Journal of Theoretical and Applied Information Technology
10th April 2016. Vol.86. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# A SURVEY ON DATABASE SYNCHRONIZATION ALGORITHMS FOR MOBILE DEVICE

[1]MOPATI B. KEKGATHETSE, [2]KELETSO J. LETSHOLO

[1] Research Scholar, Department of Computer Science & Software Engineering, Botswana International University of Science & Technology, PALAPYE BOTSWANA

[2] Lecturer, Department of Computer Science & Software Engineering, Botswana International University of Science & Technology, PALAPYE BOTSWANA

E-mail: [1]mopati.kekgathetse@studentmail.biust.ac.bw, [2]letsholok@biust.ac.bw

**ABSTRACT**

A lot of research has been done under the subject of data synchronization between a mobile database and a server-side database. To the best of our knowledge, little or no publications exist that review research problems faced under this field. In this paper, we explore the current research challenges in data synchronization between a mobile database and a server-side database. A systematic review method was employed in gathering the literature from reputable conference proceedings and electronic journals. Specifically, this paper will guide future researchers in the field to propose new solutions for the existing problems in data synchronization of mobile database.

**Keywords:** *Mobile Device, Synchronization Algorithm, Mobile Databases, Data Synchronization, Systematic Literature Review*

## 1. INTRODUCTION

Advances in technology have led to a new form of computing environment based on small, mobile devices. These devices come equipped with a lightweight database built unto them. Due to this property, processing of business information can now be achieved and as a result, more and more applications that depend on mobility of devices have emerged.

However, the mobile devices have limited processing power, rely on a finite battery source and have inadequate memory available. Moreover, they constantly access the network and this is a challenge because of limited bandwidth available to them. More often, mobile devices are occasionally disconnected from the network, this can be caused by a number of reasons; complete, unpredicted network cut, mobile data connectivity lapse or battery running low and contingency measures having to be executed as a result.

Synchronization algorithms are procedures or methods adopted to facilitate data exchange between two or more entities. Synchronization process brings data to a consistent state across all entities involved. In the context of this paper, the entities in question refer to a mobile database client and a server-side database. Research in the field of mobile databases synchronization has resulted in a number of solutions being developed. The variations in these can be attributed to different design goals and specific problems that these solutions were initially designed to address. Because mobile devices run on limited resources, it is inevitable that mobile data synchronization must account for this crucial requirement.

In this paper, we analyze existing data synchronization algorithms for mobile database between a server-side database and a mobile database. Our goal is to uncover research gaps and challenges inherent to them. This paper will guide future researchers in the field because it will aid in proposing new solutions for the existing open gaps in data synchronization of mobile database.

Journal of Theoretical and Applied Information Technology
10th April 2016. Vol.86. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The remaining part of paper is organized as follows. Section 2 describes the systematic review process applied when gathering the literature, along with the search sources and criteria used during the review process. Section 3 discusses existing literature on the subject, Section 4 describes the critical evaluation of the literature gathered based on our review criteria. Section 5 presents the open areas for further research. Finally, conclusions and directions for future work are discussed in the last section 6.

## 2. SYSTEMATIC REVIEW PROCESS

The systematic review's purpose is to build a foundation on which existing work can be compared. This review follows the guidelines for performing systematic literature reviews in presented by Kitchenham and Charters [1]. First, we discuss our search strategy, for retrieving the research papers reviewed. .Second, we describe the search string used when searching for research papers. We then look at inclusion and exclusion criteria because we cannot review all the results yielded by the search string. Finally, we present our review criteria.

### 2.1 Search Strategy

Our sources include published work ranging from year 2002 to September 2015, from the contents of IEEE Explore. Peer-reviewed journals in computer science and engineering have also been included. They are International Journal of Computer Science and Mobile Applications; International Research Journal of Engineering and Technology; International Journal of Engineering, Science and Innovative Technology; Studies from reputable proceedings of IEEE, such as, International Conference on Computing, Electrical and Electronic Engineering; International Conference on Computing Electronics and Electrical Technologies were also among our sources.

### 2.2 Search String

Searches for electronic resources are usually initiated by a combination of keywords. This string was based on keywords acquired from the main title of survey paper. The string was constructed by carefully selecting words from the title. The string used was:

{Mobile *database synchronization OR mobile data Synchronization}.*

### 2.3 Inclusion and Exclusion Criteria

The defined string has resulted in almost a thousand of papers returned. Our inclusion criteria was as follows:

1. Published papers in the year 2002 to September 2015.

2. Sometimes more than one paper describing a similar approach, published in different venues, is returned. In this case, we opt for the most recent version;

3. Papers describing different parts of a single approach, published over several papers or in different venues.

### 2.4 Review Criteria

The review criteria help to organize information from identified papers in a precise and clear manner. Our criteria are based on four properties: Security, Performance speed, Energy consumption and Memory use of mobile device.

1. *Security* – Is concerned with how secure the data being synchronized is.

2. *Performance speed* – the time-taken to synchronize data to server database.

3. *Energy consumption* – the rate of loss of battery power during synchronization.

4. *Memory usage* - Amount of memory consumed by the device during the synchronization operation.

## 3. REVIEWED APPROACHES

The research on mobile data and synchronization has been continuous since early 2000s. We have classified the approaches based on their common features. These include synchronization protocols, extensible markup language (XML) based synchronization models and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

synchronization algorithms based on message digests (SAMD). Data synchronization algorithms are implemented to facilitate transfer of photos, files, databases, videos and folders between a personal computer and mobile device. This paper focuses only on data synchronization of mobile database between a server and a mobile device.

### 3.1 Synchronization Algorithms based on Message Digests (SAMD)

Synchronization Algorithms based on Message Digests (SAMD) have been proposed and implemented in the works of [2],[[3],[[4],[5]. These algorithms were designed to facilitate data synchronization between a server side database and a mobile database. Two message digest tables are created, one at the server and the other at the mobile device/client, based on data tables. The data tables are the ones holding the actual business data. The message digest is a hash function computed to detect rows needed for synchronization, by comparing the hash values. If the values are same there is no need for synchronization, if they are different then an inconsistency exists and synchronization is necessary and progresses according to the rules of the algorithm.

These algorithms are independent of database venders, they do not use database objects such as triggers, stored procedures or timestamps. Rather they depend on standard Structured Query Language (SQL) functions for effecting synchronization procedure. SAMD algorithms are also somehow secure due to the use of a message digest. Because of the independence of database vender, they can be implemented in any combinations of server-side and mobile database. These are the disadvantages that exist in these algorithms. One downside with these algorithms is that, although they use hash functions, they do not guarantee security during transmission of data to the server because these hash values are residing in a database table at both ends. Their main purpose is to establish any data inconsistencies that may exist. Another disadvantage is that the computation of message digest is resource intensive.

### 3.2 Generic Data Synchronization Approaches

Sethia et al [6] proposed Mobile Replicated Database Management Synchronization (MRDMS). This system performs synchronization based on timestamps. These are implemented at field level rather than rows. That is, each table has its corresponding timestamp table with identical schema, maintaining timestamps for each cell. Only the modified rows are synchronized. Again, maintaining cell based timestamps instead of row based reduces chances of having a conflict because with row based timestamps unique cells may not have been altered, but the overall timestamp of each row nonetheless, may be changed. Problem with MRDMS is that it leads to multiple tables as systems grow large in terms of data requirements, assumes initial synchronization of times which might not always be the case and lastly, database insertions are only made when the device is online. This solution operates at application layer and again data security is not addressed.

A cross layer method for mobile database synchronization was implemented by Jiao et al [7]. It deals with synchronization at the transport layer rather than the application layer. The authors highlight that most systems deal with synchronization at application layer but this layer does not know information about the underlying connections. Traditional Transmission Control Protocol/Internet Protocol (TCP/IP) protocol was substituted with TCP Westwood better suited to wireless networks, because it reduces waste of communication resources due to rollback of transactions in wireless network. Touchsync has no security guarantee and is still possible to intercept the communication link. The authors [7] again have not made it clear whether the method works at same performance speed under mobile connectivity and Wi-Fi standard.

Object-oriented data synchronization (OODS) for mobile databases in mobile ad-hoc networks (MANET) was studied by Li et al [8]. With this solution, everything is observed as an object and the notion of Subscriber/Publisher is adopted. Subscribers can recognize publishing schema updates automatically and regarding Publishers, conflict resolutions and publishing schema are user defined or inherited from the schema inheritance tree. This is how data gets synchronized from publisher to subscriber. OODS provides a comprehensive solution for applications that operate on MANET but the many mobile applications operate on a non-MANET network. OODS takes advantage of both application and transport layer. Once more under this solution, evidence of data security during synchronization is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

not present.

Malhotra and Chaudhary [9] proposed an algorithm to solve the problem when all clients are relying on a single server database. If there is planned downtime or a failure arises remote workers will operate on the local storage and when connection resumes, data is sync from client system to server in serial order. The procedure is the same on file handling. If system is disconnected all files (images) uploaded by user are saved on client machine folder and when there is connection, automatically the images are transferred from client to server. This is a simple solution but it does not consider for what might happen to data during the transfer of files in online mode in case all device switches to offline, because mobile devices operate on low bandwidth, wireless networks that are not reliable. Also, this data might be of a sensitive nature and therefore might require the features of confidentiality and non-repudiation.

Synchronization between Oracle database and database inside simple mobile application was conducted by Zechmeister et al [10]. The aim is to realize concrete data transfer using extensible Markup Language (XML) documents over mobile network. Realized solution is effective and secure and XML documents are easy and simple to use. The solution can be used in similar implementing applications which use oracle database. However, XML utilizes long unnecessary tags that consume more bits, although the documents are compressed to reduce size, the decompression at mobile device consumes time. There is no flexibility offered to the application developer as they are restricted to Oracle Database Management Solution (DBMS) to implement this technique. If for example application developer is only competent in MySQL (Structured Query Language) DBMS they are at a disadvantage. In other words, the solution is proprietary and uses database dependent information specific to Oracle. Yet again application layer is the foundation for this synchronization.

### 3.3 XML Based Synchronization

Synchronization Mark-up Language (SyncML) was proposed by Lee et al [11] and it suggests a new standard for data synchronization across devices. This is a protocol which depends on XML to carry messages across network. The work was further enhanced by Li & Li [12] by introducing Huffman coding to compress data in order to improve synchronization performance. SyncML entirely depends on XML tags that wrap synchronization message, this leads to long unnecessary files that are too expensive to carry across the wireless network.

An XML based mobile data acquisition system was developed by Huang et al [13]. The mobile data acquisition system has been implemented to change the traditional way of paper-based data collection. It is based on the things that are common to different data acquisition systems under various fields like geography, agriculture, transportation. Through simple customization, this solution can meet the demands of different users under these fields. The system comprises of a data synchronization module which imports and exports data between the background database and SQLite database. First, users have to specify the requirements for their data acquisition needs. This information is then saved into an XML file, which is then copied to the device along with data dictionary, after the synchronization module is loaded. This system does not account for the security of data being acquired and shared across devices. There is a high risk of man-in-the-middle attacks and this may compromise the integrity and accuracy of data.

By using XML, a synchronization technique for mobile databases called DefferedSync, was developed by Miller et al [14]. DefferedSync transforms relational database into an XML tree structure and then makes use of deferred views in order to minimize bandwidth and storage space on mobile client. Using a view allows the client to retrieve only the data relevant to them, instead of the entire database. DefferedSync increases data transmission load because it uses XML, which is verbose by nature. A method of data synchronization based on XML in distributed heterogeneous environments was developed by Guo [15]. The method uses XML technology to realize data synchronization between different databases using a number of procedures; 1) Conversion of incremental data into XML file via mapping layer; 2) send XML file to destination in a message format. In contrast, this method uses XML to transfer plain messages, no data security is guaranteed.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

## 4. CRITICAL EVALUATION

In light of the literature review in the previous section, next, we analyze and evaluate the works presented based on our review criteria presented in section 2. A summary of this evaluation is presented in Table 1.

### 4.1 Security

According to Ghorbanzadeh et al [16] the subject of security of mobile database should be addressed under four fields' namely mobile device security, security of operating system on mobile device, security of mobile database and security of mobile network. Reviewed literature in the previous section is not explicit about the subject of security and where it fits into the above categories. SAMD algorithms however are somewhat secure due to message digest computation.

Since most of the literature on data synchronization of mobile database is not explicit about data security, it is safe to assume that most authors are relying on Transport Layer Security (TLS). This is a challenge because if the trust between the server and the client breaks down, maybe due to the Certification Authority (CA) key management lapse, this will become a big crisis therefore there is need for additional security on top of TLS.

### 4.2 Performance Speed

Performance speed in this context is the time taken for database synchronization. Saravanan [5] reviewed various synchronization algorithms and one of the drawbacks of existing system was synchronization is very slow. It was not very clear how the experiment was conducted. However, Taha and Alim's work [4] evaluated synchronization time of SAMD and SWAMD (Synchronization Wireless Algorithms based on Message Digests). The SAMD in this case is wired connectivity while the SWAMD is wireless connectivity. They used data of same type and results showed that SWAMD is slower than SAMD by 0.5 seconds.

In MRDMS [6], a matrix is used to send only the cells required instead of the entire set of row. An overhead of the matrix is added on top of data sent. The authors argue that this actually seems like bandwidth waste but in execution it saves both bandwidth and time consumed. The cross-layer method [7] has proved that indeed saving bandwidth also reduces synchronization time. The authors used TCP Westwood to decrease the times that synchronization transactions rollback and this also reduced bandwidth.

OODS [8] system was deployed over a Bluetooth Piconet. The results of three different synchronization modes revealed that OODS system has satisfying synchronization efficiency against Merge Replication (MR) technology, with an approximate synchronization delay difference of 6 seconds in average. In this work OODS was only compared to MR, further work is still needed to compare OODS with other similar technologies. In Malhotra and Chaudhary's work [9] and Zechmeister's research [10], evidence of algorithmic performance evaluations is not present.

In SyncML [11]-[12], data synchronization of photos, SMS, address book and files between mobile device and server was done. The SyncML with Huffman coding [12] has a lower synchronous time compared to the one without Huffman coding, especially in cases where synchronization data is too large. The standard SyncML has a higher synchronous time in relation to SyncML with Huffman coding. The authors argue that DefferedSync [14] reduces transmission load but there is no supporting proofs. Also, it is not explicit about the time taken to synchronize the data.

### 4.3 Energy Consumption

The finite energy of mobile device is critical to synchronization session because if there is no power, then synchronization cannot occur. Consequently, this means to conserve available power, synchronization algorithm should also account for this crucial requirement. Here we emphasize on the rate of loss of battery power during the database synchronization process with respect to size of data. The SAMD by Saravanan [5] shows little or no proof of energy consumption experiment. Similarly, Choi et al [2] did not make any efforts to measure rate of energy loss of mobile device in their work.

The developers of MRDMS algorithm [6] only put forward a mere argument that a lower computational load doesn't guarantee a better

Journal of Theoretical and Applied Information Technology
10th April 2016. Vol.86. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

protocol but due to limited battery of mobile device, high processing load on mobile device may not be acceptable. The cross layer method by Jiao et al [7] improves synchronization performance but rate of energy loss was not measured. OODS system emphasized the effect of data quantity during synchronization process and how it affects the synchronization time. The authors did not highlight how data quantity affects energy loss of mobile device either. The object-oriented data model presented in Li et al [8] accelerates synchronizing speed, nothing is mentioned about its effect on battery usage or energy consumption of mobile device. In the paper [9], the authors did not carry out any performance evaluations involving energy consumption/ rate of battery loss of mobile device. Realization of data transfer between Oracle database and mobile application database also does not show any evaluations of battery usage present [10]. The same applies for papers [12] and [14].

### 4.4 Memory Usage

In this section, our attention is directed towards use of memory available to the mobile device during transfer of data from a mobile database to a server side database. The amount of Random Access Memory (RAM) readily available for smartphones is very low compared to that of a personal computer system. With SAMD algorithms [2], [3], [4], [5], a conclusion cannot be reached with regards to amount of memory utilized during synchronization because there is no formal procedure followed to evaluate this property. Similarly, in the papers [5]-[13] there exists no evaluations or experiments focused on measuring the amount of memory being used during synchronization process. In an effort to consolidate the works reviewed and evaluated, Table 1 provides a brief summary in one piece.

## 5. DISCUSSION

### 5.1 Open Issues and Research Gaps

We have surveyed the literature and identified the following open issues and gaps in the existing solutions of mobile database synchronization:

1. Security is less, or data is not secured during synchronization session. More research work desired to address security of data in transit between a mobile database and a server-side database.

2. Most solutions operate at the application layer. Further work on database synchronization across different protocol layers still lacks.

3. Energy consumption and memory usage of mobile device is not assessed but these devices rely on finite resources so it is imperative to measure how much energy and memory is used during the database synchronization session in future work.

4. Most solutions do not measure performance speed.

The above issues are a motivation for advancing research in the area of mobile databases and synchronization.

### 5.2 Lessons Learnt

Mobile devices are vulnerable to attacks and there is absolute need for improving their security. In addition, they operate on constrained resources and it is vital that applications that run in them respect this crucial requirement. In this paper, current issues in data synchronization between a mobile database and a server-side database were investigated based on security, performance speed, energy consumption and memory usage of mobile device. We have discovered that little to no works exist that address security of data in transit. Similarly, in most of the works reviewed performance speed and resource usage of mobile device was not taken into consideration.

From this survey, we have learnt some issues pertaining to data synchronization of mobile databases;

1. Not the entire database is sent, only necessary data is transmitted.

2. XML and Java Script Object Notation (JSON) are mainly used for transfer of data over the network.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. Inconsistencies in the data are mainly in established based on individual rows using message digests or hash functions.

## 6. CONCLUSIONS

We have summarized the current issues in synchronization of data between a server-side database and a database inside a mobile device. This paper identified current research gaps and open areas for further research. In future, our research will specifically address these open issues and propose effective solutions to the problem.

## REFERENCES:

[1] B. A. Kitchenham and S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” Version 2.3. EBSE Technical Report, 2007.

[2] M. Y. Choi, E. A. Cho, D. H. Park, J. Y.Bae, C. J. Moon, and D. K. Baik, “A synchronization algorithm of mobile database for ubiquitous computing,” in *NCM 2009 - 5th International Joint Conference on INC, IMS, and IDC*, 2009, pp. 416–419.

[3] V. Balakumar and I. Sakthidevi, “An efficient database synchronization algorithm for mobile devices based on secured message digest,” in *2012 International Conference on Computing, Electronics and Electrical Technologies, ICCEET 2012*, 2012, pp. 937–942.

[4] T. A. Alhaj and M. M. Taha, “Synchronization Wireless Algorithm Based on Message Digest (SWAMD) For Mobile Device Database,” in *International Conference on Computing, Electrical and Electronic Engineering (ICCEEE)*, 2013, pp. 259–262.

[5] B. Saravanan, J, A. Mary, and J. Jagadeesan, “A DATABASE SYNCHRONIZATION ALGORITHM FOR MOBILE DEVICES,” *Int. J. Comput. Sci. Mob. Appl.*, vol. 2, no. 4, pp. 46–58, 2014.

[6] D. Sethia, S. Mehta, A. Chwodhary, Replicated Database Management Synchronization,” in *2014 International Conference on Signal Processing and Integrated Networks*, 2014, pp. 624–631.

[7] Y. Jiao, Z. Jin, and Z. Ma, “A cross-layer method to improve mobile database synchronization performance,” in *Proceedings - 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2009*,
2009, no. 1, pp. 2–5.

[8] Y. Li, X. Zhang, and Y. Gao, “Object-oriented data synchronization for mobile database over mobile ad-hoc networks,” in *2008 International Symposium on Information Science and Engineering, ISISE 2008*, 2008, vol. 2, pp. 133–138.

[9] N. Malhotra and A. Chaudhary ”Implementation of database Synchronization Technique between Client and Server,” *Int. J. Eng. Sci. Innov. Technol.*, vol. 3, no. 4, pp. 460–465, 2014.

[10] J. Zechmeister, D. Zak, T. Vana, and J. Lebduska, “Realization of Data Transfer between Mobile Device and Oracle Database,” in *Proceedings of International Conference on Advances in Communication and Information Technology 2012*, pp. 48–50.

[11] B. Lee, T. Kim, D. Kim, and C. National, “Data synchronization protocol in mobile computing environment using S y ncML,” in *5th IEEE International Conference on High Speed Networks and Multimedia Communication (Cat. No.02EX612)*, 2002, pp. 133–137.

[12] J.-L. Li and J.-P. Li, “Data synchronization protocol in mobile computing environment using SyncML and Huffman coding,” *Wavelet Act. Media Technol. Inf. Process. (ICWAMTIP), 2012 Int. Conf.*, pp. 260–262

[13] Z. Huang, Q. Chang, Z. Shen, Y. Zhou, B.Yan, and L. Liu, “Design and Implementation of an XML-Based Universal Mobile Data Acquisition System,” *2009 Fifth Int. Conf. Semant. Knowl. Grid*, 2009, pp. 456–457.

[14] K. Miller, C. Gee, R. Inaba, T. Ozyer, A. Lo, and R. Alhajj, “Synchronization of Mobile XML Databases by Utilizing Deferred Views,” in *Information Reuse and Integration,. IRI 2004. Proceedings of the 2004 IEEE International Conference on*, 2004, pp. 186 – 191.

[15] Z. Y. Lu and Z. B. Guo, “A Method of Data Synchronization Based on Message Oriented Middleware and Xml in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Distributed Heterogeneous Environments," in *Proceedings of International Conf. on Artificial Intelligence and Industrial Eng*. no. Aiie, pp. 210–212, 2015.

[16] Ghorbanzadeh, R. Malekzadeh, A. Shaddeli, and Z. Jahanbakhsh, "A survey of mobile database security threats and solutions for it," in *Proceedings 3rd International Conference on Information Sciences and Interaction Science, ICIS, 2010, PP 676-682.*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table 1: Summary of Reviewed Synchronization Approaches*

| **Approach** | **Security** | **Energy Use** | **Memory Usage** | **Performance Speed** |
|---|---|---|---|---|
| SAMD Algorithm Taha and Alim's work [4] | Message digest | Not assessed | Not measured | SAMD(wired) compared with SWAMD(wireless). SAMD has 0.6secs faster than SWAMD on average. |
| MRDMS by Sethia et al [6] | None | Not assessed | Not measured | Uses cell matrix to improve performance speed. No measurements recorded in the process. |
| Cross layer method [7] | None | Not assessed | Not measured | TCP Westwood decreases synchronization rollback of transactions leading to improved speeds. |
| OODS by Li et al.[8] | None. | Not assessed. | Not measured. | No comparisons. |
| Malhotra and Chaudhary's work [9] | None | Not assessed. | Not measured. | Not measured. |
| SyncML [11], [12] | None | Not assessed. | Not measured. | Not measured. |
| DefferedSync [14] | None | Not assessed | Not measured. | Not measured. |