International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:7, 2017

Generalization of Clustering Coefficient on Lattice Networks Applied to Criminal Networks

Christian H. Sanabria-Montaña, Rodrigo Huerta-Quintanilla

Abstract**—**A lattice network is a special type of network in which all nodes have the same number of links, and its boundary conditions are periodic. The most basic lattice network is the ring, a one-dimensional network with periodic border conditions. In contrast, the Cartesian product of d rings forms a d-dimensional lattice network. An analytical expression currently exists for the clustering coefficient in this type of network, but the theoretical value is valid only up to certain connectivity value; in other words, the analytical expression is incomplete. Here we obtain analytically the clustering coefficient expression in d-dimensional lattice networks for any link density. Our analytical results show that the clustering coefficient for a lattice network with density of links that tend to 1, leads to the value of the clustering coefficient of a fully connected network. We developed a model on criminology in which the generalized clustering coefficient expression is applied. The model states that delinquents learn the know-how of crime business by sharing knowledge, directly or indirectly, with their friends of the gang. This generalization shed light on the network properties, which is important to develop new models in different fields where network structure plays an important role in the system dynamic, such as criminology, evolutionary game theory, econophysics, among others.

Keywords**—**Clustering coefficient, criminology, generalized, regular network d-dimensional.

I. Introduction

LATTICE NETWORKS are those in which all nodes have the same number of links, and its boundary conditions are periodic. They were the first networks studied in the field of complex networks and are currently used in different areas within physics [1]-[4]. The clustering coefficient, a measurement of how many triangles are formed in a network, has been analytically calculated for this type of network [5]-[7]. However, this theoretical value is only valid below a certain link density. Vega-Redondo [5] stated that in a one-dimensional case the clustering coefficient is valid for average connectivity values (z) less than $2N/3$; we demonstrate that for a lattice network of d dimension the z value is $2dN^{1/d}/3$. This means that the clustering coefficient analytical function is not a generalized function for any link density, and is therefore incomplete. In this study, we demonstrate a generalization of the clustering coefficient for any link density, and lattice networks of any dimension. This fills a void in current theory of the clustering coefficient in lattice networks. First, we demonstrate the generalization for a one-dimensional lattice network, and second for a d-dimensional lattice network.

C. Sanabria-Montaña is with ECOPETROL S.A. Carrera 13 No 36-36, Bogotá, Colombia (e-mail: Christian.Sanabria@ecopetrol.com.co).

R. Huerta-Quintanilla is with Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Departamento de Física Aplicada Km 6 carretera antigua a progreso, Mérida, Yucatán, México (e-mail: rhuerta@mda.cinvestav.mx).

The clustering coefficient is a key property to understand tightly knit groups in the network. In consequence this concept can be specially useful to understand the criminal network structure. Calvo-Armengo and Zenou proposed a model [8] in which the expertise of criminals is shared due to interactions among gang members. In this work we modified this model using the clustering coefficient to characterize the indirect transfer of crime knowledge among the delinquents. Additionally, clustering coefficient is used to relate the probability of being caught with the gang structure [9].

II. Clustering Coefficient on Lattice Networks

A. One-Dimensional Lattice Network

In a ring or one-dimensional lattice network, the nodes are arranged in a circular configuration and each has $z = 2r$ links, which are linked to their r nearest neighbors (Fig. 1 (a)).We calculated the local clustering coefficient (C_i) for node i using the general formula:

$$C_i = \frac{\lambda_G(i)}{\tau_G(i)}, \tag{1}$$

where $\lambda_G(i)$ is the number of triangles over i, that is, the number of subnetworks with 3 links and 3 nodes, one of which is i. $\tau_G(i)$ is the number of triples in i, i.e. the number of subnetworks with 2 links and 3 nodes, one of which is i and the other two of which are connected to i. Once the local clustering coefficient is calculated, the clustering coefficient for the whole network (C) is calculated as the average over the local clustering coefficients:

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{2}$$

For lattice networks, the local clustering coefficient is the same as the network average clustering coefficient ($C_i = C$, $\forall i$). Therefore, $\lambda_G(i)$ and $\tau_G(i)$ must only be calculated for a single node, which we did at different connectivity values. In a one-dimensional lattice network, the $\tau_G(i)$ value is the combination $C(z, 2)$:

$$\tau_G(i) = C(z, 2) = \frac{z(z-1)}{2}. \tag{3}$$

By keeping the number of nodes in network N fixed, and counting the triangles in node i for different connectivity

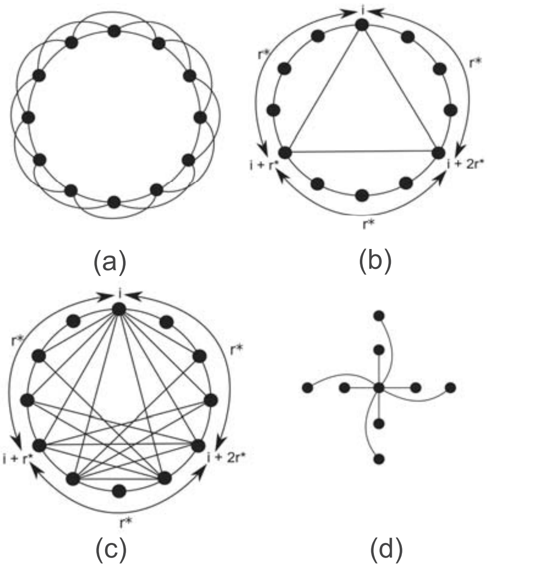

Fig. 1 Schematic illustration of one-dimensional (ring) and two-dimensional (square) lattice networks. (a) 12-node, one-dimensional network with a four-link connectivity of $z = 2r$; that is, each node links to its two first neighbors $r = 2$. (b) Triangle formed by a one-dimensional lattice network when $z = z^* = 2N/3$ (in this example, the network consists of 12 nodes, therefore $z^* = 8$, $r^* = 4$).At this z^* value the theoretical clustering coefficient is no longer valid, thus, this calculation does not consider this triangle. (c) Triangles not counted (total = 10) when $z = z^* + 2$. (d) Link arrangement for one node in a two-dimensional lattice network with N nodes and 8 links, 4 per ring. This network is the Cartesian product of two one-dimensional (or ring) lattice networks with $N^{1/2}$ nodes. Generally, a d-dimensional network is generated by the Cartesian product of d rings

values (starting with $z = 4$, then $z = 6$, $z = 8$ and successively until $z < 2N/3$), it can be shown simply that $\lambda_G(i)$ is:

$$\lambda_G(i) = 3\sum_{j=0}^{r-1} j = \frac{3r(r-1)}{2}, \tag{4}$$

Taking into account that $z = 2r$, we can rewrite (4) as

$$\lambda_G(i) = \frac{3}{8}z(z-2). \tag{5}$$

The local clustering coefficient, and therefore the network average clustering coefficient, is:

$$C_i = C = \frac{3(z-2)}{4(z-1)}, \tag{6}$$

which is well known as the clustering coefficient found in the current literature [4]-[7]. As mentioned above, this clustering coefficient is only valid for $z < 2N/3$, since in $z = z^* = 2N/3$ (5) does not count the triangle formed by nodes $(i, i + r^*, i + 2r^*)$ (Fig. 1 (b)). In terms of distance between nodes, denoted here as $d(i, j)$, (6) ceases to be valid when the distance between node i and node $i + 2r^*$ satisfies.

$$d(i, i + 2r^*) = 2r^* = N - r^*, \quad \text{then} \quad r^* = \frac{N}{3}. \tag{7}$$

Therefore, beginning with z^* and onwards, we must add the triangles not accounted for in (5) to that equation. For example, in z^* the triangle formed by the nodes $(i, i+r^*, i+2r^*)$ is not included, while in $z^* + 2$ the 10 triangles formed by a series of nodes - $\{(i, i+r^*-2, i+2r^*-1); (i, i+r^*-1, i+2r^*-1); (i, i+r^*-1, i+2r^*); (i, i+r^*, i+2r^*-1); (i, i+r^*, i+2r^*); (i, i+r^*, i+2r^*+1); (i, i+r^*+1, i+2r^*-1); (i, i+r^*+1, i+2r^*); (i, i+r^*+1, i+2r^*+1) and (i, i+r^*+1, i+2r^*+2)\}$ - are not counted (Fig. 1 (c)). The triangles not counted by (5), denoted as $\lambda^c_G(i)$, take the following form:

$$\lambda^c_G(i) = 1 + 9\sum_{j=0}^{r'} j, \quad \text{con} \quad r' = r - r^*. \tag{8}$$

Calculating the sum and replacing both r and r^*, results in

$$\lambda^c_G(i) = 1 + \frac{9}{2}\left(\frac{z}{2} - \frac{N}{3}\right)\left(\frac{z}{2} - \frac{N}{3} + 1\right), \tag{9}$$

This in turn leads to reformulation of (5) as

$$\lambda_G(i) = \frac{3}{8}z(z-2) + H_{z^*}(z)\left[1 + \frac{9}{8}(z - z^*)(z - z^* + 2)\right], \tag{10}$$

where $H_{z^*}(z)$ is the Heaviside step function. Therefore the generalized clustering coefficient for any link density is:

$$C = \frac{3(z-2)}{4(z-1)} + \frac{2H_{z^*}(z)}{z(z-1)}\left[1 + \frac{9}{8}(z - z^*)(z - z^* + 2)\right]. \tag{11}$$

It is easy to test that when $z \to N$ the clustering coefficient is $C \to 1$. This makes sense since the network is becoming a fully connected network and the clustering coefficient of a fully connected network is 1.

B. The d-Dimensional Lattice Network

To generalize the clustering coefficient to d dimensions, we must first define a d-dimensional lattice network. In general terms, a d-dimensional lattice network is the Cartesian product of d ring-type one-dimensional networks, which form toroidal topologies. The total number of nodes in the network is the product of the number of nodes in each of the rings:

$$N = \prod_{n=1}^{d} L_n, \tag{12}$$

where L_n is the number of nodes of the n-th ring. Also, connectivity z is defined as the sum of the connectivities of each ring:

$$z = \sum_{n=1}^{d} z_n, \quad \text{by definition } z_n = 2r_n, \text{ then} \quad z = 2\sum_{n=1}^{d} r_n. \tag{13}$$

In the example provided here (Fig. 1 (d)), we show the links of one node from a two-dimensional network with 8 links where each ring has 4 links ($z_1 = z_2 = 4$). Each node links to its second neighbors in each ring ($r_1 = r_2 = 2$). For this type of network, the clustering coefficient is the average of the clustering coefficients of the d rings:

$$C = \frac{1}{d}\sum_{n=1}^{d} C_n, \tag{14}$$

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:7, 2017

where C_n is the clustering coefficient of the n-th ring, as follows:

$$C_n = \frac{3(z_n-2)}{4(z_n-1)} + \frac{2H_{z_n^*}(z_n)}{z_n(z_n-1)}\left[1+\frac{9}{8}(z_n-z_n^*)(z_n-z_n^*+2)\right], \quad (15)$$

that is, the coefficient calculated in (11), but replacing $z \to z_n$ and $z^* \to z_n^*$.

In a particular case, the d rings have the same number of nodes (i.e. $N = L^d$) and each ring links to r neighbors, then $z = 2rd$, where d is the network dimension. Therefore, the network's clustering coefficient is the same as in each ring (i.e. $C = C_n\ \forall n$). We then replace $z_n \to z/d$ and $z_n^* \to z^*/d$ in (15) to produce the clustering coefficient:

$$C = \frac{3(z-2d)}{4(z-d)} + \frac{2d^2 H_{z^*}(z)}{z(z-d)}\left[1+\frac{9}{8d^2}(z-z^*)(z-z^*+2d)\right], \quad (16)$$

where $z^* = 2dN^{1/d}/3$. The first term in the above equation is the clustering coefficient of a d-dimensional lattice network, with $z = 2rd$. This can be found in the current literature [5]-[7] and is valid for values of $z < z^*$.

The maximum link value for a node in this d-dimensional lattice network (called z_{max}) is the sum of the maximum link value of this node in each of the d rings (z_n^{max}). If all the rings have the same number of nodes ($N^{1/d}$), then $z_n^{max} = (N^{1/d}-1)$ and therefore $z_{max} = d(N^{1/d}-1)$. It is important to note that due to the lattice network's topology the maximum number of links differs from $N-1$, which would be the number of links in a fully connected network, except for a one-dimensional case ($d = 1$).

The link density in this type of network is $D = z/(d(N^{1/d}-1))$ and this density tends toward 1 when $z \to z_{max}$. In (16), when the link density tends toward 1 (i.e. $z \to z_{max}$) the clustering coefficient also tends toward 1 ($C \to 1$).

To test our equation we simulated lattice networks with N nodes and different dimensions that begin with a minimum amount of links per node (z_0). Keeping N fixed, we calculated the clustering coefficient. We then increased the number of links per node by Δz and calculated the clustering coefficient after each increase. This continues until the number of links per node nears $d(N^{1/d}-1)$, which would be a fully connected, d-dimensional lattice network. In the resulting plot for one-dimensional and two-dimensional networks with $N = 7056$ nodes (Fig. 2), the points represent the clustering coefficient values calculated from the simulations and the continuous line shows the clustering coefficient generated by (16) using $d = 1$ and $d = 2$. The clustering coefficient clearly tends towards 1 more rapidly in the two-dimensional network than in the one-dimensional network; this occurs because the $z^* = 2N/3$ value in the one-dimensional network is 4074 while in the two-dimensional network the $z^* = 4N^{1/2}/3$ value is 112.

III. Application to Criminology

We applied this clustering generalization in criminology. Calvo-Armengo and Zenou developed a model [8], based on game theory and social networks, in which the delinquents learn the know-how of crime from experiences shared by criminals that belong to the same gang. The model uses regular networks. The criminal's expected payoff for each agent is a function of crime level e, marginal expected punishment cost ϕ and network's degree k.

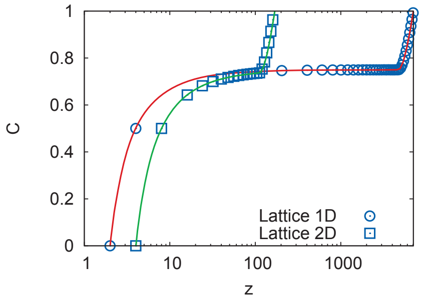

Fig. 2 Computational calculation of clustering coefficient (C) as a function of network connectivity (z) for one-dimensional and two-dimensional networks with $N = 7056$ nodes. The circles represent the one-dimensional (ring) network clustering coefficient and the squares represent the two-dimensional (square) network clustering coefficient. The continuous red line is the analytical expression generated by (16) for $d = 1$, and the continuous green line is the same analytical expression but for $d = 2$

Individuals may either be criminals or participate in labor market. Individuals in the labor force earn a wage w, while those involved in criminal market receive an expected payoff d_i given by

$$d_i(e,\phi) = e_i\left(1-\sum_{j\in N} e_j\right) - \phi e_i\left(1-\sum_{j\in N} a_{ij}e_j\right), \quad (17)$$

where a_{ij} is an element of the adjacency matrix, which takes the value of 1 if there is a link between i and j, otherwise $a_{ij} = 0$. Each agent seeks to maximize her payoff, then the expected payoff of agent i is $\pi_i(d_i, w) = \max\{d_i, w\}$.

We have modified the model proposed by Calvo-Armengo and Zenou to incorporate two facts. First, the know-how of crime is transmitted not only directly, but also indirectly. The model stated that know-how sharing is possible only if the two person know each other. Nevertheless, members of the same gang can share their experience indirectly through a third person. Second, the gang has a structure which is composed by a hard core and a clique structure. The hard core structure includes the individuals who are culturally and criminally enmeshed in the gang and are at risk of being so for life, the clique structure considers individuals who gravitate around one or more of the hard core agents [9]. Therefore the structure is related to the marginal expected punishment cost. Individuals into the hard core have a minor probability of being caught than the others which are in the clique structure.

In consequence (17) is modified to introduce a term that describes the indirect sharing of knowledge among the gang members. In addition, we have expressed ϕ in terms of the clustering coefficient as $\phi(C) = 1 - C$. When the network has a high connectivity (hard core) the probability of being caught is lower than a network with a low connectivity (clique).

Individual payoff is written as:

$$d_i(e,\phi) = e_i\left(1-\sum_{j\in N} e_j\right) - \phi(C)e_i\left(1-\sum_{j\in N}(a_{ij}+(1-a_{ij})C)e_j\right). \tag{18}$$

From (18), the crime level at equilibrium e_i^* such that $\pi_i(0,w) = \max\{0,w\}$ is

$$e_i^* = \frac{1-\phi(C)}{N-\phi k+\phi(C)C(N-k)}, \tag{19}$$

where k is network's degree. From (19) e_i^* increases with network's degree, which is comprehensible since high connectivity corresponds to a hard core structure and a low probability of being caught ϕ. Fig. 3 shows the crime level as a function of ϕ for a network with 512 nodes. The continuous line is crime level at equilibrium given by (19). This line divides the labor market from the criminal market phase. That means a value of e above of e^* makes all agents get into labor market. As ϕ increases, the crime level decreases in such a way that the best option for the agent is to choose the labor force.

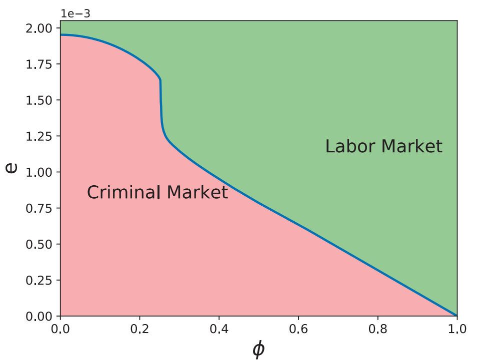

Fig. 3 Crime level e as a function of marginal expected punishment cost ϕ for a network with N = 512 nodes. The continuous line is the analytical expression generated by (19), which separates the space in two phases: criminal and labor market

IV. Conclusion

We developed a general formula to calculate the clustering coefficient for d-dimensional lattice networks with any link density. This fills a theoretical void because the current clustering coefficient theory for d-dimensional lattice networks is valid only to a connectivity value less than $2dN^{1/d}/3$. This generalization allows us to analytically demonstrate the transition from a lattice network to a fully connected network, at least in terms of the clustering coefficient. When the link density tends to 1 the clustering coefficient produced with the generalized formula also tends to 1. This is not possible using current theory since as the link density tends to 1 the clustering coefficient tends toward $3/4$.

We applied this generalization to criminology where we have modified the model proposed by Calvo-Armengo and Zenou introducing the clustering coefficient to describe not only the direct but also the indirect sharing of the know-how of crime experience and the relation of the probability of being caught with the gang structure. Our model shows that the crime level at equilibrium, which separates the labor market from criminal market phase, increases as the marginal expected punishment cost decreases. Hence individuals in the hard core structure have more probabilities to stay in criminal market that individuals in the clique structure.

This generalization of the clustering coefficient is applicable in several fields. For example, in econophysics, this generalization can be used to propose a model to understand how network structure affects the money condensation of a system. Under certain economic exchange conditions among a population of agents, one of these agents attains control of all the money in the system, a phenomenon known as wealth condensation [10]. However, when the same population is located in a spatial network under the same exchange conditions, this condensation does not occur [11].

On the other hand, in the field of evolutionary game theory [12], the players are located in a spatial structure (or network) and use a strategy (e.g. replicate, unconditional imitation, etc.) to update the action to take; to cooperate or to defect. In the prisoner dilemma, for example, it is known that under the replicator strategy the cooperator fraction is zero in a well-mixed population [13], that is, a fully connected network. However, this cooperator fraction differs from zero when players are located in a lattice network [13], [14]. This suggests a model in which the network can change its clustering coefficient within a one-dimensional lattice network, increasing its link density until it is fully connected, in such a way that the cooperator fraction varies in response to network structure modification.

Acknowledgment

We would like to thank Efrain Canto-Lugo for helping in the computer simulations. Also we thank Conacyt-Mexico for partially supporting this work.

References

[1] W. Li, A. Bashan, S. V. Buldyrev, H. E. Stanley, and S. Havlin, “Cascading failures in interdependent lattice networks: The critical role of the length of dependency links,” *Phys. Rev. Lett.*, vol. 108, p. 228702, May 2012. (Online). Available: http://link.aps.org/doi/10.1103/PhysRevLett.108.228702.

[2] M. N. Kuperman and S. Risau-Gusman, “Relationship between clustering coefficient and the success of cooperation in networks,” *Phys. Rev. E*, vol. 86, p. 016104, Jul 2012. (Online). Available: http://link.aps.org/doi/10.1103/PhysRevE.86.016104.

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports*, vol. 424, no. 4-5, pp. 175 – 308, 2006. (Online). Available: http://www.sciencedirect.com/science/article/pii/S037015730500462X.

[4] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Rev. Mod. Phys.*, vol. 74, pp. 47–97, Jan 2002. (Online). Available: http://link.aps.org/doi/10.1103/RevModPhys.74.47.

[5] F. Vega-Redondo, *Complex Social Networks*. Cambridge University Press, 2007.

[6] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princenton University Press, 1999.

[7] C. Gros, *Complex and Adaptive Dynamical Systems: A Primer*. Springer-Verlag, 2008.

[8] A. Calvó-Armengoi and Y. Zenou, "Social networks and crime decisions: The role of social structure in facilitating delinquent behavior," *International Economic Review*, vol. 45, no. 3, pp. 939–958, 2004. (Online). Available: http://www.jstor.org/stable/3663642.
[9] M. Carlie, "Into the abyss: A personal journey into the world of street gangs." (Online). Available: http://people.missouristate.edu/MichaelCarlie/site_map.htm.
[10] C. Moukarzel, S. Gonçalves, J. Iglesias, M. Rodríguez-Achach, and R. Huerta-Quintanilla, "Wealth condensation in a multiplicative random asset exchange model," *Eur. Phys. J. Special Topics*, vol. 143, pp. 75–79, 2007. (Online). Available: http://dx.doi.org/10.1140/epjst/e2007-00073-3.
[11] C. H. S. Montaña, R. Huerta-Quintanilla, and M. Rodríguez-Achach, "Class formation in a social network with asset exchange," *Physica A Statistical Mechanics and its Applications*, vol. 390, pp. 320–340, 2011. (Online). Available: http://www.sciencedirect.com/science/article/B6TVG-517J27V-4/2/318104b3a1d77c8a308251bdb8c5a1e6.
[12] G. Szabó and G. Fáth, "Evolutionary games on graphs," *Physics Reports*, vol. 446, no. 4?6, pp. 97 – 216, 2007. (Online). Available: http://www.sciencedirect.com/science/article/pii/S0370157307001810.
[13] C. P. Roca, J. A. Cuesta, and A. Sánchez, "Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics," *Physics of Life Reviews*, vol. 6, no. 4, pp. 208 – 249, 2009. (Online). Available: http://www.sciencedirect.com/science/article/pii/S1571064509000256.
[14] L. G. Moyano and A. Sánchez, "Evolving learning rules and emergence of cooperation in spatial prisoner's dilemma," *Journal of Theoretical Biology*, vol. 259, no. 1, pp. 84 – 95, 2009. (Online). Available: http://www.sciencedirect.com/science/article/pii/S0022519309000988.