Comparison between Louvain and Leiden Algorithm for Network Structure: A Review

Siti Haryanti Hairol Anuar¹,², Zuraida Abal Abas², Norhazwani Mohd Yunos², Nurul Hafizah Mohd Zakı², Nurul Akmal Hashim², Mohd Fariddudin Mokhtar¹,², Siti Azirah Asmai², Zaheera Zainal Abidin² and Ahmad Fadzli Nizam²

¹Faculty of Electrical and Electronic Engineering Technology (FTKEE), Universiti Teknikal Malaysia Melaka (UTeM), 76100 Malacca, Malaysia.
²Faculty of Communication and Information Technology (FTMK), Universiti Teknikal Malaysia Melaka (UTeM), 76100 Malacca, Malaysia.
³Faculty of Mechanical and Manufacturing Engineering Technology (FTKMP), Universiti Teknikal Malaysia Melaka (UTeM), 76100 Malacca, Malaysia.

Email: sitiharyanti@utem.edu.my

Abstract. In the real network, there must be a large and complex network. The solution to understand that kind of network structure is using the community detection algorithms. There are a lot of other algorithms out there to perform community detection. Each of the algorithms has its own advantages and disadvantages with different types and scale of complex network. The Louvain has been experimented that shows bad connected in community and disconnected when running the algorithm iteratively. In this paper, two algorithm based on agglomerative method (Louvain and Leiden) are introduced and reviewed. The concept and benefit are summarized in detail by comparison. Finally, the Leiden algorithm’s property is considered the latest and fastest algorithm than the Louvain algorithm. For the future, the comparison can help in choosing the best community detection algorithms even though these algorithms have different definitions of community.

1. Introduction

Recently, research on complex networks is becoming popular. A great variety of systems in nature, society and technology can be modelled as networks [1]. An interdisciplinary field that combines ideas from mathematics, physics, biology, computer science, social science and many more areas can be represented as a complex network. It is interesting to explore any kind of application domain [2].

Nowadays, complex networks are extremely important and could be analyzed to get the better decision making or monitoring of certain organizations [3]. Network analysis provides an effective tool to look into and investigate networks’ potential properties such as the small-world property, the scale-free property, the community structure property [4], and so on. This network analysis helps to improve the understanding of the complex world and it is challenging task [5].

Researchers have proposed many community detection algorithms with different types and scale of complex networks [6]. In this paper, the comparison between Louvain and Leiden algorithm based on modularity and hierarchical clustering in community detection are introduced and reviewed.
2. **Comparison between Louvain and Leiden Algorithm for Network Structure**

Community detection in a network structure is one of the most important network analysis tasks, or it is called community analytics. Primarily, there are two types of community detection methods, where are the agglomerative methods and the divisive methods.

![Figure 1. Taxonomy of community analytics.](image)

Based on the taxonomy in **Figure 1**, these community analytics based on modularity and hierarchical clustering is divided by two methods. For the divisive methods (refer left wing), up to down of the dendogram figure give meanings from big partition they will iteratively split by removing edges with low similarity and so-called split process [7]. However, for the agglomerative methods (refer right wing), is the other way round where from down to up. It means that from the complex network that content all nodes, they will iteratively merge if similarity becomes a high so-called merged process [8]. This paper shows the Louvain and Leiden algorithm are categories in agglomerative method. The Louvain and Leiden algorithm are based on modularity and hierarchical clustering. Modularity is a quality of function and it is denoted as $Q$ [4], [9]. Modularity function can measure the strength of communities. It is also used to mark the line of the dendogram and end the algorithm that shows the values of good partitions.

The larger values of modularity show the better communities. The values of modularity are always smaller than 1. When the value of modularity is less than 1, it shows that each node is a community itself. Subgraphs existence with small internal number of edges and many inter-community edges.

Definition of modularity,

$$Q = \frac{1}{2m} \sum_{i,j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j)$$  \hspace{1cm} (1)
Where

\[ A \] is the adjacency matrix
\[ k_i \] and \[ k_j \] are the degrees of nodes \( i \) and \( j \) respectively
\[ m \] is the number of edges
\[ C_i \] is the community of node \( i \)
\[ \delta(\square) \] is the Kronecker function: 1 if both nodes \( i \) and \( j \) belong on the same community \( \{C_i = C_j\} \),
0 otherwise [4], [9]

Hierarchical clustering, as the name suggests, is an algorithm that builds hierarchy of clusters. This algorithm starts with all the data points assigned to a cluster of their own. Then two nearest clusters are merged into the same cluster. In the end, this algorithm terminates when there is only a single cluster left. The results of hierarchical clustering can be shown using dendrogram (refer Figure 1). These algorithms have different definitions of community [10].

2.1. Louvain Algorithm
In 2008, Blondel et al. proposed the Louvain algorithm based on modularity optimization [11]. This algorithm consists of two phases. The first phase is the modularity optimization process (local moving of nodes) and the second phase is the community aggregation process (aggregation of network). Louvain provides a fast algorithm only for static network [12].

2.2. Leiden Algorithm
In 2019, Vincent Traag et al. proposed the Leiden algorithm based on improvement of the Louvain algorithm [13]. This algorithm is more complicated than the Louvain algorithm and have an accurate and fast result for the computation time. The Leiden algorithm consists of three phases. The first phase is the modularity optimization process, the second phase is the refinement of partition, and the third phase is the community aggregation process [13]. This algorithm performs well on small, medium and large-scale networks.

2.3. Comparison between Louvain and Leiden Algorithm
The difference between the Louvain and Leiden algorithm is that after the algorithm completes the first stage modularity optimization process, the refinement of partition is performed in each generated small community. Then, the community achieves a local optimum. Figure 2 shows the comparison of infographic for the Louvain and Leiden algorithm.
3. Conclusions
This paper shows the comparison between the Louvain and Leiden algorithm from several important detail of the algorithm. The basic flow of the algorithm and the scope of application of the algorithm are clarified. A few researchers have proven that the Leiden algorithm is more stable in detecting community and have efficient computation time compared to the Louvain algorithm. The Leiden algorithm is

![Figure 2. Infographic of Louvain and Leiden algorithm [13].](image)

|          | Louvain          | Leiden           |
|----------|------------------|------------------|
| Year     | 2008             | 2019             |
| Proposed by | Blondel et al.  | Vincent Traag et al. |
| Phase    | 2                | 3                |
| function in pseudocode | 4              | 0                |
| Disadvantage(s) | Communities are badly connected and disconnected especially when running the algorithm iteratively [13]. | Extra step in the process, refinement of partition. |
| Advantage(s) | Performs well on small and medium-scale networks. | Through merges process but also be able to use split process |
|           | Static networks. | Guaranteed the community are well-connected when running the algorithm iteratively. |
|           |                  | Higher quality function and take less time to run. |
|           |                  | Performs well on small, medium and large-scale networks. |
|           |                  | Temporal networks. |
typically faster than the Louvain algorithm and returns partitions of a higher quality cluster. The Leiden algorithm also performs well on small, medium and large-scale networks. Finally, the Leiden algorithm’s property is considered the latest and fastest algorithm among others of community detection algorithm. In future maybe researcher can go deep to improvise for the best techniques of identification of meaningful communities based on Leiden algorithm.

References
[1] P. Holme and J. Saramäki 2011, *Phys. Rep.* 519 97–125
[2] M. Newman, *Networks.* 2010.
[3] Z. A. Abas *et al.,* 2020 *COMPUSOFT,* 9
[4] M. E. J. Newman and M. Girvan, *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* 69 1–15
[5] D. K. Singh, R. A. Haraty, N. C. Debnath, and P. Choudhury 2020 *Proc. IEEE Int. Conf. Ind. Technol.* 2020 989–994
[6] J. Zhang, Z. Ma, Q. Sun, and J. Yan, *J. Phys. Conf. Ser.* 1069
[7] R. Cazabet and G. Rossetti 2019 *Temporal Network Theory* 181–197
[8] R. Cazabet, S. Boudebza, and G. Rossetti 2020 *Journal of Complex Network*
[9] M. E. J. Newman 2006 *Proc. Natl. Acad. Sci. U.S. A.* 103 8577–8582
[10] Z. Yang, R. Algesheimer, and C. J. Tessone 2016, *Scientific Reports,* 6
[11] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre 2008 *J. Stat. Mech. Theory Exp.* 10 1–12
[12] R. Badlani, K. Culberg, and Z. Jiang 2018, “Community Detection and Evolution in Temporal Networks,” 2018 [Online].
[13] V. A. Traag, L. Waltman, and N. J. van Eck *Sci. Rep.* 9 1–12

Acknowledgments
Appreciation to the Ministry of Education Malaysia for research grant FRGS/2018/FTMK-CACT/F00394, Malaysia Research Assessment (MyRA), Universiti Teknikal Malaysia Melaka (UTeM), Fakulti Teknologi Maklumat dan Komunikasi (FTMK) and research group of Centre for Advanced Computing Technology (C-ACT) – Optimisation, Modeling, Analytics and Simulation (OptiMAS) research group.