCLP)
  - Church and Revelle (1974) location of limited number of emergency service facilities as Maximal Covering Location Problem (MCLP)
  - Hogan and Revelle (1986) introduced the concept of backup coverage.
- Stochastic Covering Location Problem
  - Chapman and White (1974) probabilistic version of SCLP
  - Daskin (1983) MCLP considering unavailability of facilities to serve demand nodes at all times: Maximum Expected Covering Location Problem (MEXCLP)
  - Batta et. al. (1989) generalizes assumptions made in MEXCLP model

Wireless *Ad Hoc* Networks

- Network Topology: Flat vs Hierarchical Architecture
- Location Management
- Routing Protocols
  - Proactive Protocols
  - Reactive Protocols
  - Hybrid Protocols
- Clustering Algorithms
  - Graph Based Clustering
    - Highest Degree Heuristic, Lowest ID Heuristic, Node Weight Heuristic
    - Weight Based Clustering Algorithm, MOBIC Clustering Algorithm
  - Geographical Based Clustering

PROBLEM DESCRIPTION AND FORMULATION

Dynamic MEXCLP: Battlefield Scenario

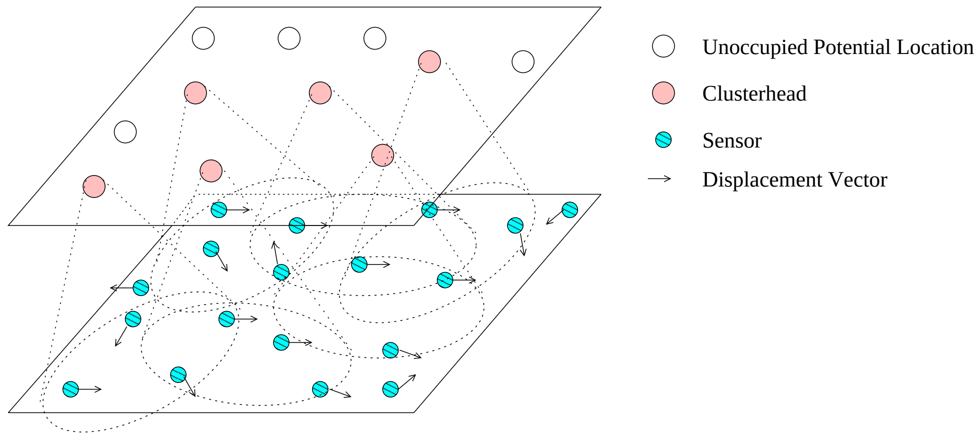

Figure 1: A snapshot of the network

Dynamic Maximum Expected Covering Location Problem Model:

- Sensors are mobile and the velocity of each sensors are known
- Potential clusterhead locations constitutes a set of discrete points
- Relocation of clusterheads takes place at discrete time periods
- The probability of failure of all links is the same

Maximum Expected Covering Location Problem (MEXCLP)

- Daskin (1983) proposed the MEXCLP model to capture the uncertainity of a facility serving demand node within its coverage
- MEXCLP associated a probability of working with each facility
- MEXCLP considers a static scenario, i.e., the demand nodes are stationary

Parameters:

Δ = set of potential facility locations.

Θ = set of demand nodes.

n = maximum number of facilities to be located.

T = maximum number of time periods in the horizon under consideration.

U = the distance beyond which a demand node is considered "uncovered".

D_{ikt} = distance between potential facility location i and demand node k at time t.

d_k = demand per period of node k.

p = probability of a link failure per period. $(0 < p < 1)$

$$r_{ikt} = \begin{cases} 1, & \text{if } D_{ikt} < U. \\ 0 & \text{otherwise.} \end{cases}$$

C = one-half of relocation cost.

The decision variables of the problem are:

$$x_{it} = \begin{cases} 1, & \text{if a facility is placed at location } i \text{ at time } t \\ 0 & \text{otherwise.} \end{cases}$$

$$y_{jkt} = \begin{cases} 1, & \text{if demand node } k \text{ is covered by at least } j \text{ facilities at time } t \\ 0 & \text{otherwise.} \end{cases}$$

w_{it} = positive difference in the number of facilities at location i between time $t-1$ and time t

Dynamic MEXCLP

Maximize $$\sum_{t=0}^{T} \sum_{k\in\Theta} \sum_{j=1}^{n} (1-p)p^{j-1} d_k y_{jkt} - \sum_{t=1}^{T} \sum_{i\in\Delta} C w_{it},$$

subject to:

$$\begin{aligned}
\sum_{j=1}^{n} y_{jkt} - \sum_{i\in\Delta} r_{ikt} x_{it} &\leq 0 \quad \forall \;\; k = 1, \ldots, |\Theta|, t = 0, \ldots, T, && (1)\\
\sum_{i\in\Delta} x_{it} &\leq n \quad \forall \;\; t = 0, \ldots, T, && (2)\\
w_{it} &\geq x_{it-1} - x_{it} \;\; \forall \;\; i = 1, \ldots, |\Delta|, t = 1, \ldots, T, && (3)\\
w_{it} &\geq x_{it} - x_{it-1} \;\; \forall \;\; i = 1, \ldots, |\Delta|, t = 1, \ldots, T, && (4)\\
x_{it} &\in \{0, 1\} \quad \forall \;\; i = 1, \ldots, |\Delta|, t = 0, \ldots, T, && (5)\\
w_{it} &\geq 0 \quad \forall \;\; i = 1, \ldots, |\Delta|, t = 1, \ldots, T, && (6)\\
y_{jkt} &\leq 1 \;\; \forall \;\; j = 1, \ldots, n, k = 1, \ldots, |\Theta|, t = 0, \ldots, T. && (7)
\end{aligned}$$

- Observation 1: Link failure and Clusterhead failure have similar modeling
- Observation 2: w_{it} and y_{jkt} takes only binary values

Clustering Sensors with Mobile Facilities

- In this model we elect clusterheads among the sensors and the clusterheads are mobile.

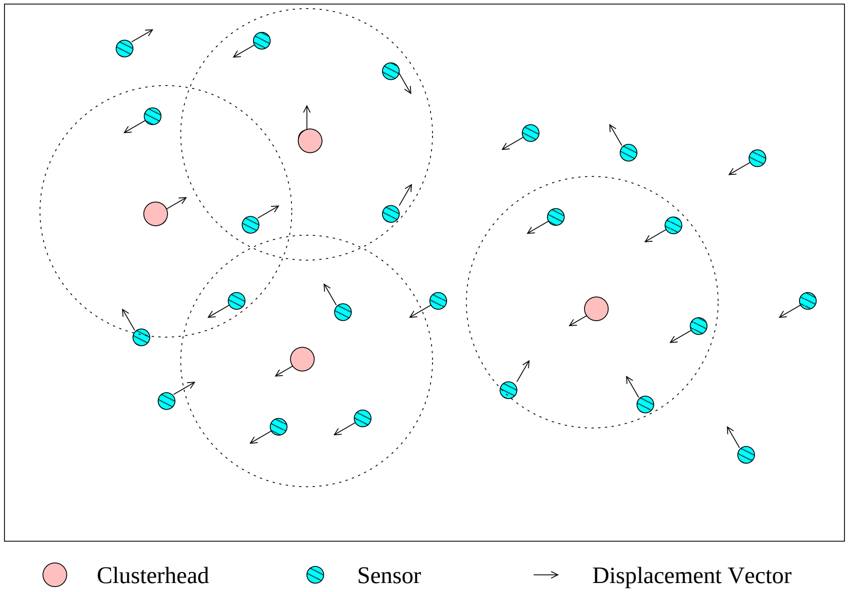

Figure 2: A snapshot of the network

SOLUTION APPROACHES

Relocation Heuristic

- Relocation Heuristic (RH) procedure is as follows:
  - Calculate demand covered by each potential location for time $t = 0$.
  - Selects the best location.
  - Updates the demand of the covered sensors.
  - Selects the second best and so on until all clusterheads are located.
  - Repeats the procedure for next time period by assigning an additional weight of relocation cost to each of the locations selected in the last time period.
- The results show that the RH performs considerably well for low relocation cost.
- The worst case analysis shows that the heuristic can perform arbitrarily bad in terms of solution quality with increasing number of time slots.

No Relocation Heuristic

- No Relocation Heuristic (NRH) does not allow relocation and the procedure is as follows:
  - Calculate demand covered by each location for all time period.
  - select the best location and place the clusterhead at that location for all time periods.
  - update the demand of the covered sensors.
  - continue the process until all clusterheads are placed
- NRH performs considerably well for high relocation cost.
- The worst case analysis shows that the heuristic can perform arbitrarily bad in terms of solution quality with increasing number of time slots.

Column Generation

- To improve heuristic solution.
- Heuristic solution serves as initial basic feasible solution for CG.
- Some addditional definitions are:

F_t = set of feasible solutions for time t.

x_{sit} = value of x_{it} if solution s is selected at time t.

y_{sjkt} = value of y_{jkt} if solution s is selected at time t.

The decision variables of the problem are:-

$$F_{st} = \begin{cases} 1 \text{ if solution } s \text{ is selected at time } t \\ 0 \text{ otherwise.} \end{cases}$$

- Dynamic MEXCLP is decomposed in to master problem and sub-problem

CG: Master Problem

- Relaxed Master Problem is solved at each iteration.
- Selects feasible solution.
- The decomposed master problem for the CG approach is as follows:

$$\text{Maximize} \quad \sum_{t=0}^{T} \sum_{k\in\Theta} \sum_{j=1}^{n} \sum_{s\in F_t} (1-p)p^{j-1} d_k F_{st} y_{sjkt} - \sum_{t=1}^{T} \sum_{i\in\Delta} C w_{it}$$

subject to

$$\begin{aligned}
\sum_{s\in F_{t-1}} F_{st-1} x_{sit-1} - \sum_{s\in F_t} F_{st} x_{sit} &\leq w_{it} && \forall\ i=1,\dots,|\Delta|, t=1,\dots,T && (8)\\
\sum_{s\in F_t} F_{st} x_{sit} - \sum_{s\in F_{t-1}} F_{st-1} x_{sit-1} &\leq w_{it} && \forall\ i=1,\dots,|\Delta|, t=1,\dots,T && (9)\\
\sum_{s\in F_t} F_{st} &= 1 && \forall\ t=0,\dots,T && (10)\\
F_{st} &\in \{0,1\} && \forall\ s, t=0,\dots,T && (11)\\
w_{it} &\geq 0 && \forall\ i=1,\dots,|\Delta|, t=1,\dots,T && (12)\\
& && && (13)
\end{aligned}$$

CG: Sub-Problem

$$\text{Maximize} \quad \sum_{k\in\Theta}\sum_{j=1}^{n}(1-p)p^{j-1}d_k y_{jkt} + \sum_{i\in\Delta}[\beta_{it}-\beta_{it+1}-\gamma_{it}+\gamma_{it+1}]x_{it}-\delta_t$$

subject to

$$\sum_{j=1}^{n} y_{jkt} - \sum_{i\in\Delta} r_{ikt}x_{it} \le 0 \qquad \forall\ k=1,\ldots,|\Theta| \tag{14}$$

$$\sum_{i\in\Delta} x_{it} \le n \tag{15}$$

$$y_{jkt} \le 1 \qquad \forall\ j=1,\ldots,n, k=1,\ldots,|\Theta| \tag{16}$$

$$x_{it} \in \{0,1\} \qquad \forall\ i=1,\ldots,|\Delta| \tag{17}$$

- The objective function of the sub-problem for $t=0$ is as follows:
  $\sum_{k\in\Theta}\sum_{j=1}^{n}(1-p)p^{j-1}d_k y_{jkt} + \sum_{i\in\Delta}[-\beta_{it+1}+\gamma_{it+1}]x_{it}-\delta_t$
- The objective function of the sub-problem for $t=T$ is as follows:
  $\sum_{k\in\Theta}\sum_{j=1}^{n}(1-p)p^{j-1}d_k y_{jkt} + \sum_{i\in\Delta}[\beta_{it}-\gamma_{it}]x_{it}-\delta_t$
- Generates feasible solution using dual multipliers from master problem.

Strategies and Termination Criteria

- Strategy I: RMP and Sub-Problem is solved sequentially.
- Strategy II: Sub-Problem for all time period is solved consecutively
- Termination criteria selected for CG is to
  - Terminate when gap between RMP and LP relaxation is within 2%
  - Terminate within limited number of CG iterations allowed.

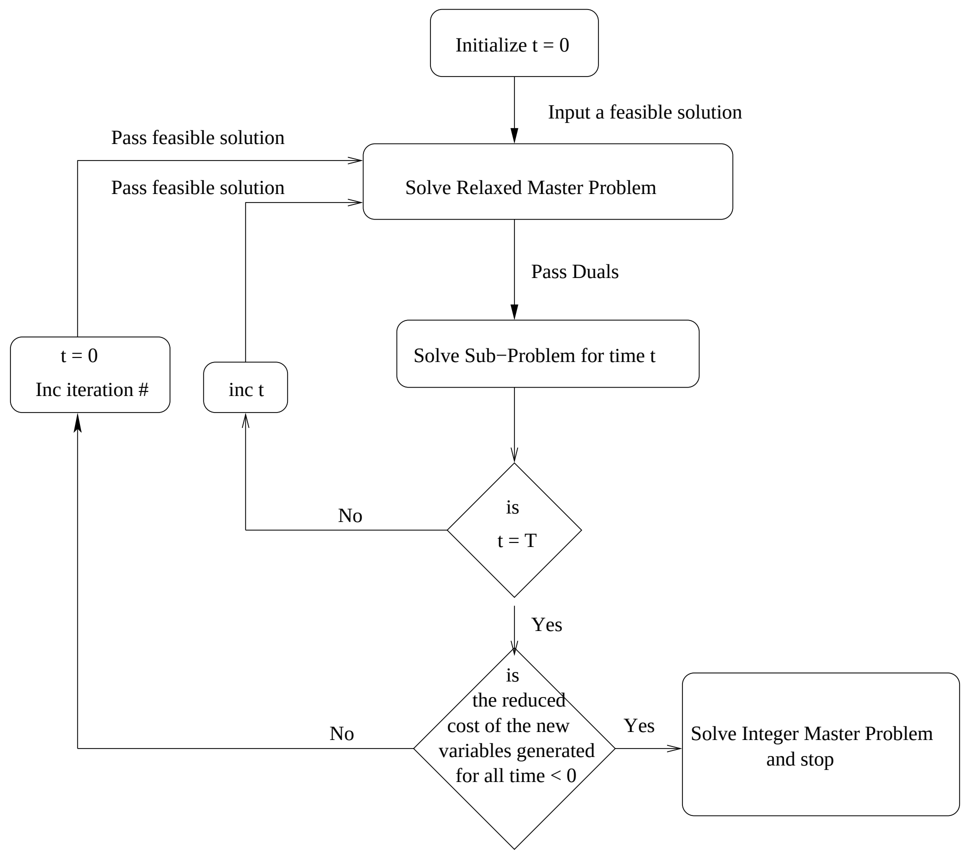

Figure 3: Strategy I

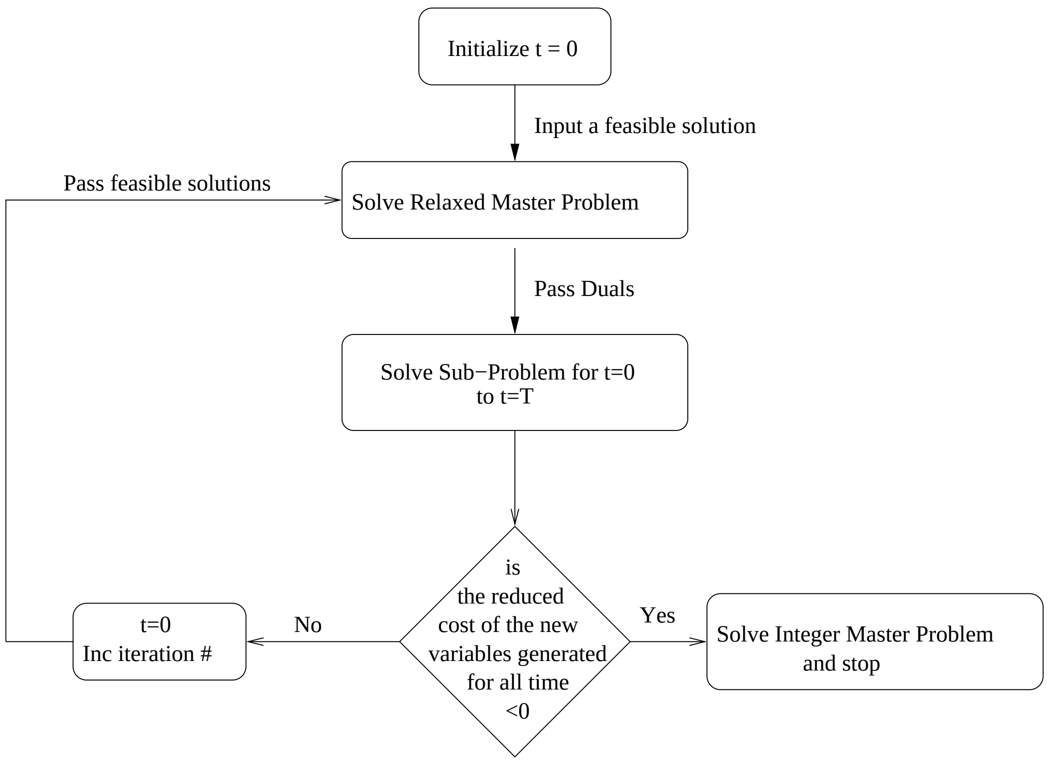

Figure 4: Strategy II

Integer Master Problem

- Integer Master Problem can be coverted into a longest path problem
- Drastically reduces the solution time to solve Integer Master Problem

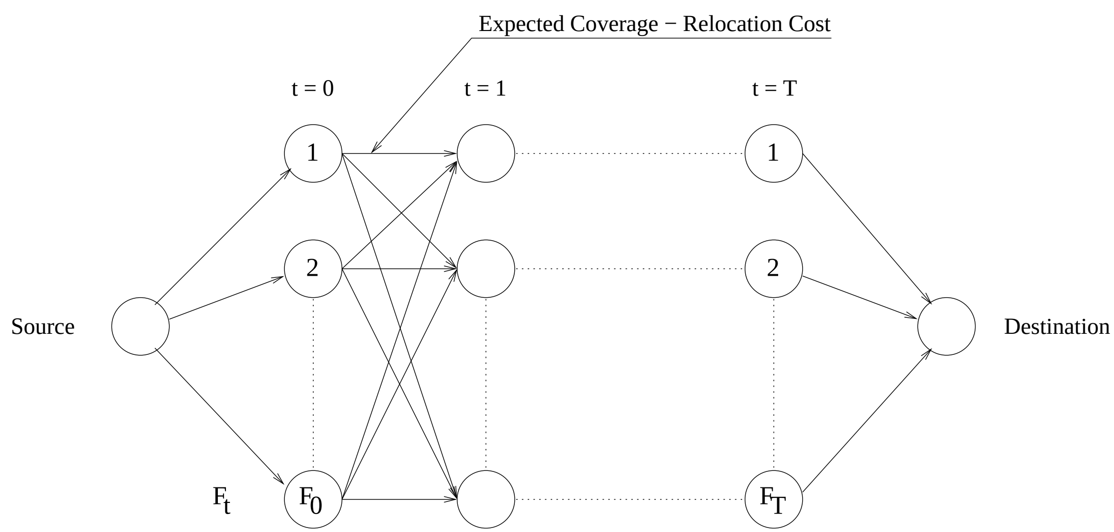

Figure 5: State Space Diagram: Integer Master Problem

NUMERICAL RESULTS

Medium Size Problem: Parameters

Parameter	Range/Value
Potential Location Range	0-1000
Sensors Location Range	0-500
Data Range	10-10
Displacement Range	0-10
n	10
$\lvert\Delta\rvert$	300
$\lvert\Theta\rvert$	200
U	100
Time slots $(T+1)$	11
p	varying between 0 and 1
C	varying between 0 and 50

Medium Size Problem: CPLEX vs CG

CPLEX solution value found within CG solution time for $C = 3$

p	CPLEX solution	CPLEX time	CG solution	CG time	% gap between (4) and (2)	CPLEX solution within CG time
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0	20820	1741	20686	4.172	0.64	NFSF
0.1	19054.44	444	18967.92	3.218	0.45	NFSF
0.2	17337.52	115	17201.6	3.109	0.78	NFSF
0.3	15597.47	91	15449.12	3.172	0.95	NFSF
0.4	13793.04	71	13641.6	3.125	1.1	NFSF
0.5	11890.25	67	11784.75	3.109	0.89	NFSF
0.6	9896.32	25	9830.88	2.938	0.66	NFSF
0.7	7765.68	4	7733.34	2.828	0.42	NFSF
0.8	5431.03	3	5399.52	2.796	0.58	NFSF
0.9	2853.13	3	2809.34	2.703	1.33	NFSF

NFSF: No Feasible Solution Found

Medium Size Problem: Improvement over RH and NRH

- The results show that CG improves over heuristic solution

RH, NRH and CG solutions for $C = 3$					
p	RH	NRH	CG	% improvement over RH	% improvement over NRH
0	20126	20330	20686	2.71	1.72
0.1	18496.41	18310.5	18967.92	2.49	3.47
0.2	16903.04	16705.6	17201.6	1.74	2.88
0.3	15256.81	15020.81	15449.12	1.24	2.77
0.4	13528.32	13326	13641.6	0.83	2.31
0.5	11766.5	11526.25	11784.75	0.15	2.19
0.6	9815.82	9641.6	9830.88	0.15	1.93
0.7	7690.77	7590.81	7733.34	0.55	1.84
0.8	5375.29	5327.04	5399.52	0.45	1.34
0.9	2786.47	2805.45	2809.34	0.81	0.14

Large Size Problem

- No feasible solution found using CPLEX within 4 hours
- Termination criteria for CG: RMP 2% off LP relaxation

Number of Facilities = 15 - 30; Number of Potential Locations = 800 - 1000
Number of Sensors = 400 - 500; Number of Time Slots = 10 -15;
Probability of Link Failure = 0.1 - 0.3; Relocation Cost = 60 - 100
Demand = 50 - 100; Coverage Radius = 80; Displacement = 0 - 60

Sr. No.	CG Solution	CG Time (sec)	CPLEX Solution	CPLEX Time	CG % gap with LP	% Improvement over heuristic solution
1	389143.47	321.77	NFSF	> 4 hrs	3.73	1.46
2	379934.40	4.99	NFSF	> 4 hrs	1.57	0.00
3	318656.64	4.34	NFSF	> 4 hrs	1.36	0.00
4	340129.02	4.84	NFSF	> 4 hrs	1.32	0.00
5	472456.44	6.77	NFSF	> 4 hrs	0.95	0.00

NFSF: No Feasible Solution Found

CONCLUSIONS AND FUTURE WORK

Conclusions

- CG Outperforms CPLEX in solution time.
- CG solution is within 4 % of the LP relaxation.
- Heuristics performs well in most of the randomly generated instances.
- ANOVA validates that CG in conjunction with heuristic outperforms CPLEX

Future Work

- Forming a connected network among clusterheads
- Load balancing
- Developing a simulation model and integrating it with the CG solution methodology.