

## HOUSEHOLDS INCOME-EXPENSES NETWORK SIMULATION

Javier Lara de Paz<sup>(a)</sup>, Idalia Flores de la Mota<sup>(b)</sup>, Gabriel Policroniades Chipuli<sup>(c)</sup>, Sashiko Shirai Reyna<sup>(d)</sup>

<sup>(a),(b),(c),(d)</sup>Facultad de Ingeniería, Universidad Nacional Autónoma de México

<sup>(a)</sup>[javier\\_lara\\_depaz@comunidad.unam.mx](mailto:javier_lara_depaz@comunidad.unam.mx), <sup>(b)</sup>[idalia@unam.mx](mailto:idalia@unam.mx), <sup>(c)</sup>[gpolicroniadesch@gmail.com](mailto:gpolicroniadesch@gmail.com),  
<sup>(d)</sup>[sashikosr@outlook.com](mailto:sashikosr@outlook.com)

### ABSTRACT

Economic systems are recognized as complex systems that can be modeled as complex networks. Analyzing and simulating with the outcoming networks parameters can be obtained the vulnerability and robustness of system's network, assisting the developing of economic policies. In present work we built up an economic weighted directed network between the catalogued labors, the productive sub-sectors, based on 4-digits NAICS-2018, and the eight common household expenses. Weighting edges with total household incomes and total amount paid for each expense. The network is constructed with a sample size of more than one hundred-thirty thousand households. With complex network metrics and adding an inflation rate to each main Expense category, the simulation was performed obtaining key economic properties such as wealth distribution per household and sectors vulnerabilities as well as system robustness.

Keywords: Complex Networks , Complex systems, simulation

### 1. INTRODUCTION

In many countries, their National Statistical Institutes carry out periodically a demographic survey system, either as a census or as a sampling survey. These surveys are designed to collect information about National Accounts, demographic characteristics, as well as economic behaviors and patterns, among many other information that can be of national interest. In some of these surveys it is described how much population earn, and how much money is spent in principal expenses, or what are the economic activities, even at a household, companies or grand productive sectors level. The information obtained is used by researchers of the academic, public or financial sectors. Even by economists, which by means of econometric, statistical or financial methods, among many others approaches, develop models to estimate either the demographic or financial behaviors. Also, it is possible to estimate the macro-economic information, the GDP, inflation and unemployment rates, economic growth, added values of different productive sectors. It can be considered that the techniques applied can be considered as in a "traditional" framework, which doesn't allow to visualize robustness of the system. A review of evolution to new trends in

economical methods can be seen in (Lara de Paz et al. 2019). Or to find weakly points that risk the system's functionality, or even to obtain clustering formations, leader entities, or structural relations within system. Some of the variables included in the generated metadata from surveys or census, are the economic activities performed by households. Also, information about labor-activities is gathered indicating in which economic productive sectors do these agents work, and how much do they earn as well as in what and how much do they spend. These data is collected through the Households Income-Expenses Survey carried out in Mexico every two years since 1992. The scope in this information provides insight about what trades and crafts are performed per sector. These functional relations can be constructed through a network architecture where economic activities, productive sectors and common expenses become nodes. Wages and amount spent correspond to the edges weight. The network topology constructed with the survey's data provides more information than an analysis performed with traditional econometric and statistical treatments. In present work we constructed a network with the economic-activities per household's variables, productive sectors and the principal expenses. Data were obtained from Mexico's official surveys about households' expenses, jobs and Economic Census for the year 2016 (ENIGH-2016, Income-Expenses National Survey (INEGI n.d.)). Mexico's National Accounts System (SCNM)(INEGI n.d.) were also consulted. We combined Complex Networks techniques with Simulation approach to obtain more insight about Mexico's socio-economic structure, (Hassani, Saporta, and Silva 2014). Zanin (Zanin et al. 2016) presented in his work some techniques about combining Data Mining methods with Complex Networks theory mainly applied to medical research. They suggested that a network analysis from the system would be an advantage to identify the robustness of the system by identifying single nodes influence over the whole network. Although they also combined Data Mining methods with Complex Networks theory, but one of their strongest points was that complex networks allow representing the system's internal structure of each individual agent relations in a simpler and more understandable way. To represent our system using Network topology let us have an overview of system

structure, formation of communities, collective behaviors and to identify the diversity of nodes and its multiple connections, representing the complexity of the system (Silva and Zhao 2016). By applying a simulation methodology allow us to find which instances are more significant to robustness and vulnerability. In the next introductory subsection some Complex Networks theory features are described to understand some of the metrics used as parameters to perform simulations.

### 1.1. Complex Networks

Interest in complex networks field has grown in many areas from physicists, mathematicians, biologists, engineering, computer scientists, among many others. The structure analysis of many different systems that can be represented as a network, are the Internet, World Wide Web, biological systems together with chemical systems, financial, social, neural, and communication networks. Although these systems can represent artificial or natural systems, they share a main feature, the complexity. Barabási, one of the more active researches in this field (*Network Science by Albert-László Barabási* n.d.) specified that an adequate manner to represent evolution of a complex system, can be through its network topology. Catanzaro and Buchanan (Catanzaro and Buchanan 2013), mentioned some disadvantages of main stream economic theories for preventing an economic crisis like in 2008, and pointed some advantages in developing financial networks. About application of network structures to economic and financial systems, Bouchaud (Bouchaud 2008) addressed to the necessity of the search and applications of different methodologies to aid main stream economic models. Schweitzer et al. (SCHWEITZER et al. 2009), examine the emergence of many economic systems described as networks and the consequences in analyzing through this approach, which is distinguished as a convenient way of data analysis. Caldarelli and Catanzaro (Caldarelli and Catanzaro 2014) made a friendly description about many networks present in Nature. The system involved corresponds to a weighted network constructed as having a branching, treelike topology, like a transportation network with transfer nodes, becoming a directed network as well.

#### 1.1.1. Complex Network Metrics

About the Complex Networks metrics and theory applied to describe our system, Boccaletti (Boccaletti et al. 2006) summarized the sufficient network's features needed for analyzing its behaviors and evolution. Also the work done by Deguchi et al. (Deguchi et al. 2014) was used to estimate HITS metrics. For modularity and community properties, we based on Newman (Newman 2006). In the work reported by Klimkova et. al ((Klimkova, Senkerik, and Zelinka, n.d.), described some methods to detect giant components in directed networks, and Levorato and Petermann (Levorato and Petermann 2011) described some techniques to identify communities in a directed weighted network, based on the modularity, allowing to detect the number of communities or

connected subgraphs in network's system. In this work we calculated the following network metrics (table 1):

Table 1: Network's Metrics calculated at scale-levels

Micro-scale	Nodes degree, in and out degree for a weighted network, eigenvector centrality
Meso-scale	Modularity, strong component number and assortativity
Macro-scale	Average path length

To estimate the system's robustness we performed the simulation adding and removing some high connected nodes according to a percolation model (D'Agostino and Scala 2014) and (Réka Albert and Barabási 2002). In complex systems robustness presents a dynamical component, while representing these systems by its topological structure underlying in complex network, these error tolerance or the vulnerability can be addressed by edges/nodes removal leading to a change on stability of such systems related to error and attack tolerances. In present work we focus only on nodes removal, since some network structures are more robust against random node failures. The node removal mechanism consists in deleting nodes randomly, the second form is by removing highly connected nodes, considering the nodes with highest node degree. Other mechanisms are by cascade or percolation method. The remaining structure allow us to obtain critical vulnerability of the system, (Reka Albert, Jeong, and Barabasi 2000).

## 2. SYSTEM'S STRUCTURE

The system considered is composed by three different class of elements, agents' activities, which are referred to the diversity of jobs, trends or crafts, registered in corresponding surveys outcomes. These categories are classified based on Mexico's National Catalogue of Jobs and Occupations (SINCO). The other class of elements corresponds to productive subsectors described at 4-digits level based on the North American Industry Classification System (NAICS-2018). The third class are the eight main expenses done per households (housing, shoes & clothing, education, personal care, food, transportation, health care and cleaning-wares). The relation between elements correspond to: between jobs and sectors are total annually incomes. Relation between jobs and expenses are amount spent by household in every expense (Fig. 1). For each household entities considered in the ENIGH-2016 survey, a sample size of more than 130 thousand basic units, it was considered the economic activities (SINCO) the relation with their economic subsector (NAICS 4-digits classification) and the total annual incomes and total annual expenses in 8 main expenses (Table 2).

Table 2: System's elements for network structure

Productive Subsector (NAICS) 2018 version.	Economic Activity (SINCO) 2011 version	ANNUAL INCOME	ANNUAL SPENT FOR EACH EXPENSES (EDGE WEIGHTS)
			<ul style="list-style-type: none"> <li>HOUSING</li> <li>FOOD</li> <li>CLOTHES &amp; SHOES</li> <li>EDUCATION</li> <li>TRANSPORT</li> <li>PERSONAL CARE</li> <li>HEALTH CARE</li> <li>CLEANING-WARE</li> </ul>
CLASS ENTITIES 1	CLASS ENTITIES 2	EDGE WEIGHTS	CLASS 3 ENTITIES

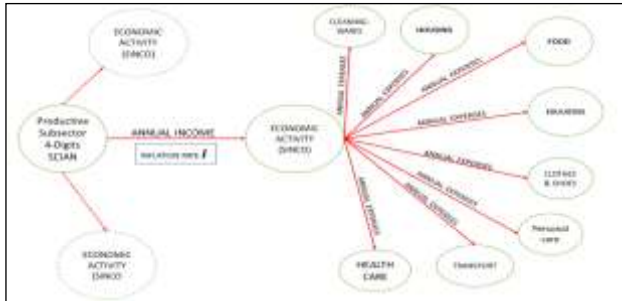


Figure 1: System's network generated by elements relations

Subsector	NAICS	SINCO	Total Income/yr	Cleaning-ware	Food	Housing	Clothes & shoes	Health care	Transport	Personal care	Education
3											
4	10003801	54112121	180986.7	15343.81	15621.89	10543.5	1888.03	983.13	3450	2870.4	890
5	10003801	54112271	180986.7	15343.81	15621.89	10543.5	1888.03	983.13	3450	2870.4	890
6	10003802	32502271	148618.16	8407.35	14837.7	9526.5	9415.7	4397.26	10367.15	18472.91	2213.11

Figure 2: Image of first 3 rows CSV Data sheet used to construct system's network.

### 2.1. Network Structure

To construct the network a directed origin-target layout was applied with origin nodes as Productive Subsectors, first target nodes as the occupational activities and directed to final nodes represented by the eight different categories of expenses. The network obtained is a directed weighted network with edges weights being the annual incomes for first branches and annual spent for second branches. For each household its corresponding branch was constructed:

Subsector → economic activity → expenses  
income                      spent

Network visualization was performed with aids of NETWORK R-package, as well as network metrics calculations. The calculated Metrics are:

- Node degree, in and out degree and weighted in and out degree. For this metric node distribution is estimated. For a weighted directed network this metric is calculated by adding (for in degree) incoming edges weights by node.
- Nodes centralities, eigenvector centrality. This measure indicates importance of each node depending on its connections and weighted links of them.
- Assortativity. This metric indicates capacity of high connected nodes to connect to other high

connected nodes (positive assortativity), or capacity of high connected nodes to connect with low degree nodes (negative assortativity). It is associated to robustness of system, as one high connected node is removed or changed, the remaining structure tends to link whether to a high connected node or to a low connected node in order to network structure to survive (Shizuka and Farine 2016)

- Modularity: this meso scale measure, indicates the formation of communities between similar node categories. It is described with a number related to the subgraphs formed by dividing network in optimal number of communities.
- Strong community number: associated to network's modularity for each node indicates the strongest association of each node to its corresponding community. This metric is obtained implementing the system's network in GEPHI software released 0.9.2 (<https://gephi.org>)
- Vulnerability and robust. A removal of nodes is done depending on different mechanism for removing nodes. This measure exhibits systems robustness, and a threshold that indicates when network starts to show fatal errors.

From these metrics simulation performance is carried out to find robustness of the system through complex network approach, which is the main scope of present work.

### 3. METHODOLOGY

Network construction and parametrization are based on the results of bi-annual Mexico's Incomes and Expenses per Households National Survey carried out on 2014. Information is available as a metadata csv format file. The data sheet contains concentrated information of about 130'000 housing with data about location, economic activities, ages, genders, detailed incomes and expenses per households, among many other relevant variables. Out from this information we are interested in economic information about earnings, expenses, economic activities and where do heads of household work. We summarized earnings in total annual incomes and 8 different categories of expenses: housing, food, clothes & shoes, education, transport, personal care, health care and cleaning-ware. Each housing is identified with a unique Id code, by this it is possible to detect same households in different official surveys. Metadata results also show a variable indicating in which productive subsector the head of household works. The coding is based on the NAICS-2013 classification system ("Sistema de Clasificación Industrial de América del Norte, México SCIAN 2013," n.d.). With this information it is possible to match each household in the survey sample with the productive subsector, the economic activity performed, annual and total amount expended in main 8 expenses. Intuitively, this matching generates a directed network (Fig. 1) with 11 different nodes categories for each corresponding column (Fig. 2)

and the edges are generated by every row, which correspond to each sampled household. Network structure is constructed under R coding, calculating network metrics, nodes degree (weighted in and out degree distribution, modularity, eigenvector centrality and vulnerability). This information allows us to have a state of the system of how it is behaving. Up to this information we start with simulation process. Network metrics results correspond to 2014 end year's behavior of economic dynamics, with corresponding information of subsequent years (2015, 2016, 2017 and 2018) through simulation, it is possible to estimate the network topology with the corresponding metrics and finally the evolution of robustness of the system. Annual accumulated inflation rate per expense categories is the changing parameter for our simulation performances which affects directly to the expenses branches which go from occupation nodes (identified by the Mexican jobs classification system SINCO) to the 8 categories of expenses. Year 2014 will be our referenced year. About incomes increasing rates, it is assumed that they remain stable through the 4 years period. Inflation rates considered are obtained from the National Consumer Price Indexes (NCPI) for the eight mentioned categories. These are obtained from official databases (Mexican Central Bank, BANXICO). Annual variations between years are multiplied to expenses edges, and network metrics are calculated for each 4 periods with year 2014 as reference. In table 2 variations of the NCPI's of Mexican economy information is showed for years 2014 to end of 2018. Simulation report will show variation of network global weighted in and out degree indicating which productive subsectors or occupations are economically more connected to higher wages or earning nodes. For in-weighted degree, it indicates which occupations or labors are better paid if for a high degree, and for expenses categories which category is more economically significant for households. The eigenvector centrality is also reported indicating what targeted nodes are more important between occupations and expenses categories. Assortativity allows us to identify systems capacity of interacting between sectors and occupations and vulnerability shows robustness of considered system. Fig. 3 shows the process followed by methodology applied.

Table 3 : Annual cumulative Mexican Inflation Rate for years 2014 to 2018\*

Expense \ Year	End 2014	End 2015	End 2016	End 2017	End 2018
HOUSING	2.02	-0.07	1.11	5.14	2.54
FOOD	6.53	2.32	4.31	7.92	5.39
CLOTHES & SHOES	2.26	2.90	3.26	4.05	2.14
EDUCATION	3.85	3.55	4.23	4.57	3.83
TRANSPORT	4.45	2.43	4.25	11.48	9.36
PERSONALCARE	6.80	4.51	4.89	5.76	5.34
HEALTH CARE	2.87	3.33	4.15	6.13	4.43
CLEANING-WARE	1.58	2.94	2.19	5.06	2.85

\*Source: INEGI

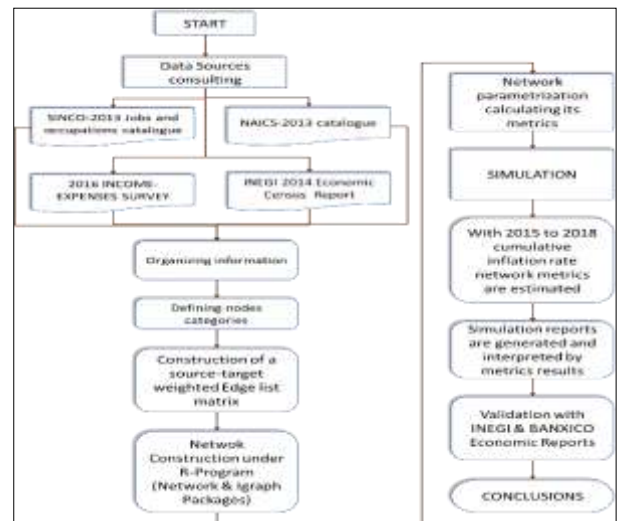


Figure 3: Methodology

#### 4. SIMULATION

To perform the simulation, parametrization of system is carried out by network construction and calculating networks node degree: weighted in and out degree, eigen centrality, assortativity and income distribution. It is identified the productive subsectors with more influence along the network structure, as wages represent edges weights from Subsector nodes to jobs/occupation target nodes, as higher out-weighted node degree a node presents, more economic influence the node has. About weighted in-degree, these correspond to earnings by transfer jobs/occupational nodes and for Expenses category nodes, as higher value it presents, more significant is the node over the network. Edges weights from jobs/occupations nodes to Expenses category nodes, as mentioned, represent the amount each household has paid for respective services. For the simulation, weights values are calculated, with the variation indexes of inflation rates corresponding to this category based on the year and on the node category, constructing four different networks models representing our system. The outcomes of network metrics are calculated in each case and interpreted under robustness insight. To finish with our comparative simulation outcomes, difference between weighted in and out degree distribution by node is analyzed, omitting the 8 nodes expenses categories, resulting only for the jobs/occupation nodes category (Fig. 15). From this distribution the Gini coefficient is obtained as an estimation to understand income distribution after paying for life costs. These metrics are calculated for years 2015-2018. Although consulted source corresponds to the 2016 survey edition (ENIGH\_2016), information was collected along 2014 and 2015, hence our base year corresponds to 2014. From this year simulation are performed. In next subsection corresponding to simulation results analysis, it is shown network metrics for each year with respective modifications to inflation rate by product for each year.

#### 4.1. System's Network parametrization

Fig. 4, shows the system's network with 594 nodes and 743409 edges. In left frame network is showed without label and nodes are sized based on its node in-degree. Bigger nodes represent the 8 Expenses categories. Right frame represents network with nodes sized based on its node out-degree.

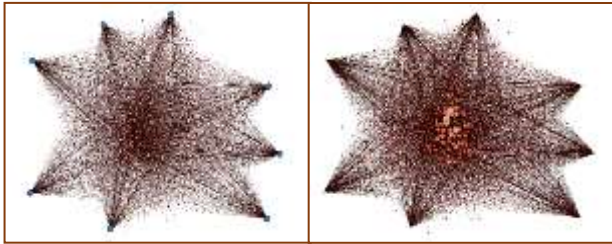


Figure 4: System's network.

Next the complex networks metrics are shown. Fig. 5 shows the degree distribution plot. This metrics indicates the average links per node. The average node **degree** = **15.59** which means that on average every node is linked with 16 different edges to its closest neighbors. It is important to underline that the 8 Expenses Categories nodes have a node **degree** = **450**. But as shown in nodes degree distribution histogram (Fig. 6), about 120 nodes present a low connection degree. In Fig. 7, it is shown the distribution of nodes degree by nodes categories. Inside frames indicate the value of the high connected nodes corresponding to productive branch "General Public Administration" with a node **degree of 143**. This means that this productive subsector is the most diversified because of the different kinds of occupations it demands. About jobs/occupations category, the most linked node with a node degree of **101** corresponds to the **9621** classification job. This code is classified in the Mexican National Occupation System Catalogue (SINCO-211), as general cleaner and sweeper (except for hotels and restaurants). This means that this occupation is the most demanded job by a high diversity of productive subsectors. In Fig. 7, nodes vs node degree is plotted omitting the 8 Expenses Category nodes and a power law is fitted. Inside labels shown node degree of most connected nodes by category. In the *x*-axe the node list is labeled, but due to space limitation it is adjusted to the font size and the complete list is not presented. Node degree indicates mainly a topological system's feature, as one of main objectives is to gain more insight on economic dynamics other metrics are considered as weighted node degree. This metric indicates the sum of edge's weight each node has, allowing us to estimate economical significance by node. As higher the value more significant is the node. For a system's directed network representation there are two other metrics related to edges' weight, weighted out and in degree. For Out-degree corresponds to the sum that a node had paid in general, as wages in the case of subsector → job/occupation edges, and for job/occupation → Expenses edges, it correspond to amount paid for living services described as Expenses.

For weighted In-degree, it indicates, for the Jobs/Occupations category nodes, the total wages earned from the different subsectors each node is connected to. For the Expenses category nodes, this metric indicates how much is expended in each of the 8 categories by jobs/occupation category nodes. This last metric corresponds to the response to our control parameter inflation rate in simulation performance. Fig. 8 shows weighted node degree distribution. In vertical axis list of nodes are plotted and in horizontal axis it is shown the value of node degree. The node with highest degree (inside label) corresponds to productive subsector category node identified as "Education supporting Services" with a value of 36080309.98. It can be interpreted as this node has paid to its 93 (node degree) different occupations related to this category the mentioned amount (units are in Mexican pesos MXN, 1USD~20MXN). In Fig. 9, a bar plot of weighted in-degree for the 8 Expenses category nodes is presented. In Fig. 10, the distribution of weighted in-degree for Jobs/Occupation Category nodes is plotted. A Power-Law fitting is shown (red dotted line) and an empirical equation with determinant coefficient is shown in upper-right inside frame. With relation to the connectivity, relevance each node presents is expressed through eigencentrality value. In Fig. 11, eigencentrality value is plotted vs Jobs/Occupations category nodes. In both Figures, it can be observed that nodes with the highest respective values are the same for both metrics. In table 4, the more diversified jobs based on mentioned metrics are described. About network metric Modularity class, network can be separated in 9 communities (Fig. 12). Respect to assortativity, the value obtained is **Assortativity** = **-0.7061681**, this means it has a dissociative behavior by high hubs nodes. Gini coefficient **GN** = **68.6** obtained from difference between weighted in and out degree, (distribution is shown in Fig. 13), shows an extremely high value. We assumed two main reasons for such a high GN value. (real one for 2014 GN=45.8), first sample size is not big enough for calculating real income distribution and there are three nodes with a weighted in degree around MXN 360 million a year. But it is able by this technique to manipulate nodes structure to get more insight

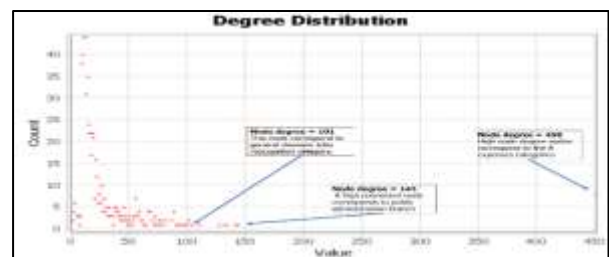


Figure 5: Node degree distribution dot plot. Inside frames indicate node degree of the more connected nodes of the 3 nodes categories.

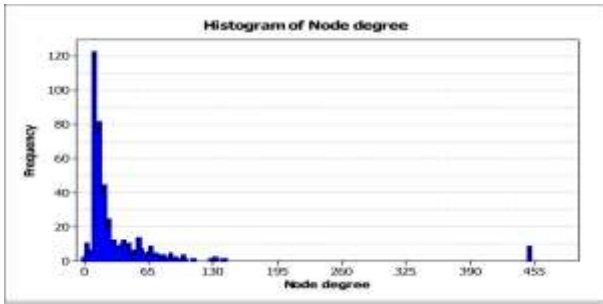


Figure 6: Histogram of node degrees

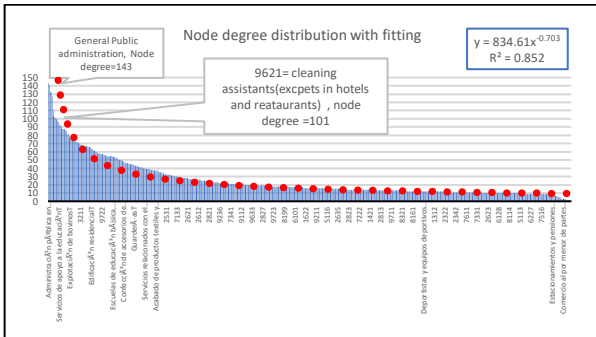


Figure 7: Node degree distribution bar-plot showing all nodes categories, except the 8 expenses categories with an potential fit. Inside frames indicate the most connected node corresponding to the productive subsector and the most connected job/occupation category node. Red dot line correspond to a power law fitting and corresponding empirical calculated model with Pearson  $R^2$  coefficient is showed in up right inside frame.



Figure 10: Distribution of weighted in-degree by jobs/occupations category nodes in horizontal axis. A Power-Law distribution fit was performed, and empirical equation is shown in upper-right frame.

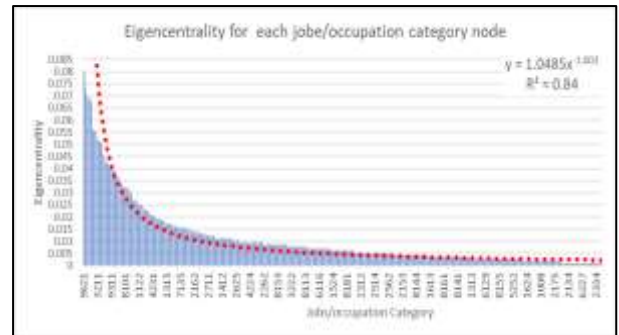


Figure 11: Eigencentality vs jobs/occupation category nodes plot. Power-Law is fitted with an determinant coefficient of 00.84 is obtained.

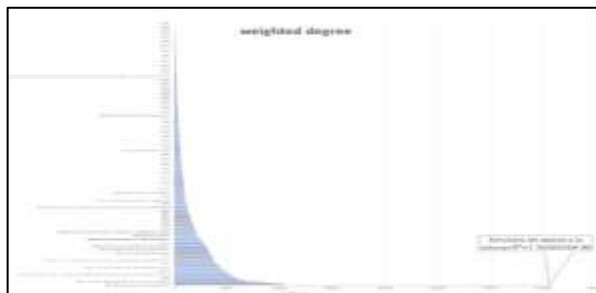


Figure 8: Weighted degree distribution plot. Node with highest degree corresponds to productive subsector “education supporting services” with a node weighted degree of 36080309.98.

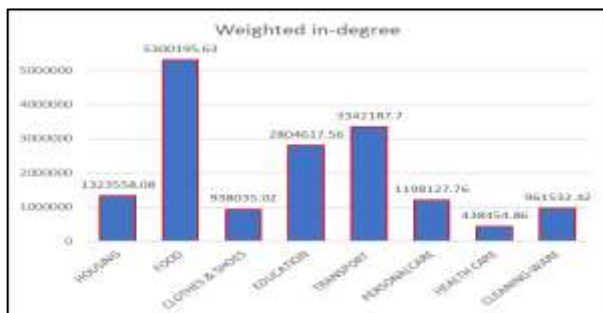


Figure 9: Column plot for weighted In-degree by Expenses (base year 2014).

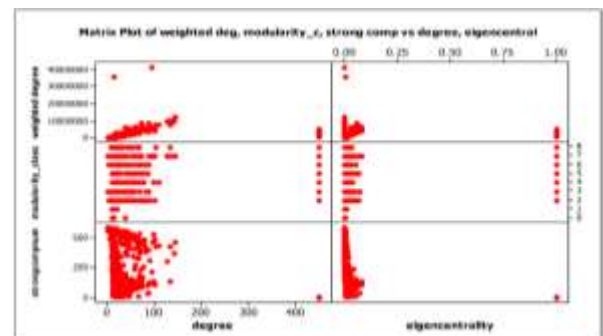


Figure 12: Matrix plot of degree and eigencentality as horizontal axes and strong, modularity and weighted degree for vertical axis.

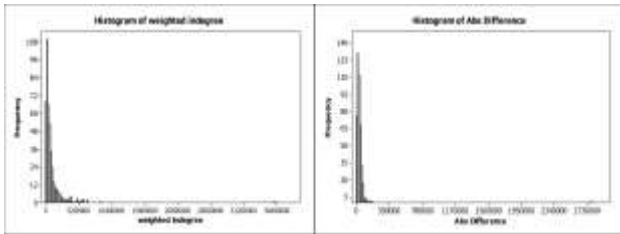


Figure 13: right plot shows weighted in-degree histogram. In right frame shows histogram of the difference between weighted in and out degree by job/occupation category nodes.

### 4.2. Simulation results

The simulation is performed changing weights of Jobs/occupations Category → Expenses edges for years 2015-2018. The outcomes' analysis are focused only in weighted degree distribution, weighted in-degree distribution and estimation of Gini coefficient. This because other metrics won't change significantly. Results are summarized in Table 5. In Fig. 15 the annual weighted degree variation for years 2014-2018 is shown, it can be observed that the food category is more significant but transport presents more variation through selected years.

Result \ Year	End 2014	End 2015	End 2016	End 2017	End 2018
Gini Coefficient	68.6	72.6	63.54	63.52	63.54
Variation %		5.38	-12.48	-0.031	0.031

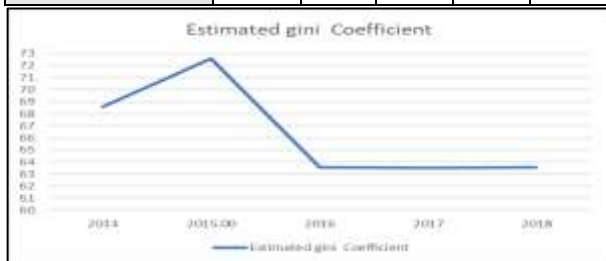


Figure 14: Estimated Gini coefficient from 2014-2016

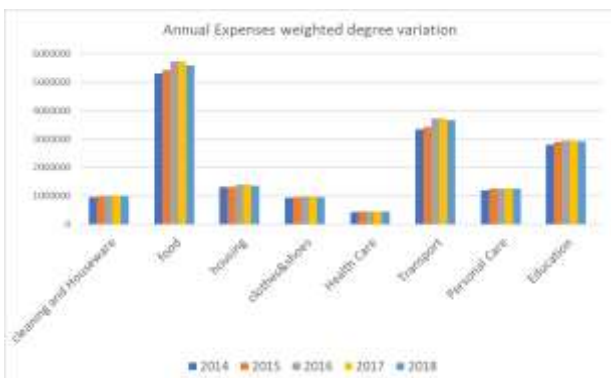


Figure 15: Simulation results for Weighted degree variation when inflation rate was introduced.

### 4.3. Validation

Our results are validated through the variation of Gini Coefficient, for corresponding years. It has to be noticed that there is no information available about years 2015

and from 2017 to date. As mentioned, our results have an estimation error of about 49 % in some years, but this can be due to lack of data, and we are only considering the analysis of jobs/occupations category. A further analysis using productive subsector should be improved. Behavior of Gini coefficient responds to shape observed by variation from 2014 to 2016. Our results can be applied to study system's behavior.

2000	2002	2004	2005	2006	2008	2010	2012	2014	2016
51.40	49.00	48.30	48.90	47.70	44.60	45.30	45.40	45.80	43.40
5.54%	-4.67%	-1.43%	1.24%	-2.45%	-6.50%	1.57%	0.22%	0.88%	-5.34%

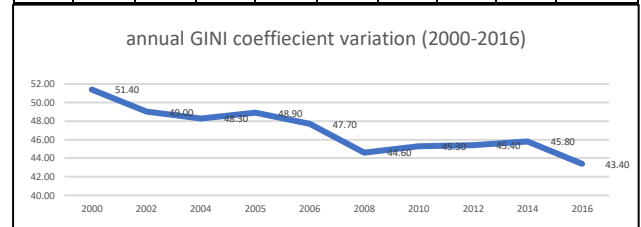


Figure 16: Gini Coefficient values and behavior from years 2000 to 2016. Source INEGI.

## 5. CONCLUSIONS

Our technique to analyze an economic system by Complex Network approach is suitable to interpret a complex system state, diversity of nodes can be tracked by its diversity of links. Simulation allows us to identify some of key nodes by categories, obtaining an estimated evolution of income distribution using Gini coefficient. A further Modularity metric scope would indicate evolution of system clusters, permitting to identify economic groups aiding to reinforce economic dynamic. Implementation of data mining approach can aid to get more insight of system, by tracking each node and its links. By means of presented method directed attacks or reinforcements of different occupations or jobs, or incentives by expenses or even aiding the design of economic policies directed to specific subsectors can be implemented.

## ACKNOWLEDGMENTS

To the PAPIIT DGAPA project IT102117, UNAM for its support to this research.

Javier Lara de Paz is a doctoral student from Programa de Doctorado en Ingeniería, Universidad Nacional Autónoma de México (UNAM) and received fellowship 593341/309027 from CONACYT.

## REFERENCES

- Albert, Réka, and Albert-László Barabási. 2002. "Statistical Mechanics of Complex Networks." *Reviews of Modern Physics* 74 (1): 47–97. <https://doi.org/10.1103/revmodphys.74.47>.
- Albert, Reka, Hawoong Jeong, and Albert-Laszlo Barabasi. 2000. "Error and Attack Tolerance of Complex Networks." *Nature* 406 (6794): 378–82. <https://doi.org/10.1038/35019019>.
- Boccaletti, S., V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. 2006. "Complex Networks: Structure and Dynamics." *Physics Reports* 424

- (4): 175–308.  
<https://doi.org/10.1016/j.physrep.2005.10.009>.
- Bouchaud, Jean-Philippe. 2008. “Economics Needs a Scientific Revolution.” *Nature* 455 (October): 1181.
- Caldarelli, Guido, and Michele Catanzaro. 2014. *Redes: Una breve introducción*. Translated by María Hernández Díaz. Edición: edición. Madrid: Alianza.
- Catanzaro, Michele, and Mark Buchanan. 2013. “Network Opportunity.” *Nature Physics* 9 (3): 121–123. <https://doi.org/10.1038/nphys2570>.
- Deguchi, Tsuyoshi, Katsuhide Takahashi, Hideki Takayasu, and Misako Takayasu. 2014. “Hubs and Authorities in the World Trade Network Using a Weighted HITS Algorithm.” *PloS One* 9 (7): e100338–e100338. <https://doi.org/10.1371/journal.pone.0100338>.
- Hassani, Hossein, Gilbert Saporta, and Emmanuel Sirimal Silva. 2014. “Data Mining and Official Statistics: The Past, the Present and the Future.” *Big Data* 2 (1): 34–43. <https://doi.org/10.1089/big.2013.0038>.
- INEGI. n.d. “ENIGH 2016.” Accessed May 15, 2019a. <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825091934>.
- . n.d. “Sistema de Cuentas Nacionales de México. Fuentes y Metodologías. Año Base 2013.” Accessed May 15, 2019b. <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825097165>.
- Klimkova, E., Roman Senkerik, and Ivan Zelinka. n.d. “Topological Structure Analysis in Directed Network.” In .
- Lara de Paz, Javier, Idalia Flores de la Mota, Gabriel Policroniades Chipuli, and O. Sashiko Shirai reyna. 2019. “WEALTH AND INCOME DISTRIBUTION: A REVIEW TOWARDS NEW TRENDS” 04 (01): 29.
- Levorato, Vincent, and Coralie Petermann. 2011. “Detection of Communities in Directed Networks Based on Strongly P-Connected Components.” In *2011 International Conference on Computational Aspects of Social Networks (CASON)*, 211–216. IEEE.
- Network Science by Albert-László Barabási*. n.d. Accessed May 15, 2019. <http://networksciencebook.com/>.
- SCHWEITZER, FRANK, GIORGIO FAGIOLO, DIDIER SORNETTE, FERNANDO VEGA-REDONDO, and DOUGLAS R. WHITE. 2009. “ECONOMIC NETWORKS: WHAT DO WE KNOW AND WHAT DO WE NEED TO KNOW?” *Advances in Complex Systems* 12 (04n05): 407–422. <https://doi.org/10.1142/s0219525909002337>.
- Shizuka, Daizaburo, and Damien R. Farine. 2016. “Measuring the Robustness of Network Community Structure Using Assortativity.” *Animal Behaviour* 112 (February): 237–46. <https://doi.org/10.1016/j.anbehav.2015.12.007>.
- Silva, Thiago Christiano, and Liang Zhao. 2016. “Complex Networks.” In *Machine Learning in Complex Networks*, 15–70. Springer International Publishing. [https://doi.org/10.1007/978-3-319-17290-3\\_2](https://doi.org/10.1007/978-3-319-17290-3_2).
- “Sistema de Clasificación Industrial de América del Norte, México SCIAN 2013.” n.d., 596.
- Zanin, M., D. Papo, P. A. Sousa, E. Menasalvas, A. Nicchi, E. Kubik, and S. Boccaletti. 2016. “Combining Complex Networks and Data Mining: Why and How.” *Physics Reports* 635: 1–44. <https://doi.org/10.1016/j.physrep.2016.04.005>.

#### AUTHORS BIOGRAPHY

**Javier Lara de Paz** studied Physics at the Faculty of Science of National Autonomous University of Mexico (UNAM), a Specialization in Applied Statistics at the UNAM. He studied his Master’s degree in Operation Research at the Faculty of Engineering of the UNAM. He has been professor of Science at the and Operation Research professor at the Engineering School at the UNAM. Currently he is studying a Ph.D. in Operational Research at UNAM.

**Idalia Flores de la Mota** She received her Ph.D. in Operations Research at the Faculty of Engineering of the UNAM. She graduated Master with honors and received the Gabino Barreda Medal for the best average of her generation. She has been a referee and a member of various Academic Committees at CONACYT. She has been a referee for journals such as Journal of Applied Research and Technology, the Center of Applied Sciences and Technological Development, UNAM and the Transactions of the Society for Modeling and Simulation International. Her research interests are in simulation and optimization of production and service systems. She is a full-time professor at the Postgraduate Program at UNAM.

**Gabriel Policroniades Chípuli** Studied Industrial Engineering at Technological Institute and Superior Studies of Monterrey, Campus Toluca (ITESM, Campus Toluca), he studied an engineering mastery with operation research orientation and has advanced metrology studies for the calculation of uncertainty, and knowledge of Lean Six Sigma Black Belt and Logistics and supply chain. Currently he is studying a Ph.D. in Operation Research at UNAM.

**O. Sashiko Shirai** obtained her bachelor’s in actuarial science in the Faculty of Science of the UNAM. She is studying her Master in Operations Research in the Faculty of Engineering of the UNAM with the CONACYT scholarship. She is an assistant professor at the Faculty of Engineering and the Faculty of Science. Her research interests lie in complex networks, statistics and operational research.