Clustering coefficient and path length in ring-lattices

Paul McCarthy

Abstract

We provide equations to calculate the mean clustering coefficient and characteristic path length of an undirected ring-lattice, of even degree $k > 2$ and order $v > k + 1$.

1 Preliminaries

Throughout this paper, we will use the notation $G = (V, E)$ to denote a graph G with vertex set V and edge set E. The symbol v is used to represent the order of G (i.e. $v = |V|$). For a node $i \in V$, the set of nodes adjacent to i is represented by N_i; the symbol n denotes the size of N_i (i.e. $n = |N_i|$, the degree of node i). Finally, the symbol k is used to represent the degree of each node in a ring-lattice.

1.1 Ring-lattice

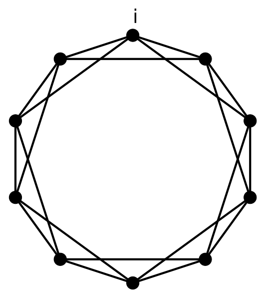


Figure 1: 4-regular ring-lattice of order 10

A *k-regular* graph is a graph in which every node has the same degree k ([1]). A *ring-lattice* is a k-regular graph in which the nodes form a ring, and each node is connected to its k nearest neighbours ($k/2$ on each side). An example of a ring-lattice is given in Figure 1.

1.2 Clustering coefficient

The clustering coefficient of a node is the ratio of the number of edges which are present between a node's neighbours to the number of possible edges between the neighbours ([3]). In other words, the clustering coefficient of a node is the density of the subgraph formed by the node's immediate neighbours and the edges which exist between them. In a graph $G = (V, E)$ of order v, the clustering coefficient c of a node $i \in V$ with n neighbours N_i may be calculated as follows:

$$c_i = \frac{2|\{(j,k) : j \neq k \in N_i, (j,k) \in E\}|}{n(n-1)} \tag{1}$$

The mean clustering coefficient is then:

$$C_G = \frac{\sum_{i \in V} c_i}{v} \tag{2}$$

In a ring-lattice, every node is identical, thus $c_i = C_G$.

1.3 Characteristic path length

The characteristic (mean) path length of a graph is the average shortest path length between all pairs of nodes ([3]). If l_{ij} denotes the minimum number of edges which must be traversed to move between nodes i and j, the characteristic path length of the graph is then:

$$L_G = \frac{\sum_{i \neq j \in V} l_{ij}}{v(v-1)} \tag{3}$$

As with the clustering coefficent, the mean path length of any one node is equivalent to that of the entire ring-lattice.

2 Derivation of clustering coefficient

We need to calculate the number of edges which exist in a subgraph formed by the neighbours of a node i, as shown in Figure 2. Initially, we restrict the derivation to lattices of order $> 3k/2$. We address this restriction in Section 2.2.

All regular graphs of degree k and order $v{:}v \leq k+1$ are complete, and thus have mean clustering coefficient 1.0. Additionally, all ring-lattices of degree 2 have mean clustering coefficient 0.0. These cases are therefore not covered by this derivation.

2.1 For order $> 3k/2$

We begin with the subgraph formed by node i and all of its neighbours. The maximum possible number of edges in this subgraph, as shown in Figure 3, is:

$$a = \frac{k(k+1)}{2} \tag{4}$$

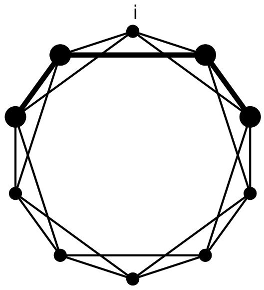


Figure 2: The subgraph formed by the neighbours of i

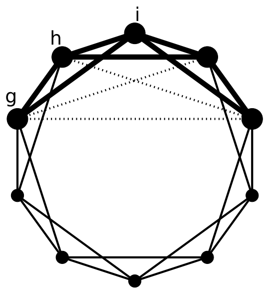


Figure 3: Maximum possible number of edges in the subgraph formed by i and its neighbours

We subtract from this the number of edges in the subgraph which do not exist (the dotted edges in Figure 3). For each node $j \neq i$ in the subgraph, the number of 'missing' edges is equal to the number of nodes on the opposite side of i, minus any nodes to which j is adjacent. For instance, node g in Figure 3 is missing two edges, and node h is missing one edge. As there are $k/2$ such nodes on each side of i, the sum of these edges becomes the sum of the sequence $(1, 2, ..., \frac{k}{2} - 1, \frac{k}{2})$. The total number of missing edges is therefore given by:

$$b = \frac{\frac{k}{2}(\frac{k}{2} + 1)}{2} \tag{5}$$

Finally, we remove node i, and the edges incident upon node i, from the subgraph, arriving at $a - b - k$, or:

$$e = \frac{k(k+1)}{2} - \frac{\frac{k}{2}(\frac{k}{2} + 1)}{2} - k \tag{6}$$

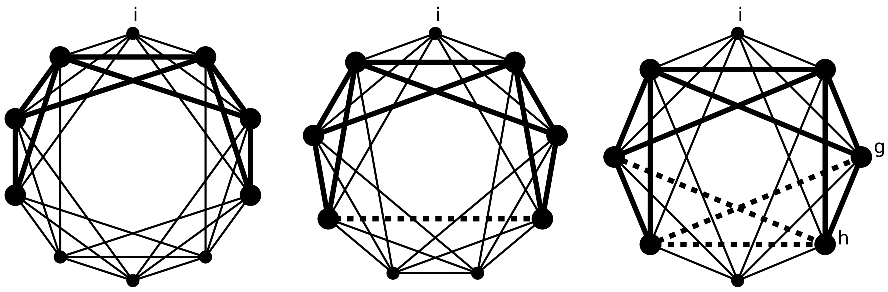
Reducing Equations (6) and (1) gives us the clustering coefficient of node i, and of the entire graph:

$$c_i = C_G = \frac{3(k-2)}{4(k-1)} \tag{7}$$

2.2 For order $\leq 3k/2$

Equation (7) is not satisfactory for ring-lattice graphs of order less than or equal to $3k/2$, as demonstrated in Figure 4.

We must account for the edges which are introduced as a result of the smaller gap across the remainder of the graph (the dotted edges in Figures 4b and 4c).



(a) Order 10 **(b)** Order 9 **(c)** Order 8

Figure 4: 6-regular ring-lattices of order 10, 9, and 8. When the order of the graph becomes $\leq 3k/2$ (9 in this example), the number of edges in the subgraph formed by node i's neighbours increases.

Using a similar approach as in Equation (5), we count the number of edges which have been added to the subgraph. In Figure 4c, we can see that one edge is added for node g, and two edges are added for node h. This sequence is obviously dependent upon the number of nodes in the 'gap' set, that is, the nodes over which the new edges must traverse. The size of this gap set, easily calculated by subtracting the size of the subgraph formed by node i and its neighbours, is $v - k - 1$.

The number of edges which must be added is then given by the sum of the sequence $(1, 2, ..., \frac{k}{2} - (v - k - 1))$:

$$\omega = \frac{(\frac{3k}{2} - v + 2)(\frac{3k}{2} - v + 1)}{2} \tag{8}$$

2.3 Calculating the clustering coefficient

Taking Equations (6) and (8), we are able to calculate the number of edges e which exist in a subgraph formed by the n neighbours N of a node i in a ring-lattice of order v and degree k:

$$\begin{aligned} e_{kv} &= \frac{k(k+1)}{2} - \frac{\frac{k}{2}(\frac{k}{2}+1)}{2} - k + \frac{\omega}{2}, \quad \text{where} \\ \omega &= \begin{cases} 0, & \text{if } v > \frac{3k}{2} \\ (\frac{3k}{2} - v + 2)(\frac{3k}{2} - v + 1), & \text{otherwise} \end{cases} \end{aligned} \tag{9}$$

The clustering coefficient of node i, and of the entire graph, is then given by:

$$c_i = C_G = \frac{2e_{kv}}{k(k-1)} \tag{10}$$

3 Derivation of path length

We need to calculate the sum of the distances from a node i to all other nodes in the lattice. A ring-lattice of degree $k \geq 2$ and order $v = k + 1$ is complete – the distance from node i to all other nodes is 1. If a single node is added to such a graph, the graph is no longer complete, and the distance from i to the new node is 2.

We can add up to $k - 1$ more nodes to this graph (partition 2 in Figure 5), each of which will have a distance of 2 from i. If we add another k nodes (partition 3 in Figure 5), the distance from i to these new nodes will be 3, and the distance from i to yet another node (partition 4 in Figure 5) will be 4.

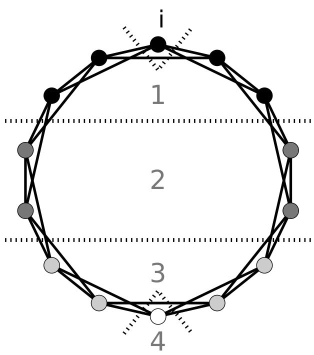


Figure 5: 4-regular ring-lattice of order 14

We may calculate the total sum of all distances by finding the sum of distances from node i to the nodes in each of the complete 'sets' (partitions $1-3$ in Figure 5), and the sum of distances from node i to the nodes in the remaining incomplete set (partition 4 in Figure 5).

It is easy to calculate the number of complete sets in the graph:

$$s = \lfloor \frac{v-1}{k} \rfloor \tag{11}$$

Then, the sum of distances from node i to the nodes in the complete sets is $\sum_{j=1}^{s} kj$, or:

$$a = \frac{ks(s+1)}{2} \tag{12}$$

Given that the number of nodes remaining in the incomplete set is $v - 1 - ks$, the sum of distances from node i to those nodes in the incomplete set is now:

$$b = (v - 1 - ks)(s + 1) \tag{13}$$

Now, $a + b$ gives us the sum of distances from node i to all other nodes.

3.1 Calculating the path length

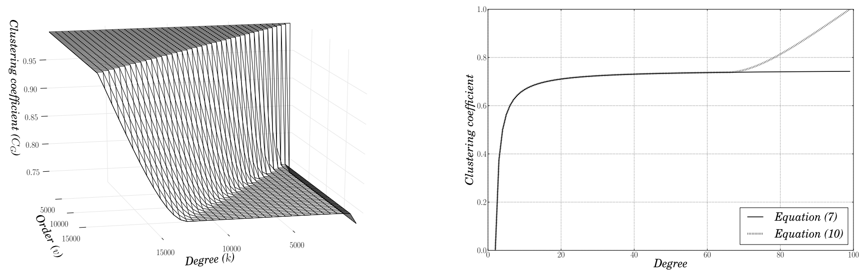
Rearranging Equations (12) and (13), we arrive at the total sum of distances d from a node i to all other nodes in a ring-lattice of degree k and order v:

$$\begin{aligned} d_{kv} &= (s+1)(v-1-\frac{ks}{2}), \quad \text{where} \\ s &= \lfloor \frac{v-1}{k} \rfloor \end{aligned} \tag{14}$$

The characteristic path length of node i, and of the entire graph, is now:

$$l_i = L_G = \frac{d_{kv}}{v-1} \tag{15}$$

4 Clustering coefficient properties



(a) Clustering coefficient over a range of v and k

(b) Clustering coefficient for a ring-lattice of order 100, and degree in the range $2 - 100$

Figure 6: Clustering coefficient properties

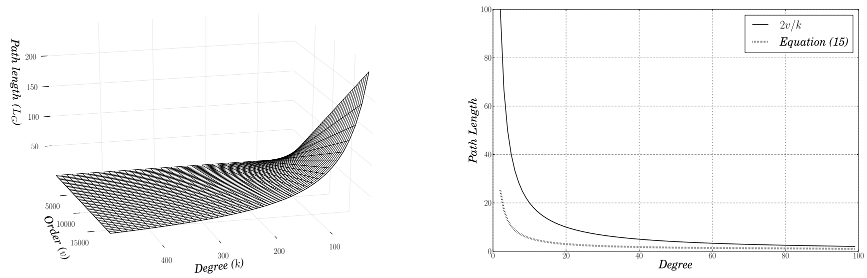
The shape of Equation (10) is shown in Figure 6a. For cases where $v \gg k$, the clustering coefficient asymptotically approaches 0.75. For ring-lattices where $k \leq v \leq 3k/2$, the clustering coefficient increases exponentially until reaching its maximum possible value of 1.0, when $v \leq k+1$.

Gros ([2]) provides Equation (7) for calculation of the clustering coefficient in a ring-lattice, with the exception that it fails to take into account the case where $v <= 3k/2$. This is demonstrated in Figure 6b.

5 Path length properties

The shape of Equation (15) is shown in Figure 7a. The path length of a ring-lattice decreases exponentially as k approaches v, reaching it's minimum value of 1.0 when $v \leq k + 1$. For a fixed k, the path length of a ring-lattice increases linearly with an increase in v.

Watts and Strogatz ([3]) provide an estimate for the path length in a ring-lattice of $2v/k$. A comparison of this estimate with Equation (15) is given in Figure 7b.



(a) Path length over a range of v and k

(b) Characteristic path length for a ring-lattice of order 100, and degree in the range $2 - 100$

Figure 7: Path length properties

6 Acknowledgments

The author would like to thank Dr. Lubica Benuskova for invaluable discussion and advice during the development of this paper. This research was funded by an University of Otago PhD scholarship.

References

[1] Gary Chartrand. *Introductory Graph Theory.* Dover Publications, New York, 1977.

[2] Claudius Gros. *Complex and Adaptive Dynamical Systems: A Primer*. Complexity. Springer, 1st edition, 2008.

[3] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.