DISTRIBUTED COMPUTING

001-P2P : A DGMA clustering algorithm for mobile ad-hoc networks

Introduction

This project aims to understand the working of a distributed group mobility adaptive clustering algorithm. The need for this kind of clustering algorithm is to reduce the network traffic in P2P systems when number of nodes becomes very large. The objective is to reduce network traffic by converting a flat P2P network to hierarchical P2P network where cluster member can interact with non-cluster member only through cluster head (CH) .

DGMA (distributed group mobility adaptive) algorithm deals with following aspects of a dynamic peer to peer system :

1. Distributed : The algorithm needs to be distributed and run identically on all nodes.
2. Group : The algorithm is designed for group mobility scenarios like military operations or search and rescue operations. Here only macro movements of group matter.
3. Adaptive : The algorithm should be adaptive to phenomenon of group mergence and partitions. These scenarios can be simulated using RRGM (Reference Region Group Mobility) model.

Design Objective Targets

In order to reduce number of cluster head reassignment or network restructuring (which requires additional cost of message exchanges) following objectives need to be met:

1. Prolonged lifetime of cluster as well as CH.
2. Adaptable to fast speed scenarios.
3. Reduce frequency of re-clustering iterations.

LDTSD : A metric for finding stability of a node given it's neighbours

To find cluster heads which are expected to have prolonged lifetime, authors introduce a metric (Linear distance based total spatial dependency) to measure the likelihood of a node to remain stable. If a node is said to be stable, then it has less tendency to drift away from it's neighbours thereby making the neighbourhood more stable.

The node first computes the mobility information using current and previous mobility information (timestamp, position). The mobility information consists of current speed and direction of node i at time t:

$$dx(t) = x(t) - x(t-1),\ dy(t) = y(t) - y(t-1)$$

$$D = \sqrt{dx^2 + dy^2}$$

$$S(i,t) = D/(T(t) - T(t-1))$$

$$tan(\varphi) = |dy(t)|/|dx(t)|$$

$$theta\ (i,t) = \begin{cases} \varphi * sign(dy(t)) & if \quad dx(t) > 0 \\ (\pi/2) * sign(dy(t)) & if \quad dx(t) = 0 \\ (\pi - \varphi) * sign(dy(t)) & if \quad dx(t) < 0 \end{cases}$$

Whenever, D is greater than a threshold, the movement of node is taken into account and the updated mobility information is exchanged with neighbouring nodes in transmission range.

<u>Algorithm</u>

1. Exchange S(i) - speed and Theta (i) with neighbours.
2. For j in range(len(neighbours)):
   a. $RD(i,j) = cos(theta(i) - theta(j))$
   b. $SR(i,j) = min(S(i), S(j))/\ max(S(i), S(j))$

3. $LDTSD(i) = RD(i)^T SR(i)$

DGMA Algorithm

The algorithm deals with three states of node :

1. Red : A node is said to be colored red if it is a clusterhead.
2. Yellow : The node is said to be colored yellow if it is a member of a cluster.
3. White : The node is inactive or is not a member of any cluster.

There are two routines which need to be run by algorithm :

1. Nomination routine : To nominate cluster heads when there are no clusterheads in neighbourhood or ratio of white nodes in neighbourhood crosses a threshold value.
2. Maintenance routine : To check and maintain membership table based on color of node (red, yellow or white).

These routines can be referred from paper.

Proofs related to Algorithm

Lemma1 : u and v are white nodes. Let v $\in$ neighbourhood(u) at the end of nomination process. Then prove that, if u.color = red $\Rightarrow$ v.color = yellow.

Proof by contradiction :

Assume v.color $\neq$ yellow, then there are two possible cases :

1. v.color is red -> After running the initial nomination process, it is not possible to have two adjacent red nodes as one of them will have higher LDTSD.
2. v.color is white -> By the end of nomination process, if v is a neighbour and u is a red node, then u would have sent a invitation request to v and v would have changed color to yellow.

Lemma2 : Worst case processing time complexity is O(n) per node.

Processing time complexity in :

1. Nomination routine : Worst case will be when all nodes are in neighbourhood. Number of iterations will be O(n).
2. Maintenance routine : For any color node, it can be observed that it is always linear time complexity O(n).

Lemma3 : Worst case message complexity is O(1) per node and over whole network it is O(n).

In maintenance routine :

Before running maintenance, every node broadcasts HELLO(nodeid, color, LDTSD). Hence, message complexity id O(1) for each node.

In nomination routine :

Only when a node is nominated, it broadcasts a invite message. Which adds a constant of O(1). Hence, message complexity still remains O(1).

Drawbacks and Possible Improvements

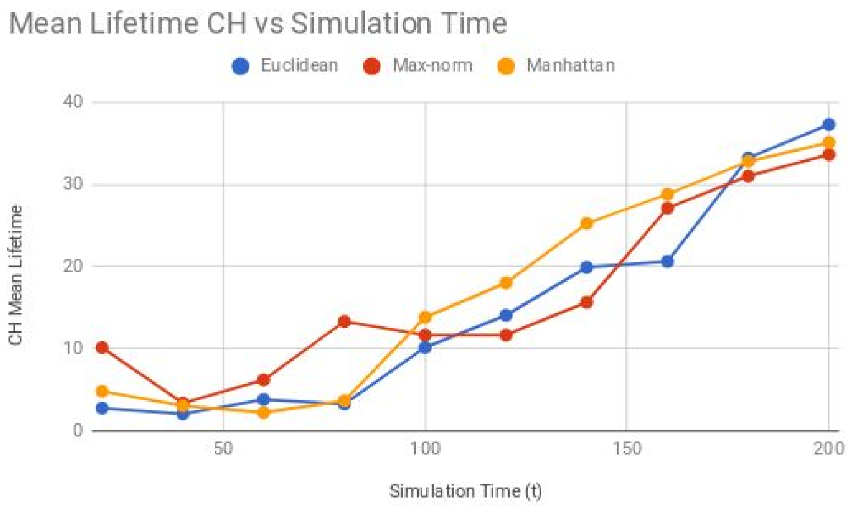
1. The stability of CH is calculated in present time. Although, the movements are random in many situations. However, it can be expected to infer future position of a node given previous positions of the node and it's neighbours using some kind of Bayesian Inference model. Once expectation of a future positions are possible, expected cluster stability can be calculated in advance. This hypothesis only works, if there is some pattern in movement which can be inferred.
2. Current DGMA algorithm only works good for group mobility simulations or movements. When movements of node are totally independent of their group center, number of CH assignments increases significantly.
3. Except after running nomination routine for first time, there is no guarantee that all nodes are covered by at least one CH or are a part of at least one cluster.
4. LDTSD is not defined when there are no neighbours in transmission range.

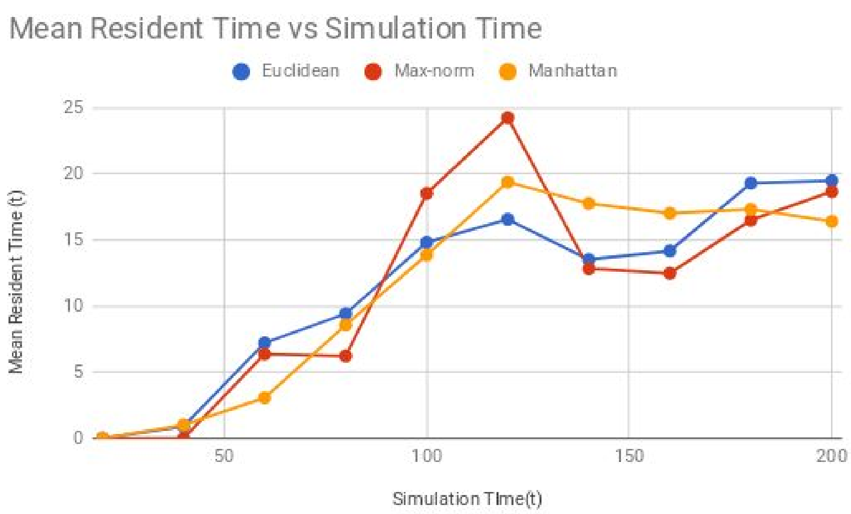
Experiments over different choice of distance metric :

I experimented with three distance metrics :

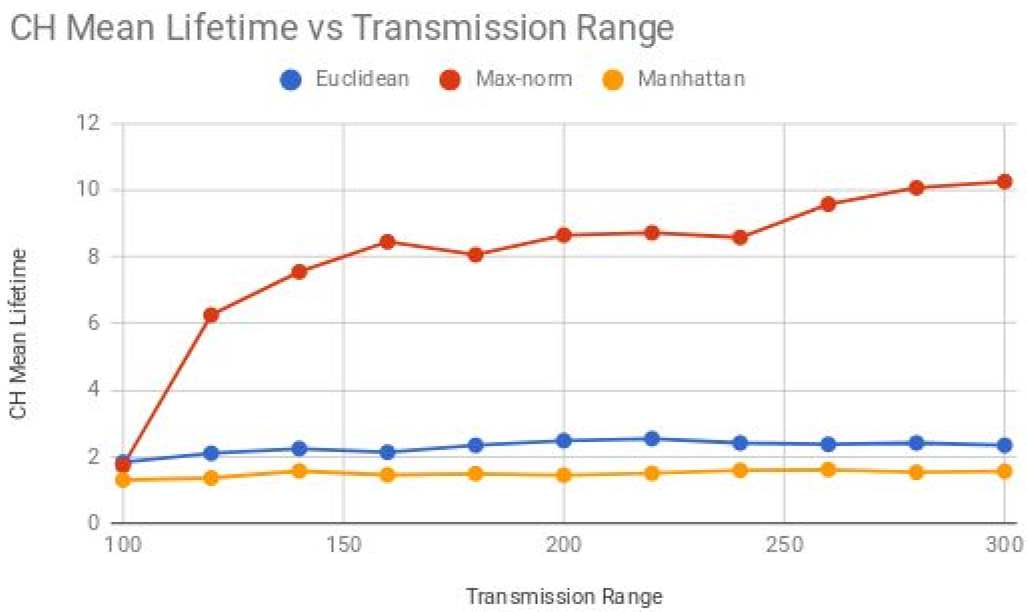
1. Euclidean distance metric (default)
2. Max-norm distance metric (max(dx, dy)) : Explaining significance of max-norm metric is difficult. It is only chosen for experimental purpose.
3. Manhattan distance metric (|dx| + |dy|)

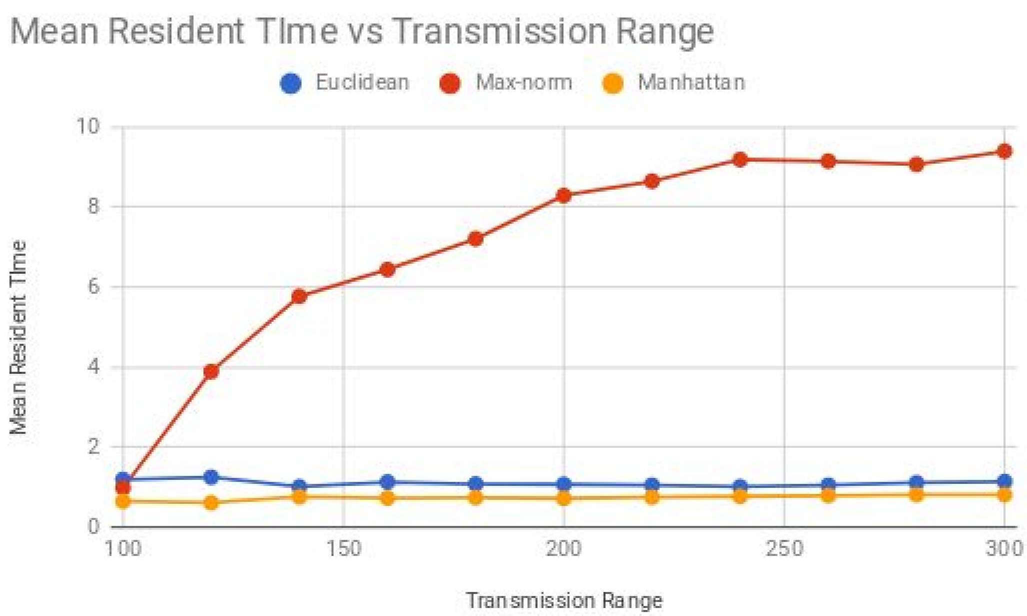
Experiment 1 : Performance parameters vs Simulation Time





Experiment 2 : Performance parameters vs Transmission Range





Experiment 1 doesn't give us any useful insights. However, in experiment 2 it is clear that stability of clusters dramatically increases as TRANSMISSION RANGE is increased when max-norm is used.