ORIGINAL RESEARCH PAPER
Pages: 201-210

Integration of WSN and RFID networks, and redundant data filtering

Mahia Samani[1] , Ahmad Khademzadeh[2]
[1]*Department of Computer Engineering, Qom branch, Islamic Azad University, Qom, Iran,*
[2] *Iran Telecom. Research center,*
m_samani_stu@qom-iau.ac.ir, zadeh@itrc.ac.ir
Corresponding author: *zadeh@itrc.ac.ir*

DOI:10.22070/jce.2021.5775.1166

Abstract- **Radio frequency identification (RFID) and wireless sensor networks (WSNs) are two important and widely applied wireless technologies with limitless future potential. RFID is used to detect object location, while WSN is used for environmental sensing and monitoring. The integration of RFID and WSN not only provides identity and location but also facilitates environmental condition sensing. However, RFID data contains excessive duplication, which results in time delay and increased energy consumption, resulting in wastage of various network resources. This paper proposes a hybrid network designed by the integration of WSN and RFID and consisting of seven types of nodes. In this architecture, the entire network is divided into clusters and this clustering is obtained by particle swarm optimization. In addition, it also proposes an algorithm to overcome the issue of redundant data on this hybrid network. Simulation results show that the proposed algorithm reduces both data redundancy and processing time compared to existing algorithms.**

Index Terms- Integration, RFID, WSN, Redundant data, Data filtering

I. INTRODUCTION

Wireless sensor networks (WSNs) comprise a sink node (or base station) and several small, light, wireless, and cost-effective devices called sensor nodes [1]-[2]. Sensor nodes are capable of data processing and sensing; they sense environmental conditions (such as temperature, pressure, light, sound, and vibrations) and capture data based on the parameters [3]. Such captured and processed data are delivered to base stations through head nodes. These sensor networks are used in numerous applications, such as environment monitoring, particularly border control, industrial control, army, and healthcare [4]-[5]. Nevertheless, such sensor networks are not capable of identifying objects around them. On the other hand, networks that integrate WSN and radio frequency identification (RFID) facilitate the simultaneous sensing of environmental conditions and identification of objects [6]. Such networks are employed by a number of facilities such as chain store management [7] and

Manuscript received 25-Sept.-2020 and revised 18-Nov.-2020, *P- ISSN: 2322-4088*
Accepted on 17-Dec.-2020 *E- ISSN: 2322-3936*

healthcare [8].

FID technology implementation comprises a data reader, tag, and software The data reader reads the tags attached to objects and saves data in its memory [9]. RFID does not support a multi-hop from one data reader to another. However, the integration of RFID and WSN can enable the transfer of data from one data reader to another by the sensor network protocol, so that it can finally be delivered to the base station. [1]-[1] 0 1

RFID and WSN integrated networks have their own challenges, such as energy consumption, data filtering, data cleaning, real-time performance, and authentication, among which, data filtering and cleaning are more important. Redundant data cleaning is a type of processing that modifies, replaces, or deletes incorrect or irrelevant data [1]-[1]. Redundant data may generally be classified into three 2 3 types: [1]-[1] 4 5

- ❖ A tag may be read by the data reader several times without any change in its data (reading tags in close time intervals).
- ❖ Some tags may be placed within the range of several readers, and the data in the overlapped area are read by several readers.
- ❖ Data level duplication occurs when in order to increase reliability and decrease the missing data rate, a few tags with one EPC may be attached to the same object [14]-[15].

The remainder of this paper is organized as follows. Section ii explains the research methodology. Section iii provides an overview of related work, while the architecture of the hybrid network is explained in Section iv. The proposed algorithm is described in Section v and simulation results are presented in Section vi. Section vii concludes the paper.

II. METHODOLOGY

The steps of the research methodology used in this work are as follows:

- ❖ Review and synthesis of existing knowledge on WSN and RFID network integration and redundant data filtering
- ❖ Investigation of existing scenarios and problems
- ❖ Formulation of a hypothesis that solves existing problems
- ❖ Hypothesis testing
- ❖ Evaluation and conclusion

III. RELATED WORK

Redundant data filtering is one of the key challenges in RFID applications. Most organizations want access to only one copy of data (i.e., data without redundancy). Many studies have been conducted with the aim of filtering redundant data in servers. The transfer of redundant data from nodes to the

TABLE I. Weak Points of evaluated approaches

APPROACHES	WEAKNESSES
INPFM [16]	some duplicate data might not meet at a filtering point computation cost is high
CLIF [17]	computation cost relatively high and inducing delays computation Overhead inter-cluster duplication may be or not filtered at an inter mediate CH
EIFS [14]	inducing delays amount overhead sometimes changing of route causes that route becomes longer

server may affect the energy of the nodes, resulting in network overhead and shortening network life. In order to reduce such unnecessary transmissions, redundant data must be filtered within the network. INPFM[1] [1] and CLIF[2] [1] filter re6dundant data ins7de the network, but these approaches involve high calculation expenses and do not remove all overheads. The INPFM algorithm employs the tree structure and filters redundant data in each K-hop reader. The problem is that if INPFM checks redundant data in each hop, calculation expenses increase; on the other hand, if such data are checked after a few steps, some redundant data will enter the network and increase the network load. CLIF applies clustering topology and includes two phases: (a) identification of redundant data, and (b) filtering of redundant data. This scheme classifies redundant data into two classes: (1) redundant data within a cluster, and (2) redundant data between clusters. If a node receiving the data is the cluster head, it checks whether the received data belongs to that cluster or the adjacent cluster and filters the data; otherwise, the data will be transferred to the next cluster. The problem with CLIF is that only the redundant data of each cluster and its adjacent cluster will be filtered, while some redundant data may remain unfiltered in overlap zones and may be transferred to the sink. This problem may be more prominent when there is more redundant data in the overlap zones. The EIFS[3] algorithm uses a clustering structure [14]. The problem in CLIF can be eliminated by transferring the redundant data to the sink without filtering, and if this process is repeated several times, the data delivery route will change (for this purpose, a threshold is taken into consideration). However, in some cases, such a route change may result in the extension of the data delivery route to the sink. A summary of the limitations of all three above-mentioned approaches is given in Table I.

IV. ARCHITECTURE:

Fig. 1 reveals the architecture of hybrid network from WSN and RFID in the proposed plan that includes the seven types of nodes as follows:

[1] In-network phased filtering mechanism
[2] Cluster-based in-network phase filtering scheme
[3] Energy Efficient RFID data filtering scheme

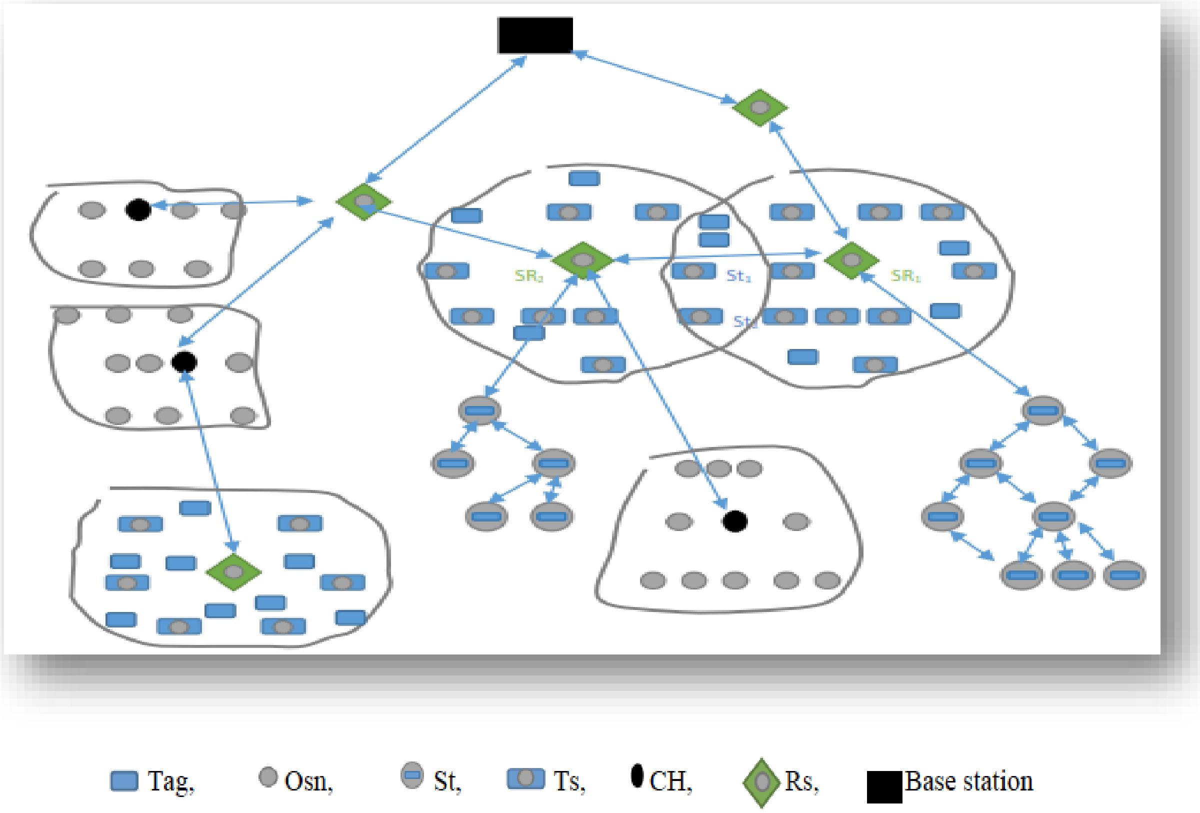

Fig. 1. The architecture of hybrid network from WSN and RFID in proposed plan.

The Ordinary Sensor node (OSN) is the usual sensor node that senses environmental conditions and relays data to the CH. The cluster head node (CH) is a type of OSN moreover, it can perform additional functions, for example: integrating and processing data. Each cluster has only one CH.

In this architecture, a TS node is composed of an RFID tag and a sensor and is capable of sensing the environmental conditions in addition to the identification of the objects. [1] 8

There are three type of TS: Active, passive and semi-active [1]- [2]-[2]. In this architecture, 9 0 1
passive TS does not use batteries, for Sensing or Communication. Sensor-tag (ST) is an integration sensor node with an RFID tag. So that the integrated sensor-tags will be able to communicate with each other as well as with other wireless devices while a TS is able to communicate only with the RFID reader.

Another component, the reader sensor (RS) is a composition of an RFID reader and a sensor node that acts as a node [2]. The RS nodes can perform 2 the following functions [10]:

- Sensing of environmental conditions.
- Reading ID number from tag.
- Wireless connection with each other and creating ad-hoc communication network.
- They are able to transfer data to another RS node or sink node as a router.

In this architecture, RS's act as the head nodes. RS's and CHs may communicate with each other may act as relays and may transfer data from one CH or RS to another, or to the base station. [21]-[2]. 3

A. Clustering algorithm

The entire network is divided into clusters. This clustering is performed by particle swarm optimization (PSO) that Kuila and Jana proposed. The PSO can be summarized as follows which are discussed in more detail in kuila and Jana (2014) [2]. 4

- ❖ Linear programming for the routing problem based on PSO with a trade-off between transmission distance and the number of data forwards.

LN: $Fitness = W_1 \times MaxDist + W_2 \times MaxHop.$ (W_1 *and* W_2 are constant) (1)

- ❖ Nonlinear programming for the clustering problem based on PSO with a Fitness function.

NLN: $Fitness = \alpha \frac{L}{AvegDist}$ (2)

Lifetime of CHs and Average cluster distance are expressed as L and *AvegDist* respectively and α is constant.

- ❖ Energy model:

The energy required by the radio to transmit a k-bit message over a distance d is calculated as follows: $E = E_T + E_R = k \times (e_t + e_{amp} \times d^n) + k \times e_r$ (3)

Where E_T and E_R are the energy consumptions of transmitting and receiving data respectively. The energy for operating the transmitter radio, transmitter amplifier, and receiver radio is shown as e_t, e_{amp} and, e_r respectively and n is the parameter of power attenuation with ($2 \leq n \leq 4$).

V. PROPOSED ALGORITHM

When an RFID data reader reads a tag, the following data may be read by the data reader:

- ➢ EPC-Code
- ➢ Time Stamp

When the reader reads two items of data (for example A, B), a repetition occurs if the following conditions are met:

1) EPC -code A == EPC -code B.

2) $|\mathbf{T_A} - \mathbf{T_B}| < \mathbf{T}$ (Constant T is a threshold for the difference between T_A and T_B.)

The duplicate data can be divided into the following two types:

1. *The redundant data inside the cluster*: All the nodes of a cluster transfer data to their own cluster head, and after eliminating redundant data, the cluster head sends information along the routing path towards the base station.

Table II. A sample of read data.

time	*Tag-EPC*
0	*8*
1	*9*
2	*11*
3	*11*
5	*9*
7	*23*
14	*27*

```
Tag {Epc ,
        t .   / * time – stamp * /
        Sensored – data }
EXTERNAL – FILTERING ( params : Tag )
{           Pos = hash ( tag.epc.)
 If arry1( pos ).epc = = tag.epc
   Then If  | arry1 [pos].t  –  tag.t |   <    T
               Then drop it / * dublicate data * /
                Else update tag to arry1 .
    Else if arry2[ pos] . epc = = tag . epc
               Then if  | arry2[ pos].t – tag.t  |   <   T
                         Then drop it .
                         Else update tag to arry2 .
               Else if  arry1 = = 0
                     Then arry1[ pos] = tag
                     Else if  arry 2 = = 0
                              Then arry2[ pos] = tag
                               Else Wrong Position.
}
```

Algorithm 1. external – filtering step .

2. *The redundant data between two clusters*: The proposed algorithm filters these duplication data at every k hop Reader. The pseudocode of the algorithm is explained in Algorithm .1. which is explained using the read data shown in table II.

Fig. 2 shows the sequences of reading data mappings. In this example, duplication data happens when Tag ID is the same and the time difference is under 2 seconds.

VI. SIMULATION RESULTS AND DISCUSSION

In this section, simulation is used to analyze the performance of the proposed algorithm. This simulation is implemented and performed under C++ language, and as we mentioned before, the nodes have classified the shape of the cluster. The detailed simulation environment is given in table III is based on presented data in [14], but the number of the clusters has been increased (almost twofold). Consequently, the number of nodes has ascended.

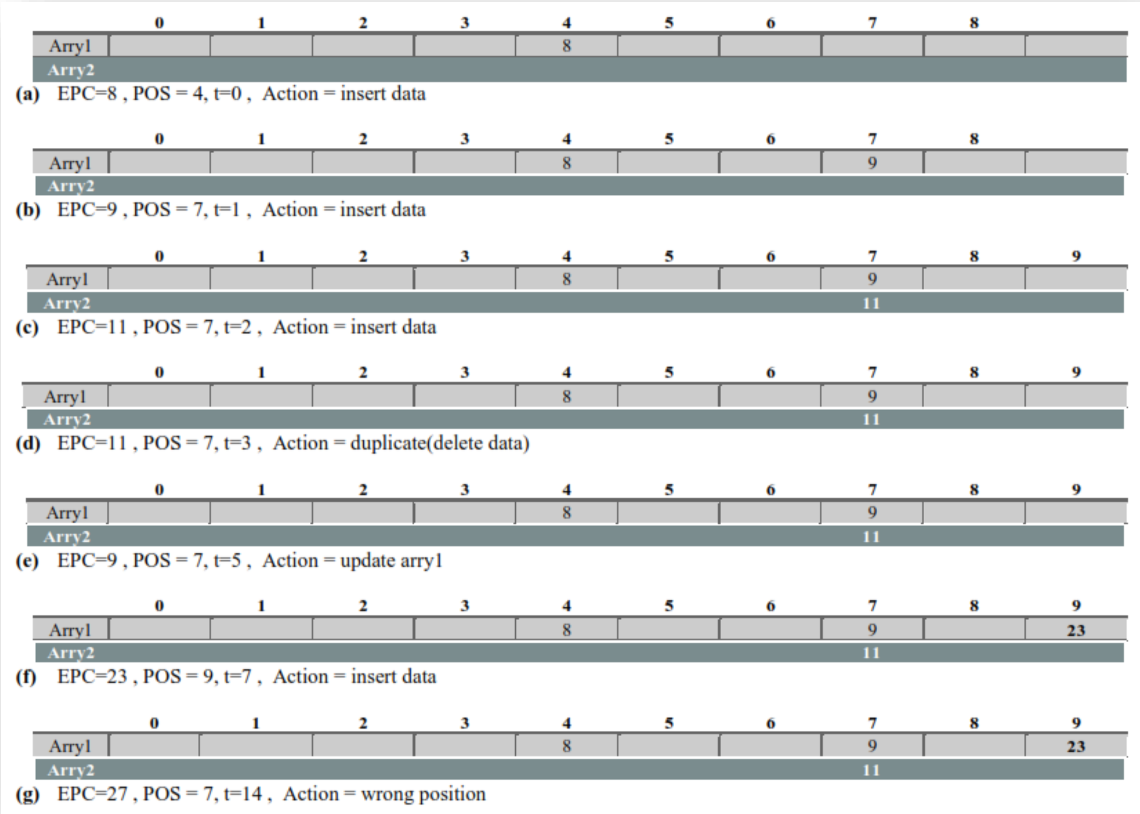

Fig. 2. the states of arrays before and after insertion of data.

Table III. Simulation environment

PARAMETERS	VALUES
FIELD AREA	$100 \times 100\ m^{2}$
NUMBER OF CLUSTER	40
MEMBER OF CLUSTER	15 - 30
NUMBER OF NODE	1000
READING RANGE	5m
TRANSMISSION RANGE	10m
READING INTERVAL	2 s
DUPLICATION RATIO	20% and 40%

A. *Wrong Position Rate (WPR)*

In this paper, WPR is the unsuccessful rate of elements in the array for a certain number of readings. Following is the equation for WPR, where “N” is the number of reading.

$$\text{WPR} = \left[\frac{FP}{N}\right]\% \tag{4}$$

In order to find the best size of the array for a certain number of “N” reading, some tests have been conducted. Fig. 3. illustrate the WPR for different number of array size i = 10000, i = 15000, i = 20000. The number of readings is changed from 1000 to 10000 with an increment of 1000 for each step. When the number of reading is 1000, WPR hit the lowest point (For all “I” values). The results show, when the array size “I” is fivefold the number of readings, the lowest WPR is achieved.

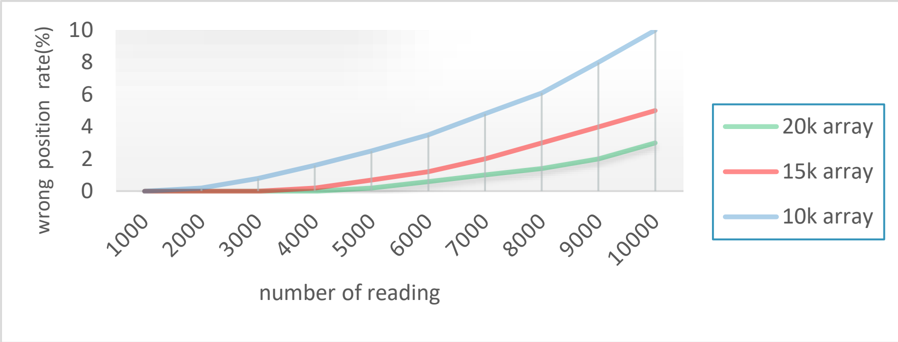

Fig. 3. Wrong Position Rate. (WPR)

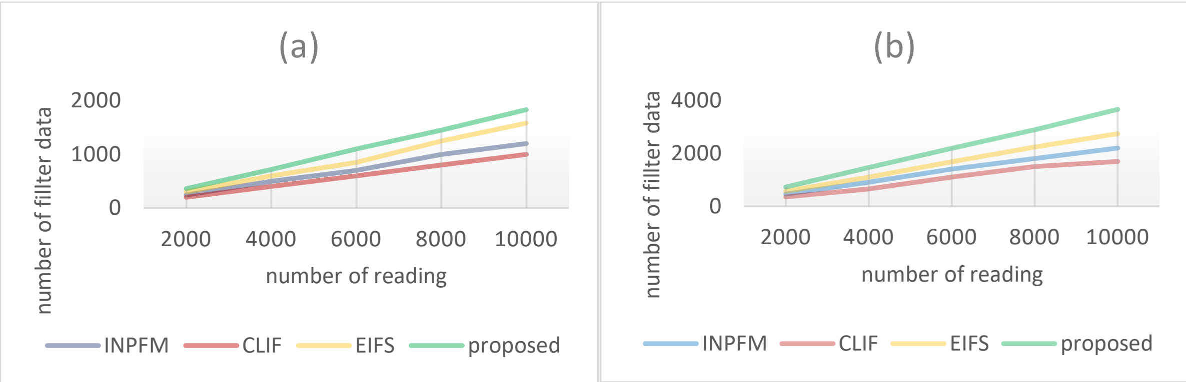

Fig. 4. Amount of data filtered (a) 20% redundant data. (b) 40% redundant data.

B. Comparative analysis of filtering performance

Use ɗtwo arrays to record information (instead of one array) that has reduced the number of wrong positions Moreover, by determining the appropriate size for the arrays (5 times the input data), the probability of losing data due to a collision reduced. Consequently, the amount of the filtered data increased.

The filtering performance of the four approaches was compared in Fig. 4. (with 20% and 40% redundant data). Our proposed algorithm filters more than other approaches (almost 92%).

C. Comparative analysis of processing time

Fig. 5 compares the processing time of the schemes to filter redundant data. The number of reading increase from 30000 to 150000 with an increment of 30000 for each step. The size of the array has been set to the maximum number of readings (150000). According to this graph, INPFM, CLIF, and EIFS have consumed more time for filtering in comparison with the proposed Scheme.

The reason for processing time reduction is that the number of comparisons in the proposed algorithm is less than the evaluated algorithms (no comparison is made to determine whether the tag-id is related

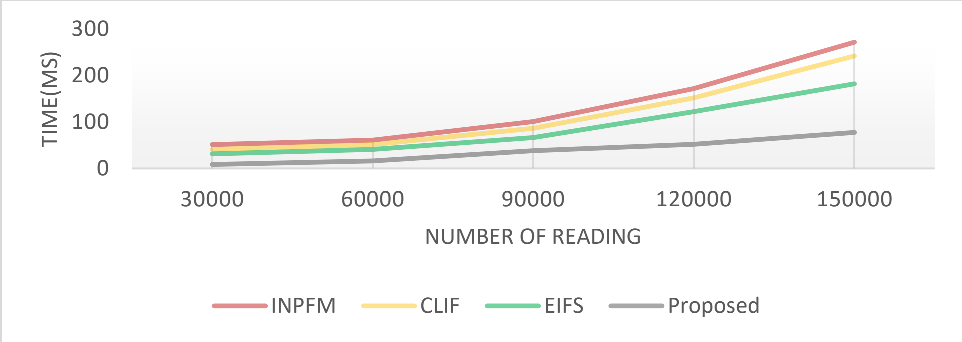

Fig. 5. Comparison of processing time for filtering redundant data.

to the cluster itself or its neighbor). Besides, duplicate information is removed near the source. As a result, the processing time has been reduced.

VII. CONCLUSION

In this study, WSN and RFID were integrated into a hybrid system that employed the advantages of both technologies. WSN and RFID can be integrated into different forms. In this paper, first, the challenges in the RFID–WSN integrated network were presented. Three schemes of INPFM, CLIF, and EIFS were investigated, and the advantages and disadvantages of each were expressed. Second, a new architecture for the RFID–WSN integrated network was proposed. The proposed network consists of heterogeneous nodes but previous works consist of homogenous nodes. Such heterogeneous networks are used in the Internet of Things. Third, the problem of redundant data in the proposed technology was considered and an algorithm was proposed to solve this problem. In this experiment, the array size was selected five times as much as the number of reading and used two arrays to record information (instead of one array), therefore the wrong position rate was very low, approximately zero. As a result, almost all duplicate data was filtered. Fourth, the algorithm has a reduced runtime since it takes only O(1) time and the number of comparisons in the proposed algorithm is less than the evaluated algorithms. Finally, our simulation results validated the efficiency of the proposed algorithm In the clustering method, owing to traffic congestion, head nodes near the base station lose their energy early (hot zone problem). In the future, the proposed scheme can be improved by providing a way to reduce the hot zone problem.

REFERENCES

[1] R. Yarinezhad, “Reducing delay and prolonging the lifetime of wireless sensor network using efficient routing protocol based on mobile sink and virtual infrastructure,” *Ad Hoc Network*, vol. 84, pp. 42–55, March 2019.

[2] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, L. H. Hanzo, “A survey of network life- time maximization techniques in wireless sensor networks,” *IEEE Commun. Surv. Tutor*, vol. 19, no. 2, pp. 828–854, Second quarter 2017.

[3] S. Jain, ”MLBC: multi-objective Load Balancing Clustering technique in Wire*less Sensor Networks,” Applied Soft Computing*. vol. 74, pp. 66–89, Jan. 2019.

[4] A. Abuelkhail, U. Baroudi, M. Raad, and T. Sheltami, “Internet of things for healthcare monitoring applications based on RFID clustering scheme,” *Wireless Networks*, vol. 27, pp.747-763, 2021.

[5] H. Landaluce, L. Arjona, A. Peraiios, F. Falcone, I. Angulo, and F. Muralter, “A Review of IOT sensing Applications and challenges using RFID and wireless sensor networks,” *Sensors*, vol. 20, no. 9, Apr. 2020.

[6] A. W. Nagpukar and S. K. Jaiawal, “An Overliew of WSN and RFID Network Integration,” *2nd International Conference on Electronics and Communication Systems (ICECS)*, 2015.

[7] G. Alfian, J. Rhee, H. Ahm, J. lee, and U. Farooq, “Integration of RFID, wireless sensor network, and data mining in an e-pedigree food traceability system,” *Journal of Food Engineering*, vol. 212, pp. 65-75, Nov. 2017.

[8] Z. Zhao, “A secure RFID authentication protocol for healthcare environments using elliptic curve cryptosystem,” *Journal of Medical System*s, vol. 38, no. 5, pp. 1–7, 2014.

[9] C. Jin, C. Xu, X. Zhang, and J. Zhao, “A secure RFID mutual authentication protocol for healthcare environments using elliptic curve cryptography,” *Journal of Medical Systems*, vol. 39, no. 3, pp. 1–8, 2015.

[10] S. Piramuthu and W. Zhau, RFID and Sensor Network Automation in the Food Industry: Ensuring Quality and Safety through Supply Chain Visibility, ch. 2, pp. 17-50, Jan. 2016.

[11] Z. Zhang, and Q. Qi, “An efficient RFID authentication protocol to enhance patient medication safety using elliptic curve cryp- tography,” *Journal of Medical Systems*, vol. 38, no. 5, pp. 1–7, 2014.

[12] B. Kochar, R. Chhillar, “An Effective Data Warehousing System for RFID Using Novel Data Cleaning, Data Transformation and Loading Techniques,” *The Intern. Arab Journal of Information Technology*, vol. 9, no. 3, May 2012.

[13] G. Kapoor, S. Piramuthu, “Single RFID tag ownership transfer protocols,” *IEEE Trans. Systems man Cybernetics Part C*, vol. 42, no. 2, pp.167-173, 2012.

[13] A. K. Bashir and S. J. Lim, ”Energy Efficient In-network RFID Data Filtering Scheme in Wireless Sensor Networks,” *Sensor Journal*, vol. 11, pp. 7004-7021, July 2011.

[14] S.E. Sarma, Towards the five- cent tag. Technical report MIT.AUTOID- WH-006, MIT Auto-ID Center: New NY , USA 2001. Available online: http://www.autoidcenter. Org/research/MIT-AUTOID-WH-006-pdf(accessed on ganuary) , 2009.

[15] W. Choi and M.S. Park, “In-network Phased Filtering Mechanism for a Large-Scale RFID Inventory Application,” In Proceedings of the 4th Intern. Conf. on IT & Applications (ICITA), Harbin, China, pp. 401-405, Jan. 2007.

[16] D.S. Kim and A. Kashif, ” Energy Effcient In-Network Phase RFID Data filtering scheme,” 5th Intern. Conf. on Ubiquitous Intelligence and Computing UIC, pp.311-322, June 2008.

[17] S. Sundaresan, R. Doss, S. Piramuthu, and W. L. Zhou, ” A robust grouping proof protocol for RFID EPC C1G2 tags,” *IEEE Trans. Information Forensics and Security* , vol. 9, no. 6, pp. 961-975, June 2014.

[18] L. Wang and L.D. Xu, ”Data cleaning for RFID and WSN Integration,” *IEEE Trans. Industrial Informatics*., vol. 10, No. 1, pp. 408-418 Feb. 2014.

[19] Zhang/RFID and Sensor Networks, AU7777_c018, May 2009, ch 18, pp.511-535. Available on internet.

[20] S. Sundaresan, R. Doss, S. Piramuthu, and W. Zhou, ”Secure ownership transfer for multi-tag multi-owner passive RFID environment with individual-owner privacy,” *Computer Communications*, vol. 55, pp.112-124, Jan. 2015.

[21] M. M. Aung, Y.S. Chang and J. Won, “Emerging RFID /USN Application and Challenge,” In Intern. Journal of RFID Security and Cryptography (LJRFIDSC), vol. 1, March-June 2012

[22] X. Wang, Y. Ji, and B. Zhao, “ An approximate duplicate- elimination in RFID data streams based on d-left time bloom filter,” in Web Technologies and Applications, vol. 8709, pp. 413–424, Sept. 2014.

[24] P. kuila and K. Jana, “Energy efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach,” *Eng. Appl. Artificial Intelligence*, vol. 33, pp.127-140 , Aug. 2014.