A Comparative Study of Various Combinatorial Approaches for Minimization of Transmission and Distribution Losses

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering
in
Computer Science and Engineering

Submitted By
Jatin
(Roll No. 801332012)

Under the supervision of

Dr. Shalini Batra
Assistant Professor
(CSED)

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR UNIVERSITY
PATIALA – 147004
July 2015

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, ***"A Comparative Study of Various Combinatorial Approaches for Minimization of Transmission and Distribution Losses"***,in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Computer Science and Engineering* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Dr. Shalini Batra* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

15/7/15
Date

(Jatin)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

16/7/15
Date

Dr. Shalini Batra
Assistant Professor
CSED
Thapar University, Patiala (Punjab)

Countersigned by

(Dr. Deepak Garg)
Associate Professor and Head
CSED
Thapar University
Patiala (Punjab)

(Dr. S. S. Bhatia)
Dean of Academic Affairs
Thapar University
Patiala (Punjab)

ACKNOWLEDGEMENT

First of all, I am extremely thankful to my respective guide ***Dr. Shalini Batra***. Assistant professor, CSED, Thapar University for his valuable guidance, advice, motivation, encouragement, moral support, sincere effort and positive attitude with which she solved my queries and provide delightful ambiance for learning, exploring and making this thesis possible. It has been a great pleaser and experience to work under his sanctuary.

I am also heartily thankful to ***Dr. Deepak Garg***, Associate Professor and Head, CSED for motivation. He sets high standards for his students and he encourages and guides them to meet those standards.

I would like to thank my family members and my friends who are dearest and precious to me for their love, encouragement, blessings and support in all respects. Most importantly, none of this would have been possible without the love and patience of my family. They are constant source of love, concern, support and strength for me all these years.

Finally I would like to thank management of Thapar University for proving a great platform for learning, not just for academics but also for music and many other creative things.

Above all, I express my thanks to the "ALMIGHTY" for blessings and kindness.

JATIN

ABSTRACT

The percentage loss of electric energy due to transmission and distribution(T&D) is notable. As the urbanization is expanding, connectivity is becoming more complex and T&D losses are further increasing. These T&D losses can be minimized significantly by optimizing the connectivity of electricity lines.

The major objective of this work is to optimize the placement of transformed in residential areas to minimize the transmission losses. To achieve the desired goal three approaches have been considered: deterministic, clustering and stochastic. In deterministic approach, brute force method is applied, in clustering K-means unsupervised clustering is applied and in stochastic approach stimulated annealing is used. A comparative analysis of all three aforesaid approaches has been done on real time data collected from PSPCL, to calculate the percentage reduction in transmission and distribution (T&D) loss by applying aforesaid approaches.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

One of the greatest marvels of science is the invention of electricity, one of the most valuable present of science which has proved to be a bonanza to humanity. It can be said without any doubt that, this gift of electricity is not a luxury but an essential commodity for the entire community as without it comfort in life is unthinkable. There is not even a single domain of life in which electricity does not leave any impact.

1.1 Electricity Delivery

The process that goes from the generation of the electricity from the different power plants to the end consumer is called electricity delivery. Which includes following four steps :-

- Generation
- Transmission
- Distribution
- Retailing

1.1.1 Generation:- The process of generating electric power from different sources of primary energy like coal, petroleum, water, is called electricity generation process. This is one of the biggest market space. The basic methods used for the generation of the electricity by major electricity production companies include movement of a loop of wires or disc of metal in between the poles of a magnet/electromagnet.

1.1.2 Transmission:- The bulk transfer of electricity from electric power plant to the different electric power substation located near the end consumer is called electricity transmission. This is different from local connectivity between high voltage subdivision and consumer. This is mainly between towns or locations which are significantly apart from each other. When different towns are interconnected with each other it becomes transmission network. This combined network is formally known as power grid, or the grid. For example in UK this combined network is called National grid.

1.1.3 Distribution:- Distribution phase is an important stage in the process of electricity delivery, where transfer of electricity takes place from electric substations to individual consumers. The connection between electric sub-station and end users is medium transmission voltage to lower voltage which broadly ranges from 2kV to 35kV with the usage of different capacity transformers. These transformers lower the voltage to the optimum utilization voltage for household appliances.

1.1.4 Retailing:- This is the final step in electricity delivery. The amount of electric energy used by domestic consumers is calculated by electric meter and as per reading, the costumer is charged accordingly. In most of the countries customers are charged monthly or bimonthly basis and sometimes these charges depends on rate of consumption, peak rate of consumption, maximum demand in kVA(Kilo volt amps) *etc*.

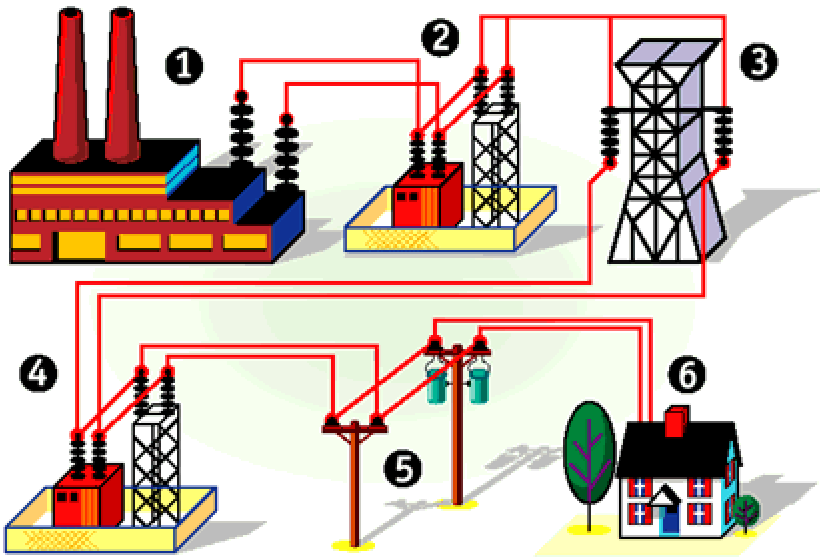


Figure 1.1 Electricity delivery system

In figure 1.1, step 1 shows the power generation, step 2 the earliest voltage step down, followed by transmission in step 3 and step 4 and step 5 and step 6 show distribution of electricity. The main emphasis of this thesis is on Transmission and Distribution optimization.

1.2 Transmission and Distribution Loss

It is a well known fact that the power generated from the large power stations is transferred through huge and complex networks like transformers, high voltage lines, cables and other complex equipments ultimately reaches to the end consumers through wires and cables laid down by electricity department. However, there is a huge difference between the amount of electric power generated from the power plant and units distributed to the end consumers and this difference is known as Transmission and distribution loss (T&D loss). Although T &D losses are quite huge, they are not paid by the consumers as the end users have actually not used that electricity.[12]

T&D Losses= (Energy Input to feeder(Kwh)-Billed Energy to Consumer(Kwh)) / Energy Input kwh x100

Transmission and distribution sectors are considered as the weakest link in the entire electric power sector. The loss due to distribution is much higher than transmission but one cannot neglect the loss due to transmission. This loss due to transmission and distribution affects not just the electric company's finances but the whole financial system of the nation. So this must be taken into consideration with utmost attention.

T&D loss can be broadly categorized into two types:-
1. Technical losses.
2. Non-Technical losses (commercial losses)

1.2.1 Technical losses

Technical losses are caused by current flowing in the electrical network. These losses are caused by the physical properties of different equipments used in the power system. The power abandoned in transmission cables and transformers due to internal electrical resistance is one of the most obvious reasons for the technical loss. In fact most of the technical losses are naturally occurring losses and consist mainly of dissipation in electrical system component such as transmission cables, transformers, measurement tools, *etc*. One can easily compute technical losses and further control technical losses by using latest less resistance alloy wires and devices instead of heavy obsolete tools. Some of the major technical losses are:-

1. Copper losses, (Inherited loss in T&D due to wear and tear).
2. Improper earthquakes.
3. Unbalanced loading.
4. Use of poor standards equipments.
5. Long one phase line.
6. Low voltage and loss due to overloading.

7. Loss due to Harmonics distortion. (Humming).
8. Non-optimized connectivity.

1.2.2 Non-Technical losses

These losses in the power system are caused by external actions such as theft, inaccuracy in the meters, poor wiring, bypassing the meter, *etc*. These non-technical losses are more difficult to measure. The most probable reasons of non technical losses are :-

1. Meter tempering, resulting in recorded of lower consumption reading.
2. Errors in technical loss computation.
3. Tapping (hooking) on LT (low tension) lines.
4. False readings by bribing meter readers.
5. Illegal connections.
6. Ignoring unpaid bills.
7. Errors and delay in meter reading.
8. Non-optimized connectivity.

It can be generalized that all technical losses can be reduced by installing better equipments and non-technical losses can be minimized by improving system transparency and by removing corruption from the system.
But one of the major reasons for the non-technical cum technical loss is non-optimized connectivity and load balancing in transmission and distribution[13].

1.3 Losses evaluation

In the Figure 1.2, the percentage T&D loss from 1976 to 2008 in India is shown. It can be concluded from the data that, T&D loss is in non decreasing fashion.

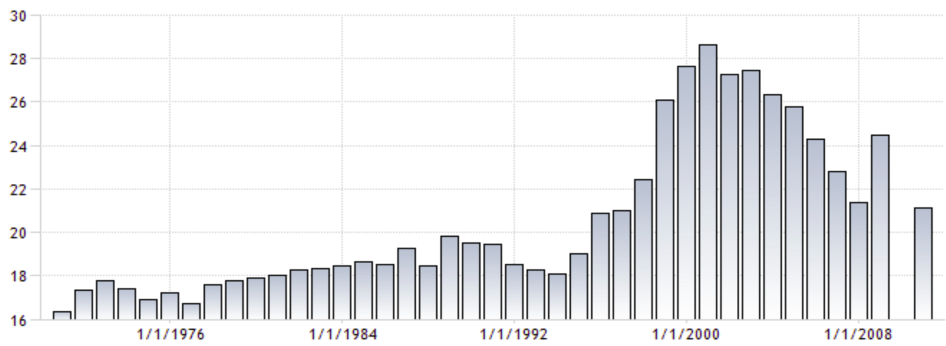


Figure 1.2 %T&D loss in India (From 1976 to 2008)

OVERALL T&D LOSSES (%)					
State	**2007-08**	**2008-09**	**2009-10**	**2010-11**	**2011-12**
Andhra Pradesh	20.3	19.2	18.1	16.1	15.3
Arunachal Pradesh	43.7	48.0	39.1	35.6	34.5
Assam	28.0	29.6	34.8	29.9	27.7
Bihar	39.1	38.0	38.3	37.0	35.0
Chhattisgarh	31.0	28.6	38.7	34.7	32.7
Goa	16.7	21.0	16.6	17.4	17.6
Gujarat	23.8	22.8	24.5	22.7	22.3
Haryana	28.1	25.7	26.8	24.4	22.7
Himachal Pradesh	13.5	13.2	14.7	14.6	14.5
Jammu & Kashmir	61.9	61.3	63.0	60.0	58.5
Jharkhand	42.3	43.0	38.5	33.5	40.8
Karnataka	25.3	23.3	21.4	20.1	19.6
Kerala	19.9	19.9	19.2	19.1	18.6
Madhya Pradesh	40.1	39.0	35.6	34.1	32.6
Maharashtra	29.1	26.5	25.2	22.5	21.6
Manipur	48.4	51.1	45.8	43.3	38.0
Meghalaya	33.4	31.2	34.0	30.0	28.4
Mizoram	24.9	32.6	37.0	35.4	34.3
Nagaland	36.4	31.0	36.5	30.8	28.1
Puducherry	13.8	13.7	13.5	13.5	13.5
Punjab	21.5	18.5	19.7	17.8	16.8
Rajasthan	35.5	31.9	29.9	27.6	24.8
Sikkim	22.3	34.0	40.6	42.4	38.8
Tamil Nadu	18.0	18.0	18.0	18.0	17.0
Tripura	23.4	24.1	24.7	20.9	20.1
Uttar Pradesh	32.6	28.6	32.3	28.9	24.4
Uttarakhand	29.7	28.0	24.5	22.5	20.5
West Bengal	24.3	23.3	23.8	23.5	22.3
Source: Planning Commission					

Table 1.1 : State wise T&D loss in India(In percentage)

In Table 1.1, State wise data is collected from the planning commission of India [2], which clearly depicts that on an average there is 20 to 40 % loss due to transmission and distribution in every state. These are official figures but actual losses may go above 45% (approx.)[15]

From the above data one can conclude T&D loss is huge and there is an immediate requirement to reduce, as it directly impact the nation's GDP.

1.4 Combinatorial Optimization

The process of finding an optimal object or solution from a pool of finite set of objects or solutions is called combinatorial optimization. In such problem, tracing of whole of the search space is not possible. The domain of these optimization problems in which it operates are: where the pool of feasible solutions is discrete or can be reduced to discrete and the objective is to find the most optimal solution. Some of the common problems under combinatorial optimization are TSP (Traveling Salesman Problem), MST (Minimum Spanning Tree problem) *etc*.

Mathematical optimization is a superset of combinatorial optimization which is related to computational complexity theory, algorithm theory and operations research. It has got some of the most important applications in several fields like machine learning, auction theory, software engineering and mathematics. According to research literature, integer programming is consistent with discrete optimization blended with combinatorial optimization, which is itself composed of optimization problems dealing with graph structures. It generally includes determining the way to effectively and efficiently allocate resources used to find optimal solutions to mathematical problems.

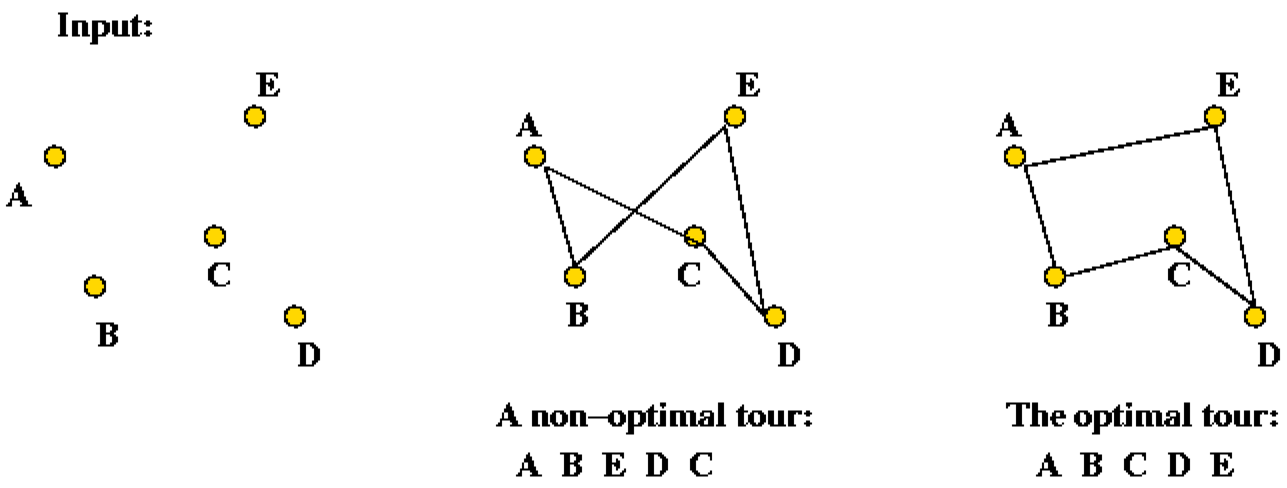


Figure 1.3: Example of combinatorial optimization.

Figure 1.3 depicts the problem in which all vertices need to be connected in optimal way with the minimal cost and by applying combinatorial optimization approach we obtain optimal result in the end.

1.5 Loss Function

In statistics, mathematical optimization, machine learning, and decision theory, a cost function or loss function is a function that relates an event or values of one or more variables onto a real number representing some "cost" related with the event. An optimization problem tries to minimize a cost function. An objective function is either a loss function or its inverse (often called a profit function, a reward function, a service function, *etc*.), which needs to be maximized.[14]

In statistics, typically a cost function is used for estimating the parameters, and the action in question is some function of the difference between true and estimated values for an instance of data. In economics scenario, for example, this is usually economic cost or loss. In optimal control, the loss is the penalty for failing to achieve a desired value. In the context of financial risk management, the function is precisely mapped to a monetary loss.

1.6 Structure of the Thesis

The rest of the thesis is organized in the following order:

Chapter 2 - Provides literature review of existing combinatorial optimization approaches.

Chapter 3 – Gives a detailed introduction about electricity transmission and distribution and its losses and problem formulation.

Chapter 4 - This chapter defines proposed approach and methodology used to solve it.

Chapter5 – Provides the implementation details and results and their comparisons.

Chapter 6 - Conclusion of thesis and suggestions for future work has been incorporated in this chapter.

Thesis concludes with references and dairy.

CHAPTER 2

LITERATURE REVIEW

As the official figures indicates, in India, on an average, 23 percent of electricity generated losses in T&D However as per studies conducted by independent agencies like TERI (The Energy and Resources Institute), it has been estimated that, transmission and distribution losses are as high as 50% in some states of India [1]. In another recent study conducted by State Bank of India, capital markets for DVB (Digital Video Broadcasting), T&D losses are estimated around 58%. So the accuracy level for the T&D loss gained much more importance, as it directly impact the cost and as a result GDP(Gross domestic product) of the nation[1]. There are multiple reasons, commercial and technical, due to which energy losses occur in the process of electricity delivery. Some of the reasons for the technical losses are energy dissipated in the conductors and other equipment used for transmission, sub transmission, energy stepping down equipments, *etc*. These technical losses are mostly inherited in the system due to already installed equipments which can be reduced to a minimum optimum level. Theses technical losses can be further classified depending upon the stage of power transmission and transformation system, like (400kV/220kV/132kV/66kV). On the other side, commercial losses are caused by pilferage, errors in meters readings, corrupted meters and wrongly estimated non-metered supply.

Mungkung *et al.* proposed a technique for technical loss analysis in Hatyai, Provincial Electricity Authority (PEA). There analysis used PSS/Adept (developed by Siemens Power Transmission & Distribution, Inc. Power Technologies International) program [2]. In this research, the technical loss in distribution system have been considered on the basis of low-voltage transformer power losses, distribution line losses, power transmission losses and transmission line losses [2], and this analysis used calculation and PSS/Adept program. They concluded that low-voltage transformer and distribution line losses can be decreased by installing better equipment for low voltage conductor, load balancing on transformer and installation of low voltage capacitor.

In today's era some countries like India faces peaking electrical energy shortages. These shortages have a very harmful effect on the overall GDP (gross domestic product) and economy of the country. The reasons cited by Navani *et al.* [3] in their research for these high losses are: lack of optimal T&D network, higher numbers transformation stages, inappropriate load distribution and poor rural electrification, *etc*. It has been concluded from their research that the sources of technical losses may be directly pushed by network investment or by network operation . The proposed technique analyzed and evaluated technical and non technical losses with the help of a case study and MATLAB simulation of power systems. The measurement of Non technical losses (NTL) and its impact on electrical power systems as a whole using already developed analytical tools is possible only if information about the Non technical losses themselves is available to the analyst. Accurately estimating losses in distribution systems is becoming significantly important. Thus this need is particularly important in developing countries, where total losses are generally high without the incorporation of the private sector.

The research also concluded that mostly technical losses are a result of inefficiency of equipments used and the size of both lines and equipments. The main important contributors are losses through resistance, excitation losses of transistors and losses due to leakage. In the AC systems, due to skin effect, the coppers losses seem to be very high. Due to the skin effect the value of the flux density is high at the centre of the conductor and hence the flow of current is also high. Due to this reason the skin effect increases the resistance and power loss. The losses of the transformer are due to the copper loss which is due to the inner impedance factor of transformer coils. Because the power systems are connected through the power transformers so, load losses which are functions of core material, lamination *etc* are not considered.

The problem is that it is precisely in these situations where data required for accurately estimating the total losses and particularly their breakdown into technical and nontechnical components are generally lacking. The information which should be include is the NTL load’s power consumption profile comparable to the legitimate loads being analyzed as well as the NTL power factor, or power factor contribution at the same time. In this paper, two bus systems have been taken with one bus as slack bus and load is on another bus. The load profiles of simple industrial area and residential area has been taken. Then a small percentage of NTL has been added to one of the load and the increased load and losses have been shown with the help of Newton-Raphson load flow method using Matlab. The power factor contributions chosen here are negative because the NTL load is assumed to be inductive. The total units supplied and total units billed have been thoroughly measured for one full month. Then their difference is used to determine the extent of non technical losses in that area.

In the history of combinatorial optimization, integer programming formed the pivot. Its initial concept was conceived by Kantorovich and Koopmans.[3] After the introduction of integer programming as generic problem and the development of the simplex method as an aid, they tried to attack almost all of the combinatorial optimization problems with integer programming techniques, which were quite often very successfully done [3].

One of the important reasons of the diversity of combinatorial optimization is that many of its problems plunge directly from practice and instances of them are attacked daily. One can think that even in very primitive communities like finding the shortest paths and searching (for instance, for hospital) is vital [1]. A traveling salesman problem comes into picture when one plans outing or sightseeing, or when a traveler or postman plans his/her tour. Similarly, assigning jobs to persons, transporting goodies, and making relations, form fundamental problems not just considered by the statisticians or mathematician.

Wayne Hong and Shirmohanunadi proposed the effective method for the reconfiguration of the distributed networks to decrease their losses due to resistive lines. This approach is very efficient and it can be easily used in planning and operative environments. The proposed system has been successfully implemented on the various distributed networks. In case of the operative studies, it is essential to fix the number of switching operations that are suitable for network configuration [4].

Quezada *et al.* analyze the impact of distribution generation on the distribution losses. They studied the different distribution generation technologies, their penetration and concentration levels. Along with various distribution networks along with topologies and patterns They discussed the impact of various parameters and concluded that the annual energy losses are function of penetration levels [5].

Two approaches are used to solve the optimization problems:

- Deterministic Approach
- Stochastic Approach

Deterministic approaches consider the integer program and algorithms for example, brute force approach. While Stochastic approaches considers the simulate annealing, genetic algorithms *etc.* for optimization. Simulate annealing is a powerful method for stochastic optimization with a deep relationship with Markov Chain Monte Carlo (MCMC). For choosing the accepting probabilities, it considers the annealing metal.[21]

Kirkpatrick *et al.* (1983) proposed an approach to simulate annealing as a combinational optimization problem. Four elements are required [6]:

I. Brief details about the configuration of the system.

II. In a configuration there should be rearrangement of elements.

III. There should be objective function which consists of trade-off.

IV. Annealing schedule for the length or times the system is evolved and value of temperature.

The annealing schedule can be developed by error and trail or by warming a system to a high rate and then slowly cooling down the system until it diffuses.
Simulate annealing is an extension of two techniques :

- Divide and Conquer technique
- Metropolis Algorithm

On the basis of temperature the various classes of the rearrangement are distinguish. So, due to these rearrangements, large changes occur in the objective function at high values of temperatures and at the low temperature values. This phenomenon of simulate annealing has been adaptive from divide and conquer.[23]
The Metropolis Algorithm has been extended in some small steps by using one configuration to other. According to previous studies it is concluded that better quality results can be obtained by using annealing schedules.

CHAPTER 3
PROBLEM FORMULATION

3.1 Problem formulation

The impact of connectivity is in any area is crucial. The way different components of any system are connected has a direct impact on the overall performance, whether it is connectivity of different components within a computer (low scale connectivity) or over the web(large scale connectivity).

Due to poor connectivity between different components, from power plant to the end user, there is a huge loss of electric energy. This loss can be reduced drastically to a minimum value if there is better, shorter and efficient connectivity.

This thesis proposes three approaches which takes current nodes which are supposed to be connected and suggest an optimal way of connectivity with minimum cost.

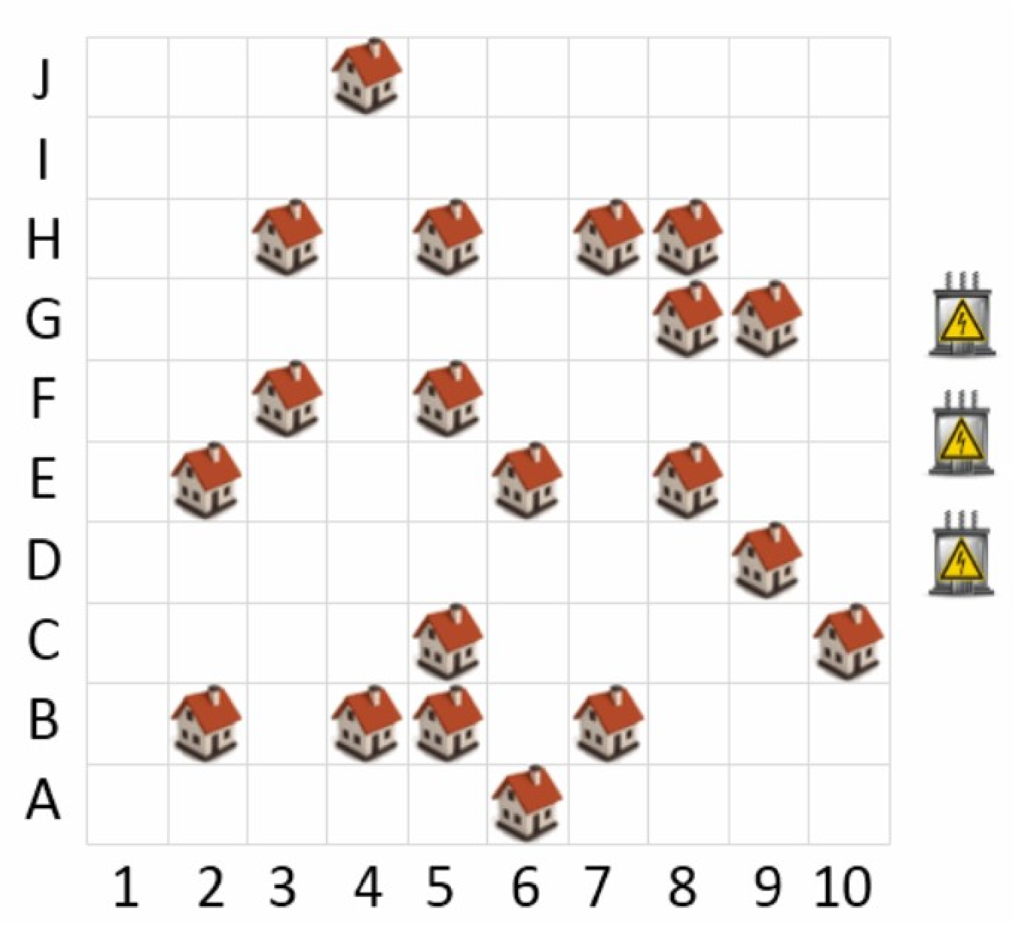


Figure 3.1 . Problem animation

The problem statement is shown above, graphically. As depicted in figure 3.1 to optimize the placement of the power plants, is to be done.

3.2 Research Gaps

Literature review shows that only mathematical optimization modeling techniques was carried out, which is insufficient to give actual picture of optimizing T&D loss, as it is based upon several assumptions. So, stochastic approach needs to be considered to optimize the T&D losses.

3.3 Objectives

- To study and analyze various approaches used for optimization of T&D losses

- To optimize the T&D losses though clustering, stochastic and deterministic approaches by applying all three approaches on the actual data.

- To compare and contrast results of clustering, stochastic and deterministic approach to find out which is the best among the three.

CHAPTER 4

PROPOSED APPROACH

4.1 Approaches

There are various combinatorial approaches available to solve the problem, and in this thesis work, three approaches have been studied and applied.

1. Deterministic approach: brute force
2. Curve-ball solution: using K-means clustering
3. Stochastic approach: simulated annealing

4.1.1 Deterministic approach: Brute Force

This is the most trivial and straightforward way to solve any problem. This approach takes all the possible solution space into consideration, calculates them and finds the optimal solution. This approach yield quick results only if the problem size is small, in which it possible to trace all the feasible solutions within nominal time. As the problem size increases, response time also increases, and in the most of the cases it grows exponentially.[10]

This approach is quite simple to be followed. For example let's consider $[N][N]$matrix and M number of power plants which is to be placed in the matrix. So there are total $C(N * N; M)$ (Combinations) possible ways to place M numbers of power plants. This deterministic approach will trace all the possible ways to calculate the optimal way to place them but as the number of locations increase it becomes almost impossible to solve the problem with this approach.

4.1.1.1 Execution flow of Deterministic approach

In Figure 4.1, The execution flow of the brute force technique is shown. Step by step execution is explained as below.

- Matrix size, number of transformers and locations of neighbors are fed as input.
- By applying basic combinatorial on matrix size and number of transformers, whole solution

space is generated.

- Each configuration will be picked from the solution space and associated cost will be calculated and stored for further process.
- After all the iterations, the minimum cost's configuration will be chosen as the optimal solution.

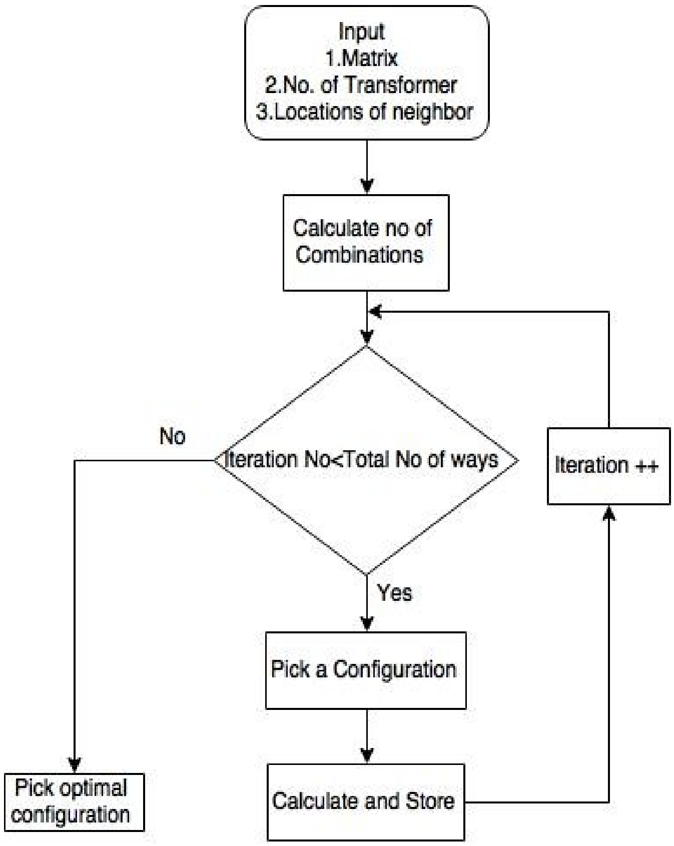


Figure 4.1 Execution flow of deterministic approach

4.1.2 Curve-ball solution: using K-means clustering

Another approach to solve the above mentioned problem is creating clusters of stationary house. K-means algorithm includes arbitrarily selecting K initial centroids where K is provided by the user, equal to number of required clusters. Each point is then assigned to a closest centroid (K) and the collection of points close to a centroid form a cluster.

With this approach, different stationary houses forms a cluster and the power plant will be placed at the centroid of that cluster, leading to optimal connectivity. This approach is better than previous brute force technique as it can handle much higher number of inputs than brute force.

4.1.2.1 Execution flow of Curve-ball solution

In Figure 4.2, The execution flow of the brute force technique is shown. Step by step execution is explained as below.

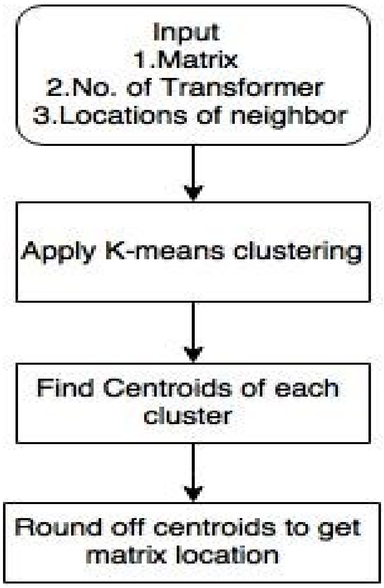


Figure 4.2 Execution flow of Curve-ball solution

- Matrix size, Number of transformers and locations of neighbors are fed as input.
- Apply K-means clustering, by taking k as number of transformers.
- Calculate centroid of each cluster.
- Round off the coordinates of centroids to the nearest matrix cell location.

Using the above method transformers locations is calculated. Once the locations of neighbors and transformers are available, optimal cost can be calculated.

4.1.3 Stochastic approach: Simulated Annealing

Simulated annealing a powerful stochastic optimization method, is a generic probabilistic meta-heuristic for the global optimization problem of locating a good approximation to the global optimum of a given function in a large search space. Steps followed to solve this problem using simulated annealing are:

- Some candidate solution is selected and variable is initialized to track the best solution encountered.

- A new candidate solution is then depicted, by altering the current candidate through some means, at each iteration state.
- If the solution proposed is better then the proposed solution is used for solving the current proposal. If strict hill climbing (a mathematical optimization technique) is adopted the, risk of getting stuck in local optimum is high aggravated.
- Probability of going towards worst proposal is in direct and positive relation with the current temperature.
- At initial level, high temperature is made to explore widely, but afterwards temperature is gradually decayed, so that good focus can be made to optimize locally.
- Re-annealing is adopted temporarily, which in turns forms flexible metals.

The closeness of optimal solutions suggested by all three approaches depends on the matrix size. As the matrix size will increase, the number of cells will increase, as a result there will be less need to do round off operations.[10] So, more closer results will be generated by all three approaches, applied on same set of data.

The implementation and the comparison between all of them is provided in the following chapter.

CHAPTER 5

IMPLEMENTATION

5.1 Modeling the data

First and foremost step to solve any problem with computer aided tools, is to map the actual data (images, maps, *etc.*) into computer understandable form. To do so, Python libraries like matplotlib, seaborn and numpy are used for modeling the actual data into matrix form. It is always better to map the data into matrix form if possible, because there is a wide scope to manipulate the matrix data, and moreover, information is projected in a better and easily understandable way. Also there are wide range of matrix functions are available in modern programming languages to do operations with on it.

Python's 2D plotting library matplotlib, visualization library seaborn and the package for scientific computing, numpy are used to model the real data into matrix form. Plotting on the matrix is done manually, by providing cell addresses on the matrix.

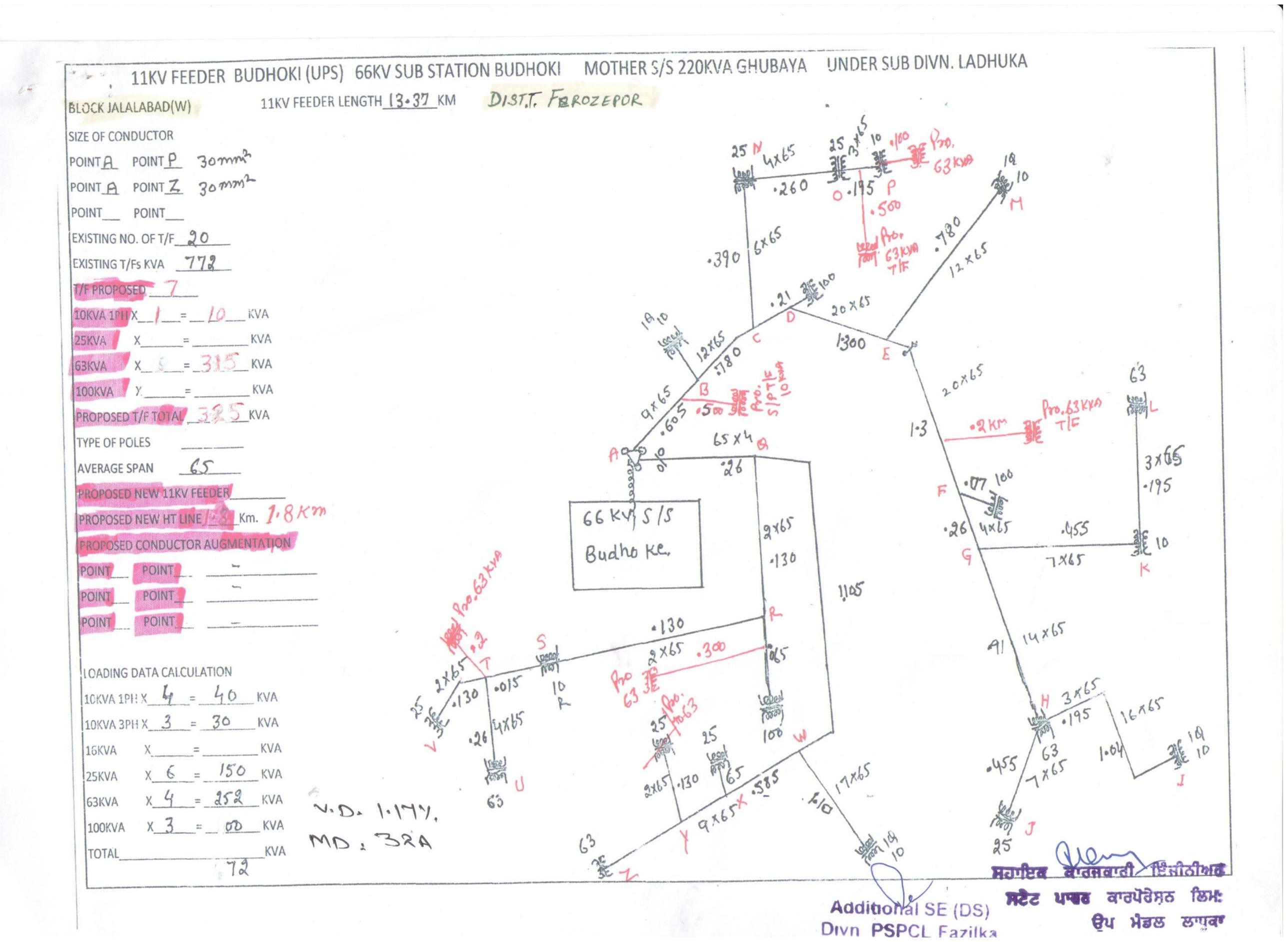


Figure 5.1 official data from PSPCL

Figure 5.1 is an official map of connectivity of ladhuka substation, Fazilka (Punjab), collected from Punjab State Power Corporation limited (PSPCL). In PSPCL almost all of the connectivity lines are documented in this form only. So at ground level, the actual problem is much more complex than it seems. The conversion of this data into matrix from is a very crucial step and in the thesis manual conversion has been done.

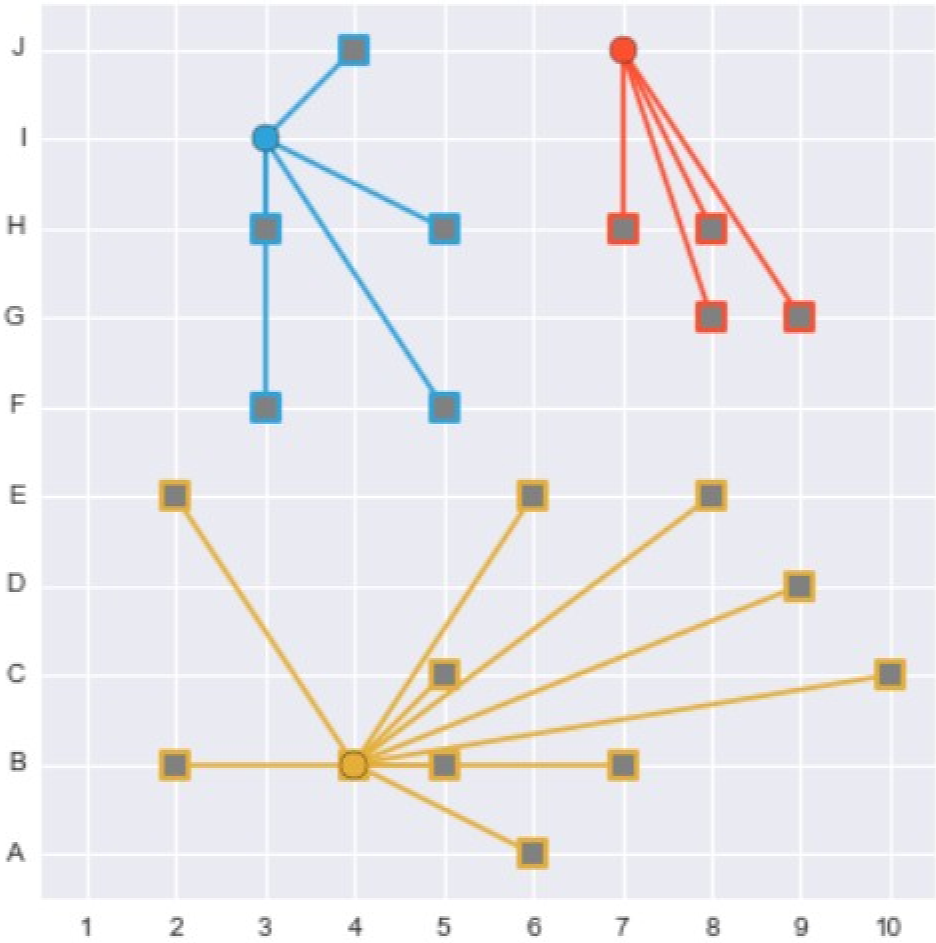


Figure 5.2 Data to Matrix Conversion

In figure 5.2, connectivity is modeled in matrix form in which circle shows the placement of transformers and square shows the placement of end consumer. The coordinates of transformers and consumer are provided manually, just to display the connectivity.

5.2 Functions Implemented

At the core of the solution of this problem, there are four functions which are implemented in all three approaches iteratively.

5.2.1 Abbreviations

Before elaborating the functions, some of the abbreviations used in functions are:-

S = An array-like point (x, y) for a single transformer

SS = An array-like of multiple (x, y) transformers placements

N= an array of neighborhood placements

5.2.2 List of Functions

1. distance_from_transformer(S, N)

2. minimum_distance_from_transformers(SS,N)

3. cost_of_transformers(SS,N)

4. Total_cost(SS,N)

5.2.2.1 distance_from_transformer(S, N)

This function takes two arguments, coordinates of a transformer and an array of neighbors. This function calculates the distance of neighbors from a specific transformer (whose co-ordinates are provided manually), and returns an array of all the distances.

5.2.2.2 minimum_distance_from_transformer(SS,N)

By using distance_from_transformer(S,N) function, it takes co-ordinates of a list of transformer and a list of neighbors, and calculates the distances between them and gives the minimum distance of each neighbor from all power plants.

5.2.2.3 cost_of_transformer(SS,N)

This function calculates the total cost of each transformer, by using distance_from_transformer function and minimum_distance_from_transformer function and by taking list of transformer and a list of neighbors as arguments. It returns cost or loss of individual transformer in the form of an array.

5.2.2.4 Total_cost(SS,N)

This function calculates the grand loss of a given configuration of transformers and neighbor**.** It also takes list of transformer and a list of neighbors to calculate total cost.

5.3 Iterative flow

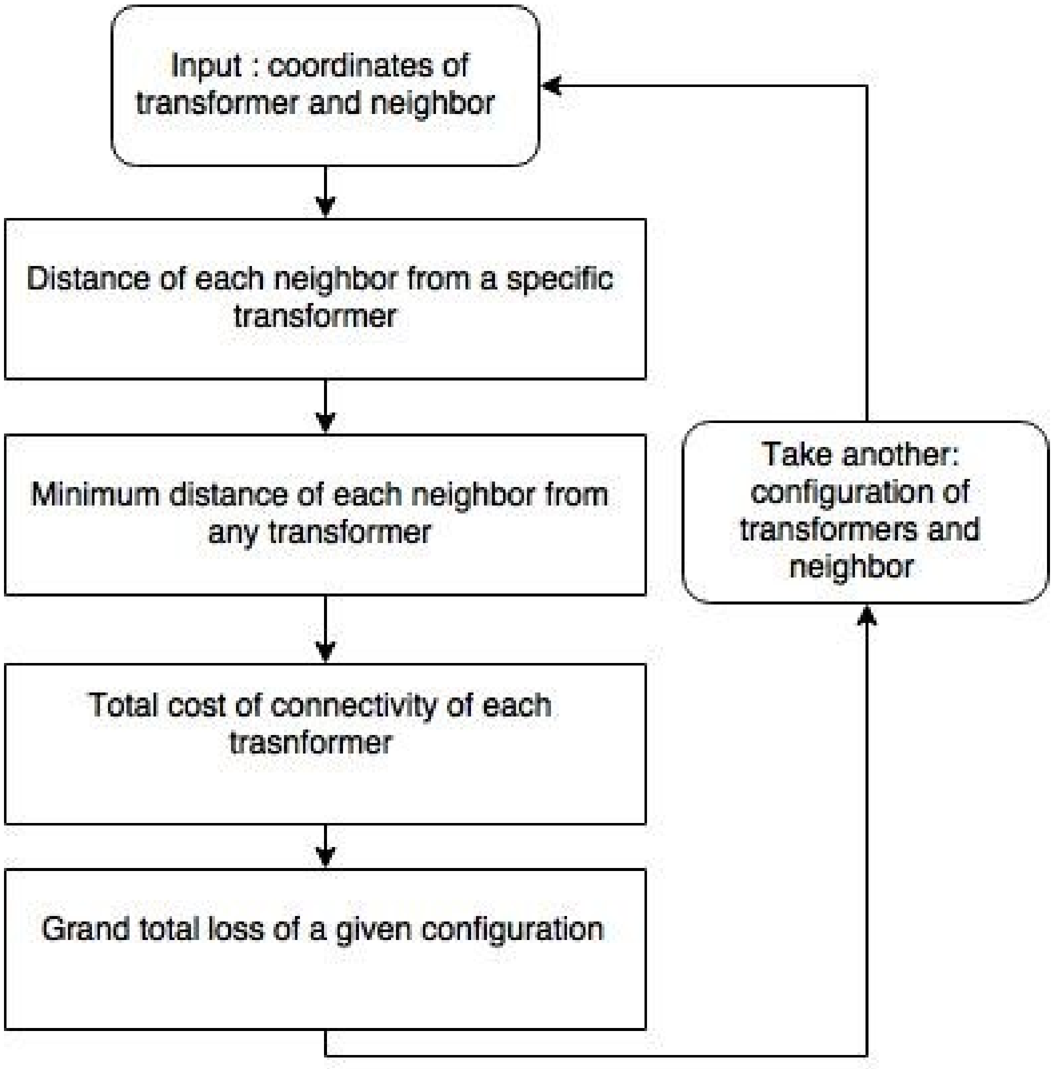


Figure 5.3 :- Cost Calculation of a configuration

In figure 5.3, four steps are used in all three approaches to calculate the cost of each configuration and to get the optimal configurations.

5.4 Setting up Prerequisites

All three approaches are elaborated by taking a manual input of list of neighbor and fixed number of transformers to connect them optimally. This helps in comparing all three approaches and providing a valid conclusion.

20 neighbors are placed on a $[10][10]$matrix with their positions:-

Array ([[9, 5],[8, 1],[8, 3],[8, 4],[8, 6],[7, 4],[7, 9],[6, 8],[5, 1],[5, 5],[5, 7],[4, 2],[4, 4],[3, 7],[3, 8],[2, 2],[2, 4],[2, 6],[2, 7],[0, 3]])

5.5 Approaches Implementation

5.5.1 Deterministic approach / Brute Force

In 10 by 10 matrix, there are 100 possible ways to place a transformer. Since there are 3 transformers, there will be $C(100;3)$(Combinations) *i.e.* 1,61,700 candidate solutions to the problem. In this deterministic approach, it traces through all of the candidate solutions to get the optimal one.

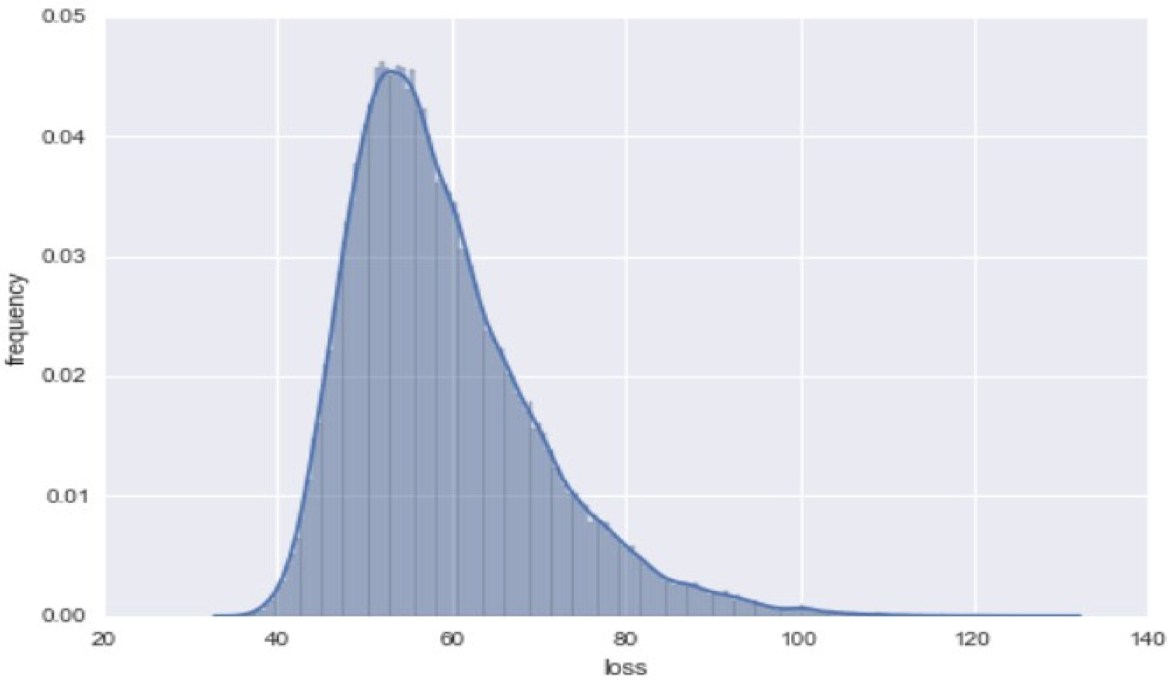


Figure 5.4 Loss vs Frequency chart

Figure 5.4, Shows the loss values of all 1,61,700 candidate solutions and their frequency. With this data, the minimum loss value can be figured out and that will be the optimal solution.

Using the above mentioned input, running time is calculated and results are :-

CPU times: 32.2 seconds

It can be observed that as the problem size grows, the number of candidate solutions also grow exponentially, so it will take to more time to get the optimal solution.

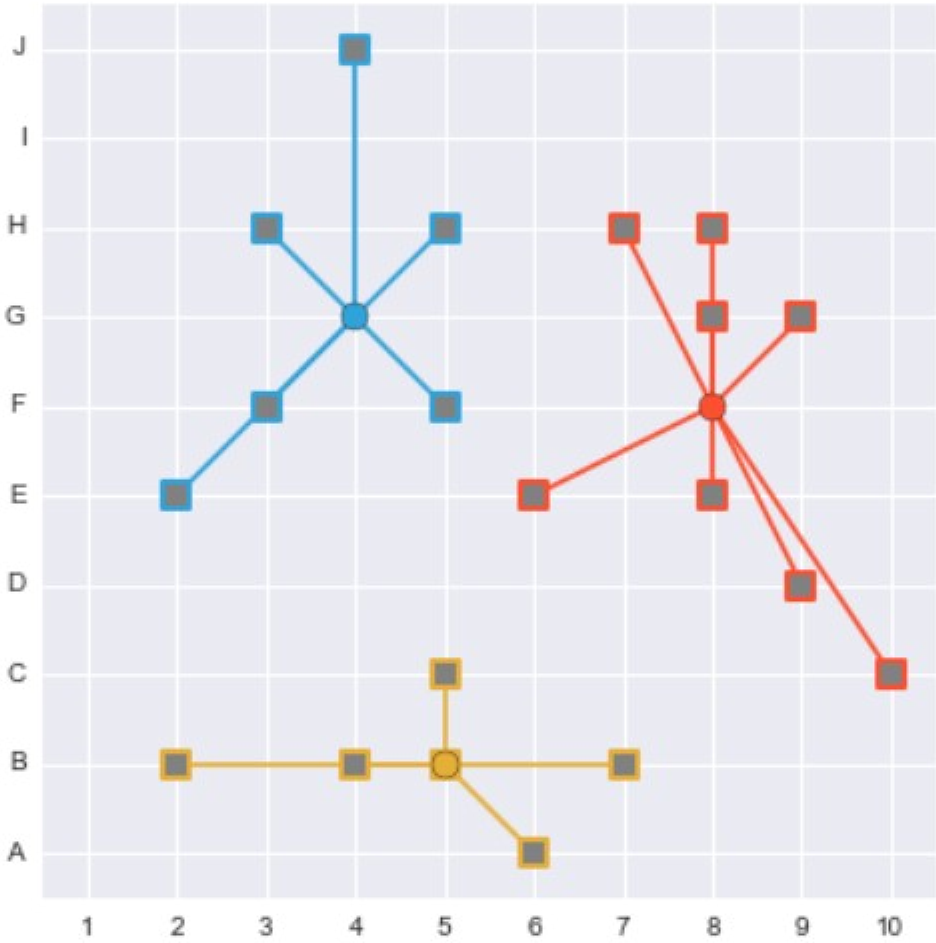


Figure 5.5 Output of the deterministic approach

The cost value associated with the both the actual configuration and optimal configuration are calculated. It can be concluded that amount of cost will be reduced with the optimal configuration.

5.5.2 Curve-ball solution: Using K-means Clustering

By applying unsupervised clustering, optimal placement of the transformers can be done. Here unsupervised clustering is done by applying K-means clustering. K-means clustering is a better option as compared to other clustering techniques because of better pre-clustering, and it is less expensive to run.

In this instance of k-means clustering k=3, is given to imitate the number of transformers and the array of neighbors is provided to do clustering. The process will break the neighbors into 3 clusters and place the transformers as close to the centroid of each cluster as possible.

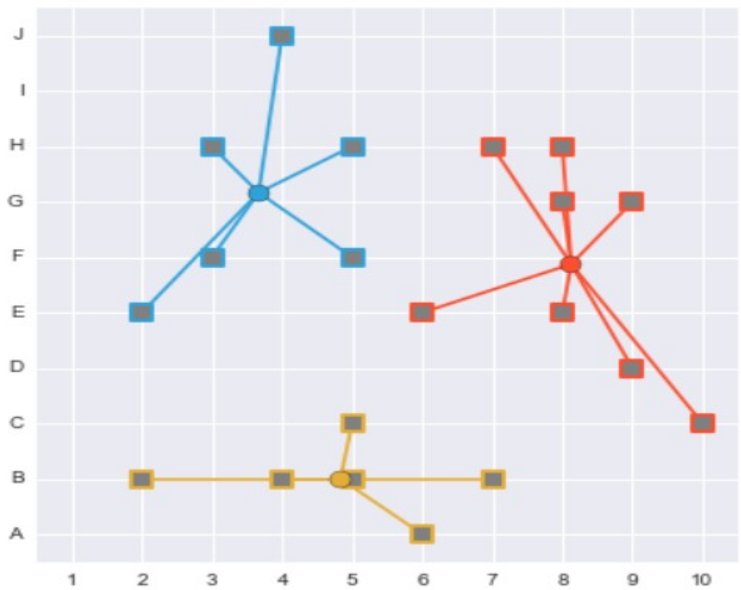


Figure 5.6 Non Rounded output of k-means clustering

In figure 5.6, the output of k-means clustering is shown where centroid depicts the placement of the transformers. It can be seen the centroids are not at the exact cell location, so round function is used to place the transformers at the nearest cell.

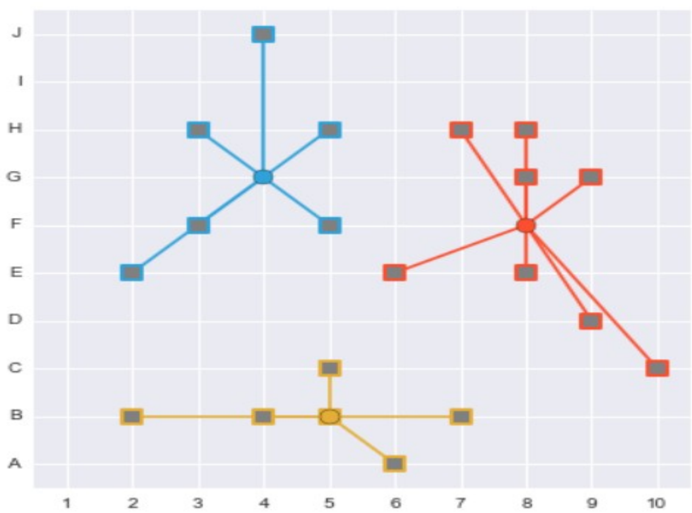


Figure 5.7 Final rounded output of the k-means clustering

The output of deterministic approach and curve-ball approach are same, so it can be concluded that the accuracy of both approaches is same.

With this input data, running time is calculated and results are :-

CPU times: 46.8 milliseconds.

Since the Time is much less than the running time of deterministic approach, this approach is much better than the previous.

5.5.3 Stochastic approach: Simulated Annealing(SA)

This is the most efficient approach to tackle the problem. SA is a powerful stochastic optimization method with deep connections to Markov Chain Monte Carlo. It uses a clever analogy to the physical process of annealing metal as inspiration for choosing acceptance probabilities in the proposal step.

SA tries to balance twin goals of (i) exploring widely and (ii) squeezing the last bit of optimization out of probable winners.

Given a loss function, a function is used to make proposals by perturbing the current state and a starting state and simulated annealing function will continue to explore the problem domain.

There are many different ways one could do this. Typically, the best ways allow enough random change to help explore the domain but preserve some similarity between the current and proposed candidates—otherwise, we're just doing a completely random exploration which defeats the purpose.

It'll perturb the current solution by randomly choosing one of the power plants and adding noise to its current position.

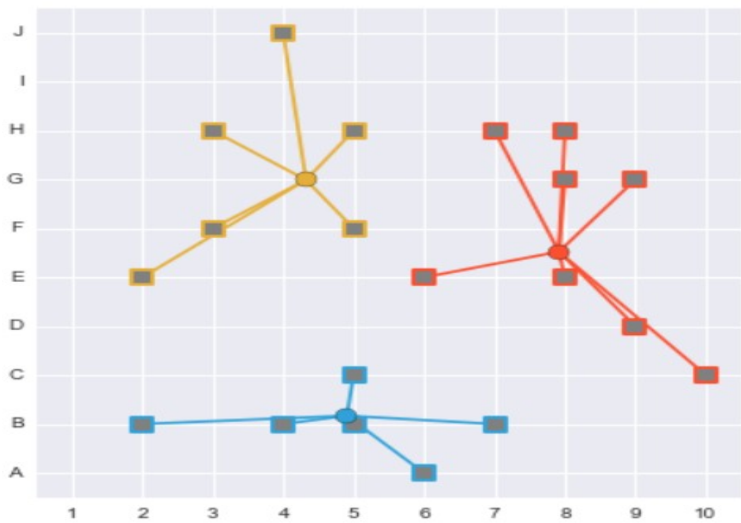


Figure 5.8 Non Rounded output of the SA approach

In figure 5.8, transformers' positions are not at the grid points of the matrix,so the transformers will be placed at the closest grid points by applying round function and the final output is:-

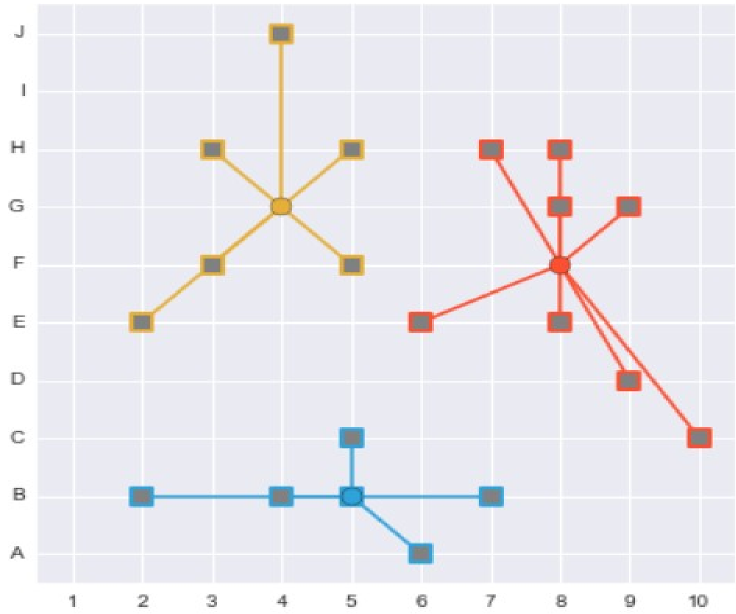


Figure 5.9 Final rounded output of SA Approach

With this input data, running time is calculated and results are :-

Now how fast it got the solution with SA.

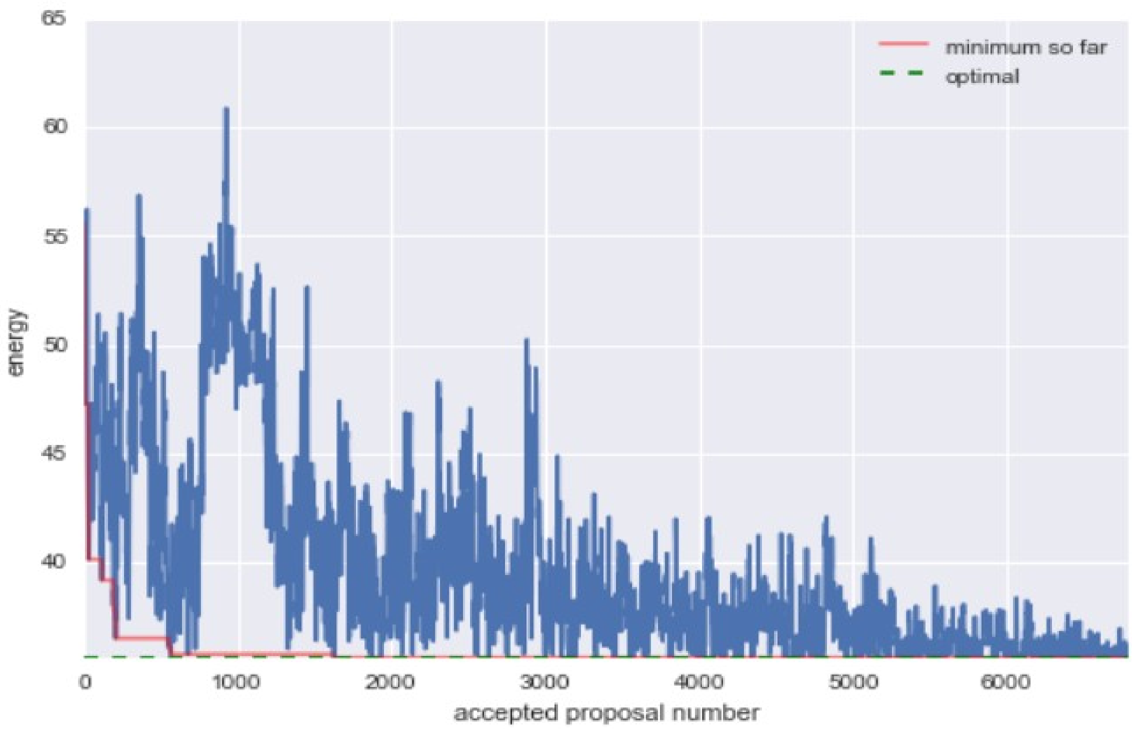


Figure 5.10 : Accepted proposal number vs energy

As we can see, this algorithm started getting *pretty good* answers almost immediately and fairly quickly converge down towards the optimum. Further it actually computes an optimal solution. Most important point of consideration is that we were not even perturbing the current candidate in a smart way.

One of the reasons why simulated annealing is so widely used is that in practice, it is robust over many possible choices of parameters, which include how to make proposals, starting temperature, number of times to re-anneal, fraction of temperature to re-anneal up to, etc.

CPU times: 4.66 seconds

The CPU time is quite less as compared to the deterministic approach and similar to that in the curve-ball's solution.

Table 5.1 list of approaches and respective execution time

Approach	**Execution Time**
Deterministic Approach	32.2 seconds
Curve-ball Solution	46.8 milliseconds
Stochastic Approach	4.66 seconds

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

The combinatorial optimization problem of transformers placement has been solved by three different approaches: deterministic, clustering and stochastic. All three approaches give almost same optimal solution with different time and space complexity. From the study and implementation of all three approaches, a conclusion can be drawn that for a smaller size of problem, all three take almost same time, but as the size grows, stochastic approach is better than others.

6.2 Future scope

These proposed solutions can be taken further by taking more numbers of parameters into consideration, to get more optimal cum realistic outcomes. Proposed solutions can be used to solve other similar problems like, designing the best road network, developing the optimal way to deliver parcels, designing more compact circuits and many more.

REFERENCES

1. Bhalla, M. S. "Transmission and distribution losses (power)." *Proceedings of the National Conference on Regulation in infrastructure Services: progress and way forward, The Energy and Resources Institute, New Delhi, India*. 2000.
2. Mungkung, Narong, et al. "Analysis of technical loss in distribution line system." *WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering*. Eds. M. Demiralp, N. A. Baykara, and N. E. Mastorakis. No. 8. World Scientific and Engineering Academy and Society, 2009.
3. Navani, J. P., N. K. Sharma, and Sonal Sapra. "Technical and non-technical losses in power system and its economic consequence in Indian economy." *International Journal of Electronics and Computer Science Engineering* 1.2 (2012): 757-761.
4. Shirmohammadi, Dariush, and H. Wayne Hong. "Reconfiguration of electric distribution networks for resistive line losses reduction." *Power Delivery, IEEE Transactions on* 4.2 (1989): 1492-1498.
5. Quezada, Víctor H. Méndez, Juan Rivier Abbad, and Tomás Gómez San Román. "Assessment of energy distribution losses for increasing penetration of distributed generation." *Power Systems, IEEE Transactions on* 21.2 (2006): 533-540.
6. Kirkpatrick, Scott, C. Daniel Gelatt, and Mario P. Vecchi. "Optimization by simulated annealing." *science* 220.4598 (1983): 671-680.
7. Antmann, P. "Reducing technical and non-technical losses in the power sector. Background Paper for the WBG Energy Strategy." *Energy Unit World Bank* (2009): 1-34.
8. Merlin, A., and H. Back. "Search for a minimal-loss operating spanning tree configuration in an urban power distribution system." *Proc. of the Fifth Power System Conference (PSCC), Cambridge*. 1975.
9. Gönen, Turan, et al. *Development of advanced methods for planning electric energy distribution systems. Final report*. No. COO-4830-3. Oklahoma Univ., Norman (USA). School of Industrial Engineering, 1979.
10. Civanlar, Seyhan, et al. "Distribution feeder reconfiguration for loss reduction." *IEEE Trans. Power Del.;(United States)* 3.3 (1988).

11. Aoki, K., T. Ichimori, and M. Kanezashi. "Normal state optimal load allocation in distribution systems." *Power Delivery, IEEE Transactions on* 2.1 (1987): 147-155.

12. Shirmohammadi, Dariush, et al. "A compensation-based power flow method for weakly meshed distribution and transmission networks." *Power Systems, IEEE Transactions on* 3.2 (1988): 753-762.

13. Karp, Richard M. "Probabilistic analysis of partitioning algorithms for the traveling-salesman problem in the plane." *Mathematics of Operations Research* 2.3 (1977): 209-224.

14. Crowder, Harlan, and Manfred W. Padberg. "Solving large-scale symmetric travelling salesman problems to optimality." *Management Science* 26.5 (1980): 495-509.

15. Ackermann, Thomas, Göran Andersson, and Lennart Söder. "Distributed generation: a definition." *Electric power systems research* 57.3 (2001): 195-204.

16. Ackermann, Thomas, and Valery Knyazkin. "Interaction between distributed generation and the distribution network: operation aspects." *Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES*. Vol. 2. IEEE, 2002.

17. Joos, G., et al. "The potential of distributed generation to provide ancillary services." *Power Engineering Society Summer Meeting, 2000. IEEE*. Vol. 3. IEEE, 2000.

18. Hoff, Tom, Howard J. Wenger, and Brian K. Farmer. "The value of grid-support photovoltaics in providing distribution system voltage support." *Proceedings of the 1994 American Solar Energy Society Annual Conference*. 1994.

19. Sur, Chiranjib, and Anupam Shukla. "Discrete Invasive Weed Optimization Algorithm for Graph Based Combinatorial Road Network Management Problem." *Computational and Business Intelligence (ISCBI), 2013 International Symposium on*. IEEE, 2013.

20. Bora, Teodoro C., Leandro dos S. Coelho, and Luiz Lebensztajn. "Bat-inspired optimization approach for the brushless DC wheel motor problem." *Magnetics, IEEE Transactions on* 48.2 (2012): 947-950.

21. Kennedy, James, and Russell C. Eberhart. "A discrete binary version of the particle swarm algorithm." *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on*. Vol. 5. IEEE, 1997.

22. Sahraei-Ardakani, Mostafa, et al. "A study of electricity market dynamics using invasive weed colonization optimization." *Computational Intelligence and Games, 2008. CIG'08. IEEE Symposium On*. IEEE, 2008.

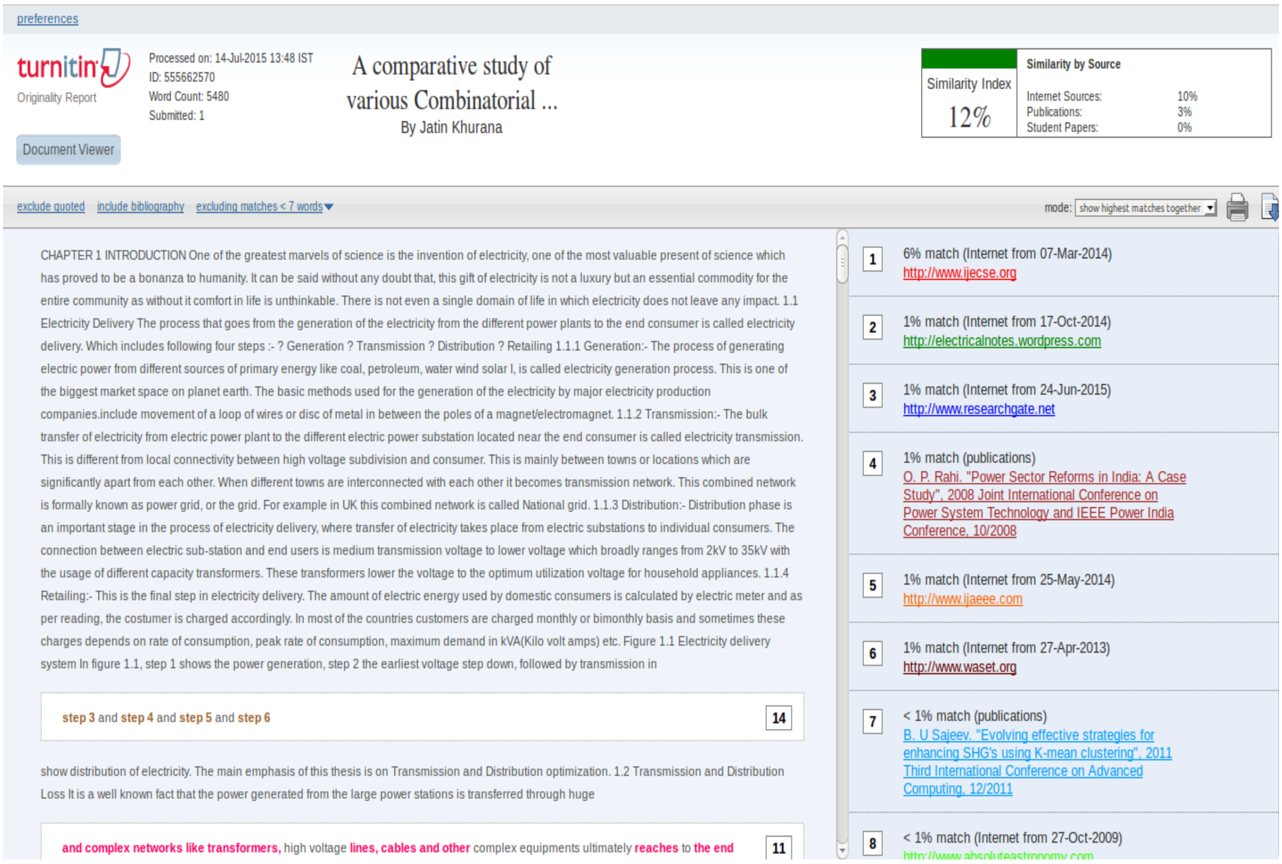
23. De Jong, Kenneth Alan. "Analysis of the behavior of a class of genetic adaptive systems." (1975).
24. Eberhart, Russ C., and James Kennedy. "A new optimizer using particle swarm theory." *Proceedings of the sixth international symposium on micro machine and human science*. Vol. 1. 1995.
25. Eberhart, Russell C., and Yuhui Shi. "Particle swarm optimization: developments, applications and resources." *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*. Vol. 1. IEEE, 2001.
26. Golberg, David E. "Genetic algorithms in search, optimization, and machine learning." *Addion wesley* 1989 (1989).
27. Eberhart, Russ C., and James Kennedy. "A new optimizer using particle swarm theory." *Proceedings of the sixth international symposium on micro machine and human science*. Vol. 1. 1995.

LIST OF PUBLICATIONS

Accepted

- Jatin and S. Batra, “MongoDB vs SQL : A case study on electricity data”, International Conference on Emerging Research in Computing, Information, Communication and Applications, (ERCICA 2015), Bangalore.

PLAGRISM REPORT



VIDEO LINK

YouTube video is at this channel :- http://youtube.com/jatinkhuranaTV