Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **4**, July **2016**

Research on Credibility Measurement Method of Data In Big Data

Xiao-Rong Cheng[1], Tian-Qi Li[1,†], Si-Zu Hou[2], Hui-Lan Zhao[1], Hong-Yu Wang[3]

[1]School of control and computer engineering,
North China Electric Power University, Hebei Baoding, 071003,China.
†Correspondence Autor
ltqwhy@163.com

[2]Department of Electronic & Communication Engineering,
North China Electric Power University, Hebei Baoding, 071003,China.

[3]College of information technology,
Hebei Normal University,Hebei Shijiazhuang, 050024,China.

Received November, 2015; revised May, 2016

Abstract. *On the basis of analytical theory of credible traditional data, this paper constructs a hierarchical trusted network model of dynamic data analysis by increasing time factor, penalty factor and other weight parameters, which puts the credibility of the analysis of the problem down to the big data combinatorial problems of between the data sources, data source and data dissemination network path. That is, by calculating the credibility between data sources, the credibility of data and data source, the paper dynamically constructs the Big Data networks of the trusted analysis through analytical data calculating the overall credibility of data. The simulation results show that the model can satisfy the requirement of the credibility of Big Data better, and provide the ideas to solve the problem of the credibility measure for prospective research is feasible.*
Keywords: Big Data; Credibility; Dynamic; Trusted Computing; Modeling and Simulation

1. **Introduction.** At present, Big Data has been a hot research topic in the academic filed. It has a typical "4V" characteristic[1]: - Volume, Velocity, Variety, Value. At the same time, it is not difficult to find the typical "HDC" attribute of Big Data, that is, Heterogeneity, Dynamic [2], Complexity [3]. Therefore, the large dataset must be filled with a large number of unreliable data, which brings a lot of risk to the users of data. If we can evaluate the credibility of the original data, we can reduce the risk effectively and improve the credibility of Big Data.

As for Big Data, people need to be more urgent and study the trusted measurements and evaluation methods of the new situation of Big Data. Therefore, the paper proposes a new model based on hierarchical model of Big Datas credibility.

2. Credibility Analysis of Big Data.

2.1. Research on data credibility.

At present, there are many research methods and some results for the trusted measurement and evaluation of data. The methods of credibility analysis are mainly divided into two categories, one is subjective trust analysis based on belief, which is a cognitive phenomenon which is the subjective judgment of the specific characteristics or behavior of the object of trust, and this kind of judgment, which

has ambiguity, uncertainty and cant be accurately described, verified and speculated, is relatively independent of the subject's characteristics and behavior [5]. Documents [5, 6, 7, 8] have proposed different subjective methods based on Probability, Fuzzy Set Theory, Cloud Theory and so on. The other is the objective trust analysis based on the evidence theory, which can be accurately described, verified and speculated. The trust relationship between the two is strictly defined by appropriate evidence. D-S evidence theory is used to calculate the credibility by documents [9, 10, 11].

Although the above research methods and classical algorithms make some contribution to the trusted measurement of common data, however, data has the typical "4V" and "HDC" attribute in the era of Big Data, which determines that the large data has the characteristics of multi-source, heterogeneous, spatio-temporal, social and high noise, which makes the methods of traditional data analysis not meeting the needs of Big Data. Most trusted computing models consider that only a part of the decision attributes, not comprehensive. In the computation of comprehensive trust, they only consider the simple weighted average of direct trust and indirect trust, and neglect the influence of environmental context, which leads not well to describing the complexity and uncertainty of trust relationship. In the process of modeling based on probability statistics, as a part of the assumption, they are more subjective, and the accuracy of forecasting results and the scientific nature of the trust decision are affected. Although the models take the dynamic interaction and randomness between the entities into account, it can not consider the impact of timeliness and malicious recommendation, and lack of flexibility. Once the weight is determined, the system is very difficult to adjust it, it will result in a lack of adaptive prediction model. Therefore, it is urgent to study the problems of the credibility measure and service of Big Data.

2.2. **Credibility description for Big Data.** Aiming at the characteristics of Big Datas "4V" and "HDC", this paper presents a model of Big Datas credibility measure of dynamic construction, which is divided into three parts - the trusted measurement model of between data sources, the trusted measurement model of data sources and the trusted measurement model of data. The credibility of between data sources is restricted by the credibility of data sources, the credibility of the data sources is restricted by the credibility of data and the credibility of the data source, the credibility of data is restricted by the credibility of between data sources and the credibility of data sources, they are interrelated and restricted each other and constitute a whole. The relationship is shown in Figure 1.

In order to understand the model in Figure 1, the relevant definitions of the method are given in this paper, which are used to explain the basic issues in the analysis of Big Datas credibility.

Data source: It refers to the provider of data in the Big Datas environment.

Data: It refers to the characteristics of multiple attributes. Its notation is denoted as Data=$\{d_1,d_2,,d_n\}$. Thereinto, d_i refers to the "i" attribute of data.

Trusted network: It refers to a network composed of data sources and directed links between them.

Definition 1. Credibility of between data sources: It is composed of the local credibility and the global credibility of between data sources. Its notation is denoted as $Trust_A(B,t)$, the meaning behind which is the comprehensive credibility of data source A relative to data source B at the "t" moment.

Definition 2. Local credibility: When there is a direct context interaction between data sources or the similarity of data or behaviors provided by between data sources exceeds a certain threshold, we believe that between data sources have a local credibility. It

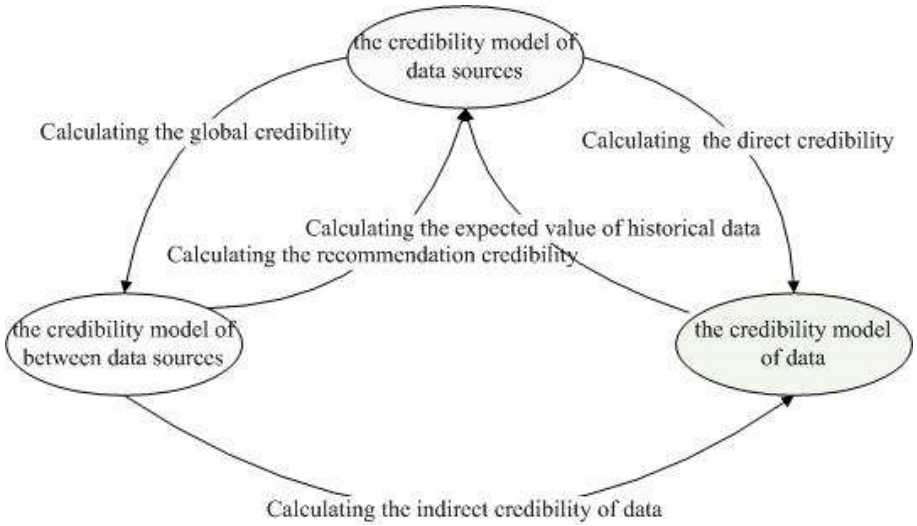

FIGURE 1. The correlation of the credibility model of between data sources, data, data source

is composed of the credibility of direct context interaction and the credibility of the similarity between data sources. Its notation is denoted as $LocalTrust_A(B,t)$, the meaning behind which is the local credibility of data source A relative to data source B at the "t" moment.

Definition 3. Global credibility: It refers to the credibility of data source in the trusted network, that is, the credibility of data source. Its notation is denoted as $GlobalTrust_A(B,t)$, the meaning behind which is the global credibility of data source A relative to data source B at the "t" moment.

Definition 4. Credibility of data source: It is composed of the expected value of the credibility of all historical data provided by data source and the recommendation credibility of data sources of each layer in the whole trusted network. Its notation is denoted as $Trust(A,t)$, the meaning behind which is the credibility of data source A at the "t" moment.

Definition 5. Recommendation credibility: It refers to the credibility of data source relative to the best path to the objective data source.
Its notation is denoted as $Recommend(A,B,t)$, the meaning behind which is the recommendation credibility of data source A relative to the best path to data source B at the "t" moment.

Definition 6. The true credibility of data provided by a data source: It refers to the comprehensive of the direct and indirect credibility of data provided by a data source. Its notation is denoted as $Trust(A,data,t)$, the meaning behind which is the true credibility of data provided by data source A at the "t" moment.

Definition 7. Direct credibility of data provided by a data source : It refers to the credibility of data source in the entire trusted network. Its notation is denoted as $DirTrust(A,data,t)$, the meaning behind which is the direct credibility of data provided by data source A at the "t" moment.

Definition 8. Indirect credibility of data provided by a data source : It refers to the credibility of this data recommended by adjacent data sources with high credibility. Its notation is denoted as $InDirTrust(A,data,t)$, the meaning behind which is the indirect credibility of data recommended by data sources associated with the data source A at the "t" moment.

Definition 9. Credibility of data: It refers to the probability of complementary events of this unreliable data provided by all the data sources which are direct or related providers in the historical records. Its notation is denoted as $Trust(data, t)$, the meaning behind which is the credibility of data in the whole trusted network at the "t" moment.

From the above definition, the relationship is shown in Figure 2.

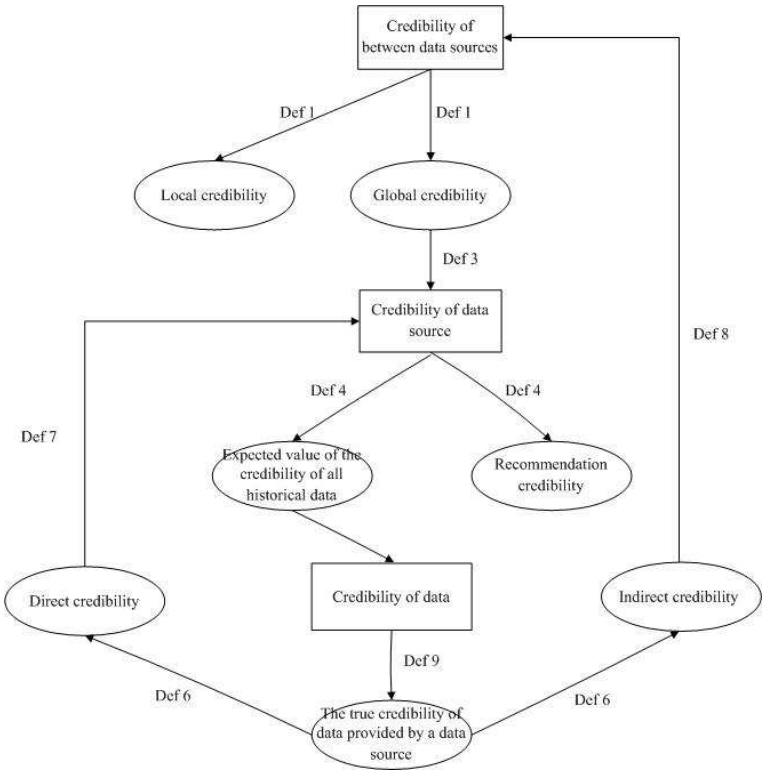

FIGURE 2. The correlation of the credibility definition of between data sources,data,data source

3. **Trusted analysis model for Big Data.** According to the relevant description of the credibility analysis of Big Data, this paper presents a dynamic network model of the credibility analysis, which is composed of a trusted network structure, when the credibility of data is calculated. In the initial time, the analysis network is composed of distributed data sources which are obtained by expert experience, and the network topology is dynamically changed. Then, the network model of the credibility analysis is constructed when the correlation between data sources is calculated . Finally, the credibility measurement of Big Data is carried out through the network model.

3.1. **Measurement model of trusted network.** When there is a direct context interaction among data sources or the similarity of data or behaviors provided by between data sources exceeds a certain threshold. The direct link of the directed graph can be established among data sources. With the expansion of the network size, the trusted network is becoming more and more stable. If it finds out that a data source is not trusted, the model can also quickly impose the penalty factor on the credibility of the provider (data source), making it less reliable for the provider for a period of time. But as time goes on, if the data source can continue to provide reliable data, its credibility will be restored. If data sources have no new context in a calculation interval in the network model of the credibility analysis, time penalty is imposed on them.

By definition 1, the credibility between data sources is calculated by the formula (1), that is, the credibility of data source A relative to data source B, as is shown below in formula (1).

$$Trust_A(B,t) = \alpha_1 \cdot LocalTrust_A(B,t) + \beta_1 \cdot GoobalTrust_A(B,t) \tag{1}$$

Thereinto, $\alpha_1 + \beta_1 = 1$.

According to definition 2, the local credibility between data sources is calculated by the formula (2), that is, the local credibility of data source A relative to data source B, as is shown below in formula (2).

$$LocalTrust_A(B,t) = \begin{cases} Random()or0, & t = 0 \\ LocalTrust_A(B,t-1) \cdot \mu_L(t), & \Delta Context(A,B,t) = 0 \\ [\alpha_2 \cdot DirTrust(A,B,Context(A,B,t),t) & \\ +\beta_2 \cdot Accept(A,B,t)] \cdot \lambda_L(t), & other \end{cases} \tag{2}$$

Notes:

a) The initial value is a random number or 0, which indicates that data source A has some trust or no trust for data source B.

b) $\mu_L(t)$ is the time decay factor at the "t" moment. If the local credibility of data source A is the same as that of data source B at the t and t-1 moment, then it is punished by the time decay factor. Thereinto, $\mu_L(t) = 1 - \frac{\Delta t}{t-t_0}$ $0 \leq \mu_L(t) < 1$. Δt is the time difference of calculation between two times. "t_0" is the starting moment of the current calculation, "t" is the current moment.

c) $\Delta Context(A,B,t)$ is whether between data source A and data source B has a new context of direct interaction at the "t" moment.

$$\Delta Context(A,B,t) = Context(A,B,t) - Context(A,B,t-1)$$

d) $DirTrust(A,B,Context(A,B,t),t)$ is the trusted value of data source A relative to data source B in the circumstances of context interaction at the "t" moment.

e) $Accept(A,B,t)$ is the recognition value of the similarity of data source A relative to data source B at the "t" moment.

$$Accept(A,B,t) = \frac{\sum_{data_a \in Data(A) \cap data_b \in Data(B)} Sim(data_a, data_b)}{Data(A) \cap Data(B)}$$

Thereinto,$Data(A)$ is a data set provided by data source A$data_a$ is a data provided by data source.

$Sim(data_a, data_b)$ is the similarity degree between $data_a$ and $data_b$.$Data(A) \cap Data(B)$ is the number of the same theme in the data sets provided by data source A and B.

f) $\lambda_L(t)$ is the penalty coefficient of local credibility of the model at the "t" moment.

$$\lambda_L(t) = \begin{cases} 1, & \Delta LocalTrust_A(B,t) \geq 0 \\ 0 \leq \lambda_L(t) < 1, & \Delta LocalTrust_A(B,t) < 0 \end{cases}$$

Thereinto,$\Delta LocalTrust_A(B,t)$ is whether the local credibility of data source A relative to data source B has changed at the "t" moment.

$$\Delta LocalTrust_A(B,t) = LocalTrust_A(B,t) - LocalTrust_A(B,t-1)$$

g) $\alpha_2 + \beta_2 = 1$.

By definition 3, the global credibility between data sources is calculated by the formula (3), that is, the global credibility of data source A relative to data source B, as is shown

below in formula (3).

$$GlobalTrust_A(B,t) = Trust(B,t) \tag{3}$$

3.2. **Trusted measurement model of data source.** According to definition 4, the credibility of data source is calculated by the formula (4), that is, the credibility of data source A, as is shown below in formula (4).

$$Trust(B,t) = \begin{cases} Random()or0, & t = 0 \\ Trust(A,t-1)\cdot\mu_S(t), & \Delta Trust(A,t) = 0 \\ [\alpha_3 \cdot \frac{\sum\limits_{data_a \in Data(A)} Trust(data_a,t)}{SUM(Data(A))} & \\ +\beta_3 \cdot (\gamma_n \cdot Recommend_n(A,t))] \cdot \lambda_S(t), & other \end{cases} \tag{4}$$

Notes:

a) The initial value is a random number or 0.

b) $\mu_S(t)$ is the time decay factor at the "t" moment. If the credibility of data source A is the same at the t and t-1 moment, then it is punished by the time decay factor. $\mu_S(t) = 1 - \frac{\Delta t}{t-t_0}$,$0 \leq \mu_S(t) < 1$.

c) $\lambda_S(t)$ is the penalty coefficient of the credibility of data source at the "t" moment.

$$\lambda_S(t) = \begin{cases} 1, & \Delta Trust(A,t) \geq 0 \\ 0 \leq \lambda_S(t) < 1, & \Delta Trust(A,t) < 0 \end{cases}$$

Thereinto, $\Delta Trust(A,t)$ is the difference of calculation for data source A at the t and t-1 moment.

$$\Delta Trust(A,t) = Trust(A,t) - Trust(A,t-1)$$

d) $Trust(data_a,t)$ is the credibility of $data_a$.

e) $Sum(Data(A))$ is the total number of data provided by data source A.

f) γ_n is a $1*n$ dimensional vector which consists of trusted weight of every layer relative to the objective data source in the trusted network.

g) $\alpha_3 + \beta_3 = 1$..

h) $Recommend_n(A,t)$ is the recommendation credibility of each layer of data sources relative to the objective data source A. Thereinto, it is a $n*1$ dimensional vector, the first element of which is the expected value of recommendation credibility of all data sources of the first layer, and the like, each vector element is the expected value for the corresponding layer. The average number of layer is set according to the accuracy and needs, the greater the number of layer is, the greater the amount of calculation is, and the credibility of the corresponding data is more accurate.

i. The recommendation credibility of a data source relative to data source A for the "i" layer of trusted network is calculated by the formula (5), as is shown below in formula (5).

$$Recommend(X_i,A,t) = Trust(X_i,t)\cdot Trust_{X_i}(Neighbor^{max}(X_i->A),t) \tag{5}$$

Thereinto, X_i is a data source X of the "i" layer. $Neighbor^{max}(X_i-> A$ is a data source with the largest credibility adjacent to X_i on the "i-1" layer.

ii. The expected value of the recommendation credibility of data sources relative to data source A for the "i" layer is calculated by the formula, as is shown below.

$$Recommend(A,t)_{(i)} = \frac{\sum\limits_{X \in Circle_i(A)} Recommend(X,A,t)\cdot Sum(Circle_i(A))}{Sum(Circle_i(A))}$$

Thereinto, $Circle_i(A)$ is a set of all data sources on the "i" layer in the trusted network. $Sum(Circle_i(A))$ is the number of all data sources on the "i" layer.

3.3. **Trusted measurement model of data.** By definition 9, the credibility of data is calculated by the formula(6), as is shown below in formula (6).

$$Trust(data, t) = 1 - \prod_{data \in Data(X)} (1 - Trust(X, data, t)) \tag{6}$$

Shows by definition 6, the true credibility of data provided by a data source is calculated by the formula (7), that is, the credibility of data source A relative to the data, as is shown below in formula (7).

$$Trust(A, data, t) = \alpha_4 \cdot DirTrust(A, data, t) + \beta_4 \cdot InDirTrust(A, data, t) \tag{7}$$

Thereinto, $\alpha_4 + \beta_4 = 1$.

By definition 7, the direct credibility of data is calculated by the formula (8), that is, the credibility of data source A, as is shown below in formula (8).

$$DirTrust(A, data, t) = Trust(A, t) \tag{8}$$

From definition 8, the indirect credibility of data is calculated by the formula (9), that is, the indirect credibility of data source A relative to the data, as is shown below in formula (9).

$$InDirTrust(A, data, t) = \frac{\sum_{X \in Neighbor_n(A)} Trust(A, X, t) \cdot Trust(X, data, t)}{n} \tag{9}$$

Thereinto, $Neighbor_n(A)$ is n data sources with high credibility adjacent to data source A.

3.4. **Algorithm analysis process.** Data source is an entity in the trusted network, denoted as "entity"; data is the data provided about a subject by the corresponding entity, denoted as "data"; the theme is the subject of data, denoted as "theme". The behaviors of an entity can be considered to provide a data for a theme in a certain period, the data belongs to the entity, denoted as "$data \in entity$", the data depends on the corresponding theme, denoted as "$data \in theme$", the theme belongs to the entity, denoted as "$theme \in entity$". An entity provides a data set, denoted as "$Data(entity) = \{theme \mid data \in entity\}$", the relationship of an entity associated with all the subjects is denoted as "$Theme(entity) = \{theme \mid theme \in entity\}$", the relationship of a theme associated with all the data sets is denoted as "$Data(theme) = \{data \mid data \in theme\}$".

As mentioned above, firstly, the credibility of an entity relative to other entities is calculated, starting from the formula (1) to calculate the credibility of between data sources. According to the formula (2) and formula (3), the contents of the two aspects are calculated, on the one hand, the formula needs to calculate the local credibility. If the entity has a context interaction (condition 1) or a new behavior (condition 2), the local credibility need updating, if there is no new behavior, time penalty is imposed on it. If the entity meets the condition 1 in the calculation process, or the entity not only meets the condition 2, the similarity of data or behaviors provided by between entities but also exceeds a system threshold, the link of the directed graph can be established among entities, thereinto, the weight of the link is the value of local credibility. On the other hand, the formula needs to calculate the global credibility.

Secondly, the credibility of an entity is calculated by the formula (4). If the expected value of the credibility of all historical data provided by the entity or the recommendation credibility of entities of each layer changes, the credibility of the entity is updated. If the credibility is not changed, time penalty is imposed on it.

Finally, the formula (6) calculates the credibility of data depending on a theme by using the probability of complementary events. The formula (7) gives the true credibility

of data provided by the entity. Meanwhile, the formula (8) gives the direct credibility of data provided by the entity, and the formula (9) gives the indirect credibility of data provided by the adjacent entities. If an entity provides some malicious and false data in the experiment, the entity will be severely punished, so that it can be a very low value in the trusted network. If its behavior is always normal, the credibility will be improved with the increase of their credit. The whole algorithm of the program flow chart is shown below in Figure 3.

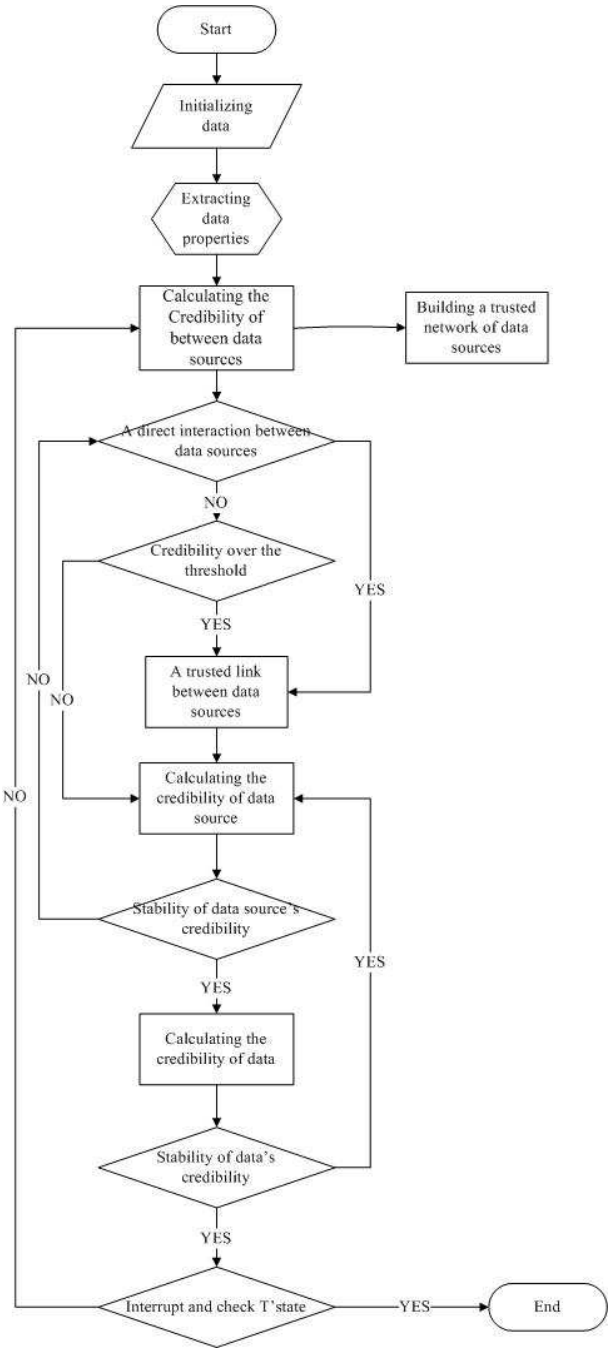

FIGURE 3. The calculation process of data's credibility

4. **Case analysis and verification.** In the paper, the simulation tool is MATLAB, and the simulation experiment selects the Big Data in social networks.

4.1. **Design of simulation experiment.** The experimental data get the goods information of commercial website, especially the collection of information on the evaluation of goods, from the current e-commerce platform through the web crawler technology, and part of the sample data are artificially labeled. Data sets include eight categories of goods, different brands of goods. In this experiment, 79,723 pieces of goods are randomly selected, the number of users reached about 300,000 people, and the evaluation information is as many as ten million. A theme refers to a piece of goods, the entity refers to a customer, and a data refers to the user's evaluation of a goods.

There are more parameters involved in the model, among which are the parameters involved in the credibility of between data sources- the number of data sources N, the local trust weight α_1, the global trust weight β_1, the local direct credibility weight α_2, the similarity weight of local trust β_2, the time decay coefficient $\mu_L(t)$, the penalty coefficient of local credibility $\lambda_L(t)$, the data or behavior similarity threshold η, the time difference of calculation Δt, the parameters involved in the credibility of data source- the credibility expectation weight of historical data α_3, the recommendation weight β_3, the time decay coefficient $\mu_S(t)$, the penalty coefficient $\lambda_S(t)$, the layer number n, the multidimensional weight vector γ_n, and the parameters involved in the credibility of data the direct credibility weight α_4, the indirect credibility weight β_4.

In this experiment, the collected data is divided into two parts. A part of data is used to establish the trusted network, which is trained repeatedly and adjusts the value of the parameters, the other part of data is to verify the stability and accuracy of the model. The settings of the parameters are as follows.

Table 1. The default parameters' list for simulation experiments

Parameter	Default value	Description
N	302412	the number of data sources
μ_L(t)	$\mu_L(t) = 1 - \frac{\Delta t}{t-t_0}$	the time decay coefficient of local credibility
λ_L(t)	$0 \leq \lambda_L(t) \leq 1$	the penalty coefficient of local credibility
α_1	0.735	the local trust weight
β_1	0.265	the global trust weight
η	0.314	the data or behavior similarity threshold
Δt	1	the time difference of calculation
α_2	0.735	the local direct credibility weight
β_2	0.265	the similarity weight of local trust
α_3	0.655	the credibility expectation weight of historical data
β_3	0.345	the recommendation weight
n	3	the layer number
γ_n	(0.64,0.27,0.09)	the multidimensional weight vector
μ_S(t)	$\mu_S(t) = 1 - \frac{\Delta t}{t-t_0}$	the time decay coefficient of the credibility of data source
λ_S(t)	$0 \leq \lambda_S(t) \leq 1$	the penalty coefficient of the credibility of data source
α_4	0.703	the direct credibility weight
β_4	0.297	the indirect credibility weight

4.2. **Experimental results and analysis.** Combined with 3.4 section, data are imported into the algorithm to verify the feasibility. In the process of experiment, we artificially set a customer's data, using the formula (1), formula (4), formula (6) for calculating

the credibility of the customer, and observe the change of its credibility with time. As shown in Figure 4.

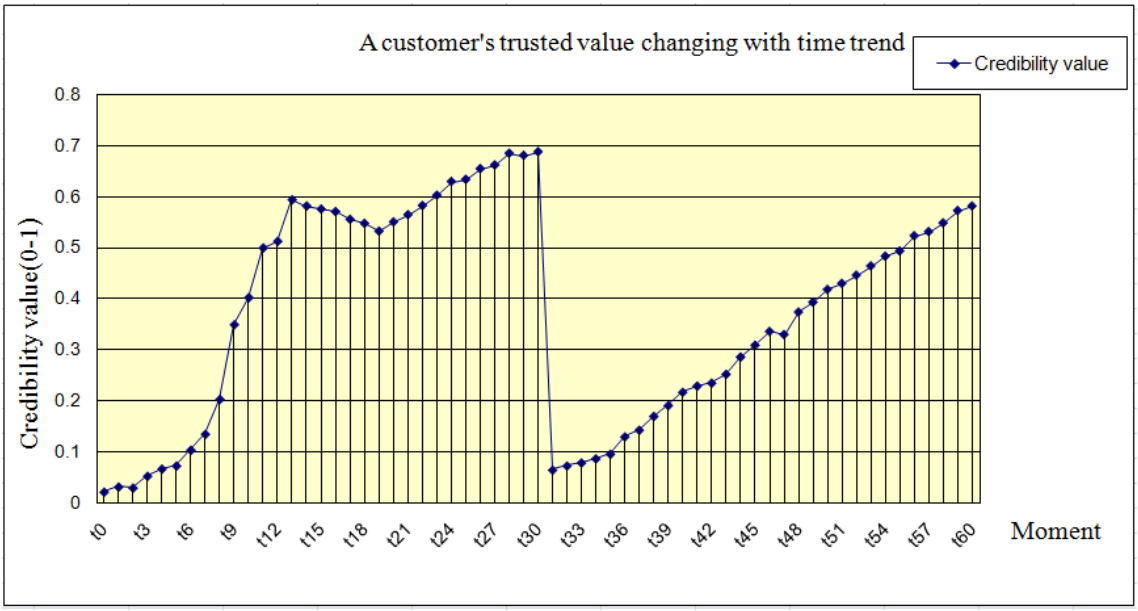

FIGURE 4. A customer's trusted value changing with time trend

From Figure 4, we can find that the customer's credibility presents a rising trend at the $T_0 - T_{30}$ moment, but the customer's credibility has a slow downward trend at the $T_{12} - T_{18}$ moment, which is mainly due to the absence of new behavior and imposes time penalty on its credibility. As a result of the customer to make an unreliable behavior at the T_{31} moment, it has been punished, resulting in its credibility dropped to 0.1. After the T_{32} moment, the customer's behavior is normal and the credibility starts recovering, but the trend is relatively slow. The trusted network topology diagram of hierarchical data sources and the transitive diagram of the credibility of hierarchical data sources relative to a data are shown in Figure 5 at a certain moment. By definition 2, the formula (2) is used to calculate the local credibility of between data sources, and the trusted network can be constructed. As shown in Figure 5 (a), a partial network topology graph is given, as shown in Figure 5 (b), the transitive credibility diagram is given for certain data. We can draw from the fact that a data not only has direct contact with the provider, but also is surrounded by a lot of data sources which are directly or indirectly linked to the data, forming a small trusted network, which can greatly improve the accuracy of a data credibility evaluation.

5. **Conclusions.** In this paper, the typical characteristics and attributes of Big Data is analyzed in detail with combination of the credibility analysis model of general data. Based on the hierarchical model, it gives the analysis model of Big Data credibility measurement. In the case of the large amount of data provided by data sources, the model can accurately analyze the credibility of data, and it is better to satisfy the requirement of Big Data. A simple instance is selected to verify the feasibility of the model. But there are still shortcomings: (1) the default values of the parameters need to be adjusted to adapt to different scenarios, which needs to be improved; (2) the method of building the credibility analysis network still needs to be improved; (3) the model doesnt take some effective measures to amend them for the unreliable data. Above three points will be the focus in further research work.

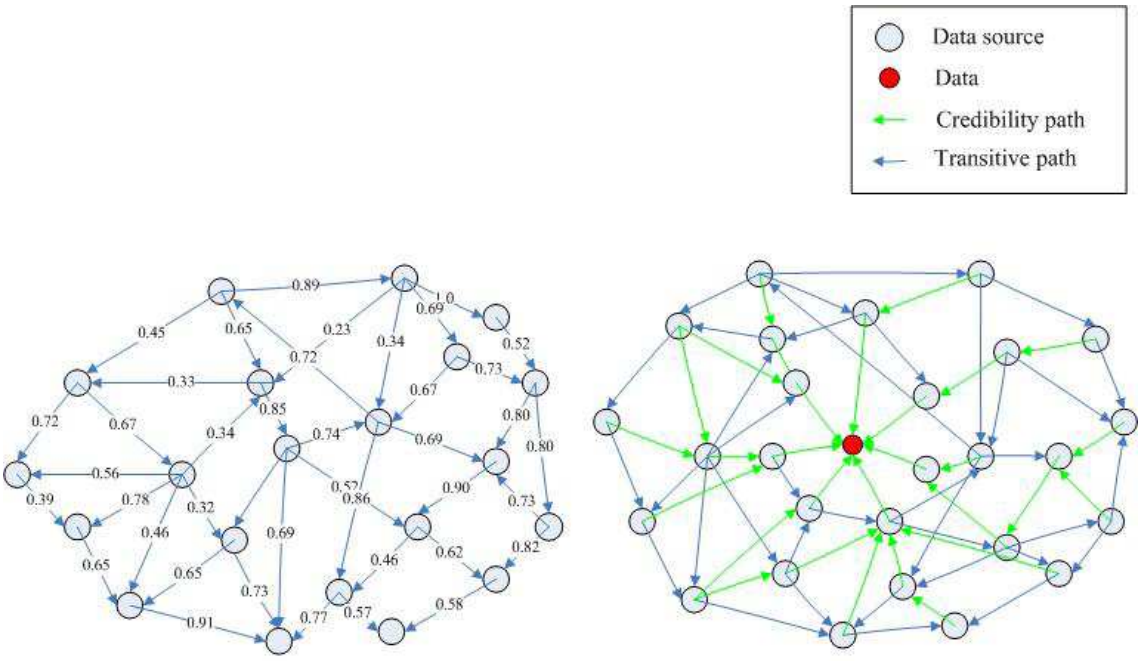

Fig. 5(a) A partial network topology graph

Fig. 5(b) Transitive credibility diagram for a certain data

FIGURE 5. The partial topological diagram of trusted network at a certain time

REFERENCES

[1] G. J. Li and X. Q. Cheng, Research Status and Scientific Thinking of Big Data, *Bulletin of Chinese Academy of Sciences*, vol. 27, No. 6, pp. 647-657, 2012.

[2] A. Y. Zhou, C. Q. Jin, G. R. Wang and J. Z. Li, A Survey on the Management of Uncertain Data, *Chinese Journal of Computers*, vol. 32, No. 01, pp. 1-16, 2009.

[3] Y. Z. Wang, X. L. Jin and X. Q. Cheng, Network Big Data: Present and Future, *Chinese Journal of Computers*, vol. 36, No. 06, pp. 1125-1138, 2013.

[4] D. G. Feng, M. Zhang and H. Li, Big Data Security and Privacy Protection,*Chinese Journal of Computers*, vol. 37, No. 01, pp. 246-258, 2014.

[5] S. X. Wang, L. Zhang and H. S. Li, Evaluation Approach of Subjective Trust Based on Cloud Model,*Journal of Software*, vol. 21, No. 06, pp.1341-1352, 2010.

[6] W. Tang and Z. Chen,Research of Subjective Trust Management Model Based on the Fuzzy Set Theory, *Journal of Software*, vol. 14, No. 08, pp. 1401-1408, 2003.

[7] X. Y. Meng, G. W. Zhang, C. Y. Liu, J. C. Kang and H. S. Li, Research on Subjective Trust Management Model Based on Cloud Model, *Journal of System Simulation*, vol. 19 No. 14,, pp. 3310-3317, 2007.

[8] H. S. Huang and R. C. Wang, Subjective trust evaluation model based on membership cloud theory, *Journal on Communications*, vol. 29, No. 4, pp. 13-19, 2008.

[9] ALMENAREZ F, MARIN A and DIAZ D, Developing a model for trust management in pervasive devices, *Proc of the 3rd IEEE Intl Workshop on Pervasive Computing and Communication Security (PerSec 2006), Washington DC, USA*. pp. 267-272, 2006.

[10] L. Zhang, J. W. Liu, R. C. Wang and H. Y. Wang, Trust evaluation model based on improved D-S evidence theory, *Journal on Communications*, vol.34 , No. 07, pp. 167-173, 2013.

[11] Q. Y. Zhao, W. L. Zuo, Z. S. Tian and Y. Wang, A Method for Assessment of Trust Relationship Strength Based on the Improved D-S Evidence Theory,*Chinese Journal of Computers*, vol. 37, No. 04, pp. 873-883, 2014.

[12] H. Q. Liang and W. Wei,Research of trust evaluation model based on dynamic Bayesian network,*Journal on Communications*, vol. 34, No. 09, pp. 68-76, 2013.

[13] Xiong L and Liu L, PeerTrust: Supporting reputation-based trust for Peer-to-Peer electronic communities, *IEEE Transactions on Knowledge Data Engineering*, vol. 16, No. 7, pp. 843-857. 2004.

[14] X. Q. Qiao, C. Yang, X. F. Li and J. L, Chen,A Trust Calculating Algorithm Based on Social Networking Service Users Context, *Chinese Journal of Computers*, vol. 34, No. 12, pp. 2403-2413, 2011.

[15] J. F. Tian, R. Z. Du and Y. L. Liu, Trust Evaluation Model Based on Node Behavior Character,*Journal of Computer Research and Development*, vol. 48, No. 06, pp. 934-944, 2011.

[16] Wohlgemuth Sven, Echizen Isao, Sonehara Noboru and Mller Gnter, On privacy-compliant disclosure of personal data to third parties using digital watermarking,*Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, No. 3. pp. 270-281, 2011.

[17] X X Yin, J W Han, P S Yu, Truth Discovery with Multiple Conflicting Information Providers on the Web,*IEEE Transactions on Knowledge and Data Engineering*, vol. 20, No. 06. pp. 796-808, 2008.

[18] Hsiao-Ling Wu and Chin-Chen Chang, A Robust Image Encryption Scheme Based on RSA and Secret Sharing for Cloud Storage Systems,*Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, No. 2. pp. 288-296, 2015.

[19] X. Y. Li and X. L. Gui, Trust Quantitative Model with Multiple Decision Factors in Trusted Network,*Chinese Journal of Computers*, vol. 32, No. 3, pp. 405-16, 2009.

[20] F. Zhang, J. Wang, Y. F. Zhao and H. Du, Trust Model Based on Group Recommendation in Social Network, *Computer Science*, vol. 41, No. 5, pp. 168-172, 2014.

[21] W. Yu, S. J. Li, S. Yang, Y. H. Hu, J. Liu, Y. G. Ding and H. Du, Automatically Discovering of Inconsistency Among Cross-Source Data Based on Web Big Data, *Journal of Computer Research and Development*, vol. 52, No. 02, pp. 295-308, 2015.

[22] A Y Chou, The analysis of online social networking; How technology is changing e-commerce purchasing decision, *International Journal of Information Systems and Change Management*, vol. 4, No. 4. pp. 353-365, 2010.

[23] S Song, K Hwang, R Zhou, et al. Trusted P2P transactions with fuzzy reputation aggregation, *Internet Computing, IEEE*, vol. 9, No. 6. pp. 24-34, 2005.

[24] Tahar Mehenni and Abdelouahab Moussaoui, Data mining from multiple heterogeneous relational databases using decision tree classification,*Pattern Recognition Letters*, vol. 33, No. 13. pp. 1768-1775, 2012.

[25] J Zhou, Q Wang, CC Hung and XJ Yi. Credibilistic Clustering: The Model and Algorithms,*Int. Journal of uncertainty fuzziness and knowledge based systems*, vol. 23, No. 4. pp. 545-564, 2015.