INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

EFFECTIVE FILTERING BASED ON LOCATION AND KEYWORD FOR PUBLISHER / SUBSCRIBER

SHWETA KATHALKAR, PROF. R. R. SHELKE

Computer Science and Engineering, H.V.P.M COLLEGE Of Engineering And Tech., Amravati, India

Accepted Date: 15/03/2016; Published Date: 01/05/2016

Abstract: Location-based services have been mostly used in many systems. Preceding systems uses a pull model or user-initiated model, where a user arrival a query to a server which gives response with location-aware answers. To offer outcomes to users with fast responses, a push model or server-initiated model is flattering an important computing model in the next generation location-based services. In the push model, subscribers arrive spatio-textual subscriptions to fastening their curiosities, and publishers send spatio- textual messages. It is used for a high-performance location-aware publish/subscribe system to send publishers ‟ messages to valid subscribers. In this paper, we find the exploration happenstances that start in manipulative a location-aware publish/subscribe system. We recommend an R-tree based index by merging textual descriptions into R- tree nodes. We design efficient filtering algorithms and effective pruning techniques to accomplish high performance. This method can support likewise conjunctive queries and ranking queries.

Keywords: LBS, Spatial-Context, MBR Filter, Token Filter, Ranking Query, R t-Tree

Corresponding Author: MS. APARNA S. KALASKAR

Access Online On:

www.ijpret.com

How to Cite This Article:

Shweta Kathalkar, IJPRET, 2016; Volume 4(9): 1642-1649

PAPER-QR CODE

INTRODUCTION

Location based services have involved important with more curiosity from correspondingly industrial and academic groups. 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 pursuit writer with their city, then they have a query "writer name" to an LBS system, which proceeds outcome based on users location and keywords. Most of the existing studies adopt a spatial keyword search based method to help users retrieve location-aware answers [7, 5]. Given a set of objects with spatial information and textual description (e.g., points-of-interest (POIs)) and a user query with location and keywords, spatial keyword search finds top-k relevant objects by considering the distance and textual relevance between the query and objects. For example, if a user wants to find a gas station nearby, she can issue a keyword query "gas station" to a LBS system, which returns the relevant gas stations by considering the user's location and keywords. Traditional spatial keyword search method requires users to type in complete keywords for finding location-aware answers. However for mobile users, typing a complete keyword is tedious and also susceptible to errors. To alleviate this problem, instance search (also known as type-ahead search or search-as-you-type) [1, 17, 16, 3, 15, 14, 18] is proposed to provide users with new search experiences, which returns relevant answers as users type in queries letter by letter. Recently many systems, e.g., Google, have been deployed to support instant search. It is very natural to extend instant search to support spatial keyword search. To this end, in this paper we study the location-aware instant search problem. In our method, as a user types in queries letter by letter, the system returns the location-aware answers on-the-fly, and provides the user with instant replies. Figure 1 provides an example of location-aware instant search over 13 POIs. At every keystroke a user types in our system, the system takes her current input string as a query and returns the relevant location-aware answers instantly. 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. This calls for an efficient filtering technique to support location-aware publish/subscribe services. To address the challenge, we propose a token-based Rtree index structure (called Rt-tree) by integrating each Rtree 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.

We make the following contributions:

(1) We introduce a new computing model and formalize the location-aware publish/subscribe problem.

(2) We propose a novel index structure, the Rt-tree, 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 our proposed indexes, we develop efficient filtering algorithms and effective pruning techniques to improve the performance.

(4) We present how to support dynamic updates efficiently

LITERATURE SURVEY

In the base paper the author presents infected fruit part detection using k-means clustering segmentation technique [1]. K-means is used to decide the natural grouping of pixels presents in the ima Reference entitled “Matching events in a content-based subscription system”, included How to professionally match high volumes of events conflicting to large numbers of subscriptions is a key issue for large-scale content-based publish/subscribe systems. In this paper we extant an efficient and applied matching algorithm that uses multi-dimensional indexing mechanism for rise a speed up constraints query and exploits the covering relations between constraints to minimize the excessive matching. Experiments show that our algorithm is considerably more efficient and scalable than other common used matching algorithms. [1] Reference entitled “Efficient filtering of XML documents for selective dissemination of information”, included Information Propagation applications are acquisition improving popularity due to dramatic improvements in communications bandwidth and ubiquity. The sheer volume of data available necessary to the use of selective nearly to propagation in order to discount emphasis users with redundant information. The preceding mechanisms for selective propagation typically rely on simple keyword matching or "bag of words" information ahead techniques. The advent of XML as a standard for information transformation and the development of query languages for XML data enables the development of more advanced filtering mechanisms that take building

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[2] Reference entitled “Models and issues in data stream systems” included In this overview paper we inspire the need for and research issues get up from a new model of data processing. In this model, data does not take the form of tenacious relations, but rather arrives in multiple, continuous, rapid, time-varying data streams. In addition to revising preceding work related to data stream systems and current projects in the area, the paper explores topics in stream query languages, new requirements and encounters in query processing, and algorithmic issues.[3] Reference entitled “Retrieving top-k prestige based relevant spatial web objects”, include The location-aware keyword query gives ranked objects that are near a query location and that have textual explanations that match query keywords. This query presents intrinsic in many types of mobile and old-style web services and applications, e.g., Yellow Pages and

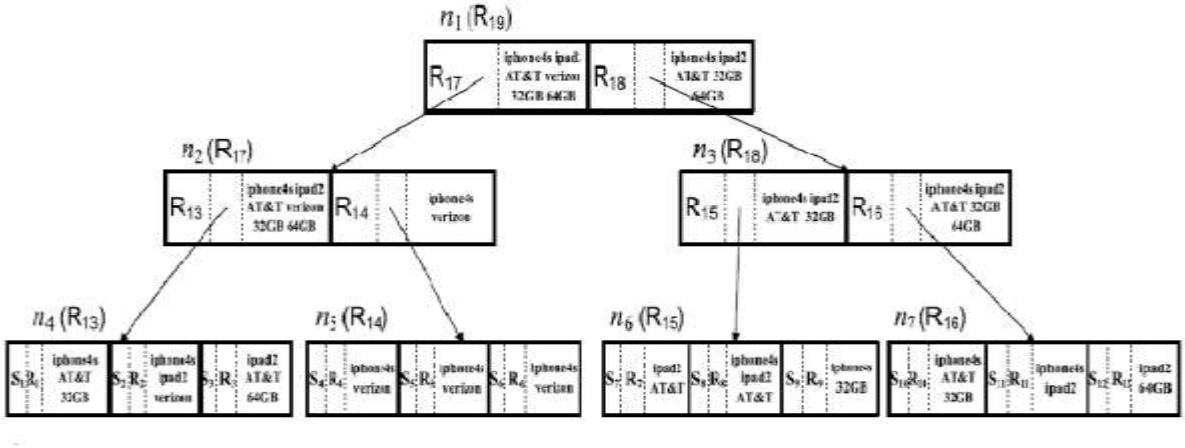

Figure 1. R t-tree index structure.

Maps services. Preceding work think about the potential outcomes of such a query as being sovereign when ranking them. However, a related outcome object with nearby objects that are also related to the query is likely to be desirable over a related object without relevant nearby objects. The paper proposes the concept of prestige-based relevance to fastening both the textual relevance of an object to a query and the effects of adjacent 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 nearness. We propose two algorithms that compute LkPT queries.[4]

PROPOSED SYSTEM

To address the encounters, a token-based R-tree index organized an idea by in escalating each R-tree node with a collection of tokens selected from subscriptions. Using the R t -tree, a filter-and-verification framework is organized for able to sending a message. To decrease the number of tokens connected with R t -tree nodes, select some high-quality illustrative tokens from subscriptions and associate them with R t -tree nodes.

Rt-TREE BASED METHOD

We first propose an Rt Tree algorithm structure and then devise efficient filtering algorithms.

1. R T -T REE A LGORITHM:

R t - Tree Indexing

Input: S, A subscription set, message m

Output: R, Outcomes of m

Step 1: Publisher publishes message m

Step 2: Build R t - tree index by collecting all message m

from „n‟ publishers

{p1, p2,...p n }

Step 3: Initialize a HashMap M

Step 4: return R t -tree ++

R t - Tree Pruning

Input: r, An R t -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 R t -tree prune to node.

L OCATION DETAILS

We consider location specific approach for publish/subscribe system. The region is considered to be rectangle, which we specify a numeric value for example 0-100 meant for one location and 100-200 meant for other location. Given a set of subscriptions S and a message m, a location-aware publish/ subscribe system delivers m to si Є S if si. R∩m. R≠f and si:T⊆ m:T.

R-TREE Indexing:

As the standard R-tree has no textual pruning power, a token-based R-tree, called R t -tree, by mounting tokens of subscriptions into R-tree nodes. R t -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 i , 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 R t - tree indexing and then filtered for further output processing.

MBR Filter:

Least bound rectangle filter for checks the location of the subscriber. This model filters the outcomes R t -tree index by examination the users 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. Subscribers curiosity is considered for token filter. This model filters the outcomes R t -tree index by checking the users 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 subscribers 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. 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.

SYSTEM ARCHITECTURE

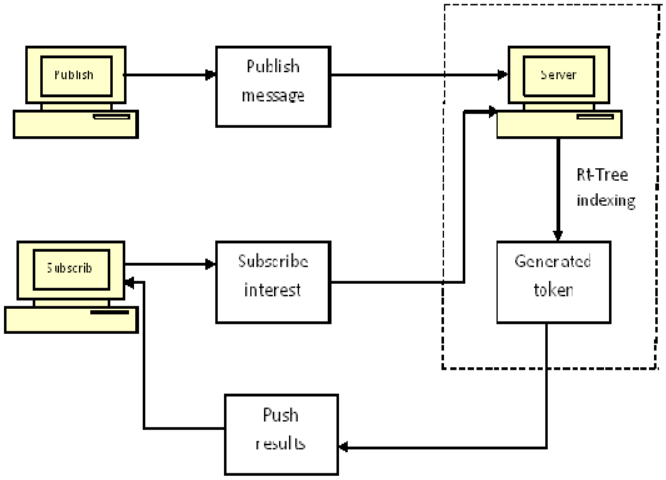

Figure 2: System Architecture

CONCLUSION

The location-aware publish/subscribe problem is studied. An effective index structure algorithm is proposed R t -tree by integrating textual description into R-tree nodes. A filter- and-verification framework is proposed and devised efficient filtering algorithms. Reducing the number of tokens in each node is proposed, which not only reduces index sizes but improves performance. An efficient algorithm to directly find answers without the verification step is defined.

REFERENCES

1. Minghe Yu, Guoliang Li, Member, IEEE, Ting Wang, Jianhua Feng, and Zhiguo Gong "Efficient Filtering Algorithms for Location-Aware Publish/Subscribe" Minghe Yu, Guoliang Li, Member, IEEE, Ting Wang, Jianhua Feng, and Zhiguo Gong.
2. 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., 1999, pp. 53–61.
3. M. Altinel and M. J. Franklin, "Efficient filtering of XML documents for selective dissemination of information," in Proc. 26 th Int. Conf. Very Large Data Bases, 2000, pp. 53–64.
4. 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.
5. 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.
6. 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.
7. 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. 1–6.
8. 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.
9. 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
10. 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.
11. G. Cugola and J. E. M. de Cote, "On introducing location awareness in publish-subscribe middleware," in Proc. IEEE Int. Conf. Distrib. Compute. Syst. Workshops, 2005, pp. 377–382.