Data Mining Approaches for Correlation and Cache Replacement in Wireless Ad Hoc Networks

Sudha Rani Y J
JNTUH, SMGOI, Hyderabad
Telangana, India
{su.joyfull@gmail.com}

Seetha M
GNITS, Hyderabad
Telangana, India
{smaddala2000@yahoo.com}

ABSTRACT: *Wireless Ad Hoc network is a collection of wireless nodes that are self-configured without fixed infrastructure. In such nodes caching is the mechanism that makes use of high speed memory known as cache memory. This memory holds frequently used data so as to reduce data transmission. It also leads to better data management. It is very useful especially in ad hoc networks that do not have fixed infrastructure. Many researchers contributed towards better data management through caching and cache replacement techniques. Base line algorithms like LRU and NRU need to be improved for better performance. Data mining technique approaches like Association Rule Mining (ARM) can be used to have better performance in cache coordination and cache replacement. The FP-Growth algorithm is used for this purpose by some researchers. However, this algorithm consumes more resources. Therefore, a light weight algorithm for discovering association rules is needed. In this paper we used FIN ARM as part of our methodology for better data management and cache replacement. Our simulation study resulted in better performance when compared with LRU and FP-Growth algorithms.*

Keywords: Wireless Ad Hoc Network, Caching, Cache Replacement, Association Rule Mining

Received: 28 May 2020, Revised 14 August 2020, Accepted 20dli August 2020

DOI: 10.6025/jnt/2020/11/4/129-135

Copyright: with Authors

1. Introduction

Wireless Ad Hoc network is the network with collection of nodes without having fixed infrastructure. On the other hand some wireless networks are with fixed infrastructure. The difference between them is shown in Figure 1. Ad hoc networks have important utility in the real world applications. Ad hoc networks have the feature known as caching. Caching helps nodes to store data that might be needed frequently in order to provide better data management. There are schemes like Least Recently Used (LRU) and Most Recently Used (MRU) techniques for cache replacement. However, these are baseline algorithms that need improvement. Therefore, in this paper, we studied data mining techniques that are widely used for different applications. Especially our focus was to use association rule mining in order to have better data management in ad hoc networks.

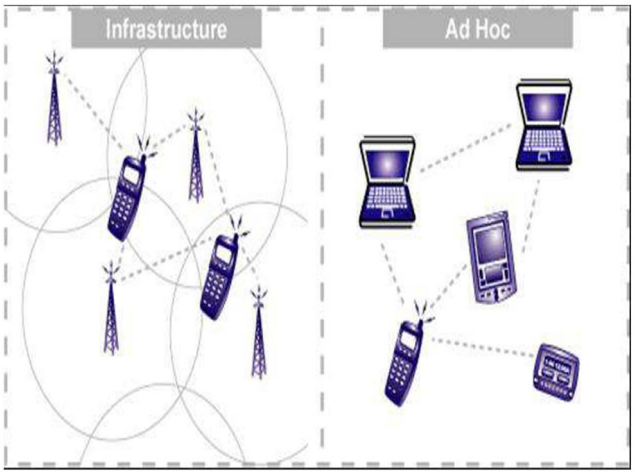

Figure 1. A typical difference between wireless ad hoc network and other network

As shown in Figure 1, it is evident that ad hoc network is a network of wireless devices. The devices are automatically configured to form a network without any fixed infrastructure. The network is formed on demand without having pre-defined infrastructure. On the other hand the other kind of wireless network (shown left in Figure 1), there is need for infrastructure for forming a network.

In the ad hoc networks caching helps in faster query processing and better data management. There is provision for reusing data which is frequently used instead of transmitting it every time. To process a query, it is possible to reuse data in cache memory. Then it is known as cache hit. When cache hit ratio is more, it is said to have better performance. Many researchers contributed towards caching and cache replacement techniques. Section 2 throws light on this. Our contribution in this paper is to have a methodology that makes use of FIN ARM for generating knowledge that can help in making cache replacement decisions. The remainder of the paper is structured as follows. Section 2 provides review of literature. Section 3 presents the proposed system in detail. Section 4 presents implementation details. Section 5 shows experimental results while section 6 concludes the paper.

2. Related Works

This section provides review of literature on different aspects of caching and cache replacement. Information density based cache replacement techniques are explored in [1] and [5]. In [2] caching techniques are used to improve information retrieval performance by using neighbour group data caching. Counter based cache replacement technique is used explored in [3]. Information solidity evaluation based approach for caching strategy is used in [4] for resource efficient access of information. On-demand cache based routing was used in [6] for improving performance of wireless ad hoc networks. Cooperative caching strategies with different case studies are presented in [7] while data caching schemes with coherence are proposed in [8]. Cache discovery over a network with multiple hops is presented in [9] with power control and better cooperation.

Cooperative caching and its implementation is done in [10] for social wireless networks that are formed by wireless devices. Different caching techniques are explored in [11], [12], [13], and [14] for efficient data management and information retrieval. FP-tree [15] is used to have association rule mining. This algorithm is used for generating association rules based on given support and confidence. The association rules are used in the Ad Hoc nodes to have better data management. However, the FP-tree based algorithm consumes more resources. It needs further improvement as the ad hoc networks are resource constrained. To overcome the drawbacks of existing algorithms that are based on FP-tree [15], we proposed a new methodology based on FIN ARM which provides more light weight processing for improving performance. The cache hit ratio is improved with the proposed system.

3. Proposed Methodology

We proposed a methodology for correlation and cache replacement. Towards this end, we proposed a framework as shown in

Figure 2. There is query manager to take care of queries of nodes and processing them. The query processing is done by hitting cache in order to perform data transmission faster. This is the case in general. Sometimes, when data is not found the cache, it needs to be transmitted explicitly causing more overhead. Improving cache hit ratio is the main aim of the proposed methodology. Cache is the high speed memory that holds frequently used data which is reused by nodes in ad hoc networks. Knowledge DB is the database containing data that can be updated by using query processing results. The FIN association rule mining algorithm has light weight processing which can help query manager to perform well and the overall system will get benefits of correlation among data items and cache replacement in an effective manner.

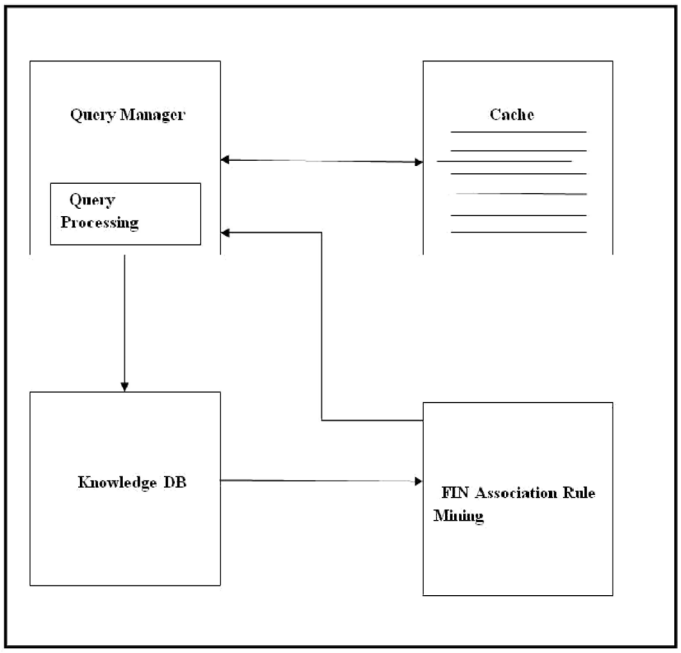

Figure 2. Proposed methodology based on FIN ARM

The proposed system is based on FIN algorithm which is more efficient as it reduces complexity in producing frequent item sets. FIN was taken from [16]. FIN makes use of POC tree for storing data that is used for association rule mining. This tree structure also make discovery of frequent items faster. This will reflect in the accurate decision making with respect to cache replacement. In order to illustrate the tree and its structure, Table 1 is used to represent data in the form of POC tree.

ID	Items	Ordered Frequent Items
1	a, c, g, f	c, f, a
2	e, a, c, b	b, c, e, a
3	e, c, b, i	b, c, e
4	b, f, h	b, f
5	b, f, e, c, d	b, c, e, f

Table 1. Sample data

As shown in Table 1, it is evident that the POC tree is generated from the data presented and the tree appears as follows.

As shown in Figure 3, the POC tree is represented with given data. This tree helps faster processing of data when compared with traditional tree-based data structures. The framework in this paper makes use of it with respect to FIN association rule mining.

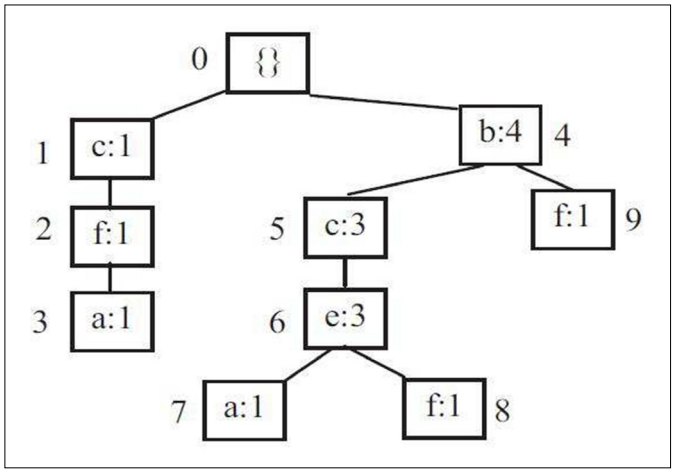

Figure 3. Generated POC-tree based on the data of Table 1

Algorithm: *FIN ARM*

Inputs: Dataset D, support sup, confidence conf

Output: Casual rules

01 Initialize vector *POC* to hold POC tree

02 Initialize vector *AR* to hold association rules

03 Initialize vector *CR* to hold casual rules

04 Construct POC from *D*

05 Find frequent 1-itemsets

06 Scan POC tree for finding frequent 2-itemsets

07 Mine all frequent (>2) itemsets

08 Mine AR rules that satisfy sup and conf

This algorithm takes data, support and confidence as inputs and generates association rules that are used to make cache replacement decisions. The support and confidence are the statistical measures that are used to have good quality association rules.

4. Experimental Results

Experiments are made with simulations done using a custom simulator. The results are observed in terms of cache hit ratio with different number of nodes in Ad Hoc network. LRU, FP-tree mining and proposed methodology are compared for performance evaluation.

As shown in Table 1, the results are presented. The results contain the hit ratio of LRU, FP-tree based ARM and FIN based ARM (proposed).

As shown in Figure 4, it is evident that there is hit ratio on different algorithms including the proposed algorithm.

As shown in Table 3, it is evident that there are results related to hit ratio with different number of nodes.

As shown in Table 5, it is evident that the hit ratio of the proposed system is better than other algorithms.

Proposed	Fp-Tree ARM	LRU
0.15	0.13	0.11
0.25	0.2	0.15
0.34	0.3	0.21
0.29	0.2	0.21
0.34	0.23	0.25
0.25	0.15	0.2
0.2	0.17	0.15
0.19	0.23	0.17
0.28	0.25	0.2
0.29	0.23	0.23
0.27	0.23	0.25
0.32	0.26	0.2
0.18	0.24	0.1
0.29	0.13	0.21
0.26	0.21	0.23
0.17	0.21	0.11
0.18	0.13	0.14
0.16	0.16	0.12

Table 2. Cache hit ratio comparison

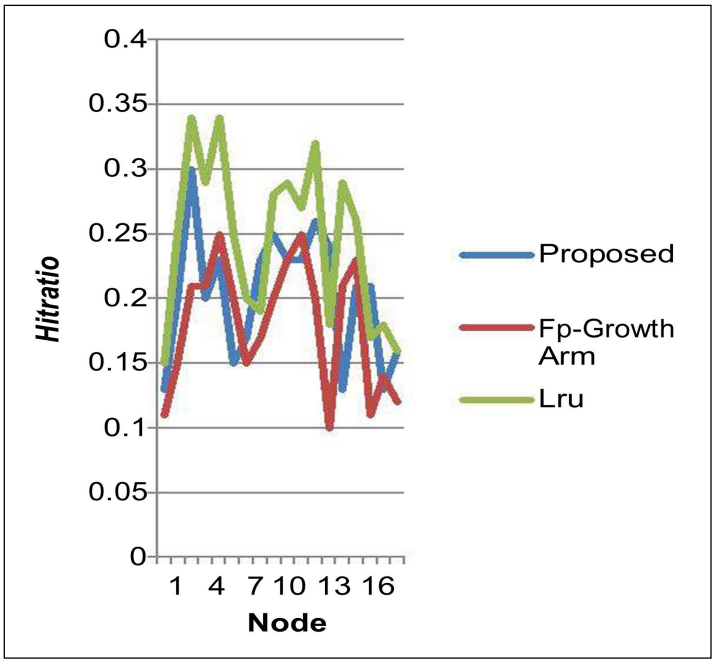

Figure 4. Cache hit ratio comparison

Proposed	Fp-Growth ARM	LRU
0.18	0.14	0.11
0.24	0.2	0.14
0.28	0.27	0.14
0.29	0.25	0.2
0.2	0.26	0.18
0.19	0.2	0.17
0.23	0.17	0.18
0.16	0.15	0.1
0.29	0.17	0.27
0.12	0.13	0.08
0.25	0.25	0.19
0.29	0.27	0.23
0.28	0.26	0.22
0.18	0.15	0.11
0.19	0.21	0.11
0.28	0.14	0.21
0.29	0.24	0.19
0.25	0.28	0.19

Table 3. Cache hit ratio comparison

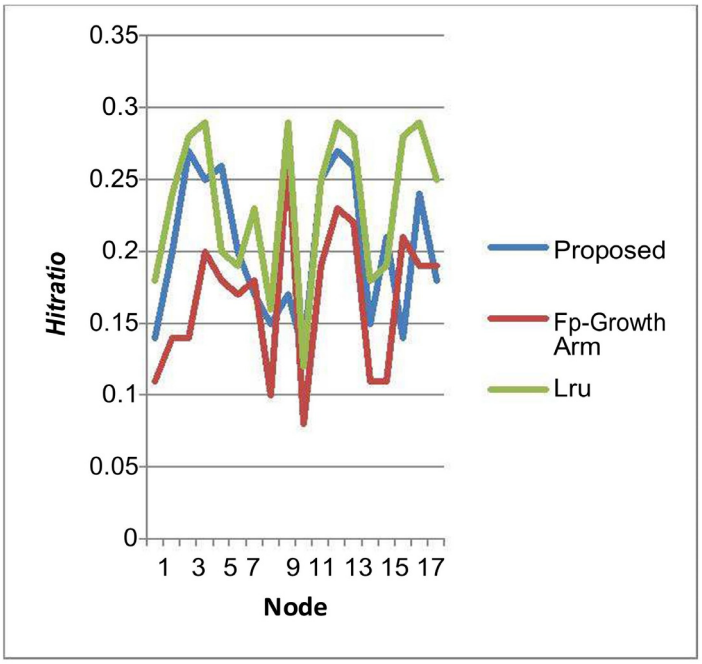

Figure 5. Hit ratio for different algorithms with different number of nodes

5. Conclusion and Future Work

In this paper, we studied the concept of caching and cache replacement in wireless ad hoc networks. It is understood that cache replacement needs improved algorithms. The traditional algorithms like LRU and NRU are not sufficient to have better data management. The FP-tree based solutions came into existence and they are suffering from the problem of resource consumption. Therefore we proposed a methodology with an underlying algorithm. The algorithm is based on light weight FIN ARM which is efficient. Thus the proposed algorithm is able to improve cache replacement with more accurate decisions. The cache hit ratio is improved further. The results of the proposed system are compared with FP-tree and base line algorithm such as LRU. The results revealed that the proposed system performs better. This research can be extended further by implementing caching mechanisms in real ad hoc networks.

References

[1] Vimala, S., Anbarasu, R. (2012). Efficient Cache Replacement Techniques Based on Information Density over Wireless Adhoc Networks. *ISSN*. 2 (3) 209-214.

[2] Shanmugavadivu, K., Madheswaran, M. (2010). Caching Technique for Improving Data Retrieval Performance in Mobile Ad Hoc Networks. *IJCSIT*. 1 (4) 249-255.

[3] RadhaVasan, A., Vinoth, N., Rahuman, Abdul., N., Silambarasan, S. (2014). Wireless Ad hoc Network using Counter Based Cache Replacement Algorithm. *JARCCE*. 3 (1) 5039-5042.

[4] Praksah, A., Durga, K. B. K. S., Divya, P., Jayanth, B. (2014). Caching Strategies Based on Information Solidity Evaluation in Wireless Ad Hoc Networks. *IJARCSSE*. 4 (10) 123-129.

[5] Nagaraju, D., Srinivasa rao, L., Nageswara Rao, K. (2012). Caching Strategies Based on Information Density Estimation in Wireless Ad Hoc Networks. *IJCER*. (6) 139-146.

[6] Dhanashri., Dhobale., D., Ghorpade, V. R. (2011). Wireless Ad-Hoc Networks with On-Demand Cache Routing. *IJCTE*. 3 (6) 785-789.

[7] Rao, Atul., Kumar, Prashant., Chauhan, Naveen. (2011). Cooperative Caching Strategies for MANETs and IMANETs. *ICACCT*, 89-95.

[8] Mandhare, V. V., Thool, R. C. (2014). Improving Performance *In*: Mobile Ad-hoc Network Using Data Caching Schemes: A Review. *IJCTA*. 5 (5) 1780-1788.

[9] Theresa Joya, Preetha., Jacob, Poulose., K. (2014). Cache Discovery Over a Multihop Wireless Ad Hoc Network. *IJIP*. 8 (2) 71-80.

[10] Suganya, S. L., Gandhi, Indra., R. (2014). Implementation of Cooperative Caching in Social Wireless Networks. *IJECS* 3 (5) 6147-6152.

[11] Zam, Abdulaziz., Movahedinia, N. (2013). Performance Improvement of Cache Management *In*: Cluster Based MANET. *IJCNIS*. 10, 24-29.

[12] Lin Yao , Jing Deng., Jie Wang., Guowei Wua. (2015). A-CACHE: Ananchor-based public key caching scheme in large wireless networks. *ELSEVIER*. 87 (0) 78–88.

[13] Badru, D., Deepthi, P., Sankaraiah, B. (2017). Analysis on Intrusion & Detection of Sybil Attacks in Mobile Adhoc Networks Using Classification. *IJARF*. 4 (3) 1-5.

[14] Hallousha, Rami., Liub, Hang., N., Dongc, Lijun., Wud, Mingquan., Radhae, Hayder. (2017). Hop-by-hop Content Distribution with Network Coding in Multihop Wireless Networks. *ELSEVIER*. 3 (0) 47–54.

[15] Islam, Shaikh. (2010). Exploring Correlation Among Data Items for Cache Replacement in Ad-hoc Networks. IEEE, p. 1-4.

[16] Deng, Z., Lv, S. (2014). Fast mining frequent item sets using Nodesets. Elsevier, Expert Systems with Applications, 41, 4505-4512.