SR-time (Corruption time + Rerank time).It will combine the corruption and ranking module together. This way re-ranking is done on-the-fly during corruption. The Ranking algorithm proposed in this method is PRMS (Probabilistic Retrieval Model for Structured Data) to get the better results of approximation algorithm. It computes the language model of each attribute value smoothed by the language model of its attribute. It assigns each attribute a query keyword-specific weight, which specifies its contribution in the ranking score. It computes the keyword-specific weight j(q) for attribute values whose attributes are Tj and query q as,

$$\mu_j(q)= P(q/T_j) /(\sum T\epsilon DB * P(q/T))$$

3.6 Performance Study

In this paper the performance of SR score and SGS Approximation is studied.

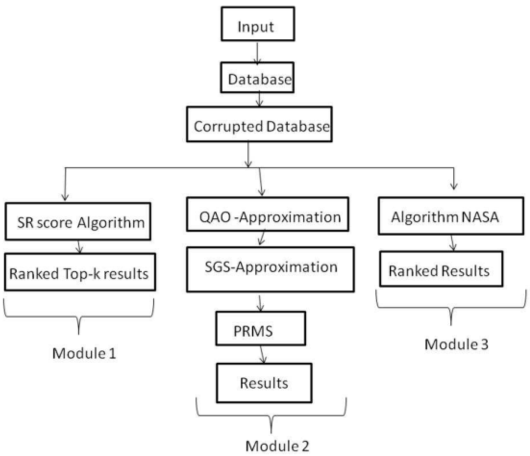

Figure 3: Execution flow of Modules.

SR Algorithm

In first module SR Algorithm implementation is done as shown in Figure.3, which is the existing in our reference [1].SR-time mainly consists of two parts: the time spent on corrupting K results and the time to re-rank the K corrupted results. We have reported SR-time using (corruption time + re-rank time) format. SR Algorithm incurs a considerable time overhead on the query processing which is higher on query processing . Figure. will gives expected result of SR score.

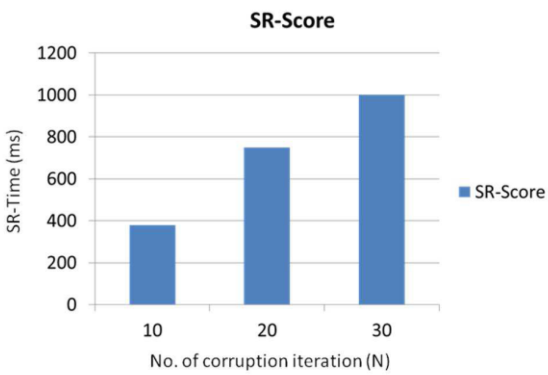

Figure 4: Result of SR-Score.

SGS-Approximation

In next module SR Algorithm implementation is done as shown in Figure 2, which is the proposed work of our reference[1].QAO measure the prediction effectiveness for smaller values Of N.SGS Approximation on the database ,re-ranking is done during the corruption. SR time is mentioned as corruption time only. Figure 5 will gives expected results of SGSApproximation. According to our performance study of QAOApproximation, SGS-Approximation, and the combined result over database ,It delivers the best balance of improvement in efficiency and reduction in effectiveness for database. It will achieves high prediction accuracy as compared with SR score algorithm. To improve performance of SR score a new technology is used named as NASA. It will reduce the SR time of execution and gives high prediction accuracy.

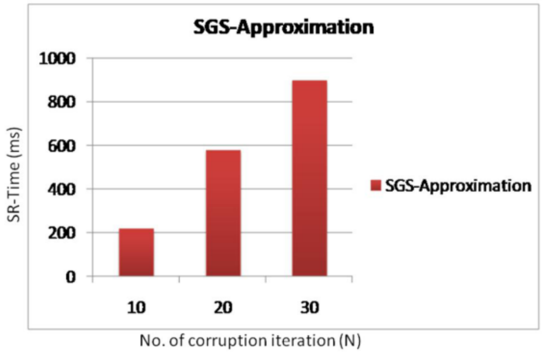

Figure 6: Result of SGS Approximation.

3.7 Disadvantage of Existing System

The average computation time of SR score by means of SR Algorithm and compare it to the average query processing time using PRMS for the queries in our query workloads. SR Algorithm incur a considerable time overhead on the query processing. This operating cost is higher for queries over the database. QAO measure the prediction effectiveness for smaller values of N using average correlation score. The results of applying SGS-Approx on the database. Since re-ranking is done on-the-fly during the corruption, SR-time is reported as corruption time only.

3.8 Problem Definition

The time to compute the SR score only depends on the top-K results, since only the top-K results are corrupted and reranked. Increasing the data set size will only increase the query processing time; the complexity of data schema could have impact on the efficiency of our model. To overcome this problem a new method named as NASA is using to reduce time complexity with the help of k- Nearest Neighbor algorithm (k- NN)

4. Proposed Work.

A computational problem instance has an input and an objective it wants to compute on that input. An algorithm is a procedure to compute the objective function. [13]. SR Algorithm incur a considerable time overhead on the query processing. The time to compute the SR score only depends on the top-K results, since only the top-K results are corrupted and re-ranked. To overcome this problem of time overhead new algorithm is going to launched called as NASA. NASA means K-Nearest Neighbor algorithm Algorithm implementing in structured Robustness Algorithm. Algorithm 1 will explain the execution flow of NASA.In pattern recognition, the k-Nearest Neighbors algorithm (or k-NN for short) is a nonparametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The neighbors are taken from a set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. k-NN has some strong consistency results. As the amount of data approaches infinity, the algorithm is guaranteed to yield an error rate. k-NN uses the previous instances as a model for future instances and prediction for the current instance is chosen as the classification of the most similar previously observed instance. Instances with correct classifications (predictions) (xi, f (xi)) are stored in memory. Given, a new instance is xq, and the prediction is most similar to is instance xk.

4.1 The execution flow of K-Nearest Neighbors is algorithm as follows

First Determine the parameter (K) number of nearest neighbors. then Calculate the distance between the query instance and samples. With use any distance algorithm. Sort the distances for all the samples and determine the nearest neighbor based on the K-th minimum distance Use the majority of nearest neighbors as the prediction value

---

Algorithm 1. NASA

Input: Query Q, Top-k, result L of Q ranking function g,Metadata M, Inverted index I, Number of Corruption iteration n, Process X, Number of Neighbour Kn(1,2,......,n-1), Database DB

---

1. SR ←0;C ← 0; // C Caches Keywors in Q
2. FOR i = 1 → N DO
3. I'←I;M' ←M;L' ←L; // Corrupted copies of I , M and L
4. For each result R in L Do
5. For each query Xl → N-1 DO
6. Find Kn Nearest Neighbor
7. Use Euclidean distance between two instances d(Xm,Xn) = Sqrt(Sumeforr = 1toN(ar(Xm) - ar(Xn))2)
8. A' ←A;// Corrupted version of A
9. For each keywords w in Q Do
10. Compute = of w in A' for keywords in Q as needed but not in C ,calculate and cache them
11. IF w varies in A' and A THEN
12. Update A',M' and entry of w in I'
13. Add R' to L'
14. Rank L' using g which returns L based on I',M'
15. SR ← Sim(L;L')
16. RETURN SR←SR/N; // AVG score over N rounds

4.1 Advantages of k-NN

1. K-Nearest Neighbors algorithm just store instances so the utilization of memory space can be less.
2. K-Nearest Neighbors algorithm can handle complex target functions and it can improve the loss of information.

5. Conclusion

The conclusion of the frame work shows Post retrieval methods predict the difficulty of a query with computing its results. The computation the query prediction based on SR algorithm gives the outcomes of corrupted database in re-ran ked form. Execution time depends on Increasing the data set size will only increase the query processing time The complexity of data schema could have impact on the efficiency of our model. k- NN algorithm helps to reduce time complexity and outcomes are in the form of ranking. So NASA will help to give a more effective result in ranking order with respect to less time of execution.

References

[1] Shiwen Cheng, Arash Termehchy, And Vagelis Hristidis, Efficient Prediction Of Difficult Keyword Queriesover Databases IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 6, June 2014.

[2] JoaquIn PErez-Iglesias And Lourdes Araujo, Evaluation Of Query Performance Prediction Methods By Range, E. Chavez And S. Lonardi (Eds.): Spire 2010, Lncs 6393, Pp. 225236, 2010.

[3] Joaqun Perez-Iglesias and Lourdes Araujo, Evaluation of Query Performance Prediction Methods by Range E. Chavez and S. Lonardi (Eds.): SPIRE 2010, LNCS 6393, pp.225236, 2010.

[4] Ben He and Iadh Ounis ," Inferring Query Performance Using Pre-retrieval Predictors", Department of Computing Science University of Glasgow.

[5] Oren Kurland1, Anna Shtok1, David Carmel2, And Shay Hummel, A Unified Framework For Post-Retrieval Query- Performance Prediction

[6] Thanh Tran And Lei Zhang Keyword Query Routing, IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 2, February 2014 363 J. Clerk Maxwell, A Treatise On Electricity And Magnetism, 3rd Ed., Vol. 2. Oxford: Clarendon, 1892, Pp.68-73.

[7] Luca Di Angelo1 And Luigi Giaccari,," An Efficient Algorithm For The Nearest Neighbourhood Search For Point Clouds," Ijcsi International Journal Of Computer Science Issues, Vol. 8, Issue 5, No 1, September 2011.

[8] Claudia Hauff,"Predicting The Effectiveness Of queries And Retrieval Systems", January 29, 2010.

[9] Xiaogang Wang, Member, IEEE , Shi Qiu, Ke Liu, and Xiaoou Tang, Fellow, IEEE,"Web Image Re-Ranking

Using Query-Specific Semantic Signatures",VOL. 36, NO. 4, APRIL 2014.

[10] Yang Kehua, Abdoullahi Diasse Hunan University, A Dynamic Materialized View Selection In A Cloud-Based Data Warehouse, Ijcsi International Journal Of Computer Science Issues, Vol. 11, Issue 2, No 1, March 2014.

[11] Yun Zhou And W. Bruce Croft, Measuring Ranked List Robustness For Query Performance Prediction, Jun 10,2007.

[12] Polynomial Time Algorithms , Computational Complexity , June 4th, 2009.

[13] Jitao Sang And Changsheng Xu, Browse By Chunks: Topic Mining And Organizing On Web-Scale Social Media, Acm Transactions On Multimedia Computing, Communications And Applications, Vol. 7s, No. 1, Article 30, Publication Date: October 2011.

[14] Leong Hou U, Hong Jun Zhao, Man Lung Yiu, Yuhong Li, and Zhiguo Gong,"Towards Online Shortest Path Computation", VOL. 26, NO. 4, APRIL 2014.

[15] Jagbeer Singh, Bichitrananda Patra, Satyendra Prasad Singh, An Algorithm To Reduce The Time Complexity Of Earliest Deadline First Scheduling Algorithm In Real-Time System, (Ijacsa) International Journal Of Advanced Computer Science And Applications,Vol. 2, No.2, February 2011.

[16] S. Cheng, A. Termehchy, and V. Hristidis,2012 , Predicting the effectiveness of keyword queries on database..

[17] Chihung Chi;Y eZhou; andXiaojun; Performance Prediction For Performance Sensitive Queries; Issnll1007-0214ll08/10llpp618-628 V ol:18;No6;December2013: