International Journal of Computer Science
and Engineering (IJCSE)
ISSN(P): 2278-9960; ISSN(E): 2278-9979
Vol. 3, Issue 4, July 2014, 79-88

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

SMART CITY ANALYSIS AND PREDICTIONS USING SPATIAL DATA AND EVOLUTIONARY COMPUTING TECHNIQUES

THOMAS JOSEPH

Assistant Professor, Department of Computer Science and Engineering,

St. Joseph's College of Engineering and Technology, Palai, Kottayam, Kerala, India

ABSTRACT

Practical problems with non-convex problems cannot be solved in a reasonable time. There is neither an easy way to classify a non-convex problems nor a shorter way to summarize the vast range of possibilities. Setting up an industry is always a Multi Objective Problem and a non-convex optimisation problem. A designer sets a pessimistic target value so the solution will not be a Pareto optimal solution. Finding a parallel development among the pareto frontier, there by achieving a tradeoff among the solutions using spatial data and predicting its sustainability, is a difficult task. Soft computing is used for solving this practical problem in a reasonable time. The prediction approaches used for parallel development are ant colony optimization and simulated annealing. The approaches give optimized solutions to set up the smart city [1] by using spatial data with low cost function.

KEYWORDS: Smartcity, Ant Based Optimisation, Prediciton Method, Simmulated Annealing, Spatial Data

INTRODUCTION

The migration of human to urban environments is occurring at an exceptional rate around the world. The citizens with a safe, healthy and sustainable environment where they can live, work and play in a smarter way should be attained [3]. A study done by an international agency shows that there will be an outbreak of human to urban environments occurring at an unprecedented rate around the world by around 2050.

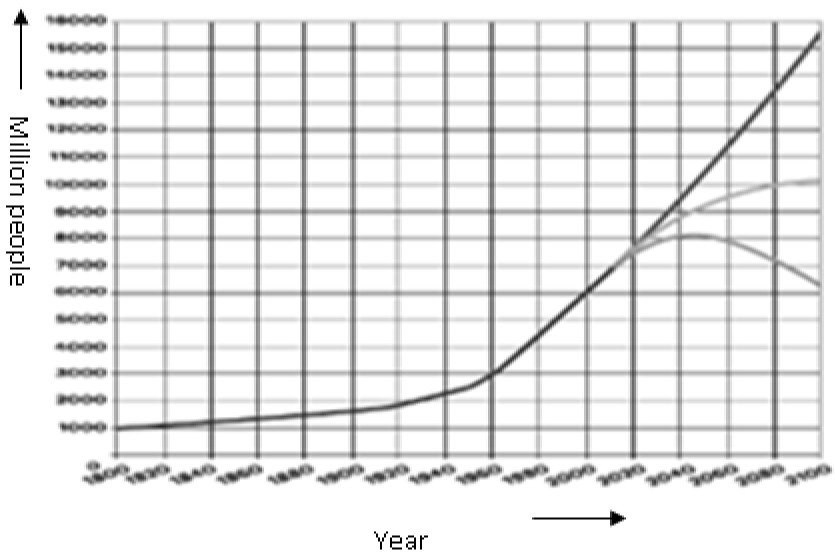

Figure 1: Graph Showing Unpredictable Outbreak of Human Population

The above fact reveals that there should be efficient planning for setting up smart cities [1] based on the spatial data to avoid exploitation of human on nature. The objective of the approach is to evaluate candidate places in smart city analysis, based on significant factors and predict optimal development in a parallel manner.

The user uses spatial data for predicting the resources in each candidate places. Using spatial data is very useful in many aspects: it provides the user the accurate information of each place, the data thus gathered can be easily processed by many GIS analysis systems. Accuracy and easy of processing the spatial data are the two main advantages of using spatial data for smart city analysis.

This work discusses two approaches for the spatial data to predict the candidate solutions. The approaches are Weighted Sum Approach [2] and Tchebycheff Approach [4]. The candidate solutions are then evaluated in a parallel manner using soft computing techniques. For predicting the optimal cost, simulated annealing and ant based optimisation techniques are used.

LITERATURE SURVEY

To solve a multi-objective optimisation problem it is necessary to convert the problem into single objective optimisation problem [4], by using adjustments such as weighted sum objectives or ε constant method. The weighted sum approach [2] gives weight to the different objectives and then factors in all these weights to form a single objective function that can be solved by using single factors of optimisation. This method is not entirely satisfactory because the weights cannot choose correctly to each objective. It indirectly leads that the approach cannot find correct solution for the original objective. The weighted sum approach [2] method was used in China, successfully to have a sustainable development of water supply network [19, 20].

In tchebycheff Approach [4] user knows the search space domain .The approach guarantee best results to the user because all candidate points are plotted in reference to one point. This guaranteed the best optimal result to the user in preprocessing method because based on a reference point all the candidate solutions are being plotted.

The constrains that been taken are weather conditions[8], policy of the government[9], infrastructure facilities, living conditions[10], power factors[11], mode of transportation, raw materials availability, communication facilities[12], boarding points and population density[13]. A sustainable smartcity [1] should consider both the hard and soft constrains as an icon. The hard constraints cannot be compromised whereas the soft constraints can be slightly compromised.

Ant based optimisation and simulated annealing are the two approaches used for predicting the sustainable icon of smart city [1]. Simulated annealing guarantee the best possible optimal distance. The worst paths will get eliminated in a faster rate at the initial temperature, later it convergence to the best optimal path gradually as temperature get decreased to nearby zero. The ant based optimisation approach is used for exploration of the search space. The pre requisite fact for the ant based approach is that the problem should be modeled as a graph. The nodes are the candidate places and the edge as the cost function between the nodes. Mapping n-dimensional cost edge to a weighted graph the user should use the euclidian distance [26]. The above approach is been used because the symmetric travel salesman problem has a euclidian distance based problem space. The above approach helps the user for the parallel development of candidate places and avoids exploitation of nature. The capital expenditure can be converted to operational expenditure, this helps to maximize the Return of Investment.

SYSTEM DESIGN

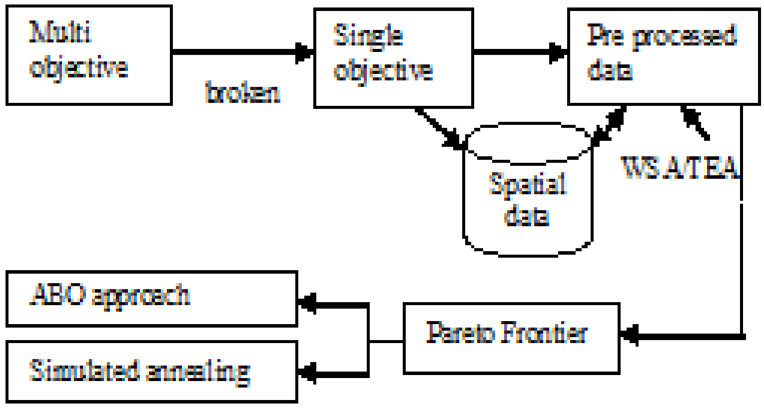

Figure 2: System Architecture

The input for the GIS application is spatial data. The data is available of two formats: scalar data and vector data[27] .Using vector type data is more significant because it contains the relevant information about the places. The spatial data along with other significant data will be get added to the spatial database for the pre-processing approach. Pre-processing is done by using two methods: weighted sum approach [2] and tchebycheff approach [4]. The processed data will be then populated on a pareto frontier. Solving a non-convex problem is neither easy nor shorter way to summarize the vast range of possibilities. The designer set a pessimistic target value in the preprocessing approach so the solution will not be a pareto optimal solution. For populating pareto set there are many methods such as random sampling, weighted sum method, distance method and constrained trade off method. Thus obtained pareto frontier solution termed as candidate solutions are evaluated using the simulated annealing and ant based optimisation approaches. The main objective of the soft computing technique is to provide the user an optimal solution with a low time complexity.

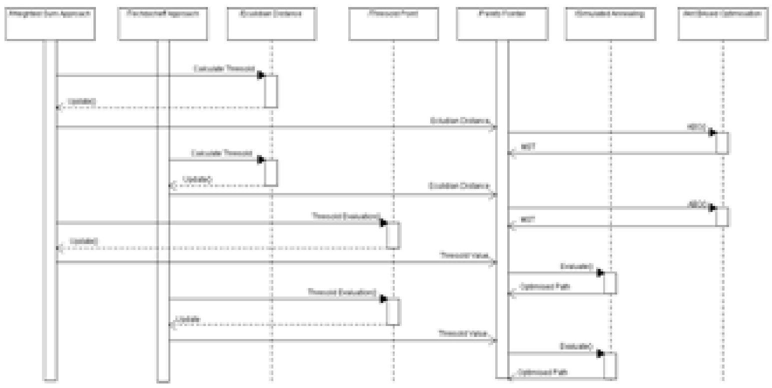

Figure 3: Sequence Diagram for Finding Optimal Solution from Candidate Places

The two other modules for processing the input data set are euclidian and threshold modules. The threshold module helps the user to process data in the desired format as simulated annealing requires. The euclidian module helps to calculate the cost function between nodes. It is frequently used in optimisation problems in which distance only need to compare. It is also referred to as quadrance within the field of rational trigonometry [26].

EXPERIMENTAL WORK

The experimental work is a prediction approach for optimal development of an area for sustainable development. The icon for sustainable city mainly involves all aspects of industry which include both hard and soft constrain.

Prediction Approach

The prediction methods used for analyzing the spatial data are weighted sum method [2] and tchebycheff approach [4]. The multi objective function is broken down into similar single objective function, which is solved using the above approaches [4]. This helps to reduce the computation and produces the result much faster than any other approaches. The basic assumption made in the approach is that the weights assigned to each constrain should be known to the user prior.

The above weights assigned should be in a [0, 1]. Let w_i be the weight of the i^{th} objective function. It is calculated by the below method.

$$wi = {}^{ri}/\sum_{k=1}^{n} \mathrm{rk} \quad (1)$$

Here r_i is the weight assigned by the user to the ith objective constrain. The multi objective function is computed by adding up all the single objective functions together.

$$F(x)_{wsa} = w_1.f(x_1) + w_2.f(x_2) + - - - + w_n.f(x_n) \quad (2)$$

Where $f_i(x_i)$ is the i^{th} objective function and w_i is the weight assigned to the i^{th} objective function. In tchebycheff approach [4] the entire search space should be known to the user. After evaluating the threshold values the user will plot the pareto frontier based on a reference place. This helps to reduce the non-convergence problem upto some extent. The advantage of this approach is that the time taken for the above preprocessed data to converge is much faster than the weighted sum approach method [2]. The generalized formula of the approach is as follows

$$F(x)_{te} = \max\{\lambda_i\ (f_i(x) - z^*_i\} \quad (3)$$

$1 \le i \le m$

Where $z* = z^*_1, z^*_n$ is the reference point of n objective functions in the multi-objective problem.

The above two approaches will preprocess the spatial data. Thus preprocessed data will be processed using soft computing techniques like ant based optimisation [21] and simulated annealing [24] for predicating the optimal cost.

Soft Computing Techniques

The two main approaches used for the soft computing are ant base optimisation technique [21] and simulated annealing [24]. The main reason for using ant algorithm is to reduce the search space there by find the optimal solution from the candidate places in an optimal manner. Ant algorithm uses mainly three data structures: the graph, Minimum Spanning Tree (MST) and ant. Graph for the input data set of nodes and the cost function associated with it, the minimal spanning tree for the partial output of the optimal path and the ant for storing the tabu list and pheromone value.

The algorithm mainly consists of three phases: exploration phase, construction phase and optimisation phase. Before exploration phase initialization phase will be done. The initial pheromone value is calculated based on the method as below.

init.phe = (4 * (maxweight - edgeweight) /3) (4)

In this phase all ants will be placed randomly at each node, based on probability value the ants will take the best route. During exploration phase all ants follows the same rules which helps to identify the edges that are potentially good for the optimal distances. The phase works on edges of the graph in the order of the decreasing pheromone level. The pheromone value should be inversely proportional to the cost function of edge. The construction phase helps to construct the low cost spanning tree.ie edges with high degree of pheromone values in the increasing order of the edge cost will be taken. The new pheromone after evaporation will be as follows.

$$e.phe = (1-\eta) * e.phe + nodevisit*init.phe \tag{5}$$

Where η the evaporation constant. A minimal spanning tree will be constructed where all the nodes should get connected with a low cost of edge function. The optimisation phase helps to reduce the edge cost further based on three major rules 1) the edge taken should not be in the tree 2)edge should have a lower cost value 3) the obtained result should not violate the degree of constrain set by the user (DCST).Based on the above three rules each edges will be get replaced until no further replacement is possible. Grid type of problems can be usually solved using the simulated annealing [24] approach. The approach is as follows

```
        Function simulated annealing ()
current_node = make_node (intial_stage (problem))
        for temp = 1 to α do
        T = schedule [temp]
if  T = 0
   return current_node
else
   next_node = a randomly selected successor of current
   λE = value[next_node] – value [current]
endif
if  λE>0
 current_node = next_node
else
  current_node = next_node only with a probability of e^λ
  (problem , schedule) returns a solution state
endif
        endfunction
```

Here input is the set of candidate places that get processed by the pre-processing approaches and schedule a mapping your time to temperature. Initial temperature is set to very high and gradually decreased to low temperature. The nodes are the candidate places of the search space. Temperature gradient factor for the approach is denoted by α.

RESULTS

The pre-processing approach helps to generate the candidate solution based on the spatial data set .Prior knowledge of weight to each constrain for the objective function is necessary in weight sum approach [2]. Tchebycheff approach [4] outperforms the weighted sum method in all cases. The below figure 4 shows the pareto frontier obtained by weighted sum approach [2] method.

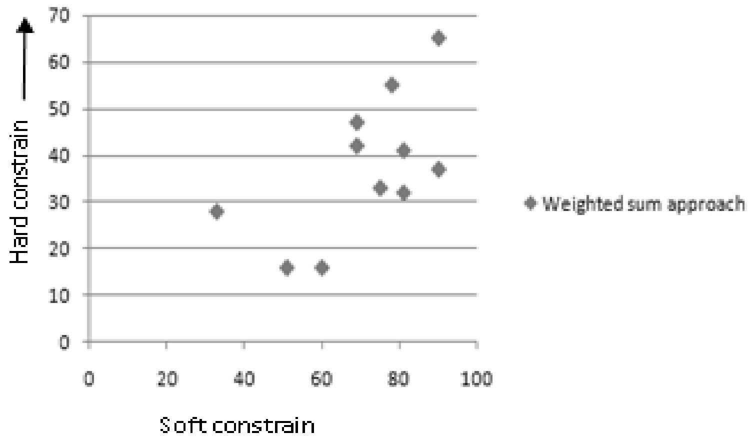

Figure 4: Pre Processing Data by Weighted Sum Method for Candidate Places

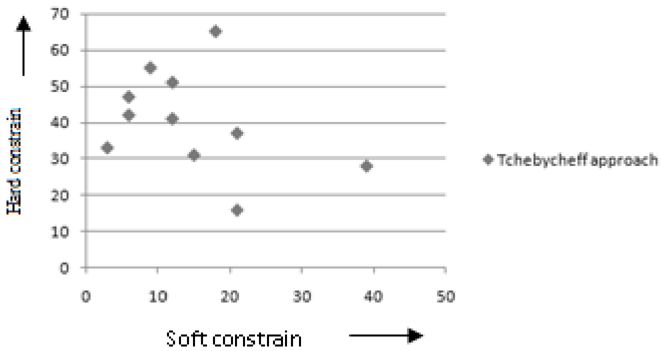

Figure 5: Pre Processing Data by Tchebycheff Method for Candidate Places

The above figure 5 shows the pareto frontier populated using the tchebycheff [4] method. The above two prediction graphs it's clear that the problem have the non-convergence solution. The tchebycheff approach outperforms the weighted sum method in the preprocessing approach.

In ant based approach will give the user better result if heuristic knowledge is used. If the initial pheromone value chosen between two places is a small positive constant [$0<\alpha<1$] and uniform throughout the edges the approach performance will be bad. [21] Heuristic knowledge of the cost function always works better than the incurious cost function on edge. The graph in figure 6 helps to compare the pheromone value on edges in ant colony optimisation. The result shows that the heuristic knowledge will outperform the incurious knowledge on edges. The optimal value remains constant throughout for 25 iterations which is taken as standard iteration number for the ant colony optimisation.

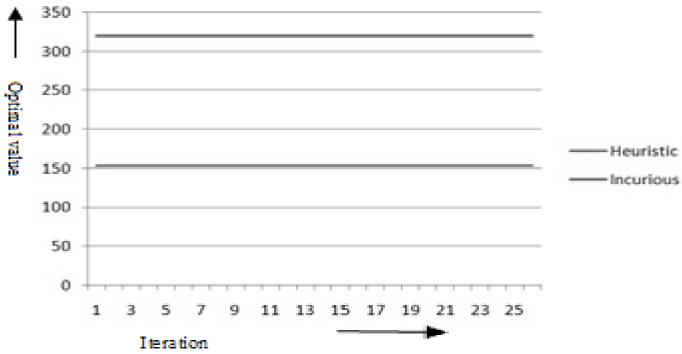

Figure 6: Graph of Ant Based Optimisation

The simulated annealing [24] works on the base of the thermodynamics. The temperature gradually gets decreased to low with respect to time function. The approach guarantees the best result for a graph type modeled problem. The objective functions for the approach is hard constrain and soft constrain. The below figure 7 shows the initial genotype configuration of candidate places in a simulated annealing approach [24].

Table of cities

#	City Name	Long. (deg)	Lat. (deg)
1	"Namakkal"	-14	31
2	"Chennai"	-38	29
3	"Ramanathap...	-19	38
4	"Coimbatore"	-3	42
5	"Erode"	-10	52
6	"Kanniyakum...	-19	16
7	"Thoothukudi"	-11	42
8	"Salem"	-3	48
9	"Vellor"	-9	55
10	"Thiruvallur"	-3	33
11	"Villupuram"	-18	65

Figure 7: Initial Genotype of Simulated Annealing

As the gradient of temperature gradually decreases the worst stages will get eliminated. The intermediate stages of the simulated annealing [24] is shown in the below figure 8 and figure 9.

Table of cities

#	City Name	Long. (deg)	Lat. (deg)
1	"Namakkal"	-14	31
2	"Chennai"	-38	29
3	"Kanniyakum...	-19	16
4	"Thiruvallur"	-3	33
5	"Coimbatore"	-3	42
6	"Salem"	-3	48
7	"Vellor"	-9	55
8	"Villupuram"	-18	65
9	"Erode"	-10	52
10	"Thoothukudi"	-11	42
11	"Ramanathap...	-19	38

Figure 8: Intermediate Genotype of Simulated Annealing

Table of cities

#	City Name	Long. (deg)	Lat. (deg)
1	"Namakkal"	-14	31
2	"Chennai"	-38	29
3	"Kanniyakum...	-19	16
4	"Thoothukudi"	-11	42
5	"Thiruvallur"	-3	33
6	"Coimbatore"	-3	42
7	"Salem"	-3	48
8	"Erode"	-10	52
9	"Vellor"	-9	55
10	"Villupuram"	-18	65
11	"Ramanathap...	-19	38

Figure 9: Intermediate Genotype of Simulated Annealing

COMPARISON

Tchebycheff [4] approach will out performs the weighted sum approach [2]. The non-convergence problem in the weighted sum approach [2] reduced to some extent. The weighted sum approach [2] is very simple in the mathematical model. The weights of each constrains need to be known by the user in prior. The computation efficiency of the approach is high and easy for pre-historic data problems. The ant based approach requires heuristic knowledge of the edge cost function. If the initial pheromone value is uniform and constant throughout all the edges of the graph the produced result will be not the best optimal result. The initial pheromone value should be less than one; it should be based on heuristic knowledge. It will outperform if the user update the pheromone value after each iteration, rather than a global update. The updating frequency of pheromone in global update done at a frequency of max_iteration/3. The evaporation rate (η) assumed as 1.5. For better result the Degree of Constrain should be always greater than two. The below comparison graph in Figure 10 shows that as the degree of constrain increases the optimal value obtained will be better than the heuristic knowledge of the edges. If the updating of the pheromone is globally done after each iterations it will outperform the Degree Constrain value. With global updating and increased degree of the tree the ant colony will guarantee best optimal solution.

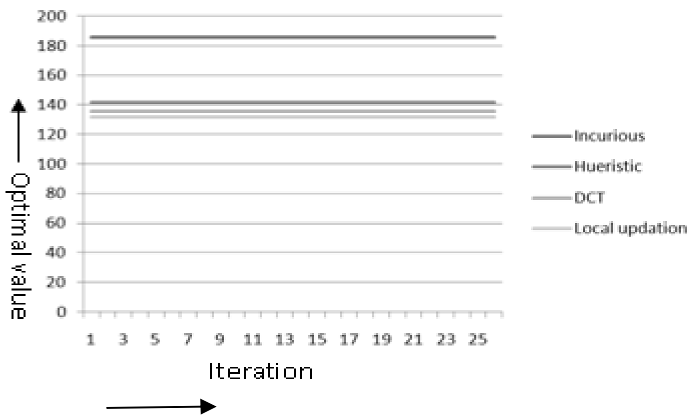

Figure 10: Comparison of Ant Based Optimisation

The accurate result of the optimal cost function is guaranteed by simulated annealing [24]. In this case also the tchebycheff approach [4] for pre-processing data outperforms the weighted method [2]. For the sample test data the optimal distance for the weighted sum method was 12,289.5 miles and that of the tchebycheff approach was 9,929.4 miles in the above soft computing method. Tchebycheff approach [4] will produce always times better result to that of weighted sum

CONCLUSIONS

In this work the prediction method help to find out the candidate places based on the constrain to set up the industry. The approaches used are weighted sum approach [2] which requires pre-historic knowledge of the weights that can be assigned to each constrains. In tchebycheff approach [4] threshold values of each place should be known to the user. This method helps to reduce the non-convergence problem of the weighted method [2]. The computational efficiency and simplicity of the weighted sum method makes it preferable. Thus populated places will be further evaluated and processed using the soft computing techniques for evaluating the optimal cost function among them. For ant based optimisation approach the prerequisite is that the problem should be modeled to a graph. The nodes are the candidate places based on the spatial data and the edges are the optimal cost distance among them. The cost function of the edge is obtained by finding the euclidian distance among the nodes [26]. The simulated annealing [24] approach guarantee the best result but the time complexity will be high compared to the ant based optimisation approach. Thus the work conducted helps to find the best optimal path for setting up smart cities[1] using ant based optimisation and simulated annealing approaches.

FUTURE ENCHANCEMENT

The exploration phase in ant based approach can be done by parallel method. This helps to reduce the time for the ant based optimisation approach [26] in the exploration phase of the approach. The other soft computing approaches can outperform the ant based approach and simulated annealing. Soft computing approach can be further improved in such a way that each candidate places can be selected based on the cluster value that need to be satisfied by it. It also guarantees best performance when compared to the other two approaches.

ACKNOWLEDGEMENTS

The authors thankfully acknowledge for the financial data set provided for the TamilNadu state for the experiment given by D K Srivastava and K R Shanmugam and Larry Engelhardt for their support in the simulated annealing approach.

REFERENCES

1. Ignasi Vilajosana, Jordi Llosa, Borja Martinez, Marc Domingo-Prieto and Albert Angles, “Bootstrapping Smart Cities through a Self-Sustainable Model Based on Big Data Flows”, IEEE Communications Magazine June 2013.
2. Ashish Ghosh and Satchidananda Dehuri, ”Evolutionary Algorithms for multi-Criterion Optimisation Survey”, International Journal of Computing & Information Sciences Vol.2,No.1,April 2004.
3. Daniele Perissin and Teng Wang, “Time-Series InSAR Applications Over Urban Areas in China“, IEEE journal of selected topics in applied earth observations and remote sensing, Vol. 4, No. 1, March 2011.
4. Liangjun Ke, Qingfu Zhang and Roberto Battiti, “MOEA/D-ACO: A Multiobjective Evolutionary Algorithm Using Decomposition and Ant Colony”, IEEE transactions on cybernetics, Vol. 43, No. 6, December 2013.
5. Zhengxing Huang, Xudong Lu, and Huilong Duan, “A Task Operation Model for Resource Allocation Optimization in Business Process Management”, IEEE transactions on systems, man, and cybernetics—part a: systems and humans, Vol. 42, No. 5, September 2012.
6. Thang N. Bui, Xianghua Deng, and Catherine M. Zrncic, “An Improved Ant-Based Algorithm for the Degree-Constrained Minimum Spanning Tree Problem” IEEE transactions on evolutionary computation, Vol. 16, No. 2, April 2012.
7. Yoon-teck bau, Chin-kuan ho and Hong-tat ewe,” Ant Colony Optimization Approaches to the Degree-constrained Minimum Spanning Tree Problem”, journal of information science and engineering 24, 1081-1094 (2008).
8. Luis M Correia, Klaus Wünstel and Alcatel-Lucent, “Smart Cities Applications and Requirements”, White Paper on European Technology Platform, May 2011.
9. Hafedh Chourabi, Taewoo Nam and Shawn Walker, “Understanding Smart Cities: An Integrative Framework”, 45th Hawaii International Conference on System Sciences, May 2012.
10. Jesse M Shapiro, "Smart cities: quality of life, productivity and the growth effects of human capital“, NBER paper series, September 2005.

11. Rafael Achaerandio and Roberta Bigliani, "Smart Cities Analysis in Spain 2012 The Smart journey“, White paper Sponsored by: Atos, BBVA, Ferrovial Servicios, Microsoft, and Urbiotica, September 2012.
12. Jose M. Hernandez, "smart cities, the silent IoT revolution" Telefonica Investigation, S.A. Unipersonal 2010, December 2010.
13. Charbel Aoun, "The Smart City Cornerstone: Urban Efficiency “, Schneider Electric White Paper, January 2013.
14. Hans Schaffers, Nicos Komninos and Marc Pallot, "Smart Cities as Innovation Ecosystems sustained by the Future Internet", FIREBALL White paper, April 2012.
15. K.C Tan, T.H.Lee and E.F.Khor, "Evolutionary Algorithms for Multi-Objective Optimisation: Performance Assessments and Comparisons", Proceedings of the 2001 IEEE Congress on Evolutionary Computation Seoul, Korea (0-7803-6657-3/01), May 2001.
16. Raymond Chiong and Ooi Koon Beng, "A comparison between Genetic Algorithms and Evolutionary Programming based on Cutting Stock Problem", Engineering Letters, 14:1, EL_14_1_14 (Advance online publication), February 2007.
17. Dilip Datta, Kalyanmoy Deb,Carlos M. Fonseca, Fernando G. Lobo and Paulo A. Condado, "Multi-objective Evolutionary Algorithm for Land-Use Management Problem", ISSN 0973-1873 Vol.3, No.4 (2007), pp. 371–384.
18. V.K Varriavamoorthy, S D Gorantiwar, A Pathirana, “Managing urban water supplies in developing countries –climate change and water scarcity scenarios”, Elsevier Physics and chemistry of earth 33 (2008)330-339, pp 330-331.
19. Last, Ewan, Mackey, Rae, “Developing a new scope model for urban sustainability”, 2007.
20. Qiang, He. Li Zhaijun, Huang, “Application of sustainable water system the demonstration in Chengdu(China)”, 2006.
21. Shu-Chuanchu, John F Roddick and Jeng-Shyang Pan, “Ant Colony system with communication stageries”, Information Sciences 167 63-76 (2004).
22. Aaarst, “Simulated Annealing and Boltzmann”, Wiley 1989.
23. Otten, “The annealing algorithm”, Kluwer Academic publishers, 1989.
24. http://en.wikipedia.org/wiki/Simulated_annealing
25. En.wikipedia.org/wiki/Traveling_salesman_problem
26. En.wikipedia.org/wiki
27. Gis.stackoverflow.com
28. www.java-forums.org
29. Resources.arcgis.com
30. Geoadmin.com