UTHM

Universiti Tun Hussein Onn Malaysia

PREMIER TECHNICAL UNIVERSITY

Modeling the Malaysia Highway Traffic Network

by:

Norhaidah Mohd Asrah

Faculty of Applied Science and Technology (FAST)

Universiti Tun Hussein Onn Malaysia

UTHM
Universiti Tun Hussein Onn Malaysia

Contents

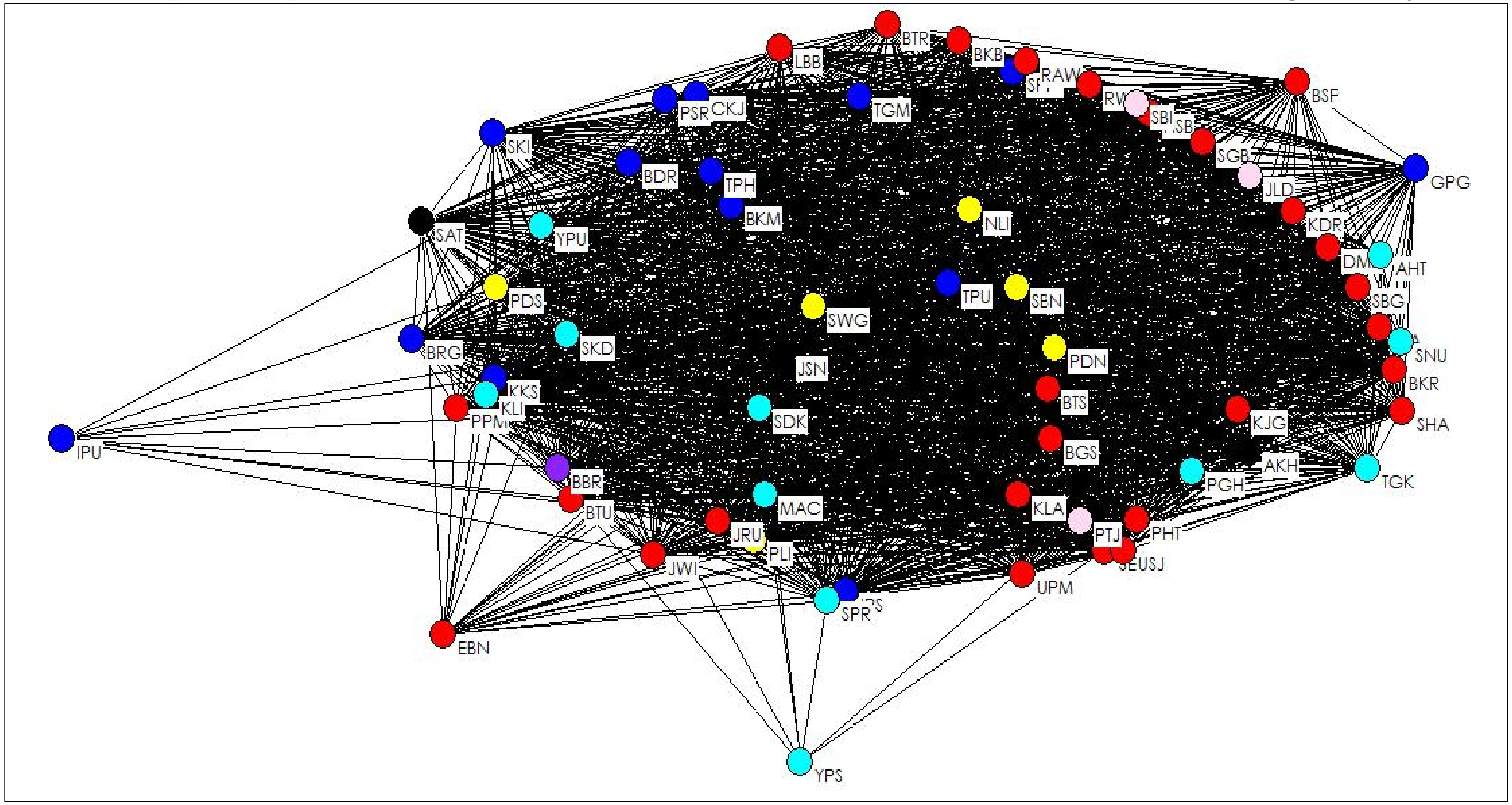
- Introduction
- Methodology
- Results & Discussions
- Conclusions & Recommendation
- References

Introduction

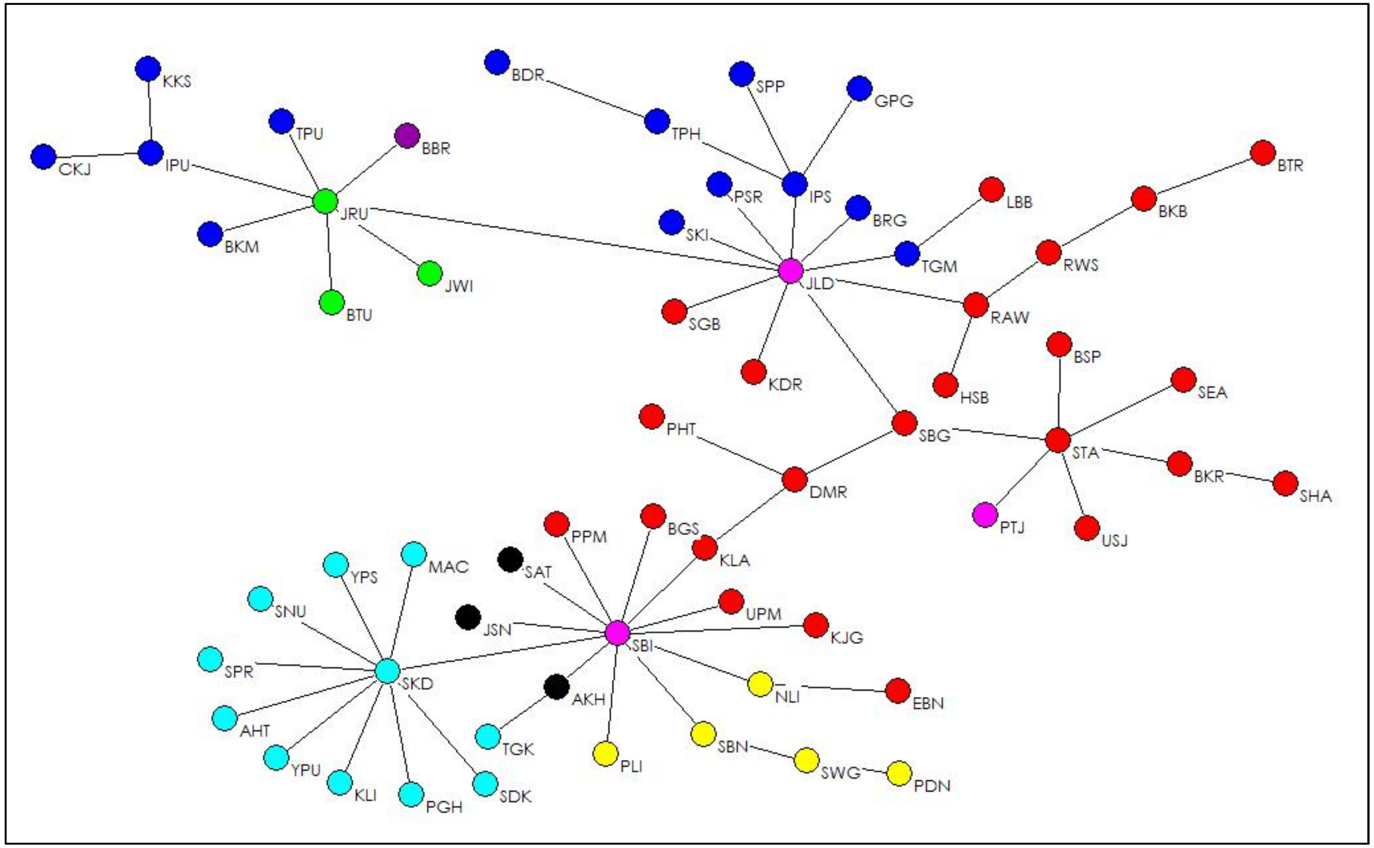
- Why the highway network in Malaysia is important?
- How to analyze the complex PLUS highway traffic network?
- The information in the PLUS highway network are
- - Modeled as weighted directed graph (WDG)
  - Visualized using theoretical graph – *nodes* (toll plazas) and *link* (traffic burden)
  - Interpreted (Social Network Analysis: Minimum spanning tree (MST) & centrality measures)
- The most important centrality measure using new measure, the *Effective Vector Variance* (EVV).

- According to DOSM (2016), the road mileage is increased from 214,813km (2015) to 236.802km (2016), 9.29% in a year & the number of registered motor vehicles also increased from 26,301,952 (2015) to 27,613,259 (2016), 4.75% in a year.
- The rising of these percentage shows that Malaysia need a good road network, and highway become crucial to support to this positive growth.
- The higher technology in development of road network can promised the instance development in many fields such tourism, foreign investments, regional development, real estates and many more.
- PLUS as the one of highway operator in Malaysia and the most busiest highway in Peninsular Malaysia, is committed to provide the effective facilities for their expressway.

Graph representation of traffic network on PLUS highway



The minimum spanning tree (MST)



Methodology

a) Data Collection

- Data collected from Toll Department of PLUS Malaysia Sdn. Bhd.

- Penang (3), Kedah (1), Perak (14), Selangor (23), Wilayah Persekutuan (3), Negeri Sembilan (4), Melaka (3) and Johor (11). Total toll plazas = 63.
- Time range is from July 2009 until December 2013.
- Traffic burden referred to the number of vehicles with their loads such as passengers and goods that coming into and exit from a toll plaza.

b) Information Filtering

- The distance matrix is used to determine the minimum spanning tree (MST). The data matrix for number of vehicles are transformed into distance matrix D using formula $D = Maximum\ value - a_{ij}$ for all $ij = 1,2, \ldots \ldots, n$. a_{ij} refers to the element (i, j) in the distance matrix and n refer to the number of toll plazas studied (63).
- MST is concept in graph theory to filter the information contained in a weighted connected graph of n objects (nodes). It is a tree with $(n - 1)$ edges (links) that minimizes the sum of the weights (distances in our case).
- MST is constructed to visualized the important information contained in the network in D in terms of topological properties.

- MST is built by linking every element in a set of *n* together in a graph characterized by a minimal distance between the nodes.
- Kruskal algorithm is used to construct the adjacency matrix using Matlab.
- Use *Pajek* software to visualize the network topology of the entire toll plazas.

UTHM
Universiti Tun Hussein Onn Malaysia

c) Centrality measures:

to understand the importance of each node relative to the others.

Degree centrality

The number of edges adjacent to the node (Borgatti, 1995). The higher the score, the more important the toll plaza is.

$$C_D(i) = \frac{\sum_{j=1}^{n} a_{ij}}{n-1}$$

where a_{ij} refers to the element of the *i*-th row and *j*-th column of adjacency matrix A.

Betweenness centrality

The number of shortest paths that pass-through a given node (Park and Yilmaz, 2010).

$$C_B(i) = \frac{1}{(n-1)(n-2)} \sum_{j,k \in G} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

where σ_{jk} denotes the number of shortest paths between nodes *j* and *k* while $\sigma_{jk}(i)$ is the number of shortest paths that pass-through node *i*.

Closeness centrality

How close this toll plaza to other toll plazas in terms of geodesic distance (Newman, 2008)

$$C_C(i) = \left[\frac{\sum_{j=1}^{n} d(i,j)}{n-1}\right]^{-1}$$

$d(i,j)$ refers to the minimum distance (shortest path) between node *i* and node *j*.

Eigenvector centrality

To measure how high the connection of an important node with other important nodes (Lohmann, 2010).

$$C_E(i) = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} a_{ij} e(j)$$

for $i = 1,2,\ldots,n$ where

$$e = (e(1), e(2), \ldots\ldots, e(n))^t$$

is the corresponding eigenvector of adjacency matrix A associated with the largest eigenvalue $\lambda_{\max}$

Overall centrality

Lee and Djauhari (2012b) introduced this measure by using the first principle component of the data matrix of size *n* x *p*, where *n* is the number of observations or samples, and *p* is the number of variables. A linear combination of the four centrality measures with maximum variance among all possible linear combinations.

$$C_O(i) = x_1 C_D(i) + x_2 C_B(i) + x_3 C_C(i) + x_4 C_E(i)$$

where $C_D(i)$, ….., $C_E(i)$ refers to the optimal linear combination of the four centrality measures based on the principal component analysis (PCA).

d) The most important centrality measure

Generalized Variance (GV)

- Multivariate dispersion measure
- $GV = |S|$, involving p random variables

  * $|S|$ refer to the determinant of the covariance structure (Anderson, 1984)
- Simple in terms of geometric and computation

Effective variance (EV)

- New measure for multivariate data (Park & Rodriguez, 2003)
- Compare two multivariate dispersions with different number of variables
- Geometric mean of all eigenvalues of S

$$EV = |S_1|^{1/p} = (\lambda_1, \lambda_2, \dots, \lambda_p)^{1/p}$$

- Largest values represent the more dispersed distribution

Vector Variance (VV)

- Handle the limitations of GV
- Sum of squares of all elements in covariance matrix, $Tr(S^2)$
- Higher value indicate more scattered the multivariate distribution around the mean vector

Effective vector variance (EVV)

- Arithmetic mean of all squared eigenvalues

$$EVV = \frac{1}{p} Tr(S^2)$$

- Compare two multivariate dispersions with a different number of variables

EV

$$EV = \frac{s_k}{\sqrt[k]{|S|}}$$

where

s_k - variance of centrality measure

k – centrality measure

EVV

$$EVV = \frac{s_k}{\frac{1}{k}Tr(S^2)}$$

where

s_k - variance of centrality measure

k – centrality measure

The most important centrality measure is the one which **maximize** the ratio (s_k and EV) and (s_k and EVV).

Results and Discussion

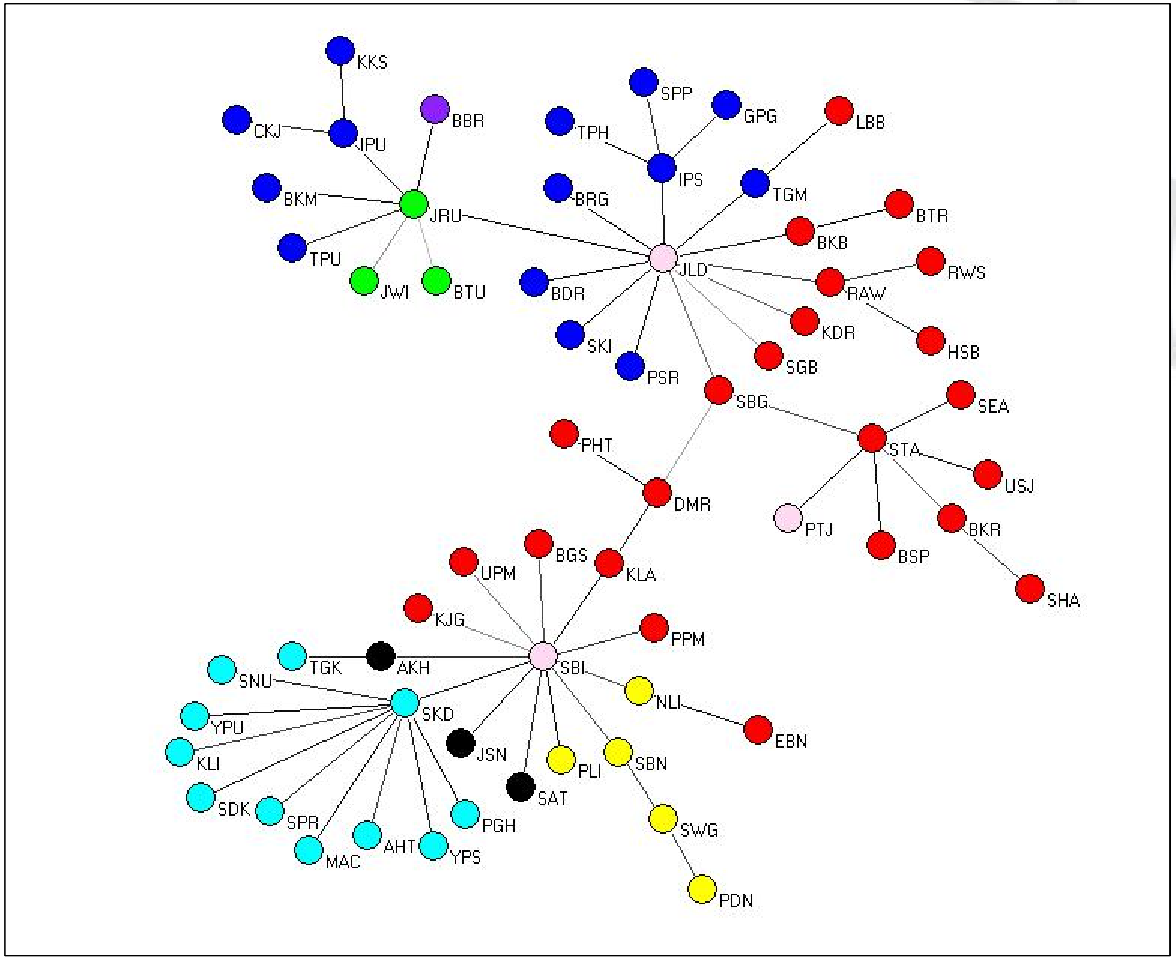


FIGURE 1. Forest representing in-coming traffic burden from 2009 until 2013

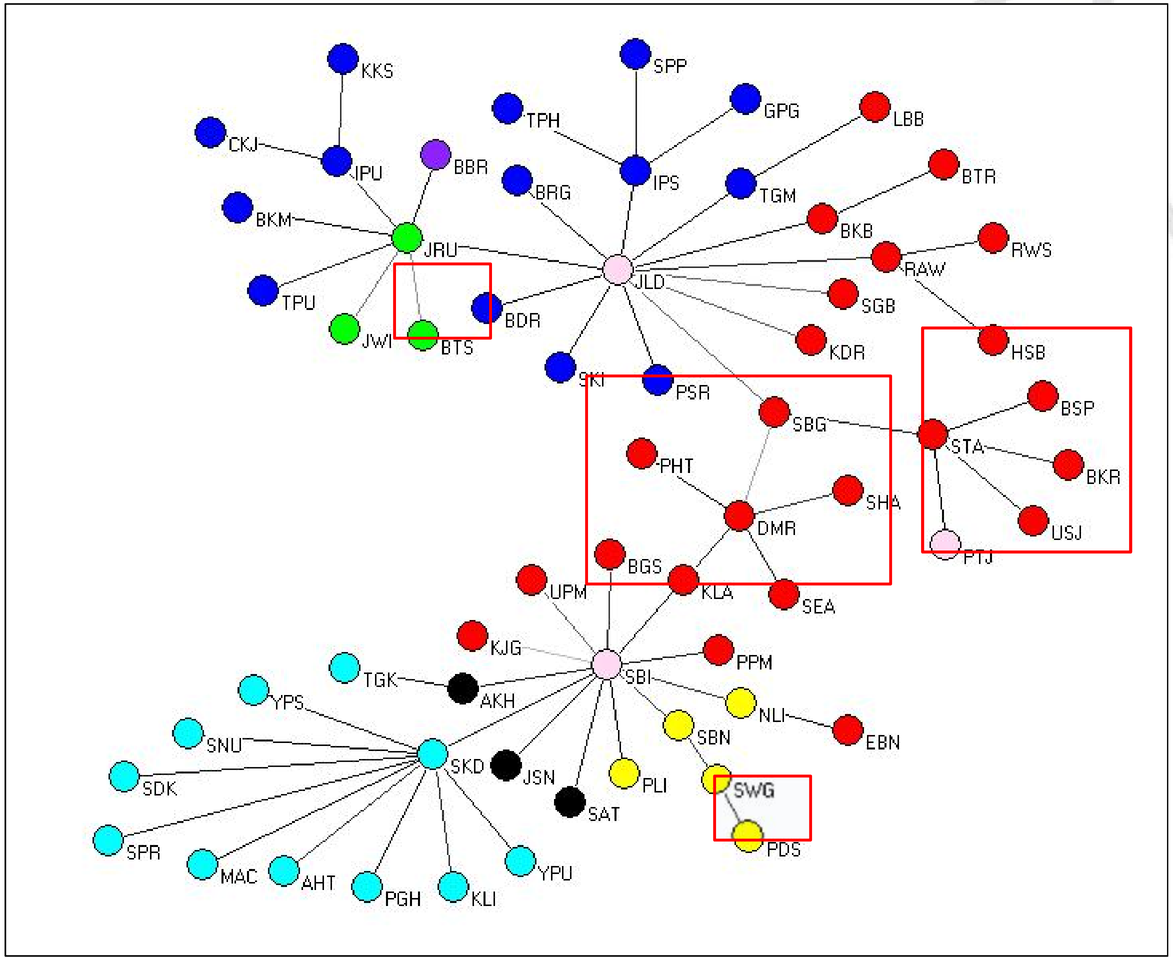


FIGURE 2. Forest representing out-coming traffic burden from 2009 until 2013

TABLE 1. Centrality measures for both in-coming and out-coming traffic burden from 2009 to 2013

Centrality Measures	In-coming	Out-coming
Degree	1. JLD & SBI (0.1905)	1. JLD & SBI (0.1905)
	2. SKD (0.1587)	2. SKD (0.1587)
	3. JRU (0.1111)	3. JRU (0.1111
Betweenness	1. JLD (0.6420)	1. JLD (0.6282)
	2. SBG (0.6034)	2. SBI (0.5695)
	3. SBI (0.6018)	3. SBG (0.5621)
Closeness	1. SBG (0.3069)	1. SBG (0.3047)
	2. DMR (0.2967)	2. DMR (0.2987)
	3. JLD (0.2938)	3. JLD (0.2928)
Eigenvector	1. SBI (0.5924)	1. SBI (0.5881)
	2. SKD (0.4356)	2. SKD (0.4310)
	3. KLA (0.1754)	3. KLA (0.1783)
Overall	1. JLD (0.0105)	1. JLD (0.0099)
	2. SBI (0.0100)	2. SBI (0.0093)
	3. SKD (0.0067)	3. SKD (0.006)

TABLE 2. Effective variance (EV) and effective vector variance (EVV)

	Effective Variance (EV)		**Effective Vector Variance (EVV)**	
	In-coming	**Out-coming**	**In-coming**	**Out-coming**
Degree	0.6017	0.6246	25.1062	25.0125
Betweenness	10.5252	10.862	439.1797	434.9986
Closeness	0.4169	0.4401	17.3938	17.6218
Eigenvector	4.1309	4.1132	172.3678	164.7111

Conclusions & Recommendation

- **JLD** is found to be the most important toll plaza for in-coming as well as out-coming traffic burden from July 2009 until December 2013. This toll plaza is scored the highest in degree centrality, eigenvector centrality and overall centrality measures.
- Meanwhile, SBI is scored the highest in degree and eigenvector centralities measures and third highest result for betweenness centrality measure.
- JLD and SBI have many connections with other toll plazas. They have very high potential to control the traffic flow in the PLUS highway, and are the most important toll plazas in the PLUS highway. They also receive the highest number of vehicles that enter from both the southern and the northern destinations.
- Both toll plazas also scored the highest ranking in betweenness centrality measure. These toll plazas have to be given the most attention by PLUS highway management in their policy and future planning.

- From EV and EVV results, it tell us that the most important centrality measures for in-coming traffic burden is betweenness, follow by eigenvector, degree and closeness centrality measures. The out-coming traffic burden also produce the same results.
- If we go through back to betweenness centrality measure, JLD, SBG and SBI are the toll plazas with higher scores. PLUS highway management have to paid more attention to these toll plaza since they produce the highest results for both EV and EVV.

References

1. E. Ivanova and J. Masarova, Importance of road infrastructure in the economic development and competitiveness, Economics and Management 18, 263-274, 2013
2. R. N. Mantegna and H. E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance (Cambridge University Press, 2000).
3. S. P. Borgatti, Centrality and AIDS. Connection. 18, 112-115 (1995).
4. S. P. Borgatti, Centrality and network flow. Social networks, 27, 55-71 (2005).
5. R. A. Hanneman and M. Riddle, Introduction to social network methods (Riverside, CA: University of California, Riverside, 2005) (published in digital form at http://faculty.ucr.edu/~hanneman/
6. S. P. Borgatti, Centrality and AIDS. Connections, 18, 112-114 (1995).
7. N. S. Yusoff, M. A. Djauhari, S. Sharif and E. S. Suleiman, Organizational commitment in Malaysian public university: An evidence via social network analysis. International Journal of Basic& Applied Sciences. 12, 17-21(2012).
8. M. A. Djauhari, S. Sharif and H. Djauhari, Network analysis on safety culture and worker's behaviour: A Forest of all minimum spanning trees, International Journal of Basic & Applied Sciences 12, 29-37 (2012).
9. S. L. Gan and M. A. Djauhari, An overall centrality measure: The case of U.S stock market, International Journal of Basic & Applied Sciences 12,99-103 (2012).
10. M. J. Naylor, L. C. Rose and B. J. Moyle, (2007) Topology of foreign exchange markets using hierarchical structure methods, Physica A 382, 199-208 (2007).
11. M. A. Djauhari, A robust filter in stock networks analysis. Physica A: Statistical Mechanics and Its Applications. 391, 5049-5057 (2012).
12. M. A. Djauhari and S. L. Gan, S. L., Minimal spanning tree problem in stock networks analysis: An efficient algorithm. Physica A: Statistical Mechanics and its Applications. 392, 2226-2234 (2013a).

13. M. A. Djauhari and G. S. Lee, Network topological property of English dialects similarity: A robust filter approach. American Journal of Applied Sciences. 10, 646-653 (2013).
14. T. Opsahl, F. Agneessens and J. Skvoretz, Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks. 32, 245-251 (2010).
15. M. E. Newman, The mathematics of networks. The New Palgrave Encyclopedia of Economics 2, 1-12 (2008).
16. K. Park and A. Yilmaz, (2010). A Social Network Analysis Approach to Analyze Road Networks. In ASPRS Annual Conference. San Diego, CA.
17. L. C. Freeman, Centrality in social networks: Conceptual clarification. Social Networks 1, 215–239 (1978).
18. V. Latora and M. Marchiori, (2004). A measure of centrality based on the network efficiency. arxiv: con-math 0402050.
19. Department of Statistics, Malaysia (2016) Statistical handbook Malaysia

UTHM
Universiti Tun Hussein Onn Malaysia

PREMIER TECHNICAL UNIVERSITY

Thank you