TEST Engineering & Management

March - April 2020
ISSN: 0193 - 4120 Page No. 3893 - 3900

A Novel Dynamic Data Scheduling in Wireless Broadcasting Environments

Seokjin Im[1]
[1]Computer Engineering, Sungkyul Univ., Anyang, Republic of Korea
imseokjin@gmail.com[1]

Article Info
Volume 83
Page Number: 3893 - 3900
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 23 March 2020*

Abstract:
Wireless data broadcast is reliable way for the scalability of information system, where a huge number of clients try to access data items. Especially, broadcasting spatial data enables various location dependent information service and IoT services from the client's current location. Data scheduling scheme is critical to the access time of clients because the scheme selects hot data items and determines the order of data items to be broadcast. This paper proposes a spatial data scheduling scheme, which considers the needs of data from the clients and the distribution of the clients simultaneously. The proposed scheme schedules dynamically through the recalculation of the access probability of items every given time slice. Through simultaneous and dynamical scheduling, the scheme reduces the average access time. The performance evaluation shows the proposed scheme outperforms the existing schemes with respect to the access time and tuning time of the clients.

Keywords:*Data Scheduling, Data Access Probability,Wireless Data Broadcast.*

I. Introduction

The advent of the high-speed communication technology and advanced devices with high computing powers makes real the vision of computing from the new concept like cloud computing or IoT (Internet of Things) [1]. Especially, the IoT services have to accommodate billions of various clients with the computing ability like various sensors, smart phones, laptop computers, smart cars, home appliances, buildings that are connected to the internet [2, 3]. In the IoT environment, the services through the networks including wireless need the efficiency in the aspect of high performances, ubiquity, reliable responses, and scalability that how many clients the services can provide the contents simultaneously and efficiently.

The scalability of the information system through the IoT platforms is critical more and more for the efficiency of the services because of the rapid increase of the number of the clients that try to access their desired data items. The wireless data broadcast system guarantees the scalability because it allows any number of clients to access data items at the same time [4, 5, 15, 16].

In the broadcasting system, a server broadcasts periodically a set of data items to a wireless channel, and then the clients access the channel and download data items that they desire to access [6, 12, 14]. Especially, the system broadcasting spatial data items allows to provide the clients with the information contents related to the clients' current locations.

The system can be categorized into flat and non-flat broadcast scheme by whether or not adopting the interest to data items of the clients into broadcasting on the channel. The flat broadcasting scheme does not take into consideration of the interest to data items of the clients. It disseminates all items in a set of data once in a broadcasting cycle. That means each data item has the same weight of the interest. In the non-flat broadcasting scheme, the system takes into consideration of the interest to data items of the clients. The non-flat system emphasizes the data items with high interest of the clients by placing them more frequently than the others on the wireless channel in a broadcasting cycle [10].

Figure 1, for example, shows a system disseminating spatial data items that adopts the non-flat scheme. In order for the server to select the high interested items, the clients transmit interested areas through the uplink channel. Using the interested areas, the server selects high interested data items, called hot data items, and broadcasts them over the downlink wireless channel in the non-flat broadcasting manner.

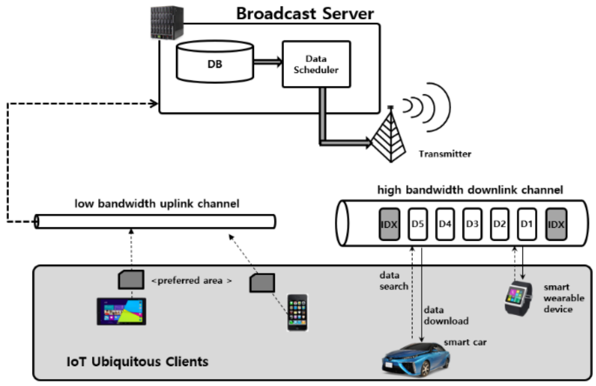

Figure 1. A Non-Flat Broadcast System

In the non-flat broadcasting, the scheduling scheme affects the clients' data waiting time, i.e., the access time, because it selects the hot data items from the set of data items to be broadcast and determines the order to be broadcast of them. In this paper, we propose a dynamic scheduling scheme for spatial data in order to support location-dependent information services in wireless data broadcast systems. The proposed dynamic scheduling scheme selects hot data items using the access probability that considers the interests to the data items of the clients as well as the distribution of the clients over data space. Also, the proposed scheme reselects periodically the hot data items by updating the access probability of the data items, then reschedules dynamically the data items on the wireless channel. That makes the proposed scheme different from the existing scheduling schemes and improves the access time of the clients.

We organize the rest of the paper as follows. In Section II, we give a summary of the related works. Section III provides the proposed dynamic scheduling scheme. In Section IV, we evaluate the proposed model through simulations and finally we conclude the paper in Section V.

II. Related Research

In the broadcasting system, it is important for the clients to download the queried data items in an energy-efficient way [13]. The system introduces an indexing scheme in order for the clients to download selectively only the queried items. In the indexing scheme, the server places an index information in interleaved manner with data items on the wireless channel. The index delivers the time information on when each data item is broadcast over the wireless channel. With the index, the clients can listen to the queried data items selectively. Thus, they process given queries energy-efficiently.

2.1 Schemes for Flat Broadcasting Model

The existing indexes for flat spatial data broadcasting can be grouped into two categories, one using the Hilbert Curve and the other using the space partition. The indexing schemes, HCI and DSI, are organized using Hilbert Curve for representing the location of each data item into an integer [8, 9, 11]. HCI has the structure of B+ tree interleaved with data items on the wireless

March - April 2020
ISSN: 0193 - 4120 Page No. 3893 - 3900

channel. DSI is constructed in a table, distributed on the wireless channel with data items. The two indexes make the clients download extra data items in addition to the queried data items because they do not use exact location of the data items. The index CEDI uses the space partition scheme that is organized with a table holding the exact location of the spatial data items [7]. That makes the clients download only queried data items, unlike DSI and HCI

2.2 Schemes for Non-flat Broadcasting

As the non-flat spatial data broadcasting, GDIN has been proposed. That uses the space partition scheme like CEDI for the flat broadcasting scheme. For selecting hot data items, GDIN defines hot cells and repeats all the data items in the cells several times in a broadcast cycle [10]. GDIN repeats the regular items within the hot cells with the hot items. That makes long the access time of the clients because it cannot help avoiding that the length of a broadcast cycle gets long. The proposed dynamic scheduling scheme in non-flat broadcasting does not repeat regular data items like GDIN. The proposed scheme selects hot data items with an effective way by considering the interest to data items of the clients and the distribution of them over the data space.

III. Dynamic Spatial Data Scheduling Sheme

We consider a system for broadcasting spatial data items in the non-flat manner, that is organized with a broadcast server, a downlink channel, an uplink channel, and a number of mobile clients, as shown in Figure1. In the system, each client sends a rectangle of region, R_{int}, as the interested area and its current location to the server through an uplink channel. Using the information from the clients, the broadcast server calculates the access probability of data items and selects hot items. Figure 2 shows the modules consisting of the server maintaining a spatial database with a set of N data items to be broadcast, called D. The data selector calculates the access probability and determines hot data items. The dynamic scheduler carries out scheduling hot items and regular ones for placing onto the downlink channel. The scheduler rearranges the data items periodically, reflecting the result from the data selector. The index generator organizes the index information for the data scheduling by the dynamic scheduler.

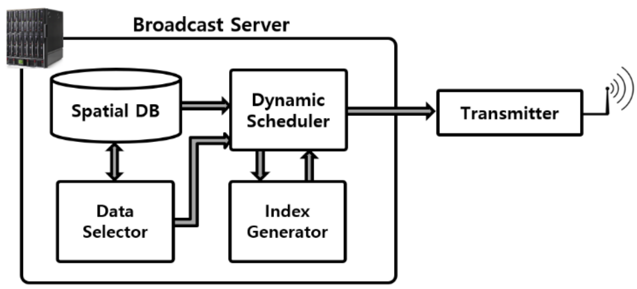

Figure 2 The Broadcast Server

For scheduling and indexing the data items to be broadcast, we adopt partitioning the data space. The broadcast server partitions the data space where the data items to be broadcast reside into grid cells. As an example, Figure 3 depicts 10 spatial data items on a data space. For the proposed dynamic data scheduling, the server partitions the data space into 4*4 grid cells as shown in Figure 3. The server allocates a cell number to each cell in row major order.

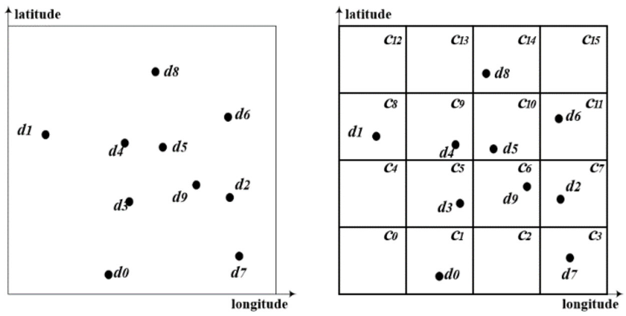

Figure 3. Spatial Data and its Grid Partition

March - April 2020
ISSN: 0193 - 4120 Page No. 3893 - 3900

3.1 Selecting Hot Data Items

In order to select hot data items from N items in D, the server calculates the data probability of each data using the information from the clients as below:

$$P_d(d_i)= N(d_i) / \sum^{k}_{j=1} N(R_j) \quad (1)$$

Here, k means the number of rectangles from the clients, same with the number of the clients. $N(R_j)$ denotes the sum of the number of data items within each rectangle from the clients. $N(d_i)$ means the number of rectangles containing data item d_i within them. The data probability $P_d(d_i)$ for data item d_i means how high the possibility the clients access d_i is.

The server calculates the region probability for each grid cell as below.

$$P_r(C_i) = N(C_i) / N_{client} \quad (2)$$

Here, $N(C_i)$ means the number of the clients residing grid cell C_i, and N_{client} means the number of the clients. The region probability shows the distribution of the clients over the grid cells.

Finally, the server calculates the access probability $P_a(d_i)$ of data item d_i in C_j, using data probability $P_d(d_i)$ of di and region probability $P_r(C_j)$ of C_j, as below:

$$P_a(d_i) = P_d(d_i) \times P_r(C_j) \quad (3)$$

Here, $P_r(C_j)$ denotes the region probability of C_j where data item di is located.The access probability means that the emphasis of the data probability, i.e., the higher region probability is, the higher possibility to be accessed is.

Algorithm1 shown in Figure4 describes the procedure for listing items in D and selecting Nhot items with high access probability as hot data items. The hot items are broadcast more frequently than the regular ones in a broadcast cycle on the wireless channel.

Algorithm 1 Listing and Selecting Hot Items

1: The input: $R = \{R_1, R_2, \ldots, R_k\}$ $D = \{d_1, d_2, \ldots, d_N\}$

2: The output: S_{hot}, S_{reg}

3: Loop for each R_j in R

4: calculate the sum of $N(R_j)$

5: update $N(d_i)$

6: Loop for each d_i in D

7: calculate $P_d(d_i)$

8: Loop for each C_i in all grid cells

9: calculate $P_r(C_i)$

10: calculate $P(d_i)$ for all data items in D

11: sorting all the data items in decreasing order of $P(d_i)$

12: select top N_{hot} data items in the sorted list and put in set S_{hot}

13: calculate $S_{reg} = D - S_{hot}$

Figure 4. Selecting Hot Data Items

In line 1 of Algorithm1, R is the set of rectangles from the clients and D is the set of data items to be broadcast. S_{hot} and S_{reg} in line 2 mean the set of data items for hot data and regular data items, respectively. The loop in line 3 is for calculating $N(R_j)$ and updating $N(d_i)$ for each rectangle from the clients. In line 6, the server calculates the data probability with Equation (1) shown above. Also, the server calculates the region probability with Equation (2) using the loop in line 8. In line 10, the server determines the access probability of data item di using the data probability and region probability. Then, the server sorts decreasingly all the data items in D according to the access probability, selects top N_{hot} data items as hot data items and determines S_{hot} with the selected hot items and S_{reg} with the rest of data items in D. Figure 5 shows an example of sorting, selecting hot and regular data items, and determining S_{hot} and S_{reg} for data items in Figure 3.

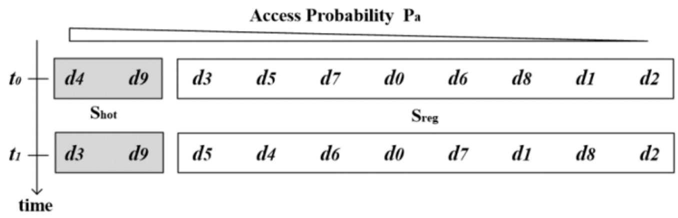

Figure 5. An Example of Access Probability

March - April 2020
ISSN: 0193 - 4120 Page No. 3893 - 3900

The server processes the algorithm periodically in order to adopt the changes of the interests from the clients to the broadcasting channel.

3.2 Dynamic Spatial Data Scheduling

The server organizes a list of data items as a schedule, called SCDL, for broadcasting every T_{slice}. Here, T_{slice} means the time period for recalculating the access probability of data items and for reorganizing the schedule of data items for broadcasting. Figure 6 depicts an example of data scheduling for the access probability shown in Figure 5. In the figure, the period for rescheduling, T_{slice}, is the time difference from t_1 to t_0. Here, hot data items are repeated several times in a broadcast cycle on the downlink channel in order that the clients access quickly them.

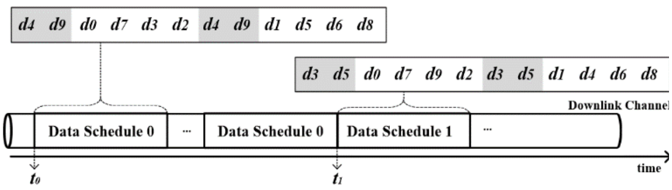

Figure 6. Data Scheduling on the Downlink Channel

Algorithm 2 Dynamic Scheduling Data

1: The input: T_{slice}, R = {R_1, R_2, …, R_k}, D = {d_1, d_2, …, d_N}
2: The output: SCDL the schedule of data items for broadcasting

3: Loop for every T_{slice}
4: Call Algorithm1 with R and D
5: $S_{hot} \leftarrow S_{hot}$ from Algorithm1
6: $S_{reg} \leftarrow S_{reg}$ from Algorithm1
7: Loop for each grid cell C_i in theset of grid cells
8: insert all data items in S_{hot} to SCDL in increasing orderof grid cell number in which each data resides
9: insert all data items in S_{reg} that reside in grid cell c_i

Figure 7. Dynamic Scheduling of Spatial Data

Algorithm 2 in Figure 7 describes the process for rescheduling the data items after recalculating the access probability of data items in D. In the loop repeated every T_{slice}, the server calls Algorithm1 for calculating the access probability and determining S_{hot} and S_{reg}. Here, T_{slice} is integer times of the time length of the broadcasting cycle. With the loop in line 7, the server schedules the hot and regular items to place them on the channel. The schedule follows the order of the cell number in row major order. Hot data items are placed repeatedly before every first cell in rows of the grid. Line 8 and 9 show the procedure of scheduling data items according to the cell order. Thus, Algorithm2 denotes the procedure that determines the order of data items to be broadcast to the channel dynamically. The algorithm reflects dynamically the change of the interests to data items of the clients.

IV. Evaluation

For the evaluation of the performance of the proposed dynamic scheduling scheme, we simulate a wireless broadcast system in the non-flat manner. The system has a broadcast server, two wireless channels, one for downlink and the other for uplink. The broadcast system accommodates 100 mobile clients. For the simulation, we use 6374 spatial data items in real skewed distribution and rectangles from the clients that follow the Zipf distribution.

4.1 Environment for Simulation

We set the parameters for the simulations as follows. Each data item has 1 KB as the size, the bucket size is 256 bytes. Here, the bucket is the data unit for delivering data items like a packet on the network. We set n to 32 for partitioning the data space into n*n grid cells. In the simulation, we adopt the indexing scheme that is transformed from CEDI, to the proposed dynamic scheduling in order to help the clients process their given queries on the wireless channel. We compare the proposed scheduling scheme in the average access time and tuning time with the two broadcast systems adopting GDIN and DSI as its indexing scheme, respectively. Here, GDIN considers only the probability that the clients access data items without the region probability, and the scheme DSI does not consider any interest to data item of the clients. Through comparing the proposed dynamic data scheduling with the two schemes GDIN and DSI, we show the proposed scheme outperforms the other schemes.

4.2 Access Time Evaluation

The access time means the time until a client obtains all data items from the downlink channel after tuning into the channel. The access time indicates how quickly the client can access queried items from the channel.

Figure 8 shows that the average access time in the unit of the number of 105 buckets of the proposed scheme and the two schemes according to the ratio of the number of hot data items to the number of entire data items to be broadcast, in the simulation condition that the query size is 0.06. The query size means the ratio of a query window to the data space. In the figure, DYNA means the proposed dynamic data scheduling scheme, and DSI and GDIN mean that the systems adopting DSI and GDIN as its indexing scheme. That depicts the proposed DYNA outperforms the other two schemes in the two values of ratio of hot data items. That is because the proposed DYNA considers the access probability by the two kind of probabilities, data probability and region probability. However, GDIN only take into consideration of data probability. It means that GDIN cannot cope with the interests to data items and the distribution of the clients over the data space. Also, DSI does not consider the interests to data items of the clients. That shows that the system by the proposed dynamic data scheduling allows the clients to access their preferred data items more quickly than other two systems.

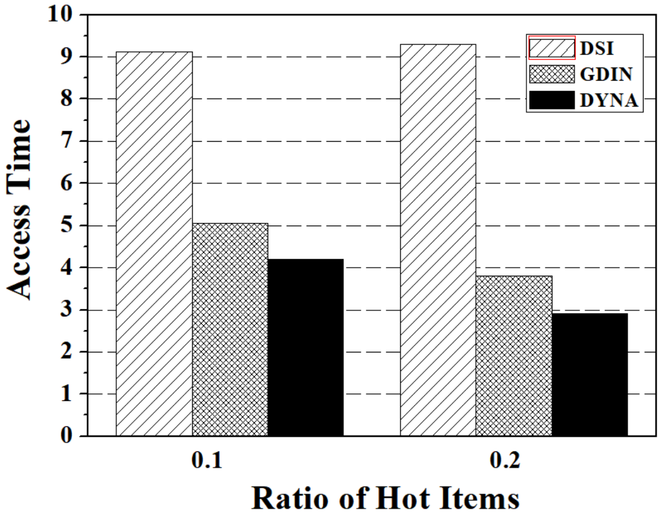

Figure 8. Access Time by Hot Item Ratio

Figure 9 depicts the average access time in the unit of the number of 105 buckets, according to the query sizes at the ratio of the hot data items of 0.2. The figure represents the effectiveness of the proposed scheme for various query sizes.

Figure 9. Access Time by Query Size

4.3 Tuning Time Evaluation

The tuning time is the performance metric indicating that how much energy a client uses for processing given query. The tuning time is measured with the number of the buckets the client accesses in energy-consuming active mode for the access time.

Figure 10 depicts the tuning time in the unit of the number of 103 buckets of the proposed scheme and the two schemes, according to the ratio of hot data items in the simulation condition that the query size is 0.06. The figure denotes that the proposed scheme has almost same with the GDIN. That is because the transformed indexing scheme from CEDI that the proposed dynamic scheduling scheme adopts is also based on the partition of grid cells. However, DYNA and GDIN outperform DSI. It means that DSI makes the clients access more buckets in the active mode on the wireless channel.

Figure 11 shows the tuning time according to the query size under the ratio of hot data items of 0.2. For the two query sizes, GDIN and DYNA outperform DSI. That is because the tuning time depends on the indexing scheme that the broadcast system adopts

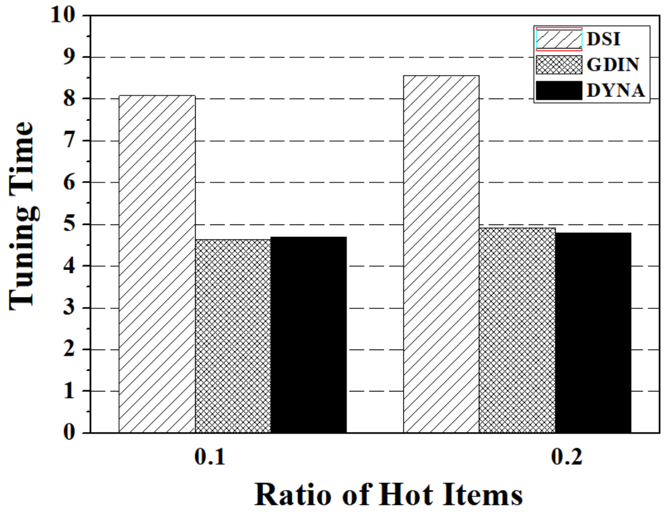

Figure 10. Tuning Time by Hot Item Ratio

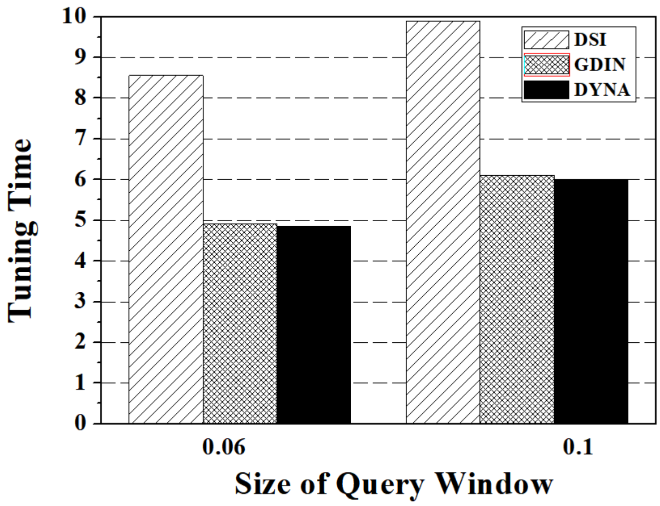

Figure 11. Tuning Time by Query Size

V. Conclusion

In this paper, we have proposed a dynamic spatial data scheduling scheme in wireless data broadcast system. The proposed scheme considers the access probability that is calculated with the two terms of data probability and region probability. It means that the proposed dynamic scheduling scheme can cope with the change of the interests to the data items of the clients and the change of the population of the clients over the data space simultaneously. The proposed scheme schedules dynamically through the recalculation of the access probability of items every given time slice. Through simultaneous and dynamical scheduling, the scheme reduces the average access time and shows the improved access time over the other broadcast systems adopting GDIN an DSI as its indexing scheme, respectively. Through the simulation study, we have shown that the proposed dynamic scheduling scheme outperforms the comparison targets, the other two broadcasting schemes.

VI. Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1D1A1B07050748).

March - April 2020
ISSN: 0193 - 4120 Page No. 3893 - 3900

VII. References

[1] A.R. Biswas and R Giaffreda. (2014). IoT and Cloud Convergence: Opportunities and Challenges. 2014 IEEE 1st World Forum on Internet of Things, Seoul, South Korea.

[2] D. Park, H. Bang, C. S. Pyo, and S. Kang. (2014). Semantic Open IoT Service Platform Technology. 2014 IEEE 1st World Form on Internet of Things, Seoul, South Korea.

[3] S. Wang, Y. Hou, F. Gao, and X. Ji. (2016). A Novel IoT Access Architecture for Vehicle Monitoring System," 2016 IEEE 3rd World Forum on Internet of Things, Reston, VA, USA

[4] Y. Gotoh, and H. Taniguchi. (2012). A Scheduling Scheme for Improving Error Error Resilience on Media Data Broadcasting. - 2015 International Conference on Network-Based Information Systems, Melbourne, Australia.

[5] S. Mavridopoulos, P. Nicopolitidis, G. Papadimitriou, and P. Sarigiannidis. (2015). Broadcast Levels: Efficient and Lightweight Schedule Construction for Push-based Data Broadcasting System. IEEE Trans. on Broadcasting, vol. 61, no. 3, pp. 470-481.

[6] Y. Yao, X. Tang, E. P. Lim and A. Sun. (2006). An energy-efficient and access latency optimized indexing scheme for wireless data broadcast. - IEEE Trans. Knowl. Data Eng., vol. 18, no.8, pp 1111-1124, Aug.

[7] S. Im, M. Song, J. Kim, S. Kang, and C. Hwang. (2006). An error-resilient cell-based distributed index for location-based wireless broadcast services. - Proc. 5th ACM Workshop MobiDE 06, pp. 59-66, Chicago. USA, June.

[8] B. Zheng, W. C. Lee, and D. L. Lee. (2004). Spatial Queries in Wireless Broadcast systems. - Wireless Netw., vol 10, no 6. Pp 723-736, Dec.

[9] W. C. Lee and B. Zheng. (2005). DSI: A fully distributed spatial index for location-based wireless broadcast services. - Proc. 25th IEEE International Conference on Distributed Computing Systems.

[10] S. Im, H. Youn, J. Choi, and J. Ouyang. (2011). A novel air indexing scheme for window query in non-flat wireless spatial data broadcast. - Journal of Comm. And Networks, vol. 13, no. 4, pp. 400-407.

[11] B. Zheng, W. C. Lee, K.C.K Lee, D. L. Lee, and M. Shao. (2009). A distributed spatial index for error-prone wireless data broadcast. - VLDB J., vol. 18. No. 4, pp. 959-986.

[12] K. Tao, Y. Kou, X. Lu, and R.Song. (2015). Design of data distribution protocols for next generation broadcast wireless systems. - 2015 10th International Conference on Communication and Networking in China, Shanghai.

[13] C. Zhan, V. C. S. Lee, J. Wang, and Y. Xu. (2011). Coding-Based Data Broadcasting Scheduling in On-Demand Broadcast. - IEEE Trans. on Wireless Comm., vol. 10, no. 11, pp.3774~3783.

[14] S. Dhanked and V.Goel. (2014). Hashing techniques for broadcasting in wireless data environment. - 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques, India.

[15] J. Kuang, and S. Yu. (2017). Broadcast-based Content Delivery in Information-Centric Hybrid Multihop Wireless Networks. - IEEE Comm. Letters, vol. 21, no. 4, pp. 889-892.

[16] A. B. Waluyo, F. Zhu, D. Taniar, and B. Srinivasan. (2014). Design and Implementation of a Mobile Broadcast System- 2014 IEEE 28th International Conference on Advanced Information Networking and Applications, Vitoria, Canada.