PAPER • **OPEN ACCESS**

Files synchronization from a large number of insertions and deletions

To cite this article: Vijayan Ellappan and Savera Kumari 2017 *IOP Conf. Ser.: Mater. Sci. Eng.* **263** 042028

IOP Conf. Series: Materials Science and Engineering **263** (2017) 042028 doi:10.1088/1757-899X/263/4/042028

Files synchronization from a large number of insertions and deletions

Vijayan Ellappan and Savera Kumari

School of Information Technology Engineering, VIT University, Vellore – 632014, Tamilnadu, India

E-mail: evijayan@vit.ac.in

Abstract: Synchronization between different versions of files is becoming a major issue that most of the applications are facing . To make the applications more efficient a economical algorithm is developed from the previously used algorithm of "File Loading Algorithm". I am extending this algorithm in three ways: First, dealing with non-binary files, Second backup is generated for uploaded files and lastly each files are synchronized with insertions and deletions. User can reconstruct file from the former file with minimizing the error and also provides interactive communication by eliminating the frequency without any disturbance. The drawback of previous system is overcome by using synchronization , in which multiple copies of each file/record is created and stored in backup database and is efficiently restored in case of any unwanted deletion or loss of data. That is, to introduce a protocol that user B may use to reconstruct file X from file Y with suitably low probability of error. Synchronization algorithms find numerous areas of use, including data storage, file sharing, source code control systems, and cloud applications. For example, cloud storage services such as Drop box synchronize between local copies and cloud backups each time users make changes to local versions. Similarly, synchronization tools are necessary in mobile devices. Specialized synchronization algorithms are used for video and sound editing. Synchronization tools are also capable of performing data duplication.

1. Introduction

This project deals with number of insertions and deletions. Customers can communicate with each other and at the same time the communication bandwidth is also reduced to achieve efficient communication channel between the customers. While inserting any data into the database, this will also create a backup in the database. This is because, if any file is deleted from the database, "administrator" will easily restore those data from the database. This will not only help the "administrator" to overcome the loss of data but also will help to save the confidential data. Making reasonable estimations to synchronize between different variations of records is a major downside with different applications. I tend to increase this algorithmic oversee in three ways: First , I tend to manage non-parallel archives . Second, these reports contain pictures picked in wander with non-uniform spreads. Provide trade-offs between the number of rounds of communication and the total amount of bandwidth required to synchronize the two files under various implementation choices.

2. Scope and objective

This paper deals with reducing the number of rounds of interaction and the bandwidth required to "synchronize" various files. The strongest protocol is shown under various scenarios of synchronization process. Errors can be founded while the communication between clients and the server is active . The main objective of developing Synchronization Algorithm is to overcome the disadvantages of previous algorithm and providing security to the important files which is supposed to be uploaded in the database. Using the key values the deleted files can be retrieve by the Administrator. Two- way authentication is available for security. The main objective in this is to provide tradeoffs between the number of rounds of communication .The total amount of bandwidth required to synchronize the two files under various implementation choices of the baseline algorithm.

2.1. Existing System

A document called load record is used to recover particular informational collections' or pictures situated inside lawful databases through particular recovery techniques actualized in the heap record. The procedure utilized as a part of existing framework is File Loading Algorithm .Recovery of documents were not conceivable in existing framework. Customary record framework don't have the alternative of reinforcement of information i.e. once the information is erased can't be recovered.

Technique:

- File Loading Algorithm

Disadvantages:

- File can't be retrieved.
- No option of backing up of data.
- No restoring technique.

2.2. Proposed System

In the proposed system, I am proposing "file synchronization algorithm" to overcome the disadvantages of the previous technique of storing file. In this algorithm once an information or table is erased, can be recovered by the default reinforcement duplicate of each table and information utilizing different set to control. i.e. no erase or read rights are given to the clients. Executive can just perform erase and get to operation on the database . This proposed idea will help the users to get into more secured and authenticated login. Users can store large amount of data Can retrieve the deleted files. Record and organizer metadata for every synchronization session is gathered and put away in a database. The latest record rendition is resolved in light of a mix of document qualities, size and time-stamp (not only the record alteration time). To guarantee the security of information, when there is instability as to which record is the latest, the client is provoked for affirmation. Various examples of All way Sync with access to a typical record framework can share normal metadata databases. In case of system, equipment, or programming disappointments amid synchronization, none of your information is harmed. Advanced to meet an elevated requirement of execution. Stores data about erased documents and envelopes in the metadata database. Does not depend on the precision of the record framework or PC clock. Makes and keeps up the correct organizer structure when expected to appropriately synchronize documents and envelopes. Synchronize information other than straightforward documents, including: registry keys, database records, messages, contacts, playlists, and so forth.

Technique :

- File Synchronization Algorithm

Advantages :

- Deleted files can be retrieved
- Default backup is created .
- No loss of confidential data

3. Problem Statement

In the case when the data is stored in cloud, the main problem persist while downloading the whole information from the cloud and checking the information which will cost more than the client's expectation. This procedure is not feasible and also it is not cost-efficient. Another problem of storing data in cloud is that, it may be lost due to some programming inefficiency and by human mistakes. The way toward investigating the standards or methodology for auditing procedure of imparted information out in the open cloud to effective client repudiation. Seek to minimize the overall communication bandwidth, measured by the total amount of information exchanged by the two users, while simultaneously maintaining a low probability of reconstruction error. Such algorithms are refer as synchronization algorithms. In the case when the useful and confidential data are deleted from the database, it is very difficult to retrieve those data and user data is lost permanently. The efficient solution is to use synchronization algorithm and creating backup.

4. Modules

4.1. User Interface Design
4.2. Synchronization
4.2. Insertion
4.3. Deletion
4.4. Back-Up

4.1. User Interface Design

In order to interact with the application, user must be a registered user of the application. To register themselves with the application, user must have to fill the registration form with some personal information's. Server will create the account for the entire user to maintain upload and download rate. Name will be set as user id. Logging in is usually used to enter a specific page. To interaction with server client must give their username and secret key then no one but they can ready to associate the server. In the event that the client as of now exits specifically can login into the server else client must enlist their points of interest, for instance, username, watchword and Email id, into the server/ and it will make the record for whole client to keep up transfer and download rate. Name will be set as client id. Signing in is generally used to enter a particular page.

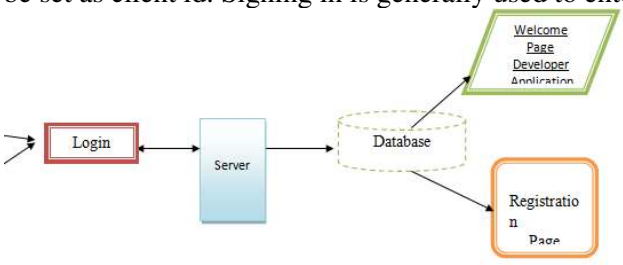

Fig.1User Interface Design

4.2. Synchronization

In this module the synchronization process takes place where the lost files can be retrieved by the help of backup process. Erased files can be restore by browsing the files from the location where the deleted files are store.

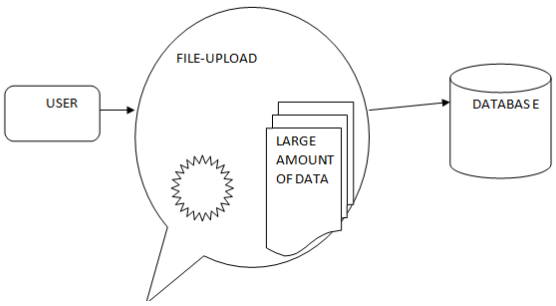

Fig. 2: Synchronization

4.3. Insertion:

In this module data is stored to database. I have differing sorts of table numbering from one to five. Table and field both are having same number .So that enormous measure of data can store to particular field . In this process signing into the application, clients has numerous alternatives. One of the alternatives: is inclusion. In this, clients can transfer their information in unthinkable arrangement into the database.

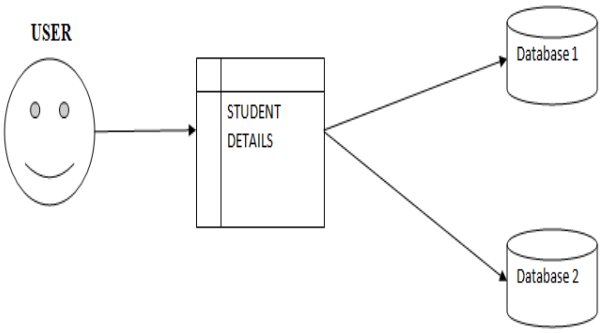

Fig 3:Insert data to database

4.4. Deletion:

The particular data and table as well as its contents can be deleted . Just need to select the desired table that need to be delete and can perform deletion on them . There are various ways by which data can be erase i.e. using drop table, delete command .This process is only performed by the administrator, no other users has the permission to delete any sort of data or can alter any table or data. Two database is available one for insertion and another database is for deletion purpose so that at the time the hacker or unauthorized user hacks the database that time the files can be retrieved by the administrator by the backup process.

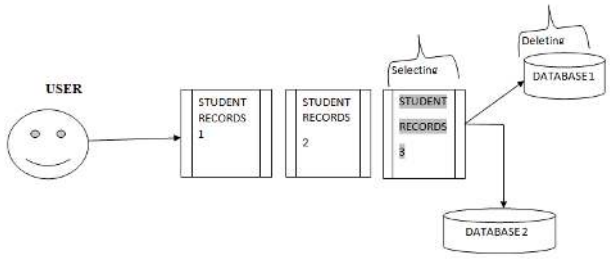

Fig. 4: Delete data from database

4.5. Back-Up:

Expansive measure of information put away database then put away copy ,then transfer information to database that time store additionally duplicate database. If obscure individual erasing all records from database. It will recover the document from copy database. By the use of "synchronization algorithm", multiple copies of files data is created in the database which can be used as backup for future reference if any of the table or data is deleted from the database and we are in need of those data.

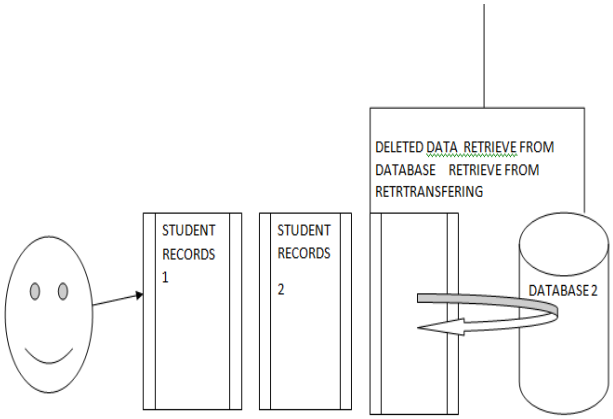

Fig. 5:Deleted data backup

5. System Architecture

.

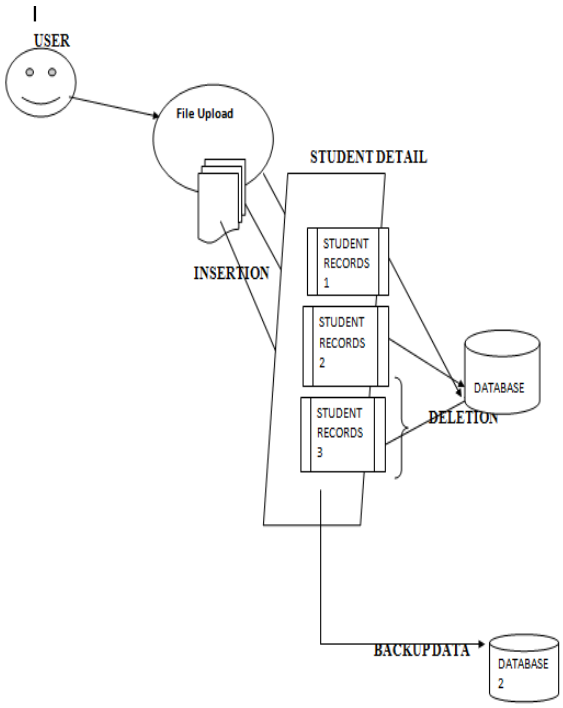

Fig. 6 System Architecture

In this paper, I developed a file synchronization protocol for efficient file handling process. In the system architecture I want to show how the entire system works. First of all, user logs in to the account. After logging in to the account user will upload file by creating table into the database. User will find the option as "create table" where the user will create a table by specifying the number of fields user require (Type of each fields are predefined). After creating the table user finds the option to insert data into the fields as "insert data". The desired table with the data is stored in the database for that particular user and this database is particularly handled by the administrator. By this process, upload of file is complete. After the upload of data, there is a option called "Backup". By choosing backup option, a copy of the desired table with its data is created on to another database which is inaccessible to the users other than the administrator. If in case any of the table is deleted accidentally, an option called "Synchronize" is provided in the front end, where administrator will upload the backed up file from the backup database and be able to synchronize the file without hassle. This will restore the deleted data table from the database and store that particular data table back to the database. The architecture defines the complete function of the work . Defines the step by step process.All the modules described above is combined in the system architecture. The user first have to be the registered user for the security purpose only the registered user have the authentication to upload the files and insert values to the database. Administrator have the full authority to see the uploaded files and can retrieve the deleted files also by the backup process which in terms known as synchronization process

5.1. Methodology

The process of analyzing the principles or procedure for Auditing process of shared data in public cloud with efficient user revocation. Break down the issue of identifying mischievous activities in view of the framework exhibitions. I ought to keep away from by utilizing decent amount finder. The main method that is used in the synchronization algorithm is backup process. The data which is deleted intentionally or intentionally may be retrieve back by using key retrieval value from the another database in which the replicated copies are present.

6. Results and discussions

After all the development of the application is done, the main focus of developers as well as users is, What outcome the application is providing? The outcome of each modules listed above must satisfy the expected result of the developer and it must satisfy the requirement given by the client.

Advantages:

- Large amount of data can be stored
- All the deleted files can be retrieve .
- File can be Synchronized

6.1. Expected result

- **User Interface Design**

Input : Username and Password.
Result : Display user home page.

- **File upload**

Input : File or Data
Result : File uploaded.

- **Insert**

Input: Insert data to database
Result : data inserted

- **Deletion**

Input: Delete data from database.
Result : Data deleted

- **Backup**

Input : Type deleted data
Result : Deleted data backup successfully.

6.2. Significant Findings

I watched the effect of the expansion of a halting rule. Ceasing criteria enhance the execution of convention by giving trade-offs amongst transfer speed and adjusts of correspondence, additionally by ensuring that anomaly cases don't wind up devouring an expansive number of rounds. Something critical to notice is that the data transfer capacity utilized with the genie-supported ceasing rule is a noteworthy change contrasted with straightforward criteria: creating heuristic halting criteria that

endeavour to approach what the genie would do may in this way enhance the convention considerably further.

7. Future enhancement

The bandwidth for the current solution is decreased by roughly 40%. In future the bandwidth can be decreased more. The search algorithm can be improved with heuristic-search for efficient search in less time. More security can be provided to the files available in database. For reducing time complexity data structure can be use for fast retrieval of files and data.

8. Conclusion

An important thing to notice is that the bandwidth used with the genie aided stopping criterion is a major improvement compared to simple criteria .By developing this project, I conclude that, the rounds of communication for the client and server interaction had considerably decreased. In order to build an efficient communication channel, the bandwidth of the communication is also low for higher efficiency. Apart from communication there is a significance of insertion and deletion in this project. I conclude that it "synchronizes" the file as well as records through which multiple copies are created in the local server. These multiple copies are termed as backup. Whenever there is unwanted deletion occurring with the file-records, Administrator can restore those records from the backup that was previously created. Full authorization is provided to the administrator so that it reduce the error and unintentional use of important file. Using key value the deleted data can be retrieve and by the mean time the bandwidth between the users is also reduced by minimizing the low probability of error. So that the user may communicate back and forth.

References

[1]. G. Cong, C. S. Jensen, and D. Wu. "Efficient retrieval of the top-k most relevant spatial web objects". PVLDB, 2(1):337–348,Year: 2009

[2]. Lara Dolecek ,Venkat Anantharame , "on communication over channels With varying Sampling Rate", Year:2014

[3]. Tuan A. Le and Hue D. Nguyen, "New Multiple Insertions-Deletions Correcting Codes for Non-Binary Alphabets", Year: 2015

[4]. Fahaad Shaon and Murat Kantarcioglu ,"A Practical Framework for Executing Complex Queries over Encrypted Multimedia Data", 233-286, Year: 2013.

[5]. Uri Pushpa , Prasad Vinod, Jilani Abdul Khader, "Insertion and Deletion on Binary Search Tree using Modified Insert Delete Pair: An Empirical Study",(IJCSNS) ,VOL.7 No.12, December 2007.

[6]. S. M. Sadegh Tabatabaei Yazdi and Lara Dolecek, IEEE Senior Member, "A Deterministic Polynomial-Time Protocol for Synchronizing from Deletion", 288–292, Year: 2013

[7]. Simone Ercoli, Marco Bertini, Alberto Del Bimbo, "Compact Hash Codes for Efficient Visual Descriptors Retrieval in Large Scale Databases", IEEE Transactions on Multimedia, Volume: PP, Issue: 99,Year: 2017

[8]. Naw Safrin Sattar , Tasmia Aqila, Rifat Shahriyar ,"Towards concurrent data structure development with Relaxed Synchronization", 9th(ICECE) Pages: 267 - 270, Year: 20l6.

[9]. Gilad Baruch; Shmuel T. Kleine ;Dana Shapira., "A Space Efficient Direct Access Data Structure",(DCC), Pages: 63-72, Year:2016.
[10].Dongming Chen,Xiaodonge Chen and Zhenhua Tan; Yu Lu; Jing Wang, "Research on file synchronization backup system based on security protocol", IEEE 3rd (ICCSN), Pages: 246 -249, Year: 2011.
[11]. T. Richardson and R. Urbanke, " Efficient encoding of low-density paritycheck codes," IEEE Transactions on Information Theory, vol. 47, no. 2,
pp. 638–656, Feb. 2001.
[12]. R. Venkataramanan , H. Zhang, and K. Ramchandran, "Interactive lowcomplexity codes for synchronization from deletions and insertions," in Proc. 48th Annual Allerton Conf. on Comm., Control, and Computing, Year: 2010.

[13]. C. Schoeny, "Efficient file synchronization," M.S. thesis, Dept. Elect.Eng., Univ. California, Los Angeles, Los Angeles, CA, USA, 2014.