Network Properties of a Typical Self-employment Agglomeration

Yong Tang ¹, Linjie He ^{2*}

¹ Management School, Hunan City University, Yiyang, 413000, CHINA ² Business School, Hunan Normal University, Changsha, Hunan, 410081, CHINA

Received 14 May 2017 • Revised 1 September 2017 • Accepted 27 September 2017

ABSTRACT

China's mass entrepreneurship and innovation policy have inspired many individuals to join incumbent agglomerations to supplement their resource insufficiency. Based on the ties of kinship and township relations among a typical agglomeration, this paper explored the individuals' preferential attachment rules and analyzes the network properties. The main findings included: self-employment agglomeration with given preferential attachment rules has a large average degree; this kind of network is scale-free, and it follows power-law degree distribution; clustering coefficient is very large; the average path length is short and it follows a power-law distribution. Our findings demonstrated that this typical self-employment agglomeration is a kind of complex network. We also discussed the mechanism of complex network's advantages on self-employment activities.

Keywords: self-employment, network properties, scale-free, complex network

INTRODUCTION

Many kinds of literature have discussed the mechanism of self-employment agglomeration in developing and developed countries, investigating the path of resource supply in a networked scope when every agent included needs more resource compensation from others (Barbieri, 2003; Gindling, 2014). From empirical studies, many scholars adopted Social Network Analysis (SNA) to test the resource acquisition among self-employment agglomeration (Elias, 2000).

By SNA, the structure analysis is also conducted as shown in **Table 1** and **Figure 2**. It's clear from **Table 1** that this selected agglomeration has an average degree of 6.728, a big number among most of business networks. We can also find out that the average distance is 3.271, which means any agent in the agglomeration will averagely need 3.271 steps to reach its targeted resource supplier. Particularly, when we conduct the clique analysis, we can see that this agglomeration has very amazing cohesion (cliques are 147 when setting a minimum number of agents as 3 and 13 when setting minimum as 4). For this reason, the clustering coefficient in this agglomeration is comparatively large (the coefficient is 0.699). Besides, **Figure 2** illustrates the cumulative degree distribution, and we can find that log(degree) and log(cumulative probability) are well exhibited the linear correlation. The amazing results make us ponder whether this kind of agglomeration has the attributes of complex network; if so, then what mechanism drives this special network a typical complex network.

SELF-EMPLOYMENT AGGLOMERATION NETWORK EVOLVING RULE

In China, the premise of people's business communication is based on "guanxi", the deep connections fostered by long-term daily communication and reciprocal interests (Peng et al., 2016). For the self-employers, they cannot afford the heavy burden to keep frequent communications only for satisfying business requirements. Accordingly, they will seek for help based on their existing relations formed by their prior township and kinship (Korur, Vargas, and Serrano, 2016; Tang, 2013; Kuah-Pearce, 2016). Under this relation structure, a new entrant will basically operate a business with his relatives and fellow villagers. Thus, in a self-employment agglomeration, every agent will select a fully connected incumbent group to form a new relationship network. Furthermore, self-employment is a typical micro-entrepreneurship, in which small-scale connections can mostly guarantee all resources supply. Based on the analysis above, we can conclude some characteristics of self-employment agglomeration:

Contribution of this paper to the literature

- It has complex network properties such as short average path length and power law distribution of degrees, which arouses our curiosity whether this kind of network is a typical complex network from its evolving mechanism.
- The main findings include: self-employment agglomeration with preferential attachment has a large average degree, and the frequency of agents' interaction is dense.
- The self-employment agglomeration under given rules is a very typical complex network.
- Under China's "guanxi" society, this kind of network has its advantage to be formed and evolved by
 necessity entrepreneurs and will be preferentially adopted by the government who advocates "mass
 entrepreneurship and innovation.

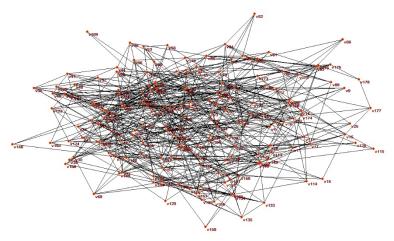


Figure 1. Topology of a Self-employment Agglomeration in China

Note: this agglomeration is located in Yiyang, Hunan Province, and has 180 businesses. This agglomeration is known as the "plastic package town" because most of businesses are package-related industries. All data were collected in 2014.

Serial number	Network Variables	Value
1	Average degree	6.728
2	Deg Centralization	0.040
3	Average length	3.271
4	Cliques (min 3)	147
5	Cliques (min 4)	13
6	Clustering coefficient	0.699

 Table 1. Overall Structures of Network

Firstly, self-employment is initiated and operated by individuals, but the agents cannot embrace all resources needed in running the complicated business; therefore they will seek for aids by entering into an incumbent group (agglomeration).

Secondly, self-employment is a kind of small-scale business, in which the resource support will arise from 3 to 4 fixed sources. For this reason, new entrant only needs to build 3 to 4 strong ties within an agglomeration.

And thirdly, the new entrants will mostly connect their peers by their kinship or fellow villagers, thus the connected ties are fully interacted prior to the new entrants.

For these reasons above, we can abstract some network rules below to reflect the process of a typical evolving self-employment agglomeration.

Step 1: at the earliest stage, there are only *s* incumbent self-employment businesses (we suppose $s \le 4$). Constrained by resource supply, these *s* businesses are resource suppliers to each other, i.e., *s* businesses are fully connected, and they form a complete network. Thus, *s* businesses are *s* vertices in the original network. This network has *s* vertices and $\frac{s (s-1)}{2}$ edges. See **Figure 3(a)**;

Step 2: new entrant will choose self-employment activity once it sees satisfactory resource supply in the network. Considering the mechanism of kinship and township relations, we suppose every new entrant will only choose the ones with whom it has a prior relation. Because of resource needs alike, it will only connect *s* incumbents. Thus after the second step, the original network has s + 1 vertices and $\frac{s (s-1)}{2} + s$ edges. See **Figure 3(b)**;

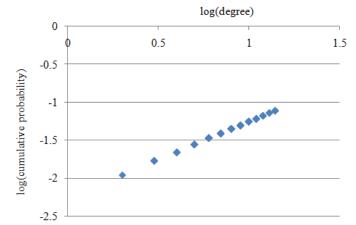


Figure 2. Cumulative Degree Distribution

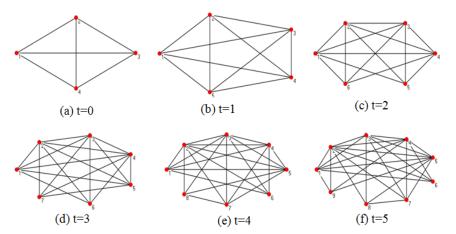


Figure 3. The Preferential Attachment and Network Evolution (s = 4)

Step 3: unlike traditional preferential attachments in BA network, the newly established self-employment will not build ties with high probability to the degreed-most incumbents. Alternatively, it will repeat the attachments to the most known incumbents. Thus, a new entrant will only choose the complete *s* -vertex sub-network in the global network and form new s + 1-vertex complete network. We can depict the evolution of network in **Figure 3** (a)-(f).

SELF-EMPLOYMENT AGGLOMERATION NETWORK PROPERTIES

Following the rules in section 2, we define a complete network G, in which *s* vertices are fully connected. Thus we have G = (V, E), |V| = s, $|E| = \frac{s(s-1)}{2}$. In order to explore its network properties, we need to calculate its average degree and average path, so that we can conclude the degree distribution and find out the clustering coefficient for this typical network. These properties are often adopted to exam whether this kind of network is a complex one (Albert and Barabási, 2002).

Average Degree

The average degree can basically describe the density of a given network, by which we can dig out whether this network is sparse or dense. Though degree distribution is uneven among all vertices in a network, it can basically tell the frequency of communication among all vertices. We firstly focus on the average degree aiming at depicting the wholeness of resource supply and demand for self-employment agglomeration network. As for an original fully connected network, its original average degree is $|\overline{E}| = 2\frac{\binom{s(s-1)}{2}}{s} = s - 1$. One more entrant will add 1 vertex and $\frac{s(s-1)}{2} + s$ edges to it, thus when the v - st new entrant enters into this network, the total vertex will be v + s, total

edges will be increased by *vs*, then we have an average degree of $|\bar{E}| = 2 \frac{\left(\frac{s(s-1)}{2}\right) + vs}{s+v}$. Note that in an agglomeration original network is much smaller than its mature network, thus *v* is much bigger than *s*,we can say in an agglomeration big enough, the average degree is 2 times of the connections every entrant has. In our agglomeration sample, the average degree is 6.728, basically 2 times of original 3- or 4-vertex complete network.

Degree Distribution

As shown in **Figure 2**, the sampled agglomeration has approximate power law distribution of degree. We now commit to modeling the agglomeration to prove this property.

We adopt mean-field theory to figure out the relations of vertex v and its degree k_v (Barabási et al., 1999). Before v enters into a network, a total number of s – complete network is $sk_v - s^2 + 2s - k_v$. When v enters into this network, the number of s – complete networks which includes the vertex of v is st + 1. We introduce time step t to record every newly added entrant, then we can construct a differential equation as (1):

$$\frac{\partial k_v}{\partial t} = \frac{sk_v - s^2 + 2s - k_v}{st + 1} \tag{1}$$

The first entrant has a degree of *s*, and the entrant of *v* will also have the degree of *s*, and then we have the initial conditions of $k_1(t_1) = s$ and $k_v(t_v) = s$ for the equation (1). Clearly, the solution to (1) can be expressed as (2):

$$k_{\nu}(t) = (s^{2} - 2s)\frac{1}{s - 1} + \frac{s}{s - 1} \left(\frac{st + 1}{st_{\nu} + 1}\right)^{1 - \frac{1}{s}}, t \ge t_{\nu}$$
⁽²⁾

From (2) we can say that the degree of any vertex *k* is the power exponent function of *t*, and power exponent is $1 - \frac{1}{s}$.

Similar to BA network, we construct a power function F(z) with the coefficient of P(k), and we define:

$$F(z) = \sum_{k=s}^{\infty} P(k)(s^2 - 2s) \frac{1}{s-1} + \frac{s}{s-1} \left(\frac{st+1}{st_v+1}\right)^{1-\frac{1}{s}}$$
(3)

Referring to Zhang et al. (2005), we obtain the cumulative degree probability function (4):

$$P(k_{\nu}(t) < k) = P\left(t_{\nu} > \frac{st+1}{s\left(1-\frac{1}{s}\right)^{s/(s-1)}\left(k-\frac{s(s-2)}{s-1}\right)^{s/(s-1)}} - \frac{1}{s}\right)$$
(4)

For the reason of uniformly entering into the incumbent network, every vertex follows uniform distribution. Then, we can construct the probability function (5).

$$P_{\nu}(t_{\nu}) = 1/(s+t)$$
 (5)

Combining function (4) and function (5), we obtain:

$$P(k_{v}(t) < k) = 1 - P\left(t_{v} \le \frac{st+1}{s\left(1-\frac{1}{s}\right)^{s/(s-1)}\left(k-s\frac{s-2}{s-1}\right)^{s/(s-1)}} - \frac{1}{s}\right)$$
(6)

The transformation of (6) is:

$$P(k_{\nu}(t) < k) = 1 - P\left(t_{\nu} \le \frac{st+1}{(s+t)s\left(1-\frac{1}{s}\right)^{s/(s-1)}\left(k-\frac{s(s-2)}{s-1}\right)^{s/(s-1)}} - \frac{1}{(s+t)s}\right)$$
(7)

Notice that:

$$P(k) = \frac{\partial P((k_v(t) < k))}{\partial k}$$
(8)

We then have:

$$P(k) = \frac{(st+1)}{(s+t)s} \left(\frac{s}{s-1}\right)^{\frac{2s-1}{s-1}} \left(k - \frac{s(s-2)}{s-1}\right)^{-\frac{2s-1}{s-1}}$$
(9)

Equation (9) shows the property of the power-law distribution of P(k). We future prove its property of scale-free. Transform equation (9) logarithmically, we have equation (10):

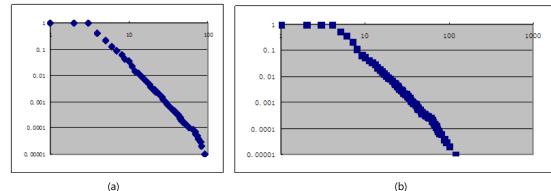


Figure 4. The Degree Distribution

$$lnP(k) = ln\left(\frac{(st+1)}{(s+t)s}\left(\frac{s}{s-1}\right)^{\frac{2s-1}{s-1}}\left(k - \frac{s(s-2)}{s-1}\right)^{-\frac{2s-1}{s-1}}\right)$$

= $ln\frac{(st+1)}{(s+t)s} + \frac{2s-1}{s-1}ln\left(\frac{s}{sk-k-s^2+2s}\right)$ (10)

Apparently, equation (10) reflects the law of power among degree distributions, and the power exponent is $\frac{2s-1}{s-1}$. Furthermore, we set $\alpha = \frac{2s-1}{s-1}$, $\beta = 1 - \frac{1}{s}$, and $s \ge 3$, then we have: $\beta(\alpha - 1) = 1, \alpha \in (2, 2.5], \beta \in (0.66, 1)$ (11)

Where the power exponents of vertex degree and probability of degree distribution follow the rule of powerlaw, thus we can determine the network under our preferential attachment rules is scale-free (Dorogovtsev S N., 2000).

Figure 4 is the simulation of network degree distribution when s=3 and s=4. We run the time step by 10000, and the simulation of two different original complete networks has the apparent property of power-law degree distribution. From the linear analysis, we can obtain α is 2.5 and 2.33 respectively, and the slope rates of these two simulations are approximately equivalent to -2.5 and-2.33. Fixed α is a very important indicator of a scale-free network.

Agglomeration Network Clustering Coefficient

In network statistics, clustering coefficient is adopted to reflect the fact of cohesion of vertices with a similar position, often defined as $C_v = \frac{number \ of \ triangles \ connected \ to \ vertex}{number \ of \ triangles \ connected \ to \ vertex}$. From this expression, bigger cluster coefficient means more convenient resource and information channels for vertices. As for the global network, clustering coefficient is the average of C_v , i.e., $C = \frac{1}{n} \sum_v C_v$. Now we study the clustering coefficient of the vertex v which has the degree of k. At the beginning, a new entrant attaching to a complete graph s –will bring sedges to the network, and $C_v = 1$. In the next time steps, new entrants sequentially attaching to incumbent vertex n will definitely attach to other s - 1 vertices with which n fully connect.

Following this rule, we can find out the law between vertex degree and its clustering coefficient:

For a network having *k* vertices, the number of triangles connected to vertex *k* is $\frac{s(s-1)}{2} + (s-1)(k-s)$, and the number of triples centered on vertex *k* is $\frac{k(k-1)}{2}$, so $C(k) = \frac{\frac{s(s-1)}{2} + (s-1)(k-s)}{\frac{k(k-1)}{2}} = \frac{(s-1)(2k-s)}{k(k-1)}$ is clustering coefficient of the global network. When *k* is sufficiently big, $C(k) \rightarrow 2(s-1)k^{-1}$. In another word, the clustering coefficient has the property of $C(k) \sim k^{-1}$.

As for the global clustering coefficient, any vertex with k degrees and related probability must be considered, thus the expression including C(k) and p(k) is denoted as:

$$C = \sum_{k} C(k)P(k) = \sum_{k=s+1}^{\infty} \frac{(s-1)(2k-s)}{k(k-1)} \frac{(st+1)}{(s+t)s} \left(\frac{s}{s-1}\right)^{\frac{2s-1}{s-1}} \left(k - \frac{s(s-2)}{s-1}\right)^{-\frac{2s-1}{s-1}}$$
(12)

C'(s) > 0 determines C(s) is an increasing function. Thus bigger *s* means bigger clustering coefficient. When s = 4, the coefficient is 0.811, basically a big coefficient when comparing with other complex networks. And the increasing function indicates that the network mentioned in this paper will have very large clustering coefficients.

From this result, it is clear that the agglomeration network has the property of large clustering coefficient in small world network.

Average Path Length

Path length determines the easiness of one vertex attaching to another incumbent vertex in a network. When examining a global network, more attention will always be paid on figuring out the average path length among all path lengths in the network. In complex network statistics, the shortest path between any two vertices is used to represent the path length. In self-employment agglomeration, shorter path implies two vertices can communicate in a more cost-saving and efficient manner. Noticing that in self-employment, every two businesses can equally communicate with each other, making their ties undirected. As a result, we can set the paths among vertices as undirected. Based on this assumption, we can compute the average path length as following.

Let d(i,j) be the shortest path length (also known as geodesic distance) between vertex pair (i,j). We define aggregated distance for a M – vertex as:

$$\rho(M) = \sum_{0 \le i \le j \le M-1} d(i,j) \tag{13}$$

Let L(M) be the average length, then:

$$L(M) = \frac{2\rho(M)}{M(M-1)}$$
(14)

Though new entrant will add the aggregate length to network, it will not affect the lengths that any incumbent vertices have; therefore L(M) is a monotonically increased function. Supposing a new entrant entering into a M – vertex network, we can set up a function reflecting the length variation between M – vertex network and (M + 1) – vertex network.

$$\rho(M+1) = \rho(M) + \sum_{i=0}^{M-1} d(i, M)$$
(15)

According to the preferential attachment, the new entrant will definitely build connections with a s – complete graph in the M – vertex network. The vertices of selected s – complete graph are labeled as w_1, w_2, \dots, w_s ; therefore, (15) can be rewritten as:

$$\rho(M+1) = \rho(M) + M + \sum_{i=1}^{M-1} D(i, w)$$
(16)

Where, $D(i, w) = min(d(i, w_1), d(i, w_2), \dots, d(i, w_s))$. Suppose the network has ψ set of $\{w_1, w_2, \dots, w_s\}$, then we need only count the distances of $\sum_{i \in \psi} d(i, w)$ when counting actual $\sum_{i=1}^{M-1} D(i, w)$. Furthermore, $\sum_{i \in \psi} d(i, w)$ can be approximately substituted by:

$$\sum_{i \in \psi} d(i, w) \approx (M - s)L(M - s + 1)$$
(17)

For the reason that $\sum_{i=1}^{M-1} D(i, w)$ is equal to $\sum_{i \in \psi} d(i, w)$, we have:

$$\rho(M+1) = \rho(M) + M + \sum_{i \in \psi} d(i, w)$$
(18)

Due to the monotonically increased function of L(M), (19) works when $s \ge 2$:

$$\frac{2\rho(M-s)}{M-s} < (M-s)L(M-s+1) = \frac{2\rho(M-s+1)}{M-s+1} < \frac{2\rho(M)}{M}$$
(19)

Noticing $\rho(M + 1) \approx \rho(M) + M + \sum_{i \in \psi} d(i, w)$, we have an inequality of (20) when combining with (19):

$$\rho(M+1) < \rho(M) + M + \frac{2\rho(M)}{M}$$
(20)

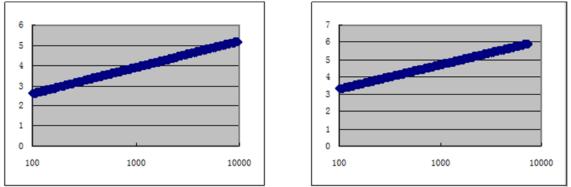
Transforming inequality (20) into (21):

$$\rho(M+1) - \rho(M) < M + \frac{2\rho(M)}{M}$$
(21)

We then have a differential equation $\frac{d\rho(M)}{dM} = M + \frac{2\rho(M)}{M}$, the solution to which is:

$$\rho(M) = M^2 lnM + \eta \tag{22}$$

Then combining (14), (21) and (22), the estimation of average path length can be approximately obtained by:



(a) s=3

Figure 5. The Relation of Average Path Length and Number of Vertices

$$\frac{2((M-s)^2\ln(M-s)+\eta)}{(M-s)(M-s-1)} < L(M) = \frac{2\rho(M)}{M(M-1)} < \frac{2(M^2\ln(M)+\mu)}{M(M-1)}$$
(23)

(b) s = 4

When $M \to \infty$, L(M) will increase approximately by the rate of $\ln(M)$, i.e., $L(M) \sim \ln(M)$.

Figure 5 is the demonstration of average path lengths when we set s = 3 and s = 4. We can observe a linear correlation between the average path and total vertices when using logarithmic transformation. When s = 3 and M = 100, the average path length is 2.7; and when M = 10000, the average path length is approaching to 5.3. Similarly, when s = 4, the average length stays at the interval of 3.2 to 6.2. Because most of self-employment agglomeration includes less than 300 vertices, network with the preferential attachment set in this paper has very short path length, and the length is linearly correlated with the logarithm of numbers of vertices, reflecting that this kind of network is a very typical small-world network.

The agglomeration with small world network property is not only from the deduction of vertices' preferential attachments but also from the essential needs of self-employment. Because of its short average path length, the agents can conveniently seek for help in this existing agglomeration, and the cost to maintain the long-term relations will be significantly alleviated. Furthermore, most agents in a self-employment agglomeration will only focus on their kinship and township relations, making this agglomeration full of "strong ties". Due to strong ties and short length path, most self-employment can share their knowledge and information after their entering into this network. Thus, people can easily master the "core competency" by knowledge spillover among agglomeration. Under this situation, business holders have no incentives to protect their business secret and make innovation in their products and technology. For this reason, most of the self-employment agglomeration is formed by low-tech businesses, and the type of entrepreneurship is called as "necessity entrepreneurship" (Tang, 2011; Hernandez et al., 2012).

CONCLUSION

China's mass entrepreneurship and innovation policy have inspired many individuals to choose selfemployment as their careers. Constrained by inborn resource and information insufficiency, many business holders choose to join incumbent agglomerations to supplement their weakness. China's tradition has forged people with strong ties by kinship and township relations, thus in a self-employment agglomeration, individuals will prefer to build the connections with their relatives or fellow villagers.

From a topology of a self-employment agglomeration, this paper finds out that it has complex network properties such as short average path length and power law distribution of degrees, which arouses our curiosity whether this kind of network is a typical complex network from its evolving mechanism.

Starting from an analysis of activities among agents when they decide to enter into an existing network, we abstract this selection mechanism into some basic rules in constructing an evolving network. And the main findings include: self-employment agglomeration with preferential attachment has a large average degree, and the frequency of agents' interaction is dense. Using mean-field theory, we observe the law of degree distribution and find out that this kind of network follows power-law degree distribution, thus this kind of network is scale-free. This property suits the requirement of BA network. We also study the clustering coefficient, and the result shows that clustering coefficient in this kind of network is very large, which reflects that this kind of network can be well clustered and has good cohesion. We finally compute the average path length in order to explore the resource supply channel. Our result proves that this kind of network has short average path length and the length follows a power-law distribution. Our findings demonstrate that the self-employment agglomeration under given rules is a very typical complex network.

The free-scale and small world network for self-agglomeration makes it more reachable to every business holders, providing them with more convenient material, information, and technology sources. Under China's "guanxi" society, this kind of network has its advantage to be formed and evolved by necessity entrepreneurs and will be preferentially adopted by the government who advocates "mass entrepreneurship and innovation. But this kind of agglomeration has its weaknesses, one of which is its fragile protection of knowledge property rights.

ACKNOWLEDGEMENTS

This paper was supported by National Social Science Foundation of China (Grant No. 13CGL022), Social Science Foundation of Hunan (Grant No. 15JD08), The Excellent Youth Scholars Project of Hunan Provincial Department of Education (Grant No. 15B046), The Overseas Study Project of Hunan Provincial Department of Education, and The General Project of Hunan Social Studies Evaluation Committee (XSPYBZZOO8).

REFERENCES

- Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47-97.
- Barabási, A. L., Albert, R., & Jeong, H. (1999). Mean-field theory for scale-free random networks]. *Physica A: Statistical Mechanics and its Applications*, 272(1), 173-187.
- Barbieri, P. (2003). Social capital and self-employment: a network analysis experiment and several considerations. *International Sociology*, 18(4), 681-701.
- Dorogovtsev, S. N., Mendes, J. F. F., & Samukhin, A. N. (2000). Structure of growing networks with preferential linking. *Physical Review Letters*, 85(21), 4633-4636.
- Elias, P. (2000). Status in Employment: A world survey of practices and problems. *Bulletin of Labour Statistics*, 1, 11-19.
- Gindling, T. H., & Newhouse, D. (2014). Self-employment in the developing world. World Development, 56, 313-331.
- Hernandez, L., Nunn, N., & Warnecke, T. (2012). Female entrepreneurship in China: opportunity-or necessitybased? International Journal of Entrepreneurship and Small Business, 15(4), 411-434.
- Kilkenny, M., & Fuller-Love, N. (2014). Network analysis and business networks. International Journal of Entrepreneurship and Small Business 15, 21(3), 303-316.
- Korur, F., Vargas, R. V., & Serrano, N. T. (2016). Attitude toward Science Teaching of Spanish and Turkish In-service Elementary Teachers: Multi-group Confirmatory Factor Analysis. *Eurasia Journal of Mathematics, Science & Technology Education*, 12(2).
- Kuah-Pearce, K. E. (2016). Migrant women entrepreneurs in the garment industry in modern China: embedding translocality and feminised Guanxi networks. *International Journal of Business and Globalisation*, 16(3), 335-349.
- Peng, H., Duysters, G., & Sadowski, B. (2016). The changing role of guanxi in influencing the development of entrepreneurial companies: a case study of the emergence of pharmaceutical companies in China. *International Entrepreneurship and Management Journal*, *12*(1), 215-258.
- Tang, Y. (2011). Exploration on Social Capital Dimensions in Necessity Entrepreneurship: Based on the Data from Migrant Workers' View. *Journal of Hunan City University*, 32(5), 49-53. (in Chinese)
- Tang, Y. (2013). The Structure Comparison of Necessity Entrepreneurial Resource Support Networks Based on SNA. *Business and Management Research*, 2(3),137-145.
- Zhang, Z. Z., Rong, L. L., & Zhou, T. (2005). An evolving model for scale-free collaboration networks. *System Engineering Theory and Practice*, (11), 55-60. (in Chinese)

http://www.ejmste.com