



## International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 4, Issue 12, December 2016

# Efficient Publish/Subscribe System using Filtering Algorithm for Location Based Service

Kiran Gaikwad, Rajesh Auti

M.E.Student, Dept. of Computer, Everest College of Engineering, BAM University, Aurangabad,  
Maharashtra, India

H.O.D., Dept. of Computer, Everest College of Engineering, BAM University, Maharashtra, India

**ABSTRACT:** The services which are location based are used in many applications. In these systems any user request for particular service through query to a server and server gives feedback according to location. For providing this facility a push model, or server initiated model come into picture. For using this facility user has to enter their interest for finding their need. Because of this model user will get related information which is going to be need in proposed method. We create index based on R-tree by collecting the tokens into R-tree nodes. They Propose an R-tree based index by integrating textual descriptions into R-tree nodes and develop efficient filtering algorithms and effective pruning techniques to achieve high performance. Using this model we can do sorting or decrease some unnecessary information by filtering algorithm which helps to increase efficiency of system but also reduces cost of such publisher/subscriber application. This method can work on both ranking queries and conjunctive queries. The results which are experimented by us can filter 500 messages/sec.

**KEYWORDS:** LBS, Spatial-Context, MBR Filter, Token Filter, Ranking Query, Rt-Tree.

### I. INTRODUCTION

Services which are location based are very useful for academic and industrial purpose. There is so many other location base services exist and they are accepted because they offer location aware Service. In this type of services user enter their choices and will get response to user with location aware answer. Many LBS services such as Foursquare and Google Maps have been broadly recognized because they can convey users with location-aware actions. The preceding LBS systems use a pull model or user-initiated model, where a user arrive a query to a server which answers with location aware outcome. For example, if a mobile user wants to search writer with their city, then they have a query “writer name” to an LBS system, which proceeds outcome based on user’s location and keywords. For providing this facility to user, one push model is designed. Using this model user has to register spatio-text for finding their interest and the publisher send text messages. This model gives a new experience to user for finding their location using queries.

One big challenge in a publish/subscribe system is to achieve high performance. A publish/subscribe system should support tens of millions of subscribers and deliver messages to relevant subscribers in milliseconds. Since messages and subscriptions contain both location information and textual description, it is rather costly to deliver messages to relevant subscribers. To address the challenge, we propose a token-based R-tree index structure (called Rt-tree) by integrating each R-tree node with a set of tokens selected from subscriptions. Using the Rt-tree, we develop a filter-and-verification framework to efficiently deliver a message. To reduce the number of tokens associated with Rt-tree nodes, we select some high-quality representative tokens from subscriptions and associate them with Rt-tree nodes. This technique not only reduces index sizes but also improves the performance. Experiments on large, real data sets show that our method achieves high performance.

This technique not only reduces index sizes but also improves the performance. Experiments on large, real data sets show that the proposed method achieves high performance. The following contributions are being made: (1) A



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 4, Issue 12, December 2016

new computing model is introduced and formalize the location-aware publish/subscribe problem. (2) A novel index structure, the  $R^1$ -tree, is proposed by integrating high-quality representative tokens selected from subscriptions into the R-tree nodes. Our method can support both conjunctive queries and ranking queries.(3) Using the proposed indexes, efficient filtering algorithms and effective pruning techniques to improve the performance are developed. (4) The work also presents how to support dynamic updates efficiently.

## II. RELATED WORK

**1. M. K. AGUILERA, R. E. STROM, D. C. STURMAN, M. ASTLEY, AND T. D. CHANDRA, IN PROC. 18TH ANNU. ACM SYMP. PRINCIPLES DISTRIB. COMPUT., 1999, PP. 53–61**

How to efficiently match high volumes of events against large numbers of subscriptions is a key issue for large-scale content-based publish/subscribe systems. In this paper we present an efficient and applied matching algorithm that uses multi-dimensional indexing mechanism to speed up constraints query and exploits the covering relations between constraints to reduce unnecessary matching. Experiments show that our algorithm is significantly more efficient and scalable than other common used matching algorithms.

**2. Efficient filtering of XML documents for selective dissemination of information M. Altinel and M. J. Franklin, in Proc. 26th Int. Conf. Very Large Data Bases, 2000, pp. 53–64.**

Information Dissemination applications are gaining increasing popularity due to dramatic improvements in communications bandwidth and ubiquity. The sheer volume of data available necessitates the use of selective approaches to dissemination in order to avoid overwhelming users with unnecessary information. Existing mechanisms for selective dissemination typically rely on simple keyword matching or "bag of words" information retrieval techniques. The advent of XML as a standard for information exchange and the development of query languages for XML data enable the development of more sophisticated filtering mechanisms that take structure information into account. We have developed several index organizations and search algorithms for performing efficient filtering of XML documents for large-scale information dissemination systems. In this paper we describe these techniques and examine their performance across a range of document, workload, and scale scenarios

**3. Models and issues in data stream systems B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom, in Proc. 21st ACM SIGMOD-SIGACT-SIGART Symp. Principles Database Syst., 2002, pp. 1–16.**

In this overview paper we motivate the need for and research issues arising from a new model of data processing. In this model, data does not take the form of persistent relations, but rather arrives in multiple, continuous, rapid, time-varying data streams. In addition to reviewing past work relevant to data stream systems and current projects in the area, the paper explores topics in stream query languages, new requirements and challenges in query processing, and algorithmic issues.

**4. Retrieving top-k prestige based relevant spatial web objects X. Cao, G. Cong, and C. S. Jensen, Proc. VLDB Endowment, vol. 3, no. 1, pp. 373–384, 2010.**

The location-aware keyword query returns ranked objects that are near a query location and that have textual descriptions that match query keywords. This query occurs inherently in many types of mobile and traditional web services and applications, e.g., Yellow Pages and Maps services. Previous work considers the potential results of such a query as being independent when ranking them. However, a relevant result object with nearby objects that are also relevant to the query is likely to be preferable over a relevant object without relevant nearby objects. The paper proposes the concept of prestige-based relevance to capture both the textual relevance of an object to a query and the effects of nearby objects. Based on this, a new type of query, the location-aware top-k prestige-based text retrieval (lkpt) query, is proposed that retrieves the top-k spatial web objects ranked according to both prestige-based relevance and location proximity. We propose two algorithms that compute lkpt queries.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

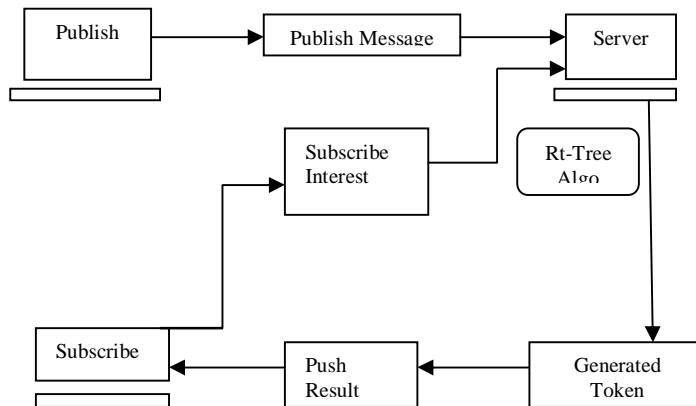
Vol. 4, Issue 12, December 2016

## III. PROPOSED ALGORITHM

### Existing Algorithm

There are many real-world applications using location aware publish/subscribe services. The first one is Group on, in which customers register their interests with locations and keywords. For each Group on message, the system provider sends the message to the customers who may be potentially interested in the message by evaluating the spatial proximity and textual relevancy between subscriptions and the message. The second one is location-aware AdSense, which extends traditional AdSense to support location-aware services. The advertisers register their location-based advertisements in the system. The system pushes relevant advertisements to mobile users based on their locations and contents they are browsing. The third one is tweet delivery. To receive feedback of their products in a specific area from Twitter, market analysts register their interests. For each tweet, the system pushes the tweet to relevant analysts whose spatio-textual subscriptions match the tweet.

To address this problem, using an Rt-tree node may have large numbers of leaf descendants and it is expensive to check whether a node is a pivotal node. Based on this observation, we propose a filter and- verification framework.



System Architecture.

### Description of the Proposed Algorithm:

YU ET AL.: EFFICIENT FILTERING ALGORITHMS FOR LOCATION-AWARE PUBLISH/SUBSCRIBE

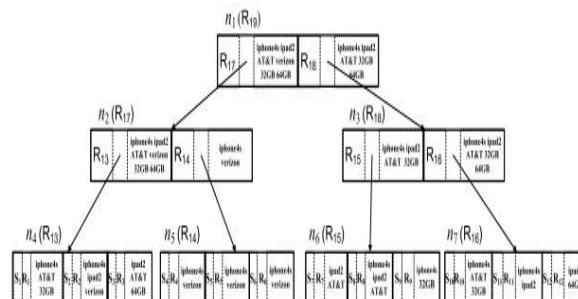


Fig. 2. R-tree index for subscriptions in Fig. 1.

### R-tree Indexing

As the standard R-tree has no textual pruning power, a token-based R-tree, called Rt-tree, by mounting tokens of subscriptions into R-tree nodes. Rt-tree is a balanced search tree. Each leaf node contains between  $b$  and  $B$  data entries, where each entry is a subscription. Each internal node is between  $b$  and  $B$  node entries. Each entry is a triple  $\langle h, Child, MBR, TokenSet_i \rangle$ , where  $Child$  is a pointer to its child node,  $MBR$  is the minimum bounding rectangle of all



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 4, Issue 12, December 2016

admissions within this child, and TokenSet is a set of tokens selected from subscriptions. The outputs for subscriber are processing using Rt- tree indexing and then filtered for further output processing. To address this problem, using an Rt-tree node may have large numbers of leaf descendants and it is expensive to check whether a node is a pivotal node. Based on this observation, we propose a filter and- verification framework.

## R<sup>t</sup>- Tree Indexing

Input: S, A subscription set, message m

Output: R, Outcomes of m

Step 1: Publisher publishes message m

Step 2: Build R<sup>t</sup>- tree index by collecting all message m from 'n' publishers {p1, p2,...pn }

Step 3: Initialize a HashMap M

Step 4: return R<sup>t</sup>-tree++

## R<sup>t</sup>- Tree Pruning

Input: r, An R<sup>t</sup>-tree node, 'm' a message, 'R' outcome of m, HashMap M

Output: R, Outcomes of m

Step 1: Visit flag = false;

Step 2: for each entry n in node r do

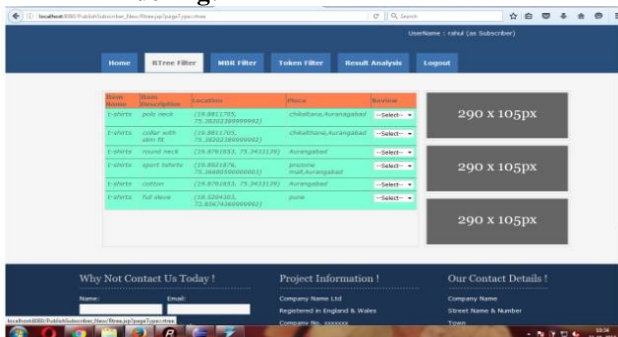
Step 3: Check location of node and filter message in location R

Step 4: Check curiosity of node and filter message of curiosity m

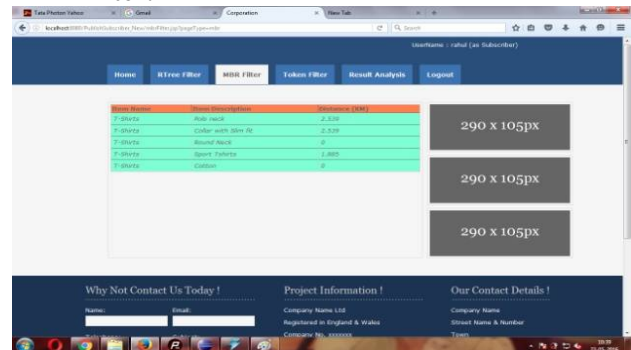
Step 5: prune outcome R and m

Step 6: Outcome of Rt-tree prune to node.

## R-TREE Indexing:



## MBR Filter:



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

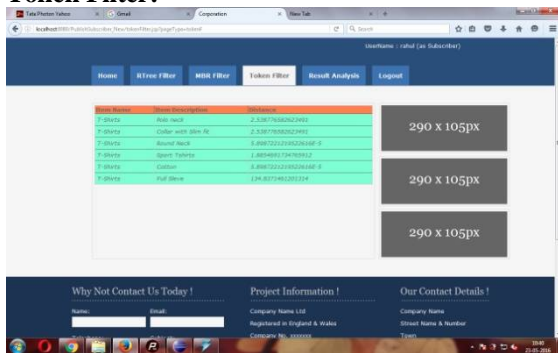
Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 4, Issue 12, December 2016

As the standard R-tree has no textual pruning power, a token-based R-tree, called Rt-tree, by mounting tokens of subscriptions into R-tree nodes. Rt-tree is a balanced search tree. Each leaf node contains between b and B data entries, where each entry is a subscription. Each internal node is between b and B node entries. Each entry is a triple h Child, MBR, TokenSet, where Child is a pointer to its child node, MBR is the minimum bounding rectangle of all admissions within this child, and TokenSet is a set of tokens selected from subscriptions. The outputs for subscriber are processing using Rt- tree indexing and then filtered for further output processing.

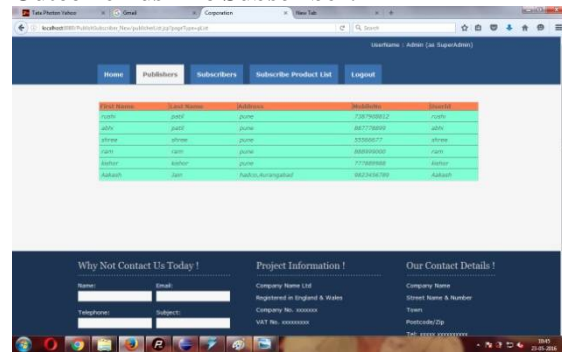
Least bound rectangle filter for checks the location of the subscriber. This model filters the outcomes Rt-tree index by examination the user’s location and publisher location. The location based outcome set conveys more location specific outcome, which does not consider the subscriber curiosities. These outcomes are used for further processing to get subscriber outcome.

## Token Filter:



It is used to checks for the textual constraint. Subscriber’s curiosity is considered for token filter. This model filters the outcomes Rt-tree index by checking the user’s location and publisher location. The location based outcome set carry more curiosity specific outcome, which does not consider the location of subscriber. These outcomes are used for additional processing to get subscriber’s location based outcome.

## Outcome Push To Subscriber:



In the push model, subscribers enter spatio-textual subscriptions to fastening their curiosities, and publishers send spatio-textual messages to the respective subscribe. The outcomes from the upstairs two methods, MBR filter and token filter, spatio- textual outcomes are filtered and send to subscriber. The server impetuses the outcome to subscriber instead of responding every time when subscriber queries.

## IV. SIMULATION RESULTS

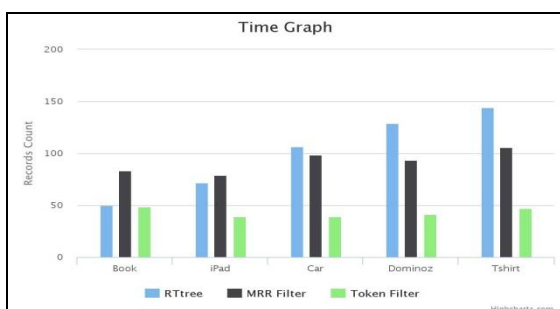


Fig.1. Time Graph

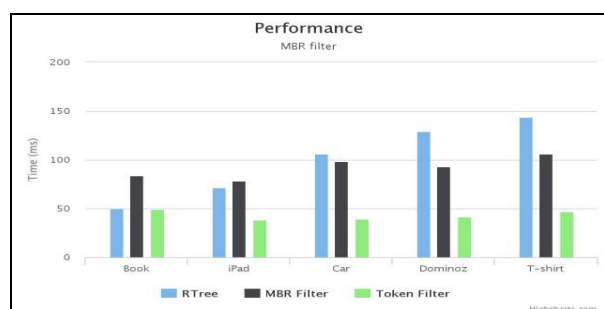


Fig. 2. Performance Graph



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 4, Issue 12, December 2016

Evaluation of the Frequency of availability against the Subscriber interest terms The Rt-tree is evaluated with different token sets, i.e., Rt-tree with token sets, Rt-tree with representative tokens, Rt-tree with multiple representative tokens.

## V. CONCLUSION AND FUTURE WORK

In this paper, we study the location-aware publish/subscribe problem. We propose an effective index structure Rt-tree by integrating textual description into R-tree nodes. We develop a filter-and-verification framework and devise efficient Filtering algorithms. We propose reducing the number of tokens in each node which not only reduces index sizes but improves performance. We devise an efficient algorithm to directly find answers without the verification step. We extend our algorithms to support both conjunctive queries and ranking queries. We discuss how to support ranking semantics. Experimental results on real datasets show our method achieves high performance and good scalability.

The location aware publish/subscribe are useful at Emergency and Safety, Communities and Entertainment, Information and Navigation and Tracking and Monitoring or private facility (eg.Buddy List) where constant monitoring and control are needed.

## REFERENCES

1. M. K. Aguilera, R. E. Strom, D. C. Sturman, M. Astley, and T. D.Chandra, "Matching events in a content-based subscription system," in Proc. 18th Annu. ACM Symp. Principles Distrib. Comput.,
2. M. Altinel and M. J. Franklin, "Efficient filtering of XML documents for selective dissemination of information," in Proc. 26th Int. Conf. Very Large Data Bases, 2000, pp. 53–64.
3. B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom, "Models and issues in data stream systems," in Proc. 21st ACM SIGMOD-SIGACT-SIGART Symp. Principles Database Syst., 2002, pp. 1–16.
4. X. Cao, G. Cong, and C. S. Jensen, "Retrieving top-k prestige based relevant spatial web objects," Proc. VLDB Endowment, vol. 3, no. 1, pp. 373–384, 2010.
5. X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi, "Collective spatial keyword querying," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 373–384.
6. X. Chen, Y. Chen, and F. Rao, "An efficient spatial publish/subscribe system for intelligent location-based services," in Proc.2nd Int. Workshop Distrib. Event-Based Syst., 2003, pp.
7. Y.-Y. Chen, T. Suel, and A. Markowetz, "Efficient query processing in geographic web search engines," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2006, pp. 277–288.
8. G. Cong, C. S. Jensen, and D. Wu, "Efficient retrieval of the top-k most relevant spatial web objects," Proc. VLDB, vol. 2, no. 1, pp. 337–348, 2009.
9. P. Costa and G. P. Picco, "Semi-probabilistic content-based publish-subscribe," in Proc. 25th IEEE Int. Conf. Distrib. Comput.Syst., 2005, pp. 575–585.
10. G. Cugola and J. E. M. de Cote, "On introducing location awareness in publish-subscribe middleware," in Proc. IEEE Int. Conf.Distrib. Comput. Syst. Workshops, 2005, pp. 377–382.
11. Y. Diao and M. J. Franklin, "Query processing for high-volume XML message brokering," in Proc. 29th Int. Conf. Very Large DataBase, 2003, pp. 261–272.
12. P. T. Eugster, B. Garbinato, and A. Holzer, "Location-based publish/subscribe," in Proc. 4th IEEE Int. Symp. Netw. Comput. Appl., 2005, pp. 279–282.
13. P. T. Eugster, B. Garbinato, and A. Holzer, "Pervaho: A specialized middleware for mobile context-aware applications," Electron. Commerce Res., vol. 9, no. 4, pp. 245–268, 2009.
14. F. Fabret, H.-A. Jacobsen, F. Llirbat, J. Pereira, K. A. Ross, and D. Shasha, "Filtering algorithms and implementation for very fast publish/subscribe," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2001, pp. 115–126.
15. J. Fan, G. Li, L. Zhou, S. Chen, and J. Hu, "Seal: Spatio-textual similarity search," Proc. VLDB Endowment, vol. 5, no. 9, pp. 824–835, 2012.
16. I. D. Felipe, V. Hristidis, and N. Rische, "Keyword search on spatial databases," in Proc. Int. Conf. Data Eng., 2008, pp. 656–665.
17. L. Fiege, F. C. G.ärtner, O. Kasten, and A. Zeidler, "Supporting mobility in content-based publish/subscribe middleware," in Proc. USENIX Int. Conf. Middleware, 2003, pp. 103–122.
18. P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," Commun. ACM, vol. 35, no. 12, pp. 51–60, 1992.
19. B. Garbinato, A. Holzer, and F. Vessaz, "Context-aware broadcasting approaches in mobile ad hoc networks," Comput. Netw.,vol. 54, no. 7, pp. 1210–1228, 2010.
20. L. Golab and M. T. Özsu, "Issues in data stream management,"SIGMOD Rec., vol. 32, no. 2, pp. 5–14, 2003.
21. R. Hariharan, B. Hore, C. Li, and S. Mehrotra, "Processing spatialkeyword (SK) queries in geographic information retrieval (GIR) systems," in Proc. 19th Int. Conf. Sci. Stat. Database Manage., 2007,p. 16.
22. K. W.-T. Leung, D. L. Lee, and W.-C. Lee, "Personalized web search with location preferences," in Proc. Int. Conf. Data Eng., 2010, pp. 701–712.
23. G. Li, J. Feng, and J. Xu, "Desks: Direction-aware spatial keyword search," in Proc. Int. Conf. Data Eng., 2012, pp. 474–485.



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

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 4, Issue 12, December 2016

24. G. Li, Y. Wang, T. Wang, and J. Feng, "Location-aware publish/subscribe," in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 802–810.
25. J. Lu, Y. Lu, and G. Cong, "Reverse spatial and textual k nearest neighbour search," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 349–360.
26. S. B. Roy and K. Chakrabarti, "Location-aware type ahead search on spatial databases: Semantics and efficiency," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 361–372.
27. H. Samet, Foundations of Multidimensional and Metric Data Structure. San Mateo, CA, USA: Morgan Kaufmann, 2006.
28. M. Sharifzadeh and C. Shahabi, "Approximate voronoi cell computation on spatial data streams," VLDB J., vol. 18, no. 1, pp. 57– 75, 2009.
29. D. Wu, M. L. Yiu, C. S. Jensen, and G. Cong, "Efficient continuously moving top-k spatial keyword query processing," in Proc. Int. Conf. Data Eng., 2011, pp. 541–552.
30. T. W. Yan and H. Garcia-Molina, "Index structures for information filtering under the vector space model," in Proc. Int. Conf. Data Eng., 1994, pp. 337–347.
31. T. W. Yan and H. Garcia-Molina, "Index structures for selective dissemination of information under the boolean model," ACM Trans. Database Syst., vol. 19, no. 2, pp. 332–364, 1994.
32. B. Yao, F. Li, M. Hadjieleftheriou, and K. Hou, "Approximate string search in spatial databases," in Proc. Int. Conf. Data Eng., 2010, pp. 545–556.
33. D. Zhang, Y. M. Chee, A. Mondal, A. K. H. Tung, and M. Kitsuregawa, "Keyword search in spatial databases: Towards searching by document," in Proc. Int. Conf. Data Eng., 2009, pp. 688–699.
34. D. Zhang, B. C. Ooi, and A. K. H. Tung, "Locating mapped resources in web 2.0," in Proc. Int. Conf. Data Eng., 2010, pp. 521–532.

## BIOGRAPHY

**Kiran S. Gaikwad** is M.E. student in computer department, Everest College of Engineering, BAM University. She received B.Tech of Computer Science degree in 2006 from BAMU, Aurangabad, MS, India.