International Journal of Scientific Research in Computer Science and Engineering
Research Paper Vol-1, Issue-3 ISSN: 2320–7639

# Efficient Computation of Range Aggregates against Uncertain Location Based Queries using Filtering-and-Verification Algorithm

S.Praveen Kumar*[1], Dr.R.V.Krishnaiah [2]

*[1] *Dept of CSE, DRK College of Engineering and Technology, Hyderabad, Andhra Pradesh, India*
[2]*Dept of CSE, DRK College of Engineering and Technology, Hyderabad, Andhra Pradesh, India*
**Available online at www.isroset.org**

Received: 26 May 2013 Revised: 28 May 2013 Accepted: 20 June 2013 Published: 30 June 2013

***Abstract:*** In many applications, including location based services, queries may not be precise. In this paper, we study the problem of efficiently computing range aggregates in a multidimensional space when the query location is uncertain. We propose novel, efficient techniques to solve the problem following the filtering-and-verification paradigm.

***Keywords:*** Uncertainty, Range Aggregates, Filtering-and-Verification

## I. INTRODUCTION

Query imprecision or uncertainty may be often caused by the nature of many applications, including location based services [1] [2]. The existing techniques for processing location based spatial queries regarding certain query points and data points are not applicable or inefficient when uncertain queries are involved. In this paper, we investigate the problem of efficiently computing distance based range aggregates over certain data points and uncertain query points as described in the abstract [3]. In general, an uncertain query Q is a multi-dimensional point that might appear at any location x following a probabilistic density function pdf(x) within a region Q.region. There is a number of applications where a query point may be uncertain

Motivating Application_: A blast warhead carried by a missile may destroy things by blast pressure waves in its lethal area where the lethal area is typically a circular area centered at the point of explosion (blast point) with radius γ

## II. CONTRIBUTIONS

We study the problem of aggregate computation against the data points is motivated by the above applications which have at least probability θ to be within distance r regarding an uncertain location based query [8] [9]. We calculate the probability for a naive way to solve this problem is that for each data point q ∈X, namely falling probability of P within r distance to q, select q against a given possibility entry, and then conduct the aggregate. This involves the computation of an essential regarding each q and P.pdf for each q ∈X; unless P.pdf has a very simple sharing such a computation may often be very exclusive and the naive method may be computationally excessive when a large number of data points is involved [10] [11]. In the paper we target the

Corresponding Author: *S.Praveen Kumar*

and γ depends on the explosive used [4] [5] . While firing such a missile, even the most advanced laser-guided missile cannot exactly hit the aiming point with 100% guarantee. The actual falling point (blast point) of a missile blast warhead regarding a target point usually follows some probability density functions (PDFs) [5] . different PDFs have been studied . where bivariate normal distribution is the simplest and the most common one. In military applications, firing such a missile may not only destroy military targets but may also damage civilian objects. Therefore, it is important to avoid the civilian casualties [6] [7] by estimating the likelihood of damaging civilian objects once the aiming point of a blast missile is determined. Points {pi} for $1 \le i \le 7$ represent some civilian objects (e.g., residential buildings, public facilities). The actual falling point of the missile, then objects p1 and p5 will be destroyed. Similarly, objects p2, p3 and p6 will be destroyed if the actual falling point is q2. In this application, the risk of civilian casualties may be measured by the total number n of civilian objects which are within γ distance away from a possible blast point with at least θ probability problem of proficiently computing range aggregates against an uncertain P for arbitrary P.pdf and P.region

The Our techniques are developed based onthe standard filtering-and-verification paradigm. We first discuss how to apply the existing probabilistically constrained regions (PCR) techniques to our problem.Then, we propose two novel distance based filtering techniques, statistical filtering (STF) and anchor point filtering (APF)[12] [13]. respectively, to address the inherent limits of the PCR technique To the best of our knowledge, we are the first to identify the problem of computing range aggregates against uncertain location based query. In this paper, we investigate the problem regarding both continuous and discrete Q [14][15]. Our principle contributions can be summarized as follows.

- We propose two novel filtering techniques, STF and APF [17] [18] respectively. The STF technique has a decent filtering power and only requires the storage of very limited pre-computed information. APF provides the flexibility to significantly enhance the filtering power by demanding more pre-computed information to be stored [19 [20]. Both of them can be applied to continuous case and discrete case.
- Extensive experiments are conducted to demonstrate the efficiency of our techniques.
- While we focus on the problem of range counting for uncertain location based queries [21] in the paper, our techniques can be immediately extended to other range aggregates. different PDFs have been studied in [22]

## III. PROBLEM DEFINITION

In the paper, S is a set of points in a d-dimensional numerical space. The distance between two points x and y is denoted by δ(x, y). Note that techniques developed in the paper can be applied to any distance metrics In the examples and experiments, the Euclidean distance is used.

For two rectangular regions r1 and r2, we have δmax(r1, r2) = max∀x∈r1,y∈r2 δ(x, y) and δmin(r1, r2) = 0 if r1 ∩ r2 _= ∅ min∀x∈r1,y∈r2 δ(x, y) otherwise (1)

An uncertain (location based) query Q may be described by a continuous or a discrete distribution as follows.

Definition 1 (Continuous Distribution). An uncertain query Q is described by a probabilistic density function
Q.pdf Let Q.region represent the _ region where Q might appear, then
x∈Q.region Q.pdf(x)dx = 1;
Definition 2 (Discrete Distribution). An uncertain query Q

## IV. INDENTATIONS

*Filtering-and-Verification Algorithm*

This motivates us to follow the *filtering-and-verification* paradigm for the uncertain aggregate query computation. Particularly, in the *filtering phase,* effective and efficient filtering techniques will be applied to *prune* or *validate* the points. The algorithm consists of two phases. In the *filtering* phase for each entry e of RS to be processed, we do not need to further process e if it is *pruned* or *validated* by the filter F. We say an entry e is *pruned* (*validated*) if the filter can claim Pfall (p, γ) < θ (Pfall (p, γ) ≥ θ) for any point p within embb. The counter cn is increased by |e| if e is *validated* where |e| denotes the aggregate value of e (i.e., the number of data points in e). Otherwise, the point p associated with e is a candidate point if e corresponds to a data entry and all child entries of e are put into the queue for further processing if e is an intermediate entry. The *filtering phase* terminates when the queue is empty. In the *verification* phase candidate points are *verified* by the integral calculations. Modules:

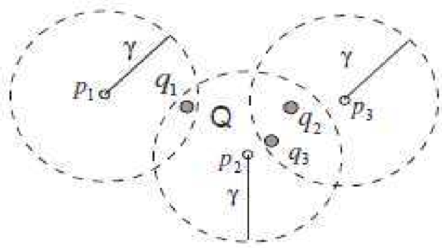
consists of a set of instances (points) {q1, q2, . . . , qn} in a d dimensional numerical space where qi appears with probability Pqi, and _ q∈Q Pq = 1;

where Q can have a non-zero probability to be absent;
that is, _x∈Q.region Q.pdf(x)dx = c or_q∈Q Pq = c for a c < 1.
For a point p, we use Pfall(Q, p, γ) to represent the probability of Q within γ distance to p, called falling probability of p regarding Q and γ. It is formally defined below.

For continuous cases,

Pfall(Q, p, γ) =_x∈Q.region ∧ δ(x,p)≤γ
Q.pdf(x)dx (2)
For discrete cases,
Pfall(Q, p, γ) =_q∈Q ∧ δ(q,p)≤γ Pq (3)

In the paper hereafter, Pfall(Q, p, γ) is abbreviated to Pfall(p, γ), and Q.region and Q.pdf are abbreviated to Q and pdf respectively, whenever there is no ambiguity. It is immediate that Pfall(p, γ) is a monotonically.




Example of Pfall(Q, p, γ)

*Filtering and verification:*

When User wants to Search any Civilian Objects like Hotels, Hospitals, Banks, etc., to generate & verify the database then gave exact details. If, he wants to Advanced Search, to give exact range values, apart from that civilian objects details to be displayed with Civilian Objects categorized and calculate the distance details.

*Query Processing:*

Admin Verify the User Requests and check the user requests is contain in our database or not. If the database contains the user requests to calculate the Range (Distance) value, and then send Response to the Appropriate Values, suppose the request doesn't contain in the database send response is 'Record Not Found' that query sent to the particular User.

*Response Results:*

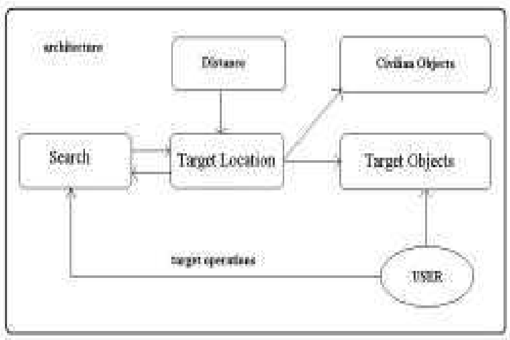
User Verify the Admin Responses and the user requests is match with database the query produce with the Correct

Results and Download the word file, otherwise not download. If the database contains the user requests to calculate the Range (Distance) value, suppose the request doesn't contain in the database received response is 'Record Not Found' but if matching results are retrieved.

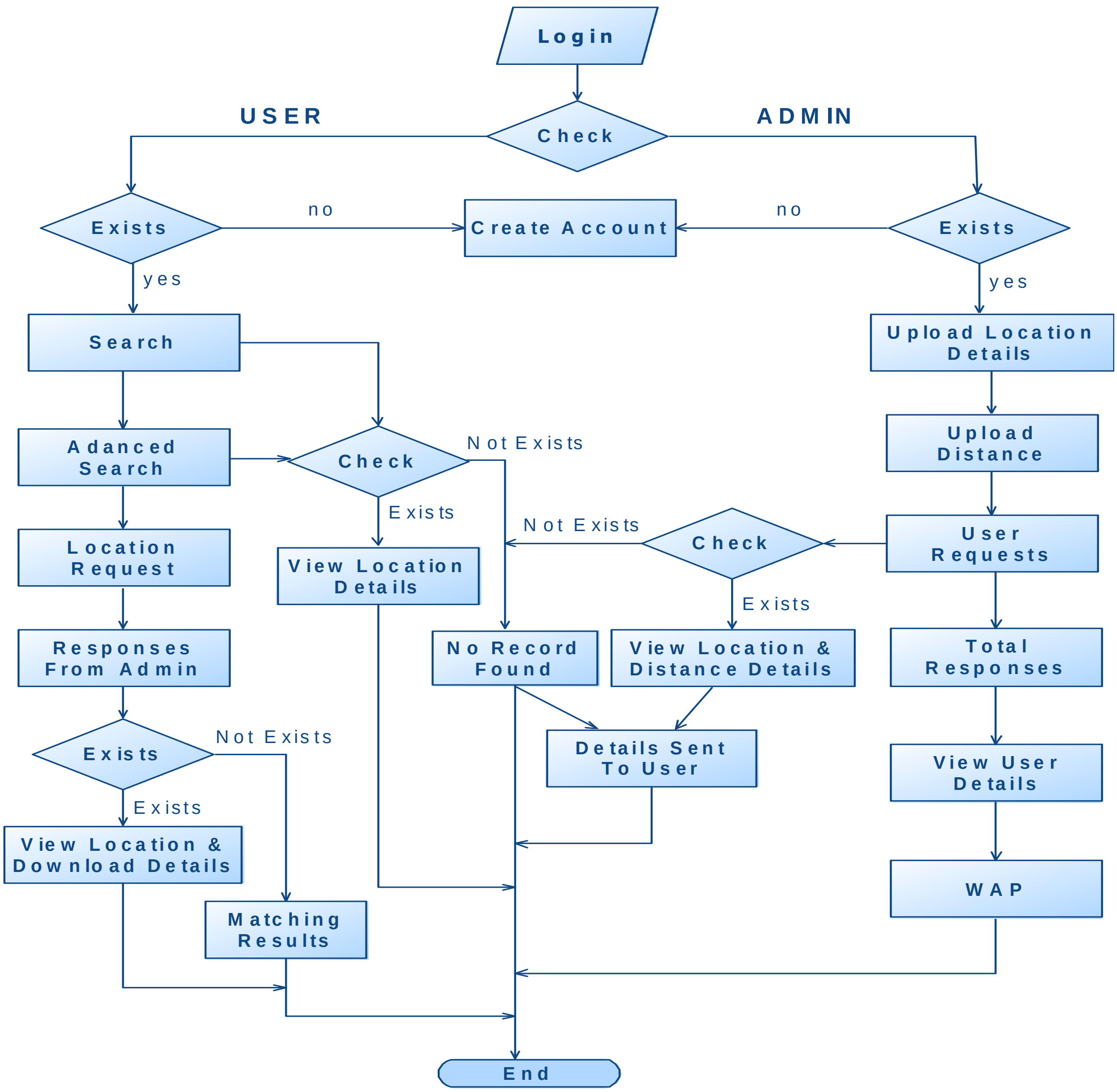
*Upload Civilian Objects with Distance:*

Admin can upload the Civilian Objects details with Correct Location (area), Address, Phone and Upload Civilian Objects Profile (.doc) file. If calculate distance in one location to another location so the admin entry the details for distance and admin can view the User Details also.

## V DIAGRAMS ARCHITECTURE




DATAFLOW DIAGRAM



## V. EXPERIMENT RESULTS

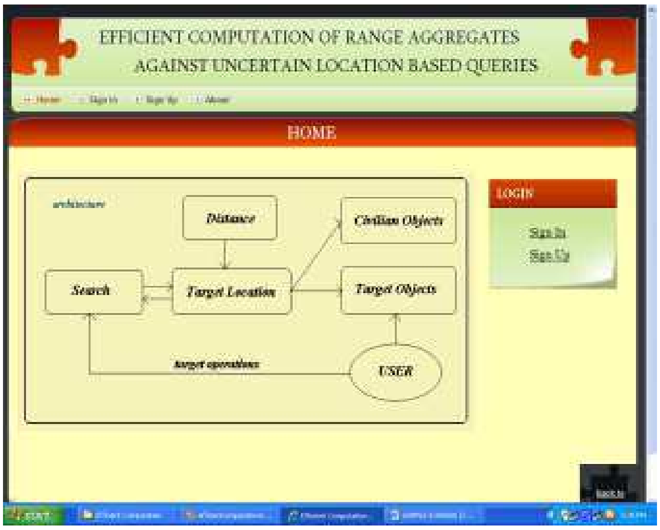



Fig 1:figure shows the home page


Fig 2:shows the login page


Fig 3:Show the signup page if the user/admin login for the first time


Fig 4:login page for user/admin


Fig 5:Shows the upload details for the admin


Fig 6:Shows the upload distance from source to destination


Fig 7:shows the search request for a location


Fig 8:shows the user search request


Fig 9:shows the location request by user


Fig 10:shows the total responses from admin


Fig 11:shows the response as record not found

## VI. CONCLUSION

In this paper, the problem of uncertain location based range aggregate query in a multidimensional space; it covers a wide spectrum of applications.To efficiently process such a query, we propose a general filtering and verification framework and two novel filtering techniques, named *STF* and *APF* respectively, such that the expensive computation and IO cost for verification can be significantly reduced. Our experiments convincingly demonstrate the effectiveness and efficiency of our techniques.

## REFERENCES


[1] P. K. Agarwal, S.-W. Cheng, Y. Tao, and K. Yi. Indexing uncertain data. In Proc. Symp. Principles of Database Systems (PODS), 2009.

[2] C. Aggarwal and P. Yu. On high dimensional indexing of uncertain data. In Proc. Intl Conf. Data Eng. (ICDE), 2008.

[3] C. Bohm, M. Gruber, P. Kunath, A. Pryakhin, and M. Schubert. Prover: Probabilistic video retrieval using the gauss-tree. In Proc. Intl Conf. Data Eng. (ICDE), 2007.

[4] C. Bohm, A. Pryakhin, and M. Schubert. Probabilistic ranking queries on gaussians. In Proc. Intl Conf. Scientific and Statistical Database Management (SSDBM), 2006.

[5] V. Bryant. Metric Spaces: Iteration and Application. Cambridge University Press, 1996.

[6] J. Chen and R. Cheng. Efficient evaluation of imprecise location dependent queries. In Proc. Intl Conf. Data Eng. (ICDE), 2007.

[7] R. Cheng, J. Chen, M. F. Mokbel, and C.-Y. Chow. Probabilistic verifiers: Evaluating constrained nearest-neighbor queries over uncertain data. In Proc. Intl Conf. Data Eng. (ICDE), 2008.

[8] R. Cheng, D. V. Kalashnikov, and S. Prabhakar. Evaluating probabilistic queries over imprecise data. In Proc. ACM SIGMOD, 2003.

[9] R. Cheng, S. Singh, and S. Prabhakar. Efficient join processing over uncertain data. In Proc. Int'l Conf. Information and Knowledge Management (CIKM), 2006.

[10] R. Cheng, Y. Xia, S. Prabhakar, R. Shah, and J. S. Vitter. Effcient indexing methods for probabilistic threshold queries over uncertain data. In Proc. Intl Conf. Very Large Data Bases (VLDB), 2004.

[11] G. W. Cordner. Police patrol work load studies: A review and critique. Police Studies, 2(3):50–60, 1979.

[12] X. Dai, M. Yiu, N. Mamoulis, Y. Tao, and M. Vaitis. Probabilistic spatial queries on existentially uncertain data. In Proc. Intl Symp. Large Spatio-Temporal Databases (SSTD), 2005.

[13] E. Frentzos, K. Gratsias, and Y. Theodoridis. On the effect of location uncertainty in spatial querying. IEEE Trans. Knowl. Data Eng., 21(3):366–383, 2009.

[14] M. Hua, J. Pei, W. Zhang, and X. Lin. Ranking queries on uncertain data: A probabilistic threshold approach. In Proc. ACM SIGMOD, 2008.

[15] Y. Ishikawa, Y. Iijima, and J. X. Yu. Spatial range querying for gaussian-based imprecise query objects. In Proc. Intl Conf. Data Eng. (ICDE), 2009.

[16] H.-P. Kriegel, P. Kunath, M. Pfeifle, and M. Renz. Probabilistic similarity join on uncertain data. In Proc. Intl Conf. Database Systems for Advanced Applications (DASFAA), 2006.

[17] H. P. Kriegel and M. Pfeifle. Density-based clustering of uncertain data. In Proc. ACM SIGKDD, 2005.

[18] X. Lian and L. Chen. Monochromatic and bichromatic reverse skyline search over uncertain databases. In Proc. ACM SIGMOD, 2008.

[19] R. Meester. A Natural Introduction to Probability Theory. Addison Wesley, 2004.

[20] W. K. Ngai, B. Kao, C. K. Chui, R. Cheng, M. Chau, and K. Y. Yip. Efficient clustering of uncertain data. In Proc. Int'l Conf. on Data Mining (ICDM), 2006.

[21] J. Pei, B. Jiang, X. Lin, and Y. Yuan. Probabilistic skyline on uncertain data. In Proc. Intl Conf. Very Large Data Bases (VLDB), 2007.

[22] G. M. Siouris. Missile Guidance and Control Systems. Springer Publication, 2004.


BIOGRAPHY:

S.Praveen Kumar has completed M.C.A from Karuna P.G College and pursuing M.Tech (C.S.E) in DRK College of Engineering and Technology, JNTUH, Hyderabad, Andhra Pradesh, India. His main research interest includes Data Mining, Information Security, network protection and security control.

Dr.R.V.Krishnaiah, did M.Tech (EIE) from NIT Warangal, MTech (CSE) form JNTU, Ph.D, from JNTU Anantapur, He has memberships in professional bodies MIE, MIETE, MISTE. He is working as Principal in DRK Institute of Science and Technology, Hyderabad. His main research interests include Image Processing, Security systems, Sensors, Intelligent Systems, Computer networks, Data mining, Software Engineering, network protection and security control. He has various publications and presentations in various national and international journals.