Network Processing of PT-TOPK Queries Using Necessary Set Based Algorithm in WSN’s

P.Vinutha, T.SaiDurga

1,2Dept. of CSE, Kakinada Institute of Engineering & Tech for women(KIET_W), Korangi, E.G.dt, AP, India.

ABSTRACT:

A wireless sensor Network consists of spatially dispersed self-directed sensors to monitor the environment conditions. This work supports in-network top-k query process over uncertain data in a distributed wireless sensor network. We develop the notion of necessary Set for efficient in-network pruning of uncertain data in a distributed setting. Based on the necessary sets, we propose, necessary set-based, algorithm for in-network processing of PT-Topk queries in a two-tier hierarchical sensor network. For providing privacy proposing locked Q protocol that that prevents attackers from gaining information from both sensor collected data and sink issued queries.

KEYWORDS: Top-k queries, distributed data management, probabilistic databases, wireless sensor networks.

INTRODUCTION:

To aid management of uncertain data research on probabilistic databases have inward rehabilitated attentions in the past few years. Most of the recent works on probabilistic data modelling suggest connecting a confidence in form of probability with a data record/tuple to detain the data uncertainty and thus carry possible worlds semantic. Accordingly system issues such as indexing techniques and query processing have been scrutinized. The information that is gathered from those physical environments is of uncertain data and there is a presence of noise. Sometimes many sensor nodes are organized in an environment to keep away from the data ambiguity for sensing accuracy. This is used in a variety of applications such as military, commerce and healthcare etc. In this networks data uncertainty and energy use is main matter when we consider in the sensor networks. The energy utilization will be high as it takes many rounds of communication. They have a number of admittance tuple and materialized search states.

RELATED WORK:

A study on processing top-k queries over dispersed uncertain database is reported only support top-k queries with the expected ranking semantic. On the divergent, our suggestion is a general approach which is appropriate to probabilistic top-k queries with any semantic. Moreover, instead of frequently requesting data which may last for several rounds our protocols are definite to be completed within no more than two rounds. These differences exclusively differentiateour effort. Probabilistic ranked queries based on uncertainty at the attribute level are studied. A unique study that ranks tuples by their probabilities gratifying the query is presented. Finally, uncertain top-k query is studied under the setting of streaming databases where a compact dataset is exploited to support proficient sliding window top-k queries.

EXISTING METHOD:

The eminence of sensors differs considerably in terms of their sensing precision, accuracy, tolerance to hardware/external noise and so on. For instance, studies show that the sharing of noise varies widely in different photovoltaic sensors, precision and accuracy of readings usually differ considerably in humidity sensors and the errors in GPS devices can be up to several meters. Nonetheless, they have mostly been studied under a centralized system setting. We look at the difficulty of processing probabilistic top-k queries in distributed wireless sensor networks. Because of sensing vagueness and environmental interferences the sensor readings are usually noisy. Thus, multiple sensors are organized at certain zones in order to get better monitoring quality. In this network, sensor nodes are collected into clusters within each of which one of sensors is selected as the cluster head for performing localized data processing.

DISADVANTAGES:

The wind station very slowly. Data is not accuracy purify. The one station to another station delays the communication rate. We explore the problem of
processing probabilistic top-k queries in distributed wireless sensor networks.

PROPOSED METHOD:

The applicability of adequate set and essential set to wireless sensor networks with both two-tier hierarchical and tree-structured network topologies have shown. There are numerous top-k query semantics and solutions proposed recently, including U-Topk and UkRanks in PT-Topk in PK-Topk in expected rank in and so on. A familiar way to process probabilistic top-k queries is to first sort all tuples based on the scoring attribute and then process tuples in the sorted order to calculate the final answer set. However while focusing on optimizing the transmission bandwidth the proposed techniques require numerous iterations of computation and communication introducing remarkable communication overhead and resulting in long latency. The objective is at developing energy proficient algorithms optimized for fixed rounds of communications. we propose a Necessary set based algorithm for in-network processing of PT-Topk queries in a two-tier hierarchical sensor network. These algorithms exploit individual and combined strengths of sufficient and necessary sets in query processing. For privacy Adopted Secure protocol which prevents Attackers to gain data

ADVANTAGES:

Additionally NSB and BB take benefit of the slanted necessary sets and boundaries among local clusters to get hold of their global boundaries respectively which are efficient for intercluster pruning. We apply sufficient set and necessary set to sensor networks with tree topology to more pick up query processing performance by facilitating sophisticated in-network filtering at the intermediate nodes along the routing path to the root. The transmission cost increases for all algorithms because the number of tuple needed for query processing is increased.

SYSTEM ARCHITECTURE:

Subsequent to acceptance of all the necessary sets, the base station combines all the received tuple into a table and finds the necessary boundary called the global boundary (GB). If GB is ranked higher than the highest ranked necessary boundary, it is finished that all the necessary data have been carried out to the base station. Thus, the base station calculates the concluding answer. Or else, entering the second phase, the base station sends the GB back to the cluster heads which return the ancillary data tuple ranked between its local necessary boundary and GB. Then, the base station computes the final answer.

PT-TOPK QUERY PROCESSING:

The PT-Topk queries endow with good conditions for the targeted dispersed processing problem. The query answer can be attained by probing the tuples in descending ranking order from the sorted table which is still denoted as T for simplicity. We can simply conclude that the highest ranked k tuples are absolutely in the answer set as long as their confidences are superior than p since their qualifications as PT-Topk answers are not reliant on the subsistence of any other tuples.

SENSOR NETWORKS:

Due to the incomplete energy budget obtainable at sensor nodes, the main issue is how to increase energy-efficient techniques to decrease communication and energy costs in the networks. Approximate-based data aggregation techniques have also been proposed. The idea is to trade-off some data quality for enhanced energy efficiency. Based on numerical modelling techniques a model-driven approach was proposed to balance the confidence of the query answer against the communication cost in the network.

DATA PRUNING:

The cluster heads are dependable for producing unsure data tuples from the gathered raw sensor readings within their clusters. To respond a query, it’s normal for the cluster heads to short-laid off uncertain data tuples before delivery to the base station in order to decrease communication and energy cost. The key matter here is how to obtain a compact set of tuples necessary for the base station to answer the probabilistic top-k queries.

STRUCTURED NETWORK TOPOLOGY:

To execute in-network query processing a routing tree is often created amongst sensor nodes and the
base station. A query is issued at the root of the routing tree and spread along the tree to all sensor nodes. Even though the perception of adequate set and necessary set introduced earlier are based on two-tier hierarchical sensor networks, they are appropriate to tree-structured sensor network.

**DATA TRANSMISSION:**

Response time is one more significant metrics to assess query processing algorithms in wireless sensor networks. All of those three algorithms like SSB, NSB, and BB execute at most two rounds of message swap there is not much difference among SSB, NSB, and BB in terms of query response time. Thus we focus on the data transmission cost in the valuation. Finally, we also conduct experiments to evaluate algorithms, SSB-T, NSB-T, and NSB-T-Opt under the tree-structured network topology.

**PERFORMANCE EVALUATION:**

The distributed algorithms for processing PT-top k queries in two-tier hierarchical cluster support wireless sensor monitoring system. Inadequate energy budget is a significant issue for wireless sensor network and radio transmission is the most control source of energy consumption. Thus we gauge the total amount of data transmission as the performance metrics. Observe that the response time is another important metrics to estimate query processing algorithms in wireless sensor networks.

**ALGORITHM:**

**AT CLUSTER HEAD:**

Compute the necessary boundary

\[ NB(T_i), N(T_i) \leftarrow \{x | x \leq NB(T_i) \land x \in T_i \} \]

Deliver \( N(T_i) \) to the base station

if cluster head receive GB from the base station then

\[ N'(T_i) \leftarrow \{x | x \leq GB \land x \in \{T_i - N(T_i)\}\} \]

Now, \( N(T_i) \) is send to the base station.

end if

**AT BASESTATION:**

It receives the tuple \( N(T_i) \) from the cluster head.

\[ (1 \leq i \leq N) \]

\[ T' \leftarrow U1 \leq i \leq N N(T_i) \]

Now, it will calculate the global boundary.

if global boundary GB is less than that of \( NB(T_i) \), then

It calculate the final necessary boundary else

\[ \text{It will broadcast GB to ci and once again it collects necessary tuple} \]
\[ T' \leftarrow U1 \leq i \leq N N'(T_i) \]

end if

Where, \( x \) is the tuple \( ci \) is the cluster head

\( N(T_i) \) is the necessary set

\( NB(T_i) \) is the necessary boundary

\( Ti \) is the records collected from the sensor

\( N \) is the number of clusters in the zone

\( T' \) is the aggregation of data sets received from the clusters.

**secureQ protocol:**

SecureQ is a Proto-filter that prevents attackers from gaining information from both sensor collected data and sink issued queries. SecureQ also allows a sink to detect compromised storage nodes when they misbehave. To preserve privacy, SecureQ uses a novel technique to encode both data and queries such that a storage node can correctly process encoded queries over encoded data without knowing their values. To preserve integrity, Truth confirmation algorithm to generate integrity verification information so that a sink can use this information to verify whether the result of a query contains exactly the data items that satisfy the query.

**SECUREQ PROTOCOL TECHNIQUE**

Select any one of the value from the original data, which is already stored into the storage node. Selected data will be computed (any basic arithmetic operation and it will be selected randomly) with remaining data. Finally, generating new encoded data set, which will be viewed by the Unauthorized user. Identify the different types of sensor. Dynamically observe the number of active sensors in the sensor networks. Based on the user request, Dynamically Storage nodes stimulate the filter to rearrange the data according to the sensor id. Dynamically SecureQ filter rearrange the data based on the value of the data. Dynamically storage node sends the accurate response data to the user with the help of Proto-filter.

**CONCLUSION:**

Derived a cost model on communication cost of the proposed algorithms and propose a cost-based adaptive algorithm that acclimatizes to the application dynamics. Although our work in this paper is based mostly under the setting of two-tier
In hierarchical network, the concepts of sufficient set and necessary set are universal and can be easily expand to a network with tree topology. The performance evaluation authenticates our ideas and shows that the proposed algorithms decrease data transmissions extensively.

REFERENCES:


Miss. P. VINUTHNA is a student of kakinada Institute of Engineering and Technology for Women (kiet-w) korangi. Presently she is pursuing her M.Tech in Computer science Engineering from this college and she received her B.Tech from SASI Institute of Engineering & Technology , affiliated to JNT University, Kakinada in the year 2011. Her areas of interest includes Computer Networks and Object oriented Programming languages, Datawarehouse and Datamining all current trends and techniques in Computer Science.

Ms. T. SAI DURGA, well known Author and excellent teacher Received B.Tech from KIET Engineering college affiliated to JNT university, kakinada in the year of 2007 and M.Tech (CSE) from KIET Engineering college affiliated to JNT university, kakinada in the year of 2011. She is working as Associate Professor, Department of B.Tech, M.Tech Computer science engineering, Kakinada Institute of Engineering Technology for women(kiet-w) korangi. She has 7 years of teaching experience in various engineering colleges. Her areas of Interest includes Data Warehouse and Data Mining, Computer Networks and other advances in computer Application.