

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(13), 2014 [6996-7002]

Differentiation information fusion-based semantic information retrieval model research

Jianhao Xu

Information Engineering, Nanning Vocational and Technical College, Nanning, Guangxi, (CHINA)

ABSTRACT

For a complicated semantic information retrieval, improve its information retrieval system user experience extent can improve its query optimization algorithm, by utilizing literature searching and other methods, it finds that present semantic detection mainly is in the ontic query technique and semantic annotation, such can only carry out surface semantic correlation simple searching but cannot go deeper into it, on this basis, the paper designs differentiation information fusion semantic analysis detection model, constructs model by three aspects, and applies the model in concrete examples, and gets the modal rationality and effectiveness, finally targeted at the model, it puts forward that the model is to be further improved, outlooks the model applications in more fields.

KEYWORDS

Differentiation information; Fusing semantics; Information retrieval; Mathematical model.



INTRODUCTION

Knowledge management and knowledge rising trigger tendency of social intellectualization, extensive application of computer overcomes institution inferiority and time barriers, let scientific researchers to connect all kinds of information and therefore construct technology-based environment system, just in such background, massive amounts of information rapidly extraction and semantic rapidly searching have become hotspot of current researches.

In semantic information retrieval aspect, lots of people have made correlation research, and been widely used, such as: in ontic information semantic detection, Yu Fan had ever proposed ontic correlation retrieval, from which it concluded ontic classification, concept, functions, as well as semantic description and so on, described semantic information retrieval model into blocks forms that was composed of every frames, and put forward overall working flow, in system key techniques, it proposed inference technique, by semantic retrieval tools and full-text retrieval tool to search, and simply introduced storing schemes and systematic visual schemes; Zhang Jing and others in multimoding information fusion video retrieval, proposed relational algebra shape model, in the model, it made full use of videos' contained text, high level semantics and other features, built used relational algebraic expression under multiple videos, and carried out fusion by experiments, in complex semantic retrieval, indicated the model advantages; Yue Jun-Ping had ever proposed concurrence analysis semantic flow that constructed by utilizing literature investigation and other methods in concurrence analysis semantic information retrieval, and targeted at world archaeology books, and made researches, finally got concurrence analysis-based semantic retrieval that had great advantages.

The paper bases on formers research, targeted at differentiated information fusion semantic information detection model, it makes analysis and researches, and establishes differentiation information fusion semantic information detection model, targeted at its performance, it detects and puts forward reasonable suggestions.

SEMANTIC RETRIEVAL MODEL RESEARCHES IN VOICE INFORMATION FUSION VIDEO

Semantic information is a kind of information that provided by any language, data, words, symbols through any path, in general is a stationary state that things owned and unknown in advance, is a kind of comprehensive media information that integrates text and voice, on the basis of semantic information possessed complex, it is hard to use one way to describe it and get good efficiency, differentiation information fusion formally handles with voice and texts as well as other information. By effective ways, it carries out fusion and gets retrieval blocks' consulting results.

Feature retrieval

Differentiation information fusion semantic information retrieval is decomposing and fusing different mode states, according to semantic information different forms, letting these retrieval models ways also not to be same structures, which need to implement most perfect retrieval result according to different methods, and meanwhile it can also use up-to-date database to execute canceling and adding as well as other functions on them so as to implement retrieval system dynamic updating function.

Text research mainly makes query for voice included text information, implements scripts information by automatic identification, text information makes comprehensive handling with it, after that through handling way that corresponds to straight line, it processes and then gets key words in established semantic texts to implement search structure and search, and ranks every result confidence level, for present research status, semantic information retrieval basic technical routes and basic conceiving ways are as following Figure 1 shows:

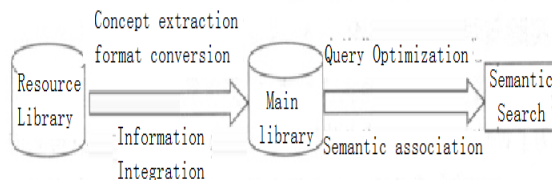


Figure 1: The basic model of semantic information retrieval

By above Figure 1, we can get semantic information format conversion, concept integration and information extraction as well as other correlation techniques storing ontic base, through applying semantic correlation and query optimization as well as other ways, it shows information in front of practitioners. Here proposes information extraction and integration and others, all are acquired by semantic annotation, by searching semantic correlation and optimization relative principles, it implements ontic-based semantic relations and query techniques.

Traditional forms semantic information retrieval, in general, it solves result by utilizing matched ways, from which in semantic information retrieval, it includes correlated all kinds of information in detection, in big era of data today, mathematical library retrieval system implements surface correlated semantic analysis detection, to implement detection results should carry out semantic correlation on important key techniques in searching semantic optimization process so that can get new retrieval flow.

DIFFERENTIATION INFORMATION FUSION SEMANTIC INFORMATION DETECTION MODEL RESEARCH

By referencing domestic and overseas correlation information, the paper designs a kind of differentiation information fusion semantic information detection model, corresponding flow chart is as following Figure 2 shows:

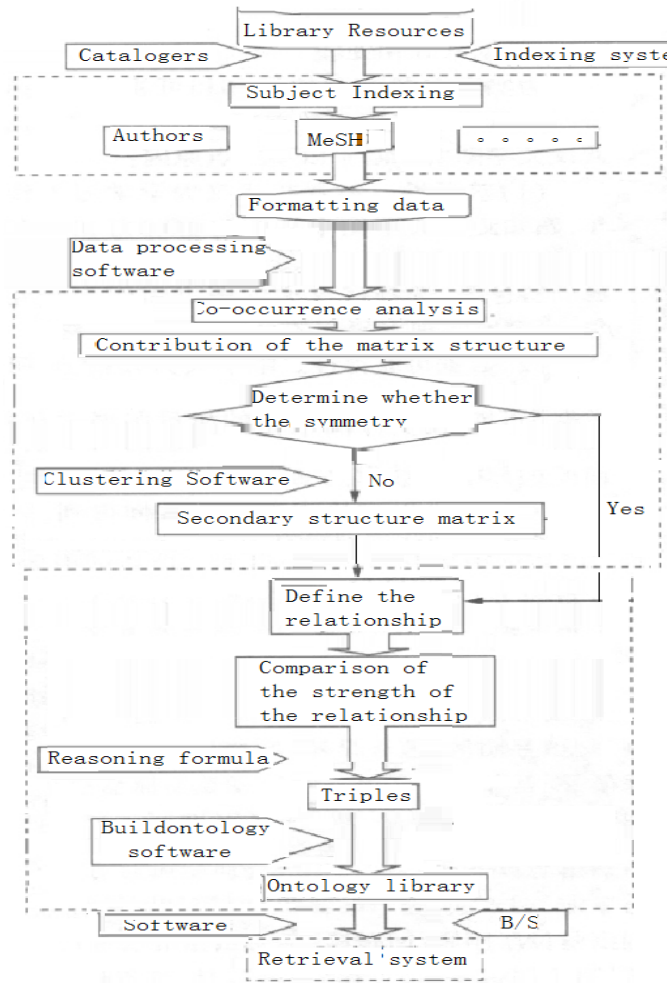


Figure 2 : Semantic information retrieval model differentiated information fusion

Differentiation information fusion semantic information detection model can divide into three parts, from which it respectively is joint composed of ontic structure, normalize and analyzer three parts, corresponding every part is composed by different designers and equipped with corresponding technical equipments. In model design, feasible retrieval semantic relation is using concepts strong and weak as its detection tool, semantic similarity quantization refers to concurrence analysis structure that uses concept features to common decide similarity, it is quantitative descriptive semantic similarity; another one is using concept extraction and convenient science, library resources as model main processing objects, making comparison by external features, and meanwhile labels external features after thematic processing, so that obtained data is also a kind of scientific data. Current researched semantic information detection, in general, it is summarizing of fuzzy words, carry out quantitative handling with semantic relations, therefore obtained detection rate is relative higher.

Differentiation information fusion semantic information retrieval model's verification

In order to verify above model, the paper introduces differentiation information fusion is in the voice detection system, by carrying out sports videos semantic contents research on segmentation and classification, and regulates upmost side one layer is out of play and in play constructed semantics in game videos, and can transfer probability, medium layer is constructed different differentiation information fusion targeted at different time sequence relations, and bottom layer is input layer observed numerical value, output is optimal time sequence value, and on the basis of above provided such model frame, we also provide flow chart regarding out of play and in play, its Figure 3 is as following shows:

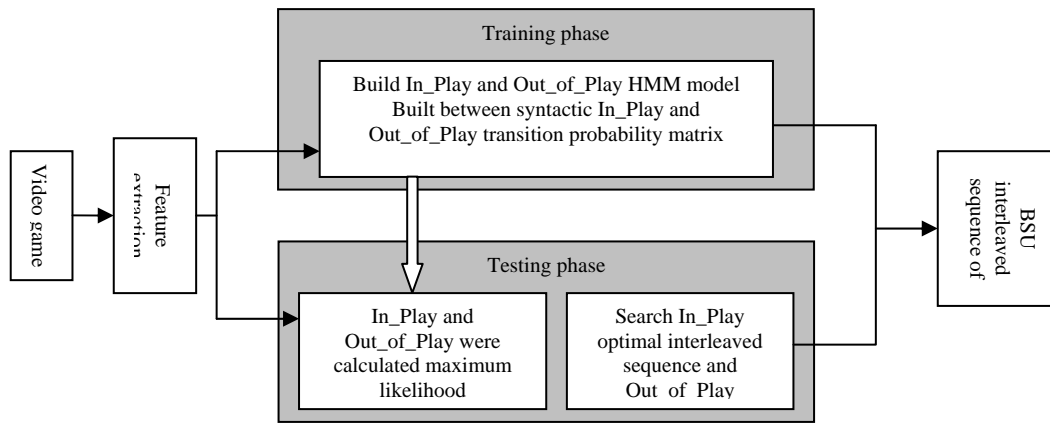


Figure 3 : Based in play and out of play model classification and segmentation of the basic processes

By above Figure 3, we can get the flow is composed of test and training two stages, so it needs to establish out of play and in play different topological structural differentiation information fusion model, in the model, HMM_o and HMM_I respectively represent out of play and in play differentiation information fusion model, that:

$$HMMs = \{HMM_I, HMM_o\} \tag{1}$$

Among them

$$HMM_I = \{HMM_{I_1}, \dots, HMM_{I_M}\}$$

$$HMM_o = \{HMM_{o_1}, \dots, HMM_{o_N}\}$$

Among them, M pieces of different In Play models are using $HMM_I = \{HMM_{I_1}, \dots, HMM_{I_M}\}$ to express, N pieces of different Out Of Play models are using $HMM_o = \{HMM_{o_1}, \dots, HMM_{o_N}\}$ to express, then corresponding state layer is:

$$S = \{Loose_view, Medium_view, Tight_view\} \tag{2}$$

In the stage, we also adopt front-end algorithm and corresponding different differentiation information fusion probability numerical values, and it should assign maximum value as node, then:

$$P(O_i / HMM_I) = \max\{P(O_i / HMM_{I_1}), \dots, P(O_i / HMM_{I_M})\}$$

$$P(O_i / HMM_o) = \max\{P(O_i / HMM_{o_1}), \dots, P(O_i / HMM_{o_N})\}$$

Finally by dynamic programming method, it carries out searching and then gets optimal staggered sequence.

By adopting K groups of training numerical values as $O = [O^{(1)}, O^{(2)}, \dots, O^{(K)}]$, respectively train Out of Play and In Play, that:

$$a_{ij} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \xi_t^{(k)}(i, j)}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \gamma_t^{(k)}(i)} \tag{3}$$

$$\sum_{l=1}^K \sum_{t=1}^{T_l} \gamma_t^{(l)}(j) \quad (4)$$

$$\bar{b}_j(k) = \frac{s.t.O_t^{(l)} = v_k}{\sum_{l=1}^K \sum_{t=1}^{T_l} \gamma_t^{(l)}(j)} \quad (5)$$

$$\bar{\pi} = \frac{1}{k} \sum_{k=1}^K \gamma_1^{(k)}(i) \quad (6)$$

$\gamma_t(j)$ and $\zeta_t(i, j)$ respectively represent on condition that observation sequence and model is given, t instant under S_i probability and S_i state $t+1$ instant probability.

For above model, it tests and by TRECVID sports video, it searches, and defines searched structure, included semantic information is previous solved differentiation information fusion syntactic transferring matrix, after that run maximum likelihood estimation so as to solve staggered sequence optimal solution, in the process, let $w \in [0.1, 0.4]$, we respectively intercept basketball and football correlation partial wonderful voice segments, and make research on it accuracy, and list out the precise concrete numerical value, as following TABLE 1 shows:

TABLE 1: Differentiation information fusion model-based semantic content precise analysis

	Test data	The paper's method
Football semantic	Soccer1	0.868
	Soccer2	.904
	Soccer3	0.892
	Soccer4	0.885
	Soccer5	0.9180
Average precise		0.8920
Basketball semantic	Basketball1	0.861
	Basketball2	0.845
	Basketball3	0.823
	Basketball4	0.8962
	Basketball5	0.891
Average precise		0.8837

By above TABLE 1, we can get the model accuracy nearly arrives at 89%, it proves the model is considerable reasonable.

Differentiation information fusion semantic information search model performance test

Differentiated information fusion semantic information retrieval is storing data by structures, and deciding knowledge compactness extent by distance, assume searched key word is a knowledge name that its distance is one, nanometer extracted bulletin, notice, decision and opinions, orders and others as key words, it respectively carries out differentiation information fusion semantic and goes ahead with key words searching, obtained searching result is as TABLE 2-3 shows:

TABLE 2 : Key words detection result

Detected files	Differentiation detection	All files	Detected files	Precision ratio	Recall ratio
89	Order	100	84	0.84	0.89
86	Decision	100	82	0.82	0.86
83	Opinion	100	83	0.83	0.83
88	Bulletin	100	86	0.86	0.88
87	Notice	100	79	0.79	0.87

TABLE 3 : Differentiation detection result

Detected files	Differentiation detection	All files	Detected files	Precision ratio	Recall ratio
98	Order	100	88	0.88	0.98
95	Decision	100	86	0.86	0.95
94	Opinion	100	83	0.83	0.94
97	Bulletin	100	84	0.84	0.97
96	Notice	100	82	0.82	0.96

In order to more clear present mutual relations, the paper draws bar figure, as following Figure 4-5shows:

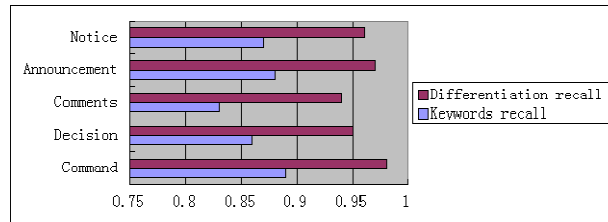


Figure 4 : Recall comparison chart

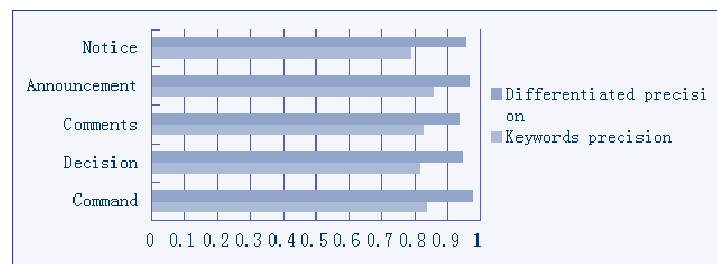


Figure 5 : Precision comparison chart

By above comparison figure, we can find that differentiated information fusion semantic information detection presents a kind of considerable significant state in recall aspect, especially in key words information detection, from previous overall precision ratio, it can also get that differentiated information fusion semantic information’s detection is obviously superior to key words detection, so it proves differentiated information fusion semantic information detection model has unique advantages.

CONCLUSION

With computer rapidly development, information amount is also ceaseless improving, detect all kinds of information is focus of current research in the field, the paper establishes differentiation information fusion semantic information detection model, the model fuses multiple information together in case integration, it shows extreme importance in complex and changeable semantic retrieval, constructs multiple semantic retrieval blocks, and combines with experiments to verify, and gets the model has stronger superiority.

But in the field, differentiated-based information fusion semantic information search model research still has more wider exploration space, the paper just takes searching semantics key words as examples to analyze, and the model still has more extensive development field and cannot be described by only one method one technology, therefore it needs more scholars to further study and excavate so as to adapt to wider fields.

REFERENCES

- [1] Guo Jia-Qi; Field ontic construction and its application research in information retrieval, Beijing University of Posts and Telecommunications, **6**, 12 (2007).
- [2] Dong Hui, Yu Chuan-Ming, JiangYing, etc; Ontology-based digital library searching model research (II)— semantic information extraction[J], Journal of Intelligence, **8**, (2006).
- [3] Sun Zheng; Ontology-based information retrieval system research in network environment, Computer engineering and design, **23**, 53921-5398 (2009).

- [4] Liu Hong-Zhe, Xu De; Ontology-based semantic similarity and correlation degree calculation research summary *Computer science*, **2**, (2012).
- [5] Fu Yi-Fan; Rete algorithm-based rules engine design and application in subject intelligence learning guidance, *Northeast Normal University*, **5**, 6-17 (2009).
- [6] Shu Yan; Ontology-based information retrieval research and application, *Gui Zhou University*, **4**, 1-9 (2007).
- [7] Wang Jin; Ontology-based semantic information retrieval research, *University of Science and Technology of China*, **5**, 9-18 (2006).
- [8] Wang Yue-Fen, Song Shuang, Miao Lu; Research on concurrence analysis application in knowledge service, *Modern Library and Information Technology*, **4**, (2006).