Tsinghua Science and Technology

Volume 23 | Issue 3 Article 8

2018

An Information Fusion Model of Innovation Alliances Based on the Bayesian Network

Jun Xia
the School of Management, Harbin Institute of Technology, Harbin 150001, China.

Yuqiang Feng
the School of Management, Harbin Institute of Technology, Harbin 150001, China.

Luning Liu
the School of Management, Harbin Institute of Technology, Harbin 150001, China.

Dongjun Liu
the Shenzhen Institute of Electronics, Shenzhen 518055, China.

TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 12/13 pp347–356
DOI: 10.26599 / TST.2018.9010079
Volume 23, Number 3, June 2018

An Information Fusion Model of Innovation Alliances Based on the Bayesian Network

Jun Xia, Yuqiang Feng*, Luning Liu*, and Dongjun Liu

Abstract: To solve the problem of information fusion from multiple sources in innovation alliances, an information fusion model based on the Bayesian network is presented. The multi-source information fusion process of innovation alliances was classified into three layers, namely, the information perception layer, the feature clustering layer, and the decision fusion layer. The agencies in the alliance were defined as sensors through which information is perceived and obtained, and the features were clustered. Finally, various types of information were fused by the innovation alliance based on the fusion algorithm to achieve complete and comprehensive information. The model was applied to a study on economic information prediction, where the accuracy of the fusion results was higher than that from a single source and the errors obtained were also smaller with the MPE less than 3%, which demonstrates the proposed fusion method is more effective and reasonable. This study provides a reasonable basis for decision-making of innovation alliances.

Key words: information fusion; innovation alliance; Bayesian networks; forecasting model; decision making; big data

1 Introduction

An innovation alliance is a union that consists of enterprises, universities, research institutes, and other institutions. These parties usually have a specific purpose and some common attributes. When information is transferred and applied in the innovation alliance, the process tends to be complicated, syncretic, and marked with other characteristics. Acquisition, classification, synthesis of the information, and decision-making require much work[1, 2]. Information fusion technology is an important tool in the field of knowledge management that can effectively organize information from different knowledge sources and form new and more comprehensive knowledge[3]. The process of information fusion in innovation alliances can collect, transfer, fusion, and apply information from alliance members and external information sources. As a result, sharing, mobility, and integration of knowledge are promoted, and collaborative innovation between each innovation subject is implemented[4].

Information fusion technology is a series of processes that enables the complete use of all types of environment and object information obtained from a plurality of sensors, analysis, and reorganization of information. Such technology allows the fusion of complete and effective information to provide support for decision-making[5–7]. In the 1970s, research institutions in the United States attempted to use the information fusion technology for multiple continuous fusion of sonar signals to determine the position of enemy submarines. This military application is the initial source of information fusion technology; thereafter, early forms of information fusion were mainly intended for data fusion. In the

• Jun Xia, Yuqiang Feng, and Luning Liu are with the School of Management, Harbin Institute of Technology, Harbin 150001, China. E-mail: sharisxia@163.com; fengyq@hit.edu.cn; liuluning@hit.edu.cn.
• Dongjun Liu is with the Shenzhen Institute of Electronics, Shenzhen 518055, China. E-mail: laueastking3168@163.com.
* To whom correspondence should be addressed.
Manuscript received: 2017-12-21; accepted: 2018-01-22

1980s, the United States set up a data fusion technical expert group and formally proposed the concept of data fusion[8]. In 1997, the same country established the Institute of Information Fusion and organized a number of special projects to study data fusion technology. With recent rapid development in computer and network communication technology as well as the growing need for military applications, information fusion technology has seen remarkable achievements[9], and its application areas have gradually expanded from the initial military field to other areas, such as intelligent robotics, image processing and analysis, earth science, agricultural applications, weather forecasting, modern manufacturing industries, and economic management[10]. However, the modeling mechanisms and application conditions of these technologies differ, and each method presents some limitations for certain prediction problems in various applications. For example, estimates may not be accurate enough or the accuracies obtained may not be high enough. Gaining complete and effective information is key to decision making, and how to integrate various methods and overcome the application limits of application and accuracy of each method is an important area of research with both theoretical and practical significance[11].

A large amount of data and information, which may be redundant and conflicting, is involved in the innovation alliance. To analyze these data and make decisions, data processing is very important[12–15]. As such, developing approaches to deal with and analyze big data in the alliance and make decisions based on these data is an essential endeavor. In the present study, an information fusion model for innovative alliance was developed based on a Bayesian network. A hierarchical model of information fusion in an innovation alliance was built, and the content and function of each layer was defined. The model was employed to study economic information prediction, and the fusion results obtained were were of high accuracy. Complete and reliable information was also obtained. This work provides a scientific basis for information management and decision-making in innovation alliances.

2 Information Fusion Hierarchical Model Based on the Bayesian Network

2.1 Bayesian network model

A Bayesian network provides a good way to represent causal information; it is a graphical model that is used to describe the connectivity and distribution of a variable set. The concept of Bayesian networks was first proposed by Howard and Matheson in 1981. Early Bayesian network model were mainly used for expert systems to describe the uncertainty of expert knowledge[16]. Since the 1990s, research on Bayesian networks has seen great progress. Today, the technology is widely used in data mining[17]. Bayesian network models use graphical methods to describe the relationship between data and present advantages such as semantic clarity and understandability. These characteristics are conducive to carrying out predictive analysis based on the causal relationship of the data[18].

A Bayesian network model can be expressed as $N = << V, E >,\ p >$, where V and E are the nodes and the edges, respectively, and p is the probability distribution of the nodes. A node with no input edge is called a root node or evidence node; this type of node has no parent node, and its probability is called prior marginal probability. A node with no output edge is called a leaf node or target node; the probability of this node contains the joint probability distribution of all nodes. Other nodes are called intermediate nodes, which are arranged in turn according to the system topology, and nodes in the same layer are ruled as having no connection to each other. A Bayesian network structure is shown in Fig. 1, where $C_i\ (i = 1, 2, 3, 4)$ is the root node that is the parent node of $B_j\ (j = 1, 2, 3)$, B_j is the parent node of Node A, and Node A is the leaf node. Bayesian networks are described by their qualitative and quantitative parts. The qualitative part refers to the directed acyclic graph topology, while the quantitative aspect refers to the conditional probability and threshold values of non-evidence nodes.

2.2 Hierarchical structure model of information fusion

The basic principle of multi-sensor data fusion technology is similar to the processes of information treatment in the human brain, which takes full advantage of multiple

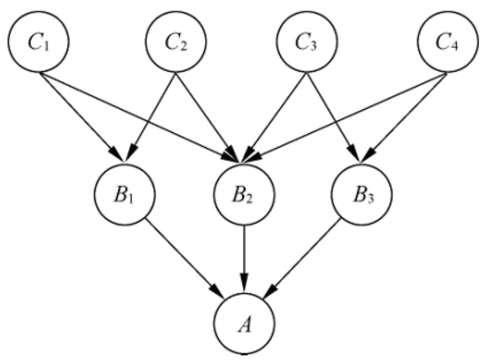


Fig. 1 Topological structure of a Bayesian network.

sensors to gather all types of information. Complementary and redundant information in time and space is reasonably controlled and used. Through some fusion algorithms, multi-feature, multi-source information can be integrated, and a comprehensive explanation and description of the observed object may be obtained. The principles and methods of information fusion technology for multi-sensors are applied to be information fusion process in the innovation alliance to integrate various information from multi-sources and obtain the final full and comprehensive information. Thus, information fusion provides a basis for information management and decision-making for the alliance.

The information fusion process of an innovation alliance is as follows: (1) Get information from all customers, suppliers, competitors, universities, government, research institutions, and enterprises and then search for all types of knowledge on the innovation activities of enterprises, intellectual property, and information on new products or technology. (2) Classify and sort the information obtained from a plurality of sources and remove excess and useless information. Information is classified according to its features in this step, and various types of information are ruled as independent of each other. (3) Fuse all types of information to achieve a final decision. The key in this step is the development of a reasonable and effective algorithm. Appropriate adjustment is necessary according to the actual situation due to the complexity and dynamics of the development process itself, as well as those of the market and policy environments.

The process of information fusion in innovation alliances is similar to that of information fusion technology from multi-sensors. Therefore, this study establishes a hierarchical structure model for information fusion based on the Bayesian network structure. The model was divided into three layers, namely, the information perception layer, the feature clustering layer, and the decision fusion layer. The Bayesian network nodes correspond to the different aspects of the information fusion model, and the root nodes of the network correspond to the various sources of information in the information perception layer, which were defined as the sensors. The parent nodes in the feature clustering layer correspond to the various features in the information fusion model, and all features are independent of each other. The leaf node in the network corresponds to the final decision in the decision fusion layer. The probability of each node could be defined to express the importance of economic information. The information fusion hierarchical structure model is shown in Fig. 2, and the content and function of each layer is described as follows.

(1) Information perception layer

Each information acquisition agency is defined as a sensor. Sensors in the innovation alliance could be defined as enterprises in the alliance, universities, research institutes, and other institutions. Decision-makers in the alliance could learn about the latest economic trends, competitor's performance, new knowledge on new technologies, new products, or other information. In the information perception layer, sensors should cover various internal and external sources of information systems and integrate the information obtained. Therefore, in the design process of the information fusion system, the sensor group should be reasonably selected according to the theoretical model and test data to ensure sufficient accuracy, timeliness, and low cost. In addition, the stability and reliability of the systems should be considered. The types and characteristics of all sensors should also be specifically analyzed to determine how to get as much knowledge as possible from large amounts of information in the sensor group.

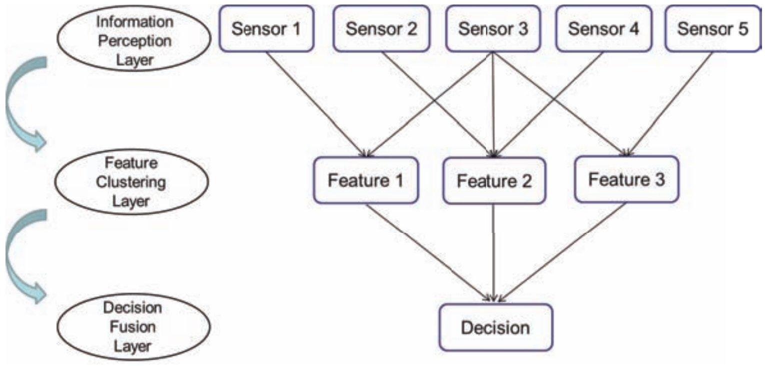


Fig. 2 Information fusion hierarchical structure model.

(2) Feature clustering layer

The information from one source is often simply the description of an object from its own perspective. It may contain different types of information on various aspects of the object but may not be comprehensive and so did each other sensor. In the multi-sensor system, information presents different features; it may be time-varying or non-time-varying, vague or precise, mutually supportive or complementary, and even contradictory and conflicting. The task of the feature clustering layer is to classify all of this information. The features of information are first identified and then extracted. Finally, information with homogeneous features is aggregated, and information with different features is considered independent.

(3) Decision fusion layer

The task of the decision fusion layer is to fuse the information that had been classified in the feature clustering layer and make a final decision. Its core workload is mainly reflected by the fusion algorithm. In terms of multi-sensor systems, the basic requirements of the information fusion algorithm are robustness and parallel processing ability because of the diversity and complexity of information. The computational speed and accuracy of the fusion algorithm, as well as coordination between the former pretreatment system and the subsequent recognition system, should also be considered. Using the fusion algorithm, complete, accurate, and efficiently integrated information could be obtained to enable decision-making in the alliance.

3 Application of the Information Fusion Model

3.1 Relevant theories and methods

(1) Trend Extrapolation Model (TEM)

When time series data show an trend of upward or downward trend, and present no seasonal changes, we can find a suitable function curve to forecast this trend[19]. TEM may be $y = F(t)$, where t represents the time as the independent variable and y represents the data as the dependent variable. At the same time, we believe that this trend may extend to the future. Thus, the forecasting value at a future time point could be obtained when a corresponding time is offered.

According to the shape of the curve of raw data, we use a modified exponential curve model in this study, as shown in Eq. (1):

$$\hat{y}_t = a + bc^t \tag{1}$$

where a, b, and c are parameters to be determined, and $0 < c < 1$. The values of the parameters a, b, and c could be determined according to the known time series data. Then a modified exponential curve prediction model could be developed by substituting the parameters a, b, and c and the trend could be predicted.

(2) Autoregressive Integrated Moving Average (ARIMA) model

The ARIMA model is based on correlation measurements in different periods of a sequence with high precision for short-term forecasting analysis[20, 21]. It is a traditional forecasting model that is mainly used for accurate prediction of time series data, especially seasonal data.

In the ARIMA model, the future values of variables can be expressed as linear functions of past values and stochastic errors[22]. The ARIMA model can be expressed as

$$\begin{cases} \Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t, \\ E(\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \tag{2}$$

where

$$\begin{cases} \nabla^d = (1-B)^d, \\ \Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p, \\ \Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q \end{cases} \tag{3}$$

In the expression, B is a backward operator, ε_t is a stochastic disturbance or stochastic error of each period, d is the difference order, p and q represent the autoregressive order and the moving average order, respectively, and X_t is the observation value $(t = 1, 2, \ldots, k)$.

An ARIMA time series model can be established in three phases, namely, sequence stabilization, model identification, and parameter estimation and model diagnosis. Through repeated processing of these three phases, an "optimal" model for forecasting can finally be determined.

(3) Artificial Neural Network (ANN) model

Since the 1980s, the ANN model has been one of the hottest research topics in the field of artificial intelligence. This model can simulate the neural networks of human brains, process information, and construct different network models based on different connections[23, 24].

A recent upsurge of global research on neural networks could be observed. Deep neural networks can promote the significance and development of the role of artificial intelligence[25]. Artificial intelligence based on neural networks has been widely used in many fields, such as speech recognition, image processing, computer vision, and robotics. AlphaGo, for example, beat world-chess masters Jie Ke and Shishi Li in several matches. Deep-learning neural networks have been used to build much larger and more complex neural networks. Many deep-learning neural network algorithms are semi-supervised algorithms that aim to deal with large data sets on the basis of a small amount of unlabeled data.

The ANN model presents the favorable characteristic of nonlinear combination and is a global approximation network. It has strong learning ability and can achieve nonlinear fitting between the input and output data[26]. In general, it consists of an input layer, a hidden layer, and an output layer; here, the hidden layer may include many middle layers. The structure of an ANN is shown in Fig. 3.

The ANN models can solve a number of practical subjects of nonlinear systems, such as function approximation and system identification. In the present study, the MATLAB neural network toolbox was utilized to develop the ANN model.

(4) Entropy weighting method

The entropy weighting method is a comprehensive method for evaluating multiple objects and indices. Its evaluation results are mainly based on objective data and virtually unaffected by subjective factors. It can avoid the interference of human factors to a great extent[27]. According to the basic principle of information theory, the information of a system is a measure of the degree of order in that system; entropy is a measure of the degree of disorder. Entropy can measure the amount of useful information of data. Entropy weight is not a coefficient of importance in the actual sense but a relative degree of intensity of each index in the sense of competition[28]. If the entropy of the index is smaller, this index can provide more information, therefore, the greater its role should be in the comprehensive evaluation, i.e., the higher its weight should be. By contrast, the greater the entropy of the index, the smaller its entropy weight. The weights should meet the conditions $0 \leqslant \lambda_i \leqslant 1$ and $\sum_{i=1}^{m} \lambda_i = 1$. The process of determining entropy weights is as follows:

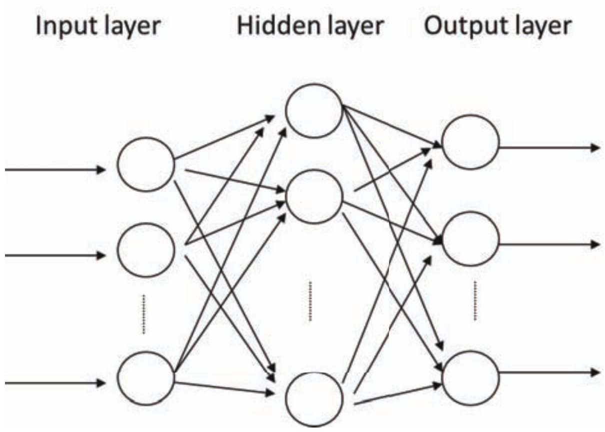


Fig. 3 Structure of the artificial neural network model.

(1) Determine the objects of evaluation, establish the evaluation index system, and construct the index level matrix.

(2) Normalize the original data of all objects to eliminate effect of dimension. The normalization equation of one benefit object is

$$r_{ij} = \frac{x_{ij} - \min_i\{x_{ij}\}}{\max_i\{x_{ij}\} - \min_i\{x_{ij}\}} \tag{4}$$

The normalization equation of a cost object is

$$r_{ij} = \frac{\max_i\{x_{ij}\} - x_{ij}}{\max_i\{x_{ij}\} - \min_i\{x_{ij}\}} \tag{5}$$

where $x_{ij} (i = 1, 2, ..., m$ and $j = 1, 2, ..., n)$ is the observation value of the j-th index on the i-th object and r_{ij} is the dimensionless normalized value.

(3) Define the entropy value p_i of the i-th object as follows:

$$p_i = -k \sum_{j=1}^{n} f_{ij} \ln f_{ij} \tag{6}$$

where $f_{ij} = r_{ij} / \sum_{j=1}^{n} r_{ij}$, $k = 1/\ln n$ $(i = 1, 2, ..., m)$. When $f_{ij} = 0$, we set $f_{ij} \ln f_{ij} = 0$.

(4) Define the weight of the i-th object λ_i as

$$\lambda_i = \frac{1 - p_i}{m - \sum_{i=1}^{m} p_i} \tag{7}$$

where the weights meet $0 \leqslant \lambda_i \leqslant 1$, and $\sum_{i=1}^{m} \lambda_i = 1$.

3.2 Information fusion algorithm

The information fusion model of the innovation alliance based on the Bayesian network was applied to study information prediction. The agencies in the innovation alliance, such as enterprises, universities, and research institutes, were defined as sensors, which obtained information and built the information perception layer. The agencies processed and classified the raw information to construct the feature clustering layer. Then, the alliance fused the information with different features and made a final decision.

To simplify the problem, we assume that various agencies used different methods to predict the time series data obtained from the sensors and that they would transmit the new knowledge to the alliance for information fusion. It would not hurt to assume that the three agencies could be defined as sensors obtaining information. These agencies were named Agencies A, B, and C and specified to utilize the gray forecasting method, the trend extrapolation method, and the exponential smoothing method, respectively, to analyze the trend of information. Thus, results predicted by different agencies would have different features. These results were submitted to the decision-maker in the alliance to fuse the information and make a decision. This process is expected to yield prediction results with higher accuracy, i.e., obtain more complete and reliable information, which could provide a basis for the alliance's information management and decision-making.

The specific process used by the sensors in the alliance employing different forecasting methods to predict the trend of the Gross Domestic Product (GDP) values they acquired and transmit this new knowledge to the alliance for information fusion and decision-making is shown in Fig. 4.

In Fig. 4, different forecasting values are shown based on the different methods applied by three agencies. Here, the information fusion problem of the innovation alliance was transformed into a problem of combination forecasting of different methods. Based on suitable weights, we combined the forecasting values from different methods at the same time point.

4 Analysis of Simulation Results

4.1 Simulation data and results

The GDP values of Guangzhou City were forecasted by three agencies and the information fusion model. The original data, which were derived from Guangzhou statistical information manual in the Guangzhou Statistical Information Network[29], are presented in the second column of Table 1.

The alliance must keep abreast of economic information and forecast the general direction of industrial development. Rational decisions could be made only when the alliance conforms to economic development trends and is aware of the details of economic changes in details. Therefore, to forecast data and research, the trends are of important practical significance.

Three agencies in the alliance predicted the GDP data using three methods, namely, TEM, the ARIMA model, and the ANN model. The results were shown in Table 1. In the fusion algorithm, the entropy weight method was used to combine the results of the three methods from the three agencies. This method could calculate the weights of the series data, and the weights obtained were $w_1 = 0.3352$, $w_2 = 0.3324$, and $w_3 = 0.3324$. The weights of each method varied slightly, thereby indicating that each method is important. The weights were used to combine the prediction results of the three methods, and the results of information fusion were obtained. These results are listed in the sixth column of Table 1.

4.2 Analysis and comparison

4.2.1 Result comparison

The ARIMA model is the traditional and typical method, and we compared the results of the ARIMA model with those of the information fusion model. The prediction curves are described as Fig. 5. In the figure, the original data began to rise from about 125.92 billion Yuan in 1995 and broke through 100 billion Yuan in 2010. In 2016, it reached 1961.09 billion Yuan. The results of various methods followed the changes in raw data, and the trends were approximately the same. The results of the two methods followed the changes of the original data, and the trends were approximately the same. However, the prediction curve of the ARIMA method was out of the curve, and the distance of the curve from that of the original data was farther from the prediction result of the information fusion model. In addition, the prediction results of the information fusion model were close to the original data.

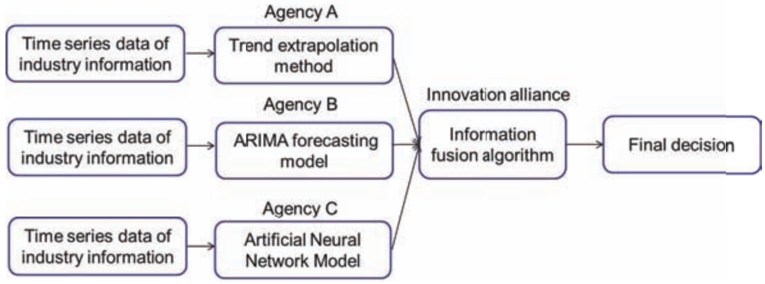


Fig. 4 Information fusion algorithm procedures.

Table 1 Forecasting results of various methods (in billion Yuan).

Year	Original data	TEM	ARIMA	ANNs	Information fusion
1995	125.92	76.53	125.92	131.18	111.11
1996	146.81	106.63	181.91	135.22	141.16
1997	167.81	139.79	182.02	154.12	158.59
1998	189.35	176.31	189.07	197.5	187.6
1999	213.92	216.56	202.99	230.54	216.69
2000	249.27	260.89	225.79	246.83	244.55
2001	284.17	309.73	264.56	267.57	280.7
2002	320.4	363.53	306.88	315.59	328.76
2003	375.86	422.8	354.32	388.28	388.56
2004	445.06	488.09	424.31	449.23	453.97
2005	515.42	560.02	509.6	507.49	525.8
2006	608.19	639.27	596.5	600.03	612.01
2007	714.03	726.56	705.98	723.31	718.64
2008	828.74	822.73	826.98	830.08	826.59
2009	913.82	928.67	954.35	923.71	935.56
2010	1074.83	1045.38	1049.29	1054.23	1049.62
2011	1242.34	1173.96	1219.3	1234.84	1209.26
2012	1355.12	1315.6	1389.32	1398.91	1367.79
2013	1549.72	1471.63	1498.38	1520.31	1496.71
2014	1670.69	1643.53	1687.07	1660.26	1663.56
2015	1810.04	1832.89	1794.41	1836.34	1821.24
2016	1961.09	2041.51	1918.31	1931.03	1963.83

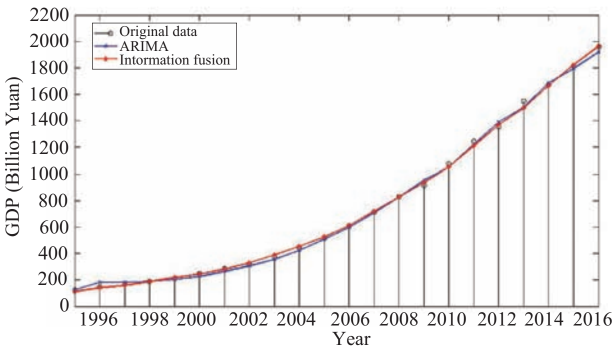


Fig. 5 Prediction curves of the ARIMA and information fusion model.

4.2.2 Accuracy analysis

To investigate the accuracy and precision of predictions, the errors of the prediction results were tested. Various prediction indices, i.e., Mean Absolute Error (MAE), Mean Percentage Error (MPE), and Root Mean Square Error (RMSE), were calculated[30] using the calculation formulas and functions presented in Table 2. MAE, MPE, and RMSE could reflect system errors and indicate the dispersion degree of the prediction results and the original data. MAE and RMSE are absolute indices, whereas MPE is a relative index and dimensionless. Table 3 reveals that the MAE, MPE, and RMSE of the fusion results were lower than those of the single-agency setup, and a higher accuracy was obtained. The MAE of the fusion results was below 120 billion Yuan, while that of the three other agencies exceeded 130 billion Yuan, the precision of the MAE of the fusion results was also enhanced by over 100% compared with the results of Agency B using the ARIMA model. The MPE of the fusion results was below 3%, which is lower than results of other forecasting methods. These findings indicate that the fusion predictions were closer to the actual data and that the information provided was more effective and reliable.

Table 2 Calculation formulas of error testing indices.

Index	Formula	Function
MAE	MAE = $\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert$	It can describe the system errors, and is absolute index.
MPE	MPE = $\frac{1}{n}\sum_{i=1}^{n}\lvert\frac{y_i-\hat{y}_i}{y_i}\rvert$	It can describe the system errors, and is a relative index and dimensionless.
RMSE	RMSE = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$	It can describe the system errors, and is absolute index.

Note: $y_i (i=1,2,...,n)$ are the actual observed values; $\hat{y}_i$ are the prediction values.

Table 3 Indices of the accuracy of prediction results.

Index	MAE (billion Yuan)	MPE (%)	RMSE (billion Yuan)
TEM	344.7770	8.26	404.6351
ARIMA	198.2622	4.16	241.0072
ANNs	136.5580	2.80	171.2258
Information fusion	118.2381	2.31	167.8423

The accuracy of results of the ARIMA and information fusion models were compared. According to Table 3, the MAE, MPE, and RMSE of the results from the fusion model were better than those of the ARIMA model. Considering the various prediction error indices, the prediction accuracy of the fusion results was higher than that of the single-agency setup, which uses the ARIMA model. These findings illustrate that complete and accurate information could be obtained through the information fusion model and that this information is comprehensive and reliable for decision-making.

5 Conclusion

An information fusion hierarchy model for innovation alliances based on the Bayesian network was presented. The decision-making process of the innovation alliance is divided into three layers, namely, the information perception layer, the feature clustering layer, and the decision fusion layer. The process of information fusion in the alliance is as follows. Information is observed via sensors in the information perception layer and clustered according to the features in the feature clustering layer. Then, comprehensive and effective information is obtained through the fusion algorithm in the decision fusion layer. The proposed technology presents the advantages of multiple sensors operating jointly, and the accuracy and reliability of the information obtained in the alliance was improved through collaborative work. The fusion algorithm improves the reliability of the fused information for decision-making in the innovation alliance.

The information fusion model was applied to research on predicting economic information, and a reasonable fusion algorithm was presented. Simulation results indicated that the results of fusion were more credible and comprehensive than those of a single-agency. As well, the prediction accuracy of fusion results was higher and errors were smaller.

The information fusion model could process and integrate information from multiple sources, achieve overall optimization of information, and form complete and comprehensive information based on a fusion algorithm. This study provides a theoretical framework of the process of information transfer and integration, as well as scientific support for big data analysis and decision making in innovation alliances.

Acknowledgment

This research was supported by the National Natural Science Foundation of China (Nos. 71472053, 71429001, and 91646105).

References

[1] J. Q. Dong and C.-H. Yang, Information technology and organizational learning in knowledge alliances and networks: Evidence from us pharmaceutical industry, *Information & Management,* vol. 52, no. 1, pp. 111–122, 2015.

[2] C. Velu, Business model innovation and third-party alliance on the survival of new firms, *Technovation,* vol. 35, pp. 1–11, 2015.

[3] E. F. Wubben, M. Batterink, C. Kolympiris, R. G. Kemp, and O. S. Omta, Profiting from external knowledge: The impact of different external knowledge acquisition strategies on innovation performance, *International Journal of Technology Management,* vol. 69, no. 2, pp. 139–165, 2015.

[4] G. Battisti, J. Gallego, L. Rubalcaba, and P. Windrum, Open innovation in services: Knowledge sources, intellectual property rights and internationalization, *Economics of Innovation and New Technology,* vol. 24, no. 3, pp. 223–247, 2015.

[5] S.-M. Chen, S.-H. Cheng, and C.-H. Chiou, Fuzzy multiattribute group decision making based on intuitionistic fuzzy sets and evidential reasoning methodology, *Information Fusion,* vol. 27, pp. 215–227, 2016.

[6] S.-M. Chen and Z.-C. Huang, Multiattribute decision making based on interval-valued intuitionistic fuzzy values and linear programming methodology, *Information Sciences,* vol. 381, pp. 341–351, 2017.

[7] G. Lin, J. Liang, and Y. Qian, An information fusion approach by combining multigranulation rough sets and evidence theory, *Information Sciences,* vol. 314, pp. 184–199, 2015.

[8] Y. Liu, Z.-P. Fan, and X. Zhang, A method for large group decision-making based on evaluation information provided by participators from multiple groups, *Information Fusion,* vol. 29, pp. 132–141, 2016.

[9] P. Braca, R. Goldhahn, G. Ferri, and K. D. LePage, Distributed information fusion in multistatic sensor networks for underwater surveillance, *IEEE Sensors Journal,* vol. 16, no. 11, pp. 4003–4014, 2015.

[10] W. Xu, M. Li, and X. Wang, Information fusion

based on information entropy in fuzzy multi-source incomplete information system, *International Journal of Fuzzy Systems,* vol. 19, no. 4, pp. 1200–1216, 2017.

[11] Z. Xu and N. Zhao, Information fusion for intuitionistic fuzzy decision making: An overview, *Information Fusion,* vol. 28, pp. 10–23, 2016.

[12] G. P. Giambrone, H. C. Hemmings, M. Sturm, and P. M. Fleischut, Information technology innovation: The power and perils of big data, *British Journal of Anaesthesia,* vol. 115, no. 3, pp. 339–42, 2015.

[13] Z. Wang, J. Xin, H. Yang, S. Tian, G. Yu, C. Xu, and Y. Yao, Distributed and weighted extreme learning machine for imbalanced big data learning, *Tsinghua Sci. Technol.*, vol. 22, no. 2, pp. 160–173, 2017.

[14] Y. Wu, Z. Chen, Y. Wen, W. Zheng, and J. Cao, COMBAT: A new bitmap index coding algorithm for big data, *Tsinghua Sci. Technol.*, vol. 21, no. 2, pp. 136–145, 2016.

[15] Y. Shen, B. Guo, Y. Shen, X. Duan, X. Dong, and H. Zhang, A pricing model for big personal data, *Tsinghua Sci. Technol.*, vol. 21, no. 5, pp. 482–490, 2016.

[16] C. Tang, Y. Yi, Z. Yang, and J. Sun, Risk forecasting of pollution accidents based on an integrated bayesian network and water quality model for the south to north water transfer project, *Ecological Engineering,* vol. 96, pp. 109–116, 2016.

[17] A. A. P. Kazem, H. Pedram, and H. Abolhassani, Bnqm: A bayesian network based qos model for grid service composition, *Expert Systems with Applications,* vol. 42, no. 20, pp. 6828–6843, 2015.

[18] G. Wang, T. Xu, T. Tang, T. Yuan, and H. Wang, A Bayesian network model for prediction of weather-related failures in rail- way turnout systems, *Expert Systems with Applications,* vol. 69, pp. 247–256, 2017.

[19] L.-X. Jian, Y.-T. Liu, and N. Wang, Trend analysis of internal structure of service industry in China, in 2010 *International Conference on Management and Service Science (MASS)*, 2010, pp. 1–5.

[20] C. A. V. Cardoso and G. L. Cruz, Forecasting natural gas consumption using arima models and artificial neural networks, *IEEE Latin America Transactions,* vol. 14, no. 5, pp. 2233–2238, 2016.

[21] Q. Yan and C. Ma, Application of integrated arima and rbf network for groundwater level forecasting, *Environmental Earth Sciences,* vol. 75, no. 5, p. 396, 2016.

[22] P. Sen, M. Roy, and P. Pal, Application of arima for forecasting energy consumption and ghg emission: A case study of an indian pig iron manufacturing organization, *Energy,* vol. 116, pp. 1031–1038, 2016.

[23] R. C. Deo and M. S.ahin, Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index uding hydrometerorological, *Atmospheric Research,* vol. 161, pp. 65–81, 2015.

[24] P. Ramasamy, S. Chandel, and A. K. Yadav, Wind speed prediction in the mountainous region of india using an artificial neural network model, *Renewable Energy,* vol. 80, pp. 338–347, 2015.

[25] P. Ganesan, S. Rajakarunakaran, M. Thirugnanasambandam, and D. Devaraj, Artificial neural network model to predict the diesel electric generator performance and exhaust emissions, *Energy,* vol. 83, pp.115–124, 2015.

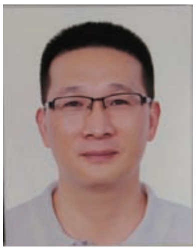
[26] A. F. Marj and A. M. Meijerink, Agricultural drought forecasting using satellite images, climate indices and artificial neural net- work, *International Journal of Remote Sensing,* vol. 32, no. 24, pp. 9707–9719, 2011.

[27] X. Fu, L. Huang, G. Su, L. Chen, C. Li, and L. Wu, Using entropy weight-based topsis to implement failure mode and effects analysis, in *Proceedings of the 6th Asia-Pacific Symposium on Internetware on Internetware*, 2014, pp. 89–96.

[28] C. Li, J. Qin, J. Li, and Q. Hou, The accident early warning system for iron and steel enterprises based on combination weighting and grey prediction model gm (1, 1), *Safety Science,* vol. 89, pp. 19–27, 2016.

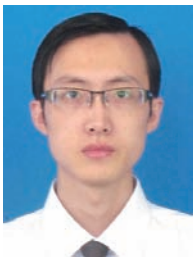
[29] Guangzhou Statistical Information Network, *Guangzhou Statistical Information Manual*. http://www.gzstats.gov.cn/ (accessed on 6 August 2017).

[30] Z. Ali, I. Hussain, M. Faisal, H. M. Nazir, T. Hussain, M. Y. Shad, A. Mohamd Shoukry, and S. Hussain Gani, Forecasting drought using multilayer perceptron artificial neural network model, *Advances in Meteorology,* vol. 2017, 2017.



Jun Xia received the MS degree in logistics and shipping economics and management from Hong Kong Polytechnic University in 2012. He is currently a PhD candidate in School of Management, Harbin Institute of Technology (HIT), China. He is the executive vice president and secretary general of Shenzhen Institute of Electronics. He has long been engaged in the work of electronic information and intelligent manufacturing industry and has many years of working experience in electronic information industry. His current research interests are multi-source information fusion and enterprise informationization.

Yuqiang Feng is a professor in the Department of Management Science and Engineering in the School of Management, Harbin Institute of Technology (HIT), China. She received the BS degree in computer science, MS degree in management engineering, and PhD degree in management science and engineering from HIT in 1984, 1988, and 1994, respectively. Her research primarily focuses on management information systems, decision support systems, negotiation support systems, and online auction. Her work has been published in *European Journal of Information Systems, Information & Management, Computers in Human Behavior, Behaviour & Information Technology, Journal of Global Information Management, Industrial Management & Data Systems, Information Technology for Development (ITD)*, *Scientometrics*, and conferences including the Hawaii International Conference on System Sciences (HICSS), the European Conference on Information Systems (ECIS), the Americas Conference on Information Systems (AMCIS), and the Pacific Asia Conference on Information Systems (PACIS).



Dongjun Liu received the PhD degree in management from Beihang University in 2013, and was postdoctoral fellow in Harbin Institute of Technology (HIT) in 2013–2016. He is the editor-in-chief of Shenzhen Institute of Electronics and has published 20 academic papers. His research interests include environmental assessment and management, integration of information technology and industrialization, and analysis of intelligent manufacturing industry.

Luning Liu is an associate professor in the Department of Management Science and Engineering in the School of Management at Harbin Institute of Technology (HIT), China. He received the BS, MS, and PhD degrees in information systems from HIT in 2005, 2007, and 2011, respectively. His research primarily focuses on enterprise information systems implementation and assimilation, e-government and social governance, mobile commerce and online platforms, big data, and business analytics. His work has been published in *European Journal of Information Systems, Computers in Human Behavior, Journal of Global Information Management, Interactive Learning Environments, Information Technology for Development, Industrial Management & Data Systems, Scientometrics*, and conferences including the Hawaii International Conference on System Sciences (HICSS), the European Conference on Information Systems (ECIS), the Americas Conference on Information Systems (AMCIS), and the Pacific Asia Conference on Information Systems (PACIS).