Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **4**, July **2017**

Bayesian Network for Motivation Classification in Creative Computation

Fuquan Zhang[1,2]
[1]School of Software, Beijing Institute of Technology
[2]Fujian Provincial Key Laboratory of Information Processing and Intelligent Control
Minjiang University
Room A9#301, Jiyuan, Juyuan, Jinshan, Fuzhou, 350008, China
8528750@qq.com

Gangyi Ding
School of Software, Beijing Institute of Technology
Room 611, Software Building, Beijing Institute of Technology,
5 South Zhongguancun Street, Haidian, Beijing, 100081, China
dgy@bit.edu.cn

Zijing Mao
Department of Electrical Engineering
University of Texas at San Antonio One UTSA Circle,
San Antonio Arrow Oaks 126, San Antonio, Texas, 78249, USA
Zijing.Mao@utsa.edu

Lin Xu
Innovative Information Industry Research Institute
Fujian Normal University
Fuqing Brunch of Fujian Normal University, Campus County 1#, Longjiang Road,
Fuqing, fuzhou,, 350300, China
xulin@fjnu.edu.cn

Xiaoyan Zheng
Foreign Language Institute
Minjiang University
Foreign Language Institute of Minjiang University , Xiyuan Gong Road 200#,
Shangjie, Minhou , Fuzhou, 350121, China
496167931@qq.com

Received January 2017; revised May 2017

Abstract. *This paper used clustered embedding features and language ontology to simulate the process of creative idea generation from human brain. Conventional creative idea generation used predefined motivation words. This study proposed a new motivation classification algorithm based on Bayesian network. We first applied the crawler algorithm to obtain motivation vocabulary corpus information.Bayesian network is applied to filter motivation words and project to a vector space for building information clustering model. Then we applied ontology theory to generating the phrase and sentence structure. By creative computing we try to discover how the idea comes from step by step and display the emergence, self-organization, self- coordination of idea creation process of human brain. From this paper, you will see the process in five steps from blurry motivations to clear creative ideas.*
Keywords: Idea creation; Creative computing; Embedding features; Ontology; Clustering; Bayesian network

1. **Introduction.** Creative idea creation is the act of creating novel and applicable ideas in which creativity is one of the kernel features. Creativity is an extremely important facet of life and is a feature of many of the tasks we perform every day. It can occur in a multitude of situations ranging from work to pleasure, from artistic portrayals to technological innovation ...[1]. Ideas and technologies have been blowout in recent years which dramatically improved our life [2, 3]. Therefore, creative computing demonstrates its importance as an inspiration for more creative ideas [4, 5].

Creative computing, belonging to one kind of software design trying to simulate the creative thinking of human being, is one of the computation algorithms emerged recently. On the proceedings of the 19th international conference on automation & computing Yang et al. put forward the theory of creative computing and give the definition, research scope and challenges of creative computing in order to develop creative computing and make it more efficiently and effectively serve the world. From the theories of Andrew Hugill [3], Creative computing contains five layers, as motivation, formulation, creation, dissemination and revision [3].Among the processes of creative computing the most difficult is the subprocess from motivation to formulation. In fact the subprocess is idea creation. An idea is a collection of thoughts, plans and images that flash through the brain. The factors such as emotion, demands, etc. influence the idea creation. If the brain is considered to be a machine presented in our body, then thinking in the reverse way implies that a machine can also create ideas. Computer is a machine meant to serve the users. Developing such a method (algorithm/ model/approach), which helps computer in creating its own sets of ideas, would provide a mind map for computer. In this paper, idea created by computer will help the method to serve any domain's users and it is a simulation of the process with emergence, self-organization and self-coordinating coming from our brains. It is related closely to the layers about motivation and formulation. The process is to finish converting from blurry motivation to clear creative idea. Jing et al's[2] explicit idea creation is a complicated process in which it requires plenty of knowledge as support for specific domain. Fig.1 shows a creative idea coming from many domains. In the process of idea creation, there are many temporary ideas emerging. But from the ideas which is the creative idea that we need, we should evaluate them and choose the best fit for us. First in the idea generation, we should choose them from syntax and semantics. Secondly in the idea evolution, we should choose them from interest, hobby and so on. At last we should evaluate them from three properties: novelty, surprise, usefulness to judge whether they are genius ideas for us. Jing et al. gives a new way to abstract idea creation from domain ontology extraction [2]. This method simulated a linear system for us to read out and generate creative ideas from personal mind. But indeed creative ideas come from our minds and thoughts which are generated by our non-linear brain–a complex system [4]. From Yang et al. [5] theory, the creativity of creative computing belongs to combinational creativity. As we can see from fig.1, the process of idea creation is a combinational creativity, because the idea involves many domains. The process is completed with diversion thinking and conversion thinking. It is shown that creative computing can be applied not only to NLP domain but to musical composition, painting and so on for many domains innovation. Therefore it has tremendous potential and vitality, with various aspects needed to be studied.

In this paper, we investigate how to generate creative ideas based on a clustering algorithms filter out irrelative motivation words. [6] We applied Support Vector Machine (SVM) algorithm to filtering motivation, then applied SNoW algorithms as a specific subject classifier, and proposed feedback and threshold filtering strategies to achieve accurate filter blog advertising purposes. However, this motivation classification is time consuming and with low accuracy. Based on heuristic rules literature [7] proposed an automatic classification method. This method can automatically classify and filter the motivation word by segmenting and extracting the word through NB classifier. But the accuracy of the method is rather low. Literature [8] proposed an classification filtering method based on latent semantic indexing and support vector machine. Based on establishing the motivation lexical information filtering model, it can automatically classify and filter the motivation words by preprocessing, feature reducting, training, filtering etc.. However, this method has the problem of incomplete information filtering. Therefore, we proposed Bayesian network based on motivation classification algorithm to improve the filtering accuracy. Our method has five steps, seeing fig.2. it can complete idea creation based on the two layers which was put forward by Andrew Hugill from motivation to formulation. Our method is proposed by the model: 1). gathering relative data of motivation 2).clustering layer 3).idea generation layer 4).idea evolution layer 5).idea evaluation layer. We have also tested our model in a case by inputting blurry motivation and the result is what we desired. The result shows that the proposed methods can actually generate creative ideas. Then we used the system of the model to compare with the questionnaire survey to show the different results in the quantity and quality.

The remaining parts of the paper are organized as follows. Material and methods are described in Section II. Test experiments and results are reported in Section III. Conclusions are drawn in Section IV. The process of general creative idea generating

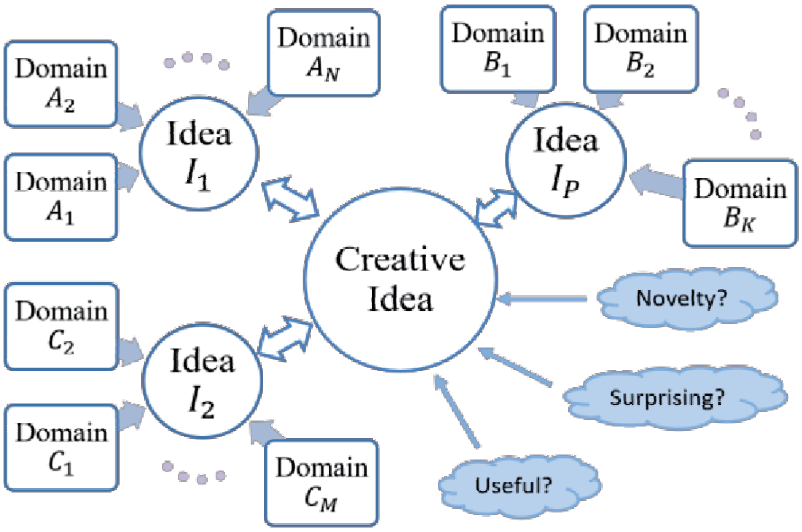

FIGURE 1. Description of creative computing pipeline

pipeline

2. Material and Methods.

In this section, we propose some effective ways to get the material data and the methods for creating creative idea. We collect text data for idea creation by crawler technology. We use the methods to make the idea creation model by following four steps.

2.1. Crawler Algorithm on Vocabulary Corpus for Motivation Information.

a). Computation of information weights. In this study, we applied improved TF/IDF algorithm to calculating motivation vocabulary information weights. We used word frequency statistics to assess the importance of a word that appears in the corpus. If the frequency

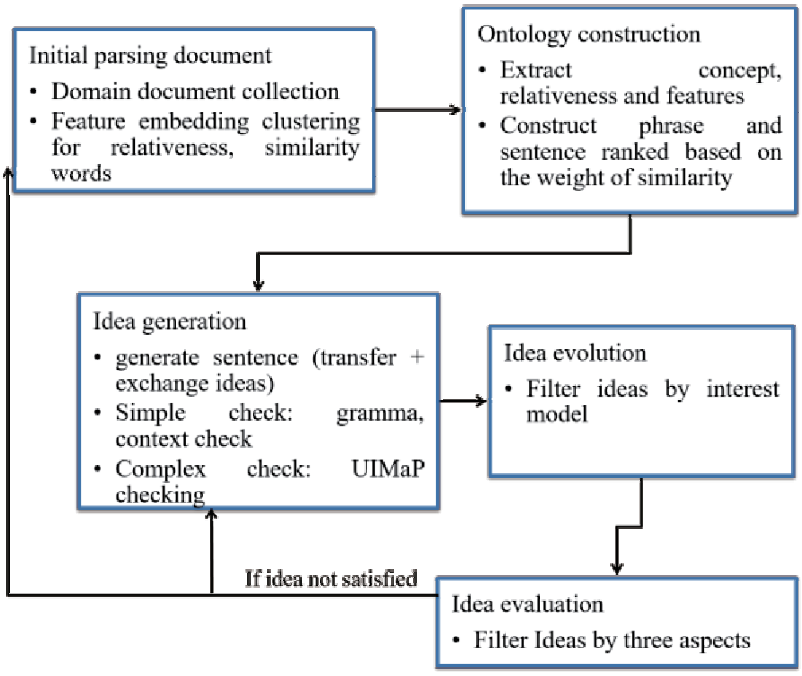

Figure 2. The creative idea generation method pipeline

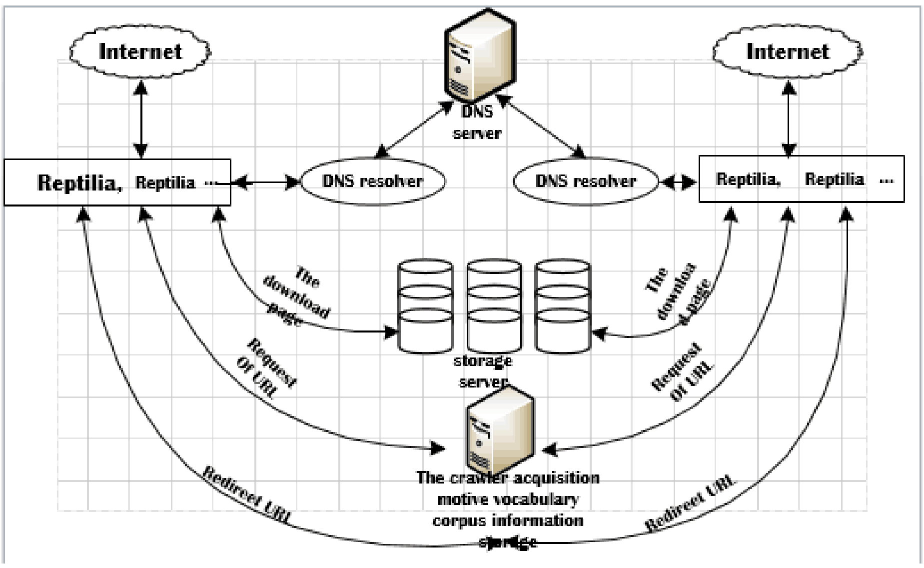

Figure 3. Overall process of collecting the text data by crawler technology

of key word A is greater than key word B in the corpus, it means the key words A is more important than key words B [9, 10]. However, this measure is not perfect, because it only reflects the local characteristics of a particular document.

From the point of the whole corpus, the high-frequency of a word doesn't mean that it's important, the weight of a key word is in inverse proportion to the total number of corpora. Assume key word set $W = (W_1, W_2, \ldots, W_n)$ the i^{th} key word W_i has x ($x \geq 1$) concepts, denotes as $T(W_i) = (t_w^1, t_w^2, \ldots, t_w^{x_i})$, transfer from key work set to concept set T_W, each key concept can be represented as $t^j_{w_i} = \left(w_i, t^j_i\right)$, $j = 1, 2, \ldots, x_i$, where the i^{th} motivation words for the j^{th} concept can be denoted as:

$$T_W = \{(w_1, t_1^1), (w_1, t_1^2), \ldots, (w_1, t_1^{x_n})\} = (t_w^1, t_w^2, \ldots, t_w^{x_n}) \quad (1)$$

And the motivation weights for key word t_k in corpus d_i can be represented as:

$$\mathrm{w_{ik}} = \mathrm{tf_{ik}} \times \log\left(\frac{\mathrm{N}}{\mathrm{n_k}}\right) \quad (2)$$

where N is the motivation word size in the corpus, n_k is the total motivation size that contains key word motivation t_k, tf_{ik} is the showing frequency of t_k in corpus d_i. And the corpus d_i can be represented as $\mathrm{d_i} = (\mathrm{w_{i1}}, \mathrm{w_{i2}}, \ldots, \mathrm{w_{in}})$.

If the information corpus in the motivation word training sample set belong to a same category, the important motivation word will present in every word corpus, which will affect the weight because of the low IDF.

When the information of the entire corpus belong to a same category, the importance of the motivation word should be in proportion to the frequency of its appearance.

A key word appears in a catalog many times, while rarely appears in other catalog , which means this key word have good ability to represent this kind of catalog. so the traditional weight calculation method has flaws. not all cases are outstanding? needs to be improved.

Assuming the motivation counts in corpus is N, the number of key word appearance in one class B_i is n_{bi}, n_{ki} represents the total motivation size that contains key word motivation t_k except class B_i, and the weight can be calculated as:

$$\mathrm{w_{ik}^{'}} = \mathrm{tf_{ik}} \times \log\left(\frac{\mathrm{n_{bi}}}{\mathrm{n_{ki}} + \mathrm{n_{bi}}} \times \mathrm{N}\right) \quad (3)$$

where $\mathrm{IDF} = -\frac{\mathrm{n_{bi}}}{\mathrm{n_{ki}}+\mathrm{n_{bi}}} \times \mathrm{N}$ and assume function $\mathrm{f}(\mathrm{x}) = \frac{\mathrm{x}}{\mathrm{x+k}}$ and $\mathrm{x_1}>\mathrm{x_2}>0$, we have

$$f(x_1) - f(x_2) = \frac{x_1}{x_1 + k} - \frac{x_2}{x_2 + k} = \frac{(x_1 - x_2)\,k}{(x_1 + k)(x_2 + k)} > 0 \quad (4)$$

It is noticeable that as x increase, IDF value will also increase. The more show time key word t_k in B_i, the less it will be shown in other classes, thus the key word t_k can represent as the feature for B_i.

b). Vocabulary corpus correlation calculation based on motivation weights. we use crawler technology, as shown in fig.3, to get materials corpus which is relative to motivation. Correlation calculations can predict relevance goal motivation vocabulary corpus, and guide the search direction of additional motives vocabulary corpus. We also compared with the pre-set weight in corpus, motivation vocabulary with weight greater than the preset will be reserved, less than the preset will be directly abandoned [11, 12]. This will not only improve motivation acquired vocabulary and corpus relevance, but also reduces local storage space. The method of correlation calculation mainly used content analysis and link structure analysis.

Bayesian network method is applied to modeling the probability of motivation word belong to a given category. We first calculated the probability for the motivation words for each class based on weights calculation, represented as (w_1,w_2,...,w_n). The probability of motivation word set d_i can be represented as:

$$w_m = P\left(w_m \left| B_j\right.\right) = \frac{1 + \sum_{i=1}^{|D|} N(w_m, d_i)}{|V| + \sum_{s=1}^{|V|} \sum_{i=1}^{|D|} N(w_s, d_i)} \quad (5)$$

where $|D|$ is the training sample size in B_j, $N(w_m, d_i)$ is the presentation frequency of motivation word w_m in d_i, $|V|$ is the total motivation count, $\Sigma_{s=1}^{|V|}\Sigma_{i=1}^{|D|}N(w_s, d_i)$ is the summation of the motivation word frequency for the given class. Then we can calculate the relativity of motivation words as:

$$P\left(B_j \middle| d_i; \hat{\theta}\right) = \frac{P\left(B_i \middle| \hat{\theta}\right) \prod_{m=1}^{n} P\left(Bj \middle| di; \hat{\theta}\right)^{N(w_m, b_i)}}{\sum_{r=1}^{|C|} P\left(B_i \middle| \hat{\theta}\right) \prod_{m=1}^{n} P\left(Bj \middle| di; \hat{\theta}\right)^{N(w_m, b_i)}} \tag{6}$$

where $P\left(B_i|\hat{\theta}\right) = B_j$ is the ratio of training sample size to total sample size, $|C|$ is the total class number, $N(w_m, d_i)$ is the word frequency of w_m in d_i, n is the total key word size; based on relativity calculation, we can obtain the desired motivation information.

c). Motivation information related corpus acquisition based on relativity. Combined with the results of correlation, we can collect relative corpus information of motivation vocabulary through crawler.

The steps of obtaining are as follows:

First, the motivation word input is acquired by other nodes selected from some seeds URL and other relevant motivation words.Second, we will download the motivation words or pass it to other nodes. If the motivation word is received from other nodes, then we will check if it already exists in the corpus; otherwise the hyperlink from the motivation words will be analyzed and the crawl process will be assigned with a new motivation word and store the previous one to database. For each new URL crawl, the weight will be calculated as (7),

$$\text{node\_num} = \text{hash}\left(\text{new\_url.host}\right) \%\text{node\_sum\_num} \tag{7}$$

While for each motivation word crawl, the weight will be calculated as (8),

$$\text{node\_num}' = \text{hash}\left(\text{new\_url\_firstword.host}\right) \%\text{node\_sum\_num} \tag{8}$$

The node numbers for both two crawl method are obtained from a mapping table maintained by each node.

Third, for each hyperlink and its logarithmic integer, if the corresponding node is numbered as an integer, it will be reassigned; otherwise the hyperlink will be sent to node.

Fourth, calculate the relevance, combined with relevance to determine whether the motive vocabulary type is a necessary one. If not, then skip; if it is, then continuing the analysis.

Fifth, read the required motivation vocabulary using regular expression matching method to find the information vocabulary motivation words, and recorded.

Sixth, the recorded vocabulary of motives is stored according to the predetermined form, so as to achieve the acquisition of vocabulary-related corpus information, the representation expression as follows:

$$Sim_{improve} = 0.8 * Sim_{\cos} + \frac{1}{n} * \sum_{i=1}^{n} Sim_{med} * \frac{(value1 + value2)}{2} \tag{9}$$

where Sim_{cos} is the motivation word feature, Sim_{med} is the motivation word information.

2.2. **Clustering of Motivation Vocabulary Information.** We applied singular value decomposition (SVD) for motivation vocabulary clustering based on the motivation corpus that has been obtained.

We first defined the average amount of information with a single motivation word appears in the corpus as:

$$W(w) = 1 + \frac{1}{\ln(n)} \sum_{i=1}^{n} P_i(w) \ln\left[P_i(w)\right] \tag{10}$$

where P$_i$(w) is the probability of single word w appear in the corpus, n is a constant. The larger W(w), more average information is contained in single word w, indicate the more common of the word which can be regarded as noise to filter out. If both of the average amount of the appearance of a motive word in the sentence and the average amount of the sentence including this motive word are large, it means that this word is very common. and the joint entropy W$'$(w) can be defined as:

$$W(w) = H(w) + H(s|w) \tag{11}$$

where H(w) annotates the average information for a single word appear in a sentence and can be defined as:

$$H(w) = -\sum_{j=1}^{n} P_j(s|w) \log(s|w) \tag{12}$$

and H($s|w$) annotates the average information of the sentence contains the single word appearing in the corpus:

$$H(s|w) = -\sum_{l=1}^{n} P_l(s|w) \log P_l(s|w) \tag{13}$$

The probability of single word w show in the corpus j is denoted as P$_j$(w):

$$P_j(w) = \frac{f_j(w)}{\sum\limits_{j=1}^{n} f_j(w)} \tag{14}$$

The probability of the sentence with the word w appear in the corpus is denoted as P$_l$($s|w$):

$$P_l(s|w) = \frac{f_l(s|w)}{\sum\limits_{l=1}^{n} f_l(s|w)} \tag{15}$$

where f$_j$(w) is the frequency of word w show in corpus, n is the motivation word number in the corpus, f$_l$($s|w$) is the frequency of sentence s containing word w show in corpus l, m is the total vocabulary size in corpus.

Second, select the motivation word features through computing the threshold Assume TF is the word frequency, represent feature t$_k$ showing frequency in the corpus. IDF is the inverse corpus frequency, denotes as IDF=log(N/n), where N represents the total motivation word number in the corpus. The basic idea for IDF is the less information for a certain feature t_k, the larger IDF, indicating the feature t_k has good classification ability. Then the TF- IDF algorithm is as follows:

$$TFIDF\left(t_k\right) = TF\left(t_k\right) \times IDF\left(t_k\right) = TF\left(t_k\right) \times \log \frac{1}{DF\left(t_k\right)} \tag{16}$$

Above equation (16) has been improved in order to reduce the influence of TF on the weights:

$$TFIDF'\left(t_k\right) = \left[0.5 + \frac{0.5 \times TF\left(t_k\right)}{TF_{\max}\left(t_k\right)}\right] \times \log \frac{1}{DF\left(t_k\right)} \tag{17}$$

In order to calculate the expectation cross entropy of each feature t_k, we choose the predefined number of best features as the subset of feature:

$$f\left(t_k\right) = P\left(t_k\right) \sum_{i=1}^{n} \left(C_i \left| t_k \right.\right) \log \frac{P\left(C_i \left| t_k \right.\right)}{P\left(C_i\right)} \tag{18}$$

where $P(t_k)$ is the appearance frequency of feature t_k, $P(C_i|t_k)$ is the probability of class C_i given the feature t_k, and $P(C_i)$ is the margin probability of class C_i.

Third, we calculate the motivation vocabulary based on the information acquistition method. When the information acquired for motivation word feature t_k greater than a given threshold, then we accept the feature, with the following formula:

$$I\left(t\right) = \sum_{i=1}^{n} p_i \times \log p_i \tag{19}$$

where n is the dimension in the feature set, p_i is the showing frequency of the current word feature. The mutual information of feature t_k and class C_i represents the correlation of feature and class. The mutual information of feature t_k is denoted as:

$$MI\left(t_k\right) = \sum_{k=1}^{n} p\left(C_i\right) \times \log \left({}^{p\left(t_k/C_i\right)}\!/_{p\left(t_k\right)} \right) \tag{20}$$

Fourth, we applied SVD for motivation vocabulary clustering. Let k be the key words elements in motivation word vector. The motivation word vector is constituted by n features. Because of the feature size differences for each motivation words, we can regard that they belong to different feature spaces, and in order to apply clustering on different motivation words, dimension reduction method is required. Therefore, SVD is used to reduce the disturbance of undesired information in the word feature vector. Let A be the word feature vector, the SVD process can transform A to A_k:

$$A_k = U_{m\times m} \sum\nolimits_{m\times n} V^T{}_{n\times n} = t \sum\nolimits_{i=k}^{k} u_i \sigma_i v^T{}_i \tag{21}$$

where u_i and v_i represent the feature vector and the semantic space of the motivation words. Similarly, in the process of calculating the context similarity, it is also required to project the word feature vector to the same k row dimensions as A_k. And the projected vector t' is as follows:

$$t' = tU_k^T \sum\nolimits_k^{-1} \tag{22}$$

After the projection, we obtained the similar vector for initial vector. Then SVD is applied to clustering. Assuming we have motivation sequences (X, s) and (X, d), where X represents a set of sample words, s and d represents the standard criterial to measure similarity or dissimilarity between samples. $Let C = \{C_1, C_2, ..., C_i, ..., C_k\}, i = 1, ..., k$, is a subset of X, we have:

$$X = C_1 \cup C_2 ... \cup C_k \tag{23}$$

For any i≠j, we have $C_i \cap C_j = \emptyset$, each C_i is a cluster. Then our clustering goal is to get higher similarity of the motivation words in the same cluster, and it can be represented as:

$$F(j) = \max\left(\sum_{C_j \in D} \sum_{d_i \in C_i} sim(d_j, s_j)\right) \tag{24}$$

In sum, after obtaining motivation corpus, we applied SVD on motivation feature vectors for clustering, which provided the basis for creative idea generation.

2.3. **Bayesian network classification for motivation word for creative idea calculation.** After clustering the motivation words, we applied Bayesian network to calculating the words relativity and use ontology [13] theory to generate creative ideas. The basic

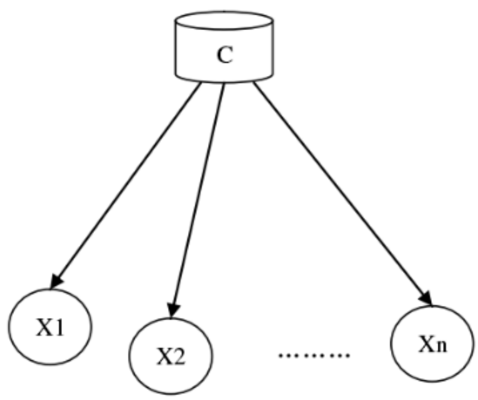

FIGURE 4. The Bayesian network model for motivation word classification

idea of Bayesian network is to use prior information and the likelihood of total motivation vocabulary information to estimate posterior information. When calculate the motivation relativity, except considering the probability of motivation word A, the conditional probability of A given word B is also required, denotes as $P(A|B)$, and the calculation formula as follows:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{25}$$

where $P(A \cap B)$ represents the joint probability and P(B) is the marginal probability of word B. Assume the probability space (Ω,F,P) contains $A_i \in \Omega$(i=1,2,...,n), $A_i \cap A_j = \emptyset\, (i \neq j)$ and $\cup_{i=1} A_i = \Omega$, then for any motivation word B∈F and P(B)>0, we have:

$$P\left(A_i \left| B\right.\right) = \frac{P\left(B \left| A_i\right.\right) P\left(A_i\right)}{\sum\limits_{j=1}^{n} P\left(B \left| A_i\right.\right) P\left(A_i\right)}\,(i, j = 1, 2, ..., n) \tag{26}$$

where P(A_i)>0 is the prior distribution and P(B|A_i) is the conditional probability, $P(B) = \sum_{j=1}^{n} P(B|A_i)P(A_i)$ is the full probability. Let C denotes the motivation class nodes, $X_1, X_2, ..., X_n$ denote n attribute nodes, then the Bayesian network can be represented in figure 4.

Then we apply the classification process of Bayesian network on motivation word. First the motivation word X can be represented as a n dimension feature vector $X = (t_1, t_2, ..., t_n)$, where $t_i, i = 1, ..., n$ represents n features for the word. Then assume we have m classes for motivation corpus denotes as $C_1, C_2, ..., C_m$, for a given word X, we calculate the maximum posterior for a class, and assign the class C_i to corresponding word if and only if:

$$P\left(C_i \left| X\right.\right) > P\left(C_j \left| X\right.\right), i \neq j \tag{27}$$

Based on the Bayes theory we have:

$$P\left(C_i \left| X\right.\right) = \frac{P\left(X \left| C_i\right.\right) P\left(C_i\right)}{P\left(X\right)} \tag{28}$$

Because of P(X) is constant for the given class, we can transform the posterior proportion to:

$$P\left(C_i \left| X\right.\right) \propto P\left(X \left| C_i\right.\right) P\left(C_i\right) \tag{29}$$

The prior distribution of P(C_i) can be estimated from the training text sample, which is P(C_i)=s_i/s, where s_i is the total trainingmotivation word size in class C_i, s is the total training motivation word size. Then we only need to maximize the conditional probability of P(X|C_i). Assume the independency of a naive Bayes network, we have:

$$P\left(X \left| C_i\right.\right) = P\left(t_1, , t_2, ..., t_n \left| C_i\right.\right) = \prod_{k=1}^{n} p\left(t_k \left| C_i\right.\right) \tag{30}$$

where probability $P(t_1|C_i), P(t_2|C_i), ..., P(t_n|C_i)$ can also be trained from the training samples, represented as the probability of t_k shown in class C_i:

$$p\left(t_k \left| C_i\right.\right) = \frac{count(t_k \left| C_i\right.)}{\sum\limits_{k=1}^{n} count(t_k \left| C_i\right.)} \tag{31}$$

where count(t_k |C_i) represents the appearance number of feature vector t_k show in class C_i, $\sum_{k=1}^{n} count(t_k|\, C_i)$ represents the total appearance number of features in class C_i. The motivation vector sparsity will result in some feature not showing in the training corpus samples, which will result in the P(X|C_i)=0 based on (30). To avoid such case, we applied Lapalace smoothing correction, represented as:

$$p\left(t_k\left|C_i\right.\right)=\frac{count(t_k\left|C_i\right.)+\delta}{\sum\limits_{k=1}^{n}count(t_k\left|C_i\right.)+n\delta} \tag{32}$$

where n is the total size of motivation feature, or the feature dimension size, δ is any non-zero constant and normally set 1.

Finally, we can calculate the maximum likelihood based on the Bayesian network for classification:

$$C=\arg\max_{i=1,2,\ldots,m} P\left(X\left|C_i\right.\right)P\left(C_i\right) \tag{33}$$

And the correlation of motivation words can be represented as:

$$A\left(S_1,S_2\right)=\max\left(1-\frac{d\left(p_i,p_j\right)}{D}\right)\left(1\leq i.j\leq 2, i\neq j\right) \tag{34}$$

where p_i and p_j are the significant feature for motivation word S_1 and S_2, D is the lateral association influence depth, $d(p_i,p_j)$ as the appearance size of motivation feature p_i in corpus p_j. The correlation between motivation words can be represented by their relativity, denotes as:

$$relate_{gloss/texe(t_1,t_2)}=\tanh\left(\frac{overlap\left(t_1,t_2\right)}{length\left(t_1\right)+length\left(t_2\right)}\right) \tag{35}$$

Then we can use ontology theory to construct creative idea phrases and sentences based on the motivation words for each classes as follows:

$$\alpha_R\left(X\right)=\frac{\left|RX\right|}{\left|\bar{R}X\right|} \tag{36}$$

where $X\neq\emptyset$, $|X|$ represents the basis of motivation word set X. When α_R (X)=1, the generated creative idea is the optimal solution.

3. RESULTS.

3.1. Experiment data and configuration. In this experiment we used Reuters-21578 as experimental corpus, which includes 22 files and 21578 motivation words. We separate the corpus into training data set with 9603 training samples, testing data set with 3299 testing samples and other 8676 samples unused. We divided the motivation words in Reuters-21578 into 135 classes. Each motivation word can be classified to maximally 14 classes and minimally 1 class. The top-10 classes with the maximum motivation word size is shown in table 1.

Given a class and a motivation word, the goal is to determine whether the word belongs to the class. We evaluate the performance with recall, precision and fallout rate, denotes as:

$$recall=\frac{a}{a+c} \tag{37}$$

$$precision=\frac{a}{a+b} \tag{38}$$

Table 1. Top-10 classes with the maximum motivation word size

Class name	Quantity of vocabulary
earnings	2709
acquisitions	1488
money-fx	460
crude	349
grain	394
trade	337
interest	289
ship	191
wheat	198
corn	160

$$fallout = \frac{b}{b+d} \tag{39}$$

where a represents the correct number of classification for motivation words, b represents the incorrect number of classification for motivation words, c represents the false negative rate and d represents the false positive rate.

3.2. **Experiment results analysis.** We tested our model on the given corpus. The results are shown in table 2. From table 2, we can see the top-10 training classes in Reuters-21578 can achieve an average recall rate at 87.6%, the average precision rate is 83.1%, where the class acquisition has the highest recall rate with 97% ,corn has the lowest recall rate at 63%, earnings has the highest accuracy rate at 93% and corn has the lowest accuracy rate at 70%. From the table, we can also observe that the precision and recall curve also increases as the training sample size increases.

Table 2. The precision and recall rate for the top-10 classes with the maximum motivation word size

Class name	Training motivation number	Testing motivation vocabulary	recall rate (%)	precision rate (%)
earnings	2709	1014	0.95	0.94
acquisitions	1488	630	0.97	0.93
money-fx	460	133	0.94	0.85
crude	349	160	0.96	0.80
grain	394	130	0.88	0.74
trade	337	106	0.92	0.85
interest	289	95	0.84	0.85
ship	191	82	0.89	0.82
wheat	198	65	0.78	0.83
corn	160	46	0.63	0.70
SUM	6575	2461	0.933	0.882

To show the feasibility of our proposed classification model, we also compared it with the existing algorithm of information feedback method and heuristic rule method. Figure 5, 6, 7 shows the recall, precision and fallout rate of the three models as the motivation vocabulary size increases.

From fig. 5 we can see when the information feedback method is applied, the average recall rate is 12%, and the recall rate decreases as the amount of motivation vocabulary increases. The average recall rate is 10% for heuristic rule method, and we observed fluctuation when the amount of motivation vocabulary is between 140 to 180; and the

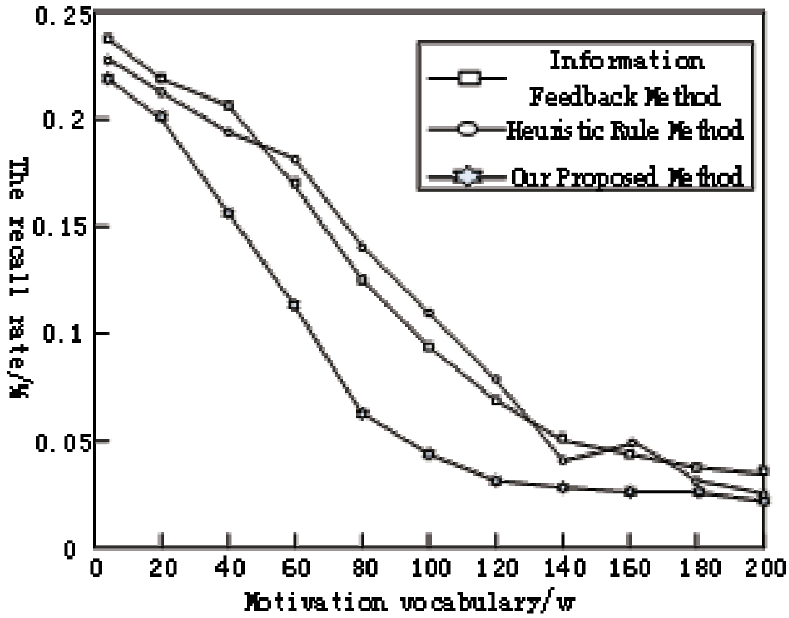

Figure 5. The recall rate of three methods

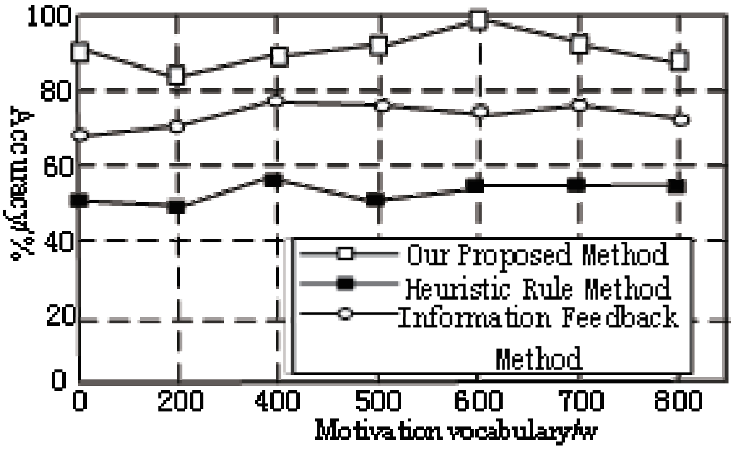

Figure 6. The precision rate of three methods

average recall rate of our proposed method is 8% and also consistently decreases as the amount of motivation vocabulary increases.

Figure 6 depicts the precision rate of the three methods. When the information feedback method is applied , the average precision rate is 76.2%, and as the amount of motivation vocabulary increases, the precision rate fluctuates between 200-400 and 600-800.

The average precision rate is 58.4% for heuristic rule method, and the performance does not vary much but suffers 17.8% degradage on average performance comparing with feedback method. While the average precision rate for our proposed method is 94.3%, improved by 18.1 percent point and 35.9 percent point compared with other two methods and the precision rate remains stable as the amount of motivation vocabulary increases.

Figure 7 shows the fallout rate of the three methods. When the information feedback method is applied, the average fallout rate is 32.2%. And as the amount of motivation

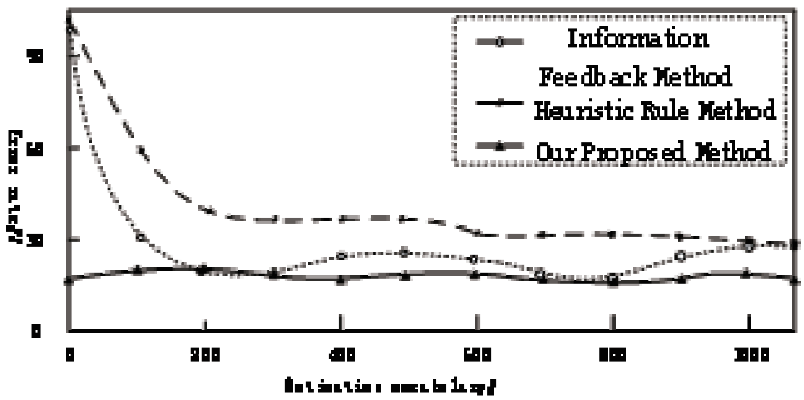

Figure 7. The fallout rate of three methods

vocabulary increases, the fallout rate first decrease rapidly before the amount of vocabulary reaches 200 and fluctuate when the amount of the vacabulary is between 200-1000. The average fallout rate is 48.4% for heuristic rule method, and the performance did not vary much after 200 motivation words but suffered 16.2% more fallout rate on average performance compared with feedback method; and the average fallout rate for our proposed method is 18.3%, improved by 13.1 percent point and 30.1 percent point compared with other two methods and the fallout rate remains stable as the amount of motivation vocabulary increases.

4. **Conclusion and Future Work.** In this paper, we have shown a feasible way to transform from blurry motivation to clear creative idea. Based on crawler technique we can obtain motivation vocabulary and then we have applied Bayesian network to classification of motivation words and use ontology theory to construct creative idea phrases and sentences. After comparing our improved method with other existing methods, we can see its high recall rate and high precision.The idea creation can be applied in many interdisciplinary domains and multidisciplinary domains. It will be better in the future. In the future we will use deep learning technology to further improve it.This technology can better imitate the emergence, self-organization, self-coordination of creative idea generated by the genuine brain.

Acknowledgment. This material is based upon work supported by the Key Technologies Research and Development Program of China Foundation under Grants No. 2012BAH38F05 and also supported by the Research Program Foundation of Minjiang University under Grants No. MYK16002. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Key Technologies Research and Development Program of China Foundation and the Research Program Foundation of Minjiang University

REFERENCES

[1] L. A. Gatys, A. S. Ecker, and M. Bethge, A neural algorithm of artistic style, *arXiv preprint* , arXiv:1508.06576, 2015.
[2] Delin Jing, Hongji Yang, and Ying chun Tian, Abstraction Based Domain Ontology Extraction for Idea Creation, *13th International Conference on Quality Software* , 2013.
[3] A. Hugill, Creative Computing Processes: Musical Composition, *IEEE 8th International Symposium on Service Oriented System Engineering*, 2014.

[4] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, Organization, development and function of complex brain networks, *Trends in cognitive sciences*, vol. 8, pp. 418-425, 2004.

[5] L. Zhang,and H. Yang, Definition,Research Scope and Challenges of Creative Computing, *Proceedings of the 19th International Conference on Automation & Computing*, 2013.

[6] Dalton L A., Application of the Bayesian MMSE estimator for classification error to gene expression microarray data, *Bioinformatics*,2011, 27(13): 1822-1831.

[7] J. GAO, B. MEI, Research on Microblog Advertisement Filtering Model Based on Text Content Analysis, *Computer Engineering*,2014, 40(5):17-20.

[8] Kossowska M, Bar-Tal Y, Need for closure and heuristic information processing: The moderating role of the ability to achieve the need for closure, *British Journal of Psychology*,2013, 104(4):457-80.

[9] Roul R K, Devanand O R, Sahay S K, Web Document Clustering and Ranking using Tf-Idf based Apriori Approach, *Eprint Arxiv*, 2014, 5(2):34-39.

[10] Trstenjak B, Mikac S, Donko D, KNN with TF-IDF based Framework for Text Categorization, *Procedia Engineering*, 2014, 69(1):1356-1364.

[11] Rangabhashiyam S, Anu N, Nandagopal M S G, et al, Relevance of isotherm models in biosorption of pollutants by agricultural byproducts, *Journal of Environmental Chemical Engineering*,2014, 2(1):398-414.

[12] Alkaoud M, Ashshohail I, Ismail M M B, Automatic Image Annotation Using Fuzzy Cross-Media Relevance Models, *International Journal of Image & Graphics*, 2014, 2(1):59-63.

[13] N. Guarino, Formal ontology, conceptual analysis and knowledge representation, *International journal of human-computer studies*, vol. 43, pp. 625-640, 1995.