Mathematical Optimization: Models, Methods and Applications

Executive Summary: Mathematical Optimization Work by Jørgen Leiros

This project focused on mathematically modeling wind turbine cable networks to minimize total cost under various constraints. The work involved formulating objective functions and constraints, then applying an optimization algorithm to determine the optimal layout of connections between turbines. The model was implemented using the Gurobi optimization library in Python.

Key areas of focus in his work include:

1. **Wind Farm Cable Routing Problem** – Developed a mathematical model to optimize wind turbine cable routing, considering constraints such as unsplittable energy flow, capacity limits, and cost minimization. Sensitivity analysis was conducted on cable capacity and the number of cables required, and alternative modeling approaches were explored.
2. **Dissimilarity-Based Clustering** – Applied clustering methods to optimize data grouping. His work extended clustering models with must-link and cannot-link constraints, explored soft constraints for practical applications, and reformulated the problem using Mixed Integer Linear Programming (MILP).
3. **Optimization in Supply Chain and Inventory Management** – Utilized advanced mathematical programming techniques such as Mixed Integer Linear Programming (MILP), Non-Linear Programming (NLP), and Stochastic Programming to address supply chain and inventory challenges.

Through this work, Jørgen has gained expertise in formulating mathematical models, sensitivity analysis, and applying optimization techniques in energy, data clustering, and supply chain management. His research showcases a data-driven approach to decisionmaking, leveraging mathematical rigor to improve operational efficiency and resource allocation.

Question 1 - The Wind Farm Cable Routing Problem

I – The Mathematical Model

The following mathematical model uses sets and parameters such as they are presented in the exercise set and only altered or added parts of the model will be presented in this report to save space for discussion. With a unit cost of 393, the cost of a potential cable installation cost c_{ij} is the Euclidean distance between two nodes multiplied by the unit cost (Britannica, 2025). Constraint (1) ensures that each turbine has exactly one outgoing connection as it is stated in the problem that the energy flow is unsplittable. Constraint (2) defines the energy flow leaving a turbine as the power generated at the turbine added to the incoming energy from any other turbines. The third constraint (3) ensures that the power flow of a single cable never exceeds its capacity and the final constraint (4) states that the substation should have no outgoing connections.

$$Eclidean\ Distance(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},\ c_{ij} = 393 * Euclidean\ Distance(i,j)$$

- $x_{i,j} \in \{0,1\}$: Binary variable, 1 if cable connection between node *i* and node *j* is installed and 0 otherwise.
- $f_{ij} \geq 0$: Continuous variable, representing power flow from node *i* to node *j*.

$$minimize \sum \sum x_{ij} c_{ij}\ {}_{i \in V\ j \in V}$$

Subject to:

$$\sum_{j \in V, j \neq i} x_{ij} = 1, \qquad \forall i \in Vt \qquad (1)$$

$$\sum_{j \in V, j \neq i} f_{ij} = P_i + \sum_{j \in V, j \neq i} f_{ji}, \qquad \forall i \in Vt \qquad (2)$$

$$f_{ij} \leq k * x_{ij}, \qquad \forall i, j \in V, i \neq j \qquad (3)$$

$$\sum_{j \in Vt} x_{0j} = 0 \qquad (4)$$

II – Sensitivity Analysis for Cable Capacity

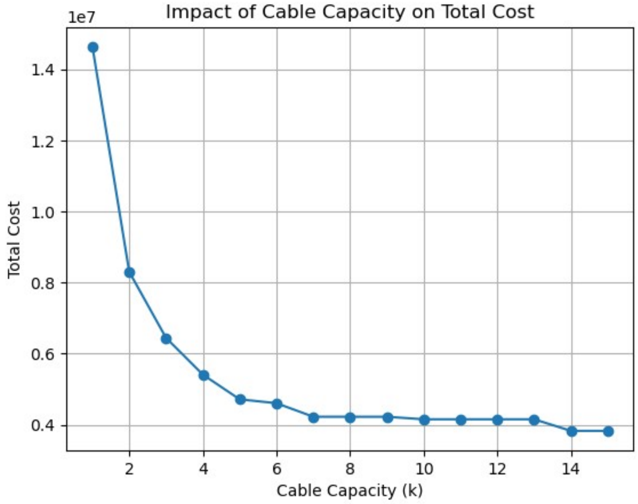


Figure 1: Graph showing the models sensitivity to changes in cable capacity.

To analyze the model's sensitivity to varying values of K for the capacity of each cable, I defined a set of values for K and found the optimal solution and related installation cost for each of these values. The results from the sensitivity analysis can be seen in figure 3, and it displays a steep downward slope for small values for K and gradual flattening of the curve as K increases. More specifically, we can see that the total cost drops significantly from a cable capacity of 1 to 4 but this decrease stagnates somewhat after for values 4 15. It should also be noted that we can see from the graph that a cable capacity of 14 results in the lowest total cost, which matches the findings in the previous section where we found that the optimal solution included all 14 turbines connected by one cable.

III – Cable Routing With 5 MW Capacity

With a maximum of C = 3 cables each with a capacity of 5 MW requires extending the mathematical model with one constraint as well as changing the value of k. The constraint states that the sum of incoming connections at the substation must be less than or equal to the maximum number of cables permitted.

$$\sum_{i \in Vt} x_{i,0} \leq C$$

Adding the constraint of a maximum count of 3 cables connected to the substation and changing the maximum capacity to 5 MW per connection results in the cable connections visualized in figure 2. The constraint of maximum 5 MW capacity forces the model to install at least three cables.

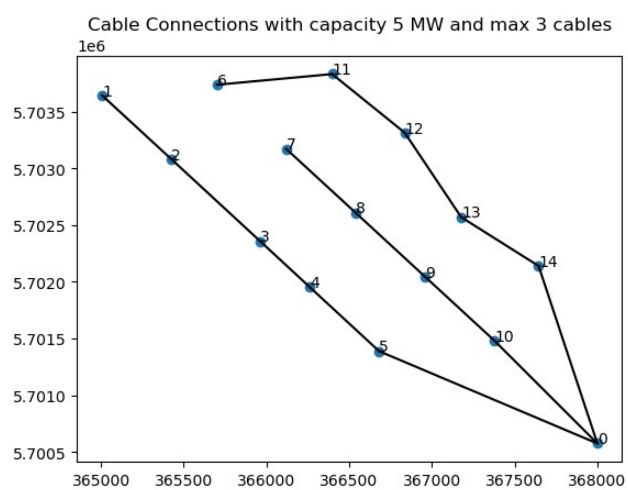


Figure 2: Optimal layout with max cable = 3 and cable capacity = 5MW

IV – Sensitivity Analysis for Number of Cable

Testing the model's sensitivity to various values for the maximum number of cables reveals that the model is completely insensitive to changes in this value, at least for values with feasible solutions. With a maximum capacity of 5 MW, it is impossible to connect each turbine to the substation using less than 3 cables and adding more than three cables will increase costs as we saw in the sensitivity analysis in section II.

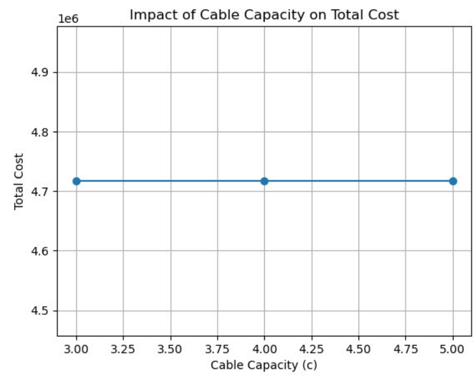


Figure 3: Total cost for various values for cable capacity C

V – Mathematical Model extended for Different Cable Types

In addition to the sets and variables defined in the exercise text I have the following:

- T: Set of cable types
- k_t: Maximum capacity for cable type t$\in T$
- u_t: Unit cost of cable type $t \in T$
- d_{ij}: Euclidean distance between nodes $i, j \in V$
- c_{ijt}: Cost of installing cable type $t \in T$ between nodes $i, j \in V$

- $x_{ij\,t} \in \{0, 1\}$: Binary decision variable, 1 if cable type t is installed between node i and j and 0 otherwise.
- $f_{ijt} \geq 0$: Continuous decision variable, representing energy flow through cable t between nodes i and *j.*

$$minimize \sum_{t\in T}\sum_{i\in V}\sum_{j\in V} x_{ijt}c_{ijt}$$

Subject to:

1. A turbine must have exactly one outgoing connection.

$$\sum_{j\in V, j\neq i}\sum_{t\in T} x_{ijt} = 1, \forall i \in Vt$$

2. A turbine can have at most one incoming connection.

$$\sum_{j\in Vt, j\neq i}\sum_{t\in T} x_{ijt} \leq 1, \forall i \in Vt$$

3. The power flow transferred to the next node is equal to the incoming power to the current node added to the power generated at the current node.

$$\sum_{j\in V, j\neq i}\sum_{t\in T} f_{ijt} = P_i + \sum_{j\in V, j\neq i}\sum_{t\in T} f_{jit}, \forall i \in Vt$$

4. The final power accumulated at the substation must be equal to the power generated by the entire network.

$$\sum_{j\in Vt}\sum_{t\in T} f_{j0t} = \sum_{j\in Vt} P_j$$

5. The substation can have no outgoing connections.

$$\sum_{j\in Vt}\sum_{t\in T} x_{0jt} = 0$$

6. The power flow of any connection must not exceed the capacity of the cable K.

$$f_{ijt} \leq capacity_t * x_{ijt}, \qquad \forall i, j \in V, \qquad i \neq j, \qquad t \in T$$

The suggested model increases the number of decision variables by a factor of the number of cable types both for *x* and *f*. This increase in decision space can lead to a more complex and computationally intensive model, especially for large number of cable types. Allowing for various cable types with corresponding cost and capacity constraints provides a more realistic representation of the real-world scenario.

VI – Alternative Modeling Approach

Wind energy is a renewable and sustainable energy source, and its role in the global energy market is rapidly expanding. As the demand for wind energy grows, designing efficient layouts for wind farms has become increasingly important. The process of designing a wind farm can be broken down into a sequence of tasks, many of which can benefit from mathematical optimization. For the wind farm cable routing problem in this exercise, the locations of the turbines have already been decided, a highly challenging task which involves minimizing the wake-effect which on average leads to 10 – 20% power loss (Barthelmie et al., 2009). As the location of the turbines have been provided the next problem becomes deciding how the turbines should be connected to the substation to minimize cost and ensuring that as much energy as possible is transferred to the substation. According to Fischetti and Pisinger (2018), this problem is still carried out manually in many companies even though mathematical optimization approaches like Mixed Integer Linear Programming (MILP) and Heuristics can provide optimal solutions.

The mathematical model such as I have defined it in part I, takes a MILP approach so to compare this to an alternative modeling approach, I will present the heuristic approach suggested by Fischetti & Pisinger (2018). The proposed model resembles the extended model presented in part V as it considers different cable types with related costs and capacities, but it adds constraints preventing crossing cables, setting a maximum number of cables connected to the substation and considers cable losses. Additionally, the authors handle cable types by introducing an auxiliary variable $y_{ij} \in \{0,1\}$ which aggregates decisions across cable types. Introducing this variable alters some constraints slightly in addition to the added constraints, so meaningful differences from my model are presented here:

Constraints:

- The sum of cable types used for a single connection is equal to the auxiliary variable which indicates whether any cable is installed for the connection. Furthermore, the auxiliary variable is defined to have a sum of one for each connection and 0 for outgoing connections from the substation.

$$\sum_{t \in T} x_{itj} = y_{ij}, \qquad i,j \in V, j \neq i$$

$$\sum_{j \neq h} y_{hj} = 1, \qquad h \in V_t$$

$$\sum_{j \neq h} y_{hj} = 0, \qquad h \in V_0$$

- Additional constraint to limit the number of incoming connections to the substation.

$$\sum_{i \neq h} y_{ih} \leq C, \qquad h \in V_0$$

Although the auxiliary variable provides for a more modular approach which can be beneficial for further extension of the model, the approach as described thus far can still be described as MILP. For the proposed framework, MILP is used to generate an initial solution to reduce the problem size when preventing crossing cables. Once the initial solution is gener ated, the model applies heuristic strategies to fix certain variables, such as selecting specific arcs $y_{ij} = 1$ to include in the solution. Fixing these variables reduces the problem size, as related constraints, like those for crossing arcs, can be dynamically excluded. For example, when $y_{ij} = 1$, crossing arcs *(h, k)* are excluded by adding the constraint: $y_{ij} + y_{hk} \leq 1, \forall crossing\ arcs$. The refined problem is then solved iteratively using the MILP solver, focusing on unfixed variables while incorporating these dynamic constraints. The best solution from the iterative process is then used as a starting point for the full MILP model subject to all original constraints, ensuring an efficient balance between exact optimization and heuristic-guided simplifications.

Question 2 – Dissimilarity Based Clustering

I – Cluster Assignments and Objective Value

Objective value: 236.63999999999993 and cluster assignment can be seen in Appendix 2A.

II – Must-Link Extended Model

To incorporate must-link constraints into the model I first defined the list of items which needed to be in the same cluster as tuples such that: $must\_link = \{(12,21),(20,13)\}$. The added constraint is as follows: $x_{ic} = x_{iic}, \forall c \in \{1,2,\dots,K\}, (i, ii) \in must\_link$ and the resulting cluster assignments can be seen in Appendix 2A.

III – Cannot-Link Extended Model

Similar to the previous exercise, a set of tuples has been defined as $cannot\_link = \{(6,18),(23,14)\}$. This extension of the model also requires an additional constraint: $x_{ic} + x_{iic} \leq 1, \forall c \in \{1,2,\dots,K\}, (i, ii) \in cannot\_link$. The new cluster assignments can be seen in Appendix 2A.

IV – Soft Constraints for Practical Settings

If a feasible solution does not exist, or if the cannot-link pairs are so similar that putting them in the same cluster produces a lower dissimilarity score despite the added penalty term, soft constraints might be more practical and lead to a better model. To accommodate soft constraints for this model, one would have to introduce a slack variable to the decision variables as well as modify both the objective function and the cannot-link constraint introduced in part III. With the extended model proposed in part II as basis, the following will be the alterations made to the mathematical model:

$$minimize \sum_{c=1}^{K} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} x_{ic} x_{jc} + \sum_{(i,ii) \in cannot\_link} p_{i,ii} * s_{i,ii}$$

Subject to (including the model constraints introduced in the problem and the added constraint in part II):

$$x_{ic} + x_{iic} \leq 1 + s_{i,ii}, \qquad \forall c \in \{1,2,\dots,K\}, \qquad \{i, ii\} \in cannot\_link$$

$$s_{i,ii} \geq 0 \qquad \forall (i, ii) \in cannot\_link$$

V – MILP Reformulation

When possible, it can be beneficial to reformulate nonlinear problems as MILP problems. The components of the mathematical model from parts I-IV which make the problem a non-linear one, are the two binary variables in the objective function indicating whether an item is assigned to a cluster, namely x_{ic} and x_{jc}. To reformulate this into a MILP problem, we can replace the two binary variables with a single variable indicating whether the pair of items are assigned to the same cluster. When doing so we also need to add certain constraints to ensure that z_{ijc} represents the product $x_{ic} * x_{jc}$.

$$minimize \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{c=1}^{K} \delta_{ij} z_{ijc} + \sum_{(i,j)\in cannot\_connect} p_{ij} * s_{ij}$$

Subject to (in addition to the constraints presented in the exercise and the hard constraint in part II and soft constraints in part IV):

$$z_{ijc} \leq x_{ic}, \qquad Z_{ijc} \leq x_{jc} \qquad Z_{ijc} \geq x_{ic} + x_{jc} - 1, \qquad \forall ijc$$

The added constraint aims to linearize the product of the two binary variables in the original problem formulation. These constraints ensure that $z_{ijc} = 1$ if $x_{ic} = 1$ and $x_{jc} = 1$, and 0 otherwise.

VI – Model Extension

To make sure the clusters are as balanced as possible, we can follow one of two approaches, similar to problems III and IV. Adding a strict constraint could be done to make sure each cluster *K* has exactly the same amount of items *N* so long as *N mod K* = 0, allowing for only the necessary deviance if not. This extension of the model could be made by adding the following:

$$n_c = \sum_{i=1}^{n} x_{ic}\,, \quad minsize = \left\lfloor \frac{n}{K} \right\rfloor, \quad maxsize = \left\lceil \frac{n}{K} \right\rceil$$

$$minsize \leq \sum_{i=1}^{n} x_{ic} \leq maxsize$$

$$\sum_{i=1}^{n} x_{ic} = minsize + v_c, \quad \forall c \in \{1,2,\dots,K\}, \qquad \sum_{c=1}^{K} v_c = n \; mod\; K$$

Alternatively, a soft constraint encourages balanced clusters but allows for some imbalance if strictly balancing clusters significantly increases the dissimilarity score. Soft constraints can be incorporated with the following alterations to the model from part IV:

$$minimize \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{c=1}^{K}\delta_{ij}z_{ijc} + \sum_{(i,j)\in cannot\_connect} p_{ij} * s_{ij} + \lambda\sum_{c=1}^{K}(s_c^{+} + s_c^{-})$$

Subject to (including the constraints from the model in part IV):

$$n_c = \sum_{i=1}^{n} x_{ic}, \quad n_c = \frac{N}{K} + s_c^{+} - s_c^{-}$$

$$s_c^{+} \geq 0, \quad s_c^{-} \geq 0, \quad \forall c \in \{1,2,\dots,K\}$$

Whether to use soft or hard constraints for this problem depends on the real-world problem we are trying to solve by clustering based on dissimilarity.

Question 3 – Mathematical Optimization in Supply Chain and Inventory Management

I – Mixed Integer Linear Programming II – Non-Linear Programming III – Stochastic Programming

References

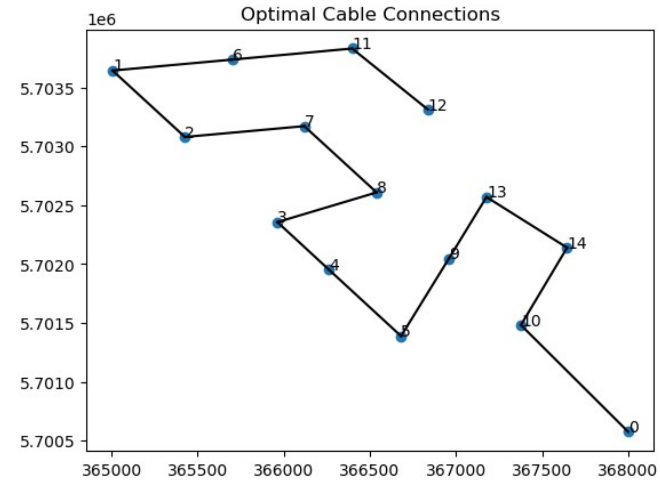
Barthelmie, R. J., Hansen, K., Frandsen, S. T., Rathmann, O., Schepers, J. G., Schlez, W., ... & Chaviaropoulos, P. K. (2009). Modelling and measuring flow and wind turbine wakes in large wind farms offshore. Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology, 12(5), 431-444.

Britannica, T. Editors of Encyclopaedia (2025, January 7). Euclidean distance. Encyclopedia Britannica. https://www.britannica.com/science/Euclidean-distance

Fischetti, M., & Pisinger, D. (2019). Mathematical optimization and algorithms for offshore wind farm design: An overview. *Business & Information Systems Engineering*, *61*, 469-485.

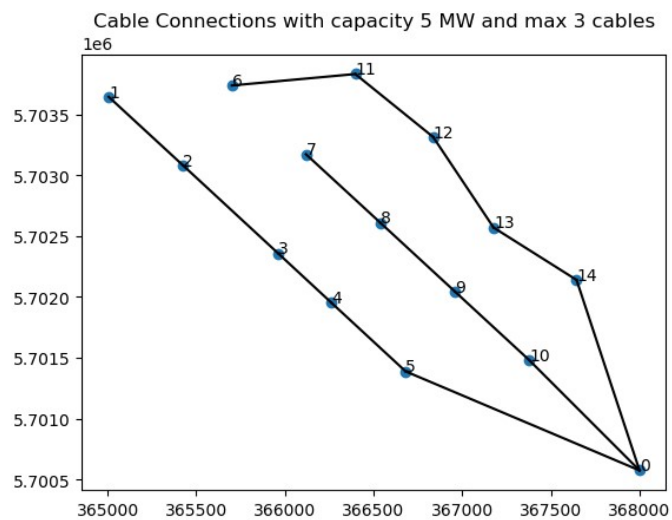
Appendix

1A. Optimal cable connections and related total cost.



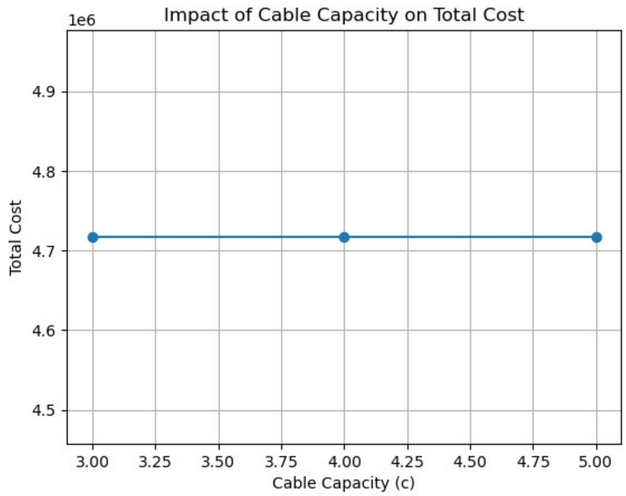

```
Total cost: 3830477.664544905
Cable from 1 to 2: Installed with flow 4.0
Cable from 2 to 7: Installed with flow 5.0
Cable from 3 to 4: Installed with flow 8.0
Cable from 4 to 5: Installed with flow 9.0
Cable from 5 to 9: Installed with flow 10.0
Cable from 6 to 1: Installed with flow 3.0
Cable from 7 to 8: Installed with flow 6.0
Cable from 8 to 3: Installed with flow 7.0
Cable from 9 to 13: Installed with flow 11.0
Cable from 10 to 0: Installed with flow 14.0
Cable from 11 to 6: Installed with flow 2.0
Cable from 12 to 11: Installed with flow 1.0
Cable from 13 to 14: Installed with flow 12.0
Cable from 14 to 10: Installed with flow 13.0
```

1B. Optimal cable connections with a capacity of 5 MW and maximum 3 cables.




```
Optimal solution found
Total cost: 4717234.7554237805
Cable from 1 to 2: Installed with flow 1.0
Cable from 2 to 3: Installed with flow 2.0
Cable from 3 to 4: Installed with flow 3.0
Cable from 4 to 5: Installed with flow 4.0
Cable from 5 to 0: Installed with flow 5.0
Cable from 6 to 11: Installed with flow 1.0
Cable from 7 to 8: Installed with flow 1.0
Cable from 8 to 9: Installed with flow 2.0
Cable from 9 to 10: Installed with flow 3.0
Cable from 10 to 0: Installed with flow 4.0
Cable from 11 to 12: Installed with flow 2.0
Cable from 12 to 13: Installed with flow 3.0
Cable from 13 to 14: Installed with flow 4.0
Cable from 14 to 0: Installed with flow 5.0
```

1C. Sensitivity of the model to various maximum cable constraints.



2A. Optimal cluster assignments to minimize dissimilarity (I-III from left to right).

Point 6 is assigned to cluster 1	Point 1 is assigned to cluster 1	Point 2 is assigned to cluster 1
Point 10 is assigned to cluster 1	Point 3 is assigned to cluster 1	Point 13 is assigned to cluster 1
Point 11 is assigned to cluster 1	Point 4 is assigned to cluster 1	Point 16 is assigned to cluster 1
Point 12 is assigned to cluster 1	Point 5 is assigned to cluster 1	Point 17 is assigned to cluster 1
Point 14 is assigned to cluster 1	Point 8 is assigned to cluster 1	Point 18 is assigned to cluster 1
Point 21 is assigned to cluster 1	Point 9 is assigned to cluster 1	Point 19 is assigned to cluster 1
Point 24 is assigned to cluster 1	Point 15 is assigned to cluster 1	Point 20 is assigned to cluster 1
Point 27 is assigned to cluster 1	Point 25 is assigned to cluster 1	Point 22 is assigned to cluster 1
Point 29 is assigned to cluster 1	Point 26 is assigned to cluster 1	Point 23 is assigned to cluster 1
Point 2 is assigned to cluster 2	Point 6 is assigned to cluster 2	Point 28 is assigned to cluster 1
Point 7 is assigned to cluster 2	Point 10 is assigned to cluster 2	Point 30 is assigned to cluster 1
Point 13 is assigned to cluster 2	Point 11 is assigned to cluster 2	Point 6 is assigned to cluster 2
Point 16 is assigned to cluster 2	Point 12 is assigned to cluster 2	Point 8 is assigned to cluster 2
Point 17 is assigned to cluster 2	Point 14 is assigned to cluster 2	Point 10 is assigned to cluster 2
Point 18 is assigned to cluster 2	Point 21 is assigned to cluster 2	Point 11 is assigned to cluster 2
Point 19 is assigned to cluster 2	Point 24 is assigned to cluster 2	Point 12 is assigned to cluster 2
Point 20 is assigned to cluster 2	Point 27 is assigned to cluster 2	Point 14 is assigned to cluster 2
Point 22 is assigned to cluster 2	Point 29 is assigned to cluster 2	Point 21 is assigned to cluster 2
Point 23 is assigned to cluster 2	Point 2 is assigned to cluster 3	Point 24 is assigned to cluster 2
Point 28 is assigned to cluster 2	Point 7 is assigned to cluster 3	Point 27 is assigned to cluster 2
Point 30 is assigned to cluster 2	Point 13 is assigned to cluster 3	Point 29 is assigned to cluster 2
Point 1 is assigned to cluster 3	Point 16 is assigned to cluster 3	Point 1 is assigned to cluster 3
Point 3 is assigned to cluster 3	Point 17 is assigned to cluster 3	Point 3 is assigned to cluster 3
Point 4 is assigned to cluster 3	Point 18 is assigned to cluster 3	Point 4 is assigned to cluster 3
Point 5 is assigned to cluster 3	Point 19 is assigned to cluster 3	Point 5 is assigned to cluster 3
Point 8 is assigned to cluster 3	Point 20 is assigned to cluster 3	Point 7 is assigned to cluster 3
Point 9 is assigned to cluster 3	Point 22 is assigned to cluster 3	Point 9 is assigned to cluster 3
Point 15 is assigned to cluster 3	Point 23 is assigned to cluster 3	Point 15 is assigned to cluster 3
Point 25 is assigned to cluster 3	Point 28 is assigned to cluster 3	Point 25 is assigned to cluster 3
Point 26 is assigned to cluster 3	Point 30 is assigned to cluster 3	Point 26 is assigned to cluster 3