Volume 1 No. 3, July 2011 ISSN-2223-4985
International Journal of Information and Communication Technology Research

http://www.esjournals.org

Data Retrieval and Gaining Optimistic Nepotism Samples

[1]SarathChand P.V, [2]Balraju M, [3]Bhukya Shankar Nayak,
[4]Rambabu pemula, [5]Suryaprakash G
[1]Author, Indur Institute of Eng. & Technology,Siddipet, Medak Dist.A.P. India
[2]Co-author, Madhira Engineering College,Hyderabad. A.P. India
[3] Co-author, Madhira Educational society group of Institutions
[4] Co-author, Nimra Institute of Engineering & Technology, Ongole , Andhra Pradesh, India.
[5] Co-author, Department of IT, ACE Engineering college,Ghatkesar, Andhra Pradesh, India.

ABSTRACT

Data ware house mostly concentrates on transactional processing and operational systems. Data mining tools helps in automating day to day operations of organizations where data is locked up with in the each transactional system for inaccessible of the required data sample. In accessing these situations, sometimes data mining tools may be disconnected to the data ware house. The data is locked up with in the each transaction processing system and could not be used effectively for organization wide information retrieval and decision support functions. To avoid these problems a new algorithm was discovered whereby the organizations extract the data samples from their informational assets through the use of spatial stores called Data ware houses. The algorithm which is described in this paper is a subject oriented, integrated, time variant and non-volatile. It is created by both the internal systems and external sources of data.

Keywords: *Biased Samples, hits, Randomization, Data assets, Customization, Streams*

1. INTRODUCTION

The term Nepotism constitutes the favorite data samples for the retrieval operations. It mainly deals with gathering of data from multiple transaction processing systems and the external sources which ensure the data quality, organizing the data for information processing and providing information retrieval and analyzing through on-line analytical processing (OLAP), reporting, web based and data mining tools. Data ware house was quickly evolved in a popular business application area. The builders of data ware house consider the system to be the key components for the informational retrieval. Many examples can be cited of highly successful data ware house developed and deployed for different business types. The data retrieval is major task for the organizations to analyze and track business dimensions.

2. OPTIMISTIC NEPOTISM SAMPLES

The computational elements that constitute the most optimistic Data retrieval models are called as Optimistic Nepotism Samples (ONS). These are often referred as nodes, units, or processing elements. The processing elements in ONS are one-to-one relationship with actual samples. A single processing element (PE) as a representative of the collectively activity of a group of samples. This model is similar to the brain models, and also it will make the problem more traceable when we are attempting to model the behavior of some sort of complexes data retrieval. The processing elements is shown in the figure1

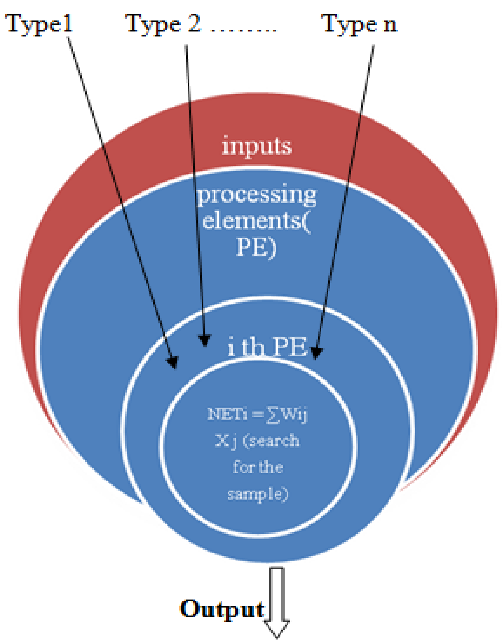

Figure 1: The Processing Method

The structure represents the single PE in the form a data samples. The input connections are modeled as arrows from the user or organization. Every input is associated with a quantity w_{ij} weight. The single out-put value which is the retrieval from the ware house and can be fan out for other PEs. Each PE is numbered with **i** th, the PE has many inputs but has only one output, which can be transferred to many PEs in the Distributed Databases Network. The input the **i** th receives from the **j** th PE as indicated by $\mathbf{x_i}$. This value is the output of the **j** th node as output generated by the **i** th node is labeled as $\mathbf{x_i}$**.** Each connection with the **i** th PE is associated with a weight or strength. The weights of the retrieval form one node to another node are denoted $\mathbf{w_{ij}}$. Each PE determines a net input values based on the inputs samples and in the absence of Special sample, typically calculate the net input by summing the input values by multiplying with their corresponding weights. The input of the **i** th unit can be written as

$$\mathbf{NET_i = \sum_j Wij\ X\ j}$$

where the index **j**, runs over all the samples to be retrieved by the PE. This sum of the products plays an important role in the retrieval process of the data samples because there is often a very large number of interconnection exists in the data ware house. The speed at which the calculation can be performed usually determines the performance of any given Ware house. Once the net input is calculated it is converted to an activation value or simply activation for the PE. The activation is described by the value as

$$\mathbf{a_i(t)=S_i(a_i(m-1),NET_i(t))}$$

which denotes the activation is an explicit function of the NET inputs. Generally the current activation depends on the previous inputs of the activation $\mathbf{a(m-1)^2}$. Once the PE value is calculated the output value is determined by applying the function

$$\mathbf{X_i = F_i(a_i)\ where\ a_i = NET_i}$$

And the function is normally written as

$$\mathbf{X_i = F_i(NET_i)}$$

3. DATA RETRIEVAL BY ONS

The ONS has been initialized so that the internal data structures have been allocated and contains the valid information. The user has to set the outputs from the input samples to a normalized vector to be propagated through the set of samples.. Once the PE generates the output, the user application reads the output vector from the appropriate arrays and uses the output accordingly. The ONS starts from the hidden samples since the input samples are assumed as normalized and available in the ware house. The retrieval begins by calculating the total input samples by each unit on the hidden layer. The units are N > 1 with the largest aggregate input is declared as the **hit**, the required data is found and the output from that unit is set to 1. The outputs from all the losing units are set to 0.Once the processing on the hidden layer is completed the rate of the output data is calculated by the performance of another sum of products at each node.

3.1 ONS production algorithm

This algorithm describes to enter to the hidden data and retrieval from the hidden data as an output. The hidden data is Hidden_Data and the retrieval data is Retrieval_Data.

3.1.1 Procedure for Hidden_Data

```
(NET: ONS FIRST, SECOND, THIRD: INTEGERS)
Var Units: float [];
//pointer to output units
Invec : float[]
//pointer to input units
Connects : float[]
//pointer to array for iterations
I,j : integer
Begin
Units = NET.Hidden_Data.OUTS
// locate the output array
For i=1 to length (units)
Do
Units[i] = 0;
Invec = NET.INPUTS.OUTS;
// for all hidden data and to locate the //array
Connects = NET.Hidden_Data.WEIGHTS[i];
//locating the inputs
For j=1 to length do
Units[i] = units[i] + connects[j] * Invec[j];
End do;
Rank (FIRST, SECOND, THIRD)
End do
Compete (NET.Hidden_Data.OUTS, FIRST, SECOND, THIRD)
End procedure;
```

3.1.2 Procedure for Processing Elements

```
Procedure Compete
(UNITS: float [] , FIRST, SECOND, THIRD: INTEGER)
//generates output for the UNITS by //using compete function
Var outputs : float []; //output array
Sum : float;
Hit , Area, show :float; //store output
I : integer
Begin
Outputs = UNITS;
Sum = Outputs (FIRST);
Hit = Outputs (FIRST);
If (SECOND!= 0 )
Then
Sum = Sum + Outputs (SECOND);
Area = Outputs (SECOND);
If(THIRD !=0)
```

Volume 1 No. 3, July 2011 ISSN-2223-4985
International Journal of Information and Communication Technology Research

http://www.esjournals.org

```
Then
Sum = Sum + Outputs (THIRD);
Show = Outputs(THIRD);
End if;
End if;
For i= 1 to length (units) // for all hidden //data
Do
Outputs[i] =0;
End do;
Outputs [FIRST] = Hit / sum;
If (SECOND!= 0 )
Then
Outputs [SECOND] = place / sum;
If (THIRD! =0)
Then
Outputs [THIRD] = show / sum;
End if;
End if;
End procedure;
```

3.1.3 Procedure for Retrieval_Data

```
(NET: CPN;: FIRST, SECOND, THIRD: INTEGER)
//generates the output
Var units: float [];
Hidvec : float [];
Connects : float [];
I : integer;
Begin
Units = NET.OUTPUTS.OUTS;//starts //the output array
Hidvec = NET.HIDDENS.OUTS;//start //of hidden array
For i=1 to length (Units) // for all output //units
Do
Connects = NET.OUTPUTS.WEIGHT[i];
Units[i] = Hidvec[FIRST] * connects[FIRST];
If (SECOND !=0) //if there second hit
Then
Units[i] = units[i] + hidvec[SECOND]* connects
[SECOND];
If (THIRD!=0)//if there is second hit
Then
Units[i] =Units[i]+ hidvec[THIRD] * connects[THIRD];
End if;
End if;
End do;
End procedure;
```

4. IMPLEMENTAION OF ONS

Instead of dealing with the entire Data Ware house, we can think of sampling the steams at periodical levels by using the ONS technique. To obtain the unbiased samples we must know the length of the data stream in advance. If the length is not available we have to modify the approach by using the Reservoir Samples. This is also a part of the ONS technique and is used to select the unbiased data samples of **'n'** number of elements without replacements. It is to maintain a sample of size at least **'n'** called the "Reservoir'. From which the random sample of **'n'** can be generated. The Reservoir should be maintained by a set of **'n'** candidates in the reservoir, which are the true random samples of the elements seen so far in the streams. As the data stream flows every new element has certain probability of replacing an old element in the reservoir. The probability is **n/N.** It maintains the invariant that the set of **'n'** candidates in the reservoir forms random samples of the elements so far. The ONS operates in the form of random sampling and sketching, is often deal with the massive and high dimensional data samples. The ONS algorithm is generally used for randomization which leads to a simpler and more efficient algorithm in comparison to known deterministic algorithms. The ONS returns the right answer.

Let Y be a random data to be retrieved with a mean value of µ and the standard derivation is $\boldsymbol{\sigma^2}$. The ONS says that

$$\mathbf{P(|Y-\mu| > s) <= \sigma^2/s^2.}$$

For any positive real number s. This inequality can be used to bind the variance of the random data. In much case the multiple random data sample are used to boost the confidence in our results. The graph can be shown the availability of the random data sample to retrieved by the ONS algorithm is shown in the **chart 1**

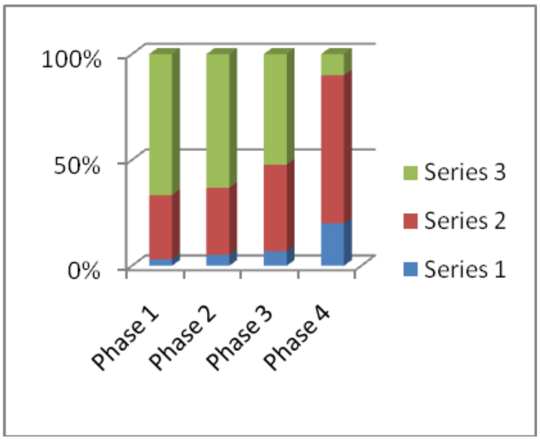

Chart 1

In the chart the randomized data retrieval increased by the function of ONS. The red color indicates the data retrieval rate of percentage which at the most of 99% and the green, blue indicates the massive presence of the data in the data ware house.

5. CONCLUSION AND FUTURE WORK

The ONS (Optimistic Nepotism Samples) is proposed to the wide range of information retrieval in the form of data samples as they apply for data ware housing. The proposed work integrates to nay business purpose and enterprise architecture of data ware house solutions. The algorithm constitutes the security measures by the randomized samples, in addition to being a critical factor and a challenging factor. The independence of the algorithm is preserved to a very wide variety of customer and business solutions are possible. In the ONS structure there may be chance of time delay at reservoir level so the work needs to be done to decrease the time delay for to match the information delivery to the user as early as possible.

Many of the Top level managements generally may prefer web based and report based information display. While data ware housing currently provides data based and derived information. In future the author has to improve the ONS in a efficient retrieval of data samples in the web based and report generated information. The time delay may be the constraint for the algorithm which should be improved. The out sourcing data ware housing is the major problem which requires to have the ability to monitor the data quality of the ware house by the outsourcer as well as the client.

REFERENCES

[1] Neural Networks algorithms, applications and programming techniques by James A. freeman and David M.Skapura pp 248-259 1991

[2] 27 Aug 2008 ... Freeman - Neural Networks Algorithms Applications Programming Techniques torrent FREE Download

[3] Data Mining Concepts and Techniques second edition by Jiawei Han and Micheline Kamber pp 468-474

[4] 6 Jul 2007 ... Download Free eBook: Data Mining: Concepts and Techniques

[5] Data Mining: Concepts and Techniques, 2 Edition ... August 26, 2010.

[6] "iSAX 2.0: Indexing and Mining One Billion Time Series," icdm, pp.58-67, 2010 IEEE International Conference on Data Mining, 2010

[7] In Proceedings of the 12th ACM SIGKDD in the International conference on Knowledge discovery and data mining (KDD '06). ACM, New York, NY, USA, 748-753.

[8] M.J.Zaki Efficient Enumeration of frequent sequences. In proc 7th int. conf information and knowledge Management 98 pages 68-75 washing ton DC, nov 1998.

[9] Fabio Aioli, Ricardo Cardin, Fabrizio Sebastiani, Alessandro Sperduti,"Preferential Text Classification: Learning Algorithms and evaluation measures", Springer – Inf Retrieval 2009.

AUTHORS

Prof.P.V.SarathChand
B.E (C.S.E),M.Tech(CS)
Pursuing PhD (CSE)
He is having 14 years of Teaching Experience. He is working as a Professor and Head of I.T Dept for Indur Institute of Engineering and Technology,Siddipet, Medak, Dist.

Dr.M.BalRaju
BE (E.C.E)
M.Tech(CSE),PhD(CSE)
and have 19 years of Teaching Experience. He is working as a Principal for Madhira Educational society group of Institutions (integrated campus), Hyderabad.

Mr. Bhukya Shankar Nayak
B.Tech(C.S.E) from JNTUniversity Hyderabad and M.Tech (IM&SC) from IIT Madras, Chennai. Presently working as a Assoc. Prof & Head of the department of Computer Science & Engineering at Madhira Educational society group of Institutions. He is having 7 years of teaching.

P.Rambabu
B.Tech(cse)
M.Tech(se) from JNTUH and working as Assistant Professor in Nimra Institute of Engineering & Technology, Ongole , Andhra Pradesh, India.

G.Suryaprakash
M.Sc., M.Tech.(CSE), working as an Associate professor in the Dept. of IT in ACE Engineering College, Ghatkesar,Andhra Pradesh, India. He is having 17 years of teaching experience.