

DATA BROADCASTING WITH TRUTH FINDING ALGORITHM

¹Rajeev Ranjan Kumar Tripathi, ²Ruby Pandey, ³Rohit Sharma

³*Rohit Sharma, Student, Department of CSE, Buddha Institute of Technology
GIDA, Gorakhpur, (India)*

ABSTRACT

We have digitized the data and frequently we are using these data to take decisions. We have data which are tightly coupled with time, time changes data changes. Such data is known as temporal data. Temporal data is very dynamic in nature. On the other hand we have some data that does not change its value with time. We refer those as non temporal data. Non temporal data is static in nature. Data broadcasting ensures availability of data to a person or a group of person. Sometimes a person is interested in a particular type of data only; it is the responsibility of the data provider to provide only data of interest to that person. This type of broadcasting is known as pull based. There is one more type of data delivery scheme where data is provided to the users whenever any changes are noticed by the provider without considering the interest of the person. This type of broadcasting is known as push based. In data broadcasting we have some data enjoying very high priority and its broadcasting is required very rapidly.

Keywords: Temporal Data, NonTemporal Data, Data Broadcasting, Data of Interest, Push Based, Pull Based

I. INTRODUCTION

Raw facts are known as data. When data is refined we get information. Conversion of data into information is a cyclic process. User changes notion of data and information changes. In this paper we will treat data and information as interchangeable words. Information can be represented into following data types: text, audio, image, graphics and animation. Value of information depends on the followings: source, user and time. Basic principle of communication says that “delivery of message from a source to a targeted receiver must be done in time without any loss of integrity of message”. These facts are always maintained. However there is a contradiction that “all facts that are available in public domain cannot be true”. Several replicas of information are available at different vendors and it is a challenge to find the correct one. For example if someone asks for the length of the river Ganga, different sites are providing different information. Truth discovery is a process

which is often used to find a correct version of information. In most of the cases it is assumed that the information having very high degree of occurrence is correct. While discussing the data broadcasting we cannot ignore all these things. Conventionally we have used followings broadcasting mediums: Print media as news paper and electronic media mainly radio stations and TV stations. In this era of mobile computing where we are browsing information while on move. Fundamentally mobile computing have major challenges like: bandwidth, connectivity and battery. Last challenge is not a matter of great concern now a days but the first two cannot be ignored if we are talking about India where more than one billion subscribers are there. Next section of this paper is shedding light on a model which deals with data delivery based on interest of user.

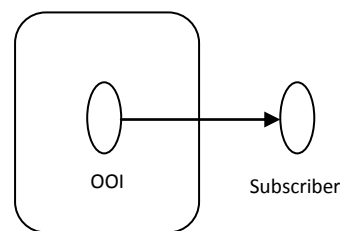


Fig1. Event Notification Model

The OOI (Object of Interest) is the object in which subscriber has interest, any change taking place in this object must be notified to the subscriber immediately. In the above event notification model, responsibility of OOI is tremendously increased because there may be many subscribers which have interest in the same object. To introduce the decoupling in between subscriber and OOI an observer is introduced. This observer observes the OOI any when any change is noticed, it is acknowledged to the subscriber. Publisher is an object that declares that it will publish a notification when a particular type of event occurs at OOI. For an OOI, there may be n types of events and for each type, there is a separate publisher. Observer receives the output of publisher as an input. Further observer may notify a type of change to a subscriber for which subscriber has subscribed or notify all the changes. The first one is called pull based and the second is called push based. Whenever a subscriber wants to unsubscribe for these services it sends a un subscription request to the observer, observer deletes the records of subscriber from its list. Next section of this paper is shedding light on data broadcast model.

II. GENERIC DATA BROADCAST MODEL

Here we are considering wireless environment so that user can browse information anytime, anywhere even on move. In this mobile environment, we cannot ensure the connectivity of a subscriber with service provider all the time. The data broadcasting system is continuously broadcasting the data and because of link failure, subscriber may miss the data. To fix this type of problem, retransmission is required periodically. A high priority data means data which is being transmitted again and again. Data broadcasting model is being represented in the figure 2. We have a disk containing N records. This disk is revolving from an angle 0^0 to 360^0 . Let this disk completes its one rotation in T time and we are considering that there is no latency that is as soon as data is read by the head, it is broadcasted. Then t , the time requirement in broadcasting a record is given by.

$$t = \frac{T}{N} \quad (1)$$

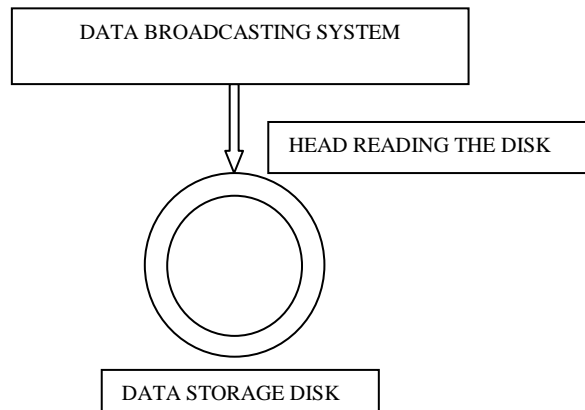


Fig2.Generic Data Broadcast Model

If a subscriber misses a data, after t same data will be available to him. This is a generic model for data broadcasting where every data is enjoying same priority. In real life data is assigned priority. The next section of this paper is shedding light on cyclic repetition of data broadcast model which is being derived from the generic data broadcast model.

III. CYCLIC REPITION OF DATA BROADCAST MODEL

In this variant, n disks are maintained with varying rotational speed. Disk₁ has highest rotational speed. As we move from disk 1 to n , rotational speed gradually decreases.

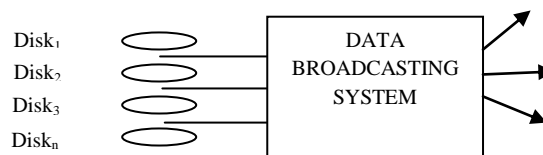


Fig3.Generic Data Broadcast Model

High temporal data is kept on first disk. Static data is kept on n^{th} disk. High temporal data is enjoying the highest priority. Again we have assumed that there is no latency. On i^{th} disk, every record is enjoying same priority. Disks in upper side rotate faster than the disks in lower side. Hence records located in upper disks are transmitted more frequently than the records of lower disks. In this way high temporal data is broadcasted very rapidly in comparison to the lower temporal data or static data. Instead of maintaining many disks, generic model can be modified in three ways. Next section is describing these variants.

IV. CYCLIC REPITION OF DATA BROADCAST MODEL

The generic model can be modified if the designer wants to use single disk instead of several disks. Here designer can assign priority to data on the basis of broadcasting frequency.

4.1. Flat Disk Model

In flat disk model, each block of record is pushed after equal interval of time. This model assumes that all the records are enjoying same priority.

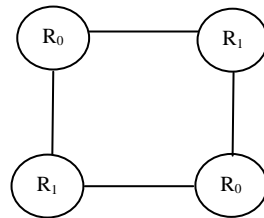


Fig4.Flat Disk Model

Here, we have considered only two records, R_0 and R_1 . They are located at constant distance in time domain on the disk. In a rotation, each record is broadcasted twice.

4.2. Multi Disk Model

In this model records are kept at equal distances in the time domain. Record having high priority is inserted more frequently than others. Let we have three records, R_0 , R_1 and R_2 . R_0 has the highest priority, then R_1 and then R_2 . R_0 is inserted on the disk more times than R_1 and R_2 .

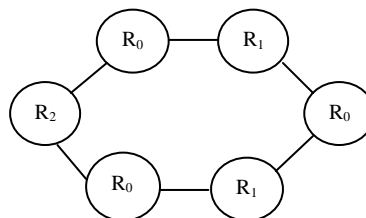


Fig5.Multi Disk Model

Here, R_0 is broadcasted thrice, R_1 is broadcasted twice and R_2 is broadcasted only once. It is to be noted that consecutive broadcasting is prohibited here.

4.3. Skewed Disk Model

It is same as multi disk model with a variation that consecutive broadcasting of a record takes place. If R_0 is broadcasted thrice continuously, R_1 is broadcasted twice continuously and R_2 is broadcasted only once the multi disk model is treated as skewed disk model.

V. CLASSIFICATION OF DATA DELIVERY MECHANISM

We can further classify data-delivery mechanism in three different ways:

5.1. Push Based Mechanism

In this mode, if any information is updated, it is broadcasted to the subscriber without keeping the interest of subscriber.

5.2. Pull Based Mechanism

In this mode, if any information is updated, it is broadcasted to the subscriber by keeping the interest of subscriber.

5.3. Hybrid Based Mechanism

It is a mixed mode operation of both push based and pull based.

VI. TRUTH FINDING ALGORITHM IN DATA BROADCASTING

As stated in the introduction section that available data is not always consistent. If data broadcasting is dependent on the data which is provided by the third party, its integrity must be cross checked. This paper is recommending to use truth finding algorithms in data broadcasting system so that data provided by the data broadcasting system can be used by subscriber in decision takings.

VII. CONCLUSIONS

This paper is an attempt to summarize the data broadcasting techniques. Highly available data cannot be considered consistent. Truth finding algorithms must be used by the system if data broadcast system is using data from other vendors. Highly temporal data needs to be broadcasted very frequently. This paper is emphasizing on the fact that source of information should be trusted so that decisions may be taken without any ambiguity. In wireless communication, connectivity cannot be ensured thus in this scenario cyclic repetition is necessary.

REFERENCES

- [1] Juhn Tseng, "Flat data broadcasting and receiving scheme for popular video services," IEEE, Vol.44 Issue-1, March 1998.
- [2] Huang and Chen, "Dependent Data Broadcasting for Unordered Queries in a Multiple Channel Mobile Environment," IEEE Transaction on Knowledge and Data Engineering, Vol.16, No-1, Feb-2003.
- [3] Pang and Chan, "Efficient Channel Allocation Tree Generation for Data Broadcasting in Mobile Computing Environment," Wireless Networks 9,117-129,2003.
- [4] Chen, Wu and Yu, "Optimizing Index Index Allocation for sequential Data Broadcasting in Wireless Mobile Computing," Wireless Networks 9,117-129,2003.
- [5] Dalia Attia Waguih, Laure Berti-Equille, "Truth Discovery Algorithms: An Experimental Evaluation," Qatar Computing Research Institute, Submitted on 23 Sep 2014
- [6] Raj Kamal, Mobile Computing, Oxford Higher Education, 2007.
- [7] R.R.K. Tripathi, Mobile Computing, Dhanpat Rai, 2nd Edition, 2014.