ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

A Hybrid BTF Approach for Temporal Query Processing In the Patented Medical Database

T. Yogameera[1], Dr.D.Shanthi Saravanan[2]

PG Student, Department of CSE, P.S.N.A College of Engineering and Technology, Tamilnadu, India[1]

Professor, Department of CSE, P.S.N.A College of Engineering and Technology, Tamilnadu, India[2]

ABSTRACT: The Intellectual Query Processing has become mandatory for efficient Information retrieval and Management system. The traditional approaches are not sufficiently enough in dealing with the issues of upcoming temporal events, in the context of MIS in data mart or in enterprise databases. The proposed system provides an effective methodology with three streamlined techniques namely Topic Relevant Query Suggestion with extended Query Augmentation to enhance the functionality of patented data search in high dimensional databases besides anonymizing the sensitive data. On posting a query in the patented database by experts, the reference records are first clustered into topics and classes and the highly coherent cluster partitions are obtained as result. The upshots in each coherent cluster are combined generating top K relevant answers for the examiners. By going through a detailed analysis on different literatures of processing and retrieval of information, the proposed approach amplifies the user's intention contour and enhances the retrieval time with more efficient memory management. The system provides the information in a graphical part with additional filter criteria. Also the proposed system puts forward the inference control strategy to avoid privacy leakage thereby performing double encryption in the datasets based on the severity. Further investigation of efficient scheduling of intermediate data sets in cloud is considered to update the dataset in-case of dynamic change. This technique when implemented in patented medical database will give better results in Query processing in terms of accuracy and cost economy in Electronic Data Interchange (EDI).

KEYWORDS: Automatic Error Correction, Query Augmentations, Query Analysis, Referential Medical information management, patented database structure, inference control, expert opinion, knowledge sharing, privacy.

I.INTRODUCTION

The intention of this paper is to render referential medical information management system. The proposed system deals with query processing in multidimensional data from reference sites for expert analysis, satisfying the relevancy of search. The sensitive data publishing is dealt with anonymized temporal events. The proposed three proven streamlined techniques viz. Automatic Error correction, Topic Relevant Query suggestion, query augmentation helps in précised query contour and quick retrieval of referential data from the patented medical databases in the hospital data mart which holds big data of a medical forum.The proposed system provides the ranked list of reference records when searched for opinion on treatment of acute cases. With careful analysis the patient can be treated with correct and available advanced technologies. This model is based on the notion of differential privacy that is applicable to k-anonymized data sets. It puts forward the inference control strategy to avoid privacy leakage thereby performing double encryption in the datasets based on the severity. Further to support this special technique efficient scheduling of intermediate data sets using temporal pattern matching algorithm is suggested in order to preserve the dataset dynamically. The Patented Medical Databases is now being used for referential report generations with detailed and analysed metadata structure to implement the referential analysis with the automated machines, so that the records must be with accurate analyzation for the treatment after observations recorded. In acute syndrome cases and other immediately diagnostic needed cases like echo cardio problem, cancer etc. we are in need of proper less time consuming appraised information retrieval system to be designed. The common approaches like As You Type, Prior Art Relevancy, Type Ahead Search, Click through Data, SVM Ranking have not proven efficient in this era of medical treatments. To solve this problem, the patented medical database is

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

now being opted for referential report generations with detailed and analysed metadata structure. The Research is underway to implement the referential analysis with the automated machines and with human robots, when implemented in cloud it will help for experts to make decision based on foreign enhanced metadata. Privacy is mandatory for sensitive medical data since these are sources for the allocation of public funds, medical research and statistical trend analysis.

II. EXISTING SYSTEM

As pattern search help the end users to find previously published associated patterns it has become more and more popular. It is performed by both business enterprises and academic communities. But most of the pattern-search users have very less knowledge about the available patents. Hence they tend to use a try-and see approach or prior art relevancy by frequently issuing queries and checking solutions. Hence it is a time consuming process.

Disadvantages of Existing System: The major disadvantage is the relation between a particular sensitive data with the other data cannot be identified properly with edit distance alone. Frequent access pattern on the data may get changed in timely manner with updated technologies becomes the problem statement. More-over, no proper constrain satisfaction is obtained.

III. LITERATURE SURVEY

A User-Friendly Patent Search Paradigm [1]**,** the paper proposes techniques like error correction, prior art query suggestion, and query expansion, to improve the usability of patent search, also how to efficiently find relevant answers from a large collection of patents. First partition patents into small partitions based to their classes. Then, given a query, highly relevant partitions are used to answer the query. Finally, the answers are combined from each partition and generate top-k answers of the patent-search query. (Yang Cao et.al, 2013)**,** though the author has discussed about automatic error correction with the partial query keyword and Prior-art search has been a milestone approach in patent right provision, the problems with increased number of results takes more time to analyze the obtained data for initial processing itself; moreover no privacy preservation has been discussed. Query rewriting is yet to be considered in this paper. Efficient Pattern Matching over Event Streams [2], Pattern matching over event streams is increasingly being employed in many areas including financial services, click stream analysis, and electronic health systems. While regular expression matching is well studied, pattern matching over streams presents new challenges like richer languages for regular expression matching. Furthermore, efficient evaluation of these pattern queries over streams requires new algorithms and optimizations. The conventional wisdom for stream query processing (i.e., using selection-join-aggregation) is inadequate. This paper, presents a formal evaluation model that offers precise semantics for this new class of queries and a query evaluation framework permitting optimizations in a principled way. The runtime complexity of query evaluation using this model and develop a suite of techniques that improve runtime efficiency by exploiting sharing in storage and processing. (J. Agrawal et.al, 2011), but this paper has no syntactic pointers to most recent events selected into the buffer in this run. The dependency of shared processing and memory may make the system sluggish .Towards Expressive Publish/Subscribe Systems [3]. Traditional content based publish/subscribe (pub/sub) systems allow users to express stateless subscriptions evaluated on individual events. However, many applications require the ability to handle stateful subscriptions, this paper introduces, a stateful pub/sub system based on nondeterministic finite state automata (NFA). It allows users to express subscriptions that span multiple events, and it supports powerful language features such as parameterization and aggregation, provides non-ambiguous subscription semantics as well as unique opportunities for optimizations. But Query rewriting for multiple query optimizations is yet to be considered. The number of edges that each instance could potentially traverse can be very large. If we index the dynamic predicates, index maintenance becomes much more expensive. Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data [4] with the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search thus enabling an encrypted cloud data services is of paramount importance. Considering the larger number of data users and documents in the order of their relevance to these keywords, related works on searchable encryption focus on single key word search or Boolean keyword search, and rarely sort the results. This paper defines and solves the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE). Proposed system establishes s a set of strict privacy requirements for such a secure cloud data utilization system. By choosing the efficient similarity measure of coordinate matching, i.e., as many matches

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

as possible, to capture the relevance of data documents to the search query. The problem here arises when time complexity has been increased for semantic analysis for all matching documents. Encryption method chosen may not be efficient which may drop the whole aim of the system. Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding for data mining [5], a cloud storage system, consisting of a collection of storage servers, provides long-term storage services over the Internet. Storing data in a third party's cloud system causes serious concern over data confidentiality. General encryption schemes protect data confidentiality, but also limit the functionality of the storage system because a few operations are supported over encrypted data. Constructing a secure storage system that supports multiple functions is challenging when the storage system is distributed and has no central authority. This paper introduces a threshold proxy re-encryption scheme and integrates it with a decentralized erasure code such that a secure distributed storage system is formulated. Also lets a user forward his data in the storage servers to another user without retrieving the data back. The main technical contribution is that the proxy re-encryption scheme supports encoding operations over encrypted messages as well as forwarding operations over encoded and encrypted messages. This method fully integrates encrypting, encoding, and forwarding.

Need of the hour: The above discussed references shows that there is a need of enhanced medical information management system, the criteria to be considered are query processing time for relevantly ranked reference records retrieval, enhanced prediction of possible sufferings for experts analysis and opinion on precautious living. The spatial data analysis reveals as the possibility of disease that may affect as a fortune, so that people can prevent them earlier instead of spending money and time along with physical suffering. More over the reference site creation helps in expert's knowledge and experience sharing so that it provides a way for appropriate treatment in time. Clouding such service software needs to be patented by a standard forum, and avail on demand. The currently mile stoned prior art approach deals only with the presorted and stored datasets for references with no prediction strategies but the need of the hour is to deal with temporal updates with evolving technologies in medical science. Existing system also lacks identifying the intermediate sensitive data for encryption in data publishing. The proposed system deals with these problems as core and has proposed a novel approach. The outcome of this project is a MIS (Management of Information System) software product overcoming the above issues discussed, satisfying the need of the hour.

PROPOSED SYSTEM:

The proposed system introduces a user friendly interface component is used to capture the user's intention and find the relevant data with automated query correction, supported by query suggestion, query expansion. A novel method to select top k relevant results with ranks based on aggregated topic relevancy, query keyword relevancy and prior art relevancy along with efficient inverted indexing approach. Privacy preservation delimits the scope of the proposed system. To help end users to easily find related patterns in their contour, the foremost step is to capture user's intention. That is providing a topic based keyword suggestion to find relevant records and filtering from the list of matching records forms our goal and the proposed system results in a software product with high level filter criteria like age, gender and location, in-depth analysis of disease statistics for reference sites with factor based temporal matching algorithm for matching the relevant contents from the historical databases.

Advantange of Proposed System: It provides a way for appropriate treatment to that patient along with the ease of maintenance of data records in terms of dynamic changes. Sensitive data identification and constrain satisfying records are ranked high. An efficient knowledge management and sharing is achieved.

IV. IMPLEMENTATION AND METHODOLOGY

Proposed System Design involves identification of classes their relationship as well as their collaboration. The information management structure in the medical database used in appraised query processing has been modelled. The posted query by the expert will first be verified and automatic error correction is performed to enter into correct consolidated record links then the topic based suggestion such as echo, cancer, liver disorder etc based on the temporal events updated is provided. Further the query is expanded and rewritten by the system which uses inverted index to capture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

the upshots. The top k answers are retrieved based on the citation edges and query keyword, and visualized to the expert. The existing technologies discussed in the related works need to be further concentrated with the temporal high dimensional database events. The clustering and classification technique is the perfect management strategy to handle the upcoming big data that are topically related. Though the recently available trends mines the query results based on the content, it is indeed necessary to step into the smart mining with improved vision on intelligent information retrieval from the engineered database. Thus the points to concentrate on includes the following

- Relevant results within accurate user's contour.
- Speedup of appraisal in the big data analyzation
- Clouding the structured patented databases
- Finding positive partition classes for the given query.
- Data visualization in ranked order with pattern matching made more efficient.

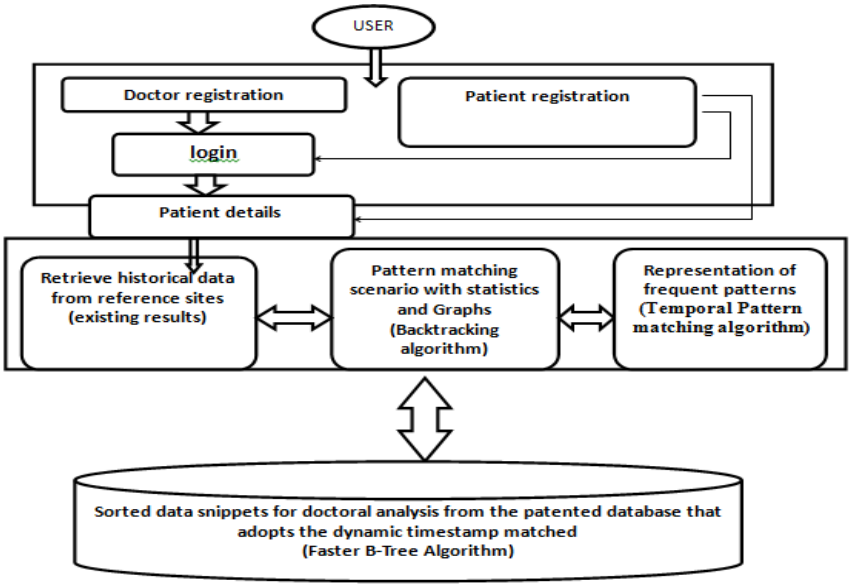


Fig 1: BTF algorithm implemented in medical database

Our paper concentrates in these points. The qualified technique Automatic Error correction now not mostly available in medical database is quintessential for accurate result retrieval. The Backtracking algorithm is used for a quick test whether the partial solution can be a valid solution or not. The recursive depth first search strategy in our tree helps in pruning the irrelevant search space and in determining whether the branch is valid or not thus on going with Nano unit time operations. Topic based query suggestion can be achieved by our proposed temporal pattern search algorithm that arrays the events of same type with the time stamp value, this is useful to skip unnecessary histories and encapsulate the users recall view of records. The recently enhanced Faster B-Tree algorithm is used to make a sorted data storage and perform the updating and retrieval of the temporal events in logarithmic time, thus helps in query expansion and suggestion that makes a friendly interface for the user.

4.1 MODULE DESCRIPTION:

Implementation is the stage of the project when the theoretical design is turned out into a working system. Our modules are

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

4.1.1 Input Data Sets Module [Fig 4.1]**:** Information are captured to find out the dataset and an end user application is created in order to access accordingly. The spatial information obtained like area of living, helps in identification of epidemic disease and acts as a filtering criteria along with age, ender etc.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

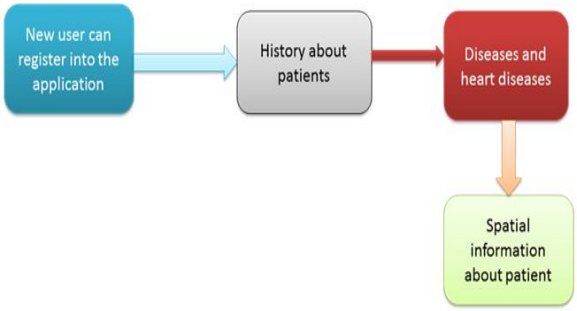


Fig 4.1 Input Data Sets Module

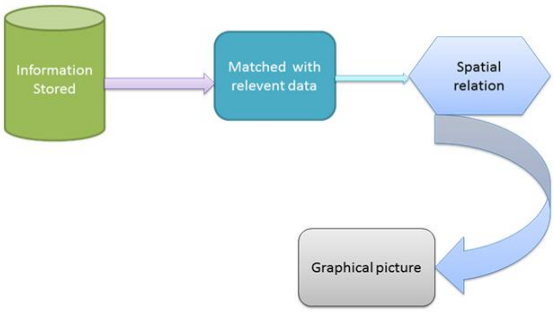


Fig 4.2.Frequent pattern Data Sets Identification Module

4.1.2 Frequent pattern Data Sets Identification Module[Fig 4.2.**]:** In this module, the high severity data accessed by the requestor from the data owner through cloud are analyzed. The temporal pattern matching algorithm helps in ordering the time stamps, So that the recent updates can be identified instead of searching unnecessary histories. The retrieved result is structured into a graphical structure for faster analysis.

4.1.3. Entity Relational Data Sets Identification Module[Fig 4.3]**:** In this module, the datasets in the table that refers other tables are identified in order find out its severity as well as the leakage that may acquire of it. Table join operations can be performed if needed, that co-ordinates reference available globally.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

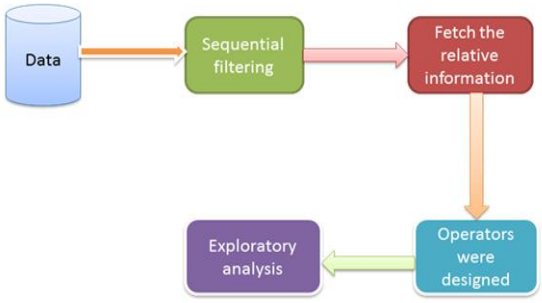


Fig 4.3.Entity Relational Data Sets Identification Module

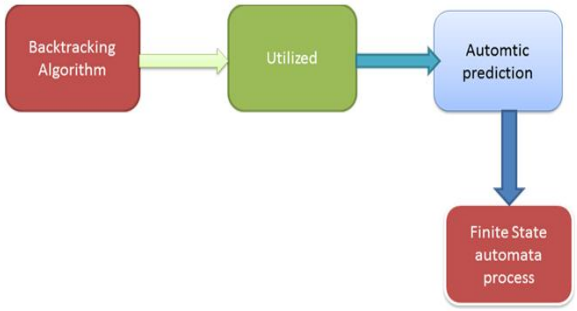


Fig 4.4 Anonymizing Data Sets Module

4.1.4. Anonymizing Data Sets Module[Fig 4.3] **:** In this module, datasets that are unpredictable are added in order to make the system more secure. Backtracking helps in validation of the data branch in a sorted faster b-tree. Finite state Automata constructed finally helps in determined result records or experts opinion sharing

4.1.5. Scheduling Frequent pattern Module: In this module, the dynamic changes that may acquire are analyzed for each and every time stamp and checked against its severity. So that unnecessary search spaces can be pruned so that a low overhead can be achieved in time and space utilization.

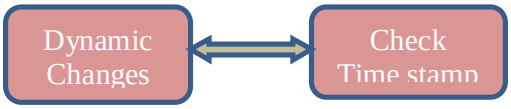


Fig 4.5 Scheduling Frequent pattern Module

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

4.2 ALGORITHM:

Backtracking from a given node n:

```
boolean solve(Node n) {
if n is a leaf node {
if the leaf is a goal node, return true
else return false
} else {
for each child c of n {
if solve(c) succeeds, return true
}
 return false
 }
 }
 boolean solve(Node n) {
 if n is a leaf node {
 if the leaf is a goal node {
 print n
 return true
 }
 else return false
 } else {

 for each child c of n {
 if solve(c) succeeds {
 print n
 return true
} }
 return false}}
```

Faster B-Tree algorithm

```
Faster B-Tree Create(T)
X←Allocate-node()
Leaf[x]←True
N[x]<__0
Disk-Write(x)
Root[T]←x
r←root[T]
if n[r]=2t-1
then s←Allocate-Node()
root[T]←s
leaf[s]←FALSE
n[s]←0
C1[s]←r
B-Tree-Split-child(s,l,r)
B-Tree Insert-Non full(s.k)
else B-tree Insert-non Full(r,k)
```

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

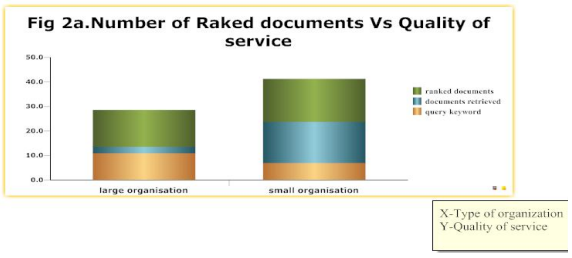
Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

B-tree Search (x, k)
i←1
while i<=n[x] and k>key i[x]
doi←i+1
if i<=n[x] and k=keyi[x]
then return(x,i)
if leaf[x]
then return NIL
else Disk-Read(ci[x])
return B-Tree-Searc(ci[x],k).

V. EXPERIMENTAL RESULTS

This section investigates the efficiency of the proposed algorithm against the efficiency of two other algorithms in terms of accuracy and relevancy of results. The dataset we have consists of 135 time stamped record each of which represents the historical reference records of certain disease in a certain health unit (clinic) for a period of consecutive months from 1997 to 2012. The prediction process is applied to each of the last 12 points in every time series. The reason of choosing is that I relevancy is ,because it is dimensionless and can represent a combination of opinion of many different experts having different specialisation Also it can compare the performance of different forecasting methods. .The performance of our proposed algorithm is compared to other popular algorithms that are: Prior art relevancy and as you type approach if given a time series, then the values of its exponential smoothing are Best results for the exponential Time stamp. It if found that the top k answers retrieved were ranked high in other approaches also, but with time stamp ordering the accuracy and time consumed found to be reduced by P(0.8%) which stages the proposed system.We have implemented our proposed techniques. We compared with the prior art technique and SVM ranking in the retrieval of information in the simulated medical database. The obtained results were satisfying in the advanced information retrieval in respect of relevancy, quality and ranking within the bounded time factor.

5.1 Relevancy of documents: We evaluate the effect on k (the number of selected partitions). We partitioned the records into topics and classes and evaluated the effectiveness and quality by varying the value of k. To evaluate the result quality, we used the milestone technique p@k, where the precision is the ratio of the number of retrieved relevant results to the number of retrieved results, and p@k is the precision of the top-k results.



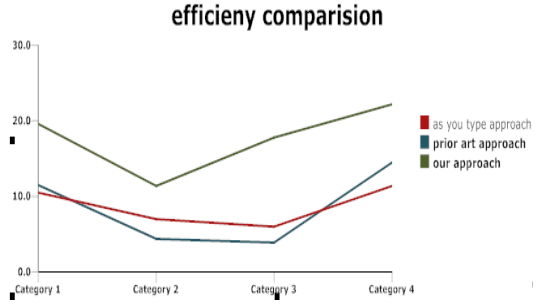


Fig2b.Comparision with different approaches (with p@k)

Fig. 2a shows the results. The main reason is as follows: First, the more records used to answers a query, the more relevant answers, and thus the higher precision. Second, as each query usually belongs to limited number of topics, only several partitions are relevant to the query. Thus, the precision is stable when 'k' is large enough. For example, where k '>' 10, they achieved nearly the same precision. Then, we evaluated the efficiency. Fig. 2b shows the results. We see that with the increase of k, the elapsed time increased. This is because the more the categories used to

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

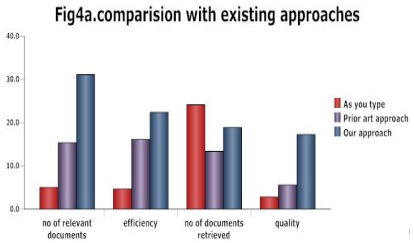
answer a query the efficiency increases.

5.2 Precision: In this section, we compare the result quality. Table 1 shows the experimental results. We can see that our three techniques can improve the result quality. For example, for p@50, error correction improved to 0.83, query suggestion increases the precision to 0.82, and query expansion can improve the precision to 0.88. More importantly, our method by combing the three methods can improve the precision to 0.88, and the improvement ratio is about 0.84, achieved by query rewriting. The main reason is as follows: First, the automatic error correction can provide users accurate keywords based on users inputs. Second, the query expansion can suggest relevant keywords. Third, topic-based query suggestion can provide users topic- relevant keywords.

5.3 Efficiency

Table 1: Quality comparison

Approaches compared	p@k	Query Suggestion	Query Expansion	Automatic Error Correction	Overall efficiency	Ranking of techniques
As you Type	10	0.77	0.66	0.62	0.6833	3
Prior art search	15	0.79	0.70	0.78	0.7566	2
BTF approach (proposed approach)	25	0.82	0.88	0.83	0.8433	1



In this section, we compare the efficiency. We partitioned the data into 24 partitions. We used three computers to manage the data and each node managed three partitions. For each partition, we built the corresponding inverted indexes. We first compared the two methods by varying the number of keywords. We can see that for different numbers of keywords, our method always outperformed the existing method SVMPR, with the speedup ratio about 8. This reflects that our method achieved high efficiency since we employ an effective partition-based method. We then evaluated the two methods by varying the number of returned answers. We can see that for different values, our method always out performed SVMPR. This is because, we partition the data into eight partitions and each partition was inversely indexed and searched by different cores. More importantly, our partition-based method can prune the search space and thus can improve the performance significantly.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

5.4 Scalability: We also evaluate the scalability of our method. Fig. 4a shows the experimental results by varying the number of patents. Fig. 4a shows the scalability on quality and we can see that with the increase of the number of records, the precision of query suggestion and query expansion increased slightly. This is because we can utilize more data to select the topic of each record and find more relevant keyword pairs. The more data used the higher precision of the topic model. For query correction, the precision nearly kept the same as we only used the trie structure to correct the keywords but the quality increased as we proposed the prefix based depth first search. On the other hand, our method scaled very well.

VI. CONCLUSION AND FUTURE WORK

The user friendly expert analysis query processing techniques discussed so far had made milestones in the domain of information retrieval and management. But our proposed suggestion may make a perfect benchmark. The patented data from the sources are first clustered into topics and classes, when given a query the highly coherent cluster partitions are recovered. The upshots in each coherent cluster are combined generating top K relevant answers for the examiners from the database. With the simulated results the on- going implementation of our method had shown high efficiency and increased quality of the partitioned based search strategy The techniques streamlined namely query suggestions, Topic relevancy, query augmentation along with privacy preservation when handshake with existing prior art relevancy may prove marvellous enhancements in knowledge engineering.

FUTURE ENHANCEMENT: In future preservation of privacy and cost optimization of datasets that are accessible through cloud by considering many factors such time span of usage, availability of servers are to be concentrated to further increase efficiency.

ACKNOWLEDGEMENT

It is my pleasure to acknowledge Mr.J.Venkatesan Prabhu, Head, Kaashiv- InfoTech, Chennai for his support in implementation part and for his guidance throughout this course of work.I express my hearty tanks to my parents for the constant guidance and encouragement.

REFERENCES

[1] A User-Friendly Patent Search Paradigm Yang Cao, Ju Fan, and Guoliang Li IEEE Transaction on knowledge and data engineering, vol. 25, no. 6, 2013.
[2] J. Agrawal, Y. Diao, D. Gyllstrom, and N. Immerman, "Efficient Pattern Matching over Event Streams," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 147-160, 2011.
[3] A. Demers, J. Gehrke, M. Hong, M. Riedewald, and W. White, "Towards Expressive Publish/Subscribe Systems," Proc. 10th Int'l Conf. Extending Database Technology (EDBT), pp. 627-644, 2011.
[4] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data,"Proc. IEEE INFOCOM '11, pp. 829-837, 2011
[5] H. Lin and W. Tzeng" A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding for data mining,", IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 6, pp. 995-1003, June 2012
[6] D. Zissis and D. Lekkas, "Addressing Cloud Computing Security Issues" Future generation Computer Systems, vol. 28, no. 3, pp. 583-592, 2011
[7] R.S. Boyer and J.S. Moore, "A Fast String-Searching Algorithm," Comm. ACM, vol. 20, no. 10, pp. 762-772, 1977. [3] R. Cox, 2010.
[8] X. Xue and W.B Croft, "Automatic Query Generation for Patent Search," Proc. ACM Conf. Information and Knowledge Management (CIKM), pp.2037-2040, 2009.
[9] L.S. Larkey, "A Patent Search and Classification System," Proc. Fourth ACM Conf.Digital Libraries, pp. 179-187,2011.
[10] J. Fan, H. Wu, G. Li, and L. Zhou, "Suggesting Topic-Based Query Terms as You Type," Proc. Int'l Asia Pacific Web Conf. (APWEB), pp. 61-67, 2010.
[11] G. Li, J. Feng, and C. Li, "Supporting Search-As-You-Type Using SQL in Databases," IEEE Trans. Knowledge and Data Eng., vol. 25, no. 2, pp. 461-475, Feb. 2013.
[12] S. Bashir and A. Rauber, "Improving Retrievability of Patents in Prior-Art Search," Proc. European Conf. Information Retrieval (ECIR), pp. 457-470, 2010.
[13] Y. Guo and C.P. Gomes, "Ranking Structured Documents: A Large Margin Based Approach for Patent Prior Art Search," Proc. Int'l Joint Conf. Artificial Intelligence (IJCAI), pp. 1058-1064, 2011.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Author details:

Ms.T.Yogameera completed her B.E Computer science and Engineering in 2012 with first class from R.V.S College of Engineering and Technology, Dindigul, Tamilnadu and pursuing M.E. Computer Science and Engineering in P.S.N.A College of Engineering and Technology, Dindigul, Tamilnadu. She has secured best academician award 2 times during her school days in Shri Maharishi Vidya Mandir(CBSE), Dindigul, Tamilnadu, she has grabbed the first best project award in the internationalscience and technology contest She has participated in two international symposiums presenting her own paper and won second prize in both, she has attended 3 national level workshops and 1 international workshop, 2 national seminars, and 15 days internship program on mobile operating system, has published one international journal. Her domains of interest are data mining, information retrieval, query processing, database system, currently concentrating on information management in medical systems.

Dr.D.Shanthi Saravanan received her B.E Computer Science and Engineering in 1992 from Thiagaraja College of Engineering, Madurai, Tamilnadu and M.E Computer Science and Engineering from Manonmaniam Sundaranar University, Tamilnadu and PhD in Soft Computing from Brila Institute of Technology, Ranchi. She is currently working as a Professor in Department of Computer Science and Engineering, PSNA College of Engineering and Technology. She has more than 21 years of Teaching and Programming Experience. She is a member of various professional societies like IEEE, CSTA, IAENG and IACSIT. She has published more than 25 papers in international journals and conferences and also 5 books in computing and applications. She is a reviewer of various Her research interest includes Genetic Algorithm, Neural Networks, and Technology, Ranchi. She is currently working as a Professor in Department of Computer Science and Engineering, PSNA College of Engineering and Technology. She has more than 21 years of Teaching and Programming Experience. She is a member of various professional societies like IEEE, CSTA, IAENG and IACSIT. She has published more than 25 papers in international journals and conferences and also 5 books in computing and applications. She is a reviewer of various Her research interest includes Genetic Algorithm, Neural Networks, and Intelligent Systems, Image processing, Embedded System, Machine Learning and Green Computing