Gyancity Journal of Electronics and Computer Science,
Vol.8, No.1, pp. 48-60, March 2023
ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003

Dynamic RFID Data Filtering and Application

Khaled Mohammed[1, 2]

[(1)] *Programming Technology Department, Sana'a Community College, Yemen*
[(2)] *Administration Information System Department, Faculty of Economic & Administration Science, Queen Arwa University, Yemen,*
Khaled.Mohammed@scc.edu.ye

Abstract

Radio-frequency Identification "RFID" Technology uses the radio frequency waves to transfer the RFID data between RFID readers and tags which are used to identify objects/ employees without line of sight. The RFID data which is captured by the tag reader may contain false readings, noise, and duplicates which implies data filtering and cleaning. Therefore, it is necessary to develop efficient processing algorithms of RFID data. This paper presents a dynamic technique to filter the RFID data, eliminate duplicates and filter noise. Data filtering during employee identification in the workplace enhances the performance of employee attendance management systems. The proposed system compared to De-noising and duplication Elimination approach under different arrival rates at a rate 0.1tag/sec and under the noise rate at rate between 0.085-.01 tag/sec

Keywords: *Filtering technique, Duplicate Elimination, RFID system, De-noising RFID data, Dynamic RFID Data Filtering and Employee Attendance Management*

1. Introduction

Recently, researchers are focusing on enhancing the performance of RFID systems by solving tag collisions, reader collision and data filtering problems.

The unique identifier EPC readings received by an RFID reader from tags are subject to errors. Due to the existence a large number of tags in the coverage region of a reader, some readings are incorrect, and others are ignored or include noise. These problems cause a large congestion in the spooler process in memory that is leading to reduction of the efficiency and effectiveness of the system. This paper proposes an effective approach to solve these problems. Moreover, we seek dramatically to reduce the size of the false readings, noise and duplicates compared to the actual readings, through considering significant readings in a particular interval to be the right readings. This approach will facilitate the integration of data and storage process after being classified and handled properly.

Gyancity Journal of Electronics and Computer Science,
Vol.8, No.1, pp. 48-60, March 2023
ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003

The paper is organized into six sections. Section 2 presents the background about RFID data filtering. Section 3 reviews and discusses the different RFID data filtering techniques. Section 4 discusses the proposed approach. Section 5 presents the system performance evaluation. Finally, the conclusion will be discussed in section 6.

2. Background

RFID readers in RFID Employee Attendance Management Systems (EAMS) usually read tags multiple times. Memory hungry duplicate elimination methods such as Time Bloom Filter (TBF) waste memory for storage of multiple tag detection times. Experimental results of real-time application of the proposed method in Employee Attendance Management (EAM) indicate excellent performance.

Data stream filtering and cleaning in RFID systems [1] is indispensible. The most important characteristics of data streams such as uniformity, huge amount of data, reading times, erroneous data and requirement of integration should be considered for enhancing system effectiveness.

It is necessary to carry out RFID data filtering [2-5]. There are three classes of undesirable types of data that need to be filtered out:

(1) False negative readings (FNR): This case occurs when the RFID reader reads a part of the RFID tags and neglects the other. The RFID reader reads much RFID tags frequently at the same time; this process sometimes causes tag collision, the RFID reader cannot be reading all the RFID tag at the same time. The second reason is wireless frequency interference.

(2) False positive readings (or noise) (FPR) : This case occurs when the RFID reader reads the RFID tags outside the scope of regular reading. For example, the reader reads the items that have been read in an adjacent state. The second reason is unknown reasons for the reader or environment.

(3) Duplicate Readings (DR): RFID reader reads all the RFID Tags that are existing for some time within his domain; the RFID reader is installed to cover a vast area, the tags get multiple reading in the overlapped domain by multiple readers. Another reason for reading tags intentionally is to enhance reading accuracy.

In practical terms, the multi-reading for readers with RFID tags are performed in several cycles to achieve the highest recognition rate [5]. We seek to reduce the size of the false readings or noisy readings compared to the actual readings through significant readings in a particular time interval to ensure right readings.

3. Overview of RFID Data Filtering Techniques

3.1. Sliding Windows Technique [6]

The main objective of this approach is reducing the noisy readings. Bai et, al. [6] proposed a sliding window technique. This method uses the FIFO queue "win-buffer" to solve the problem and store the data. The win-buffer holds the sliding windows. Win-Buffer contains two main parameters before adding the incoming tag queue to it. It includes win-size and time that is described in the model of [t1, t1+win-size], after a period T, the model will become [t1+T, t1+win-size+T].

RFID reader reads the data rows that have entered into the sliding windows queue with a count Threshold of EPC values to reduce noise. If the count of readings with the same tag EPC value is equal to or bigger than the threshold, then the EPC value is not noise. It continues to be forwarded to further processing. Otherwise, the reading is ignored. They used two parameters: win-size of a sliding time and threshold to detection the noise.

The RFID reader uses three parameters: header id, tag id, timestamp. Every parameter has a particular task, the tag id refers to the EPC of tag, and header id refers to EPC of the reader and timestamp. In their algorithm, it is possible to consider the pair of keys (header id, tag id) for reading. It uses multiple readers to observe same tags by tag id.

3.2. Merge Baseline Technique [6]

RFID system used a network sensor for reading and transferring data. The reader, can read the data from time to another via the network several times. Bai et, al. [6] proposed the algorithm that has been named Marge Baseline. The goal of this algorithm is deleting or minimizing the redundancies.

It is noted that the algorithm has one parameter: max-distance. The reader compares between the current and previous readings. If a reading is within max-distance in time, have the same key, then the reading is found as a duplicate.

The algorithm presented in [6] compares between the tag id key that exists in the win-buffer queue and the same id tag key that is existing in the incoming queue. If nothing is found, that means, not duplicate has been found, and it is output as a new reading, and then the reader has added the data from incoming queue to the win-buffer queue. Otherwise, the reader deletes the id tag and his data from income queue and adds the rest of the tags to the win-buffer queue.

3.3. De-noising Technique [7]

The de-noising technique has been implemented for processing and detection of noisy readings. The technique works under two conditions. The first condition is checking for every reading from the tag. If the reading is not found in the list, then add the reading to the list and store detection time: the initial and latest time of reading the tag. The second condition is calculating the number of repeated readings of the tag in

Gyancity Journal of Electronics and Computer Science,
Vol.8, No.1, pp. 48-60, March 2023
ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003

the RFID reader. If the count of repeated readings is greater than or equal a count-threshold and time length of repeat reading the tag is less than or equal time-threshold, then reading is correct. Otherwise, the reading is noise.

3.4. Duplication Elimination Technique [7]

The duplicates elimination technique is implemented in [7] for removing the duplicate data. The duplicates elimination algorithm works in a similar way to the de-noising algorithm with some modifications. More details could be found in [7].

3.5. Multiple Readers Filtering Technique [8]

This Technique addresses one of the false readings problems which occurs by of misreading between the reader and Tags. There is a lot of reasons for this problem such as jamming signals, pressure on readers to read more than his capacity at the same time. In addition to the security problems and penetration. It is necessary to control multiples readings through orientation of the signals. The signals are visible to some readers and invisible to another. Identification of the tagged object is necessary to be confirmed if it gets the positive reading. Otherwise, the reader uses the probabilities P for identifying the tagged object [8].

3.6. Other Techniques

Bai et al. [6, 9, 10] proposed a new method that has named a sliding window for de-noising. This method has a low occurrence rate. It used the FIFO queue “win-buffer” to solve the problem and stored the data. The win-buffer holds the sliding windows of reading the size of sliding windows that pass with the times. The method used Threshold of EPC values to reduce noise. The performance of this proposed has been studied under different noise ratios. Each tag is repeated ten times, with 200 milliseconds, the overall tag arrival rate is 1/sec. The results indicated noise ratios between 1, to 50%. The output readings of their method are not in the correct time order.

Mahdin et al. [7] proposed a de-noising algorithm that used the count-threshold and time-threshold parameters to indicate the number of frequent reading and time is set for the number of frequent reading that should be achieved in order correct reading respectively. The proposed technique keeps the recent readings of the tags in the list and calculates the length of time by the difference between the initial time and the latest time of reading the tag. The performance of the proposed technique is reducing the processing of time reading from 0.308 to 0.125, and the algorithm is performed under different arrival rate. Every tag reading has been repeated for ten times. The applied noise ratio varied between 10 to 50%. Therefore, the noise ratio is very high.
Bai et al. [6] proposed effective and efficient algorithms for RFID data filtering. It included noise removal and duplicate elimination. They developed algorithms that

were compared to baseline implementation and worked more efficiently. It requires less buffer space for history storage to reduce the De-noising and duplicate elimination. The performance of the proposed method is dependent on the average of arrival rates for the tags. The highest rate of reading arrived in 1000 tags/sec, but the reading rate after 500 readings/sec resulted in longer delays.
Mahdin et al. [7] proposed the de-noising with elimination duplication algorithm to improve performance and get rid of the redundancy. The algorithm has studied output readings while checking for the readings which has been output before. If the reading not been output before then the state of the output is set to true and no duplication is declared. Otherwise, delete the reading. The performance of the algorithm has focused on the small list of distinct readings. Their results showed increase in processing time whenever the arrival rate increased.

Roozbeh et al. [11] proposed a method that processed data to improve performance and get rid of redundancy. The method keeps an initial timestamp for all existing items that are detected in the reader area. It compared between the initial timestamp for all existing RFID tags and the tags that are coming after each reading. If RFID reader receives any different reading with the existing one, the difference explains that the RFID reader has added the tag to the list of reading or removing the tag from the reader area. The method cannot work well with passive RFID tags because of the occurrence of false readings.

Kamaldin et al. [12] proposed a method for processing duplicate data. The method modified the Bloom approach to improving the performance. The approach has used the single hash function. They compared their results with the three previous approaches Bloom filter, d-left time Bloom filter. The result has shown that filtering of redundant data is better than other approaches in the three types of processing such as true positive rate, false positive rate and execution time. The Bloom approach is very complex and needs a lot of time [13].

Peng et al. [14], proposed an approach which works with multiple RFID tag readers for cleaning the false negative and false positive readings. The idea of this approach is that the RFID reader compares the current reading with the reading from the previous reader and the next reader if the reading is the same than the reading is correct. Otherwise, the reading will be removed from the list. The approach does not work for one RFID reader.

Gyancity Journal of Electronics and Computer Science,
Vol.8, No.1, pp. 48-60, March 2023
ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003

4. Dynamic RFID Data Filtering Algorithm

The RFID reader receives enormous amounts of data. It cannot distinguish between the true and false readings. The following algorithm explains an algorithm to solve the previous problems. Step 1 starts to check the reading of tag, if the tag reading has been red before then it is ignored. Step 2 compares the word numbers of RFID tag reading and word-threshold, if it equals then, the reading is correct. Otherwise, the RFID tag does not contain any data, or the data is incomplete and remove tag reading from the list automatically before entering to process.

The advantage of this condition is reducing the noise and improving the way, the reading that are coming to RFID reader are stored in the list to avoid the risk of memory overflow that produce much noise readings.

It is noted that we have tested the state of output RFID tag at the beginning of the algorithm. The Algorithm made the checkpoint to prevent repeated readings when the state of output RFID tag is true. Therefore, remove the tag record from the list until the repeated reading does not cause noise, duplication increases the processing time for data. Step 3 checks every reading from the tag. If the reading is not found in the list, then add the reading to the list and store the initial timestamp or latest timestamp of tag reading. Otherwise, increment the counter by one and update the time detection in reading the Latest timestamp. Step 4 calculates the number of repeated reading of tag in the RFID reader. If the number of readings is greater than or equal a count-threshold, then the reading is correct. Otherwise, the reading is noise and is removed from the list. Also, the algorithm calculated the time of repeating the tag in step 5, if time length is greater than time-threshold, then output the reading with Latest timestamp and remove the tag record from the list. Otherwise, the reading duplication and remove the tag record from the list.

Algorithm 1: Dynamic RFID data filtering (DRDF)

Input: Parameters: tag ID, detection-time count-threshold, time-threshold
and Word-threshold

Outputs: Filtering out of noise and Duplicates Elimination

1. For Every incoming reading from the tags Do,
2. Step1: IF State of Output is True before then
3. Delete the tag record from the list.
4. Else
5. Step2: IF the word numbers of the tag is not equal the word-threshold then
6. Delete the tag record from the list
7. Else
8. The reading is correct,

Gyancity Journal of Electronics and Computer Science,
Vol.8, No.1, pp. 48-60, March 2023
ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003


```
Step 3:     IF the reading is not on the list then
                 Add the reading to the list
                 Determine the detection-time
                 Save  this time in reading initial and  last time stamp
                 Increment the reading counter by one.
            Else
               Increment the reading counter by one.
          Update the Latest timestamp in detection-time
          Output the reading with the Latest –time
        End if.
Step 4:   IF the count of repeated reading is greater than or equal the count-
 threshold then
               IF  the Reading is repeated then
                  Delete the reading
              Else
                Output the Reading  with Initial timestamp
               Update the output state of reading to true.
            End if.
      Else
       IF  time length is greater than or equal the time-threshold then
            Remove the reading from the list
         End if.
      End if.
   End if.
  End For.
 For every record of tag  that is found in the list DO,
Step 5: IF the current time for latest timestamp  is greater than or equal the
time-threshold then
          IF  State of Output is True   then
             Output the reading with Latest  timestamp
             Remove the tag record from the list
           End If.
         Else
         Remove the tag record from the list
      End If.
   End For.
```

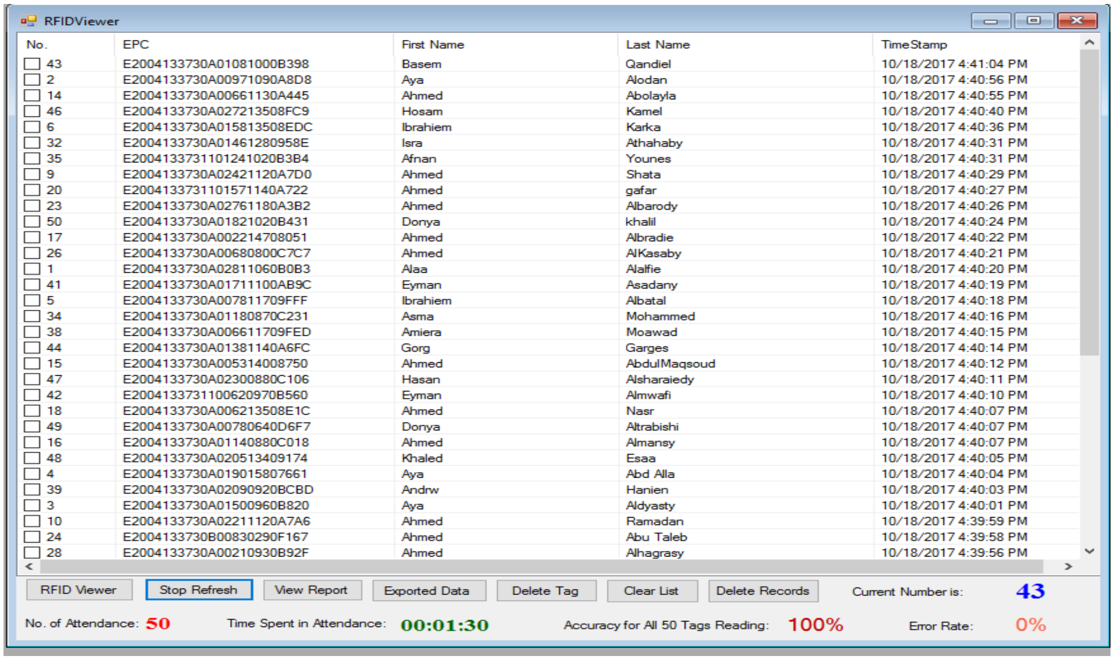
The proposed approach has been implemented and tested in an EAMS setting at Sana'a Community College in Sana'a Yemen. Figure 1 shows the RFID system layout.

Figure 1
The RFID system layout.

Figure 2 shows the real-time test results of the system. The output attendance sheet for a group of employees shows the Electronic Product Code (EPC) for each attendant, his name, the date and time of place work entry. The RFID system automatically gets the accuracy, and error rate in real time.

Figure 2
The employee's RFID based sample attendance sheet

ISSN: 2446-2918 DOI: 10.21058/gjecs.2023.81003



5. System Performance Evaluation

In this section, we explained the experimental work that measures the performance the proposed approach under noise rate and time execution under different arrival rates. We compared the performance of this approach with the de-noising and duplication elimination approach. The results shown in Table 1 indicate that the proposed approach is better than de-noising and duplication elimination in a lot of the tasks. The dynamic RFID data filtering approach is capable of effectively filtering the noisy the data, elimination of duplicates and processing the RFID data at high speed.

Table 1

Test result and comparison between de-noising and duplicates elimination and dynamic RFID data filtering.

Comparison of algorithms through completion rate and characteristics	De-noising & Duplicates Elimination	Proposed Approach (DRDF)
Add reading to the list with conditions	No	Yes
Accept missing value	Yes	No
Time threshold	Yes	Yes
Counter threshold	Yes	Yes
Word threshold	No	Yes
State of output the reading before applied rules	No	Yes
State of output record before saved in the Reading list	No	Yes

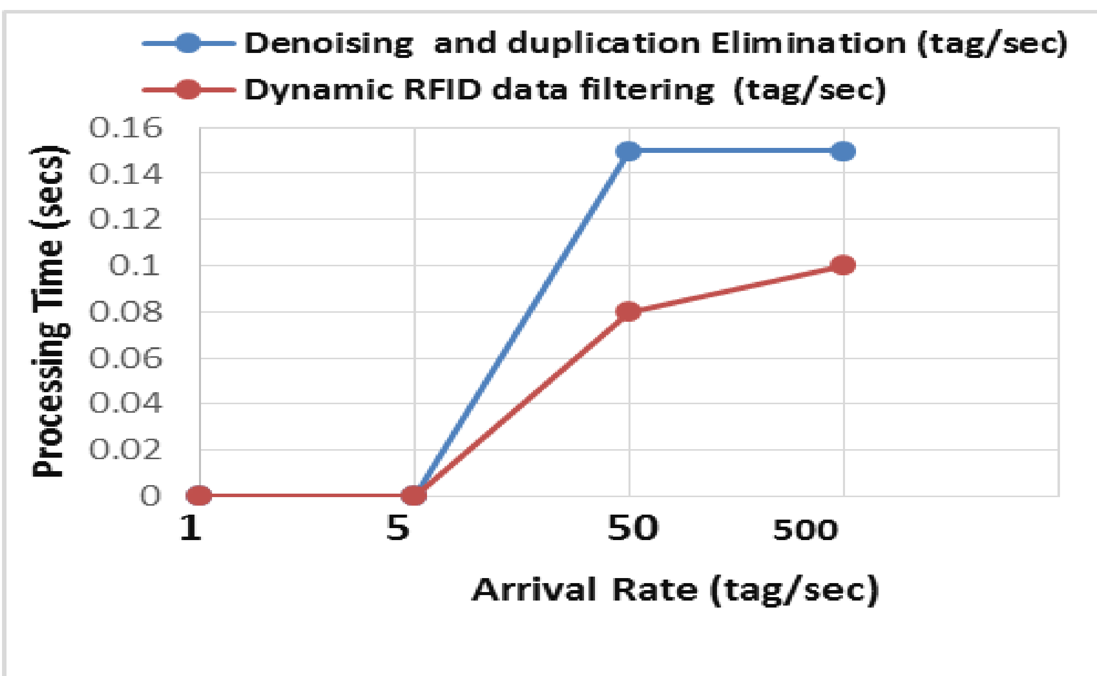
Delete record from the list	Only after	Before & After
Noise readings rate (sec)	0.15	Less than 0.1
Arrival rate (tag/sec)	0.125	0.085 -0.01

5.1. Result and Discussion

The aim of this experiment is to measure the performance regarding the execution time under different arrival rates. We tested the algorithm at different arrival rate such as 1 tag/sec, 5 tag/sec, 50 tag/sec, 500 tag/sec and repeated every tag ten times. Figure 3 shows the variability in test results. The reading has been tested on 1,5,50 and 500 tags/sec. The RFID reader receives the readings very fast at rates of 1 and 5 tags/sec. The difference between the performance of algorithms when the RFID reader read the reading on 50 and 500 tags/sec is also tested. The average of the arrival rate for the dynamic RFID data filtering is Less than 0.1 tag/sec while the average of the arrival rate for the de-noising and duplication elimination is 0.125 tag/sec. The experiments show that the dynamic RFID data filtering is better than the other methods regarding the arrival rate.

Figure 3

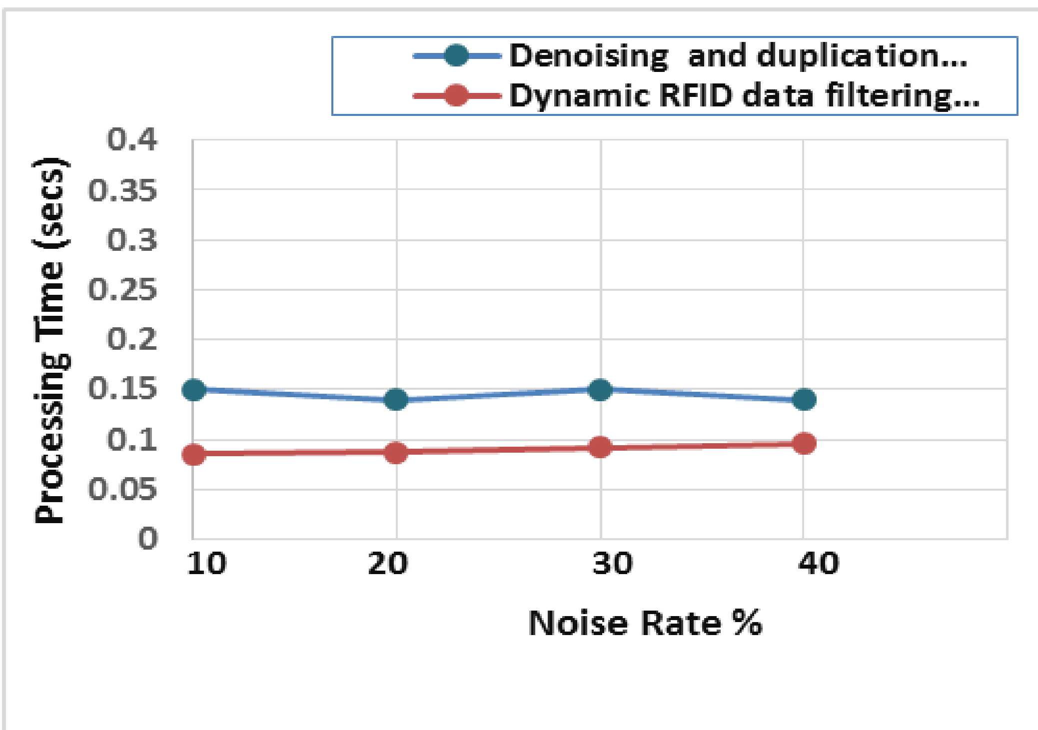
Execution time under different arrival rates.



We measured the performance this proposed approach under the noise ratio in some stages ranging from 1 tag/sec, 5 tag/sec, 50 tag/sec, 500 tag/sec and repeated every tag 10 times. The results showed that the noise ratio is less than 3%. Figure 4 shows the variability in test results. The RFID reader did not get any noisy readings in all algorithms when the RFID reader reads 1 tag/sec or 5 tag/sec. We noted the difference between the performance algorithms when the RFID reader read 50 tags/sec or 500 tag/sec. The average of the noise rate for the dynamic RFID data filtering ranges between 0.085-0.01 tag/sec while the average of the noise rate for the de-noising and duplication elimination is 0.15 tag/sec. We found that the dynamic RFID data filtering is better than the other methods in reducing the noise rate.

Figure 4

Time execution under noise rates



6. Conclusions

The paper has explained the various filtering techniques for RFID data filtering and implemented a new approach to reduce the noise, and eliminate duplication. The proposed approach uses a dynamic list for the process of filtering and data organization to avoid process risk of memory overflow that facilitated to reduce the noise and duplication in a very large. The proposed approach works under certain conditions such as word threshold, a time threshold, counter threshold and repeated readings for many times for each tag to get the correct reading which has reduced processing time for the RFID data. Our experiments have achieved a significant reduction in the time for reading the RFID data, processing and storing the records of the RFID tags under different noise rates and arrival rates. The experimental results also showed a significant reduction in noise and duplication elimination compared to the de-nosing and duplication approach. Implementation in a work setting for management of employee's attendance resulted in promising results in de-noising and duplicate elimination.

References

1. O. Mylyy, "RFID Data management, aggregation and filtering," Hasso lattner Institute at the university of Potsdam, 2006.
2. C. Floerkemeier, D. Anarkat, T. Osinski, and M. Harrison, "PML core specification 1.0," Auto-ID Center White Paper, 2003.
3. C. C. Aggarwal and J. Han, "A survey of RFID data processing," in Managing and Mining Sensor Data, ed: Springer, 2013, pp. 349-382.
4. J. Brusey, C. Floerkemeier, M. Harrison, and M. Fletcher, "Reasoning about uncertainty in location identification with RFID," in Workshop on Reasoning with Uncertainty in Robotics at IJCAI, 2003, pp. 23-30.
5. H. Vogt, "Efficient object identification with passive RFID tags," in Pervasive computing, ed: Springer, 2002, pp. 98-113.
6. Y. Bai, F. Wang, and P. Liu, "Efficiently Filtering RFID Data Streams," in CleanDB, 2006.
7. H. Mahdin and J. Abawajy, "An approach to filtering RFID data streams," in Pervasive Systems, Algorithms, and Networks (ISPAN), 2009 10th International Symposium on, 2009, pp. 742-746.
8. S. Tyagi, A. Ansari, and M. A. Khan, RFID Data Management: INTECH Open Access Publisher, 2010.

9. S. Tyagi, A. Ansari, and M. A. Khan, "Dynamic threshold based sliding-window filtering technique for RFID data," in Advance Computing Conference (IACC), 2010 IEEE 2nd International, 2010, pp. 115-120.

10. K. Bashir, M.-S. Park, S.-I. Lee, J. Park, W. Lee, and S. C. Shah, "In-network RFID Data Filtering Scheme in RFID-WSN for RFID Applications," in Intelligent Robotics and Applications, ed: Springer, 2013, pp. 454-465.

11. R. Derakhshan, M. E. Orlowska, and X. Li, "RFID data management: challenges and opportunities," in IEEE International conference on RFID, 2007, pp. 175-182.

12. H. Kamaludin, H. Mahdin, and J. H. Abawajy, "Filtering Redundant Data from RFID Data Streams," Journal of Sensors, vol. 2016, 2015.

13. H. Yongsheng and G. Zhijun, "Redundancy removal approach for integrated rfid readers with counting bloom filter," Journal of Computer Information Systems, vol. 9, pp. 1917-1924, 2013.

14. Z. Ji, Z. Luo, E. Wong, and C. T. X. Peng, "A P2P Collaborative RFID Data Cleaning Model," in Hong Kong the 3rd International Conference on Grid and Pervasive Computing-Workshops2008 IEEE, 2008.