J Glob Optim (2011) 51:285–300
DOI 10.1007/s10898-010-9598-5

Maturity, distance and density (MD^2) metrics for optimizing trust prediction for business intelligence

Muhammad Raza · Omar Khadeer Hussain · Farookh Khadeer Hussain · Elizabeth Chang

Received: 13 August 2010 / Accepted: 19 August 2010 / Published online: 2 September 2010

Abstract The modelling and management of trust between interacting parties are crucial parts of the overall business intelligence strategy for any organization. Predicting trust values is a key element of modelling and managing trust. It is of critical importance when the interaction is to be conducted at a future point in time. In the existing body of work, there are a few approaches for predicting trust. However, none of these approaches proposes a framework or methodology by which the predicted trust value can be considered in light of its accuracy or confidence level. This is a key element in order to ensure optimized trust prediction. In this paper, we propose a methodology to address this critical issue. The methodology comprises a suite of metrics—maturity, distance and density (MD^2) which are capable of capturing various aspects of the confidence level in the predicted trust value. The proposed methodology is exemplified with the help of case studies.

Keywords Trust · Trust prediction · Optimization · Prediction · Business intelligence · Optimized business intelligence

1 Introduction

The modelling of trust and reputation among business entities is of extreme importance for organizations. It forms a crucial part of the business intelligence strategy of any business

M. Raza · O. K. Hussain · F. K. Hussain (✉) · E. Chang
Digital Ecosystems and Business Intelligence Institute, Curtin University of Technology, Bentley, WA 6102, Australia
e-mail: farookh.hussain@cbs.curtin.edu.au

M. Raza
e-mail: muhammad.raza@postgrad.curtin.edu.au

O. K. Hussain
e-mail: o.hussain@cbs.curtin.edu.au

E. Chang
e-mail: elizabeth.chang@cbs.curtin.edu.au

organization. Given the importance of this research area, it has received a huge amount of research attention and correspondingly there is a plethora of literature on various aspects of modelling trust. A key element of modelling and managing trust is being able to accurately predict future trust values [1,4,5]. This is of particular importance when the trusting agent needs to make a decision about the trusted agent at a future point in time. The concept of *predicting* or *forecasting* values is not new. Different models and theories such as the Markov Model [6], Kalman Filter Theory [7], Holt-Winter forecasting method [10], Neural Networks [22], Bayesian Networks [25] etc, have been proposed in the literature for forecasting and have been widely used for various application domains such as energy forecasting, weather forecasting … etc.

Some of these approaches have also been used in predicting trust and reputation values. Hussain et al. [4] proposed one of the widely used approaches for trust prediction in the literature. Their approach leverages the previous trust values of an entity along with intelligent pattern-matching techniques and extrapolation techniques to determine the future trust value. An underlying assumption of this model is that the previous trust values will match certain pre-defined patterns (such as a line, curve etc...). An inherent shortcoming of this approach is that it fails to model the confidence level that one may associate with the predicted trust value. This is important in a dynamic environment, where the behaviour of the entity in question may not necessarily be static. Additionally, there might be some approximations or limitations made to the underlying trust values based on which the prediction process will be conducted. In order to account for these, we need to account for the confidence level of the predicted trust value. If the decisions regarding business interactions are made solely on the basis of predicted trust value with no consideration given to the associated confidence in the predicted value, this would result in non-optimal decisions and non-optimal business intelligence.

In order to ensure that optimal business decisions are made as a part of the overall business intelligence strategy, we need a mechanism whereby the predicted trust value is considered along with its associated confidence level. Similar to the optimization methods proposed in [39], the optimization process in this case would non-linear and multi-objective. The objective of this paper is to propose such a framework for optimal business intelligence.

This is achieved by proposing a set of metrics which are capable of capturing the confidence level in the predicted trust values, regardless of the prediction model being employed. These metrics are the *maturity*, *distance* and *density* abbreviated as MD^2. We propose *maturity* as a means of capturing the length of time that the entity in question has been in existence. We propose *distance* as a means of capturing how far into the future (in terms of time) the prediction process is to be carried out. We propose the *density* to capture the number or volume of trust values based on which the prediction would be carried out. Subsequently, based on these three metrics, we determine the confidence level (*clevel*) in the predicted trust value.

Finally, in order to enable optimized business intelligence, the decision regarding interaction is made by taking into account *both* the predicted trust value and the confidence level (*clevel*).

In Sect. 2, we provide an overview of the existing body of work on prediction in general and trust prediction in particular. In this section, we consider the various approaches that are used in the literature for prediction. In Sect. 3, we give formal definitions of the terms trust modelling, trust prediction, trust determination and confidence level. In Sect. 4, we offer a case study to illustrate in detail the problem that we are tackling in this paper. In Sect. 5 and Sect. 6, we introduce and discuss the proposed metrics for computing the confidence level. In Sect. 7, we demonstrate how our proposed model would result in optimized trust prediction. Finally, in Sect. 8 we conclude our work and outline future directions for research.

2 Prediction or forecasting

In this section, we present an overview of the existing body of work in predicting future values. The discussion in this section is divided into two parts. In Sect. 2.1, we examine the various techniques or approaches that have been used to predict future values. In Sect. 2.2, we examine the existing works on trust prediction and highlight their shortcomings.

2.1 Background

Forecasting or prediction is a key research challenge across a number of domains such as power forecasting [33,34], weather forecasting [35], stock market forecasting [30–32,36,37] etc. Various approaches have been proposed to predict future values. One of the most common approaches for predicting future values is the Markov model [6] which has been successfully used in Time series prediction [17]. There are other derivatives of the Markov model such as the Hybrid-order tree-like Markov model which is able to precisely predict user Web access, providing high coverage and good scalability [23]. Klepis et al. [38] have proposed an approach for predicting the three-dimensional structure of proteins.

Neural network-based approaches and models have also been used to predict values. Due to better accuracy and performance, neural networks are considered to be one of the most reliable methods for prediction. Drossu et al. [22] show the successful implementation of neural networks for time series prediction and address the effectiveness of neural networks in constructing approximations for unknown functions by learning. Liu et al. [19] have investigated the importance of aspects such as time delays between inputs and input dimensions for time series prediction using neural networks [19]. A neural network can capture any type of non-linear relationship between input and output data through training which makes it ideal for prediction in any domain such as time series prediction [18,20,21].

The use of Holt-Winter exponential smoothing has also been investigated in forecasting and prediction [10–12]. Bermudez et al. [24] proposed a generalized Holt-Winter exponential smoothing scheme for forecasting the future demand. An interesting aspect of their proposed method is that optimization is achieved by unifying the stages of estimation of the parameters and model selection. The forecasting accuracy in [24] is achieved through symmetric mean absolute percentage.

Besides the aforementioned approaches, other approaches such as Kalman Filter Theory [7,8] and Bayesian networks [25] have been used for predicting the future values.

It is important to note that all of these approaches propose a framework by which one can predict the future value. Making decisions based on just the predicted trust would result in non-optimal decisions. This is because, although the predicted values are based on factual past information or history, the predicted trust value is not a fact before the actual occurrence of the event. Additionally, there might be some patterns in the underlying history or past information that may influence (either positively or negatively) the reliability of the predicted value. Hence, in order to optimize the prediction process, one needs to know the level of confidence in the predicted value. By taking into account both the predicted value and the confidence level in the predicted value, one can ensure optimized decision making.

2.2 Trust prediction

One of the critical aspects of trust management and trust modelling is the prediction of the trust value at a future point in time. Ma et al. [14] focus on trust relationship prediction that is considered as a classification problem in which a trustor-trustee pair is assigned either a *trust* or *no-trust* label on the basis of user-generated features. Similar to this is the case in [13,15,26,27] where the "trust existence" is predicted between two users. The objective of these approaches is to determine the existence of indirect trust relationships between two entities in a social network, based on the existence of direct trust relationship between some other entities in the network. Trust prediction based on similarity measures [28] also focuses on the existence of trust among entities which follows the *web of trust* model.

Considerable work on trust prediction can be found in [29] who use the fuzzy regression-based approach. In a service-oriented environment, a *service trust vector* is proposed which consists of a set of values. The three components of the proposed vector are *final trust level*, *service trust trend* and *service performance consistency level*. Since prediction is carried out for the future, the service trust gives an indication of the future Quality of Service. The consistency level shows the level of quality maintained by the entity.

As discussed earlier, the Kalman filter is one of the approaches frequently used for forecasting and prediction. The Kalman filter is also used for trust prediction in pervasive systems [9]. Being a set of recursive mathematical equations, the Kalman filter is suitable for pervasive systems due to its lightweight nature, in terms of both memory and computation load [9]. Provided with a set of previous observations t, the Kalman filter can predict the value for $t + 1$.

The Markov chain is another approach for carrying out the prediction process. Hussain et al. [4] have proposed the use of discrete Markov Chains in order to predict future trust and reputation values. Their proposed model is unique in the sense that it can model three different kinds of data: data with seasonal variations, data with trend, and random data cases (noise).

From the above discussion, it is obvious that there is some existing research on predicting the future trust value of an entity. The existing literature on trust prediction can be grouped into two broad classes. One class of the work on prediction is actually on the prediction of "existence of trust" among entities. In other words, it focuses on determining the existence of trust between two entities in a social network. The other class of the trust prediction work is on the prediction of "trust values" either qualitatively and quantitatively in a future time spot. In this study, our focus is on the latter.

The major shortcoming of all the existing work on trust prediction is that all of them propose a framework for predicting only trust values. None of the existing approaches proposes a framework or a methodology by which one can determine the confidence level in the predicted trust value. It is crucial that, for optimized business intelligence and for optimized trust prediction, we be able to consider the predicted trust value in light of the confidence level in the predicted value. Failure to do so would result in non-optimal business intelligence and non-optimal selection of business partners. This paper is a first step in that direction. In this paper, we propose a framework by which the trusting agent can determine the confidence level in the predicted trust value (Tables 1, 2, 3).

Table 1 Trust values of ABC Transportation Company

Time spot (or year)	1999	2002	2003	2004	2005	2006	2007	2008	2009	Predicted trust value for 2012
Trust value	3	5	5	5	5	5	5	5	5	5

Table 2 Maturity level and its associated quantitative and qualitative expressions

|tc – tf|	Qualitative representation of *m* metric	Quantitative representation of *m* metric
Less than ds	Max	3
Equal to ds	Normal	2
Greater than ds	Min	1

Table 3 Quantitative and qualitative representation of distance (*ds*) metric

$(tp - tc)/m$	Qualititative representation of *ds* metric	Assigned quantitative value to ds metric for computing confidence level (*clevel*)
Less than 1	Min	3
Equal to 1	Normal	2
Greater than 1	Max	1

3 Formal definitions of trust modeling, trust prediction and trust determination

Comprehensive work on defining and managing trust and reputation can be found in [1–3,5,16]. However, the existing literature does not provide a distinction between the terms '*trust modelling*' and '*trust prediction*'. It is important to note that the concepts of trust modelling and trust prediction are different from each other and the primary distinctive difference between these concepts is the time dimension. In this section, we clearly define the concepts of '*trust modelling*', '*trust prediction*' and '*trust determination*'.

3.1 Trust

In our work, we adopt the definition of trust proposed by Chang et al. [1], which states that trust is *"the belief the trusting agent has in the trusted agent's willingness and capability to deliver a mutually agreed service in a given context and in a given time slot"*.

3.2 Trust modeling

We define trust modelling as *"The process of determining the trust value of an entity either quantitatively or qualitatively"*.

Mathematically, trust modelling can be expressed as follows:

$$\mathrm{T} = \mathrm{f}\,(\text{Entity A}) \tag{1}$$

where;

'*Entity A*' denotes the identity of the entity whose trust value is being modelled

'T' denotes the trust value of Entity A

'$f\,()$' denotes the mathematical function for modelling the trust value of Entity A

We classify trust modelling into two broad categories based on the time at which the trust modelling is carried out. These categories of trust modelling are *trust prediction* and *trust determination*. Their formal definitions are presented below.

3.3 Trust determination

We define trust determination as *"The procedure of expressing the trust value of an entity, either quantitatively or qualitatively, either in the past or during the current point in time"*.

The process of trust determination may result either in a quantitative or qualitative expression of the trust value. Mathematically, trust modelling can be expressed as follows:

$$\mathrm{T} = \mathrm{g}\,(\text{Entity A, t}) \tag{2}$$

where;

'*Entity A*' denotes the identity of the entity whose trust value is being determined

'T' denotes the *determined* trust value of Entity A

'$g()$' denotes the mathematical function for determining the trust value of Entity A

't' denotes the point in time (either in the past or during the current time) for which the trust value has been determined.

3.4 Trust prediction

We define trust prediction as *"The process of determining the trust value of an entity, either quantitatively or qualitatively, at a future point in time"*.

Similar to trust determination, this process may result either in a quantitative or qualitative expression for the predicted trust value. Mathematically, trust prediction can be expressed as follows:

$$\mathrm{T}_{\mathrm{future}} = \mathrm{s}\,(\text{Entity A, } \mathrm{t}_{\mathrm{future}}) \tag{3}$$

where;

'*Entity A*' denotes the identity of the entity whose trust value is being predicted

'T_{future}' denotes the trust value of Entity A at a point in time in the future

'$s()$' denotes a mathematical function for predicting the trust value of Entity A

'$\mathrm{t}_{\mathrm{future}}$' denotes the point in time in future for which the prediction process is being carried out.

3.5 Confidence

We define confidence as *"The process of determining and expressing the accuracy or the reliability of the trust level"*.

Mathematically, confidence can be expressed as follows:

$$\mathrm{c} = \mathrm{r}\,(\text{Entity A,}\mathrm{T}_{\mathrm{future}}) \tag{4}$$

where;

'*Entity A*' denotes the identity of the entity whose trust value is being predicted

'T_{future}' denotes the trust value of Entity A at a point in time in the future

'$r()$' denotes a mathematical function for determining the reliability of the predicted trust value of Entity A

'c' denotes the confidence in the predicted trust value.

In this paper, the confidence level *(clevel)* is expressed as a real number between the [0–1], where 0 corresponds to 'no confidence' and 1 corresponds to 'full confidence'. Other values of the confidence level are expressed as real numbers in the range [0–1].

It is important to note that the confidence value or level may be associated with a trust level, which could have been the outcome of either the trust determination process or the trust prediction process. In this paper, we are concerned only with the confidence level of the predicted trust value.

For optimized business intelligence, it is important that we consider the predicted trust value (or the determined trust value) in combination with the corresponding confidence level. In any given interaction, trust technology is used as a means of determining the reliability of an unknown interacting party (say B). It relies on modelling the previous or existing behaviour of entity B (as reported by other parties). However, the reliability of the underlying data or the underlying behaviour should be modelled in order to enable optimized business intelligence.

This paper is a first step in that direction. In this paper, we propose a set of metrics that can be used to determine the reliability of the underlying data on which the process of trust prediction is being carried out.

4 Case study

In order to further elucidate the research issue discussed in the previous sections, let us consider the following case study:

Customer A has to move his personal goods from Perth to Sydney in 2012. He previously made use of ABC Transportation Company in 2007, when he moved from Adelaide to Perth. At that point in time, he had found ABC Transportation Company to be trustworthy and had assigned it a trust value of 5.

In order to assist him with his decision-making process, Customer A issues a reputation query about the trusted entity, gathers reputation values about the trusted entity and determines the following trust values for the ABC Transportation Company.

Using one the trust prediction methods, Customer A determines the trust value of the trusted entity to be 5 in 2012.

The issue for Customer A is to determine the confidence level in the predicted trust value for 2012. The confidence level could be determined based on the patterns of available past trust information such as the following:

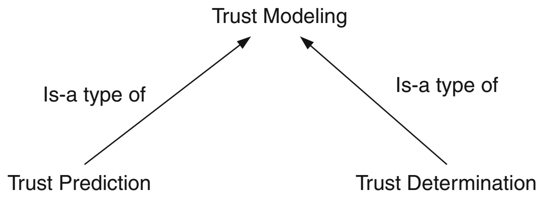
(a) Duration (of time) for which the trusted entity has been in operation
(b) How far (as a function of time) into the future the prediction process is being carried out (assuming the current time spot or year to be 2010)
(c) The pattern of the distribution of trust values in the past.

These patterns can directly impact on the predicted trust value. As a result of this, it is extremely important to capture these patterns of the underlying trust information as a part of the computed confidence level. Customer A would then be able to make a trust-based decision regarding interaction, taking into account both the predicted trust value and its corresponding confidence level. In the following section, we present a methodological framework by which this can be achieved (Fig. 1).

5 MD^2 metrics for generating confidence level

As mentioned in the previous sections, for optimized business intelligence, the predicted trust value should be considered along with its associated confidence level. The confidence level

Fig. 1 Relationship between Trust Modelling, Trust Prediction and Trust Determination



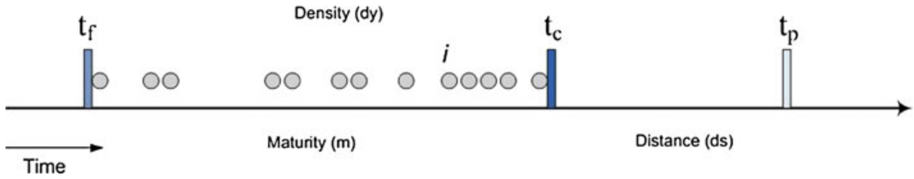


Fig. 2 Pictorial representation of the three confidence metrics

in the predicted trust value would depend on various aspects or attributes of the available trust history. In the following section, we present and discuss the attributes and the impact that they would have on the confidence of the predicted trust value. These attributes are: (a) *Maturity* (b) *Distance* and (c) *Density* (Fig. 2).

5.1 Maturity

One of the metrics for calculating the confidence level in the predicted trust values is *'Maturity'*. We define maturity (m) of an entity as *"the total life span of an entity for which it has been in existence"*. This metric captures and expresses the duration of time for which the entity in question has been operational. Mathematically, it can be computed as the difference between the current time spot (tc) and the time spot for the first interaction (tf), represented as follows:

$$m = \mid tc - tf \mid \tag{5}$$

The next step is the scaling of the maturity level. In order to do this, we compare the numerical value of the maturity metric with the numerical value of the distance metric. The distance metric (explained in the following section), captures and expresses how far into the future the prediction is carried out. It is computed as the difference between the current time spot and the predicted time spot.

The maturity value (m) has a great impact on the confidence level of the predicted trust value. We believe that the greater the maturity of the entity, the higher would be the confidence in the predicted value and vice versa. The reason for this is that the higher the maturity value, the more information we would have at our disposal to model the behaviour of the entity; hence, we would be able to predict the future value more accurately. Mathematically, this relation can be expressed as follows:

$$clevel \propto m \tag{6}$$

The algorithmic formulation of the maturity level is shown in Fig. 3.

As shown in the above algorithmic formulation of the maturity computation metric, if the time difference between the current time spot (tc) and the first time spot (tf) is greater than the distance metric, then qualitatively, we consider that the maturity value of the entity in

```
m = maturity level
tc = current time spot
tp = predicted time spot
tf = time spot for first interaction
ds = |tp - tc|

If |tc – tf| > ds then
        m = max
else if |tc – tf| = ds then
        m = normal
else
        m = min
End If
```

Fig. 3 Qualitative formulation of maturity metric

question is at maximum. In such a case, quantitatively, we consider that the maturity value of the entity in question is 3. This is due to the fact that in this scenario, the entity in question has been in existence for a longer duration of time, relative to the predicted time spot. This is because in such a scenario, ignoring other factors, there would be a large number of data points based on which the prediction process can be conducted. An alternative way to explain this would be as follows: assuming one interaction per time spot, in the above scenario, there would be a large number of trust values on which the prediction would be based.

Alternatively, if the time difference between current time spot*(tc)* and the first time spot*(tf)* is equal to the distance metric, then qualitatively, we consider the maturity value of the entity in question as *normal.* In this scenario, quantitatively, we consider that the maturity value of the entity in question is 2.

However, if the time difference between the current time spot*(tc)* and the first time spot*(tf)* is less than the distance metric, then qualitatively we consider the maturity value of the entity in question as *low*. The corresponding quantitative expression would be 1. This is because, again ignoring other factors, in this case the prediction is being carried out based on a small number of data points and correspondingly the confidence level in the predicted trust value would be *low.*

Referring to the case study presented in Sect. 3, we find that the length of time between the current time spot and the first time spot is 11 years. This length of time is much greater than the length of time between the current time spot and the predicted time spot (which is 2 years). Assuming a distribution of one trust value per time spot, we can see in this case that there is a large amount of trust information or trust history (relative to that present after the current time spot) based on which the prediction process would be conducted. Based on just this factor, and ignoring other factors used for confidence level computation, the qualitative representation of the *maturity (m)* metric would be *Max* and its corresponding quantitative representation would be 3.

5.2 Distance

The second metric in confidence level calculation is the *Distance* metric. We define distance (*ds*) as *"a metric that expresses how far into the future the prediction process is being carried*

```
Let ds denote the distance metric
tc = current time spot
tp = predicted time spot
tf = time spot for first interaction
m = |tc – tf|

If    (|tp – tc|/ m) > 1 then
            ds = max
else if    (|tp – tc|/ m) = 1 then
            ds = normal
else
            ds = min
End if
```

Fig. 4 Algorithmic formulation of distance level

out". The distance (*ds*) between current time spot *(tc)* and the time spot at which the prediction is to be carried out (*tp*) would have a great impact on the confidence level of the predicted trust value. Ignoring other factors, the farther the predicted time spot (*tp*) is from the current time spot (*tc*), the less would be the confidence in the predicted trust value and vice versa. In order to quantify the distance metric, we express it relative to the maturity *m* of the trusted entity.

Mathematically, the distance metric can be computed as follows:

$$ds = \frac{tp - tc}{m} \tag{7}$$

Mathematically, the relationship between the confidence level (*clevel*) and the distance metric can be expressed as follows:

$$clevel \propto \frac{1}{ds} \tag{8}$$

As shown above in Fig. 4, if the numerical value of (|tp – tc|/m) is equal to 1, then qualitatively the value of the *ds* metric will be *Normal* or vice versa. This means that, ignoring other factors, if the distance between the time spot at which the prediction is to be carried out (or predicted time spot) and the current time spot is equal to the maturity of the trusted entity, we can be fairly confident about the predicted trust value.

Similarly, if the numerical value of (|tp – tc|/m) is less than 1, then qualitatively the value of the *ds* metric would be *Min*. This means that if the distance between the predicted time spot (*tp*) and the current time spot (*tc*) is less than the maturity *m* of the trusted entity, then we can be highly confident about the predicted trust value (ignoring other factors). This is due to the fact that the length of the time difference between the prediction time spot and the current time slots is less than the maturity of the trusted entity.

Similarly, if the numerical value of (|tp – tc|/m) is greater than 1, then we qualitatively assign the value of the *ds* metric as *Max*. This means that, if the distance between the predicted time spot (*tp*) and the current time spot (*tc*) is more than the maturity *m* of the trusted entity, then we will have little confidence in the predicted trust value.

Continuing with the case study presented in Sect. 3, we find that the length of time between the current time spot and the predicted time spot is two years. This length of time is much less than the length of time between the current time spot and the first time spot (which is

```
dy = density
tc = current time spot
tp = predicted time spot
tm=median time spot
tf = time spot for first interaction
tm = (tc + tf)/2
n(tf - tm) = number of interactions between tf and tm
n(tm - tc) = number of interactions between tm and tc

If n(tm - tc) > n(tf - tm) then
        dy = MRI
else if n(tm - tc) = n(tf - tm) then
        dy = Normal
else
        dy = LRI
End if
```

Fig. 5 Qualitative computation of density metric

11 years). Since the prediction process is being carried out at a point in time which is not as far into the future relative to the time for which we have trust history or information, the qualitative representation of the *distance (ds)* metric would be *Max* and its corresponding quantitative representation would be 3 (Fig. 5).

5.3 Density

The third metric used for the calculation of the confidence level in the predicted trust value is known as Density (*dy*). We define the density or the distribution of the previous interactions within the maturity of the trusted entity as *"the frequency of the various interaction values recorded at different time spots over its life span"*.

The purpose of the density metric is to capture or express the distribution of the number of trust values of a given entity over the maturity.

We consider three different types of distributions of the trust values over the maturity time space as follows:

- LRI (Least Recent Interactions): In this type of distribution, the density of trust values of the trusted entity is greater in the first half of the maturity time space than in the second half of the maturity time space. In this scenario, the confidence level of the predicted trust value of the trusted entity would be *low* as we do not have more recent interactions in order to model the behaviour of the trusted entity. In this scenario, quantitatively the density *(dy)* metric is represented as *low* and its corresponding quantitative expression is 1.
- MRI (Most Recent Interactions): In this type of distribution, the density of trust values of the trusted entity is greater in the second half of the maturity time space than in the first half of the maturity time space. In this scenario, the confidence level of the predicted trust value of the trusted entity would be *High* as the prediction process is being carried out based on recent behaviours. In contrast to the above scenario, in this case, quantitatively the density (*dy*) metric is represented as *high* and its corresponding quantitative expression is 3.
- Normally Distributed: In this type of distribution, the density of trust values of the trusted entity is evenly distributed throughout the maturity time space. In this scenario, the confidence level of the predicted trust value of the trusted entity would be less than that in the

Table 4 Qualitative and quantitative expression of density metric

Distribution of trust values	Qualititative representation of density (*dy*) metric	Assigned quantitative value to *dy* metric for computing confidence level (*clevel*)
LRI	Low	1
Normally distributed	Medium	2
MRI	High	3

scenario of MRI and more than that of LRI. In this scenario, quantitatively the density (*dy*) metric is represented as *medium* and its corresponding quantitative expression is 2.

The density levels along with their corresponding qualitative and quantitative expressions are shown in Table 4.

Continuing with the above case study presented in Sect. 3, we find that the median time spot (tm) for the previous information or history occurs in 2005. The number or volume of trust information available between the first time spot (1999) and the median time spot is 4 (excluding the median time spot). The amount or volume of trust information available between the median time spot (including the median time spot) and the current time spot is 5. In this scenario, a greater amount of trust history or information is available to model the recent behaviour of the trusted entity, which would result in a more reliable predicted value (ignoring other factors), which we consider to be the most recent interactions.

6 Confidence level

In order to calculate the confidence level in the predicted trust value of an entity, we consider the summation of the above proposed three metrics (maturity, distance and density), relative to the case of summation which would correspond to the highest possible level of confidence.

The confidence level (*clevel*) in the predicted trust value is computed as follows:

$$clevel = \frac{m + dy + ds}{\text{highest possible level of confidence}} \tag{9}$$

Equation 9 may be re-written as follows:

$$clevel = \frac{m + dy + ds}{\text{max(m)} + \text{max(dy)} + \text{max(ds)}} \tag{10}$$

Where,

'm' corresponds to the quantitative value of maturity

'dy' corresponds to the quantitative value of density

'ds' corresponds to the quantitative value of distance

'max(m), 'max(dy)' and 'max(ds)' corresponds to the quantitative value of maturity, density and distance correspondingly which would lead to the highest possible level of confidence

It is important to note that the purpose of this research is not to propose guidelines or devise a framework for the acceptability of the '*clevel*' values for optimized trust prediction. On the contrary, the objective of this research is to give the trusting agent (or the domain expert) an insight into the confidence level that can be associated with the predicted trust value. Depending on (a) the domain of the business interaction, and (b) the business value of the interaction, different levels of confidence values (*clevel)* could be considered acceptable. It is not possible to establish a definite benchmark value for *clevel* as an acceptable standard across various domains and across varying levels of the monetary value of the interaction. For example, in the medical domain, one would like the confidence level in the predicted quality of service delivered by a doctor or a hospital to be as high as possible (95% or higher confidence level), since the lives of people may be at stake. Conversely, if we are trying to predict the trust value of a logistic service provider who would move goods worth AUD 100 from Sydney to Perth, then we do not necessarily require such a high degree of confidence level in the predicted trust value [40,41].

7 Decisions based on optimized trust values

The computed confidence level *(clevel)* can be considered along with the predicted trust value as shown in Eq. 11.

$$OTV = ATV \otimes clevel \tag{11}$$

where;

ATV (Actual trust value) is the predicted trust value

clevel is the computed confidence level for ATV

OTV is the optimized trust value taking into account the corresponding *clevel*

$\otimes$is a metric for weighting the actual trust value (ATV) with the confidence level *(clevel)to determine the optimized trust value* (OTV). In this paper, for simplicity purposes, we use this metric as the multiplication operator leading to Eq. 12 as shown in Eq. 12.

$$OTV = ATV^{*}clevel \tag{12}$$

If only the discrete trust values in the range [1–6] are being considered, then Eq. 12, may be re-written as follows:

$$OTV = Round(ATV^{*}clevel) \tag{13}$$

Continuing with the above example of the interaction between Customer A and ABC Transportation Company, the confidence level computed using (10) would be 1. Corresponding to a confidence level of 1, the optimized trust value computed using (12) would be 5. Based on the computed optimized trust value, Customer A would be able to make an interaction-based decision with the trusted entity.

In order to further elucidate our proposed methodology, let us assume that Customer B has to decide which one of the following three software firms—SEC1, SEC2 and SEC3—it would like to outsource. Customer B has found out that the actual predicted value of all of these three firms is 5. Let us furthermore assume that the values for the maturity, distance and density metrics for these three business entities are as shown in Table 5.

It is interesting to note from Table 5, that the actual predicted trust value of SEC1, SEC2 and SEC3 is 5. If the trusting agent is making a decision based on just the predicted trust value, it might result in the selection of SEC2. However, as we can see from Table 5, there is little trust information available for SEC 2 (corresponding to maturity level of 1) and the prediction process is being conducted at a time farther into the future, compared with the duration of time for which there is available trust information for SEC2. The selection of SEC2 as the most reliable interacting partner of the three available interacting partners would

Table 5 Quantitative expressions and confidence levels for SEC1, SEC2 and SEC3

	Quantitative expression for maturity metric	Quantitative expression for distance metric	Quantitative expression for density metric	Predicted trust value	Computed confidence level
SEC1	3	2	3	5	0.88
SEC2	1	1	3	5	0.55
SEC3	2	3	1	5	0.66

clearly result in a non-optimal business decision, particularly when there are other business entities such as SEC 1 available for interaction.

However, based on the metrics presented in this paper, the optimized trust values for SEC1, SEC2 and SEC 3 would be 4, 3 and 3 respectively. This would result in the selection of SEC1 as the interacting partner as, of the three business entities, it has the highest optimized trust value. As we can see from the above discussion, the optimized trust value takes into account the confidence level of the predicted trust value. Hence, an optimal selection (SEC1) is made.

8 Conclusion and future work

Besides trust modelling and trust management, trust prediction is also one of the crucial challenges in business intelligence. However, in the existing literature, the process of carrying out the trust prediction is not optimized as no work takes into account the confidence level in the predicted trust value. This is a major issue in business decision making [40,41]. In this paper, we have proposed a framework by which the trusting agent can quantify and determine the confidence level in the predicted trust value. Here, we propose three metrics, namely, *maturity*, *distance* and *density*, by which the confidence value can be generated. These metrics inherently model and represent the reliability of the underlying data used for prediction purposes in an open and distributed environment. Taking into account the generated confidence level and the predicted trust value, the trusting agent can compute the actual trust value for optimized trust prediction and optimized business intelligence. Our proposed model is simple in structure and can easily be implemented on top of any trust prediction model. In order to demonstrate the effectiveness of our methodology, we illustrate a case study where a trusting agent makes use of optimized trust values in contrast to predicted trust values.

Our future work involves investigating the following aspects: (a) deployment of this model in a real business environment, and (b) development of a fuzzy-logic framework or a methodology by which the trusting agent can factor both the predicted trust value and the generated confidence level to determine the actual trust value.

References

1. Chang, E., Dillon, T., Hussain, F.: Trust and Reputation for Service-Oriented Environments: Technologies for Building Business Intelligence and Consumer Confidence. Wiley, West Sussex (2006)
2. Abdul-Rahman, A.: The PGP trust model. J. Electron. Commerce (1997)
3. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: HICSS 2000, pp. 1769–1777. (2000)

4. Hussain, F.K., Chang, E., Hussain, O.: A robust methodology for prediction of trust and reputation values. In: ACM Workshop on Secure Web Services (SWS 2008) with 15th ACM Conference on Computer and Communications Security, pp. 97–108. ACM, Virginia (2008)
5. Hussain, F., Chang, E., Dillon, T.: Markov model for managing dynamic trust. In: INDIN 2005: 3rd International Conference on Industrial Informatics, Frontier Technologies for the Future of Industry and Business, pp. 725–733. IEEE, Perth (2005)
6. Hassan, M.R., Nath, B.: StockMarket forecasting using hidden Markov model: a new approach. In: Proceedings of the 5th International Conference on Intelligent Systems Design and Applications, pp. 192–194. Wroclaw, Poland (2005)
7. Kalman, R.E.: A new approach to linear filtering and prediction problems. Trans. ASME J. Basic Eng. **82**(Series D), 35–45 (1960)
8. Brockwell, P.J., Davis, R.A.: Introduction to Time Series and Forecasting. Springer, Berlin (2002)
9. Capra, L., Musolesi, M.: Autonomic trust prediction for pervasive systems. In: Proceedings of the 20th International Conference on Advanced Information Networking and Applications, Vienna, Austria, pp. 481–488 (2006)
10. Holt, C.C.: Forecasting seasonals and trends by exponentially weighted moving averages. Int. J. Forecast. **20**, 5–10 (2004)
11. Kotsialos, A., Papageorgiou, M., poulimenos, A.: Long-term sales forecasting using Holt–Winters and neural network methods. J. Forecast. **24**, 353–368 (2005)
12. Croux, C., Gelper, S., Fried, R.: Computational aspects of robust Holt-Winters smoothing based on M -estimation. Appl. Math. **53**, 163–176 (2008)
13. Matsuo, Y., Yamamoto, H.: Community gravity: measuring bidirectional effects by trust and rating on online social networks. In: Proceedings of the 18th International Conference on World Wide Web. Madrid, Spain, pp. 751–760 (2009)
14. Ma, N.: Trust relationship prediction using online product review data. In: CNIKM 2009. ACM. Hong Kong, China, pp. 47–54 (2009)
15. Carbo, J., Garcia, J., Molina, J.M.: Subjective trust inferred by Kalman filtering vs. a fuzzy reputation. In: Proceedings of 23rd International Conference on Conceptual Modeling, Shanghai, China, vol 3289, pp. 496–505 (2004)
16. Hussain, F.K.: A methodology for trust management in service oriented environment. PhD Thesis, Curtin University of Technology, Australia (2006)
17. Wang, X., Whirgham, P., Deng, D.: Time-line hidden Markov experts for time series prediction. In: Proceedings of Neural Networks and Signal Processing. IEEE Computer Society, pp. 786–789. (2003)
18. Shiblee, M., Kalra, P.K., Chandra, B.: Time Series Prediction with Multilayer Perceptron (MLP): A New Generalized Error Based Approach. Lecture Notes in Computer Science, 5507. Springer, Berlin/Heidelberg (2009)
19. Liu, F., Quek, C., Ng. G.S.: Neural network model for time series prediction by reinforcement learning. In: IEEE International Joint Conference on Neural Networks. IEEE Computer Society. Montreal, Canada (2005)
20. Coyle, D., Prasad, G., McGinnity, T.M.: A time-series prediction approach for feature extraction in a brain–computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. **13**(4), 461–467 (2005)
21. Sfetsos, A., Siriopoulos, C.: Combinatorial time series forecasting based on clustering algorithms and neural networks. Neural Comput. Appl. **13**(1), 56–64 (2004)
22. Drossu, R., Obradovic, Z.: Rapid design of neural networks for time series prediction. IEEE Comput. Sci. Eng. **3**(2), 78–89 (1996)
23. Dongshan, X., Junyi, S.: A new Markov model for Web access prediction. Comput. Sci. Eng. **4**(6), 34–39 (2002)
24. Bermúdez, J.D., Segura, J.V., Vercher, E.: SIOPRED: a prediction and optimisation integrated system for demand. TOP **16**(2), 258–271 (2008)
25. Hou, L., Wang, L., Yang, J.: Evolutionary prediction of online keywords bidding. In: E-Commerce and Web Technologies, pp. 124–133. Springer, Berlin/Heidelberg (2008)
26. Borzymek, P., Sydow, M., Wierzbicki, A.: Enriching trust prediction model in social network with user rating similarity. In: CASON 2009, pp. 40–47. IEEE Computer Society, Fontainebleau, France (2009)
27. Liu, H.: Predicting trusts among users of online communities—an Epinions case study. In: ACM Conference on Electronic Commerce, pp. 310–311. ACM, Chicago, IL, USA (2008)
28. Skopik, F., Schall, D., Dustdar, S.: Start trusting strangers? Bootstrapping and prediction of trust. In: WISE 2009, pp. 275–289 Springer, Berlin (2009)
29. Li, L., Wang, Y., Varadharajan, V.: Fuzzy regression based trust prediction in service-oriented applications. In: Autonomic and Trusted Computing 2009, pp. 221–235. Springer, Berlin (2009)

30. Hassan, M.R., Nath, B.: StockMarket forecasting using hidden Markov model: a new approach. In: Proceedings of the 5th International Conference on Intelligent Systems Design and Applications, pp. 192–196 (2005)
31. Mańdziuk, J., Jaruszewicz, M.: Neuro-evolutionary approach to stock market prediction. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2007), Orlanodo, FL, USA, pp. 2515–2520. IEEE Press, Los Alamitos (2007)
32. Hu, C., He, L.T.: An application of interval methods to stock market forecasting. Reliable Comput. **13**(5), 423–434 (2007)
33. Al-Hamadi, H.M., Soliman, S.A.: Fuzzy short-term electric load forecasting using Kalman filter[J]. IEEE Proc.Gener. Trans. Distrib. **153**(2), 217–227 (2006)
34. Taylor, J.W., McSharry, P.E.: Short-term load forecasting methods: an evaluation based on European data. IEEE Trans. Power Syst. **22**, 2213–2219 (2007)
35. Lee, R., Liu, J.: iJADE WeatherMAN : a weather forecasting system using intelligent multiagent-based fuzzy neuro network. IEEE Trans. Syst. Man Cybern. C **34**, 369–377 (2004)
36. Mockus, J.: On simulation of optimal strategies and Nash equilibrium in the financial market context. J. Global Optim. ISSN 1573-2916 (2009)
37. Nenortaite, J., Simutis, R.: Development and evaluation of decision-making model for stock markets. J. Global Optim. **36**(1), 1–19 (2006)
38. Klepeis, J.L., Floudas, C.A.: Ab initio tertiary structure prediction of proteins. J. Global Optim. **25**, 113–140 (2003)
39. Chinchuluun, A., Pardalos, P., Migdalas, A., Pitsoulis, L. (eds.): Pareto Optimality, Game Theory and Equilibria. Springer, Berlin (2008)
40. Zopounidis, K., Pardalos, P.M., Baourakis, G. (eds.): Fuzzy systems in management, economy and marketing. World Scientific, Singapore (2001)
41. Pardalos, P.M., Resende, M. (eds.): Handbook of Applied Optimization. Oxford University Press, Oxford (2002)