

# Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector

Youngseok Choi<sup>1</sup> · Habin Lee<sup>1</sup>  · Zahir Irani<sup>1</sup>

Published online: 17 August 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The prevalence of big data is starting to spread across the public and private sectors however, an impediment to its widespread adoption orientates around a lack of appropriate big data analytics (BDA) and resulting skills to exploit the full potential of big data availability. In this paper, we propose a novel BDA to contribute towards this void, using a fuzzy cognitive map (FCM) approach that will enhance decision-making thus prioritising IT service procurement in the public sector. This is achieved through the *development of decision models that capture the strengths of both data analytics and the established intuitive qualitative approach*. By taking advantages of both data analytics and FCM, the proposed approach captures the strength of data-driven decision-making and intuitive model-driven decision modelling. This approach is then validated through a decision-making case regarding IT service procurement in public sector, which is the fundamental step of IT infrastructure supply for publics in a regional government in the Russia federation. The analysis result for the given decision-making problem is then evaluated by decision makers and e-government expertise to confirm the applicability of the proposed BDA. In doing so, demonstrating the value of this approach in contributing towards robust public decision-making regarding IT service procurement.

**Keywords** Big data analytics · Fuzzy cognitive map · Decision modelling · IT service procurement · Simulation

## 1 Introduction

Decision-making and planning regarding procurement, as a part of Supply Chain Management is a fundamental and essential business process that relates to the economic efficiency of overall supply chain associated with service and product delivery. Sadrian and Yoon (1994)

---

✉ Habin Lee  
habin.lee@brunel.ac.uk

<sup>1</sup> Brunel Business School, Brunel University London, Kingston Lane, Uxbridge UB8 3PH, UK

explain that procurement decision-making is considered a necessity in supply chain management (SCM) due to the uncertainties associated with demand, procurement budget, and the impact of procurement. To cope with the ever-changing global market, it has been crucial for firms to be able to exploit and develop their competitive advantage by achieving effective and efficient procurement management practices (Sharif and Irani 2006a; Lau et al. 2005; Piotrowicz and Irani 2010). While the importance of procurement decision in public sector has received increasing attention, there is still a scarcity of studies regarding public procurement (Love et al. 2012; Preuss 2009).

Public procurement is concerned with how public sector organizations spend tax-payers' money on goods and services. Public procurement is generally guided by principles of transparency, accountability, and achieving value for money for tax payers, i.e., citizen (Sharif et al. 2010; Walker and Brammer 2009). In particular, the effectiveness of public service procurement has been emphasised as a major challenge in recent years (Grudinschi et al. 2014) and it is here where the authors of this paper seek to make a contribution to the normative literature.

e-Government has been widely spread and implemented by local and central government, IT service and infrastructure has been one of the major public procurements and investments (Rose et al. 2015; Tucci and Poulin 2015; Osman et al. 2014; Irani et al. 2008). Once the demand for IT service and infrastructure is defined and specified, decision makers in public sector need to find the best way of spending their budget to fulfil the needs from public because many different public IT service and infrastructure for the same purpose can have different effect and impact. Moreover, the public services can make multi-dimensional impact through various economic and social factors, so, it is not easy to predict through modelling, and then evaluate the impact of the decision on the procurement of public service. In this regard, decision-making for the prioritisation of public IT service procurement in public sector is very important especially during austere times. This is especially the case given local governments need to demonstrate best “value for money” given their limited budgets thus demonstrating the best levels of efficiency and effective procurement decision-making is ever sought with heightened levels of transparency.

Big data analytics (BDA) can play an important role in this type of decision-making, especially in regards to public service procurement; with prediction techniques using large data sources to evaluate what would have happened under different circumstances (Waller and Fawcett 2013; Schoenherr and Speier-Pero 2015; Sharif and Irani 2006b). In addition, BDA can dramatically improve operational and supply chain decision-making by evaluating the strategy and improving data-driven forecasting (Sanders 2014) and also can provide the sustainability for supply chain by analysing the relevant data (Papadopoulos et al. 2016). However, contemporary BDA tools are difficult to use for public sector decision makers who are more familiar with model based decision-making that are influenced by smaller data sets and more qualitative contributions. This paper proposes a model driven data analytics approach to support decision-making in supply chain management for public sector thus allowing the wide spread adoption of big data sets. A fuzzy cognitive map; an easy to use decision modelling framework, is proposed as a new approach that integrates big data analytics to support evidence based decision-making and impact analysis. Such approaches have been used to support the modelling of logistics of information management.

Big data usually indicates data sets with sizes beyond those easily manipulated by commonly used tools and requires a more dedicated and sophisticated approach to analytics (Snijders et al. 2012). Much discussion about the definition, history, and properties of big data from industry and academia has emerged since big data received large attention following clarity around its potential. The explosion of data is nothing new since human invented

electronic device to compose and store information. Even in 1960s, [Marron and de Maine \(1967\)](#) noted the need of handling explosive amount of data as saying “‘the information explosion’ noted in recent years makes it essential that storage requirements for all information be kept to a minimum”. However, the nature of data, use and information explosion is very different from that of past as big data nowadays is characterised with more dimensions notably 5V (volume, variety, velocity, variability, and veracity) and 1C (complexity) ([Hilbert 2013](#)). There is also a greater sense of potential to realise competitive advantage through carefully data-mined and exploited data that when structured provides meaningful management information upon which robust strategic, tactical or operational decisions can be taken.

In particular, big data in public sector has different attributes from private sector. While big data in private sector is characterised as 5V and 1C, as we stated, the big data in government and public sector can be characterised with 2S (Silo and Security) and 1V (Variety) ([Kim et al. 2014](#)). As enormous amount of data sources in legacy database are dispersed in different public authorities and organisation, organising these various data source is very critical for gaining competitive advantage from BDA in public sector.

For this reason, the lack of control tower for BDA and dispersed silo have been considered as the big challenges for BDA in government. By developing big data supporting systems such as data control tower and portal with highly secured technology, BDA can utilise the data for public decision making with competitive advantage ([Amankwah-Amoah 2015](#); [Lu 2014](#)). To overcome these critical challenge for BDA in public sector, European Union, United States, and some Asian countries like Singapore and Korea have shown an active interest in building data portal that provides integrated view on relevant data sources to specific topics such as health care and local governments ([Kim et al. 2014](#)). The BDA platform based on fuzzy cognitive maps proposed in this paper is aimed at supporting integrated view of diverse open data sources.

BDA is increasingly seen as having the potential to deliver a competitive advantage throughout the supply chain ([Sanders 2016](#)) and data-driven decision support in the supply chain context has started to appear in the literature, see [Hilletoft et al. \(2016\)](#). The efficiency of decision-making for supply chain management can be improved by collecting and analysing those data that can provide a better basis for understanding the causality of environmental variables throughout the supply chain ([Gimenez and Ventura 2005](#); [Lummus and Vokurka 1999](#)). Accordingly, there has been a significant demand for sophisticated decision support in the supply chain context based on data analytics. In doing so, making decisions efficient and effective throughout the supply chain ([Hilletoft et al. 2016](#)). To do so however, requires appropriate and accurate information ([Hazen et al. 2014](#)) or data sets thus allowing managers to predict the outcomes of decisions and how these may affect the entire supply chain ([Hilletoft et al. 2010](#)) that is often made possible through modelling techniques.

However, while the potential benefits of BDA are exponential for procurement decision-making across the public sector, governments and data providers face steep practical, legal and ethical barriers when seeking to exploit big open data ([Brown et al. 2011](#)). Most emergent BDAs are focusing on the data-driven approach and methods to convey the analytic result to decision makers ([Kambatla et al. 2014](#)) but decision makers often struggle to interpret the results and hidden analytic process ([Labrinidis and Jagadish 2012](#)). This represents one of the key drawbacks of data-driven decision modelling. For this reason, much BDA research points out the importance of visualisation and presentation of big data and its analytic results ([Cuzzocrea et al. 2011](#); [McAfee and Brynjolfsson 2012](#); [Miller and Mork 2013](#)), however, visualisation is still limited to cover the data summary and report ([Hashem et al. 2014](#)) and offers limited interaction and ability to assess causality. This drawback of BDA, which mainly

depends on data-driven approach can be critical to supporting decision-making in the public sector in which large numbers of decision variables are inter-related through causal-effect relationships (Dunn 2015). Decision makers are more familiar with easy to use diagram based policy models like cognitive maps (Axelrod 2016). Such decision models, however, in most cases, are developed based on subjective opinion of decision makers and lacks linkage to factual data. Therefore, there are increasing needs to integrate visualised decision models with factual data for evidence-based decision-making.

Given the void in the literature, this study proposes an innovative framework for decision modelling and impact analysis for efficient and effective IT service procurement in public sector based on open big data. More specifically, this paper offers a fuzzy cognitive map (FCM) approach, which is particularly suited to modelling complex and dynamic social problems (Mago et al. 2013) by integrating with BDA techniques and has been used in many sectors. This innovative approach allows decision makers to *develop decision models and evaluate the impact of options by capturing the strengths of both data analytics and intuitive qualitative approach*. It is here where the authors add to the normative literature. As a qualitative modelling technique, an FCMs have traditionally been applied to model the decision-making problems that exist in various fields such as medicine, politics, environmental science, etc (Papageorgiou and Salmeron 2013). FCMs are easy to understand and intuitive for describing decision-making problems. However, BDA has strength in identifying the structure of target problems through sound formalisms (Esposito et al. 2014).

The integration of FCMs with BDAs offers an academic contribution to the decision support discipline. The two main approaches for decision modelling are data-driven and model-driven decision modelling. Data-driven decision modelling is a decision support approach mainly using the data and analytic algorithm to find a final decision. It is an appropriate approach to well defined and structured decision-making with vast amounts of data that human cannot manually view and check (Power 2008). For this reason, most BDA BI have complex data-driven decision modelling modules behind their dashboard (Kambatla et al. 2014). In contrast, model-driven decision modelling uses qualitative models or formal representation to describe the relevant variable to decision-making (Power and Sharda 2007; Bhargava et al. 2007). It is suitable to model the dynamic interaction among decision variable by giving decision makers the opportunity to tune their decision model for final decision-making (Morton et al. 2003). As existing BDA and business intelligence (BI) depend on data-driven decision modelling (Power et al. 2015), they fail to integrate the strength of model-driven decision modelling. Models in a model-driven decision modelling provide a simplified representation of a situation that is understandable to a decision maker (Bonczek et al. 2014) so this approach makes a non-technical specialist easily accessible to decision-making (Power and Sharda 2007). An FCM approach is based on the utilisation of big data, providing a novel direction to decision support by combining model-driven and data-driven decision modelling and this integration can make BDA capture the strength of both approach.

However, the integration of FCMs with BDAs poses a significant academic challenge due to the characteristics of big data. Existing efforts to integrate BDA in FCMs have mainly focused on applying learning algorithms to fine tune FCMs but have limitation in scalability and applicability in big data context. Public sector which is facing with the sheer volume of public open big data demands more scalable and simple approach to BDA (Jin et al. 2015). The learning process for the weight matrix calculation of an FCM using data in this study is newly devised using simple optimisation problems while previous studies on FCM learning depend on the non-scalable learning method such as Hebbian learning. This simplicity ensures the scalability and applicability of FCM approach to complex decision-making process using big data. The proposed approach is applied to a real decision-making problem regarding

IT service supply for information society in cooperation with a local government in Russia Federation.

The paper is organized as follows. The literature reviews on FCMs in the context of data utilisation are briefly introduced in Sect. 2. The details of the proposed BDA method for decision-making regarding decision modelling and impact evaluation are presented in Sect. 3. In Sect. 4 we validate proposed method by showing the real application example to local government decision-making regarding IT service procurement for e-Society building. Discussion and conclusion with future study are presented in Sect. 5.

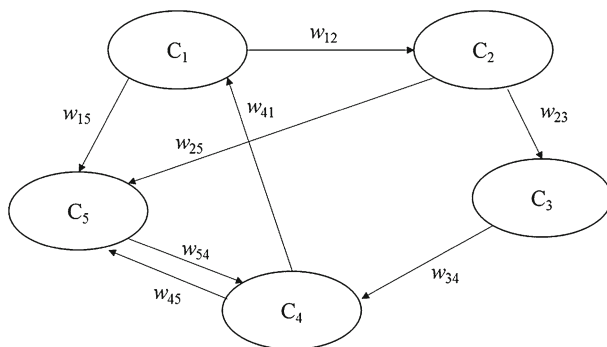
## 2 Fuzzy cognitive maps for decision modelling and impact simulation

### 2.1 Basic concept of FCM

FCMs are fuzzy signed graphs with feedback (Stylios and Groumpos 2000). An FCM is a representation of a system in a given domain (Kok 2009). It comprises concepts ( $C_i$ ) representing key drivers of the system, joined by directional relationships between concepts.

Each connection has a weight that quantifies the strength of causal relationships. An FCM models a dynamic complex system as a collection of concepts and cause–effect relationships between the concepts (Stylios and Groumpos 1999). A simple illustration of an FCM consisting of five node concepts is depicted in Fig. 1. A weight  $w_{ij}$  describes the strength of causality between two concepts. A weight takes a value in the interval  $[-1, 1]$ . The sign of the weight indicates positive causality if  $w_{ij} > 0$ , which means that an increase in the value of concept  $C_i$  will cause an increase in the value of concept  $C_j$ . Similarly, a negative value of  $w_{ij}$  indicates negative causality. When no relationship exists between two concepts, then  $w_{ij} = 0$ . The value of a concept is usually fuzzified by mapping linguistic measure (i.e., very low, low, middle, high, and very high for 5 scale measure) to a fuzzified value in the interval  $[0, 1]$ . According to the scale of the fuzzification scheme, every fuzzified concept is given with a fuzzy value. The fuzzification of linguistic measures allows decision makers transform qualitative measures into quantitative values. Thus, a cognitive map can be used as a multivariate time series prediction model.

FCMs emerged as a technique to model social, political, business, engineering and public policy issues and support corresponding decision-making processes. Andreou et al. (2003) use an FCM to find and evaluate alternative solutions for the political problem of Cyprus



**Fig. 1** A simple fuzzy cognitive map

by collecting opinion of related experts. Using a multiple scenario analysis, the value of a hybrid method is demonstrated in the context of a model that reflects the political and strategic complexities of the Cyprus issue as well as the uncertainties involved. [Giordano and Vurro \(2010\)](#) propose a methodology based on an FCM to support the analysis of stakeholders' perceptions of drought, and the analysis of potential conflicts. [Georgiou and Botsios \(2008\)](#) apply an FCM to learning style recognition. They propose a three-layer FCM schema to allow experienced educators or cognitive psychology to tune up the system's parameters to adjust the accuracy of the learning style recognition. FCMs are reported to be a worthy tool for learning-style recognition as they are effective in handling the uncertainty and fuzziness of a learning style diagnosis. [Lee et al. \(2013\)](#) apply FCM to long-term industrial marketing planning in business and management discipline.

One of the strengths of FCMs for decision makers lies on its simulation capability that allows decision makers assess the impact of changes on some of the concept values on other consequence variables. The simulation of an FCM is a process of quantifying the impact of changes on some of the concept values, based on change evaluation functions across the FCM. More specifically, the value of each concept at time  $t$  is calculated by applying the calculation rule of the equation below, which computes the influence of other concepts on the target concept:

$$x_i(t) = f \left( \sum_{j=1, j \neq i}^n x_j(t-1) w_{ji} \right) \quad (1)$$

where  $x_i(t)$  is the value of concept  $C_i$  at time  $t$ ,  $x_j(t-1)$  is the value of concept  $C_j$  at time  $t-1$ ,  $w_{ji}$  is the weight of the relationship between concept  $C_j$  and  $C_i$ , and  $f$  is the activation function. At each time step, the values of all concepts in FCM change and recalculate according to this equation. The calculation results in each iteration reflect the state of each concepts. This nature of FCM simulation enables it provide the long-term perspective of decision-making by showing the impact and change of state for each concepts. This simulation process shows not only the final value of each concepts, but also the progress how each decision variables can approach the idle state, which can be very critical information to develop a new decision and its impact.

The values of concepts in an FCM at time  $t$  can also be expressed as a matrix form. Assuming that vector  $\mathbf{X}(t)$  is the  $n$  by 1 vector that gathers the value of  $n$  concepts, then the matrix  $\mathbf{W}$  is an  $n$  by  $n$  matrix representing the weights between  $n$  concepts:

$$\mathbf{X}(t) = f(\mathbf{W}^T \mathbf{X}(t-1)) \quad (2)$$

An activation function is borrowed from artificial neural networks. It is a function that calculates the output of a concept based on its inputs, usually using a total sum operator. The output of activation function has usually  $+1/-1$  as its upper/lower bound. The most common type of activation function in FCMs is the sigmoid function, which is a reciprocal of negative natural logarithms with few parameters. In addition to this function, tangent hyperbola and linear type activation functions are applied to diverse applications.

Based on the definition of equation and activation function, a state vector that contains the values of all concepts at time  $t$  can be calculated. In a simulation of an FCM, the calculation of the state vector is iterated until the steady state is reached, indicating that no changes occurred in the state vector at that point. Not all simulation results reach to steady states. In few cases, values of concepts may fluctuate as iterations proceed, and both initial vectors of concepts and the structure of an FCM can cause unstable simulation results ([Carvalho](#)

and Tome 2002). In case an FCM simulation result fails to reach to a steady state, then it is advised to modify the structure of the FCM.

## 2.2 Fuzzy cognitive maps (FCMs) for data utilisation

As illustrated above, the simulation results of an FCM are highly dependent on the state vector ( $X$ ) and the weight matrix ( $W$ ) of the FCM. In early stage of FCM studies, the derivation of state vectors and weight matrix were based on the opinion of human experts. However, in some domains where enough and relevant data is available, algorithms for automatic learning of FCM model structure were proposed. Due to the similarity between FCM approach and neural network, most of FCM learning studies to calculate weight matrix of an FCM have their basis on Hebbian learning method.

For example, Kosko (1994) proposed a learning model by using simple Differential Hebbian Learning law (DHL). The learning process modifies weights of edges in an FCM in order to find a desired weight matrix. Papageorgiou et al. (2004) propose another extension to Hebbian algorithm, called Active Hebbian Algorithm (AHL) method, that not only determines a desired set of concepts, initial structure and the interconnections of an FCM structure, but also identifies which concepts should be activated. Another approach to learning weight matrix of FCM is application of genetic algorithms or evolutionary algorithms. Koulouriotis et al. (2001) apply the Genetic Strategy (GS) to learn FCM weight matrix.

Stach et al. (2005) applied real-coded genetic algorithm (RCGA) to calculate FCM weight matrix from a set of historical data. Konar and Chakraborty (2005) use reasoning and unsupervised learning for a special type of cognitive maps based on Petri nets. Ghazanfari et al. (2007) use Simulated Annealing and Genetic algorithm in FCM learning and compare the performances of two algorithm to find the former is superior to the latter for FCMs with more concepts. They also introduce a new method to learn weight matrix rapidly. In their study, heuristic algorithms are used to learn FCM matrix. Papageorgiou et al. (2011) apply the fuzzy decision tree that develop the fuzzy value based decision tree and then based on the path length from node to leave weight is modified.

However, the Hebbian approach has drawback on its scalability due to the potential saturation and “catastrophic forgetting” (Amin et al. 2012). The bottleneck of the RCGA method for FCMs is also due to the scalability, as the number of parameters that have to be established grows quadratically as the number of concepts increases. Furthermore, genetic optimization is time consuming when employed to problems with large number of variables (Stach et al. 2007). Therefore, existing algorithms for learning weight matrix have limitations to be applied to BDA for decision modelling and impact analysis. In this study, we estimate a weight matrix by decomposing an FCM into partial cognitive maps and applying simple parameter optimisation, which is hinted by weight calculation (Soulié and Hérault 1990; Polk and Seifert 2002) of single layer neural network. This approach enables analytics method to be simple but easily scalable to cope with FCM with many concepts and data.

## 2.3 Fuzzification and fuzzy time series

Assigning fuzzy values to concepts in FCM is the first task to be conducted for an FCM based simulation. For data-driven decision modelling and simulation via FCMs, developing a fuzzification method that matches numeric values from open data into linguistic measure, i.e. fuzzy values, is a pre-requisite. In previous studies on FCM learning, simple membership function with equal fuzzification scheme has been used without any sophisticated fuzzifica-

tion scheme. By dividing the difference of maximum and minimum value into the number of scale, each sub-interval can be easily mapped into fuzzy values.

Some studies regarding value fuzzification can be found from fuzzy time series and fuzzy set study. Fuzzy time series was introduced by [Song and Chissom \(1993\)](#) and is based on fuzzy set approach ([Zadeh 1965](#)), which consists of three main stage: Fuzzification (focus of this section), defining the fuzzy relationship and defuzzification. The definitions of terms used in fuzzy time series are given as follows: Let  $U$  the universe of discourse, where  $U = u_1, u_2, \dots, u_b$ . A fuzzy set  $A_i$  of  $U$  is defined as  $A_i = f A_i(u_1)/(u_1) + f A_i(u_2)/(u_2) + \dots + f A_i(u_b)/(u_b)$ , where  $f A_i$  is the membership function of the fuzzy set.  $u_a$  is a generic element of fuzzy set  $A_i$ ;  $f A_i(u_a)$  is the degree of belongingness of  $u_a$  to  $A_i$ ;  $f A_i(u_a) \in [0, 1]$  and  $1 \leq a \leq b$ .

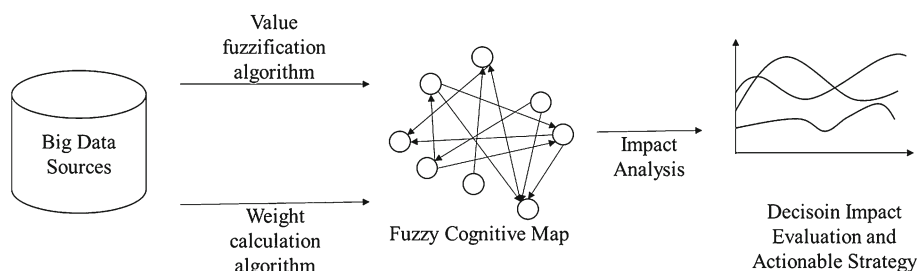
[Sullivan and Woodall \(1994\)](#) propose a method based on Markov model. They use a probability distribution function to obtain the linguistic labels. The basic idea is to assign the linguistic measures after defining interval of time series data. The results of the time-invariant Markov model is compared with those of time-invariant fuzzy time series models ([Song and Chissom 1993](#)). [Chen \(1996\)](#) propose a randomly chosen length of an interval for a fuzzification. It is based on the distribution-based length and several length intervals are applied to identify the best forecasting results. [Huang et al. \(2011a\)](#) point out that an interval length influences the performance of forecasting performance and propose two methods which are based on average and the distribution, for defining the length. [Egrioglu et al. \(2010\)](#) apply the golden section search and parabolic interpolation based algorithm to identify the best interval for a fuzzification. In the optimization process, a MATLAB function called “fminbnd” which minimizes MSE is used. The function “fminbnd” is used to find minimum of a single-variable function on a fixed interval. The optimal interval provides increased accuracy of the forecast.

[Kuo et al. \(2009\)](#) apply particle swarm optimization approach to Chen’s interval forecasting model. Later, their work is extended by a novel hybrid forecasting model based on aggregated fuzzy time series, and particle swarm optimization is developed to adjust the length of each interval in the universe of discourse ([Huang et al. 2011b](#)). [Wang et al. \(2013\)](#) apply a fuzzy clustering to forming the subsets of given range for the fuzzification intervals. Their method is validated via Alabama University enrolment and Germany’s DAX stock data. There are other fuzzification methods that are also based on fuzzy clustering in which no interval is used and instead the data is fuzzified to the cluster centres ([Bulut et al. 2012](#); [Chen and Tanuwijaya 2011](#)). However, if concept values have outliers during long period and breach the assumption, the fuzzy values can be skewed

A drawback of the methods is the lack of consideration of determining a reasonable universe of discourse and the length of intervals ([Chen et al. 2014](#)). Simple fuzzification scheme with equal interval, which are adopted in most FCM studies, cannot cope with the data sets. For this reason, this study innovatively introduces the data normalisation-based fuzzification method that can cope with the problem due to outliers that cause the skewness in equal-length interval fuzzification.

### 3 Research methodology: big data-driven decision-making using FCM approach

The first step of the research methodology is starting from obtaining relevant data from big data sources to get fuzzy values and weight calculation for the FCM. After building an



**Fig. 2** Research framework: data utilisation for decision-making and impact evaluation using FCM

FCM with calculated fuzzy values and weights, we can conduct a decision impact analysis by comparing simulation result with different initial values of important decision variables. Figure 2 shows the framework for big data utilisation using FCM.

### 3.1 Make linkages between data and FCM: data fuzzification

As presented in Fig. 1, an FCM contains concepts and relationships, which describe interactions among concepts in a system. In most FCM studies, the values of concepts are usually fuzzified by mapping linguistic measures to fuzzified values in the interval  $[0, 1]$  based on the knowledge of domain experts. The weight values of relationships can also be defined by human experts. Focused group interview (Özesmi and Özesmi 2004) and group discussion (Jetter and Schweinfart 2011) are most common methods to assign initial fuzzy values to FCMs. However, the higher human intervention in FCM modelling is very time consuming and not efficient to cope with dynamic decision-making situation in reality. More importantly, weight matrices developed based on subject opinion of human experts may well reflect the reality and far from recent demand on data-driven decision-making (Nishisato and Ahn 1995). In this regard, open data can be used to complement the subjective opinion of human experts on fuzzy values of concepts and weight matrices.

Fuzzification function for historical data proposed in the literature is relatively simple. Most of fuzzification methods are based on simple categorical scheme. Let  $V = \{v_1, v_2, \dots, v_n\}$  be the set of real valued variables that are observed in time series. Let  $C = \{c_1, c_2, \dots, c_n\}$  be a superset of fuzzy sets  $c_i$ , where  $n = \text{cardinality}(C)$ . At time step  $t \in [0, 1, 2, \dots, t_e]$ ,  $t_e \in \chi$  is constant parameter that limits the considered time period. For example, if the concept is the historical data observed during 10 years,  $\chi$  is 10. Every value of  $v_1(t)$  is mapped by the fuzzification function  $\mu_i$  to a fuzzy value in set  $c_i$ , which means  $c_i(t) = \mu(v_i(t))$ . According to the fuzzy sets theory, the construction of fuzzification functions  $\mu_i$  is a complex task and usually done by domain experts (Lee 1990). However, in most practical cases, fuzzification function  $\mu_i$  is constructed by assuming a simple linear normalization:

$$c_i(t) = \frac{v_1(t) - \min(v_i)}{\max(v_i) - \min(v_i)}$$

Based on the equation, vector  $C(t)$  is constructed to describe the state of data at time  $t$ .

Even though the simple fuzzification function is easy to implement and intuitive, the fuzzification results can be far from reality. If there is a significant change (for example outliers) on data during the observed period, the fuzzified results can be skewed into lower

**Table 1** Sample time series data

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Value	139	167	156	231	213	224	245	204	211	225	237	218	234

or higher scales. For more sophisticated and realistic fuzzification, we apply the *time series normalisation-based fuzzification* scheme as shown below.

Let be  $Y = \{y_1, y_2, y_3, \dots, y_n\}$ ,  $y$  is subset of  $R^n$ , which is time-series vector and we can make another vector by sorting the series in ascending order, i.e.

$$y_k > y_l > y_m \cdots > y_z$$

$Y_{\text{sorted}} = \{y_k, y_l, y_m, \dots, y_z\}$  with one to one correspondence between components in vector  $f: R^n \rightarrow R^n$ .

Then the sequence of  $y_{\text{sorted}}$  can be considered as monotonic series, which is non-stationary with trend. In this case, the differencing series of  $y_{\text{sorted}}$ ,  $y'_{\text{sorted}}$  is stationary random process.

and its differencing series  $y'_{\text{sorted}}(t) = \{y'_1, y'_2, y'_3, \dots, y'_n\}$  can be considered as stationary process and can be written as

$$y'_1 = y_l - y_k, y'_2 = y_m - y_l, \dots \text{ and } Y'_{\text{sorted}}(t) \sim e^t$$

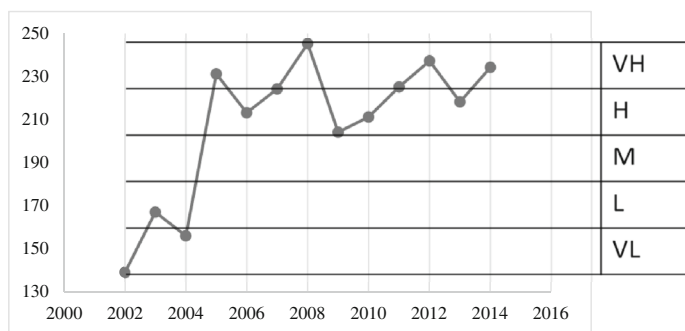
based on an assumption that the differencing following the normal white noise distribution.

Based on this, we can select outliers that are abnormally larger or smaller than other data using the statistics. There are two options for identifying outliers; first option is to identify outliers based on the normality assumption (Elliott and Stettler 2007; Dang et al. 2009) and second one is conducting Grubbs' test (Grubbs 1950), which identify only one outlier. The outlier can be replaced with mean of differencing series and we can reconstruct adjusted  $y_{\text{sorted}}$ ,  $Y_{\text{adjusted}}$ . Then the original series  $Y$  can be replaced with  $Y_{\text{adjusted}}$  using  $f^{-1}: R^n \rightarrow R^n$ . The overall fuzzification process can be summarized as follows.

- 1 st Step Building stationary process using the differencing of sorted time series data.
- 2 nd Step Selecting outliers in differencing and replacing them with the differencing mean.
- 3 rd Step Imputing the original time series and applying simple fuzzification scheme

Following example shows how the process is implemented step by step. Table 1 shows the sample time series data that is skewed. Without any modification, the simple fuzzification scheme can be applied as shown below in Fig. 3. Due to the skewness of data values, any data in time series cannot be assigned to “Middle” in 5-scale fuzzification scheme (VH: very high, H: high, M: middle, L: low, VL: very low). Due to the sharp rise between 2004 and 2005, values after 2005 are assigned to high and very high categories. To realise more realistic fuzzification, data normalisation is necessary. The basic idea of imputation in this study is to normalise the value differences by detecting and replacing large values that can be considered as outliers with mean value.

The first step of data imputation is to generate stationary process for outlier detection. The differencing of trend data can be seen as stationary process. To generate the process, we can sort the data in ascending order to make the time series has monotonic trend. The 3rd column of Table 2 presents the difference and its mean along with standard deviation. The second step is identifying outliers among the difference values under the assumption of stationary process. The first option for identifying outliers is detecting the difference values



**Fig. 3** A simple fuzzification scheme example

which are larger than mean + 2sigma (i.e., 28.59) and consider those values as outliers. The values bigger than 28.59 belongs to the top zone which is higher than 97.72 % from the top. Based on this criterion, we can conclude that the difference value “37” is an outlier and we can replace it with the mean of difference, 8.833333.

The second option is to test if “37” is an outlier using a statistical test, i.e. *Grubbs’ test* for identifying an outlier. By applying a two side Grubb’s test, both minimum and maximum value can be tested whether they are outliers or not.

To test whether the minimum value is an outlier, the test statistic is

$$G = \frac{\bar{y} - y_{\min}}{s}$$

with  $y_{\min}$  denoting the minimum value. To test whether the maximum value is an outlier, the test statistic is

$$G = \frac{y_{\max} - \bar{y}}{s}$$

with  $y_{\max}$  denoting the maximum value.

The critical value can be defined as

$$G_{crit} = \frac{(n-1)t_{crit}}{\sqrt{n(n-2+t_{crit}^2)}}$$

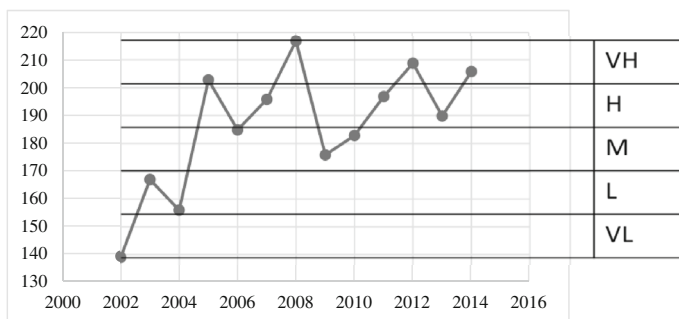
where  $t_{crit}$  is the critical value of the t distribution  $T(n-2)$  and the significance level is  $\alpha/n$ . Thus the null hypothesis is rejected if  $G > G_{crit}$ . In this case, we need to test only for maximum outlier so we can calculate the g statistics using the formula. For the maximum value “37”, the g value is 2.848 and the critical value Gcrit is 2.285 with 5 % level. So we can reject the null hypothesis and conclude that “37” is outlier. Based on the second option, we can also find that “37” is an outlier and it can be replaced with mean difference, 8.8333.

Finally, we can impute the original value using normalised difference as presented in the 5th column.

Then we can apply the simple equal width fuzzification scheme for imputed data and revised fuzzification results can be obtain as shown in Fig. 4.

**Table 2** Adjusting time series data based on detecting outlier in its differencing

Year	Value	Difference	Normalised difference	Adjusted value
2002	139			139
2004	156	17	17	156
2003	167	11	11	167
2009	204	37	8.833333333	175.8333333
2010	211	7	7	182.8333333
2006	213	2	2	184.8333333
2013	218	5	5	189.8333333
2007	224	6	6	195.8333333
2011	225	1	1	196.8333333
2005	231	6	6	202.8333333
2014	234	3	3	205.8333333
2012	237	3	3	208.8333333
2008	245	8	8	216.8333333
	Mean	8.833333		
	SD	9.888775		

**Fig. 4** Fuzzification result using proposed scheme

### 3.2 FCM weight calculation using data

As many strategic variables are inter-related each other even for single decision-making, decision makers need to consider various data regarding those variables as much as they can. As we stated earlier, many weight calculation methods have been proposed but most of them has critical drawback in terms of their scalability as their basis algorithms are coming from Hebbian approach. To cope with more complex model with sheer volume of data, we obtain its scalability by partitioning the FCM into single layer perceptron problem and applying simple optimisation instead of learning and test approach.

Let  $X_t$  be the status-vector whose elements denote the concept values of an FCM at time  $t$  and  $W$  be the weight matrix of the FCM. Then the general calculation method of FCM can be expressed in vector form as follow:

$$X_t = f(X_{t-1} \cdot W + X_{t-1})$$

For the estimation of the weight matrix using historical data, we can consider the object function

$$\min E_{total} = (X_t - f(X_{t-1} \cdot W + X_{t-1}))^2$$

that minimises the sum of squared errors from each concept presenting the difference between real historical value and calculated value. This estimation process is performed using a back propagation technique which is used for neural network as the FCM can be seen as a combination of single layer artificial neural network if we decompose the FCM into multiple single neural networks.

To decrease the total error of each partitioned FCMs, we can adjust weights among concepts using the formula

$$w_i^+ = w_i^{initial} - \eta * \frac{\partial E_{total}}{\partial w_i^{initial}}$$

where  $\eta$  is learning rate and the derivative of  $E_{total}$  error with respect to  $w_i^{initial}$  can be decomposed using chain rule:

$$\frac{\partial E_{total}}{\partial w_i^{initial}} = \frac{\partial E_{total}}{\partial x_i} * \frac{\partial x_i}{\partial E_{x_i}} * \frac{\partial E_{x_i}}{\partial w_i^{initial}}$$

where  $x_i$  denotes the estimated output target concept of weight and  $E_{x_i}$  denotes the error generated from the concept.

After repeating this estimation process, we can get a final weight matrix for each partitioned FCM using historical data. For the initial values for the weight matrix, we will use the random number between 0 and 1.

### 3.3 Impact analysis method: network and complexity analysis perspective

The impact of value changes in a group of concepts is measured by comparing the simulation results with historical time series data. The iterative simulation with regards to different initial values for the concepts provides decision makers with meaningful insights. However, the impact evaluation becomes complicated as the number of concepts and their relationships in an FCM increases and becomes a complex task (Stylios and Groumpos 2004). In this case, it is difficult for a decision maker identify concepts in the FCM that may have a significant impact on the other concepts. If the decision maker considers only decision variables and target decision variables, s/he may not cope with side effects from other relevant concepts. For a comprehensive and efficient analysis of the impact of different decision, decision makers need to take into account major mediating concepts in a decision model as well. More sophisticated evaluation method is needed.

The impact due to the change of concept values can be determined by two factor; the amount of changes in concepts and their topological positions in an FCMs (a network). The first factor, the amount of a value change, is important for activation function. As most of activation functions have positive slopes, the output usually increases as the input values increase and vice versa. The second factor, topological position, needs speculation from network and system perspectives as FCMs are bidirectional networks (Khan and Quaddus 2004). Each concept in an FCM has different influence on the whole system according to its network property such as centrality. In this study, we adopt the concept of centrality for more sophisticated impact analysis. Bonacich (1972) suggests that the eigenvector of the

largest eigenvalue of an adjacency matrix can be used as a good network centrality measure. Unlike degree centrality, which weighs links to other nodes equally, the eigenvector weighs links according to linked nodes' centralities. The eigenvector centrality can also be seen as a weighted sum of not only direct connections but also indirect connections mediated by other concepts. Thus, it takes into account the entire patterns into account in the network (Bonacich 2007) and it is more applicable to FCMs as we are interested in the whole impacts of a change on a concept (node). Eigenvector centrality is defined as the principal eigenvector of an adjacency matrix representing a network. Equation (3) describes eigenvector centrality  $x$  in two equivalent ways, as a matrix equation and as a sum. The centrality of a node is proportional to the sum of the centralities of the nodes to which it is connected (Bonacich 2007).

$$\lambda x = Ax, \lambda x_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, n \quad (3)$$

where  $A$  is the adjacency matrix of the graph,  $\lambda$  is the largest eigenvalue,  $n$  is the number of nodes, and  $x$  is the eigenvector. Using the formula above, we can calculate the Bonacich centrality for each concept in an FCM, and concepts with higher centralities can be considered to be more influential to other concepts. The centrality information enables user to recognise which concepts need to be adjusted to obtain expected impacts more easily.

#### 4 Experiments: ABC local government IT service procurement decision case

The result of FCM learning and its simulation can be evaluated based on how well the calculated weight matrix and simulation result reflects on the real phenomena, as the aim of the FCM is to model the social phenomena and problem. The evaluation of FCM derived based on data-driven learning methods is not addressed in the literature, however, we can refer the evaluation approach from a knowledge representation and modelling literature. An FCM as a knowledge representation model can be evaluated in four level; golden standard, application-based, data-driven, and assessment by human (Brank et al. 2005). Golden standard based evaluation usually can be applied for the situation with absolute truth or permanent domain knowledge. For example, a conceptual model describing the relationship between school and students as “have/belong to” can be evaluated based on golden standard such as domain knowledge representation or definition of each concepts. Application based evaluation can be done by showing the applicability of a model to a given situation. Data-driven evaluation uses the data related to the model to check the fitness of model to data. The best evaluation option is an assessment by human (i.e., experts) as it can cover the other three approaches. But this evaluation approach needs proper experts who have enough domain knowledge on the phenomena that the model describes. For this reason, only few studies on knowledge representation and modelling adopt the human expert-based evaluation. In the context of an FCM modelling, golden standard based evaluation cannot be applied as FCMs do not have golden standard<sup>1</sup> such as knowledge source implemented based on formal language. For the reason, most of FCM learning methods try to evaluate their learning methods rather than the effectiveness of FCMs as a knowledge representation method. Data-driven approach

<sup>1</sup> For example, the conceptual model describing the relationship between school and students as “have/belong to” can be evaluated based on golden standard such as domain knowledge representation or definition of each concepts.

also has limitations in verifying FCM's structure as data can be used only to validate the correlation among concepts rather than confirm the direction of causal relationship. In this regard, the assessment by human expert can be the best evaluation method if proper experts can be invited for evaluation. Human experts can evaluate an FCM weight matrix based on their domain knowledge. In addition, they can confirm the applicability of FCM approach to decision-making and assess how the FCM simulation and impact analysis results are valuable for practical decision-making. To authors' best knowledge, this study is one of the first studies that evaluate the FCM learning and impact analysis results based on domain experts' opinion.

In this study, the evaluation for FCM with calculated fuzzy values and weigh matrix is done by human experts in ABC (we are anonymising the real name of the authority) public authority. The evaluation criteria are to verify the suitability of fuzzy value assigned to each concept and weight matrix values. In this evaluation scenario, twelve concepts' fuzzy values and nineteen weight values have been estimated based on concepts' historical data and then validated by comparing with the values derived from the experts group. In addition to this quantitative evaluation, we evaluate the strategic implications derived through an impact analysis based on the FCM simulation by asking experts from ABC regional public authority for their opinion on the quality of the implications through an in-depth interview.

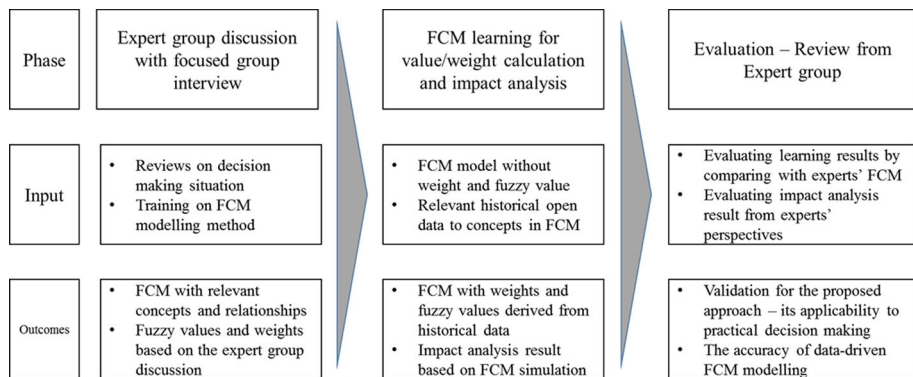
#### 4.1 Prioritising IT services procurement

ABC local government is managing the regional program "Development of the Information Society in A region in 2014–2018" which is the successor of a series of federal and regional programs devoted to the promotion of e-governance in various fields of public administration in the region in 2002–2013. To develop the information society, the supply of IT service procurements such as IT infra and IT education that enables citizens to utilise the IT for access to various social service (Hong and Huang 2005).

Public service procurement, as a part of public service supply chain of local government, has been recognised as an important body for the program. Decision-making on IT service procurement should be implemented based on the prioritised possible IT service procurement options that can realise the goals of the program due to the limited budget for the supply of public IT service. However, the decision makers in ABC local government have some major problems and difficulties in prioritising their possible public IT service procurement options, as described below.

- Difficulties in evaluating the impact of each IT service procurement execution due to the multidimensional effect of IT service to social and economic variables.
- Lack of analytical support for decision-making that is mainly due to the absence of analytical tools
- Specific traditional management style, which requires certain actions, but does not require any practical results or impact.

There are four possible IT service procurement options that can be supplied to the citizen of ABC region; e-workflow introduction, increasing the e-service provision, supporting the broadband penetration, and opening the programme that helps citizens' e-skills. ABC local government wants to analyse how each of IT service procurement options can make an impact to the development of information society under the e-government context.



**Fig. 5** Overall evaluation process

## 4.2 Evaluation process

We firstly provided ABC public authorities with a training session to allow them be familiarised with FCM concepts in collaboration with the e-Governance Centre of the ITMO University. The overall evaluation process is shown below in Fig. 5.

We invited six decision makers who are in charge of IT service procurement decision in ABC local government for a focused group discussion. Firstly, based on the focused group discussion and the tutorial session on FCM, the group generated an FCM model for IT funding decision of ABC region. The expert group assigned fuzzy values that can be used for validation of the data driven-FCM learning method to concepts and causal relationships in the FCM model. Secondly, developed FCM model was used for the input to the FCM learning phase. In FCM learning phase, no value from the expert group were used for an FCM learning, which only depends on historical data for fuzzy values of the concepts in the FCM model.

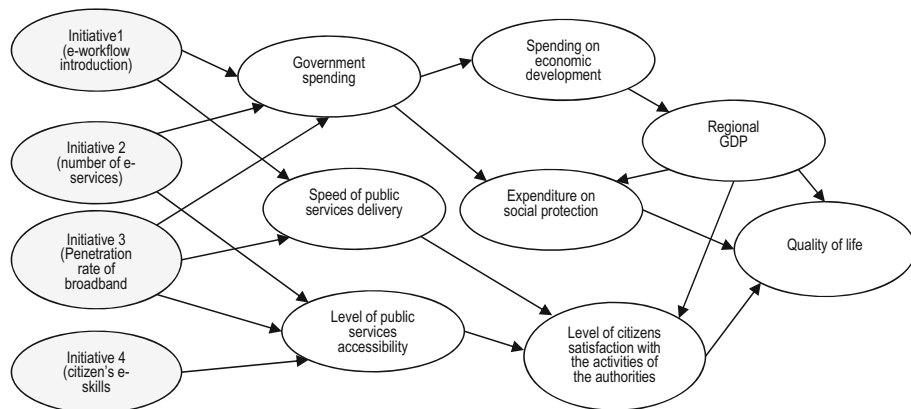
The outcomes of the FCM learning were estimated weights and fuzzy values for the concepts. Lastly, the calculated fuzzy values based on the historical data for all concepts are evaluated by comparing with experts' opinion which is collected by voting approach. The simulation results for the impact analysis of different decision using the estimation were evaluated and reviewed by the experts group again to verify the proposed methods and the simulation results.

Figure 6 shows the initial FCM that was developed based on four strategic decision variables as well as eight other relevant variables that present the causal-effect relationships and the social impact of funding decisions. The relevant data that represent each concept was provided by ABC local government and also collected from the open data portal managed by Russia federation and European Commission. The concepts in FCM and their data sources are provided in Table 3. For the FCM simulation, all the data have been imported from open big data portal managed by Russia local government<sup>2</sup> and European commission<sup>3</sup> to the FCM simulation platforms. This platform,<sup>4</sup> as a big data platform, has played a role as a big data control tower which can facilitate and control the various public open data sources so

<sup>2</sup> Russia federation statistical systems (<http://fedstat.ru/>), Russia federation open data porta (<http://mvd.ru/opendata/>), ABC local government open data porta.

<sup>3</sup> European Open Data Portal (<https://open-data.europa.eu/>).

<sup>4</sup> PolicyCompass (<http://policycompass.eu>).



**Fig. 6** Initial FCM for evaluating the impact of different IT service procurement options

that they can easily accessible and usable for supporting various decision-makings in public sector. As we discussed in introduction section, this type of big data portal as a data control tower and integration tool among different silos is necessary for utilisation of disperse big open data.

### 4.3 Estimation of fuzzy values and impact analysis with simulation results

Proposed learning algorithm needs first initial weight matrix that will be adjusted through the iterative error minimization sequence. We obtained the initial weight matrix using the correlation matrix among concepts using historical data. To calculate the error, the most recent values of the historical data will be the output values and values just before recent will be used for input values for estimation. The revised FCM weights and its simulation parameter (i.e., weights and fuzzy values for concepts) based on dataset provided by ABC government are shown in Fig. 7 and Table 4. We used 5 scale values between 0.2 (very low) and 1.0 (very high) for concept value fuzzification and 8 scale values between  $-1.0$  (strong negative) and 1.0 (strong positive) for weight matrix values. The estimated fuzzy values for concepts are mostly consistent with the expert's opinion except the value for "Penetration rate of broadband" as it is at all time high but experts thought it was middle level comparing to the other developed countries.

The calculated fuzzy values for "Government spending" and "Regional GDP" are slightly overestimate comparing to the experts' opinion (very high vs. high) as two concepts reach their peak currently. But some experts think the regional economy has room to grow the GDP and spending so 4 experts assigned "high" while 2 experts did "very high" for these two concepts.

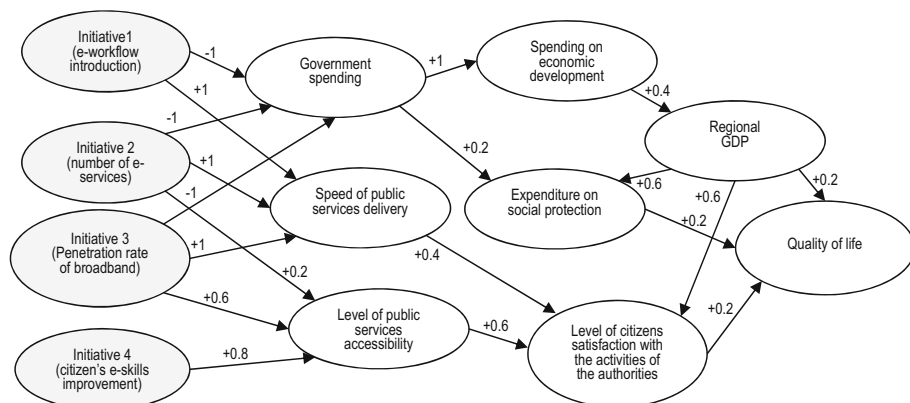
The estimated fuzzy weights were also validated by comparing the values with those based on experts opinion. All calculated 19 weights among concepts have no difference from experts' opinions in terms of their sign and no big differences are found for value (less than 1 scale difference in most cases). Using the estimated FCM, we conduct a base simulation that shows the steady state concept values when current fuzzy concept values interact with each other based on the weight matrix values. Figure 8 shows the steady state concept values.

**Table 3** Dataset for FCM and relevant data sources

Concepts	Relevant document	Data sources
Quality of life	Information Society in the ABC region 2014–2018 Regional Economic Strategy Regional Administrative Reform Strategy Department for Informatisation and Communication Policy	Administrative reform portal: <a href="http://ar.gov.ru/">http://ar.gov.ru/</a> United interagency information-statistical system: <a href="http://fedstat.ru/indicators/start.do">http://fedstat.ru/indicators/start.do</a>
Level of citizens satisfaction with the activities of the authorities	Regional Social Strategy Departmental Quality Management Policies	Regular social surveys DB (2 years only: 2013–2014)
Government spending	Departmental Saving Resources and Efficiency Improvement Policies	Regional Budget (as a separate document not machine readable, only this year budget electronic but previous one are hard copies)
Regional GDP	Regional Economic Strategy Industrial development strategy Agricultural Development Strategy Entrepreneurship Development Strategy Federal and Regional Strategy for IT Sector Development	Departmental Performance Reports DB Regional Budget Reports of the responsible committees United interagency information-statistical system: <a href="http://fedstat.ru/indicators/start.do">http://fedstat.ru/indicators/start.do</a>
Spending on economic development	Regional Economic Strategy Industrial Development strategy Agricultural Development Strategy Entrepreneurship Development Strategy Federal and Regional Strategy for IT Sector Development	Regional Budget Reports of the responsible committees

**Table 3** continued

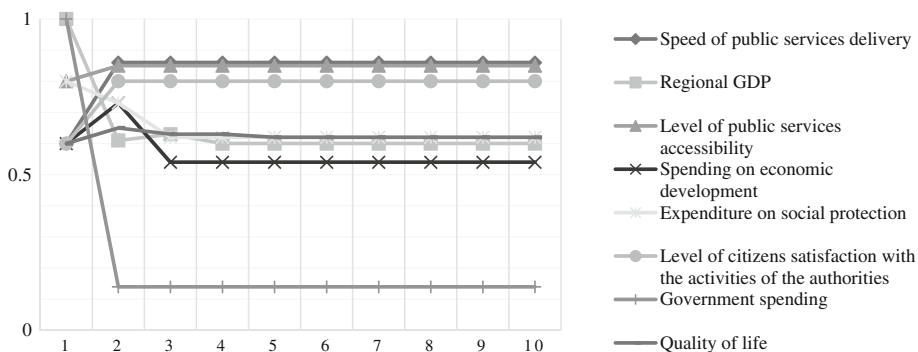
Concepts	Relevant document	Data sources
Expenditure on social protection	Regional Economic Strategy Regional social policy Committee on Social Policy Activity Plan Committee on Healthcare Activity Plan	Regional Budget Reports of the responsible committees United interagency information-statistical system: <a href="http://fedstat.ru/indicators/start.do">http://fedstat.ru/indicators/start.do</a> Reports of the responsible committees Regular social surveys DB
Speed of public services delivery	Information Society in the ABC region 2014–2018 Regional Administrative Reform Strategy Department for Informatisation and Communication Policy	Reports of the responsible committees Regular social surveys DB (might be a problem on integrity)
Level of public services accessibility	Information Society in the ABC region 2014–2018 Regional Administrative Reform Strategy Department for Informatisation and Communication Policy	Reports of the responsible committees Regular social surveys DB (might be a problem on integrity)
Historical Information	Adoption of new regulations  Planned events (elections, publication of significant reports, opening of large enterprises etc.)  Unplanned significant events (recession, administrative changes, closing of large enterprises, actions of foreign states etc.)	United interagency information-statistical system: <a href="http://fedstat.ru/indicators/start.do">http://fedstat.ru/indicators/start.do</a> Information on the activities of the Ministry of Internal Affairs of the Russian Federation, which is located in the network "Internet" in the form of open data: <a href="http://mvd.ru/opendata/">http://mvd.ru/opendata/</a> Administrative reform Portal: <a href="http://ar.gov.ru/">http://ar.gov.ru/</a> Federal and regional regulatory databases Results of the analysis of the reaction to the events in social networks



**Fig. 7** FCM with calculated weights from historical data

**Table 4** Fuzzy values for FCM—comparison between experts opinion and historical data based values

	Current fuzzy value from experts	Calculated current fuzzy value from historical data
Quality of life	High	High
Level of citizens' satisfaction	Middle	Middle
Government spending	High	Very high
Regional GDP	High	Very high
Spending on economic development	Middle	Middle
Expenditure on social protection	High	High
Speed of public services delivery	Middle	Middle
Level of public services accessibility	High	High
e-workflow introduction	Very low	Very low
Number of e-services	Middle	Middle
Penetration rate of broadband	<b>Middle</b>	<b>Very high</b>
Citizen's e-skills	High	High



**Fig. 8** FCM simulation results

**Table 5** Bonacici centrality of concepts in FCM

Concepts	Centrality
Speed of public services delivery	2.122
Government spending	2.122
Level of public services accessibility	1.132
Level of citizens satisfaction with the activities of the authorities	1.013
Expenditure on social protection	0.518
Quality of life	0.498
Regional GDP	0.306
Spending on economic development	0.283

The simulation results indicate that government spending is expected to decrease while public service accessibility and satisfaction level expected to be improved in time. But this simulation result only shows the future values of each concept if all the concepts keep the current trend so more in-depth analysis is needed for the strategic decision regarding the concept. For this reason, we need to conduct an impact analysis that shows the impact of decision variables to other important concepts by changing the initial value of decision variables.

Before conducting the impact analysis, we assess which concept is most influential to the whole FCM by calculating Bonacici centrality to prioritise the impacts of 4 decision variables. This step is very significant and useful to identify important variables apart from the 4 decision variables. According to Eq. (3) in Sect. 3.3, Table 5 shows the Bonacici centrality values of the concepts. The result of centrality analysis presents some major concepts (i.e., Speed of public service delivery, Government spending, and Level of public services accessibility), which can be considered as major influential factors for the decision-making problem. The results provide decision makers with a reference to prioritise different decisions options as explained in the next step.

Impact analysis seeks to scrutinise those decisions that can make significant impact to derive more desirable outcomes in the future. In this problem domain, we have four decision variables regarding the funding decision and we need to find out where the ABC government should allocate limited funding according to the priority. For this work, we investigate the impact of a decision variable by changing its fuzzy value while locking the other decision variables and repeat the simulation for other decision variables (or combination of more than one decision variable). The results of the impact analysis are shown in Table 6.

As we can see from Table 6, four decision variable can have different impact on the important concepts such as “Level of public services accessibility”, “Speed of public services delivery” and “Government spending.” The speed of public services delivery can be improved by enhancing e-workflow introduction, number of e-services, or penetration rate of broadband while citizen’s e-skills improvement does not have any impact on public service speed. Accessibility level of public service can be improved by increasing penetration level of broadband or citizen’s e-skills. Government spending can be also affected by three

decision variables except citizen's e-skill improvement and increasing these three variables have positive impact on decreasing government spending.

#### 4.4 Action plan development for ABC government with the interpretation of impact analysis results

The final stage of the scenario-based evaluation is to derive action plans from the impact analysis results. The derived action plan for ABC funding decision is evaluated by experts who initially provided the FCM model for the decision scenario.

Table 6 shows the impact of each decision concepts by changing initial value of specific concept while the other concepts values are controlled. Four different types of funding decision for IT service supply can be prioritised according to their impacts on important concepts such as speed of public service delivery, government funding, and service accessibility, which have high centrality in the FCM.

The impact analysis results imply that opening IT class for citizens' skill improvement is most efficient way to improve the public service accessibility while it only has marginal impact on the other concepts. If the citizen's e-skill improvement increases to 1.0, the accessibility level of public service will be 0.77, which is highest value among other options. Providing infrastructures and e-service is positive impact on improving the level of government spending and public service accessibility. This finding shows that all decision concepts for information society building has different impact according to the specific target variables. The decision maker from ABC regional government also agreed with the result of impact analysis and they decided to consider to open more classes for citizens' IT skill improvement and elevation of broadband penetration rate in ABC region as they agree with the importance of service accessibility of citizens and saving government spending simultaneously. Throughout the overall evaluation process, proposed method successfully provides the comprehensive perspectives on the decision-making situation based on the data analysis. The suggested impact analysis results are consistent with the experts' opinion.

## 5 Conclusions

Modern day political rhetoric and commitment has created an expectation on making data open and accessible to the public, together with clarity around *how* decision-making takes place. However, there is disconnection between availability and the ability to utilise these data, though it can be fully utilised for procurement decision including various SCM decision situation. A reason for this mismatch is a lack of robust and yet simple analytical methods with which the public can use with modest levels of skills. In response to this void, this study proposed a novel approach to decision modelling and impact analysis, which is applicable to the decision-making for SCM regarding IT service in particular.

The authors of this paper have developed an innovative framework for decision modelling and impact analysis that based upon open big data and when applied to IT service procurement in the public sector. Underpinning this approach is a Fuzzy Cognitive Map, which is particularly suited to modelling complex and dynamic social problems. This research has sought to exploit BDA to enhance decision-making thus *developing decision models that capture the strengths of both data analytics and the established intuitive qualitative approach*. The approach was verified through an application to the evaluation of decision-making on IT service supply for information society building and makes a meaningful contribution to the

**Table 6** Impact analysis table—the result of simulation with regard to the each decision variable's initial value change

	Initial value of Initiative 1 (e-workflow introduction)					Initial value of Initiative 2 (number of e-services)				
	0.2	0.4	0.6	0.8	1	0.2	0.4	0.6	0.8	1
<b>Speed of public services delivery</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>	<b>0.8</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>	<b>0.8</b>
Regional GDP	0.61	0.61	0.6	0.6	0.6	0.61	0.61	0.6	0.6	0.6
<b>Level of public services accessibility</b>	<b>0.6</b>	<b>0.6</b>	<b>0.6</b>	<b>0.6</b>	<b>0.6</b>	<b>0.6</b>	<b>0.61</b>	<b>0.62</b>	<b>0.63</b>	<b>0.65</b>
Spending on economic development	0.59	0.58	0.57	0.56	0.55	0.59	0.58	0.57	0.56	0.55
Expenditure on social protection	0.63	0.63	0.63	0.62	0.62	0.63	0.63	0.63	0.62	0.62
Level of citizens satisfaction	0.75	0.75	0.75	0.76	0.76	0.75	0.75	0.76	0.76	0.77
<b>Government spending</b>	<b>0.35</b>	<b>0.31</b>	<b>0.27</b>	<b>0.23</b>	<b>0.2</b>	<b>0.35</b>	<b>0.31</b>	<b>0.27</b>	<b>0.23</b>	<b>0.2</b>
Quality of life	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
<b>Initial value of Initiative 3 (penetration rate of broadband)</b>										
	0.2	0.4	0.6	0.8	1	0.2	0.4	0.6	0.8	1
<b>Speed of public services delivery</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>	<b>0.8</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>
Regional GDP	0.61	0.61	0.6	0.6	0.6	0.61	0.61	0.61	0.61	0.61
<b>Level of public services accessibility</b>	<b>0.6</b>	<b>0.63</b>	<b>0.67</b>	<b>0.7</b>	<b>0.73</b>	<b>0.6</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>
Spending on economic development	0.59	0.58	0.57	0.56	0.55	0.59	0.59	0.59	0.59	0.59
Expenditure on social protection	0.63	0.63	0.63	0.62	0.62	0.63	0.63	0.63	0.63	0.63
Level of citizens satisfaction	0.75	0.75	0.76	0.77	0.78	0.75	0.75	0.76	0.76	0.77
<b>Government spending</b>	<b>0.35</b>	<b>0.31</b>	<b>0.27</b>	<b>0.23</b>	<b>0.2</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>
Quality of life	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
<b>Initial value of Initiative 4 (citizen's e-skills improvement)</b>										
	0.2	0.4	0.6	0.8	1	0.2	0.4	0.6	0.8	1
<b>Speed of public services delivery</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>	<b>0.8</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>
Regional GDP	0.61	0.61	0.6	0.6	0.6	0.61	0.61	0.61	0.61	0.61
<b>Level of public services accessibility</b>	<b>0.6</b>	<b>0.63</b>	<b>0.67</b>	<b>0.7</b>	<b>0.73</b>	<b>0.6</b>	<b>0.65</b>	<b>0.69</b>	<b>0.73</b>	<b>0.77</b>
Spending on economic development	0.59	0.58	0.57	0.56	0.55	0.59	0.59	0.59	0.59	0.59
Expenditure on social protection	0.63	0.63	0.63	0.62	0.62	0.63	0.63	0.63	0.63	0.63
Level of citizens satisfaction	0.75	0.75	0.76	0.77	0.78	0.75	0.75	0.76	0.76	0.77
<b>Government spending</b>	<b>0.35</b>	<b>0.31</b>	<b>0.27</b>	<b>0.23</b>	<b>0.2</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>
Quality of life	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62

**Table 7** Existing research on public procurement decision-making

Research on procurement decision-making	Strategy for prioritising alternative decision	Applicability to Other decision-making for SCM	Problem domain and findings
Preuss (2009)	Guidelines for public service procurement	Yes—guidelines of decision-making for sustainable public procurement	Exploratory approach based on literature reviews for procurement decision of general public sector
Amann et al. (2014)	N/A	No—focus on public procurement	Identifying the importance of social responsible goal for public procurement in EU member states based on structural equation modelling
Grudinski et al. (2014)	N/A	No	Empirical model based on the survey to minimise the risk for procurement decision for health care domain
This research	FCM approach by utilising big data analytics	Yes—applicable to various decision-making situation with data	Analytics that enables prioritising various public IT service procurement options by evaluating different options with data

normative literature through tackling an academic and practical challenge. Specifically, the contributions claimed in this paper are as follows.

Firstly, this paper provides analytics for decision-making regarding IT service procurement in public sector based on the connection between FCM approach and BDA. There have been a few attention on the importance of procurement decision and prioritising public service procurement options, however, few tackled the utilisation of big open data for the decision modelling and impact analysis. Most of studies on decision-making for public service procurement focused on finding behavioural factors and guidelines for decision based on the exploratory and conceptual research rather than providing practical and actionable analytics (see Table 7). By providing the evaluation of each possible IT service procurement options with their impacts, the proposed approach can provide the decision makers on IT service procurement with the insight for selecting efficient and effective IT service procurement.

Secondly, this paper also articulates an approach that integrates an easy to use decision model with BDA. Though FCM is widely used to model social and political decision problems, no study tackles integration of big open data and FCMs for data and model driven decision modelling and impact analysis. Through combining an FCM and data-driven approach, the proposed approach enables the decision makers in public sector not only to easily understand and interpret decision models and analytic results but also to utilise big data to strengthen the causal-effect relationships and fuzzy concept values of FCMs for a strong evidence of decision-making.

Thirdly, existing FCM learning approaches using historical data have limitations for big data based decision-making and evaluation due to a scalability issue. This makes the

**Table 8** The summary of existing FCM learning research

FCM research using data	Method for weight calculation	Method for value fuzzification
Kosko (1994)	Differential Hebbian learning	N/A
Papageorgiou et al. (2004)	Active Hebbian learning	Simple fuzzy value mapping with normal fuzzification scheme
Koulouriotis et al. (2001)	Genetic strategy	N/A
Stach et al. (2005)	Real coded genetic algorithm	N/A
Ghazanfari et al. (2007)	Simulated annealing with genetic algorithm	N/A
Papageorgiou et al. (2011)	Decision tree	Simple fuzzy value mapping with normal fuzzification scheme
This research	Partitioning based optimisation using single layer perceptron approach	Time-series normalisation-based fuzzification with outliers detection

approaches difficult to be applied to BDA. Moreover, the approaches struggle with data sets that have drastic changes in trends and breach normal distribution assumption for fuzzification. As demonstrated in Table 8, existing FCM studies using historical time series data use traditional learning approaches such as Hebbian learning, genetic algorithm, and decision tree that have limitation in scalability.

Thirdly, existing fuzzification methods have limitation in coping with the drastic changes to derive accurate fuzzy values from long historical data. The proposed method in this paper is scalable for fuzzification of concepts and weight matrix estimation for data sets with drastic changes. To obtain the scalability for weight calculation method, we adopted the partition-based optimisation approach using the concept of single layer approach. The value fuzzification scheme in this study return more meaningful fuzzy values by using time-series normalisation-based fuzzification, which detects the outliers and reconstructs the time series data using random process. The estimated FCM based on the method was verified via an application to a real world decision-making problem. The impact analysis based on simulation results turned out to be useful for a group of experts who have been working on the domain problem for a long period. This is one of the first studies that empirically verify the usefulness of an FCM estimated based on fuzzification and weight learning algorithms by comparing with human expert opinions.

Finally, the proposed approach has significant academic contribution by proposing novel approach to integrating model-driven and data driven decision modelling approach. The simple and intuitive nature of FCM enhances the understanding of decision problem for decision maker and scalable weight calculation method enables an FCM to utilise the big data. While existing BDA and BI for decision support solely depend on data-driven decision modelling, this research show the novel approach to combine model-driven and data-driven decision modelling to combine all strengths from two approaches. By doing so, we extend the applicability of BDA to model-driven decision modelling and analysis and clarify how the big data can be used for quantitative approach as well as qualitative approach.

The practical implications of this study are apparent. The proposed approach enables decision makers implement data-driven decision-making based on open big data. As decision

makers of public IT service supply from ABC region for the evaluation pointed out, there have been difficulties in prioritising the various IT service supply due to the complex cause-effect relationships among large number of social and economic factors that need to be considered even though the list of factors can be found from the literature. Therefore, decision makers tend to use their subjective insights in making decisions on the priority of available decision options. The use of qualitative method like FCMs can help organising different factors and relate them to see impact analysis via a systematic simulation. However, the lack of linkage with real big data of such qualitative methods has a limitation in implementing data-driven and evidence-based decision-making. The proposed approach that integrates quantitative open data with a qualitative decision model (FCM) provides decision makers with new opportunity to realise data driven decision-making. This was supported by the testimony from the field trial participants who provided very positive feedbacks about the simulation results they obtained via impact analysis stage of the evaluation.

The proposed method can be used not only for validating and confirming the public authorities' decision, but also for the simulation to expect which different decision variable can make an impact and how they interact each other. Though we didn't use massive amount of data for the validation of proposed method though field trial in Russia regional government, it would be also applicable to analyse more complex decision situation with massive data considering the scalability and simplicity of proposed BDA.

Future studies can deal with the applicability of proposed method to the business decision-making. In addition to the decision-making for SCM context, the proposed approach can be very useful to design the future strategy of firm with big data. Also, various decision modelling and impact analysis in public sector can be covered using the proposed method. By capturing the strength of data analytics and qualitative approach, the research can be applied to complex decision-making problems that have relevant big data.

**Acknowledgements** This study was partially funded by EU FP7 project Policy Compass (Project No. 612133). The authors acknowledge the support received from the Policy Compass project members including Governance Centre of St. Petersburg National Research University who provided data and helped the evaluation of the proposed approach.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Amankwah-Amoah, J. (2015). Safety or no safety in numbers? Governments, big data and public policy formulation. *Industrial Management & Data Systems*, 115(9), 1596–1603. doi:[10.1108/IMDS-04-2015-0158](https://doi.org/10.1108/IMDS-04-2015-0158).
- Amann, M., Roehrich, J. K., Eßig, M., & Harland, C. (2014). Driving sustainable supply chain management in the public sector: The importance of public procurement in the European Union. *Supply Chain Management: An International Journal*, 19(3), 351–366.
- Amin, A. H. M., Khan, A. I., & Nasution, B. B. (2012). *Internet-scale pattern recognition: New techniques for voluminous data sets and data clouds* (Vol. 20). Boca Raton: CRC Press. <https://books.google.com/books?id=tTjegWHsIgC&pgis=1>. Accessed March 17, 2016.
- Andreou, A. S., Mateou, N. H., & Zombanakis, G. A. (2003). The Cyprus puzzle and the Greek-Turkish arms race: Forecasting developments using genetically evolved fuzzy cognitive maps. *Defence and Peace Economics*, 14(4), 293–310. doi:[10.1080/10242690302931](https://doi.org/10.1080/10242690302931).

- Axelrod, R. (1976). *Structure of decision: The cognitive maps of political elites*. Princeton: Princeton University Press. <https://books.google.com/books?hl=en&lr=&id=aKh9BgAAQBAJ&pgis=1>. Accessed April 1, 2016.
- Bhargava, H. K., Power, D. J., & Sun, D. (2007). Progress in Web-based decision support technologies. *Decision Support Systems*, 43(4), 1083–1095. doi:10.1016/j.dss.2005.07.002.
- Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1), 113–120. doi:10.1080/0022250X.1972.9989806.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564. doi:10.1016/j.socnet.2007.04.002.
- Bonczek, R. H., Holsapple, C. W., & Whinston, A. B. (2014). *Foundations of decision support systems*. Amsterdam: Elsevier Science. <https://books.google.com/books?hl=en&lr=&id=bzGjBQAAQBAJ&pgis=1>. Accessed March 29, 2016.
- Brank, J., Grobelnik, M., & Mladenic, D. (2005). A survey of ontology evaluation techniques. In *Proceedings of the conference on data mining and data warehouses* (pp. 166–170).
- Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of “big data”. *McKinsey Quarterly*, 4, 24–35.
- Bulut, E., Duru, O., & Yoshida, S. (2012). A fuzzy time series forecasting model for multi-variate forecasting analysis with fuzzy C-means clustering. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 6(3), 671–677.
- Carvalho, J. P., & Tome, J. A. B. (2002). Issues on the stability of fuzzy cognitive maps and rule-based fuzzy cognitive maps. In *2002 annual meeting of the North American Fuzzy Information Processing Society proceedings. NAFIPS-FLINT 2002* (Cat. No. 02TH8622) (pp. 105–110). IEEE. doi:10.1109/NAFIPS.2002.1018038.
- Chen, S.-M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, 81(3), 311–319. doi:10.1016/0165-0114(95)00220-0.
- Chen, S.-M., & Tanuwijaya, K. (2011). Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. *Expert Systems with Applications*, 38(8), 10594–10605. doi:10.1016/j.eswa.2011.02.098.
- Chen, Y. S., Cheng, C. H., & Tsai, W. L. (2014). Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting. *Applied intelligence*, 41(2), 327–347.
- Cuzzocrea, A., Song, I.-Y., & Davis, K. C. (2011). Analytics over large-scale multidimensional data. In *Proceedings of the ACM 14th international workshop on data warehousing and OLAP—DOLAP '11* (p. 101). New York, NY: ACM Press. doi:10.1145/2064676.2064695.
- Dang, X., Serfling, R., & Zhou, W. (2009). Influence functions of some depth functions, and application to depth-weighted L-statistics. *Journal of Nonparametric Statistics*, 21(1), 49–66. doi:10.1080/10485250802447981.
- Dunn, W. N. (2015). *Public policy analysis*. London: Routledge. <https://books.google.com/books?hl=en&lr=&id=lPE5CgAAQBAJ&pgis=1>. Accessed April 1, 2016.
- Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R., & Basaran, M. A. (2010). Finding an optimal interval length in high order fuzzy time series. *Expert Systems with Applications*, 37(7), 5052–5055. doi:10.1016/j.eswa.2009.12.006.
- Elliott, M. R., & Stettler, N. (2007). Using a mixture model for multiple imputation in the presence of outliers: The “Healthy for life” project. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56(1), 63–78. doi:10.1111/j.1467-9876.2007.00565.x.
- Esposito, C., Ficco, M., Palmieri, F., & Castiglione, A. (2014). A knowledge-based platform for Big Data analytics based on publish/subscribe services and stream processing. *Knowledge-Based Systems*, 79, 3–17. doi:10.1016/j.knsys.2014.05.003.
- Georgiou, D. A., & Botsios, S. D. (2008). Learning style recognition: A three layers fuzzy cognitive map schema. In *2008 IEEE international conference on fuzzy systems (IEEE World Congress on Computational Intelligence)* (pp. 2202–2207). IEEE. doi:10.1109/FUZZY.2008.4630675.
- Ghazanfari, M., Alizadeh, S., Fathian, M., & Koulouriotis, D. E. (2007). Comparing simulated annealing and genetic algorithm in learning FCM. *Applied Mathematics and Computation*, 192(1), 56–68. doi:10.1016/j.amc.2007.02.144.
- Gimenez, C., & Ventura, E. (2005). Logistics-production, logistics-marketing and external integration. *International Journal of Operations & Production Management*, 25(1), 20–38. doi:10.1108/01443570510572222.
- Giordano, R., & Vurro, M. (2010). Fuzzy cognitive map to support conflict analysis in drought management. *Fuzzy cognitive maps*. [http://link.springer.com/chapter/10.1007/978-3-642-03220-2\\_17](http://link.springer.com/chapter/10.1007/978-3-642-03220-2_17). Accessed October 30, 2015.

- Grubbs, F. (1950). Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, 21(1), 27–58.
- Grudinschi, D., Sintonen, S., & Hallikas, J. (2014). Relationship risk perception and determinants of the col-laboration fluency of buyer-supplier relationships in public service procurement. *Journal of Purchasing and Supply Management*, 20(2), 82–91. doi:10.1016/j.pursup.2014.03.004.
- Hashem, I. A. T., Yaqoob, I., Badrul Anuar, N., Mokhtar, S., Gani, A., & Ullah Khan, S. (2014). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98–115. doi:10.1016/j.is.2014.07.006.
- Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72–80. doi:10.1016/j.ijpe.2014.04.018.
- Hilbert, M. (2013). Big data for development: From information- to knowledge societies. *SSRN Electronic Journal*,. doi:10.2139/ssrn.2205145.
- Hilletoft, P., Hilmola, O.-P., & Wang, Y. (2016). Simulation based decision support systems in the supply chain context. *Industrial Management & Data Systems*,. doi:10.1108/IMDS-11-2015-0477.
- Hilletoft, P., Ujvari, S., Lattila, L., & Hilmola, O.-P. (2010). Agent-based decision support for maintenance service provider. *International Journal of Services Sciences*. <http://www.inderscienceonline.com/doi/abs/10.1504/IJSSCI.2010.032223>. Accessed June 3, 2016.
- Hong, J., & Huang, L. (2005). A split and swaying approach to building information society: The case of Internet cafes in China. *Telematics and Informatics*, 22(4), 377–393. doi:10.1016/j.tele.2004.11.005.
- Huang, Y.-L., Horng, S.-J., He, M., Fan, P., Kao, T.-W., Khan, M. K., et al. (2011a). A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, 38(7), 8014–8023. doi:10.1016/j.eswa.2010.12.127.
- Huang, Y., Horng, S., & Kao, T. (2011b). An improved forecasting model based on the weighted fuzzy relationship matrix combined with a PSO adaptation for enrollments. *International Journal of Innovative Computing Information and Control*. [https://www.researchgate.net/profile/Jui-Lin\\_Lai/publication/267844677\\_An\\_improved\\_forecasting\\_model\\_based\\_on\\_the\\_weighted\\_fuzzy\\_relationship\\_matrix\\_combined\\_with\\_a\\_PSO\\_adaptation\\_for\\_enrollments/links/54b65fde0cf24eb34fd119f.pdf](https://www.researchgate.net/profile/Jui-Lin_Lai/publication/267844677_An_improved_forecasting_model_based_on_the_weighted_fuzzy_relationship_matrix_combined_with_a_PSO_adaptation_for_enrollments/links/54b65fde0cf24eb34fd119f.pdf). Accessed March 21, 2016.
- Irani, Z., Love, P. E. D., & Jones, S. (2008). Learning lessons from evaluating eGovernment: Reflective case experiences that support transformational government. *Journal of Strategic Information Systems*, 17(2), 155–164.
- Jetter, A., & Schweinfort, W. (2011). Building scenarios with fuzzy cognitive maps: An exploratory study of solar energy. *Futures*, 43(1), 52–66. doi:10.1016/j.futures.2010.05.002.
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59–64. doi:10.1016/j.bdr.2015.01.006.
- Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, 74(7), 2561–2573. doi:10.1016/j.jpdc.2014.01.003.
- Khan, M. S., & Quaddus, M. (2004). Group decision support using fuzzy cognitive maps for causal reasoning. *Group Decision and Negotiation*, 13(5), 463–480. doi:10.1023/B:GRUP.0000045748.89201.f3.
- Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-data applications in the government sector. *Communications of the ACM*, 57(3), 78–85. doi:10.1145/2500873.
- Kok, K. (2009). The potential of fuzzy cognitive maps for semi-quantitative scenario development, with an example from Brazil. *Global Environmental Change*, 19(1), 122–133. doi:10.1016/j.gloenvcha.2008.08.003.
- Konar, A., & Chakraborty, U. (2005). Reasoning and unsupervised learning in a fuzzy cognitive map. *Information Sciences*, 170(2–4), 419–441. doi:10.1016/j.ins.2004.03.012.
- Kosko, B. (1994). Fuzzy systems as universal approximators. *IEEE Transactions on Computers*, 43(11), 1329–1333. doi:10.1109/12.324566.
- Koulouriotis, D. E., Diakoulakis, I. E., & Emiris, D. M. (2001). Learning fuzzy cognitive maps using evolution strategies: A novel schema for modeling and simulating high-level behavior. In *Proceedings of the 2001 Congress on Evolutionary Computation* (IEEE Cat. No.01TH8546) (Vol. 1, pp. 364–371). IEEE. doi:10.1109/CEC.2001.934413.
- Kuo, I.-H., Horng, S.-J., Kao, T.-W., Lin, T.-L., Lee, C.-L., & Pan, Y. (2009). An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, 36(3), 6108–6117. doi:10.1016/j.eswa.2008.07.043.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032–2033. doi:10.14778/2367502.2367572.

- Lau, H. C. W., Ning, A., Pun, K. F., Chin, K. S., & Ip, W. H. (2005). A knowledge-based system to support procurement decision. *Journal of Knowledge Management*, 9(1), 87–100. doi:[10.1108/13673270510582983](https://doi.org/10.1108/13673270510582983).
- Lee, C. C. (1990). Fuzzy logic in control systems: Fuzzy logic controller. II. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2), 419–435. doi:[10.1109/21.52552](https://doi.org/10.1109/21.52552).
- Lee, K. C., Lee, H., Lee, N., & Lim, J. (2013). An agent-based fuzzy cognitive map approach to the strategic marketing planning for industrial firms. *Industrial Marketing Management*, 42(4), 552–563.
- Love, P. E., Edwards, D. J., & Irani, Z. (2012). Moving beyond optimism bias and strategic misrepresentation: An explanation for social infrastructure project cost overruns. *IEEE Transactions on Engineering Management*, 59(4), 560–571.
- Lu, Y. Q. (2014). Research on E-government model based on big data. *Advanced Materials Research*, 989–994, 4905–4908. doi:[10.4028/www.scientific.net/AMR.989-994.4905](https://doi.org/10.4028/www.scientific.net/AMR.989-994.4905).
- Lummus, R. R., & Vokurka, R. J. (1999). Defining supply chain management: A historical perspective and practical guidelines. *Industrial Management & Data Systems*, 99(1), 11–17. doi:[10.1108/02635579910243851](https://doi.org/10.1108/02635579910243851).
- Mago, V. K., Morden, H. K., Fritz, C., Wu, T., Namazi, S., Geranmayeh, P., et al. (2013). Analyzing the impact of social factors on homelessness: A fuzzy cognitive map approach. *BMC Medical Informatics and Decision Making*, 13(1), 94. doi:[10.1186/1472-6947-13-94](https://doi.org/10.1186/1472-6947-13-94).
- Marron, B. A., & de Maine, P. A. (1967). Automatic data compression. *Communications of the ACM*, 10(11), 711–715.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, 90(10), 60–66, 68, 128. <http://europepmc.org/abstract/med/23074865>. Accessed March 31, 2016.
- Miller, H. G., & Mork, P. (2013). From data to decisions: A value chain for big data. *IT Professional*, 15(1), 57–59. doi:[10.1109/MITP.2013.11](https://doi.org/10.1109/MITP.2013.11).
- Morton, A., Ackermann, F., & Belton, V. (2003). Technology-driven and model-driven approaches to group decision support: Focus, research philosophy, and key concepts. *European Journal of Information Systems*. <http://eprints.lse.ac.uk/2562/>. Accessed April 1, 2016.
- Nishisato, S., & Ahn, H. (1995). When not to analyze data: Decision making on missing responses in dual scaling. *Annals of Operations Research*, 55(2), 361–378. doi:[10.1007/BF02030867](https://doi.org/10.1007/BF02030867).
- Osman, I. H., Anouze, A. L., Irani, Z., Al-Ayoubi, B., Lee, H., Balci, A., et al. (2014). COBRA framework to evaluate e-government services: A citizen-centric perspective. *Government Information Quarterly*, 31(2), 243–256.
- Özesmi, U., & Özesmi, S. L. (2004). Ecological models based on people's knowledge: A multi-step fuzzy cognitive mapping approach. *Ecological Modelling*, 176(1–2), 43–64. doi:[10.1016/j.ecolmodel.2003.10.027](https://doi.org/10.1016/j.ecolmodel.2003.10.027).
- Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S. J., & Fosso-Wamba, S. (2016). The role of big data in explaining disaster resilience in supply chains for sustainability. *Journal of Cleaner Production*. doi:[10.1016/j.jclepro.2016.03.059](https://doi.org/10.1016/j.jclepro.2016.03.059).
- Papageorgiou, E. I., Markinos, A. T., & Gemtos, T. A. (2011). Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. *Applied Soft Computing*, 11(4), 3643–3657. doi:[10.1016/j.asoc.2011.01.036](https://doi.org/10.1016/j.asoc.2011.01.036).
- Papageorgiou, E. I., & Salmeron, J. L. (2013). A review of fuzzy cognitive maps research during the last decade. *IEEE Transactions on Fuzzy Systems*, 21(1), 66–79. doi:[10.1109/TFUZZ.2012.2201727](https://doi.org/10.1109/TFUZZ.2012.2201727).
- Papageorgiou, E. I., Stylios, C. D., & Groumpos, P. P. (2004). Active Hebbian learning algorithm to train fuzzy cognitive maps. *International Journal of Approximate Reasoning*, 37(3), 219–249. doi:[10.1016/j.ijar.2004.01.001](https://doi.org/10.1016/j.ijar.2004.01.001).
- Piotrowicz, W., & Irani, Z. (2010). Analysing B2B electronic procurement benefits: Information systems perspective. *Journal of Enterprise Information Management*, 23(4), 559–579.
- Polk, T. A., & Seifert, C. M. (2002). *Cognitive modeling*. Cambridge: MIT Press. [https://books.google.com/books?hl=en&lr=&id=FJbIV\\_iOPjIC&pgis=1](https://books.google.com/books?hl=en&lr=&id=FJbIV_iOPjIC&pgis=1). Accessed March 24, 2016.
- Power, D. J. (2008). Understanding data-driven decision support systems. *Information Systems Management*, 25(2), 149–154. doi:[10.1080/10580530801941124](https://doi.org/10.1080/10580530801941124).
- Power, D. J., & Sharda, R. (2007). Model-driven decision support systems: Concepts and research directions. *Decision Support Systems*, 43(3), 1044–1061. doi:[10.1016/j.dss.2005.05.030](https://doi.org/10.1016/j.dss.2005.05.030).
- Power, D., Sharda, R., & Burstein, F. (2015). *Decision support systems*. <http://onlinelibrary.wiley.com/doi/10.1002/9781118785317.wcom070211/full>. Accessed March 29, 2016.
- Preuss, L. (2009). Addressing sustainable development through public procurement: The case of local government. *Supply Chain Management: An International Journal*, 14(3), 213–223. doi:[10.1108/13598540910954557](https://doi.org/10.1108/13598540910954557).

- Rose, J., Persson, J. S., Heeager, L. T., & Irani, Z. (2015). Managing e-Government: Value positions and relationships. *Information Systems Journal*, 25(5), 531–571.
- Sadrian, A. A., & Yoon, Y. S. (1994). A procurement decision support system in business volume discount environments. *Operations Research*, 42(1), 14–23.
- Sanders, N. R. (2014). *Big data driven supply chain management: A framework for implementing analytics and turning information into intelligence*. Upper Saddle River: Pearson Education.
- Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, 58(3), 26–48. doi:10.1525/cmr.2016.58.3.26.
- Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, 36(1), 120–132. doi:10.1111/jbl.12082.
- Sharif, A. M., & Irani, Z. (2006a). Applying a fuzzy-morphological approach to complexity within management decision making. *Management Decision*, 44(7), 930–961.
- Sharif, A. M., & Irani, Z. (2006b). Exploring fuzzy cognitive mapping for IS evaluation. *European Journal of Operational Research*, 173(3), 1175–1187.
- Sharif, A. M., Irani, Z., & Weerakkoddy, V. (2010). Evaluating and modelling constructs for e-government decision making. *Journal of the Operational Research Society*, 61(6), 929–952.
- Snijders, C., Matzat, U., & Reips, U.-D. (2012). “Big data”: Big gaps of knowledge in the field of Internet. *International Journal of Internet Science*, 7, 1–5.
- Song, Q., & Chissom, B. S. (1993). Forecasting enrollments with fuzzy time series—Part I. *Fuzzy Sets and Systems*, 54(1), 1–9. doi:10.1016/0165-0114(93)90355-L.
- Soulié, F. F., & Héroult, J. (Eds.). (1990). *Neurocomputing*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-76153-9.
- Stach, W., Kurgan, L., & Pedrycz, W. (2005). A survey of fuzzy cognitive map learning methods. In: Grzegorzewski, P., Krawczak, M., & Zadrozny, S. (Eds.), *Issues in soft computing: Theory and application, exit*, (pp. 71–84).
- Stach, W., Kurgan, L., & Petrycz, W. (2007). A framework for a novel scalable FCM learning method. In *Proceedings of the 2007 symposium on human-centric computing and data processing*. [https://scholar.google.ae/scholar?hl=en&q=A+Framework+for+a+Novel+Scalable+FCM+Learning+Method&btnG=&as\\_sdt=1%2C5&as\\_sdt=#0](https://scholar.google.ae/scholar?hl=en&q=A+Framework+for+a+Novel+Scalable+FCM+Learning+Method&btnG=&as_sdt=1%2C5&as_sdt=#0). Accessed March 17, 2016.
- Stylios, C., & Groumpos, P. (1999). Mathematical formulation of fuzzy cognitive maps. In *Proceedings of the 7th Mediterranean conference on control and automation*.
- Stylios, C., & Groumpos, P. (2000). Fuzzy cognitive maps in modeling supervisory control systems. *Journal of Intelligent and Fuzzy Systems*, 8(1), 83–98. <http://kic.teiep.gr/stylios/pdf/FuzzyCognitiveMapsInModelingSupervisoryControlSystems.pdf>. Accessed September 14, 2015.
- Stylios, C., & Groumpos, P. P. (2004). Modeling complex systems using fuzzy cognitive maps. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 34(1), 155–162. doi:10.1109/TSMCA.2003.818878.
- Sullivan, J., & Woodall, W. H. (1994). A comparison of fuzzy forecasting and Markov modeling. *Fuzzy Sets and Systems*, 64(3), 279–293. doi:10.1016/0165-0114(94)90152-X.
- Tucci, C. L., & Poulin, D. (2015). Introduction to the special issue on electronic government: Investment in communities, firms, technologies and infrastructure. *Electronic Commerce Research*, 15(3), 301–302. doi:10.1007/s10660-015-9195-2.
- Walker, H., & Brammer, S. (2009). Sustainable procurement in the United Kingdom public sector. *Supply Chain Management: An International Journal*, 14(2), 128–137.
- Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. doi:10.1111/jbl.12010.
- Wang, Z., Song, Q., Soh, Y. C., & Sim, K. (2013). An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation. *Computer Vision and Image Understanding*, 117(10), 1412–1420. doi:10.1016/j.cviu.2013.05.001.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.