A New Method for Identifying the Central Nodes in Fuzzy Cognitive Maps using Consensus Centrality Measure

Mamoon Obiedat *, Sandhya Samarasinghe * and Graham Strickert b

* Centre for Advanced Computational Solutions (C-fACS), Lincoln University, Canterbury, New Zealand.
 b Centre for Global Water Security, University of Saskatchewan, Canada

Email: mamoon.obiedat@lincolnuni.ac.nz

Abstract: The Fuzzy Cognitive Map (FCM) provides a robust model for knowledge representation. FCM is a fuzzy signed weighted directed graph that depicts the knowledge of the domain as nodes representing the factors of the domain and arcs representing the connections among these factors. The centrality of a node in FCM, also called the importance of a node in this paper, is considered the most important index of all the graph theory indices applying to FCM which helps decision makers in analysing their FCM models. By finding the centrality values of the nodes in FCM, the important (central) nodes, which are the focal point for decision makers, are determined. The highest centrality value of a node in FCM is the most important one.

Little research has addressed the centrality of the nodes in an FCM using only the degree centrality measure. The degree centrality measure only accounts for the direct connections of the node. Although the degree centrality index is considered an important measure in determining the centrality of a node in an FCM, it is not sufficient and has significant shortcomings; it ignores the importance of the indirect connections, the role of the node’s position and flow of information through that node, i.e., how a node is close to other nodes and how the node contributes to the flow of information (communication control) through that node.

In the literature, there are other centrality measures that can handle direct and indirect connections to determine the central nodes in a directed graph. This paper presents a new method for identifying the central nodes in an FCM. In order to achieve that, we provide, in addition to the degree measure, new important measures to overcome the above drawbacks. These new centrality measures are: betweenness and closeness measures. In this paper, we calculate and normalize the three centrality measures values for each node in the FCM. These values are then transformed into linguistic terms using 2-tuple fuzzy linguistic representation model. We use the 2-tuple model because it describes the granularity of uncertainty of the fuzzy sets and avoids the loss of information resulted from the imprecision and normalization of the measures. The calculated centrality measures values for each node in the FCM are then aggregated using a 2-tuple fuzzy fusion approach to obtain consensus centrality measure. The resulting aggregated values are then ranked in descending order to identify the most central nodes in the FCM, and this would improve the decision-making and help in simplify the FCM by removing the least important nodes from it. Finally, a list of future works related to this paper is suggested.

Keywords: Fuzzy cognitive maps, centrality measures, central nodes, 2-tuple fuzzy linguistic representation and fusion models.
1. INTRODUCTION

Kosko (1986) proposed FCM as an extension of cognitive map (CM) introduced by Axelord (1976) by adding two characteristics. Firstly, the connections among nodes in an FCM take fuzzy values in the interval [-1, 1] or linguistic values such as low, medium, and high etc. Secondly, the FCM allows feedback loops. Figure 1, an example of CM, depicts the causal knowledge on the Middle East peace policy (Kosko, 1986). A positive sign of the connection in the CM between nodes $C_i$ and $C_j$ means that increase/decrease in node $C_i$ causes increase/decrease in node $C_j$. A negative sign of the connection in the CM between nodes $C_i$ and $C_j$ means that increase/decrease in node $C_i$ causes decrease/increase in node $C_j$. In Figure 1, Syrian control of Lebanon will be increased if Soviet imperialism increases and it will be decreased if PLO Terrorism increases. The CM can be coded into adjacency matrix as in Table 1 for easy mathematic calculations, where the positive sign (+), the negative sign (-) and no connection between two nodes in CM are represented by 1, -1 and 0, respectively in the adjacency matrix.

FCM model supports realistic and robust means for the decision makers to depict and analyse the domain knowledge and structure in order to make their decisions in an effective and accurate manner (Dickerson and Kosko, 1993; Ozesmi and Ozesmi, 2004). The degree centrality of a node, one of the graph theory measures (Harary et al., 1965), reflects the importance of the node i.e., how much an important role a node plays or contributes in the FCM (Kosko, 1986).

The degree centrality measure has been used to analyze the structure of social FCMs by characterising their most important nodes (Ozesmi and Ozesmi, 2004; Strickert, 2009). Altay and Kayakutlu (2011) utilized the degree centrality measure of a node in FCMs to prioritize and rank the factors (criteria) for decision making in complex applications and used this rank to reduce the excess number of criteria in order to make it a realistic and robust decision making process.

FCM is a dynamic system and it allows feedback loops (cycles). This means that the effect of change in one node on other nodes in turn and through indirect connections can affect the node initiating the change. Thus, the indirect paths between each pair of nodes, namely, the shortest paths must also be considered for identifying the importance of a node. The degree centrality measure of a node gives a strong indication about the contribution of a node in FCM and articulates how this node is connected directly to other nodes. In order to involve the indirect paths in calculating the central nodes in FCM, we consider the betweenness and closeness centrality measures in this calculation.

The betweenness centrality measure of a node introduced by Freeman (1977) and based on the idea in Bavelas (1948) counts the number of shortest paths going through that node. That means it emphasizes the value of the communication control (Del Pozo et al., 2011). The closeness centrality measure of a node presented by Sabidussi (1966) sums the shortest path distances from that node to every other node. That means it emphasizes the value of the quickest paths.

In this paper, we use the above three measures to obtain a consensus centrality measure based on 2-tuple fuzzy linguistic representation and fusion. This paper highlights the steps used in achieving this objective as follows: in section 2, we review in some detail the concept of the centrality measures and 2-tuple fuzzy representation and fusion methods. In section 3, we present our proposed approach. We apply the proposed method using a real life example and discuss the results in section 4. Finally, we conclude the paper and suggest some future works in section 5.

<table>
<thead>
<tr>
<th>Table 1. The adjacency matrix of Figure 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>C1</td>
</tr>
<tr>
<td>C2</td>
</tr>
<tr>
<td>C3</td>
</tr>
<tr>
<td>C4</td>
</tr>
<tr>
<td>C5</td>
</tr>
<tr>
<td>C6</td>
</tr>
</tbody>
</table>
2. PRELIMINARIES

2.1. Centrality Measures

The centrality measures have been used for analysing social networks and identifying central nodes of these networks (Alvin Chin, 2007; Del Pozo et al., 2011; Freeman, 1978). Such networks are represented by directed or undirected and weighted or unweighted graphs. A directed graph, also called digraph \( G \), which is our attention in this paper, is represented by \((V, E)\) pair, where \( V \) is a set of nodes and \( E \) is the set of directed connections among nodes. It is worth mentioning that a digraph can represent any kind of social or cognitive structure in the form of FCM. The most used centrality measures to identify a central node in a digraph, which we will review in the following subsections, are: degree, betweenness and closeness.

Degree Centrality Measure

The degree centrality is a local centrality measure because it is determined by only its directed connections. The degree centrality of a node is the summation of its absolute incoming (indegree) and outgoing (outdegree) connection weights:

\[
D_C(v) = \sum (id(v) + od(v))
\]  

(1)

where the indegree \( id(v) \) is the summation of connection weights entering node \( v \), and the outdegree \( od(v) \) is the summation of connection weights exiting node \( v \). With regard to FCM, the degree centrality is a measure of how strong direct connections of a node is with other nodes in the FCM.

Betweenness Centrality Measure

The betweenness is a global centrality measure because it is based on the shortest paths between node pairs in the graph. The betweenness centrality of a node is determined by summing the proportion of shortest paths between node pairs that go through that node (Freeman, 1977). Given the directed graph \( G^m(V, E) \), the betweenness centrality of a node \( v \) is defined as:

\[
B_C(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}
\]  

(2)

where \( \sigma_{st} \) is the number of shortest paths from node \( s \) to node \( t \) and \( \sigma_{st}(v) \) is the number of shortest paths from \( s \) to \( t \) that passes through node \( v \). With regard to FCM, the betweenness is a measure of how much a node controls communication between other nodes and spreads information through the FCM.

Closeness Centrality Measure

Like betweenness, the closeness is a global centrality measure based on the shortest paths concept. It measures how much a node is close to all other nodes in the graph. For a directed graph \( G^m(V, E) \), the closeness centrality of a node \( v \) is defined as:

\[
C_C(v) = \frac{1}{\sum_{t \neq v} d_G(v, t)}
\]  

(3)

where \( t \neq v \), and \( d_G(v,t) \) is the shortest path between nodes \( v \) and \( t \). With regard to FCM, the closeness is a measure of how quickly a node communicates with other nodes in the FCM.

2.2. A 2-Tuple Fuzzy Linguistic Representation Model

A 2-tuple fuzzy linguistic representation model (Herrera and Martínez, 2000a) was developed based on a symbolic linguistic model (Delgado et al., 1993). Let \( S = \{s_0, ..., s_g\} \) be a linguistic term set such as low, medium, and high etc., where \( s_i < s_j \) if and only if \( i < j \) and \( g + 1 \) is the number of linguistic terms. The symbolic model represents the linguistic information by an ordered linguistic term set like \( S \) and uses the ordered index \( i, i \in [0, g] \), of the linguistic term set \( S \) to execute the calculations.
In the 2-tuple, the linguistic term is represented by a 2-tuple \((s, \alpha)\) where \(s\) is a linguistic term and \(\alpha \in [-0.5, 0.5)\) is a symbolic translation. Based on the above, the 2-tuple model represents the linguistic information in continuous approximation process and so, it handles any counting of information resulting from symbolic aggregation in the universe of the discourse, and consequently, the results from any fusion process will match one of the initial linguistic terms without loss of information.

To clarify the 2-tuple model, let \(S = \{s_0, \ldots, s_g\}\) be a symbolic linguistic term set where each linguistic term \(s_i\) is represented by triangular parameters \((a_i, b_i, c_i)\) and \(i = 0, \ldots, g\) is the index of \(s_i\), and \(\mu_{s_i}(n)\) is membership function values associated with \(s_i\) for a particular value \((n)\) of a centrality measure and computed using the following Equation:

\[
\mu_{s_i}(n) = \begin{cases} 
\frac{n - a_i}{b_i - a_i}, & a \leq n \leq b \\
\frac{c_i - n}{c_i - b_i}, & b < n \leq c \\
0, & \text{otherwise}
\end{cases}
\]  

(4)

Let \(\beta \in [0, g]\) be the result of a weighted aggregation of the labelled indices in the linguistic term set \(S\) as in Equation 5. Let \(i = \text{round}(\beta)\), \(i \in [0, g]\), and \(\alpha = \beta - i\), then \(\alpha\) is called a symbolic translation of a linguistic term \(s_i\) that supports the "difference of information" between a counting of information \(\beta \in [0, g]\) resulting from a symbolic aggregation operation and the closest value in \([0, \ldots, g]\) that indicates the index of the closest linguistic term \(s_i\) in \(S\) (Herrera and Martinez, 2000a). Based on the above, the equivalent 2-tuple \((s_i, \alpha)\) to \(\beta\) denoted as \(\Delta(\beta)\) and the equivalent \(\beta\) to 2-tuple \((s_i, \alpha)\) denoted as \(\Delta^{-1}(s_i, \alpha)\) are obtained by Equations 6 and 7, respectively.

\[
\beta = \frac{\sum_{i=0}^{g} i \cdot \mu_{s_i}}{\sum_{i=0}^{g} \mu_{s_i}}
\]

(5)

\[
\Delta(\beta) = \begin{cases} 
s_i, & i = \text{round}(\beta) \\
\alpha = \beta - i, & \alpha \in [-0.5, 0.5)
\end{cases}
\]

(6)

\[
\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta
\]

(7)

Example, let \(\beta = 2.6\) (Figure 2) is the value of the result of a symbolic aggregation assessed in the above linguistic term set \(S\). Then \(i = \text{round}(2.6) = 3\), \(a = 2.6 - 3 = -0.4\), and the representation of \(\beta = 2.6\) in 2-tuple is: \(\Delta(2.6) = (s_3, -0.4)\).

3. PROPOSED METHOD

The main objective of this paper is to obtain a consensus centrality measure from three aforementioned centrality measures for identifying the central nodes in an FCM and then ranking them according to their importance. In order to do that, we firstly compute the values of each measure for the nodes separately. Secondly, we convert numeric centrality measures into symbolic fuzzy linguistic terms, and then we use the
2-tuple fuzzy linguistic representation model to transform the linguistic terms into 2-tuple fuzzy linguistic set. Thirdly, we use the 2-tuple fusion approach and arithmetic mean operator to combine the 2-tuple fuzzy linguistic terms of the measures. Consequently, combined value of the node centrality represents a node’s importance in comparison with other nodes in FCM, i.e., the larger the node centrality value is the higher the node’s importance. Finally, we rank the nodes in descending order according to their importance. We can use this list to simplify the FCM by removing the least important nodes based on a threshold that determines the number of required nodes in the FCM. Moreover, these centrality values of the nodes give an indication of the most important nodes in FCM that helps the decision maker in analysing the structure of the FCM and making decisions.

To compute the degree centrality of a node, we represent the FCM in an adjacency matrix $E$ of dimension $N \times N$, where $N$ is the number of nodes in the FCM and each element in the matrix $e_{ij}$ represents the connection weight from node $i$ to node $j$. Then, we compute the indegree $id(i)$ and outdegree $od(i)$ for each node in FCM using the Equations 8 and 9, respectively: Then, we apply Equation 1 to compute the degree centrality for all nodes in FCM.

$$\text{id}(i) = \sum_{j=1}^{N} e_{ji}$$

$$\text{od}(i) = \sum_{j=1}^{N} e_{ij}$$

For the betweenness and closeness centrality measures, we use simple computations to calculate the shortest paths between each pair of nodes in terms of the number of steps and save them in a distance matrix $D$. The element of the distance matrix $d(v,t)$ represents the shortest path between node $v$ and node $t$ and it takes an infinite value if there is no path between them. Then, we apply Equation 2 to compute the betweenness centrality for all nodes in FCM.

For closeness, the infinite values in the distance matrix $D$ cause a problem in computing these centrality values. To avoid that, we convert the distance matrix $D$ to a converted distance matrix ($C$) (Botafogo et al., 1992), where the elements of matrix $C$ are computed as follows:

$$c(v,t) = \begin{cases} d(v,t), & \text{if } (d(v,t) \neq \infty) \\ K, & \text{otherwise} \end{cases}$$

where $K$ is a finite conversion constant that usually takes a value related to the number of nodes. In this paper, we consider the value of $K$ equals to number of nodes in FCM $N$. Then, we apply Equation 3 to compute the closeness centrality for all nodes in FCM.

Then, we normalize each centrality measure using min-max normalization method to scale its values between 0 and 1. Then, we convert these normalized values into symbolic fuzzy sets (Figure 3). Those sets are then transformed into 2-tuple fuzzy sets using the approach in Herrera and Martinez (2000b). Then, we use the 2-tuple fuzzy fusion approach in Herrera (2000) and arithmetic mean aggregation operator to combine the 2-tuple linguistic terms of the centrality measures. The results from fusion process are then sorted in descending order in order to identify which nodes in the FCM are the most important and which ones are the least important. This ranking leads to further analysis of the FCM and its simplification.

### 4. Real Life Example and Discussion

Now, we present a real life FCM from Strickert (2010) to illustrate the proposed method. This FCM represents a stakeholder’s perceptions on natural hazard (earthquake) mitigation in ski areas. It comprises 23 nodes (Table 2) and 256 connections. After computing and normalizing the value of centrality measures, each centrality measure values was represented individually by a symbolic linguistic term set $S = \{s_0, ..., s_6\}$. We selected the following seven triangular memberships (Figure 3) to express the

![Figure 3. The linguistic term set $S$ and its semantics.](image-url)
representations of these measures according to Equation 6 are: \(6 \times 1 / 1 = 6\), \(6 \times 1 / 1 = 6\) and \((5 \times 0.23 + 6 \times 0.77) = 5.77\), respectively. Second, the 2-tuple semantics of the linguistic term set \(S\): \(\beta\) Finally, we obtain the consensus centrality measure of the Communication node by applying the arithmetic values of this node are: 1, 1, and 0.96, respectively. We represent each of these numerical values \(n\) by a symbolic linguistic term set \(S\) to compute their membership values associated with the seven linguistic terms of \(S\) using Equation 4. Thus, the symbolic linguistic representation \((s_\beta, \mu_{s_\beta}(n))\) of the three measures of the Communication node in \(S\) are: \(\{s_0, 0\}, (s_1, 0), (s_2, 0), (s_3, 0), (s_4, 0), (s_5, 0), (s_6, 0)\}\) and \(\{s_0, 0\}, (s_1, 0), (s_2, 0), (s_3, 0), (s_4, 0), (s_5, 0.23), (s_6, 0.77)\}, respectively.

Then, we transform the symbolic linguistic terms into 2-tuple linguistic sets. First, we use Equation 5 to calculate \(\beta\). The \(\beta\) values of the degree, betweenness and closeness centrality measures of the Communication node are: \(6 \times 1 / 1 = 6\), 6 * 1 / 1 = 6 and \((5 \times 0.23 + 6 \times 0.77) = 5.77\), respectively. Second, the 2-tuple representations of these measures according to Equation 6 are: \(\Delta(6) = (s_{round}(0), 0 - round(6)) = (s_0, 0)\), \(\Delta(6) = (s_{round}(0), 6 - round(6)) = (s_6, 0)\) and \(\Delta(5.77) = (s_{round}(5.77), 5.77 - round(5.77)) = (s_0, -0.23)\), respectively.

Finally, we obtain the consensus centrality measure of the Communication node by applying the arithmetic mean operator to combine \(\beta\) values \([6 + 6 + 5.77] / 3 = 5.92\) and using Equation 6 again to obtain the 2-tuple value equivalent to combined \(\beta\) value, \(\Delta(5.92) = (s_{round}(5.92), 5.92 - round(5.92)) = (s_0, -0.08)\), (Table 2).

Based on Table 2, the first, second and third most important nodes in the FCM are: Communication, Safety and Emergency, respectively. The first one is the most important node according to degree and betweenness measures, while it was ranked as the second important node according to closeness measure. The most important node according to closeness measure is Recovery of Ski Area which is the fourth important node according to consensus measure. In contrast, all the centrality measures agree that the two least important nodes on this FCM are Equipment and Bad Timing. The High Intensity was the third least important node in accordance with betweenness, closeness and consensus measures, while the third least important node in accordance with degree measure is Hazard Awareness.

We note from this case that the ranking of the nodes according to the three centrality measures is very similar to the ranking of the consensus measure with a slight difference. But this does not necessarily mean that this applies to all cases. The result is determined by several factors, i.e., number of connections, the frequency that a node lies on the path between other nodes and the ability of a node to communicate quickly with the other nodes.

When the FCM is large or complex, it is very difficult to gain insight and understand its structure and it involves expensive calculations too. For easy calculations and achieving a simple structure of FCM, it is essential to simplify it into a smaller one with few nodes and connections. In this paper, we utilize the
ranking of nodes according to their importance to remove the excess nodes and their connections from the FCM. We consider that the least important nodes are the excess nodes and use a threshold to remove them. For example, as we have seen from Table 2, the High Intensity, Bad Timing and Equipment are the least important nodes and could be removed from the FCM.

Another advantage of identifying the central nodes in an FCM is to use them as a motivation for what-if scenario simulations. For example, the FCM can be used as a dynamical system to simulate policy options where central nodes are clamped at high values to assess their effect on the other nodes when the system has reached a steady state.

5. CONCLUSION AND FUTURE WORKS.

In this paper, we combined the degree, betweenness and closeness centrality measures to reach a consensus centrality measure for identifying central nodes in FCM. The degree centrality measure is a local measure characterizing the direct connections, but the betweenness and closeness centrality measures are global measures describing the role of the node’s position and depend in their calculations on the shortest paths between nodes. Using a 2-tuple fuzzy linguistic representation of the centrality values followed by 2-tuple fuzzy fusion for combination, we obtained a consensus centrality measure and highlighted the approach using a real life example. Also, we ranked the nodes in the FCM from the most central node to the least central one. This ranking can be used to simplify a complex FCM. As future works that will be our attention, we plan to apply the centrality measures for identifying the central connections in an FCM.

6. REFERENCES.


Strickert, G., S. Smarasinghe, and C. Doscher. (2009). Mixing with the mountains: A triangulated super-physical true-social approach for policy advancement to improve resilience to low probability high consequences catastrophic natural events for mountainous communities. Journal of Natural Resources Policy Research,