doi: 10.14355/ijcsa.2014.0303.01

Recent Developments on Statistical and Neural Network Tools Focusing on Biodiesel Quality

Delano B. Marques*[1], Alex O. Barradas Filho[2], Alexandre R. S. Romariz*[1], Isabelle M. A. Viegas[3], Djavania A. Luz[3], Allan K.D. Barros Filho[2], Sofiane Labidi[2], Antonio S. Ferraudo[4]

[1]Department of Electrical Engineering, University of Brasilia, Brasilia-DF, Brazil

[2]Department of Electrical Engineering, Federal University of Maranhão, São Luis-MA, Brazil

[3]Department of Chemistry (LAPQAP/NEPE), Federal University of Maranhão, São Luis-MA, Brazil

[4]Department of Sciences, São Paulo State University Júlio de Mesquita Filho, Jaboticabal- SP, Brazil

[1]delanobm@gmail.com; [1]romariz@ene.unb.br

Received 16 March 2014; Revised 4 May 2014; Accepted 9 May 2014; Published 8 August 2014

Abstract

The performance of both the traditional linear regression and Artificial Neural Network (ANN) techniques has been compared to check the validity to predict the properties of biodiesel and mixtures of diesel and biodiesel. We present on this paper a review on statistical and ANN applications to the Biodiesel quality. A case study is also presented showing the prediction of oxidative stability of Biodiesel using, for the first time, other official quality parameters instead of the chemical composition as input data. In this sense, our hope is that this paper would complement a series of recent review papers and catalyze future research in this rapidly evolving area.

Keywords

Quality Assessment of Biodiesel; Prediction; Multivariate Analysis; Statistical; Neural Network

Introduction

Most standards adopted for biodiesel quality follows an international standard usually according to the EN 14214 (European) and ASTM D6751-11B (American). Table 1 shows different quality standards for biodiesel from different countries. Among the parameters established by all international norms, are those from the standardization of mineral diesel and those that were derived from analysis of vegetable oils, commonly used in the oleochemical industry (Monteiro et al., 2008).

Various studies have been published on analytical methods for biodiesel, and previous reviews directed for biodiesel analysis (Monteiro et al., 2008; Knothe, 2006; Zhang, 2012) showed the main analytical techniques developed until now. Among these analytical methods and other analytical needs stated earlier, the commonly used analytical methods for analyzing biodiesel are chromatography and spectroscopy (Zhang, 2012; Ghisi et al., 2011; Balabin et al., 2011).

Among the new methods involving the determination of inorganic contaminants, electroanalytical area is growing and showing many publications (Santos et al., 2012; Trindade et al., 2012; Tubino and Aricetti 2013; Martiniano et al., 2013).

The properties of pure biodiesel or diesel-biodiesel mixtures, when determined through instrumental analytical methods allow the gathering of much information about the biodiesel samples (Monteiro et al., 2008; Lôbo et al., 2009) but is a tedious process and as well as very time consuming (Kumar, 2010). Although there is a considerable number of articles on the application of ANN for Biodiesel quality, there are still many gaps about this subject and there are also several official parameters that have not yet been addressed. So the application of other tools in order to estimate these properties is considered of immense importance.

Chemometry based on mathematic statistical models has been widely used to assess the quality of complex samples such as biofuels, however, these tools are not always satisfactory in terms of accuracy. This fact explained the expansion of the applications of ANN and the large increase in publications in recent years, including fuels and biofuels.

In the present work we organized a text that included concepts, informations and data from the literature, related to the application of prediction methods directed to the quality of biodiesel.

The literature shows that, in almost all studies about prediction of Biodiesel quality parameters, input data into the ANN typically involves the chemical composition of Biodiesel associated with at least one quality parameter, measured experimentally. This work proposes a new and alternative way, which is the use of other quality parameters (*ester content, density, viscosity, acid value and iodine number*) of biodiesel as input data, which are official parameters, described on official norm (EN 14112).

Our intention also highlights the quality aspects of biodiesel and present recent data on statistical and ANN tools, applied to biodiesel.

The need for these mathematic and computational tools consolidate the chemometrics, which is an area designed specifically for the application of Multivariate Analysis (MVA) in the chemistry area.

TABLE 1 - QUALITY STANDARDS FOR BIODIESEL FROM DIFFERENT COUNTRIES

Property	Units	Brazil	United States	Europe	Germany	Australian	India	Japan	South Africa
		ANP	ASTM	EN	DIN	(1)	IS	JASO	SANS
		14/2012	D6751-11B	14214:2010	51606		15607	M360	1935
Density	kg/m^3	850 - 900 at 20°C	-	860 - 900 at 15°C	875 - 900 at 15°C	860 - 890 at 15°C	860 - 900 at 15°C	860 - 900	860 - 900 at 15°C
Viscosity at 40°C	mm^2/s	3.0 - 6.0	1.9 - 6.0	3.5 - 5.0	3.5 - 5.0	3.5 - 5.0	2.5 - 6.0	3.5 - 5.0	3.5 - 5.0
Water and sediment, max.	mg/kg	(2)	500	500	300	500	500	500	500
Total contamination, max.	mg/kg	24	-	24	20	24	24	24	24
Flash point, min.	°C	100	93	101	110	120	120	120	120
Ester content, min.	% mass	96.5	-	96.5	-	96.5	96.5	96.5	96.5
Carbon residue, max.	% mass	0.05 in 100% sample	0.05 in 100% sample	0.30 on the 10% distillation residue of the sample	0.03 in 100% sample	0.05 in 100% sample	0.05 in 100% sample	0.30 on the 10% distillation residue of the sample	0.30 on the 10% distillation residue of the sample
Sulfated ash, max.	% mass	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02
Sulfur, max.	mg/kg	10	15	10	10	15 for S15 grade; 500 for S500 grade	50	10	10
Sodium and potassium, max.	mg/kg	5	5	5	5	5	Report	5	5
Calcium and magnesium, max.	mg/kg	5	5	5	-	5	Report	5	5
Phosphorous, max.	mg/kg	10	10	4	10	10	10	10	10
Copper corrosion, 3h at 50 °C, max.	-	1	3	1	1	1 for sulfur < 10 mg/kg; 3 for sulfur > 10 mg/kg	1	1	1
Cetane number, min.	-	Report	47	51	49	51	50	51	51
Cold filter plugging point, max.	°C	(3)	-	(3)	-	Report	-	-	-4 (winter); +3 (summer)

Table 1 (cont.)

Property	Units	Brazil ANP 14/2012	United States ASTM D6751-11B	Europe EN 14214:2010	Germany DIN 51606	Australian	India IS 15607	Japan JASO M360	South Africa SANS 1935
Cloud point, max.	°C	-	Report	-	-	-		-	-
Cold soak filter ability, max.	Seconds	-	360	-	-	-		-	-
Acid value, max.	mg KOH/g	0.50	0.50	0.50	0.50	0.80	0.50	0.50	0.50
Free glycerin, max.	% mass	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Total glycerin, max.	% mass	0.25	0.24	0.25	0.25	0.25	0.25	0.25	0.25
Monoglyceride content, max.	% mass	0.80	-	0.80	0.80	-	-	0.80	0.80
Diglyceride content, max.	% mass	0.20	-	0.20	0.40	-	-	0.20	0.20
Triglyceride content, max.	% mass	0.20	-	0.20	0.40	-	-	0.20	0.20
Methanol and/or ethanol content, max.	% mass	0.20	0.20 or Flash point < 130°C	0.20	0.30	-	(4)	0.20	0.20
Iodine value, max.	$gI_2/100g$	Report	-	120	115		Report	120	140
Oxidation stability at 110°C, min.	Hours	6	3	6	-	-	6	-	6

(1) Barabás and Todoruț (2011)

(2) Water and sediment maximum limit is 380 mg/kg;

(3) Cold filter plugging point varies according to area of commercialization of biofuel;

(4) Alcohol content maximum is 0.20% in mass for methanol or 0.20% in mass for ethanol.

There are many MVA tools applied to the diagnosis and prediction of properties.One suggestion for the classification of these studies in the literature could involve linear and nonlinear models (Nia et al., 2014; Morsy and Sun, 2013).

The main linear models are: Multi Linear Regression (MLR) (Balabin et al., 2007), Principal Component Analysis (PCA) (Abdi and Williams, 2010), Principal Component Regression (PCR) (Balabin et al., 2007; Dodge, 2003; Jolliffe, 1982) and Partial Least Squares Regression (PLS) (Balabin and Safieva, 2011).

PLS is a statistical (multivariate) method with a close relationship to PCR (Balabin et al., 2007). The main difference of the PLS method is that both the input (X) and output (Y) data are projected to new spaces. In the case of biodiesel analysis, X and Y are spectra and properties, respectively. Because of its simplicity and lower requirements for computational resources, PLS is widely used in the analysis of very different datasets. In addition, this approach can be regarded as a linear standard method for data analysis (Balabin et al., 2011).

One of the main non-linear calibration models are neural networks (Balabin et al., 2007; Balabin et al., 2011). ANNs (Schalkoff, 1997) are a functional abstraction of the biological neural structures of the central nervous system (Anderson, 1983; Akkurt et al., 2003) and they can exhibit a surprising number of the characteristics of human brain, e.g., learning from experience and generalizing from previous examples to solve new problems (Öztaş et al., 2006). Typically, a network consists of a set of sensing units (source nodes) that constitute the input layer, one or more hidden layers of computational nodes. The input signal is propagated forward through the network, layer by layer. Generally, neural networks are adjusted, or trained, in order to achieve a particular target for a give output.

The multilayer perceptrons (MLP) have been successfully applied to solve many difficult problems through their training in a controlled manner with very popular algorithm known as error back propagation algorithm (Rumelhart et al., 1986). An example of a network topology multilayer perceptron is shown in Figure 1 containing an initial layer with four input variables, two hidden layers with three neurons in each one and an output layer.

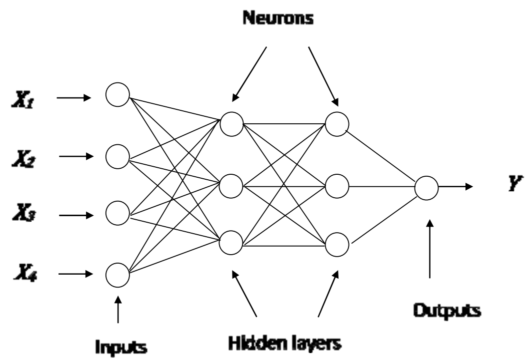

FIGURE 1 – TOPOLOGY OF A MULTILAYER PERCEPTRON NEURAL NETWORK

In terms of applications in the analytical chemistry area, computational tools, including statistic and ANNs is gaining more interest mainly focusing on complex matrices as petroleum and biofuels.

An overview of ANNs models has been provided over the last years by various authors (Rumelhart et al., 1986; Wu et al., 2014; Gang and Wang, 2013) who conducted research involving mathematical description of ANNs models, training algorithms such as supervised/unsupervised learning.

A very small number of papers (Canha et. al., 2012 and Meira et al., 2011; De Lira et. al., 2010) were found in the literature on application of chemometric tools to predict the oxidative stability of biodiesel. This motivated the present work, which proposed the use of ANN instead of chemometrics tools to predict the oxidative stability that is an important biodiesel quality parameter.

Comparison Between Statistical and Ann Tools

The main difference between ANN and statistical regression parameters is related to the structure and the degree of difficulty of each of these tools. The results generated by a statistical regression equation are simple to be analyzed and used to identify input-output relationships and the effectiveness of each input variable. On the other hand, when using ANNs, the network should firstly learn the relationships between variables (Kisi, 2005), the information is then stored in convoluted network weights and biases, which are difficult to interpret. Despite of ANNs which are extremely versatile to perform the mapping of complex and nonlinear relationships, some disadvantages are described on the literature (Sablani, 2008).

The main success and advantages of the expansion of neural networks methods, nowadays, are due to the association between modern instrumental analytical methods and the efficient regression tools provided by modern mathematics. This association also makes possible the creation of accurate and robust methods for the prediction of object properties possible (Balabin and Lomakina, 2009; Balabin, 2009; Manly, 2005; Hobro and Lendl, 2009; Sun et al., 2010; Balabin et al., 2008). Therefore, progress in chemometrics has a direct effect on the field of analytical chemistry. A featured paper recently published (Nia et al., 2014) presents a comprehensive comparison of nonlinear methods (Kernel PLS (KPLS), support vector machines (SVM), least-squares SVM (LS-SVM), relevance vector machines (RVM), Gaussian process regression (GPR), ANN, and Bayesian ANN (BANN)) where a PLS regression is used as a linear benchmark.

Numerous authors have compared performance of statistical and neural networks models on specific problems (Manjunatha et al., 2010; Sarle, 1994; Schumacher et al., 1996; Wu, 1992; Hruschka, 1993). Balabin and his col l aborators have been those among the leading researchers who have published on the relationship between statistical methods and neural networks comparing them in the application of these techniques in predicting parameters of Biodiesel quality. An important part of the matters presented and discussed here comes from their work. A larger and more detailed view on the subject can be found in their papers (Balabin et al., 2011; Balabin and Smirnov, 2011; Balabin et al., 2007; Balabin et al., 2008; Balabin, 2009; Syunyaev and Balabin, 2007; Syunyaev and Balabin, 2008; Balabin and Lomakina, 2011; Balabin and Smirnov, 2012).

Recent Prediction Methods Developed for Biodiesel Quality

Many of the well established instrumental methods, applied to biodiesel, combined with multivariate analysis, have been studied in order to establish calibration models, with many advantages including acceptable precision and accuracy, still being quick and not very laborious (Ghesti et al., 2006; Blanco et al., 2004; Brásio et al., 2013; Cheenkachorn, 2006).

Several multivariate calibration models are presented in the literature (Zhang, 2012; Lôbo et al., 2009; Balabin and Lomakina, 2009; Balabin, 2009; Balabin and Lomakina, 2011; Balabin and Smirnov, 2012), where it is shown that the data obtained by instrumental techniques are treated by multivariate analysis, the quantification of analytes being possible even in the presence of interferents (Lôbo et al., 2009).

Cheenkachorn (2006) used statistical models and ANN tools to predict several properties of biodiesels using the fatty acid compositions of various vegetable oils. In this work the authors used data mainly collected from previous literature. The input variables of the models included different types of fatty acids such as palmitic acid, stearic acid acid, oleic acid, etc. and the predicted properties included viscosity, high-heating value (HHV), and cetane number. The results showed that both statistical models and the ANNs precisely predicted the properties of biodiesel derived from certain vegetable oils.

Balabin et al. (2011) compared the performance of linear and non-linear calibration techniques (MLR, PCR, PLS, polynomial PLS (Poly-PLS) and Spline-PLS versions, and ANN) for prediction of biodiesel properties from near infrared spectra. The model was created for four important biodiesel properties: density (at 15 ºC), kinematic viscosity (at 40 ºC), water content, and methanol content. They concluded that the ANN approach was superior to the linear (MLR, PCR, PLS) and ‘’quasi’’-non-linear (Poly-PLS, Spline-PLS) calibration methods.

Balabin and Smirnov (2011) compared the performance of 16 different feature selection methods for the prediction of properties of biodiesel fuel, including density, viscosity, methanol content, and water concentration. A comparison with a non-linear calibration model, (MLP-ANN) was provided. Among the feature selection algorithms tested were MLR, Interval PLS (iPLS), back-propagation ANN (BP-ANN) and others. Two linear techniques for calibration model building (MLR and PLS) were used for the evaluation of biofuel properties. The results of other spectroscopic techniques application, such as Raman, ultraviolet-visible (UV-Vis), or nuclear magnetic resonance (NMR) spectroscopies, could be greatly improved by an appropriate feature selection choice.

Balabin and Smirnov (2012) in their recent study made a rather general comparison of linear, such as partial least squares or projection to latent structures (PLS); quasi-nonlinear, such as the polynomial version of PLS (Poly-PLS); and intrinsically non-linear, such as artificial neural networks (ANNs), support vector regression (SVR), and least-squares support vector machines (LS-SVM/LSSVM), regression methods in terms of their robustness. Biodiesel was one of the systems chosen as representative examples of real-world samples in this study.

Response Surface Methodology (RSM) and ANN were employed to study the relationship between process variables and free fatty acid conversion and for predicting the optimal parameters (Talebian-Kiakalaieh et al., 2013). The highest conversion found was 88.6% at optimum condition being 14h, 65°C, 70:1 and 10wt% for reaction time, reaction temperature, methanol to oil molar ratio and catalyst loading, respectively. The RSM and ANN could accurately predict the experimental results, with R^2=0.9987 and R^2=0.985, respectively.

Studies involving statistical and neural network tools showed predictions on the biodiesel content in biodiesel/diesel blends. This theme has been one of the aplications most studied involving chemometrics and ANN (Balabin and Lomakina, 2011; Ramadhas et al., 2006; Pimentel et al., 2006; Kalogeras et al., 2010; Fernandes et al., 2011) mainly including spectroscopic techniques.

Predictions involving engine emission characteristics, and engine performance was another themeexplored (Kumar et al., 2013). Good correlation were obtained recently (Kumar et al., 2013; Prasad et al., 2010; Kumar et al., 2012) between the ANN predicted values and the desired values for various engine performance values and the exhaust emissions.

Many other important properties of biodiesel and biodiesel/diesel blend also were objects of studies on predictions in the last years, such as flash point, fire point (Kalogeras et al., 2010; Saldana et. Al., 2011; Kumar and Bansal, 2007), methanol content (Balabin et al., 2011; Felizardo et al., 2007); cloud point, pour point, volatility at different temperatures 250 °C, 350 °C and 360 °C, cetane index, sulfur, flash point and cold filter plugging point (CFPP) (Kalogeras et al., 2010); water content (Balabin and Lomakina , 2011; Balabin et al., 2011; Felizardo et al., 2007); cold filter plugging point (Baptista et al., 2008); vapor pressures (Freitas et al., 2012); biodiesel feedstock (Zawadzki and Shrestha, 2009); oxidative stability (Canha et al., 2012); concentration of triglycerides (Oliveira et al., 2007); biodiesel production process modelling (Hui, 2012); engine torque, specific fuel consumption and exhaust gas components (Najafi et al., 2007) discrimination of methyl biodiesel from different sources (Flores et al., 2012).

Table 2 presents a selection of prediction methods applied to standard parameters (cetane number, viscosity, ester content, density, flash point, methanol and/or ethanol content, water content, oxidative stability, iodine value, cold filter plugging point, cloud point, pour point and blend) related to the quality of biodiesel, available in the current literature.

As we can see, there are a number of the standard parameters that still have not been addressed with this type of statistical or computational tools. On the other hand, many other studies involving biodiesel quality and applications of statistical and computational tools also have been found in the literature.

TABLE 2 –PREDICTION METHODS APPLIED TO STANDARD QUALITY OF BIODIESEL

Characteristic	Official Method	Prediction Method	Tools	References
Cetane Number	EN ISO 5165; ASTM D613	Artificial Neural Network	multi-layer feed forward, radial base function, generalized regression, recurrent network, backpropagation algorithm	Ramadhas et al., (2006) Nadai et. al. (2013)
		Others	linear regression, best subset method, empirical equation, principal component analysis	Cheenkachorn (2006) Bamgboye and Hansen (2008) Sivaramakrishnan and Ravikumar (2012) Nadai et. al. (2013)
Viscosity	EN ISO 3104; ASTM D445; ABNT NBR 10441	Artificial Neural Network	multi-layer feed forward, backpropagation algorithm, levenberg–marquardt algorithm, kohonen algorithm, batch gradient descent with momentum algorithm, scaled conjugate gradient algorithm	Cheenkachorn (2006) Baptista et al. (2008) Kumar and Bansal (2010) Balabin et al. (2011) Balabin and Smirnov (2011) Saldana et al. (2012)
		Others	linear regression, best subset method, multiple linear regression, principal component regression, partial least squares regression, poly partial least squares, spline partial least squares, eyring's equation transformed, support vector machine, stepwise multiple linear regression, interval partial least squares regression, backward interval partial least squares regression, forward interval partial least squares regression, moving window partial least squares regression, (modified) changeable size moving window partial least squares, searching combination moving window partial least squares, successive projections algorithm, uninformative variable elimination, including uninformative variable elimination, simulated annealing, genetic algorithms, projection to latent structures	Cheenkachorn (2006) Krisnangkura et al. (2010) Kumar and Bansal (2010) Balabin et al. (2011) Balabin and Smirnov (2011) Zhang (2012) Saldana et al. (2012)
Ester Content	EN ISO 14103; ABNT 15764	Artificial Neural Network	modular feedforward	Oliveira et al., (2006) Oliveira et al., (2007)
		Others	partial least squares regression, principal component regression	Oliveira et al., (2006) Oliveira et al., (2007)
Density	EN ISO 3675; EN ISO 12185 ASTM D1298; ASTM D 4052; ABNT NBR 7148; ABNT NBR 14065	Artificial Neural Network	forward backpropagation network, multi-layer feed forward, levenberg–marquardt algorithm,kohonen algorithm, batch gradient descent with momentum algorithm, scaled conjugate gradient algorithm	Baroutian et al. (2008) Kumar and Bansal (2010) Balabin et al. (2011) Balabin and Smirnov (2011) Saldana et al. (2012)
		Others	linear mixing rules, gcvol group contribution method, extended gcvol group contribution method, support vector machine, multiple linear regression, principal component regression, principle of least squares regression, partial least squares regression, poly partial least squares regression, spline partial least squares regression, synergy partial least squares stepwise multiple linear regression, interval partial least squares regression, backward interval partial least squares regression, forward interval partial least squares regression, moving window partial least squares regression, (modified) changeable size moving window partial least squares, searching combination moving window partial least squares regression, successive projections algorithm, uninformative variable elimination, including uninformative variable elimination, simulated annealing, genetic algorithms, projection to latent structures	Baptista et al. (2008) Kumar and Bansal (2010) Ferrão et al. (2011) Balabin and Smirnov (2011) Saldana et al. (2012)
Flash Point	EN ISO 3679; ASTM D93; ABNT NBR 14598	Artificial Neural Network	batch gradient descent with momentum algorithm, levenberg-marquardt algorithm, scaled conjugate gradient algorithm	Kumar and Bansal (2010)
		Others	partial least squares regression, interval partial least squares regression, synergy partial least squares regressions, principle of least squares regression	Ramos (2009) Kumar and Bansal (2010)

				Ferrão et al. (2011)
Methanol And/Or Ethanol Content	EN 14110; ABNT NBR 15343	Artificial Neural Network	multi-layer perceptron, levenberg–marquardt algorithm, backpropagation algorithm, kohonen algorithm	Balabin et al. (2011) Balabin and Smirnov (2011)
		Others	multiple linear regression, principal component regression, partial least squares regression, poly-partial least squares regression, spline-partial least squares regression, stepwise multiple linear regression, interval partial least square regression, backward interval partial least square regression, forward interval partial least square regression, moving window partial least squares regression, changeable size moving windows partial least squares, searching combintion moving windows partial least squares, modified changeable size moving windows partial least squares, successive projections algorithm, uninformative variable elimination, including uninformative variable elimination, simulated annealing, genetic algorithms, genetic algorithm interval partial least squares.	Balabin et al. (2011) Balabin and Smirnov (2011)
Water Content	EN ISO 12937; ASTM D 6304	Artificial Neural Network	multi-layer perceptron, levenberg–marquardt algorithm, backpropagation algorithm, kohonen algorithm	Balabin et al. (2011) Balabin and Smirnov (2011)
		Others	multiple linear regression, principal component regression, partial least squares regression, poly-partial least squares regression, spline-partial least squares regression, stepwise multiple linear regression, interval partial least square regression, backward interval partial least square regression, forward interval partial least square regression, moving window partial least squares regression, changeable size moving windows partial least squares, searching combintion moving windows partial least squares, modified changeable size moving windows partial least squares, successive projections algorithm, uninformative variable elimination, including uninformative variable elimination, simulated annealing, genetic algorithms, genetic algorithm interval partial least squares.	Balabin et al. (2011) Balabin and Smirnov (2011)
Oxidative Stability, 110 °C	EN 14112	Artificial Neural Network	-	-
		Others	partial least squares regression, interval partial least-squares regression, multiple linear regression, successive projections algorithm, principal component analysis	De Lira et al. (2010) Meira et al. (2011) Canha et al. (2012)
Iodine Value	EN 14111	Artificial Neural Network	-	-
		Others	multiple linear regression, principal component analysis, partial least squares regression, poly-partial least squares regression, spline-partial least squares regression	Baptista et al. (2008) Gopinath et al. (2009) Balabin and Safieva (2011)
Cold Filter Plugging Point	EN 116; ASTM D6371; ABNT NBR 14747	Artificial Neural Network	-	-
		Others	principal component analysis, partial least squares regression, universal quasichemical model, poly-partial least squares regression, spline-partial least squares regression	Baptista et al. (2008) Lopes et al. (2008) Balabin and Safieva (2011)
Cloud Point	EN 23015; ASTM D2500	Artificial Neural Network	-	-
		Others	universal quasichemical model	Lopes et al. (2008)
Pour Point	ASTM D-97; ASTM D-5949	Artificial Neural Network	-	-
		Others	universal quasichemical model	Lopes et al. (2008)
Blend	EN 14078; ABNT NBR 15568	Artificial Neural Network	-	-
		Others	partial least squares regression, interval partial least squares regression, partial least squares by exclusion regression, multiple linear regression, successive projections algorithm, principal component analysis, multivariable calibration, root mean squared error of prediction, least-squares statistical regression.	Pimentel et al. (2006) De Oliveira et al. (2009) Mahamuni and Adewuyi (2009) Francesquett et al. (2010) Fernandes et al. (2011) Moser (2012) De Vasconcelos et al. (2012)

Prediction of the Oxidative Stability of Biodiesel Using Ann Based on its Physicochemical Properties: A Case Study

Oxidative stability is an important parameter in evaluating the quality of biodiesel fuel. Its determination is usually arduous, expensive, analysis is very time consuming, and the results obtained are not always accurate due to experimental errors . This work is aimed at predicting the oxidative stability of biodiesel based on its physicochemical properties.

The changes that occur due to the biodiesel oxidation, such as the formation of hydroperoxides, which may polymerise with other free radicals to form insoluble sediments and gums, can compromise fuel quality, which means, effects on physicochemical parameters such as ester content, viscosity, density, iodine number and acid value. These alterations as a result of oxidation reactions can also cause the corrosion of fuel system components, hardening of rubber components and fusion of moving components (Monyem, 2001; Ortech Corporation, 1995; EN 14112/2003).

Rancimat is an official test method (EN 14112, EN 15751) used in almost all countries to determine oxidative stability of Biodiesel. It is an accelerated oxidation test in which the biodiesel to be tested is run at elevated temperatures (110 °C) whilst exposing the sample to a stream of purified air (10 L/h) accelerating the oxidation process of the biodiesel. After passing through the biodiesel, the air is fed into a collection flask containing distilled water and a probe to measure conductivity.

As the biodiesel sample degrades, the volatile organic acids produced are carried to the collection flask, and the conductivity of the solution is recorded by the probe. Oxidation stability will be given by the induction period, defined as the time between the start of the test and the sudden conductivity increase of the solution in the collection flask. This results in auto-oxidation in a few hours, instead of months.

Materials and Methods

Selected data of pure biodiesels related to the physicochemical properties, ester content, iodine number, viscosity, density, acid value and oxidative stability, used in the present work, were obtained from the interlaboratory official program of the Brazilian government, coordinated by The Brazilian National Agency of Petroleum, Natural Gas and Biofuels (from Portuguese: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis - ANP)].

The data corresponded to the results from the analysisthat were carried out, and provided through reports, by different testing laboratories, participating in this important program in Brazil. These laboratories followed the official ANP procedures (ASTM D6751-11B, EN 14.214:2010 and NBR RANP 14/2012) for Biodiesel analysis. The biodiesel used was provided by manufacturers approved by ANP, and was obtained via methylic route.

The automatic module of the artificial neural networks of the Statistics 10.0 software was used in this study.

The MLP used in this study, consisted of 5 input variables (ester content, density, viscosity, acid value and iodine number), one hidden layer, in which the number of neurons varied from 3 to 30 during the training phase, and an output layer (induction period). The learning algorithm employed was the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method that is a variation of the backpropagation.

BFGS is a quasi-Newton method that uses the quadratic Taylor approximation of function objective, whose idea is to make an iterative approximation of the matrix Hessian inverse (H_i). The current approximation H_i is used in each iteration to define the next descending direction of the method. Ideally, the approximations converge to the inverse of the Hessian matrix and they are carried out using Equation 1, where q_i and p_i are vectors determined by other equations (Silva, 1998).

$$H_{i+1} = H_i + \frac{p_i p_i^T}{p_i^T q_i}\left[1 + \frac{q_i^T H_i q_i}{p_i^T q_i}\right] - \frac{H_i q_i p_i^T + p_i q_i^T H_i}{p_i^T q_i}$$

(Eq. 1)

In this work, the following activation functions were tested: identity, logistic, hyperbolic tangent, sine and exponential.

In the supervised learning paradigm, the output values generated by the network are ~~is~~ compared to the desired values ~~output~~ using Equation 2, which is the squared error cost function to be minimized (Sum of Squares (SOS)), where $d^{(1)}, d^{(2)}, \ldots, d^{(p)}$ are the output values of the training examples (desired output) and $y^{(1)}, y^{(2)}, \ldots, y^{(p)}$ are the values calculated by the network. The training happens until the error is below a threshold predefined or until the number of iterations or cycles is reached (Jain, Mao and Mohiuddin, 1996).

$$E = \frac{1}{2}\sum_{i=1}^{p} \| y^{(i)} - d^{(i)} | \|^2$$

(Eq. 2)

Results and Discussion

The experiments were done to determine the best number of neurons in the hidden layer and to evaluate the performance of ANNs in predicting induction period of biodiesel. Under supervised learning, the desired output for the input variables was given to the network.

453 samples and their respective physicochemical parameters were used to generate the evaluated database. Several samples of this database were incomplete with respect to the physicochemical analysis . Missing data were supplemented with the median of each parameter. Besides the size of database, the type of data is also extremely important for predictive studies.

Database was divided into subgroups: 80% of the data were used for training purposes and 20% for testing purpose and also to verify the generalizability of the trained network. This was because, in order to make the prediction, the network learnt a rule using the training examples. The network was fed with the five parameters established before, as the main parameters influenced by the biodiesel oxidation, seconded the literature (Canha et al., 2012; Hoekman et al. (2012).

Results for the performance of the ten best ANNs with different numbers of neurons are presented in Table 3 for the database.

As we can see all networks showed good performance to predict the oxidative stability of biodiesel, however, preference should be given for models with simpler topologies with minimum error and maximum value of the correlation coefficient.

TABLE 3 - TOPOLOGY, CORRELATION COEFFICIENT, ERROR AND ACTIVATION FUNCTIONS OF THE 10 BEST MLP NEURAL NETWORKS TRAINED.

Network	Topology	Correlation Coefficient		Error		Activation function	
		Training	Test	Training	Test	Hidden layer	Output
1	5-26-1	0.778388	0.889690	0,002324	0,002254	Tanh	Tanh
2	5-6-1	0.785254	0.884360	0,002262	0,002456	Tanh	Logistic
3	5-14-1	0.733176	0.885739	0,002728	0,002345	Tanh	Sine
4	5-10-1	0.788032	0.890015	0,002238	0,002323	Tanh	Logistic
5	5-16-1	0.775358	0.884420	0,002361	0,002383	Tanh	Logistic
6	5-13-1	0.801687	0.891741	0,002107	0,002237	Tanh	Logistic
7	5-28-1	0.800537	0.884146	0,002119	0,002409	Tanh	Tanh
8	5-8-1	0.816125	0.890790	0,001970	0,002216	Tanh	Tanh
9	5-6-1	0.772202	0.889273	0,002385	0,002477	Tanh	Identity
10	5-22-1	0.781152	0.884469	0,002303	0,002346	Tanh	Tanh

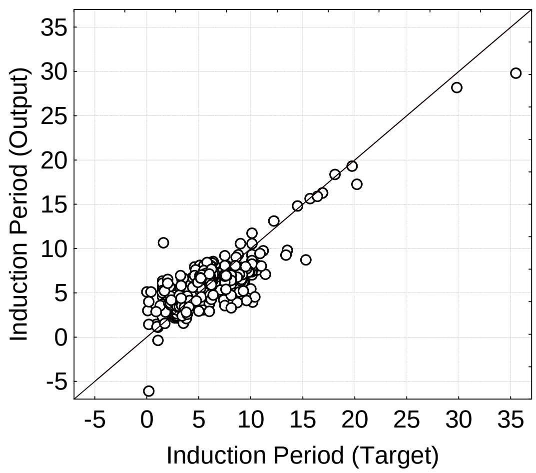

FIGURE 2 - CORRELATIONS BETWEEN EXPERIMENTAL VALUES (TARGET) AND PREDICTED (OUTPUT) INDUCTION PERIOD BY ANN/BFGS METHOD.

This work confirmed the general dependence of oxidative stability property (induction period) on physicochemical properties, ester content, iodine number, viscosity, and density.

Among the 10 networks that presented the best performance, one containing 8 neurons in the hidden layer was highlighted due to the best coefficients presented, whose medium correlation coefficient was 0.85 to this network.

Figure 2 shows the results obtained after application of the ANN chosen to predict oxidative stability. The performance of the model was carried out simultaneouly for both sets, calibration and validation.

From this figure it is possible to observe that although some samples present anomalous behavior (outliers), in the general the predicted values are fairly uniform and well distributed along the straight line of correlation, with a higher number of samples presenting an oxidative stability lower than 10 h and a medium induction period value of 6.16 h. These results correspond to region and value expected for the most of the Biodiesel biofuel samples. Predicted values for all samples range from 0.00 to 35.5.

Furthermore, in spite of its large calibration range, the model allows obtaining an excellent agreement between the values predicted by the model and those determined by the reference method. Therefore, these results confirm that this model can be used to predict the oxidative stability of biodiesel with values similar to the Rancimat ones.

Conclusions

The aim of this work was to present an update of the literature about the application of ANN to predict quality parameters of biodiesel. The work also presented a case study on the prediction of oxidative stability.

A first important conclusion is clear, namely that these computational tools are very important to solve complex problems in analytical chemistry.

The published results show that both statistical models and the ANNs precisely can predict the properties of biodiesel.

The number of publications presented here is neither complete and can appear exhaustive but this is merely a sample of publications that demonstrate, on the one hand, the usefulness of applications of ANNs and statistical tools currently and, on the other hand also show how large the interest in biodiesel is and is growing increasingly.

The existence of many standards that have not yet been addressed by these tools and, on the other hand, the eficiency and importance observed of the application of statistical and computational tools directed to the quality parameters of Biodiesel indicated clearly the importance of continuity of these studies.

A case study showed the prediction of oxidative stability of Biodiesel. The medthod used 4 other official quality parameters as input data which avoided the need of additional experiments to determine the mass fractions of FAMEs by chromatography, usually required by the prediction methods for biodiesel quality parameters.

Results indicated that the proposed ANN method was able to predict the induction period with the medium error of 0.002, considered an adequate and small value. In terms of accuracy, the prediction method presented a good value whose medium correlation coeficient was 0.85 for the network number 8. The application of the model presented a medium value of IP of 6.6h. Good correlation was observed between the experimental values (target) and the predicted induction period (output) for all biodiesel samples by the method proposed.

Based on the results presented here, both the data from the recent literature as the case study on the prediction of oxidative stability of Biodiesel, indicated that ANNs offered a very important alternative method for Biodiesel quality, which should not be under-estimated, mainly considering that ANN was extremely versatile for mapping complex and nonlinear relationships.

ACKNOWLEDGMENT

The authors thank Brazilian National Agency of Petroleum, Natural Gas and Biofuels (ANP/CPT) from Brazil by the supplying the database on biodiesel analysis for this research and UFMA-LAPQAP/NEPE for the support and colaborations during the development of this research. The authors also thank the Brazilian agencies FAPDF, FINEP, and CAPES for financial support and provision of research fellowships and scholarships.

REFERENCES

Abdi H, Williams L.J. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics 2010, 2(4):433-59.

Akkurt S, Ozdemir S, Tayfur G, Akyol B. The use of GA-ANNs in the modelling of compressive strength of cement mortar. Cement Concrete Res. 2003;33(7):973-9.

Anderson JA. Cognitive and psychological computation with neural models. IEEE Trans. Sys. Man Cybern. 1983,13(5):799-814.

Balabin RM, Lomakina EI, Safieva RZ. Neural network (ANN) approach to biodiesel analysis: analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy. Fuel 2011,90(5):2007-15.

Balabin RM, Lomakina EI. Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. J. Chem. Phys. 2009,131:074104.

Balabin RM, Lomakina EI. Support vector machine regression (SVR/LS-SVM) - an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. Analyst 2011, 136:1703-11.

Balabin RM, Safieva RZ, Lomakina EI. Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. Chemometr. Intell. Lab. Syst. 2007,88:183-8.

Balabin RM, Safieva RZ, Lomakina EI. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. Chemometr. Intell. Lab. Syst. 2008, 93:58-63.

Balabin RM, Safieva RZJ. Near-infrared (NIR) spectroscopy for biodiesel analysis: fractional composition, iodine value, and cold filter plugging point from one vibrational spectrum. Energy Fuels 2011,25:2373-82.

Balabin RM, Smirnov SV. Interpolation and extrapolation problems of multivariate regression in analytical chemistry: benchmarking the robustness on near-infrared (NIR) spectroscopy data. Analyst 2012,137:1604-10.

Balabin RM, Smirnov SV. Variable selection in near-infrared spectroscopy: benchmarking of feature selection methods on biodiesel data. Anal. Chim. Acta. 2011,692(1-2):63-72.

Balabin RM. Conformational equilibrium in glycine: experimental jet-cooled Raman spectrum. J. Phys. Chem. Lett. 2009,1(1):20-3.

Bamgboye AI, Hansen AC. Prediction of cetane number of biodiesel fuel from the fatty acid methyl ester (FAME) composition. International Agrophysics 2008;22:21-9.

Baptista P, Felizardo P, Menezes JC, Correia MJN. Multivariate near infrared spectroscopy models for predicting the iodine value, CFPP, kinematic viscosity at 40°C and density at 15°C of biodiesel. Talanta 2008,77:144-51.

Barabás I, Todoruț IA. Biodiesel Quality, Standards and Properties. Biodiesel-Quality, Emissions and By-Products 2011:3-28.

Baroutian S, Aroua MK, Raman AA, Sulaiman NM. Estimation of vegetables oil-based ethyl esters biodiesel densities using artificial neural networks. Journal of Applied Sciences 2008,8(17):3005-11.

Blanco M, Beneyto R, Castillo M, Porcel M. Analytical control of an esterification batch reaction between glycerine and fatty acids by near-infraredspectroscopy. Anal. Chim. Acta 2004;521(2):143-8.

Brásio ASR, Romanenko A, Leal J, Santos LO, Fernandes NCP. Nonlinear model predictive control of biodiesel production via transesterification of used vegetable oils. Journal of Process Control. 2013,23:1471-9.

Canha N, Felizardo P, Menezes JC, Correia MJN. Multivariate near infrared spectroscopy models for predicting the oxidative stability of biodiesel: effect of antioxidants addition. Fuel 2012,97:352-7.

Cheenkachorn K. Predicting properties of biodiesels using statistical models and artificial neural networks. As. J. Energy Env. 2006,7(2):299-306.

De Lira LFB, De Albuquerque MS, Pacheco JGA, Fonseca TM, Cavalcanti EHDS, Stragevitch L, et al. Infrared spectroscopy and multivariate calibration to monitor stability quality parameters of biodiesel. Microchem. J. 2010,96(1):126-31.

De Oliveira IK, Rocha WFC, Poppi RJ. Application of near infrared spectroscopy and multivariate control charts for monitoring biodiesel blends. Anal. Chim. Acta 2009,642:217-21.

De Vasconcelos FVC, De Souza JR. PFB, Pimentel MF, Pontes MJC, Pereira CF. Using near-infrared overtone regions to determine biodiesel content and adulteration of diesel/biodiesel blends with vegetable oils. Anal. Chim. Acta. 2012,716(24):101-7.

Dodge Y. The Oxford Dictionary of Statistical Terms: Oxford University Press, 2003.

EN 14112/2003. Fat and oil derivatives, Fatty acid methyl esters (FAME), Determination of oxidation stability (accelerated oxidation test).

Felizardo P, Baptista P, Menezes JC, Correia MJN. Multivariate near infrared spectroscopy models for predicting methanol and water content in biodiesel. Anal. Chim. Acta 2007,595(1-2):107-13.

Fernandes DDS, Gomes AA, Da Costa GB, Da Silva GWB, Véras G. Determination of biodiesel content in biodiesel/diesel blends using NIR and visible spectroscopy with variable selection. Talanta 2011,87:30-4.

Ferrão MF, Viera MS, Pazos REP, Fachini D, Gerbase AE, Marder L. Simultaneous determination of quality parameters of biodiesel/diesel blends using HATR-FTIR spectra and PLS, iPLS or siPLS regressions. Fuel 2011,90(2):701-6.

Flores IS, Godinho MS, Oliveira AE, Alcantara GB, Monteiro MR, Menezes SMC, et al. Discrimination of biodiesel blends with 1H NMR spectroscopy and principal component analyses. Fuel 2012,99:40-4.

Francesquett JZ, Viera MS, Lovato RH, Fachini D, Gerbase AE, Costa AB, et al. Modelos de regressão multivariada empregando seleção de intervalos para a quantificação do biodiesel em blendas biodiesel/diesel. Tecno-Lógica 2010,14(2):87-92.

Freitas SVD, Oliveira MB, Lima AS, Coutinho JAP. Measurement and prediction of biodiesel volatility. Energy Fuels 2012,26:3048-53.

Gang, W. and Wang, J. Predictive ANN models of ground heat exchanger for the control of hybrid ground source heat pump systems. Applied Energy 2013, 112: 1146-53.

Ghesti GF, Macedo JL, Braga VS, Souza ATCP, Parente VCI, Figuerêdo ES, et al. Application of Raman spectroscopy to monitor and quantify ethyl esters in soybean oil transesterification. J. Am. Oil Chem. Soc. 2006,83:597-601.

Ghisi M, Chaves ES, Quadros DP, Maques EP, Curtius AJ, Marques ALB. Simple method for the determination of Cu and Fe by electrothermal atomic absorption spectrometry in biodiesel treated with tetramethylammonium hydroxide. Microchem. J. 2011,98:62-5.

Gopinath A, Puhan S, Nagarajan G. Theoretical modeling of iodine value and saponification value of biodiesel fuels from their fatty acid composition. Renew. Energ. 2009,34(7):1806-11.

Hobro AJ, Lendl B. Stand-off Raman spectroscopy. Trends Anal. Chem. 2009,28:1235-42.

Hoekman, S. K. et al. Review of biodiesel composition, properties, and specifications. Renew. Sust. Energ. Rev. 2012, 16: 143-169.

Hruschka H. Determining market response functions by neural network modeling: A comparison to econometric techniques. European Journal of Operational Research. 1993,66(1):27-35.

Hui LW. Case study of artificial neural network modeling on catalyzed and enzymatic transesterification process for biodiesel production. J. Appl. Sci. Res. 2012,8(3):1672-81.

Jain AK, Mao J, Mohiuddin KM. Artificial Neural Networks: A Tutorial. Computer 1996,29(3):31-44.

Jolliffe IT. A note on the Use of Principal Components in Regression. Journal of the Royal Statistical Society: Series C (Applied Statistics) 1982,31(3):300-3.

Kalogeras K, Bezergianni S, Kazantzi V, Pilavachia PA. On the prediction of properties for diesel /biodiesel mixtures featuring new environmental considerations. Comput. Aided Chem. Eng. 2010,28:973-8.

Kisi O. Daily River Flow Forecasting Using Artificial Neural Networks and Auto-Regressive Models. Turkish J. Eng. Env. Sci. 2005,29:9-20.

Knothe G. Analyzing biodiesel: standards and other methods. J. Am. Oil Chem. Soc. 2006,83(10):823-33.

Krisnangkura K, Sansa-Ard C, Aryusuk K, Lilitchan S, Kittiratanapiboon K. An empirical approach for predicting kinematic viscosities of biodiesel blends. Fuel 2010,89(10):2775-80.

Kumar J, Bansal A. Application of artificial neural network to predict properties of diesel – biodiesel blends. Kathmandu University Journal of Science, Engineering and Technology 2010,06(2):98-103.

Kumar J, Bansal A. Selection of best neural network for estimating properties of diesel-biodiesel blends. In Proceedings of the 6th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases 2007:16-9.

Kumar S, Pai PS, Shrinivasa RBR. Radial-basis-function-network-based prediction of performance and emission characteristics in a bio diesel engine run on WCO ester. Advances in Artificial Intelligence 2012,2012:1-7.

Kumar S, Srinivas PP, Shrinivasa RBR. Influence of injection timings on performance and emissions of a biodiesel engine operated on blends of Honge methyl ester and prediction using artificial neural network. J. Mech. Eng. Res. 2013,5(1):5-20.

Lôbo IP, Ferreira SLC, Cruz RS. Biodiesel: parâmetros de qualidade e métodos analíticos. Quim. Nova. 2009,32(6):1596-608.

Lopes JCA, Boros L, Krähenbühl MA, Meirelles AJA, Daridon JL, Pauly J, et al. Prediction of cloud points of biodiesel. Energy Fuels 2008,22(2):747-52.

Mahamuni NN, Adewuyi YG. Fourier transform infrared spectroscopy (FTIR) method to monitor soy biodiesel and soybean oil in transesterification reactions, petrodiesel-biodiesel blends, and blend adulteration with soy oil. Energy Fuels 2009,23(7):3773-82.

Manjunatha R., Narayana P B, Reddy K H C. Application of Artificial Neural Networks for Emission Modelling of Biodiesels for a C.I Engine under Varying Operating Conditions. Modern Applied Science 2010,04(03) 77-89.

Manly BFJ. Multivariate Statistical Methods: a Primer. 3rd ed.: Chapman and Hall, 2005.

Martiniano LC; Gonçalves VRA; Yotsumoto Neto S; Marques EP ;Fonseca TCO; Paim L; Gouveia AS; Stradiotto NR; Aucelio RQ; Cavalcante GHR; Marques EP; Marques ALB. Direct simultaneous determination of Pb(II) and Cu(II) in biodiesel by anodic stripping voltammetry at a mercury-film electrode using microemulsions. Fuel 2013, 103: 1164-67.

Meira M, Quintella CM, Tanajura DSADS, G. HR, Fernando JDS, Neto PRDC, et al. Determination of the oxidation stability of biodiesel and oils by spectrofluorimetry and multivariate calibration. Talanta 2011,85(1):430-4.

Monteiro MR, Ambrozin ARP, Lião LM, Ferreira AG. Critical review on analytical methods for biodiesel characterization. Talanta 2008,77:593-605.

Monyem, A. G. J. V. The effect of biodiesel oxidation on engine performance and emissions. Biomass Bioenerg 2001, 20(4):317-325.

Morsy N, Sun D. Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef. Meat Sci. 2013, 93(2):292-302..

Moser BR. Efficacy of specific gravity as a tool for prediction of biodiesel–petroleum diesel blend ratio. Fuel 2012,99:254-61.

Nadai DV, Simões JB, Gatts CEN, Miranda PCML. Inference of the biodiesel cetane number by multivariate techniques. Fuel 2013,105:325-30.

Najafi G, Ghobadian B, Yusaf TF, Rahimi H. Combustion analysis of a CI engine performance using waste cooking biodiesel fuel with an artificial neural network aid. Am. J. Appl. Sci. 2007,04(10):756-64.

Nia W, Nørgaarda L, Mørup M. Non-linear calibration models for near infrared spectroscopy. Anal. Chim. Acta 2014,813: 1–14.

Oliveira FCC, Brandão CRR, Ramalho HF, Costa LAF, Suarez PAZ, Rubim JC. Adulteration of diesel/biodiesel blends by vegetable oil as determined by Fourier transform (FT) near infrared spectrometry and FT-Raman spectroscopy. Anal. Chim. Acta 2007,587:194-9.

Oliveira JS, Montalvão R, Daher L, Suarez PA, Rubim JC. Determination of methyl ester contents in biodiesel blends by FTIR-ATR and FTNIR spectroscopies. Talanta 2006,69:1278-84.

Ortech Corporation. Operation of a cummins N14 diesel on biodiesel: performance, emissions and durability. 1995.

Öztaş A, Pala M, Özbay E, Kanca E, Çagˇlar N, Bhatti MA. Predicting the compressive strength and slump of high strength concrete using neural network. Constr. Build. Mater. 2006,20(9):769-75.

Pimentel MF, Ribeiro GMGS, Cruz RS, Stragevitch L, Pacheco Filho JGA, Teixeira LSG. Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration. Microchem. J. 2006,82(2):201-6.

Prasad TH, Reddy KHC, Rao MM. Performance and exhaust emissions analysis of a diesel engine using methyl esters of fish oil with artificial neural network aid. J. Eng. Technol. Manage. 2010,02(1):23-7.

Ramadhas AS, Jayaraj S, Muraleedharan C, Padmakumari K. Artificial neural networks used for the prediction of the cetane number of biodiesel. Renew. Energ. 2006,31:2524-33.

Ramos MJ, Fernández CM, Casas A, Rodríguez L, Pérez A. Influence of fatty acid composition of raw materials on biodiesel properties. Bioresource Technol. 2009,100(1):261-8.

Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation, parallel distributed processing Cambridge: MA:MIT Press, 1986;1.

Sablani SS. Status of Observational Models Used in Design and Control of Products and Processes. Compr. Rev. Food Sci. Food Saf. 2008, 130-136.

Saldana DA, Starck L, Mougin P, Rousseau B, Pidol P, Jeuland N, and Creton B. Flash point and cetane number predictions for fuel compounds using quantitative structure property relationship (QSPR) methods. Energy

& Fuels 2011, 25(9): 3900-08.

Saldana DA, Starck L, Mougin P, Rousseau B, Ferrando N, Creton B. Prediction of density and viscosity of biofuel compounds using machine learning methods. Energy Fuels 2012,24(4):2416-26.

Santos AL, Takeuchi RM, Munoz RAA, Angnes L, Stradiotto NR. Electrochemical determination of inorganic contaminants in automotive fuels. Electroanalysis. 2012,24:1681-91.

Sarle WS. Neural Networks and Statistical Models. In Proceedings of the Nineteenth Annual SAS Users Group International Conference 1994,Cary:1-13.

Schalkoff R. Artificial Neural Networks McGraw-Hill: Clemson University, 1997.

Schumacher M, Robner R, Vach W. Neural networks and logistic regression: Part I. Comput. Stat. Data an. 1996,21(6):661-82.

Silva LNDC. Análise e Síntese de Estratégias de Aprendizado para Redes Neurais Artificiais. Universidade Estadual de Campinas. Campinas 1998,250.

Sivaramakrishnan K, Ravikumar P. Determination of cetane number of biodiesel and its influence on physical properties. ARPN Journal of Engineering and Applied Sciences 2012,7(2):205-11.

Sun F, Ma W, Xu L, Zhu Y, Liu L, Peng C, et al. Analytical methods and recent developments in the detection of melamine. Trends Anal. Chem. 2010,29:1239-49.

Syunyaev RZ, Balabin RM. Frequency dependence of oil conductivity at high pressure. J. Dispersion Sci. Technol. 2007,28:419-27.

Syunyaev RZ, Balabin RM. Polarization of fluorescence of asphaltene containing systems. J. Dispersion Sci. Technol. 2008;29:1505-11.

Talebian-Kiakalaieh A, Amin NAS, Zarei A, Noshadi I. Transesterification of waste cooking oil by heteropoly acid (HPA) catalyst: Optimization and kinetic model. Appl. Energy 2013,102:283-92.

Trindade JM.; Martiniano LC, Gonçalves VRA, Gouveia AS, Marques ALB, Baugis GL, Fonseca TCO., Song C, Zhang J, Maques EP. Anodic stripping voltammetry coupled with design of experiments for simultaneous determination of Zn^{+2}, Cu^{+2}, Pb^{+2}, and Cd^{+2} in gasoline. Fuel 2012, 91:26-32.

Tubino M, Aricetti JA. A green potentiometric method for the determination of the iodine number of biodiesel. Fuel 2013,103:1158-63.

Wu FY, Yen KK. Applications of neural network in regression analysis. Comput. Ind. Eng. 1992,23(1-4):93-5.

Wu, W., Dandy, G.C and Maier, H.R. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling Review Article. Environ. Modell. Softw. 2014, 54: 108-27.

Zawadzki A, Shrestha DS. Biodiesel feedstock and blend level sensing using visible light spectra and neural network. Transactions of the ASABE 2009,52(2):539-42.

Zhang WB. Review on analysis of biodiesel with infrared spectroscopy. Renew. Sust. Energ. Rev. 2012,16:6048-58.

doi: 10.14355/ijcsa.2014.0303.02

Towards a Socio-technical Ontology Engineering Methodology

Dana Indra Sensuse*[1], Mesnan Silalahi[2], Indra Budi[3]

Laboratorium e-Government, Fakult of Computer Science, University of Indonesia
Kampus Depok-UI, Indonesia

*[1]dana@cs.ui.ac.id; [2]mesnans@yahoo.com; [3]indra@cs.ui.ac.id

Received 4 December 2013; Accepted 15 January 2014; Published 8 August 2014

Abstract

The development of the semantic web technology has enabled the integration of existing systems and resources. The implementation can be seen such as in the Linked Open Data initiatives, semantic search and semantic web portal. Ontology has been playing the crucial role for the implementation of the semantic web technology, but the development of ontology itself has lagged behind. Methodologies have been developed for ontology engineering, but most of them focused on the technical rather than the socio-technical aspect. Only a few has approached socio-technically but only partially. This article will describe aspects towards a socio-technical approach in ontology engineering methodology more comprehensively. This approach is important because of the complexity and the social nature of the ontology building process. A comparative study will expose the plus and minus points of methods used within the existing methodologies from the socio-technical point of view. Then a workbench will be proposed that take socio-technical aspect within the ontology engineering methodology into account.

Keywords

Ontology; Ontology Engineering; Methodologi; Socio-technical Approach

Introduction

Knowledge is considered the most important asset for organizations and knowledge management has become an important issue nowdays. There are many knowledge based and information systems developed and implemented for the effective and efficient operations of an organisation. One of the problems encountered during the implementation was the difficulty in searching an accurate information and/or the right knowledge. The cause of the problem was identified as the data segmentation within islands of information. In general this problem was caused by the low level of the interoperability. The key to the development of the interoperability is through the development of the ontology. Ontologies have long been argued as one approach for capturing and representing domain knowledge. Currently the information presented on the internet has very limited semantic values and is not machine understandable hence not processable. The data are heterogenous with various structures and formats. To have semantic value the data have to be annotated with ontology which has specifications for the concepts used and relations between the concepts. The development of semantic web has driven the ontology building in many domains. In order to ensure that the ontology will be used widely there must be a consensus within the related parties. And to reach a consensus effectively there is a need in the approach to the ontology enginering methodology that take into account socio-technical aspects. The socio-technial approach is needed because of the complexity in the ontology building process. This paper tries to answer the question and key factors in the socio-technical approach along the ontology lifecycle and how to operationalize the socio-technical approach effectively in the ontology engineering processes.

The socio-technical approach had been inserted into some of the existing methodologies in ontology engineering, but the methodologies have approached it partially. In this paper a methodological model will be proposed with a comprehensive socio-technical approach. The social aspect will be considered in balance with the technical aspect through the phases of the ontology building process, from planning phase to implementation and evaluation phase. This will include the considerations of the social aspects in the methods, techniques and tools of the proposed methodology.

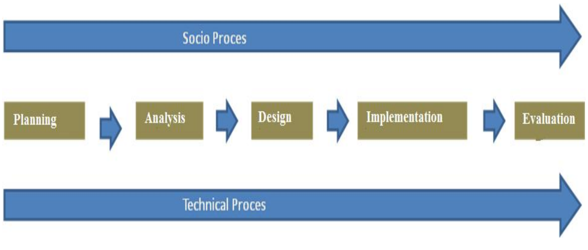

FIG. 1 THE PHASES OF A SOCIO-TECHNICAL APPROACH IN ONTOLOGY ENGINEERING

Semantic Web Application

The semantic web architecture proposed in (Berners Lee, et al, 2001) contains ontology sublayer which plays an important role in semantic web implementation. With the availability of an ontology we can give meaning to the data representation on a webpage. Within the semantic web architecture, the data will be stored in the RDF (Resource Description Framework) files. On top of it, the semantic value (ontology) is given by the OWL (Web Ontology Language) sublayer. OWL is the official standard of W3C used for ontology building with various levels of expressivity.

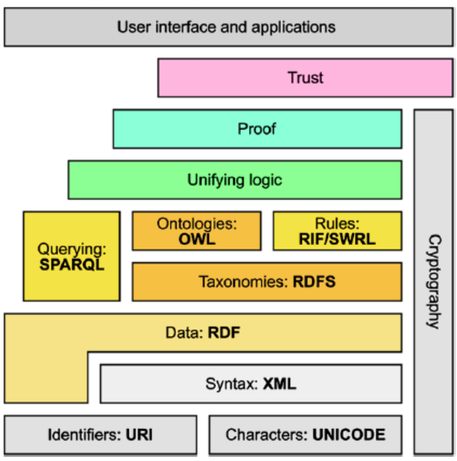

FIG. 2 THE SEMANTIC WEB ARCHITECTURE

In the development of a semantic web application, e.g the semantic search application, there are various ways to integrate data from various source locations. This integration has to be carried out because of the heterogeneity of the data source formats. One of the approach is hybrid ontology integration whereby there will be a local ontology for each data source which represents the schema of the local database. Then there will be a global ontology whereby all the local ontologies mapped into, that will be used as an entry point for designing query strategy-for searching. The semantic search is supported by SPARQL engine as the standard in query syntaxes in semantic web application.

Methodologies in Ontology Engineering

Various methods and methodologies had been proposed in conducting the ontology enginering. These various attempts were carried out in order to make the ontology building process running effectively and efficiently. Informally there are two approaches in the ontology building methods i.e the top-down approach that is heavyweight and mostly is carried out by large parties. The other is a bottom-up approach that is lightweight and usually is carried out by many related parties that later could be mergered. In general, there are three main activities in building an ontology (Tempich et al, 2006): ontology management, ontology development that covers the domain analysis, conceptualization and implementation, and the support and maintenance of the ontology. A state of the art in the ontology methodology such as the METHONTOLOGY (Gómez-Pérez et al., 2003) describes the ontology building processes, activites, and tasks in details. This methodology has been known widely but not the use. ONIONS (ONtologic Integration Of Naive Sources) is a methodology that describes various methods to integrate sources of information (ontology network) with the emphasis on the domain ontology. OTK (On-To-Knowledge) (Staab et al., 2001) was developed with the approach using the common sense found on other methodologies. DILIGENT (Tempich et al., 2004) build an argumentation framework in modeling ontology engineering supported by wiki based tools. UPON (United Process for ONtologies) uses a UML base approach with use cases and has incremental and iterative characteristics. Below is a list of the plus and minus points of some important methodologies.

A Socio-Technical Approach

According to Studer, et .al (1998) ontology is defined as: “A formal explicit specification of a shared conceptualization of a domain of interest”. From that definition we can infer the characteristics of an ontology as formal, explicit, and conceptual specification of a domain. We can thus say that an ontology has to be built based on a consensus of the stakeholders and the related parties and is implicitly meant the wide acceptability. The true value of the ontology itself is, when it is shared and used publicly. In fact the development of an ontology is more and

TABEL 1 MINUS AND PLUS POINTS OF SOME EXISTING METHODOLOGIES.

Methodology name & Author	Plus points	Minus Points
• OTK • Staab et al.	• Cover the whole lifecycle • Use of ontology pattern • Contains strict rules for update/insert/delete ontology • Integrate participants early in the process to identify use case and competency questions • Rather detailed in building ontology from scratch • Propose ontology learning	• Too formal for small scale application. • Focus depend on the application (enterprise ontology) • Not address the collaboration aspect. • Target User are ontology engineers
• Enterprise Ontology • Uschold & King (1995)	• Application independent • Middle-out strategy in identifiying concept.	• Not proposing a feasibility study or prototyping. • No clear description on the techniques and activities. • Not proposing a lifecycle.
• TOVE • Grüninger & Fox (1995)	• Proposes identifying intuitively the main motivating scenarios	• No clear division in the phases. • The activities as well as the techniques used not described into details. • No selection proces in life-cicle model • Domain limited in business.
• METHONTOLOGY • Gomez Perez et al.	• Very detailed process • Application independent. • Adoption of middle-out strategy. • An evoluted prototype. • A mature methodology • Contain whole lifecycle • Has an integrating part. • Based on IEEE standard for software development.	• Details in predevelopment proces not sufficient • Too complex for small scale application • Ontology reuse in the end development phase. • lifecycle model does not respect any usage-oriented aspects • Not address collaboration. • Target users are ontology engineers
• SENSUS	• Addressing the distributive setting • Ontology Reuse	• A top-down approach • Not mention the lifecycle • Not too detailed • No particular technique in details.
• Ontology 101 • Noy & McGuinness	• Address the naming convension • Based on best practices • Iterative process • Has an integrating part in begin phase	• Has some missing part of the whole lifecycle (evaluation and implementation)
• HOLSAPPLE • Holsapple & Joshi	• Collaborative building process. • Focus on reaching consensus. • Iterative prosess	• Only specify the effort on reaching a consensus
• DILIGENT • Tempich et al.	• Focus on the collaboration process • Based on the good argumentation to reach consensus. • Iterative process • Distributive environment setting • The approach focused on the human • Targeted for domain experts and end-user • Start using the ontology early in the process • A centralised approach by introducing a control board	• No proposal for activity in the specification of the ontology requirements • No proposal for reusing and reengineering knowledge resouces to speed the process. • The balance of the controll board not addressed.
• HCOME	• A distributive environment • The approach focused on the human	• The supporting tool is not clear
• DOGMA • Mustafa Jarrar, Robert Meersman	• Separate clearly the base facts in domain (*lexons*) from constaints, rules, identification, derivation to imporve reuse • A collaborative approach	• Centralised approach.

more a complex social process rather than a technical design activity (Kings & Davies, 2009). Mika (2005) stated that the creation and maintenance of ontology is purely a social activity. To understand the social existence is crucial in understanding how the ontology evolutes and accepted. Thus there are still a lot of things to investigate in ontology engineering, such as the socio-technical context in order to develop a methodology that will enhance the successful implementation of the developed ontology. The low level of success in system implementation and the use of ontology in particular, has driven the development of the socio-technical model in the information system in general and in the field of ontology engineering in particular. In general, this approach is aimed to a greater acceptability and to smoothen the consensus building. According to Cherns (1987) there are ten principles to the socio-technical approach in system development, i.e: compactibility, minimum critical specification, variance control, location boundary, information flow, power and authority, multifunctional, support coungruency, transitional organisation and incompleteness principle. These principles can be taken into account for modeling the ontology development. In the field of information system development, Mumford (1983) developed an ETHICS model to look for a socio-technical solution of problems in the implementation of an information system. This model focused on the participatory mechanism of the stakeholders. This approach ensures the feedback that will expose the user requirements which eventually ensure the acceptability. Below is the list of some methodology with the focus on–technical aspect aiming to automating the ontology building process and the partially socio-technical approach that focused on some aspect of a social approach to enhance the success of the ontology building process.

TABEL 2 SOCIO-TECHNICAL COMPARISON ON EXISTING METHODOLOGY

Focus	Name	Characteristics	Ontology	Central Actor
Technical				
Merging/ Integration	Pinto et al, KACTUS, ONION			Knowledge Engineer
Layered (n-tier)	DOGMA	The split between the static and the dynamic parts	Centralized	Knowledge Engineer
Learning	Ontolearn			
Rapid Development	ROD			Knowledge Engineer
Comprehensive	METHONTOLOGY	Based on SE methodology, very detailed	Static Ontology	Knowledge Engineer
Common Sense	OTK			Knowledge Engineer
	UPON	based on UP and UML	Large Scale Domain Ontology	Ontology Engineer
Partly Socio-technical				
Argumentation Based	DILIGENT	Distributive setting, More Role for Domain Expert	Dynamic Domain Ontology	Control Board in Consensus Building
Human Centered	HCOME/HCONE	Nominal Group Technigue	Decentralized	Knowledge Worker
Consensus	HOLSAPPLE	First Initial Ontology developed further with Delphi Method	Static Ontology, centralized	Knowledge Engineer
Inter-organisational	DOGMA-MESS	Database like inpsired architecture	Dynamic	Domain Expert

Research Methodology

The proposed methodological model will approach the ontology engineering from the point view of socio-technical design. The model will take into account the existing best practices in software engineering as well as in ontology engineering. Fig.2 describes the proposed components for input in the development of the methodology. The methods that will be applied in the design model will take into account concepts in meta-design, concepts in virtual workgroup collaboration, and the argumentation model in consensus building. In the planning phase account will be taken on how to handle the seeds of the ontology design such as the goal and intention of the ontology, and the scoping and the user requirement elicitation in complex system.

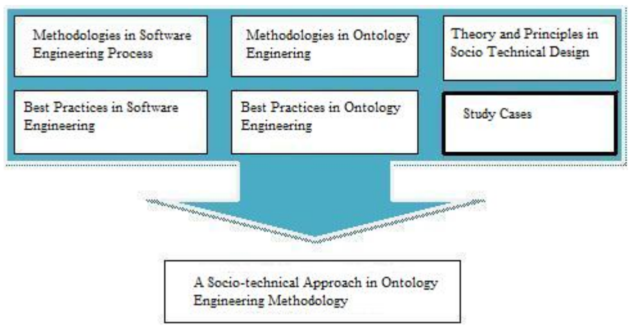

FIG 3. INPUT FOR THE METHODOLOGY DEVELOPMENT

The tool is web based which is build for supporting the socio-technical approach that has the functionalities for: the capabilities for supporting collaboration with a user-friendliness in inserting people to participation, contribute, put arguments and voting (or some other conflict resolution mechanism) and/or for just to put comments. Furthermore the communication channel will be built on various methods such as mailing list, blogs, chatting, short messages, or a functionality to shouting. More importantly is the integration with tools for the core functionalities such as ontology editing, ontology browsing and ontology visualization. Furthermore accounts have to be taken on the supporting features such as documentation and tracking ontology versions history.

Below is the list of some of the tools that are reviewed before integrating into the proposed workbench. The tools selected will be integrated within the portal. The portal is based on a CMS software that supports the semantic web application (such as Drupal) and has capabilities for extention and adaptability such as integration with the more powerful existing tools for ontology editing, ontology browsing and ontology visualization.

FIG. 6 REVIEWED TOOLS FOR THE SOCIO-TECHNICAL ONTOLOGY ENGINEERING

Conclusions

The socio-technical approach has been proposed to be comprehensively used as a basis in developing a methodology for the ontology engineering. This methodology is proposed based on the partially socio-technical approaches that have been tempted in order to increase acceptability of the new system in general and the ontology in particular. This will eventually boost the wider use of the ontology and eventually can contribute to the higher interoperability of the exsting systems and data sources. By including the methods, tasks, and tools that have the socio-tehnical characteristics from past results in ontology engineering methodology development, it will add socio-technical value of the new proposed methodology. This methodology has a modular property in which the new methods, tasks or tools can later be added to fully covered the whole range of the ontology engineering process with the socio-technical value.

ACKNOWLEDGMENT

We thank the Ministry of Education of Republic of Indonesia for supporting this research within the scheme of BOPTN 2013.

REFERENCES

A. De Nicola, M. Missikoff, R. Navigli. A proposal for a Unified Process for ONtology building: UPON. In Proceedings of 16th International Conference on Database and Ex-pert Systems Applications (DEXA) (2005).

C. Tempich, H. S. Pinto, Y. Sure, and S. Staab. An Argumentation Ontology for Distributed, Loosely-controlled and Evolving Engineering processes of ontologies (DILIGENT). In ESWC 2005, pages 241–256, 2005.

Cherns, A., 1987. Principles of sociotechnical design revisited. Human Relations, 40, 3, 153–162.

De Moor, A., De Leenheer, P., and Meersman, R. (2006) DOGMA-MESS: A Meaning Evolution Support System for Interorganizational Ontology Engineering. In Proc. of the 14th Int'l Conference on Conceptual Structures (ICCS 2006) (Aalborg, Denmark), LNAI 4068, Springer, pp. 189-203.

Elena Paslaru Bontas Simperl and Christoph Tempich, 2006. “Ontology Engineering: A Reality Check,” in Proceedings of the 5th International Conference on Ontologies, Databases, and Applications of Semantics ODBASE 2006.

Elio Toppano, A Communication-based Model of Ontology Design and (Re)Use, International Conference on Intelligent Semantic Web-Services and Applications (ISWSA 2010) , June 14–16, 2010, Amman, Jordan.

Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2003) “Ontological Engineering”. Springer Verlag. Advanced Information and Knowledge Processing series. ISBN 1-85233-551-3. November 2003.

Gruber TR (1993a) A translation approach to portable ontology specification. Knowledge Acquisition 5(2):199–220.

Gruninger M. and Fox M. S. (1995) Methodology for the Design and Evaluation of Ontologies, IJCAI Workshop on Basic Ontological in Knowledge Sharing, Montreal, Canada.

H. S. Pinto, S. Staab, and C. Tempich. DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolving Engineering of ontologies. In R. L. de Mántaras and L. Saitta, editors, ECAI, pages 393–397. IOS Press, 2004.

Jarrar, M., Meersman, R.: Formal Ontology Engineering in the DOGMA Approach. In: On the Move to Meaningful Internet Systems: CoopIS, DOA, and ODBASE, LNCS, Springer Verlag, pp. 1238-1254 (2002).

Kings, N. J., Gale, C. & Davies, J., (2007) Knowledge Sharing on the Semantic Web. In Franconi, E., Kifer, M. & May, W. (Eds.) European Semantic Web Conference (ESWC) 2007. LNCS 4519 ed. Innsbruck, Springer-Verlag.

Klein, M., Noy.: A component-based framework for ontology evolution. Technical Report IR-504, Department of Computer Science, Vrije Universiteit Amsterdam. March (2003).

Mumford, E. (1983). Designing human systems for new technology: The ETHICS method, Manchester, UK: Manchester Business School.

Noy, N. F., Chugh, A., and Alani, H. The CKC Challenge: Exploring Tools for Collaborative Knowledge Construction, IEEE Intelligent Systems 23 (1) 2008.

P. De Leenheer, Towards community-based ontology evolution, PhD Thesis, Vrije Universiteit Brussel, Brussels, Belgium, 2009.

Peter Mika: Ontologies Are Us: A Unified Model of Social Networks and Semantics. International Semantic Web Conference 2005: 522-536.

Rahwan, I., Ramchurn, S.D., Jennings, N.R., McBurney, P., Parsons, S., Sonenberg, L. 2004. Argumentation-based negotiation. The Knowledge Engineering Review, Vol. 18 no. 4, 343-375.

Simone Braun et al. Ontology Maturing: a Collaborative Web 2.0 Approach to Ontology Engineering, 2007.

Smith, B. "Ontology", (2003), in L. Floridi (ed.), Blackwell Guide to the Philosophy of Computing and Information, Oxford: Blackwell, 155–166.

Staab, S., Schnurr, H.P., Studer, R., Sure, Y. (2001) "Knowledge Processes and Ontologies". IEEE Intelligent Systems 16 (1):26–34.

Stephan Bloehdorn, Peter Haase, Zhisheng Huang et al. (2008) Ontology Management, 3-20. In Semantic Knowledge Management.

Studer R, Benjamins VR, Fensel D (1998) Knowledge Engineering: Principles and Methods. IEEE Transactions on Data and Knowledge Engineering 25 (1-2):161–197.

Tharam Dillon, Elizabeth Chang, Maja Hadzic, Pornpit Wongthongtham (2008) "Differentiating Conceptual Modelling from Data Modelling, Knowledge Modelling and Ontology Modelling and a Notation for Ontology Modelling", Wollongong, Autralia.

Tim Berners-Lee, James Hendler & Ora Lassila: "The Semantic Web", Scientific American 284 (5):34-43 (May 2001).

Uschold, M and Gruninger, M. (1996), Ontologies: Principles, Methods and Applications, Knowledge Engineering Review, vol. 11, pp. 96-137.

Uschold, M. King, M. Towards a Methodology for Building Ontologies. Workshop on Basic Ontological Issues in Knowledge Sharing (1995).

Wendy Olphert, Leela Damodaran, Citizen Participation and engagement in the Design of e-Government Services:The Missing Link in Effective ICT Design and Delivery, Journal of the Asociation of Information System, Volume 8, Issue 9, Article 4, pp. 491-507, September 2007.

Dana I. Sensuse received his B.Sc in Soil Science (Bogor Agricultural University, Indonesia, 1985), M.LIS from Dalhousie University (Canada, 1994), and Ph.D in Information Studies from University of Toronto (Canada, 2004). Now he works as a teaching staff in Faculty of Computer Science, University of Indonesia, and Head of e-Government Laboratory. His research interests include E-Government, Knowledge Management, and Information Systems.

Mesnan Silalahi received his M.Sc in Materials Science (TU Delft, Netherland, 1994) and Magister in Information Technology from University of Indonesia (Indonesia, 2007). He is now taking his Ph.D in Computer Science (University of Indonesia). He works as a researcher at the Indonesian Institute of Sciences. His research interest includes Ontology Engineering, Semantic Web and Knowledge Management.

Indra Budi received the Doctoral degree in Computer Science in 2008 from University of Indonesia. Currently, he is an academic staff at the University of Indonesia. His

research interest includes Information Extraction, Natural Language Processing, Named Entity Recognition and Information Retrieval. He was the first PIC of Content Development Activity, Program B Fasilkom UI, Grant from DIKTI in 2004 and involved in Program CISA Review Course in 2004.

doi: 10.14355/ijcsa.2014.0303.03

The Research of Single_Node Risk Spread in Supply Chain Complex Network Based on Fixed Risk Values

Lei Wen[a], Yachao Shi[b]

Department of Economics and Management, North China Electric Power University,

Baoding Hebei Province, 071003, China

[a]ncepuwxzj@163.com; [b]syc_0319@163.com

Received 3 December 2013; Accepted 15 January 2014; Published 8 August 2014

Abstract

Complex networks cannot only describe the complicated, asynchronous system, but also can be used to model and analyze network topological properties. Sophisticated methods for network construction and analysis exist in other fields. But until recently, researchers have few focused on the risk spread of a supply chain network. In this paper, a supply chain risk network based on single_node risk spread is modeled and the static network statistics are analyzed, including degree distribution, risk distribution, average path length and clustering coefficient. The simulation results indicate that supply chain complex network is a small-world network with short average path length and high degree of clustering, and its degree distribution and risk distribution follow a double power law. In addition, the average risk tends to decrease with the total number of risk node increase in supply chain risk spread network.

Keyword

Supply Chain; Complex Network; Single_node Risk Spread; Modelling and Simulation

Introduction

Supply chain is an integrated network where suppliers, sellers and customers are highly interconnected through material/product flows, information flows, and financial flows. In supply chain, it has many tiers which provide goods or services to the next similar level tier. Moreover, each tier may have multiple components or members, and the flow of goods is linear.

With the rapid development of economic globalization, various factors of uncertainties create risks which can impact the proper functioning of the enterprise such as procurement risk, production risk, financial risk, marketing risk and outsourcing risk. As a kind of effective mode in the market competition, the supply chain has already been accepted by a lot of enterprises. Due to its structural characteristics, there exists great risk. As a potential threat, supply chain risk may take advantage of its vulnerability to destruct the supply chain system, and finally bring damage and loss for the entire supply chain upstream and downstream enterprises. Therefore, risk is the main problem of supply chain enterprises.

Complex network describes a wide range of systems in nature and society. It is an important tool to describe and analyze the changes between each factor. The relationship of supply chain members provides a channel to transmit the risk variations of suppliers and customers and therefore the liquidity risks are rampant along supply chains. A lot is infected if one object in it is infected. To describe this phenomenon, we employ computer software to analyse the laws of risk spread evolution based on single_node risk.

The rest of this paper is organized as follows: In the next section, there is a brief literature review relating supply chain risk and complex network risk research. In Section 3, we design a supply chain risk network and describe the risk spread based on single_node risk and its connection mechanism. In section 4, some basic concepts of the complex network are introduced in this paper, such as degree distribution, risk distribution, average path length, and clustering coefficient. In Section 5, a simulate experiment of single_node risk spread to observe indicators changes is showed. Finally, we offer some concluding remarks in Section 6.

Literature Review

Supply Chain Risk

Supply chain is an integrated network wherein a number of various business entities work together in an effort to acquire raw materials, convert into final products, and deliver them to customers. As a complex system, it is especially vulnerable to a variety of risks. Risk sometimes interpreted as the probability of loss due to danger, damage, loss, injury or any other unreliable and uncertain consequences. In view of its great significance, there has been increased attention focused on supply chain risk from various directions.

Dina Neigera, Kristian Rotarua and Leonid Churilovb put forward a novel value-focused process engineering methodology for process-based supply chain risk identification. Chaudhuri A, Mohanty B K and Singh K N propose a group decision making approach using numeric and linguistic data in supply chain risk assessment. Cigdem Sofyalıoglu and Burak KartalFuzzy determine risk management strategies by the method of the analytic hierarchy process. Walid Klibia and Alain Martela develop a risk modeling which can facilitates the evaluation and the design of supply chain networks operating under uncertainty.

However, the risk is random, independent, and spread along the upstream to downstream node enterprise. Recently, many researches pay more attention to the spread in supply chain.

Fabrizio Natalea et al build a movement network to simulate epidemic spread. They analyze degree, in-degree and out-degree distribution of the nodes and show average number of infected nodes. Monique A. van der Gaaga Fred Vosa design a detailed stochastic state transition model and analyze its characteristics. Mathilde Paula et al focus on the spread process of HPAI H5N1 and provide a stochastic state-transition simulation model which described the spread of Salmonella from multiplying through slaughter. Houyuan Ye point out that the risk spread are divided into five types: bubble burst, elements scarce, collapse of structure, tsunami and chain reaction. Guoping Cheng and Qin Liu research on the occurred conditions of close coupling and the changes in the amount of risk caused by close coupling.

In summary, the research of the risk spread is still little. Most scholars focused on the risk spread of supply chain in a certain respect. It is an undeniable fact that it has a certain reference value to the similar enterprises, but the application range is parochialism. In this paper, we analyse the rules of risk spread evolution in the entire supply chain, through a definite transmission routes.

Complex Network Risk

Networks provide a new angle of view to express relationships between constituent elements how to change. Graphs can be used to create a visual image of network interior movements.

Disease remained an overarching concern for all scientists. To use complex network is one of the effective methods to study the spread of disease. The two most common models are SIR and SIS in a wide variety of disease transmission model. However, each of these has particular features, so corresponding extended models arise.

Mathieu Moslonka-Lefebvrea, Marco Pautassoc and Mike J. Jegerc establish in directed networks at four levels of competence in six network structures (local, small-world, random, one-way, uncorrelated, and two-way scale-free networks). Mario Ventresca and Dionne Aleman estimate contagion spread using social network characteristics. They study the effect of the six mitigation strategies on the degree distribution and local clustering coefficients. D. Mayera, J. Reiczigela and F. Rubela simulate the airborne spread of foot-and-mouth disease virus using Lagrangian particle model which take into account the realistic orographic and meteorological conditions.

Like the disease spread, the information spread is the object of researchers' focus either, especially in the epoch of information. M. Nekoveea and Y. Morenob et al [25] introduce a general stochastic model for the spread of rumors, and use some equations to analyze the threshold behavior and dynamics of the model on several models. Laijun Zhao, Hongxin Cui et al redesign a flow chart of the rumor spreading process with the SIR model, and simulate the process of rumor spread in the new media age. Laijun Zhao, Qin Wang et al study a rumor spreading model called LiveJourna which provide a more detailed and realistic description of rumor spreading process with combination of forgetting mechanism and the SIR model of epidemics.

The network conceptual framework and graph-theory-analytic framework have been very useful for the description and analysis of the disease and rumour spread. This approach could be applied in many directions. Such the mature approach provides reference experience for the spread of supply chain risk.

The Supply Chain Risk Model

Basic Element

The network conceptual framework has been very useful for the description and analysis of the risk of transmission. Now, we will present the basic definitions of the components of risk network:

i) Risk source: risk source is the node enterprise which generates risk. The origin of risk by classification can be divided into external risks which caused by external environment and internal risk which caused by internal operations.

ii) Risk nodes: in supply chain network, each risk enterprise can be called as risk node. According to the difference of risk source, risk nodes can be divided into conduction node and risk receiver.

iii) Risk carriers: risk carriers are "mediums" or "bridges" which can transmit risk from one node to another. For different properties of risk carriers and risk source attribute, risks have had transmitted dependent on different medium.

iv) Conduction pathways: it can be defined as the logical distance of risk transmitted from one node to another. The supply chain risk conduction path is based on "Supplier --- Manufacturer --- Dealers --- Consumers".

v) Risk threshold: the risk threshold is a value that determines whether the node enterprises can normally transmit material/product, information and finance. In the supply chain network, each enterprise itself has a constant risk threshold value regardless of how much it received risk.

The Risk Supply Chain Network Model

1) Supply Chain Network

To observe the rules of risk spread in a network, a network is generated randomly. A weighted network is represented as a graph A=(V,E) with a set V of n node enterprises connected by M edges which are randomly chosen from the n(n-1)/2 possible edges. New nodes, which connected to m existing nodes with a certain probability, are added sequentially until the network reaches the required size.

It is worthy of note that the risks are spread along with the material/product flows, information flows, and financial flows. That means there are business relationships among each node enterprise. One risk node enterprise will establish business relation with others. One randomly chosen node was initially designated as a source risk. At each time step, each node was considered in a fixed order. The spread of risks begins in suppliers, passing sellers and finally deliver to customers. If exceed risk threshold, the node could potentially infect any or all of its neighbours, with a given probability. Once infected risk, a part of risk node enterprise will control them, and others continue to repeat the process.

2) Definition of Node

In the supply chain network, each node is assumed to be a risk enterprise, and each edge was assumed to be a method of spread of the risk from one risk node enterprise to another. Each business entity is represented by a node. The node of supply chain complex network is defined as vi= (ID, Category, Level), ID is the number of node in supply chain. Category is the label of node; the nodes are divided into three categories as suppliers, sellers and customers. The level of node shows the business order in supply chain, the customer's level is bigger than seller's and the seller's level is bigger than the supplier's.

Generation Process

1) The Selection of Local World

According to the categories and level of nodes, there are three ways to select:

a) When the new added node is supplier, it will establish business relation with other suppliers or sellers, so all the suppliers and sellers are selected as its local world.

b) When the new added node is seller, it will establish business relation with suppliers; customers or other sellers, so all nodes of the network are selected as its local world.

c) When the new added node is customer, it will establish business relation with sellers, so all the sellers are selected as its local world.

2) Connection Mechanism

At each step of the evolution, generate a number of edges and connect them together by adding new node, or, to put it another way, a new added node can establish business relation with k nodes by immanent probability.

In this paper, one must take into consideration that the node's risk value may affect its probability proportional of node connection. A new node connects to existed node within its local world by probability $p(k) = \frac{\prod_{l \in A}^{m} r(l) - r(k)}{(m-1) * \prod_{l \in A}^{m} r(l)}$.

Here, m is the total number of risk nodes, r(k) is the risk value of new node k, r(l) is the risk value of all its associated node. A is the ultimate supply chain network.

Modelling Process

The supply chain network model based on risk connection is generated by the following algorithm:

1) Start with a small number m0 of nodes and confirm the risk of node, the edges, direction and its risk spread probability.

2) A new coming node enters into the supply chain network, select its local world from the existing network.

3) The new coming nodes connected with existed nodes by probability p (k).

4) Repeat 2), 3) until the number of nodes in the network reaches to N.

Statistical Characteristics for the Parameters

Degree

The degree which be described by k_i is one of the most important topology attributes of complex network. k_i is the number of nearest neighbours a node. It has two components: in-degree k_{in} (the number of incoming edges) and out-degree k_{out} (the number of outgoing edges). In supply chain risk network, k_{in} means the supply channel and k_{out} means the sale channel.

The distribution function P (k) is the probability that a node has degree k. It describes the network's topology properties and consequently can expound the mechanisms of structure's evolution. $P(k)_{in}$ is the probability that a node has k incoming edges directed to itself (probability of supply chain from k supply channel), $P(k)_{out}$ is the probability that a node has a total of k edges to other nodes (probability of supply chain to k sale channel). Most networks have degree distributions, which have a power law tail P (k) ~$k^{-\theta}$.

Risk

Risk sometimes also known as vulnerability, disruptions or disturbances. All those can have negative impacts. Risk probability interpreted as a chance of risk emerging. It is a measurement data.

The change of risk probability obeys three distributions: normal distribution, triangular distribution and β distribution. The difference between the three applications is the variable property. Normal distribution represents the variation of general economic variables, the change of random variable is described by β distribution and triangular distribution can indicate any of them.

Any type of risk is uncertain. Before analyze them, we should make a general judgment to significance and the most intuitive approach is to compare the size of risk probability. According to the change of risk personality, variation tendency can be seen on the whole network.

Average Path Length

The average effective path length L_s of the network represents the minimum distance between a lowest-level supplier and a customer. It's defined to be the average value: $L_s = \frac{1}{c}\sum_{i \neq j} d_{ij}$

Here, d_{ij} is the average shortest effective path length from lowest-level supplier (k_{in}=0) to customer, c is the number of links between them.

From the above formula, it can be seen clearly that with c decrease L_s increase. L_s indicates the network's performance. The smaller L_s is, the shortest reaction time there are. In supply chain risk network, the flow of product, information or financial will be interrupted under the influence of risk. They may not have an opportunity to pass down and then choose another path. Therefore, its shortest path length is changed. In this paper, the average path length is actually minimum distance in sub-supply chain.

Clustering Coefficient

In supply chain network, a positive clustering coefficient indicates that nodes with similar degrees tend to be connected with each other closer. On average, the likelihood of a node being infected by risk increased with a positive clustering coefficient.

The Clustering coefficient C_i is defined as the proportion of number of links shared by the neighbor node, which is adjacent to each other.

$$C_i = \frac{1}{k_i - 1} \sum_{j,k} \frac{w_{ij} + w_{jk}}{2} \left(a_{ij} + a_{jk} + a_{ki} \right)$$

Where k_i is the degree of node i; w is the weight of edges between two nodes, including w_{ij} and w_{ik}; $a_{ij}a_{ik}a_{ki}$ respect there is an edge between nodes i, j and k. The default value is 1 if there is a connection between two nodes, otherwise it is 0.

Average clustering coefficient C measures the prevalence of local connections in supply chain risk network. It is defined by:

$$C = \frac{1}{N} \sum C_{\mathrm{i}}$$

Under the influence of the risk, the effective clustering coefficient will wave violently, and the value of average clustering coefficient perhaps not by much.

Simulation Experiment of Single_Node Risk Spread

Preferences

The enactment of parameters in supply chain risk network model is generated by the following algorithms:

a) Set up 30 nodes in the initial network, and the number of suppliers, sellers and customers' edge is set as 3, 5, 3,respectively;

b) Set up a threshold for every node at random;

c) Suddenly increase the risk value of a single node over 50%;

d) Perform the simulation with the Initial Parameters. Each experiment is performed 50 times and took the average;

Specific parameters will be described in TABLE 1.

TABLE 1 THE SUPPLY COMPLEX NETWORK EXPERIMENTAL DATA

Experiment nunumber number	Suppliers	Seller	Customers
1	120	200	320
2	180	300	480
3	240	400	640
4	300	500	800
5	360	600	960
6	420	700	1120
7	480	800	1280

Degree Distribution

FIG. 1 is the double logarithmic distribution graphs of supply chain risk network node degree. We can see that most of node enterprises degree is relatively small, yet a minority of the degree of node enterprises is very big.

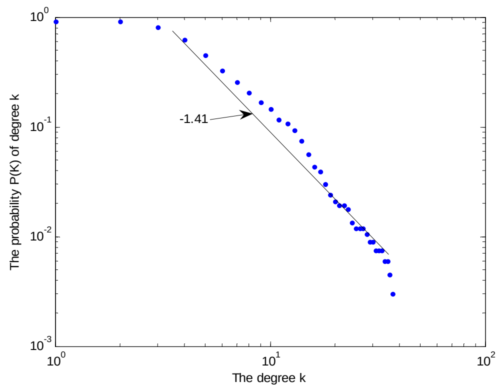

FIG. 1 THE DEGREE DISTRIBUTION OF SUPPLY CHAIN NETWORK BASED SINGLE_NODE RISK

In FIG. 1, we find that the degree distribution of supply chain network follows a power law distribution. The power-law exponent of all nodes degree distribution $\gamma_{\mathrm{degree}} = -1.41 \pm 0.05$. That means the most business entities have relatively less business relation with other entities, but few have a large number of business relation with other entities.

In supply chain risk network, the degree of a node defined to show its ability and prominence. The larger degree, as a supplier, means it has power supply capacity; as a seller, means it has more supply channel and sales leads; as a customer, means it has more purchase funnels and stronger purchase capitally. Above all, the larger the degree of a node and the closer it connect with others. The node with larger degree has more opportunity to build business relation with new added enterprise and finally become "core enterprise".

Risk Distribution

We can see that the node enterprise risk distribution obeys the power-law distribution which manifest in FIG. 2. It shows that the distribution of the ''risk'' of the nodes of risk movement network is heavy-tailed. Most of node enterprises degree is relatively small.

In FIG. 2, we find that the risk distribution of supply chain network also follows a power law distribution. The power-law exponent of all nodes risk distribution $\gamma_{risk} = -1.44 \pm 0.05$. That means the most business entities have relatively lower risk, but few have a large risk. The node with large risk will cause the unstable of supply chain.

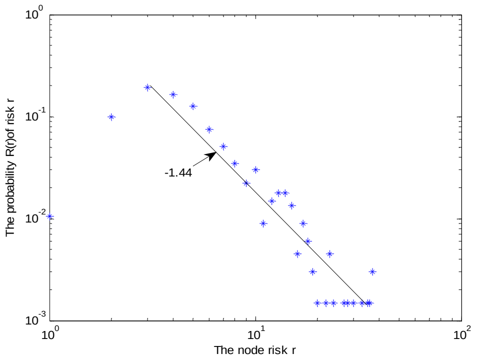

FIG. 2 THE RISK DISTRIBUTION OF SUPPLY CHAIN NETWORK BASED SINGLE_NODE RISK

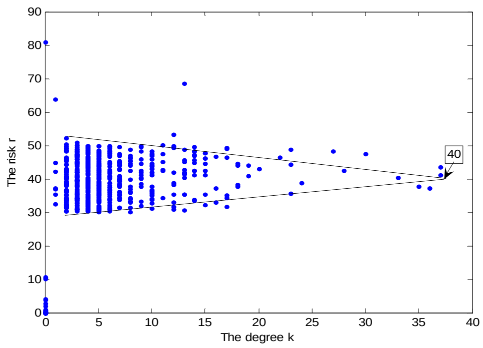

FIG. 3 THE RELATIONSHIP BETWEEN DEGREE AND RISK

From FIG. 3, we can find that there is no distinctive linear relationship between degree and risk. It reflects two sides: (1)the most business entities have relatively lower risk and its risk value trend fluctuates above and below 40, but few have a large risk; (2)the most business entities have relatively lower degree (k<20), but few have a large degree (k>20). That means the degree and risk distribution follow a power law distribution.

The Scale of Node Risk

From FIG. 3, we can find that the number of risk nodes is increasing, and the total numbers of risk nodes are of small proportion from 0.14% to 7.9%. That means risk will travel along conduction carriers to other enterprise nodes.

From the FIG. 4, we can easily come to the conclusion that the number of risk nodes is increasing first, then decreasing and finally remain unchanged in the same time period.

At the beginning of the risk spread of supply chain, most of the associated risk nodes are especially vulnerable to the single risk node. A supply chain network with a small scale will exhibit fast risk spread. However, each node enterprise has a certain risk thresholds. The growing risk nodes cannot be increased without limitation. At each step, a proportion of risk will be control until they complete be dissolved. So overall, the total numbers of risk nodes are increasing and finally tended to be stable.

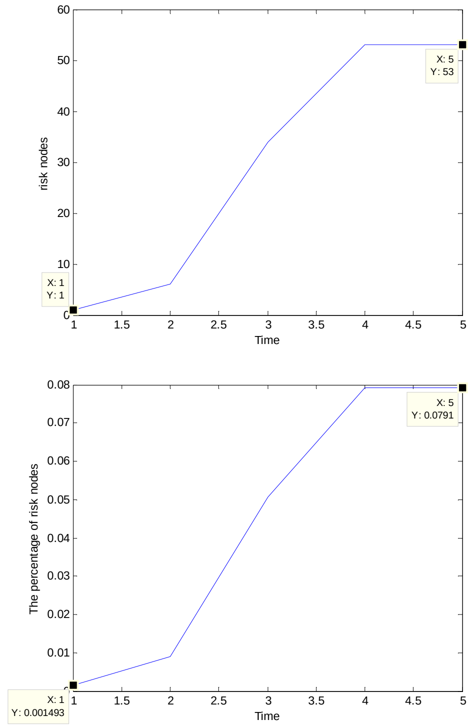

FIG. 4 THE NUMBER OF RISK NODE AND ITS PERCENTAGE BASED SINGLE_NODE RISK

Average Risk

FIG. 5 shows the tendency of average risk. Average risk increases with time and its slope performs increasing previously and decreasing later.

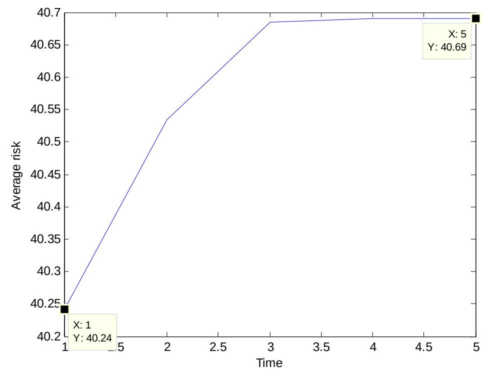

FIG. 5 THE AVERAGE RISK IN SUPPLY CHAIN NETWORK BASED SINGLE_NODE RISK

From FIG. 5, we can find that the size of average risk is decreased from 40.24 to 40.69. That means a part of risk will be control by them, and others continue to attack their associated risk nodes. The invasive risk will cause the unstable of supply chain.

Average Path Length

The path length indicates the typical number of intermediate links between any two randomly chosen vertices. From TABLE 2, we can find the average supply chain path length is nearly to 4.2541 with scale of supply chain increasing. We can see that while the network scale grows larger, the average effective path length tend to become smaller, which means that the risk node enterprises have faster reflect and shorter response time.

TABLE 2 THE SUPPLY CHAIN RISK NETWORK AVERAGE EFFECTIVE PATH LENGTH CALCULATED RESULTS

Experiment number	average path length path length
1	3.4423
2	3.6452
3	3.8859
4	3.954
5	4.092
6	4.1005
7	4.2541

Average Clustering Coefficient

Clustering has been shown to be a network feature relevant to risk spread. From the network structure, different scales of supply chain have different average clustering coefficient. With the scale of supply chain increased, the average clustering coefficient C is nearly to 0.0015. The result show the supply chain has a low average clustering coefficient C which manifest TABLE 3.

TABLE 3 THE RISK SUPPLY CHAIN NETWORK CLUSTERING COEFFICIENT CALCULATED RESULTS

Experiment number	average clustering coefficient
1	0.0113
2	0.0097
3	0.0065
4	0.0057
5	0.0054
6	0.0017
7	0.0015

Conclusions

Networks are not only relevant for many current issues but have found application in a variety of supply chain systems. In this paper, we used a computer model to simulate the risk spread based on single_node risk in supply chain network. By means of analytical statistic characteristics and numerical results, we show that: (1) Degree distribution and risk distribution has the same variation rule that all of them follow a double power law; (2) There is a positive correlation between the number of risk node and average risk. (3) This supply chain risk network is a small-world network which characterized by a short average path length and high average clustering. The interaction between the statistical properties of the network and the results of risk spread is of certain reference value for similar network.

ACKNOWLEDGMENTS

This paper is supported by Humanity and Social Science Youth foundation of Ministry of Education under Grant Nos. 10YJC630271.

REFERENCES

Benita M Beamon. Supply chain design and analysis::: Models and methods [J]. International journal of production economics, 55 (1998) .pp.281-294.

Chaudhuri A, Mohanty B K, Singh K N. Supply chain risk assessment during new product development: a group decision making approach using numeric and linguistic data [J]. International Journal of Production Research, 51 (2013) .pp. 2790-2804.

Douglas M Lambert, Martha C Cooper. Issues in Supply Chain Management [J]. Industrial Marketing Management, 29(2000). pp. 65-83.

Fabrizio Natalea et al.Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread [J]. Preventive Veterinary Medicine, 4 (2009). pp. 341-350.

Guoping Cheng and Qin Liu. The risk management of supply chaint based on risk conduction [J].logistics engineering and management. 1 (2009). pp .22.

Harland C, Brenchley R, Walker H. Risk in supply networks [J]. Journal of Purchasing and Supply management, 9 (2003) .pp. 51-62.

Hokey Mina, Gengui Zhoub. Supply chain modeling: past, present and future [J]. Computers & Industrial Engineering, 43(2002).pp.231-249.

Houyuan Ye. The particularities of six manners and their main features in Risk Conduction [J]. Modernization of Management. 5 (2005). pp. 38-40.

Kathrin Büttnera, Joachim Krietera, et al. Static network analysis of a poultry supply chain in Northern

Germany—Characterisation of the potential spread of infectious diseases via animal movements [J].Preventive Veterinary Medicine. 110(2013) .pp.418-428.

Klibi W, Martel A. Scenario-based supply chain network risk modelling [J]. European Journal of Operational Research, 2012. pp.

Mayer D, Reiczigel J, Rubel F. A Lagrangian particle model to predict the airborne spread of foot-and-mouth disease virus [J]. Atmospheric Environment. 42 (2008). pp. 466-479.

Mitchell V W. Organizational risk perception and reduction: a literature review [J]. British Journal of Management, 6(1995) .pp.115-133.

Monique A. van der Gaaga Fred Vosa et al. A state-transition simulation model for the spread of Salmonella in the pork supply chain [J]. European Journal of Operational Research, 156 (2004). pp. 782-798.

Moslonka-Lefebvre M, Pautasso M, Jeger M J. Disease spread in small-size directed networks: epidemic threshold, correlation between links to and from nodes, and clustering [J]. Journal of theoretical biology. 260(2009). pp. 402-411.

Nagurney A, Cruz J, Dong J, et al. Supply chain networks, electronic commerce, and supply side and demand side risk [J]. European Journal of Operational Research, 164 (2005) .pp. 120-142.

Nakamaru M, Levin S A. Spread of two linked social norms on complex interaction networks [J]. Journal of Theoretical Biology. 230 (2004). pp. 57-64.

Neiger D, Rotaru K, Churilov L. Supply chain risk identification with value-focused process engineering [J]. Journal of Operations Management, 27 (2009) .pp. 154-168.

Nekovee M, Moreno Y, Bianconi G, et al. Theory of rumour spreading in complex social networks [J]. Physica A: Statistical Mechanics and its Applications. 374 (2007). pp. 457-470.

Paul M, Baritaux V, Wongnarkpet S, et al. Practices associated with Highly Pathogenic Avian Influenza spread in traditional poultry marketing chains: Social and economic perspectives [J]. Acta tropica, 126 (2013). pp. 43-53.

Pinior B, Konschake M, Platz U, et al. The trade network in the dairy industry and its implication for the spread of contamination [J]. Journal of dairy science, 2012.

Santoso T, Ahmed S, Goetschalckx M, et al. A stochastic programming approach for supply chain network design under uncertainty [J]. European Journal of Operational Research, 167 (2005). pp. 96-115.

Sofyalıoglu C, Kartal B. The Selection of Global Supply Chain Risk Management Strategies by Using Fuzzy Analytical Hierarchy Process–A Case from Turkey [J]. Procedia-Social and Behavioral Sciences, 58 (2012). pp.1448-1457.

Ventresca M, Aleman D. Evaluation of strategies to mitigate contagion spread using social network characteristics [J]. Social Networks. 2013. pp. 75–88.

Wu D, Wu D D, Zhang Y, et al. Supply chain outsourcing risk using an integrated stochastic-fuzzy optimization approach [J]. Information Sciences, 2013.pp.242–258.

Wu T, Blackhurst J, Chidambaram V. A model for inbound supply risk analysis [J]. Computers in industry, 57(2006) .pp. 350-365.

Zhao L, Cui H, Qiu X, et al. SIR rumor spreading model in the new media age [J]. Physica A: Statistical Mechanics and its Applications, 2012.

Zhao L, Wang Q, Cheng J, et al. Rumor spreading model with consideration of forgetting mechanism: A case of online blogging LiveJournal [J]. Physica A: Statistical Mechanics and its Applications. 390 (2011). pp. 2619-2625.

Lei Wen was born July 1973. The management master degree is gained from economic and management department of North China Electric Power University on 2000 at Baoding city of Hebei province of china. The management doctor degree is gained from management department of Tianjin University on 2004 year at Tianjin city of china. His research interests focus on supply chain risk management and simulation, intelligent optimization algorithms etc.

He is the professor and vice director at economic and management department of North China Electric Power University of china.The main publisher include: (1) Statistic charicteristics analysis of directed supply chain complex network, the international journal of advancements in computing technology, volumn4, nov30, 2012. (2) The modeling and simulation of supply chain based on directed complex network. Journal of information & computational science, volumn10, dec 10, 2013. (3) Risk-immune spread in supply chain network based on single_node risk. Asian

Conference on the Social Sciences, 2013.

Yachao Shi was born October 1989. At present, she is studying for a master's degree at North China Electric Power University. She is majoring in logistics engineering. Her research interests focus on supply chain risk management and simulation, Intelligent optimization algorithms etc.

The main publisher include: (1) The modeling and simulation of supply chain based on directed complex network. Journal of information & computational science, volumn10, dec 10, 2013. (2) Risk-immune spread in supply chain network based on single_node risk. Asian Conference on the Social Sciences, 2013.

doi: 10.14355/ijcsa.2014.0303.04

ASAMO: Authentication and Secure Communication using Abstract Monitoring Objects for Mobile Grid Computing

Parveen Begam H*1, Maluk Mohamed M A[2]

Software Systems Group, M.A.M College of Engineering, Tiruchirappalli, India

*1ssg_parveen@mamce.org; [2]ssg_malukmd@mamce.org

Abstract

The prominent feature of mobile grid computing is the collaboration of multiple entities to perform collaborative tasks using mobile devices. Currently, most of the security solutions for mobile grid environment use static set of algorithms and protocols. Despite the increased usage of mobile computing, exploiting its full potential is difficult due to its inherent problems such as resource scarcity, frequent disconnections, and mobility. Our focus is to provide security and efficient power management. In this paper, a framework of dynamic secure routing protocol called Select Successive Hop Routing (SSHR) algorithm using Abstract Monitoring Objects (AMOs), Secure Service Certificates (SSC) and layer based encryption for each data transfers to improve the security level, anonymity and dynamicity of the data are proposed. This includes (i) Authentication of the mobile node (ii) Security flow between mobile nodes (iii) Saving battery power (iv) Hand-over of authentication information. The proposed framework solves the issues like asymmetry in network connectivity, computing power and mobility. The outcome is compared with certain existing security protocols and demonstrates how the proposed protocol is able to adapt itself over different conditions of the mobile device, and how it can provide a performance gain in the execution.

Keywords

Mobile Grid Computing; Authentication; Secure Communication; Secure Routing; Digital Signature; ECDSA.

Introduction

Mobile Grid computing extends the traditional Grid computing paradigm to include a diverse collection of mobile devices that communicate using radio frequency, infrared, optical and other wireless mechanisms. The prominent feature of mobile grid computing is the collaboration of multiple entities to perform collaborative tasks using mobile devices that rely on two fundamental functions: communication and resource sharing. The fundamental function is to enrich one another and provide new solutions that solve many of the limitations and problems found in different technologies, such as reduced CPU performance, limited secondary storage, heightened battery consumption sensitivity, and unreliable low-bandwidth communication. Although the individual computing devices may be resource-limited in isolation, as an aggregated sum, they have the potential to play a vital role within grid computing.

Security is a very important factor in mobile grid Computing and is also difficult to achieve owing to the open nature of wireless networks and heterogeneous and distributed environments. Since Internet is not security-oriented by design, there exist various threats, in particular, malicious internal and external users. Securing communication and fine-tunning controlling access to shared resources are the important issues for mobile grid services. Currently, most of the security solutions for the mobile grid environment use a static set of algorithms and protocols. The countermeasures against threats utilize encryption/decryption for confidentiality, message authentication code for integrity, digital signature for authentication, undeniable digital signature for non-repudiation, access control for authorization, and intrusion detection/defense for availability/DOS. The Internet-based grid computing encounters similar threats and involves all the security requirements discussed above. As a result, the threats to the mobile grid systems may become more serious and defending them becomes more difficult. For instance, centralized authentication is generally unavailable and multiple-site co-authentication is difficult to implement due to the distributed and heterogeneous features of grid computing systems. Thus, the Single-Sign-On authentication comes into play. As long as communication and information exchange is conducted over the Internet, communication messages should be encrypted to provide confidentiality.

Initially, the integration of mobile wireless consumer devices into the Grid, seems unlikely, due to the inherent limitations typical of mobile devices such as reduced CPU performance, limited secondary storage, heightened battery consumption sensitivity, and unreliable low-bandwidth communication. However, the millions of laptops and PDAs sold annually suggest that this untapped abundance should not be prematurely dismissed. Given that the benefits of combining the resources of mobile devices with the computational grid are potentially enormous, one must compensate for the inherent limitations of these devices in order to utilize them successfully in the Grid. The research challenges arising from this problem, propose the vision of a potential architectural solution. Proxy-based, clustered system architecture with favorable deployment, interoperability, scalability, Adaptivity, and fault-tolerance characteristics as well as an economic model to stimulate future research in the emerging field has been suggested.

A survey of the current state of wireless grid computing, includes a discussion of the cooperation between wired and wireless grids, including ways in which wireless grids extend the capabilities of existing wired grids. It also discusses many of the new capabilities and resources available to wireless grid devices and a sampling of several applications of these new resources. It provides a sampling of many current research endeavors in the wireless grid arena and an examination of a number of the potential challenges resulting from the unique characteristics of wireless grid devices.

The Surrogate Object (SO) is a software entity that acts as a representative of a particular Mobile Host (MH) in the wired network and maintains application specific data structures and methods. The major advantages of using a surrogate object are: Maintaining location information about each mobile device, acting as a place-holder that can realize local caching for faster information access, handling message delivery for the MH, when it is out of reach from the MSS and acts as data sink that can collect data from diverse sources and delivering appropriate data to the MH depending upon the current position of the mobile devices to progress in a mobile grid environment in a secured manner. It is very important that the downstream node may or may not be malicious to provide secure communication.

A development methodology for Secure Mobile Grid Systems is proposed in which the security aspects are considered from the first stages of the life-cycle and in which the mobile Grid technological environment is always presented during each activity. The analysis activity, in which the requirements (focusing on the grid, mobility and security requirements) of the system are specified and driven by reusable use cases through which the requirements and needs of these systems are defined. These use cases have been defined through a UML-extension for security use cases and Grid use cases which capture the behavior of these kinds of systems.

Mobile cloud computing can address some security problems by executing mobile applications on resource providers external to the mobile device. An extensive survey of mobile cloud computing research, while highlighting the specific concerns in mobile cloud computing, will be done. The paper is concluded with a critical analysis of challenges that have not yet been fully met.

The security service, which works as a middleware, with the ability to dynamically change the security protocols used between two peers, is proposed. These changes can occur based on variations on wireless medium parameters and system resource usage, available hardware resources, application-defined QoS metrics, and desired data ‘‘security levels’’. The compared solution to certain widespread static security protocols demonstrates how the middleware is able to adapt itself over different conditions of medium and system, and how it can provide a performance gain in the execution of cryptographic primitives, through the use of data semantics.

The security support for a middleware framework supporting Mobile Grid Services is proposed. Mobile Grid Services, the extension of the original static Grid services, are characterized by the ability of being able to move from nodes to nodes during execution. By combining an existing mobile agent system (JADE) and a generic grid system toolkit (Globus), the Mobile Grid Services framework is achieved. Security is a critical issue when deploying Mobile Grid Services to the real-world applications, where the resources of the services must be protected in potentially hostile environments. The Mobile Grid Services Framework is not practical, if no security measures are provided. The details of security mechanisms consisting of authentication, authorization, message integrity and confidentiality, agent permission and agent protection as well as their realization in the MGS API are presented.

In this paper, a framework of secure routing protocol using Abstract Monitoring Objects (AMOs), Secure Service Certificates (SSC) and layer based encryption in each mobile device to improve the security level and dynamicity of the data is proposed. The Surrogate Object (SO) is a software entity that acts as a representative of a particular Mobile Host (MH) which helps in storing the service certificates and also in solving issues like asymmetry in network connectivity, computing power and mobility. In wireless environment, the downstream node may be authorized or malicious. ECDSA (Elliptic Curve Digital Signature Algorithm) is used for anonymous communication with the receiver and the encryption of sender's information to maintain the secrecy of the message. Usage of hop to hop TTL increases the reliability and time constraints of the message. The proposed system overcomes the existing security related drawbacks upto a certain extent.

The rest of the paper is organized as follows. Section 2 describes the importance of security in a mobile grid environment over which the Abstract Monitoring Object (AMO) model has been built to provide effective authentication and secure communication. Section 3 describes the AMO model in detail, including Secure Service Certificates, Layered Routing. Section 4 analyzes security-related measures, robustness and complexity of the proposed architecture. Section 5 compares the performance with some existing methods. Section 6 concludes the paper and provides directions for future research.

Brief Overview of the Security Issues in Mobile Grid

Security is a very important factor in Mobile Grid Computing and is also difficult to achieve, owing to the open nature of wireless networks and heterogeneous and distributed environments. In a distributed job execution environment, the potential risks rise for both the integrity of the application and the resource provider. The prime motivation of combining the mobile and grid computing is to carry out the user's work while on move. Due to various constraints to the wireless network, the environment and the user may rapidly change their environment from stationary to mobile and location dependent. The combination of mobile and the grid may lead to a lot of security issues like authentication of a mobile node and mobile code, prevention of attacks in the base station, secure communication between two mobile nodes, communication cost of constructing the session keys and computing complexity of authenticity and security. Particularly in a mobile grid environment, some of the risks like the integrity of processing results are compromised by malicious mobile node, unnecessary free riding of MN without contributing the resources which in turn reduces the system utility, spoiling the data confidentiality, destruction of files and applications by malicious mobile code and acting as owner are to be solved. Providing a countermeasures for the security issues is crucial.

The countermeasures under investigation include the authentication, confidentiality, secure communication, access control and so on. In each situation, our aim is to balance functionality, performance and security while achieving solutions and without imposing restrictions (eg. Increased power consumption) on the personal use of a mobile device. In order to provide security in a holistic manner, a new and efficient dynamic secure routing for mobile grid environments with the help of secure service certificates and Abstract Monitoring Objects along with layered routing for mobile grid computing is proposed.It is as follows.

The Proposed Security Object-based Model

The supervisory host model for mobile grid systems defines an architecture that helps in generating the Secure Service Certificates (SSC) and digital signatures for mobile hosts to have efficient security and power management. Some of the existing methods to generate digital signatures and certificates are Symmetric or Asymmetric algorithms, Identity based algorithms, biometric-based algorithms. In the existing systems, mobile hosts depends on Certificate Authority (CA) for generation of certificates using various certificate generation methods. Issues arise when the CA becomes malicious and when it consumes more battery power. The focus is to provide security and power management, this includes (i) Authentication of a mobile node - achieved by providing certificates with digital signature stored in the Surrogate Object(SO), (ii) Security flow between two mobile nodes - achieved by providing layered routing to the original data and preventing it from hackers, (iii) Saving the battery power - achieved by introducing a Supervisory Host (SH) (a static node located at Mobile Support Station (MSS) of each Mobile Grid Network (MGN) and it takes care of certificate generation and signing) and (iv) Hand-over of authentication information - done using ECDSA which has 168 bit key size which is far less when compared to RSA algorithm. The introduction of SH will enhance the security, anonymity and power consumption.

Certificate Generation

Any mobile node (MHi) entering into Mobile Grid Network (MGNi) should perform single sign-on process for authentication. A node (say MHs) from one grid establishing communication with other node (MHd) in another grid, must notify the receiver that it is an authenticated mobile node. To achive this, the Secure Service Certificate (SSC) is used. Figure 1 explains the certificate generation process with the help of supervisory host which resides in the base station. The mobile node which communicates with another node, firstly requests for the public key to the common (public key information) pool of that grid network. The pool is a database with a collection of public keys of all the mobile grid nodes that exists in Mobile Station Subsystem (MSS). After receiving the public key, the mobile host prepares the necessary components to generate its Secure Service Certificate. The mobile host sends those components to the Supervisory Host (SH). The Supervisory Host is a special host residing in static environment generating a digital certificate and sending it to the mobile grid node. After verification, the certificate is accepted by mobile node. There are individual Supervisory Hosts for each individual mobile grid network (MGNi).

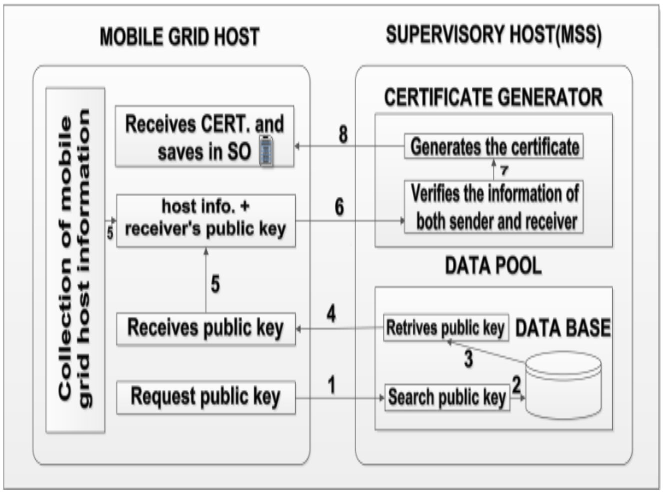

FIG. 1 CERTIFICATE GENERATION PROCESS USING SUPERVISORY HOST

The Mobile Host issues the necessary components to SH in an encrypted format. After receiving the contents, SH starts generating the certificate along with a digital signature with the help of sender's and receiver's public key. Figure 2 shows the Secure Service Certificate model. The components of the certificate are: serial number, certificate identifier algorithm, issuer's unique id, validity period, subject id, subject public key, algorithm identifier, TTL and Nonce. (IDs||TTL||Kpu(s)||IDd||N1). Certificate identifier algorithm helps to identify the encryption algorithm in which the certificate has been encrypted. Issuer's unique id is the identity of the supervisory host. The validity period is the time for which the certificate is valid. Subject id is the id of the receiver. Subject public key is receiver's public key. The sender node communicates with the receiver using this public key. Algorithm identifier is a type of algorithm where the data are encrypted. Finally, SH makes a copy of the SSC and sends it to the corresponding mobile node.

FIG. 2 SECURE SERVICE CERTIFICATE MODEL

Authentication

1) *Storage of Certificate in SO*

After generating the certificate, it is very necessary to store the certificate in a secure location such that no intruder can access the certificate. Figure 3 explains the procedure to store the certificate in its SO. To achieve this, each mobile node stores the SSC in its Surrogate Object (SO). The Surrogate Object is a database which stores all the details about the mobile node in a secure manner. The mobile nodes move out of grid network anytime and anywhere. If a mobile node has an incomplete process and if that node is not present in that particular grid, the process becomes tedious and remains undone. To overcome such situations, the SO acts as a virtual mobile node and performs all the jobs performed by that particular mobile node. To store a certificate in SO, the mobile node sends the request message to the SO.

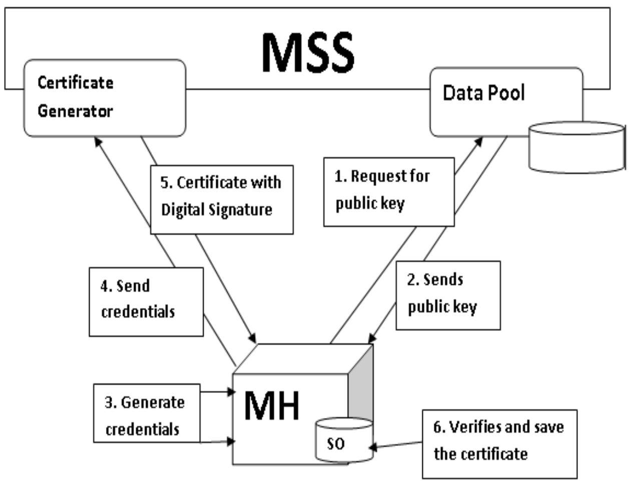

FIG.3 CERITIFTAE STORAGE PROCESS IN SO

The request message contains all the attributes of the certificate such as size, access rights and so on. When the Surrogate Object receives the request message, it checks for authentication and stores the certificate. Since the mobile devices are not stable in a mobile grid environment, there is a possibility for an intruder to forge as if it is the original mobile node. To avoid this, the Surrogate Object verifies the validity of the certificate and stores it. If authentication fails, the SO will reply with an error message. The mobile node sends the certificate to SO and it is stored secure. If the storage is successful, the SO sends another ACK signal indicating the successful storage. For unsuccessful storage of certificate, the SO sends a negative acknowledgement (NACK) signal. When a NACK is received, the mobile nodes redo all the storage process.

2) Certificate Verification

In the mobile grid, since the mobile devices are not stable, there is a possibility for an intruder to forge as if it is the original mobile node. To avoid this, the Surrogate Object verifies whether the certificate is valid or not and then stores it. To make this, the Surrogate Object sends the details of the certificate to the SH and gets the certificate verified. First, the Surrogate Object directly communicates with the SH and sends the details of the certificate along with the details of its own mobile node. The SH searches its database and matches with the details provided by the Surrogate Object. If it matches with the given data, the SH returns an ACK signal. When an ACK signal is received, then it is clear that the certificate is valid and original. Now, the Surrogate Object stores the certificate securely for future use.

Secure Communication

A communication between two mobile nodes is said to be secure, only if the data sent from one node to another cannot be opened by a malicious node and the sent data reaches the destination safely. The data are forwarded to the next node only after the authentication of the node. The authentication is done by the SSC. The process of checking the authenticity of next node is done by a monitoring object called Abstract Monitoring Object (AMO). This object monitors the entire data transmission and ensures the security and reliability of data. This also triggers secure routing process for secure and fast transfer of data to the destination.

1) AMO Processing

Figure 4 shows the complete architecture of AMO (Abstract Monitoring Object) which is used to monitor the nearest nodes and locate intruders. The mobile host sends a dummy data to the nearest mobile node. This dummy data are called as AMO. When the nearest node is able to decrypt the AMO, then the node is an authorized node. If the node is authorized, the mobile node sends the original data to that node. The primary step in this module is AMO generation. To generate an AMO, the mobile node encrypts a dummy data and makes it look like the original message. This AMO is transferred to the nearest mobile node. The mobile node receives the AMO and tries to decrypt the message. When the mobile node succeeds in decrypting the message, an automated ACK signal is received at the sender node. After that, the sender node sends the original data. If the nearest node is an intruder, the node cannot decrypt the data. When the Time To Live (TTL) expires, an automated NACK signal is received. After that, the mobile node redirects the path of the data transfer to the net nearest node. The process is repeated until the data reach its destination successfully.

When the ACK signal is received by the sender node, it indicates that the nearest node is an authorized node. Then the mobile node encrypts the original data with ECDSA algorithm with 168 bits of key size. After encryptions, the data are padded with the certificate along with the digital signature and the data are transferred securely to the nearest node from which the ACK signal is received. If the receiver node is the destination, the node decrypts the data directly and uses it. If the node is not the destination, the entire operation is repeated.

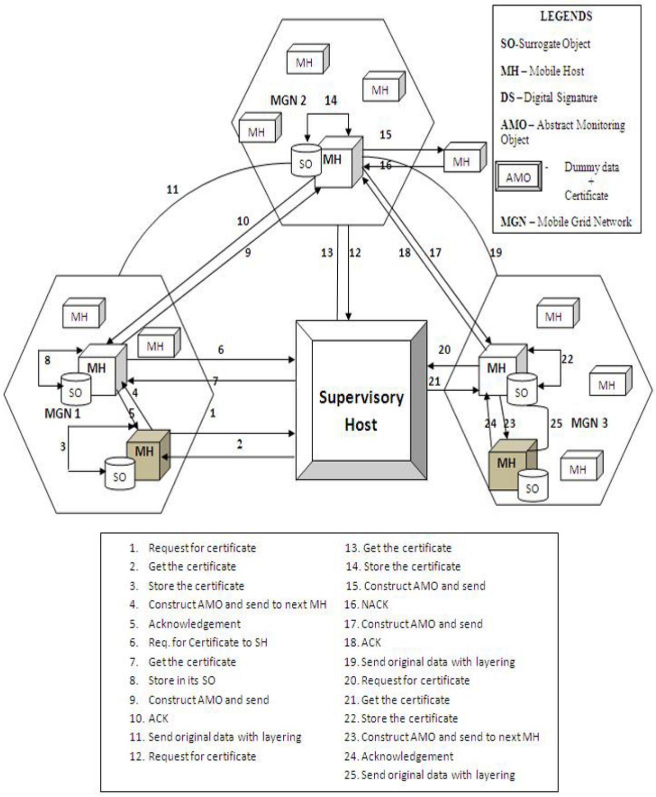

FIG. 4 AMO (ABSTRACT MONITORING OBJECT) ARCHITECTURE DIAGRAM

2) *Select Successive Hop Routing Algorithm*

The AMO is created just after the successful verification and storage of the SSC. The AMO consists of an encrypted dummy data along with the SSC. As soon as the AMO is created, it triggers the Select Successive Hop Routing algorithm (SSH) to find the next nearest node. After finding the next nearest node, the AMO consisting of dummy data and the SSC is transferred to the nearest node to check whether it is authenticated or not. If the node is authenticated, the AMO automatically returns an acknowledgement signal (ACK) to the upstream node. If the nearest node is unauthorized or when the TTL expires, a negative acknowledgement (NACK) is automatically sent to the upstream node. In a mobile grid environment, all devices are dynamic and unstable. Hence maintaining a routing table and finding the destination becomes a major issue. Thus to ensure and enhance the security and reliability of data that are transferred through this mobile grid environment, a new routing algorithm called Select Successive Hop (SSH) Routing algorithm is designed. It is designed in such a way that a drastic enhancement in reliability and security within a short transmission time is explicitly seen.

Without AMO:

1. A mobile node (say MHi) enters a Mobile Grid Network (say MGNi) using Single Sign-on authentication
2. Before the source mobile host (MHs) communicating with the destination mobile host (MHd), a certificate from Certificate Authority (CA) is received by sending a request message as (IDs||TTL|| Kpu(s) ||IDd||N1)
3. CA receives the request message and checks for the authentication
    a. If yes, the CA returns the public key of the destination host (Kpu(d)) to MHs
    b. Else, CA returns an error message to MHs
4. After receiving the Kpu(d) from CA, MHs integrates all necessary components and generates an X.509 Digital Certificate using RSA algorithm
5. MHs & MHd performs Mutual Authentication
6. After the mutual authentication, the original data transfer takes place

With AMO:

1. A mobile node (say MHi) enters a Mobile Grid Network (say MGNi) using Single Sign-On authentication
2. Before communicating with destination mobile host, the source mobile host (MHs) contacts Supervisory Host (SH) and gets a certificate
3. To achieve this, MHs gets the public key of the destination (Kpu(d)) from a public key database (data pool) located at MSS
4. Then it generates all the necessary components (IDs||TTL|| Kpu(s) ||IDd||N1) for SSC (Secure Service Certificate) and sends it to the SH
5. SH receives the request message and checks for authentication
    a. If yes, SH generates and provides the SSC to MHs
    b. Else, SH returns an error message to MHs

6. Each MHi has an individual Surrogate Object (SOi) which acts
   a. As a cache memory and stores all the details about the corresponding MHi securely
   b. As an object, the SOi does the process of its corresponding MHi when it is unavailable in the corresponding MGNi
7. The MHs stores the SSC in its SO in a secure manner
8. Communication between two mobile nodes can be done using the Abstract Monitoring Object (AMO)
9. As soon as the AMO is created, it automatically triggers the Select Successive Hop Routing algorithm (SSHR) to locate the next nearest node
10. After knowing the location, the AMO senses the corresponding direction for a nearest neighbor
11. When a nearest host (MHn) is found, the AMO sends a dummy data to that node
12. After receiving the dummy data, the MHn tries to read the data
    a. If the receieved data are authorized, an ACK is sent to MHs and original data transfer takes place
    b. If the data are unauthorized, then NACK will be sent to MHs. Then the path is redirected to next nearest node
13. MHs encrypts the original data and forms a layer wrapping along with the certificate
14. MHn receives the data and reads the data
    a. If the MHn is an original destination, it removes the layer and reads the original message
    b. If the node is not the destination, it removes the existing layer, makes another layer of encryption and transfers it to the next nearest neighbor. The process is repeated from step 10 until destination (MHd) is reached

The different possible scenarios that could occur are:

Case 1: "Best Case"

The next nearest node MHi is the destination node, i.e.,. (i=d) or all the MHi are authorized nodes. Thus the constraint is t(n) ≥ MH(g(n)) where t(n) is the time constraint and the MH(g(n)) is the mobile host function

Case 2: "Average Case"

There are some considerable amounts of unauthorized nodes in the network where the constraint is MH2(g(n)) ≤ t(n) ≤ MH1(g(n)) where MH is the mobile host

Case 3: "Worst Case"

All the nearest nodes MHi are unauthorized nodes. Hence the time constraint will be maximum or infinite. Thus the method fails as t(n) ≤ MH(g(n)) where t(n) is the time constraint and the MH is the mobile host

Performance Analysis

To study the performance of the model, the data loss probability scenario explained in the previous section were considered, and were implemented over a simulated model of a Mobile Grid network. The scenario was studied over both the Abstract Monitoring Object and Supervisory Host along with and without Abstract Monitoring Object and Supervisory Host. Throughout this section, the algorithm without Abstract Monitoring Object and Supervisory Host is referred to as the old model and the one with Abstract Monitoring Object and Supervisory Host is termed as the new model. Typical performance studies include the actual number of packets lost for different number of mobile hosts in both the old and the new models, such as proportionate increase in the packet loss for different number of mobile nodes in both the models; impact of movement of Abstract Monitoring Object and Supervisory Host in terms of packet loss and message traffic. Query time both in the old and the new model is also taken into account for performance comparison.

Comparison of the Old and the New Models

The comparisons of the query time in the caching application over both the old and the new models were discussed. Query time was the overall time taken from an action request and the reception of the reply by a source node. Figure 5 shows the query time in the old model. The systems were implemented from a source to destination TTL. If the data were lost in between or they were sent to any unauthorized node, the sender would have to wait until the TTL elapsed. Thus the query time was much higher and there was a higher magnitude in waiting time.

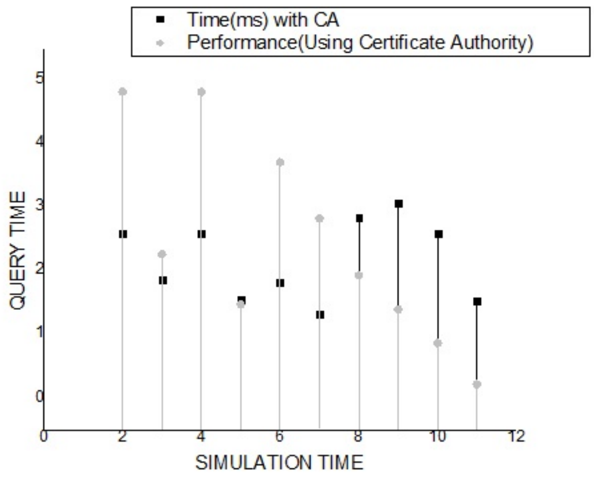

FIG. 5 QUERY TIME COMPARISON IN THE OLD MODEL

Figure 6 shows the query time in the new model. In the model, hop to hop TTL is used addition to source to destination TTL.In hop to hop TTL, the magnitude of TTL is low. Thus, the sender need not wait until the TTL elapses, if the data were lost in between or they were sent to any unauthorized node. Thus the query time was much reduced and there was a lower magnitude in waiting time. The immediate reply can be obtained if any unauthorized node is detected. Hence, the query time is far less than the existing model.

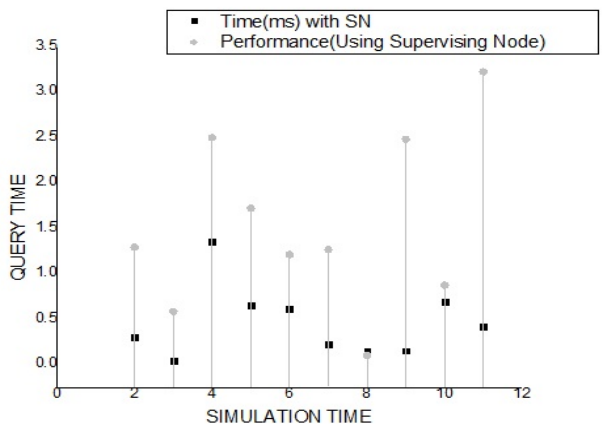

FIG. 6 QUERY TIME COMPARISON IN NEW MODEL

One of the key advantages of the security objects (Surrogate object and Abstract Monitoring Object) model is that it is free to migrate to any node in the wired network. This flexibility is desirable for various reasons, including failure recovery, security, load balancing and network latency reduction. The effects of the query latency for various migration frequencies were discussed. Figure 7 shows the migration of security objects across the many numbers of nodes. The query latencies are considerably high. This is due to the increase in the percentage of time spent for generation of certificates and verification of certificates by certificate authority itself.

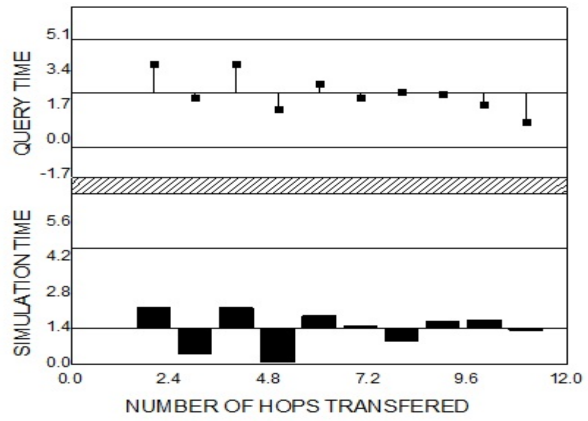

FIG. 7 QUERY TIME BASED ON NUMBER OF NODES TRAVELLED IN OLD MODEL

Figure 8 shows the migration of security objects across many numbers of nodes. The query latency is considerably less, because of the increase in the percentage of time spent on the generation of certificates and the verification of certificates by supervisory host itself. But as the move frequency is reduced, the query time improves. Surprisingly, further change in the move frequency seems to have almost no impact on the query time. A closer study of the network traffic generated as a result of performance, reveals the reason for providing migration freedom for the security objects.

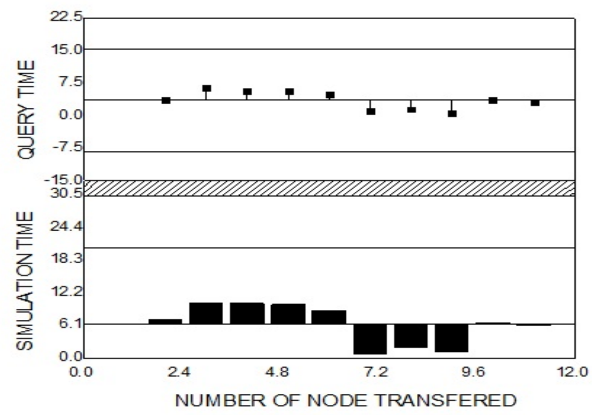

FIG. 8 QUERY TIME BASED ON NUMBER OF NODES TRAVELLED IN NEW MODEL

Figure 9 shows the comparison of query time in both Old and New models in terms of number of messages exchanged versus simulation time for various move frequencies.

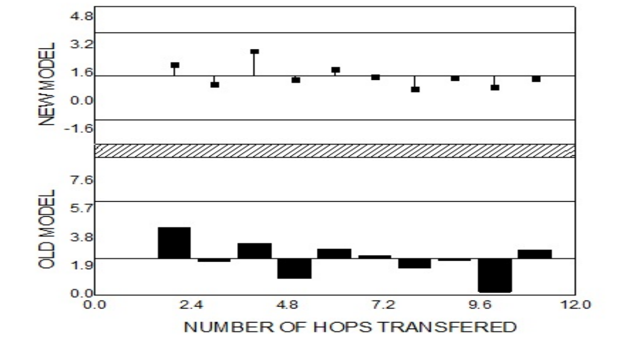

FIG. 9 PERCENTAGE INCREASE IN QUERY TIME WITH INCREASING NUMBER OF NODES TRAVELLED

Figure 10 shows the comparison of the packet loss probabilities in the caching application over both the old and new models. The packet loss probability is plotted against the number of nodes travelled in the network for different simulation time intervals. In the old model, the TTL is used between the source and the destination. If the TTL elapses, then the packet will be lost and NACK is sent. If the packet is lost in the network due to any link break or node failure, the packet will remain in the network (until TTL elapses) which increases the possibility of congestion. The major drawback is rectified by hop to hop TTL, which allows the lost packet to be destroyed quickly because of less magnitude of the TTL.

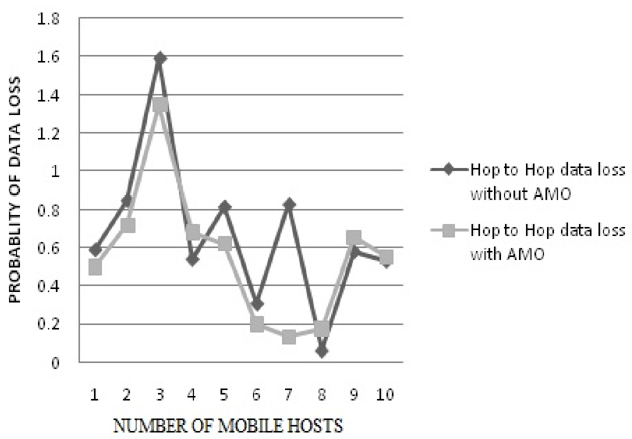

FIG.10 PERCENTAGE INCREASE IN QUERY TIME WITH INCREASING PACKET LOSS PROBABILITY

Conclusions

This paper proposes the secure routing based model for providing security and faster routing without drastic fall in performance of data transfer. Authentication of the mobile node can be achieved by providing certificates with digital signature stored in the Surrogate Object (SO). Security between two mobile nodes can be achieved using AMO (Abstract Monitoring Objects). Mobile devices are battery powered. Sobetter power consumption of mobile devices can be achieved through Supervisory Host (SH) where SH is located at Mobile Support Station (MSS) and takes care of certificate generation and signing. The hand-over of authentication information can be done using ECDSA (168 bit key size) wherethe key size is far less when compared to RSA algorithm.

It is felt that it is the first practical model in the literature for providing security and reliability under various conditions, and the first integrated model in all of the literature to consider individual variability and behavior of each mobile host using security objects for the enhancement of security and reliability. The routing based security model is used to improve security without degrading the performance of the system.

REFERENCES

A. Frier, P. Karlton and P. Kocher, \The SSL3.0 Protocol Version 3.0", see http://home.netscape.com /eng/ssl3/.

Arash Habibi Lakshkari, PIR Mohamed seyed dhanesh, Behrang samadi,"A Survey on Wireless Security Protocols"2009, IEEE.

Bruno P.S.Rocha DanielN.O.Costa, RandeA.Moreira, CristianoRezende,Antonio.F.Loureiro,AzzedineBoukerche,Adaptive security protocol selection for mobile computing, Journal of Network and Computer Applications 33 (2010) 569– 587,2010 Elsevier.

Bruno P.S.Rocha, DanielN.O.Costa, RandeA.Moreira, CristianoG.Rezende, Antonio A.F.Loureiro,, Azzedine Boukerche, Adaptive security protocol selection for mobile computing, FGCS,2010,Elsevier.

Budiarto, Shojiro Nishio, Masahiko Tsukamoto "Data management issues in mobile and peer-to-peer environments", 2002 Elsevier Science.

Chang-Seop Park, Dankook University ,On Certificate Based Security Protocols for Wireless Mobile Communication Systems ,0890-8044/97, 1997, IEEE Network September-October 1997.

Dario Bruneo, Marco Scarpa, Angelo Zaia, and Antonio Puliafito. Communication Paradigms for Mobile Grid Users. In Proceedings of the 3rd IEEE/ACM International Symposium on Cluster Computing and Grid. Tokyo, Japan, june 2003.

David G. Rosado a, Eduardo Fernández-Medina a,*, Javier

López b, Mario Piattini, Analysis of Secure Mobile Grid Systems: A systematic approach, FGCS, 2012, Elsevier.

Elliptic Curve Cryptography-An Implementation Tutorial-Anoop MS,Tata Elxsi Ltd, Thiruvananthapuram, India, 1997.

Enhancing Grid Security Infrastructure to Support Mobile Computing Nodes-K. Chae and M. Yung (Eds.): WISA 2003, LNCS 2908, pp. 42–54, 2004.c Springer-Verlag Berlin Heidelberg 2004.

Ersan Kayan and Ozgur Ulusoy, "An Evaluation of Real-Time Transaction Management Issues in Mobile Database Systems" Department of Computer Engineering and Information Science, Bilkent University, Turkey,1999.

Fei Yan, Huanguo Zhang, Zhidong Shen, Liqiang Zhang, Weizhang Qiang,"An Improved Wireless GSI based on Trusted Computing Technology", IEEE 2006.

Fei Yan, Huanguo Zhang, Zhidong Shen, Liqiang Zhang, Weizhang Qiang,"An Improved Wireless GSI based on Trusted Computing Technology", IEEE 2006.

Giuseppe Anastasi, Alberto Bartoli, and Francesco Spadoni. A Reliable Multicast Protocol for Distributed Mobile Systems: Design and Evaluation. IEEE Transactions on Parallel and Distributed Systems, 12(10):1009–1022, October 2001.

K Vedder, "Security Aspects orMObile Communication," Computer Security and fndustrrol Crypfography-State of the Art and Evolution, Springer-verloa. May. 1991, DD. 193-21 0

Key Sizes Selection in Cryptography and Security Comparison between ECC and RSA- M.J. Wiener, Efficient DES key search, manuscript, ell-Northern Research, August 20, 1993.

Konstantinos Katsaros and George C. Polyzos, Towards the Realization of a Mobile Grid, CoNEXT'07, December 10-13, 2007, New York, NY, U.S.A. Copyright 2007 ACM.

Kwok-yan Lam, Xi-Bin Zhao, Siu-Leung chung, Ming Gu and Jia-Guangum, "Enhancing Grid Security Infrastructure to support Mobile Computing nodes" ,Springer verlog Berlin Heidalberg,2005.

M.A. Maluk Mohamed, D. Janaki Ram and Mohit Chakraborty, "Surrogate Object Model: A New Paradigm for Distributed Mobile Systems", Proceedings of the 4th International Conference on Information Systems Technology and its Applications (ISTA'2005), May 23-25, 2005 - New Zealand, pp.124- 138.

Niroshinie Fernando , Seng W. Loke , Wenny Rahayu," Mobile cloud computing: A survey",FGCS,2012,Elsevier.

Randy Butler, Von Welch, Douglas Enhert, Ian foster, Steven Tweke, John volmer, Carl Kesselman "A national scale Authentication Infrastructure", 2000,IEEE.

Reynold Cheng, Yu Zhang, Elisa Bertino, and Sunil Prabhakar, "Preserving User Location Privacy in Mobile Data Management Infrastructures" Springer-Verlag Berlin Heidelberg 2006.

Sander and C. Tschudin, "Towards mobile cryptography," in Proc. of the IEEE Symposium on Security and Privacy, Oakland, CA, USA, 1998.

Sanjay P.Ahua, Jack R.Myers, "Survey on wireless Grid Computing", Journal on Super computing,2006,Springer Science.

Sharad Jaiswal and Animesh Nandi ," Trust No One: A Decentralized Matching Service for Privacy in Location Based Services", MobiHeld 2010, ACM, August 2010.

Speeding up Secure Web Transactions Using Elliptic Curve Cryptography- Vipul Gupta, Douglas Stebila, Nils Gura, Hans Eberle Sun Microsystems,Inc.

Sze-Wing Wong and Kam-Wing Ng, "Security Support for Mobile Grid Services", International Journal of Web Services Practices, Vol.2, No.1-2 (2006), pp. 11-20.

The SSL Landscape – A Thorough Analysis of the X.509 PKI Using Active and Passive Measurements Ralph Holz, Lothar Braun, Nils Kammenhuber, Georg Carle, IMC'11, November 2–4, 2011, Berlin, Germany. Copyright 2011 ACM.

Thomas Phan, Lloyd Huang,Chris Dulan, Challenge: Integrating Mobile Wireless Devices Into the Computational Grid MOBICOM'02, September 23–26, 2002.

Thomos pan, Llyod Huang, Chris Dulan, "Challenge:Integrating Mobile Wireless Devices into the Computational Grid", MOBICOM'02,ACM September 2002 .

Vipul Gupta, Sumit Gupta, Sheueling Chang Performance Analysis of Elliptic Curve Cryptography for SSL, WiSe'02, September 28, ACM 2002, Atlanta, Georgia, USA.

Weiqi Dai,T.Paul Parker,Hai jin,"Eworthiness via Assured Digital Signing" IEEE Transaction on Dependable and Secure Computing,Vol 9,No6.2012.

Xukai Zou, Yuan-Shun Dai , Xiang Ran,"Dual –Level Key Management for secure Grid Communication in Dynamic and hiearchical groups",FGCS 2006 Elsevier.

Xukai Zou, Yuan-Shun Dai , Xiang Ran,"Dual –Level Key Management for secure Grid Communication in Dynamic and hiearchical groups",FGCS 2006 Elsevier.

Yue-Hsun Lin, Ahren Studer,Hsu-Shun, Jonathan, King-Hang Wang, "SPATE: Small-group PKI-less Authenticated Trust Establishment", MobiSys'09, Poland, 2009 ACM.

Yuliang Zheng,"An Authentication and Security protocol for mobile computing",1999,IEEE.

Yun Zhou, Xiaoyan Zhu and Yuvang Fang, IEEE"MABS:MultiCast Authentication Based on Digital Signature", IEEE Transaction on Mobile Computing, Vol 9,No 7,July 2010.