

## A RFID Data-Cleaning Algorithm Based on Communication Information among RFID Readers

Yunhua Gu, Bao Gao, Jin Wang, Mingshu Yin and Junyong Zhang

*Jiangsu Engineering Center of Network Monitoring, School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing, 210044, China, {yhgu6655@163.com, gtb900206@163.com, wangjin@nuist.edu.cn}*

### **Abstract**

*For the insufficient of existing algorithms of redundant-reader elimination, the paper proposes a data-cleaning algorithm based on communication information among readers. This algorithm is no longer dependent on the tag to send information between the RFID readers, but to transmit information directly through communication information among the readers to determine which reader the tag belongs to and whether the reader is redundant or not. This paper does a lot of experiments to verify the effectiveness of the algorithm.*

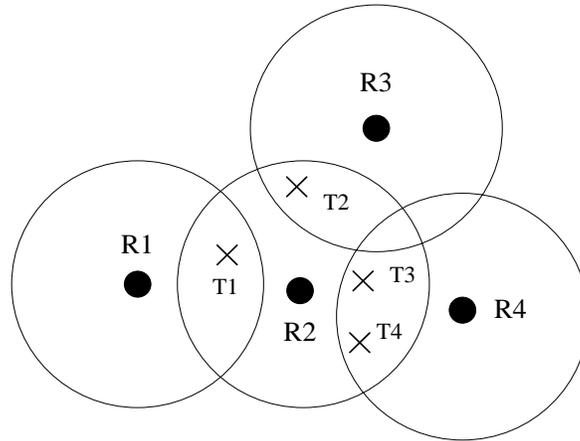
**Keywords:** *RFID, redundant reader, reader communication*

### **1. Introduction**

RFID [1] (Radio Frequency Identification, RFID) is a non-contact automatic identification technology. It uses radio frequency signals to access to the relevant data of objects, and to identify the objects. RFID technology enables to provide data information for the application real-time and efficiently, so it is widely used in various fields. However, the readers may be densely deployed in some practical applications, so the detection range of the RFID readers overlap, which produces a lot of redundant data and limits the large-scale implementation of an RFID system. So before the original data was sent to the high-level applications, the redundant data must be cleaned.

The redundant data in the RFID systems can be divided into two categories: the redundant reader and the redundant tag data. The redundant reader belongs to spatial redundancy, which means that the detection range of the readers overlap spatially and result in redundant data. This is to say, in an RFID system of readers densely deployed, all the tags in the detection range of a reader be read by at least one other reader. As shown in Figure 1. The black dots represent the readers, and the circles represent the detection range of the readers. In Figure 1, the detection range of the four readers overlap with each other, and each tag can be read by two readers or more at the same time. In fact, we only need the reader  $R_2$ , because it can be read all of the tags. So the reader  $R_1$ ,  $R_3$  and  $R_4$  are considered to be redundant readers. The redundant tag data belongs to temporal redundancy, which emphasize that a tag is read many times by the same reader, thus a large number of duplicate records emerge.

This paper studies the problem of the redundant reader, which is an important reason to affect the performance of an RFID system. And our purpose is to detect the maximum redundant readers and we can turn off them simultaneously.



**Figure 1. Schematic Diagram of Reader Redundancy**

There are some existing classical methods of redundant reader elimination, such as RRE [2] (Redundant Reader Elimination) algorithm, proposed by Carbutar *et al.*, and LEO algorithm [3] (Layered Elimination Optimization), proposed by Hsu *et al.* RRE algorithm is an approximation algorithm that it can extend the lifetime of RFID reader network by preserving network coverage and eliminating redundancy. A significant improvement of the LEO algorithm is that amount of “write-to-tag” operations could be largely reduced during the redundant reader identification phase. And LEO algorithm is a distributed approach which does not need to collect global information for centralizing control, leading to no communications or synchronizations among RFID readers. There are also some improved algorithms based on the above algorithm [4-7].

The rest of this paper is organized as follows. In Section 2, this paper will firstly study and analyze the RRE algorithms and LEO algorithm, and then point out their deficiencies. An algorithm based on RFID reader communication to eliminate the redundant reader will be proposed and discussed in Section 3. This algorithm requires the reader can directly communication with its adjacent readers. The reader that holds the maximum number of tags will lock the tag in its detection range and then notify its adjacent readers. The adjacent reader eliminates the data if the tag be locked by other reader. Until all tags have been locked, the readers that do not lock the tag are redundant reader. Section 4 gives the performance analysis and simulation comparisons. Finally, some concluding remarks are made in Section 6.

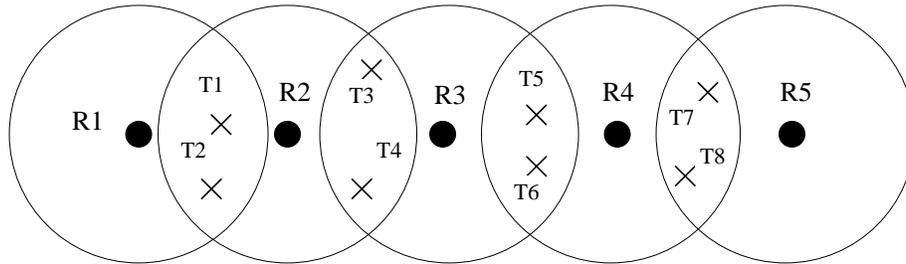
## **2. Redundant Reader Elimination Algorithm and Its Analysis**

RRE algorithm can preserve network coverage and eliminate redundancy and the specific steps of RRE algorithm are divided into three steps.

- 1) Each reader detects the tags in their detection range;
- 2) Each reader statistics the number of tags in its detection ranges, and write its tag count and reader identifier on to all its covered tags.
- 3) The reader who holds the maximum number of tags locks the tag that in its detection range. Until all tags are locked, the readers that don't lock tags are considered to be redundant reader.

As shown in Figure 1, the reader  $R_2$  holds the maximum number of tags, so the reader  $R_2$  lock all the tags covered by  $R_2$ . After that, all the tags are locked, so the reader  $R_1$ ,  $R_3$ ,  $R_4$ , which do not lock any tag, are considered to be redundant reader.

RRE algorithm solves the problem that multiple readers read a tag simultaneously and upload the data to the data center, which results in redundancy. But RRE has some shortcomings. First, the RRE algorithm requires the readers to write to the tags in their detection range continuously, which will result in high time complexity. If there is an RFID system that has  $n$  tags and  $m$  readers, and the readers can read the tags and write to the tags. So the complexity of writing operation of the reader is  $O(nm)$ . Second, RRE algorithm assumes that the relative position of the reader and the tag does not change over a long period of time, but this assumption is not established for RFID applications that have the dynamic tags. Third, in some scenarios of RFID systems, readers whose interrogation zones overlap equal numbers of tags. In this case, RRE algorithm does not obtain the optimal results. As shown in Figure 2.



**Figure 2. Readers whose Interrogation Zones Overlap Equal Numbers of Tags**

In Figure 2, the reader  $R_2, R_3, R_4$  all have 4 tags in their detection range. According to the RRE algorithm, the three readers will be all reserved and activated, and the readers that are considered to be redundant are the reader  $R_1$  and  $R_5$ . But it is not difficult to find that we only need the reader  $R_2$  and  $R_4$  to be reserved, so  $R_3$  is also a redundant reader.

LEO algorithm makes the improvements to the RRE algorithm for its insufficiency. To reduce the time complexity of the RRE algorithm, LEO algorithm adopts the approach that the first reader communicate with the tag will lock the tag. The time complexity of the process drops from  $O(nm)$  to  $O(n)$ , but the LEO algorithm depends on the order that reader reads the tag, which leads to unstable results and the reliability is poor. For Figure 1, the LEO algorithm does not obtain the optimal results. As shown in Table 1.

**Table 1. LEO Aalgorithm (Figure 1)**

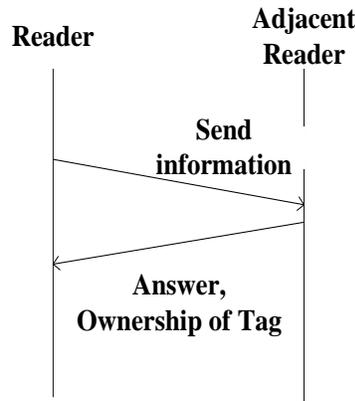
	T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>	T <sub>4</sub>
R <sub>1</sub>	R <sub>1</sub>			
R <sub>2</sub>		R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>
R <sub>3</sub>				
R <sub>4</sub>				
Final	R <sub>1</sub>	R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>
Redundant Reader		R <sub>3</sub> , R <sub>4</sub>		

### 3. A Data-Cleaning Algorithm Based on Communication Information among Readers

In this paper, according to the insignificance of RRE algorithms and LEO algorithm, an algorithm is proposed based on the RFID reader communication redundancy to eliminate redundant reader.

### 3.1. Algorithm Prerequisite

- (1) The Topology of RFID system is not limited.
- (2) Before using this algorithm, the tag conflicts and readers conflicts in the RFID system have been resolved.
- (3) The adjacent readers can communicate with each other directly in the RFID system [8]. This paper will present a communication protocol between the adjacent readers, as shown in Figure 3, and we will make a explanation for it.



**Figure 3. The Communication Protocol of the Adjacent Readers**

- 1) The reader sends information to its adjacent reader, and the format of the information is <reader identifier, tag identifier, tag count>.
- 2) If the adjacent readers also contain a tag that the reader sent to the adjacent readers in its detection range, the adjacent reader will answer. Depending on the tag count within the detection range to determine which reader can lock the tag, and thus we can judge the redundant reader and the reserved reader.
- 3) The readers that are marked as redundant reader and reserved reader are no longer involved in the response.

### 3.2. Specific Steps of Proposed Algorithm

Based on 3.1, we will give the flow diagram and the specific steps of the algorithm that we propose. The flow diagram is shown in the Figure 4.



**Figure 4. The Flow Diagram of Algorithm Based on the RFID Reader Communication to Eliminate Redundant Reader**

Step 1: Each reader detects tags within its detection range, and statistics the number of tags within its detection range.

Step 2: According to the sequence of the tag identifier, we select one of the readers that hold the first tag to send information to its adjacent readers firstly, and the format of the information is <reader identifier, tag identifier, tag count>.

Step 3: Determine whether the tag is read by the adjacent readers simultaneously. If it is read, go to step 4, and if it is not read, go to step 5.

Step 4: The adjacent readers that contain the tag will answer. If they are marked as redundant reader or reserved reader, they do not make response although they cover the tag. The information of responses depends on the tag count within their detection range, and determines the ownership of the tag in turn. If the tag count of the reader is smaller than the tag count of adjacent reader, this tag is be considered to belong to adjacent reader, and the tag count of reader subtract one; If the tag count of the reader is greater than or equal the tag count of adjacent reader, the reader is the owner of this tag and the tag count of the adjacent reader subtract one.

Step 5 if the tag is not included in any adjacent reader, it means that the tag is read only by the reader, so the reader is certainly not redundant reader and need to be reserved. If this reader also cover other tags, the reader send information to inform the adjacent readers that it is the owner of these tags, so the adjacent readers will update their tag count.

Step 6 Once the tag count of a reader becomes 0, it is marked as redundant reader. If not , we continue to send information to its adjacent from the next tag according to the sequence of the tag identifier.

Moreover, if the relative position of the reader and the tag changes, the reader will send a message to its adjacent readers, and go back to step one. Therefore we can redefine the dynamic changes of ownership of the tag.

### 3.3 Analysis of Algorithm Instance

Take Figure 2 for example, we make an algorithm analysis of the instance. Firstly we give an algorithm instance analysis table, as shown in Table 2. And then we explain the procedure in detail.

**Table 2. LEO Algorithm (Figure 2)**

	T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>	T <sub>4</sub>	T <sub>5</sub>	T <sub>6</sub>	T <sub>7</sub>	T <sub>8</sub>
R <sub>1</sub>								
R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>				
R <sub>3</sub>								
R <sub>4</sub>					R <sub>4</sub>	R <sub>4</sub>	R <sub>4</sub>	R <sub>4</sub>
R <sub>5</sub>								
Final	R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>	R <sub>2</sub>	R <sub>4</sub>	R <sub>4</sub>	R <sub>4</sub>	R <sub>4</sub>
Redundant Reader		R <sub>3</sub> , R <sub>4</sub>						

Tag T<sub>1</sub> and T<sub>2</sub> are in the detection range of the reader R<sub>1</sub>. R<sub>1</sub> will send information <R<sub>1</sub>, T<sub>1</sub>,2> and <R<sub>1</sub>, T<sub>2</sub>,2>to its adjacent reader R<sub>2</sub>. Tag T<sub>1</sub> and T<sub>2</sub> are also in the detection range of R<sub>2</sub>, and the tag count of R<sub>2</sub> is 4, so tag T<sub>1</sub> and T<sub>2</sub> belong to R<sub>2</sub>. The tag count of R<sub>1</sub> becomes 0, and the reader R<sub>1</sub> is marked as redundant reader. At the same time R<sub>2</sub> is determined to be reserved, and the tag T<sub>3</sub> and T<sub>4</sub> are in the detection range of R<sub>2</sub>,so they also belong to R<sub>2</sub>. Next , R<sub>2</sub> sends information to its adjacent reader to inform that it is the owner of tag T<sub>3</sub> and T<sub>4</sub>, so the tag count of the adjacent reader R<sub>3</sub> subtract 2. R<sub>3</sub> continues to send information <R<sub>3</sub>, T<sub>5</sub>,2> and <R<sub>3</sub>, T<sub>6</sub>,2>to its adjacent readers. R<sub>2</sub> is a reserved reader and do not need to answer. R<sub>4</sub> covers tag T<sub>5</sub> and T<sub>6</sub>, and the tag count of R<sub>4</sub> is 4, so the tag T<sub>5</sub> and T<sub>6</sub> belong to the reader R<sub>4</sub>. This time the tag count of R<sub>3</sub> becomes 0, so R<sub>3</sub> is marked as redundant reader. And R<sub>4</sub> is a reader determined to be reserved.R<sub>4</sub> also contains the tag T<sub>7</sub> and T<sub>8</sub>, therefore R<sub>4</sub> sends the information to its adjacent readers to inform that it is the owner of T<sub>7</sub> and T<sub>8</sub>. The tag count of R<sub>5</sub> subtract 2. And then the tag count of R<sub>5</sub> becomes 0, then R<sub>5</sub> is marked as redundant reader.

### 3.4. Algorithm Analysis

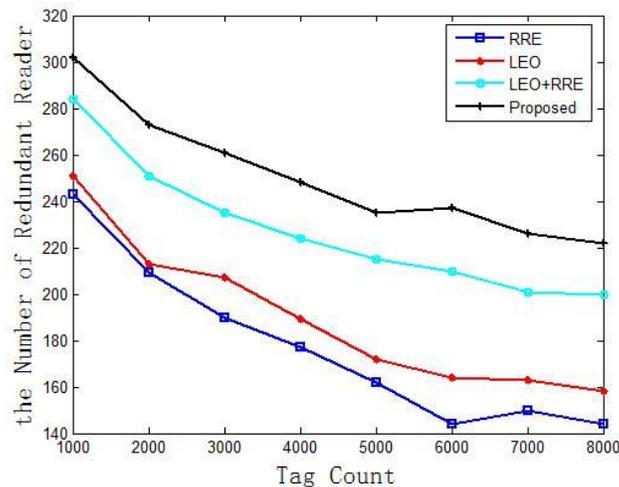
The proposed algorithm in this paper base on reader communication directly, so the readers don't write on tags. The time complexity of algorithm is reflected in the exchange of communication between the readers. The reader through an interaction to determine the ownership of each tag. If an RFID system with  $n$  readers, the time complexity of the algorithm is  $O(n)$ . Compared to the RRE algorithm, the time complexity of the algorithm can be reduced, at the same time the algorithm is more stable than LEO algorithm. This is to say, the reliability of the algorithm is better. And the algorithm also overcomes the problem that LEO algorithm and RRE algorithms require the unchanged relative position of the reader and tag for a long time. So the proposed algorithm can better determine the ownership of dynamic tag and be adapted to the practical application of RFID systems.

## 4. Experiments

In this paper, the experimental environment is generated by RFID systems network based on simulation and experimental parameters are set consistent with the literature [2].

Experiment One:

The experimental range is  $1000 \times 1000m^2$ , 500 readers and 1000-8000 tags are randomly deployed in this range . The results of each algorithm to eliminate the redundant readers as the change of the number of tags are shown in Figure 5. The algorithm proposed in this paper has high redundant reader detection rates than RRE algorithm, LEO algorithm and LEO + RRE algorithm. The main reason is that the RRE algorithm, LEO algorithm and LEO + RRE algorithm has the problem of misjudgment. But the proposed algorithm does not have the problem.

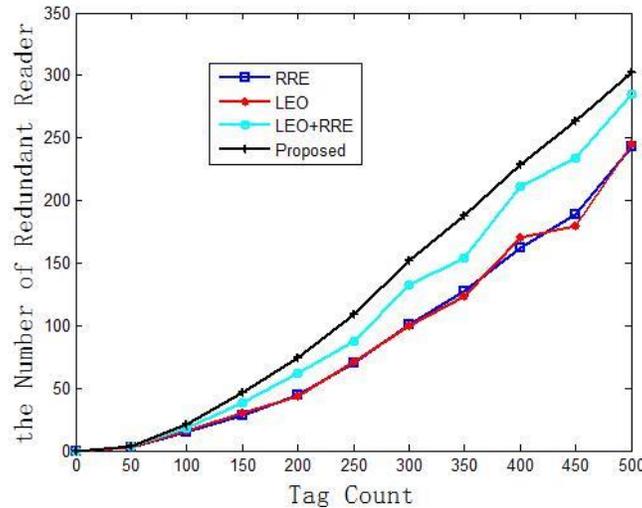


**Figure 5. The Results of Each Algorithm to Eliminate the Redundant Reader, when the Number of Readers is Constant, and the Number of Tags is Change**

Experiment Two:

The experimental range is  $1000 \times 1000m^2$ , 0-500 readers and 1000 tags are randomly deployed in this range . In the process of the number of readers changing from 0 to 500 , the number of redundant readers detected by the RRE algorithm, LEO algorithm, LEO + RRE algorithm and the algorithm proposed in this paper is also increasing respectively, as shown in Figure 6. The proposed algorithm is almost the same with the other three algorithms, when the number of readers is less than 100. But when the number of readers is more than 100, the

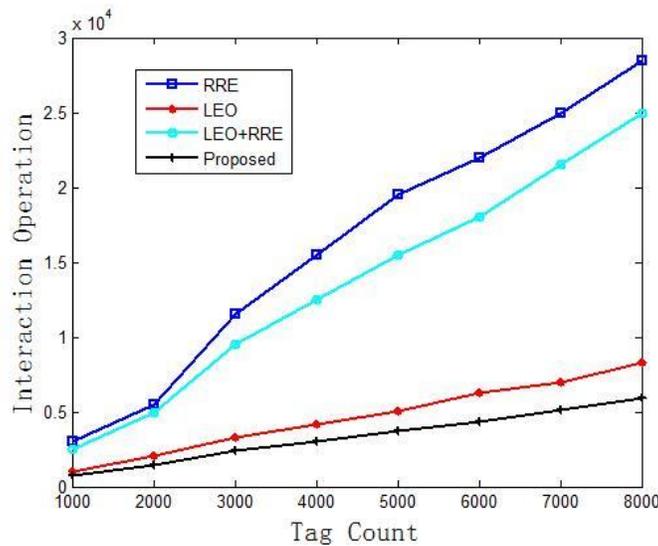
number of redundant readers detected by the algorithm proposed in this paper begins to appear obvious growth .And with further increase in the number of readers, redundant readers detected by the algorithm proposed in this paper increase corresponding .



**Figure 6. The Results of Each Algorithm to Eliminate the Redundant Reader, when the Number of Tags is Constant, and the Number of Readers is Change**

Experiment Three:

The experimental range is  $1000 \times 1000m^2$ , 500 readers and 1000-8000 tags are randomly deployed in this range. The purpose of this experiment is to compare the time complexity of



**Figure 7. Comparison of the Time Complexity of Each Algorithm**

The algorithm proposed in this paper with the RRE algorithm ,the LEO algorithm and the LEO + RRE algorithm, as shown in Figure 7. We select interaction operation to represent the time complexity. The interaction operation is the operation of writing information to the tags by readers for the RRE, LEO and LEO+RRE algorithm, and for the algorithm proposed in

this paper, it is the number of interactions of information among the readers. LEO algorithm is significantly better than the RRE algorithm and LEO+RRE algorithm in terms of reducing the writing operation. But the three algorithms are inseparable from the operation of writing information to the tags by readers. By means of direct reader communication, the algorithm proposed in this paper reduces the time complexity, and is more stable than LEO algorithm.

## 5. Conclusions

This paper proposes an algorithm based on the RFID reader communication to eliminate the redundant reader directly. The readers do not need to run the operation of writing information to the tags, but exchange the information among the readers directly in this algorithm. Experimental results show that the algorithm reduces the time complexity greatly, and increases the rate of detection of redundant reader, and overcome the problem of the change in the relative position between the reader and the tag. Eliminating redundant reader can optimize the deployment of RFID systems, and then the system performance can be improved.

## Acknowledgements

The work was supported by the National Natural Science Funds research project (61103142). It was also supported by Industrial Strategic Technology Development Program (10041740) funded by the Ministry of Trade, Industry and Energy (MOTIE) Korea. Prof. Yunhua Gu is the corresponding author.

## References

- [1] R. Derakhshan, M. E. Orłowska and X. Li, "RFID Data Management Challenges and Opportunities [C]", IEEE First International Conference on RFID, (2007), Grapevine, Texas, USA.
- [2] B. Carbutar, M. K., Ramanathan and M. Koyuturk, "Redundant Reader Elimination in RFID Systems [C]", In Proceedings of the 2nd Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, (2005), Santa Clara, Calif., USA.
- [3] C.-H. Hsu, Y.-M. Chen and H.-J. Kang, "Performance-effective and low-complexity redundant reader detection in wireless RFID networks [J]", EURASIP Journal on Wireless Communications and Networking, vol. 22, (2008).
- [4] I. Nazish and M. C. E. Yagoub, "Efficient algorithm for redundant reader elimination in wireless RFID networks [J]", International Journal of Computer Science Issues, vol. 7, no. 3, (2010).
- [5] The simulation and analysis of algorithms for redundant reader elimination in RFID system [A], Proceedings of the 3rd UK Sim European Symposium on Computer Modeling and Simulation, (2009), Washington, DC: IEEE Computer Society Press.
- [6] A density-based algorithm for redundant reader elimination in a RFID network [A], Proceedings of the 2nd International Conference on Future Generation Communication and Networking, (2008), Washington, DC: IEEE Computer Society Press.
- [7] S.-L. Lv and S.-Z. Yu, "A algorithm of redundancy reader removal based on RFID middle-ware system [J]", Journal of Electronics, vol. 40, (2012).
- [8] A. Solanas, "A Distributed Architecture for Scalable Private RFID Tag Identification [J]", In Computer Network, vol. 51, (2007).

## Authors



**Yunhua Gu** Professor Yunhua Gu received the B.E. degree in the Computer Science and Technology from Shanghai Jiaotong University in 1987, and M.E. degree in the Computer Science and Technology from Southeast University in 2000. Now, she is a professor in the Computer and Software Institute, Nanjing University of Information Science and Technology. Her main research interests are computer network, database, GIS systems. She is a member of ACM and a senior member of CCF.



**Bao Gao** She received the B.E. degree in the Software Engineering from Nanjing University of Information Science and Technology China, in 2010. Now, she is a master degree candidate of Nanjing University of Information Science and Technology, major in computer science and technology. Her main research areas are RFID Data Management.



**Jin Wang**, he received the B.S. and M.S. degree in the Electronical Engineering from Nanjing University of Posts and Telecommunications, China in 2002 and 2005, respectively. He received Ph.D. degree in the Ubiquitous Computing laboratory from the Computer Engineering Department of Kyung Hee University Korea in 2010. Now, he is a professor in the Computer and Software Institute, Nanjing University of Information Science and technology. His research interests mainly include routing method and algorithm design, performance evaluation and optimization for wireless ad hoc and sensor networks. He is a member of the IEEE and ACM.