



# A distributed architecture for scalable private RFID tag identification

Agusti Solanas \*, Josep Domingo-Ferrer, Antoni Martínez-Ballesté, Vanesa Daza

*Rovira i Virgili University of Tarragona, Department of Computer Engineering and Maths, Av. Països Catalans 26, E-43007 Tarragona, Catalonia, Spain*

## Abstract

The fast growth of Radio Frequency IDentification (RFID) implies a deployment challenge, namely how to keep this technology scalable without renouncing security and privacy features. This paper focuses on combining tag privacy and scalability using the hash locks mechanism. Our contribution is twofold: (i) a cell-based architecture is proposed in which the readers co-operate in order to conduct tag identification in a private and scalable way; (ii) a communication protocol for the proposed architecture is defined and assessed. The proposed architecture and protocol solve the scalability shortcomings of private RFID identification pointed out by Juels and Weis.

© 2007 Elsevier B.V. All rights reserved.

*Keywords:* RFID; Scalability; Security; Privacy; Identification

## 1. Introduction

The development of the radio-frequency identification (RFID) technology is now a reality and pervasiveness of this technology within the next few years is a realistic situation. A few years ago, RFID tags were only used for tracking expensive items such as shipping containers, automobile parts, or even live cattle. The value of such tracked goods fully justified the investment in RFID tags, which initially had a significant cost. Nowadays, the cost of RFID components is decreasing and their low

prices have opened a broad range of possibilities involving mass deployment of tags.

A key factor in the spectacular market push of RFID technology is the interest by large retailers (e.g., Wal-Mart<sup>1</sup>), important manufacturers (e.g., Gillette, Procter & Gamble, etc.) and governments. As a result, almost any object in our daily life is liable to carry a RFID tag, at least during a period of its life (e.g., during the manufacturing process, during its distribution, at the shop, etc.). Due to this important revolution many information technology manufacturers have entered the RFID arena (e.g., Philips Semiconductors, Intel, Intermec, Tyco, etc.).

Some authors [10] depict the future of RFID with a mixture of optimism and concern. On the good

\* Corresponding author.

*E-mail addresses:* [agusti.solanas@urv.cat](mailto:agusti.solanas@urv.cat) (A. Solanas), [josep.domingo@urv.cat](mailto:josep.domingo@urv.cat) (J. Domingo-Ferrer), [antoni.martinez@urv.cat](mailto:antoni.martinez@urv.cat) (A. Martínez-Ballesté), [vanesa.daza@urv.cat](mailto:vanesa.daza@urv.cat) (V. Daza).

<sup>1</sup> Wal-Mart started to explore the RFID technology in 2003 and devoted at least three billion dollars to implement it [10].

side, the RFID technology will eventually replace bar codes. It can be used on a variety of fields and serves many purposes, namely goods tracking, supply chain optimization, reduction of theft and loss during the distribution and selling processes, technical equipment tracking, inventory management, firemen team location in fire departments, etc. On the bad side, RFID systems are a cause of privacy concern for consumers, which might seriously hamper their deployment. Clearly, RFID technology must not violate privacy nor civil liberties. In [11] the most threatening points are identified as: hidden placement of tags, uniqueness of the identifiers, massive data aggregation, possibility of hidden readers and individual tracking and profiling.

The above privacy threats have unleashed radical opposition in some cases. Consumer Against Supermarket Privacy Invasion and Numbering (CASPIAN [8]) criticized the plan of Benetton to attach tags to its products, leading to a boycott of these products in 2003 [4,5]. Something similar happened to Tesco [16] in 2005. Some authors [1] are completely against the use of supermarket cards and, accordingly, they are against the deployment of RFID systems [2] in the sense that they would help infringing on personal privacy. They argue that the improper use of this technology represents a massive push toward global surveillance being driven by the retail sector.

Some of the aforementioned privacy problems can be overcome by using cryptographic techniques such as the ones described in [12,13]. Juels and Weis have given a definition of strong privacy for RFID in [13]. However, their notion of strong privacy seems to be in conflict with scalability: private tag identification in their model involves decryption of the ID of the tag being identified by exhaustive search. The same authors acknowledge this shortcoming when they point out that

For RFID tags capable of only symmetric-key cryptography, we believe that our definition may require the reader to perform brute-force search to identify tags[. . .] Such a system scales poorly.

The scalability of RFID privacy is a major challenge to be satisfied in order for pervasive tag deployment to be possible. This paper intends to advance in that direction. A solution based on the distribution of tag ID information with a communication protocol between readers is presented. For the sake of concreteness, we focus on privacy based on hash locks.

However, our approach can be used to mitigate the scalability problems of any RFID privacy-preserving technique. In Section 2, the architecture of the system is described and the notation and assumptions are presented. In Section 3, an information sharing protocol suite is detailed. The results of the protocol simulation are given in Section 4. Finally, Section 5 is a conclusion.

## 2. The architecture

Our system can be defined in terms of its components (i.e., tags and readers), the location and capabilities of those components and their privacy functionalities. The proposed architecture is cell-based. Cell-based approaches have successfully been used to alleviate traffic in dense ad hoc networks [7] and are proposed here to scale computation in private RFID identification.








### 2.1. Main components

The main components in our architecture are RFID readers and tags (see Table 1 for a graphical representation of these components).

The tags considered in this article are passive devices that can answer queries from readers by using the power transmitted by the readers. Each tag has a unique identifier and returns it when queried by a reader. Returning the identifier without any sort of protection could jeopardize tag security. Thus, we assume that the tag is able to compute a simple hash function to protect its identity from an eavesdropper (these security aspects are elaborated further in Section 2.3). Also, it is assumed that tags can change their location at any time. No other additional features are considered for the tags.

In our model, the readers are static devices intelligently distributed to cover the area in which the tags are roaming. In contrast to the tags described above, readers are active devices capable of detecting tags by

Table 1  
Symbols used in the architecture description

| Symbol                                                                              | Meaning                       |
|-------------------------------------------------------------------------------------|-------------------------------|
|  | Reader covering a cell        |
|  | Tag in a cell                 |
|  | A message between two readers |
|  | Tag ID in the cache           |
|  | System Access Point (SAP)     |
|  | System Exit Point (SEP)       |
|  | Movement of a tag             |

emitting a signal with a certain frequency. *It is assumed that a reader has some computational and storage capabilities* (e.g., a reader must be able to store a number of identifiers in a cache, compute hash functions, generate pseudo-random numbers, etc.). The coverage range of the readers is a system parameter to be taken into account during the deployment process: the readers must be able to locate the tags in the system and no shadow areas<sup>2</sup> are allowed. Without loss of generality, we will assume that each reader covers a specific square area, which will be called a *cell*. Moreover, all readers have communication capabilities and can exchange information with other readers using a secure channel (e.g., an encrypted wireless connection).

## 2.2. Covering the space

We now describe the spatial distribution of readers and the distribution of IDs among them to increase scalability. Consider an area  $\Psi$  that can be covered through the use of a number of readers. Assume that tags enter and leave  $\Psi$  through designated points called System Access Points (SAP) and System Exit Points (SEP), respectively.

Readers are placed according to a grid pattern depicted in Fig. 1. Let  $A_i$  be the square cell covered by the  $i$ th reader  $R_i$ . For the sake of simplicity, we consider that all readers have the same coverage range, so that all cells have the same size. Further, we consider cells to be disjoint and to span the entire area  $\Psi$ . Formally,

$$\bigcup A_i = \Psi \quad \forall i, \quad (1)$$

$$A_i \cap A_j = \emptyset \quad \forall i, j | i \neq j. \quad (2)$$

It is assumed that the readers are able to locate a tag by collaborating. Although the technologies for locating a tag are beyond the scope of this work, we next discuss some relevant issues about tag location.

Let  $D$  be the radius of the smallest circle containing a square cell and  $d$  the radius of the greatest circle inscribed in a square cell. Depending on the location of the tag, readers can face three different situations (see numbers in Fig. 1):

1. The distance between the tag and the reader  $R_i$  is less than  $d$ : in this case, the tag is located in the square area covered by  $R_i$ .

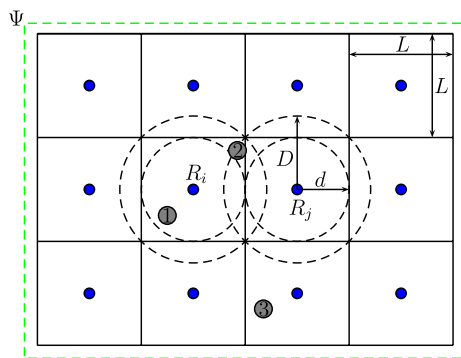


Fig. 1. Scheme of the coverage of a set of readers. The numbers are used to indicate different tag location situations.

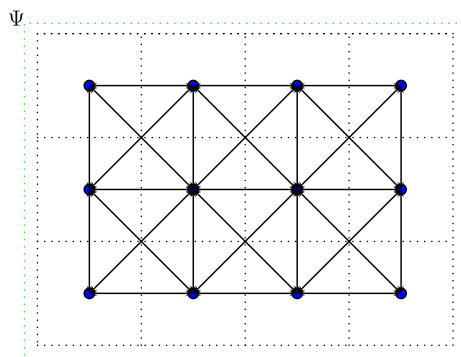


Fig. 2. Communication graph.

2. The distance between the tag and  $R_i$  is less than  $D$  and greater than  $d$ : in this case  $R_i$  needs the help of an adjacent reader  $R_j$  to determine the location of the tag. Note that it is not necessary to exactly locate the tag: determining its current cell is enough. Thus, it is possible to fulfill this task with only two readers.<sup>3</sup>
3. The distance between the tag and  $R_i$  is greater than  $D$ : in this case, the tag is off range of  $R_i$ .

In order to perform properly, the readers must be able to communicate between them. To that end, we propose the use of a network topology that can be represented as a graph where nodes are readers and edges are connections between readers. Let  $R_i^{\text{adj}}$  be the set of readers which are adjacent, i.e., connected, to  $R_i$ . As an example, Fig. 2 shows the communication graph for a regular grid with 12 readers.

<sup>2</sup> A shadow area is defined as a region where tags are not detected by any reader.

<sup>3</sup> If the exact location were needed, at least three readers ought to be used.

### 2.3. Privacy

When the RFID technology was initially deployed, the population of tags was quite small. However, the number of outstanding tags is growing fast and this trend is likely to continue with the replacement of optical bar codes with electronic product codes (EPC). Clearly, the possibilities of harming the privacy of consumers of products tagged with EPC will increase.

A variety of information can be stolen from consumers by reading without authorization the identifiers of the tags attached to the products that they buy. It would be possible to determine the amount of money that a consumer spends, the products that he/she prefers, his/her eating habits, etc. An eavesdropper would be able to gather unprecedented amounts of information. Thus, privacy is a very important factor that has to be tackled if a wide and complete deployment of the RFID technology is desired.

In [13], Juels and Weis propose a definition of privacy. However, basic RFID tags lack or have little cryptographic functionality, so researchers have had to develop a farrago of lightweight methods offering security and privacy with as little cryptography as possible. Examples are the method by Ohkubo, Suzuki and Kinoshita (OSK) [15] and its evolutions [3], the method by Nohara, Inoue, Baba and Yasuura (NIBY) [14], the YA-TRAP scheme of Tsudik [17], the zero-knowledge scheme of Engberg, Harning and Jensen (EHJ) [9] and O-TRAP [6].

As mentioned above in Section 2.1, in this article we consider a class of RFID tags capable of simple cryptographic operations, such as computing a one-way hash function. Thus, a tag can securely send its ID to the reader by implementing the improved randomized hash locks described in [13], which have been shown to be a secure evolution of the deterministic hash locks and randomized hash locks proposed in [18].

The basic operation of the improved randomized hash locks is depicted in Fig. 3 and is next briefly described:

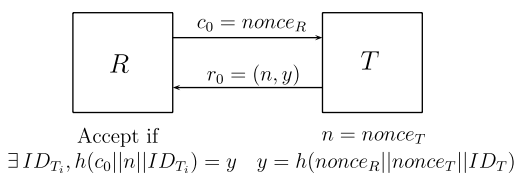


Fig. 3. Diagram of Juels–Weis improved randomized hash locks.

1. A reader  $R$  sends a challenge  $c_0$  to a tag  $T$ , where  $c_0 = \text{nonce}_R$  is generated uniformly at random.
2.  $T$  generates its own nonce  $\text{nonce}_T$  and hides its unique identifier  $ID_T$  by sending a response  $r_0 = (\text{nonce}_T, h(\text{nonce}_R || \text{nonce}_T || ID_T))$ .
3. To determine  $ID_T$ ,  $R$  must perform an exhaustive search of the IDs in its database to compute  $r_i = (\text{nonce}_T, h(\text{nonce}_R || \text{nonce}_T || ID_{T_i}))$  and compare the result with  $r_0$ . Once  $R$  finds an  $ID_{T_i}$  that satisfies  $r_i = r_0$ , the tag is identified.

In [13] it is proven that improved randomized hash locks offer strong tag privacy in the face of eavesdroppers. The main limitation of this technique is scalability: indeed, the authors of [13] express their belief that, for RFID tags capable of only symmetric-key cryptography, their definition of strong privacy may require the reader to perform brute-force search to identify tags, which scales poorly. They also point out the need of definitions and protocols for RFID privacy that are weaker, but more *practical* and *useful*. We absolutely agree with their remarks and, in this paper, we concentrate on the definition of a protocol permitting the practical use of privacy schemes like the improved randomized hash locks in a scalable manner.

### 3. Information sharing protocol suite

Our method is based on sharing information between readers. The shared information is the tag ID and the ID of the reader which covers the cell in which a certain tag is located. The readers use three kinds of messages to share information: tag arrival, tag roaming and tag departure (see Table 2). This information is stored in the local cache of each reader involved in the message exchange.

In order for information to be shared without unnecessary replication, each reader removes from its cache the information related to tags which are no longer in the reader's cell or in any cell adjacent to the reader's cell.

The suite consists of three protocols corresponding to the life cycle of a tag with respect to the system (i.e., arrival, roaming and departure):

Table 2

Message types used by the information sharing protocol suite

| Message        | Meaning                                      |
|----------------|----------------------------------------------|
| $(T, \oplus)$  | Tag $T$ enters the system for the first time |
| $(T, R_i)$     | Tag $T$ enters the cell of reader $R_i$      |
| $(T, \ominus)$ | Tag $T$ leaves the system                    |

1. *Arrival protocol*. This protocol starts when a new tag enters the system for the first time. Upon arrival of a tag in the system, a number of messages are sent in order to propagate the tag ID and the ID of the reader that acted as a SAP (entry point).
2. *Roaming protocol*. This protocol is used when a tag moves from the cell of a reader into the cell of another reader. As a result, a number of ID propagation messages are generated and a request of ID deletion is also sent to the readers which are no longer accessible by the tag from its new location (the tag ID should only be kept by the readers of the current cell and the adjacent cells).
3. *Departure protocol*. This protocol is responsible for managing the departure of tags from the system. It generates a number of ID deletion notifications to rid the readers of the IDs of the departed tags.

### 3.1. Arrival protocol

The number of SAPs and their location are variable and user-definable, i.e., they depend on the nature and lay-out of the facility (airport, factory, store, etc.) served by the RFID system.

In a wholesale distribution center, each type of good enters the system through a designated gate, to which a specific SAP can be associated. Each SAP is supposed to know all the possible tags which can enter the system through it. For example, fresh fish enters through gate 1 served by SAP1, cleaning products enter through gate 2 served by SAP2, etc. Thus, SAP1 only needs to know the information about fresh fish and SAP2 only needs to know the information about cleaning products. Thanks to this division, the amount of information stored in the SAPs scales better and goods can enter the system in an orderly fashion.

We assume that a SAP consists of a reader connected to a computer that can efficiently access a database of tag IDs. Regarding the remaining readers (those which are not SAPs), they can be very simple devices with little storage and computational capabilities.

**Note 1.** It should be noticed that our approach substantially differs from a centralized scheme in which all readers are connected to a back-end computing system. In our approach, only SAPs need a connection to the back-end and only the

incoming tags are considered, which increases scalability.

For the sake of simplicity we describe the protocol with a single SAP and a single reader  $R_{in}$  that receives the new tags entering the system. The generalization to multiple SAPs and readers is straightforward.

The arrival protocol is as follows:

#### Protocol 1 (*Tag arrival*).

1. The protocol starts when a tag  $T$  is detected by a SAP (see Step 1 of Figs. 4 and 5). If  $T$  is not found in the SAP database then the SAP raises an alarm to inform that there is an unidentified tag trying to enter the system without permission.
2. If the tag is correctly identified, the SAP sends a message to  $R_{in}$  in order to inform the reader that a new tag  $T$  is going to enter the reader's cell<sup>4</sup> (see Step 2 of Figs. 4 and 5).
3.  $R_{in}$  adds the tag  $T$  to its cache; thus, when  $T$  reaches the cell of  $R_{in}$ , the reader is able to authenticate  $T$ . After authentication,  $R_{in}$  sends a message to all readers  $R_{in}^{adj}$  in adjacent cells to inform them that  $T$  has entered  $R_{in}$ 's cell and can roam to any adjacent cell (see Step 3 of Figs. 4 and 5).
4. In response to that message, the adjacent readers  $R_{in}^{adj}$  add  $T$  to their caches and record the name of the message originator (i.e., in this case  $R_{in}$ ).

### 3.2. Roaming protocol

The roaming protocol performs the task of sharing the tag ID information needed for updating the caches of the readers. By properly updating their caches, the readers can authenticate the tags in their cells, and can also leverage their resource management (memory allocation, computational power). To that end, the ID information of the tags in a cell must be shared with adjacent readers, and non-adjacent readers must remove the ID information from their caches. Such a removal averts the uncontrolled growth of the reader caches and makes the system scalable in terms of computational cost and memory space.

<sup>4</sup> Once a tag enters the system, its ID can be removed from the SAP database because it is transferred to the readers in the system and the tag will not be allowed to leave the system through a SAP.

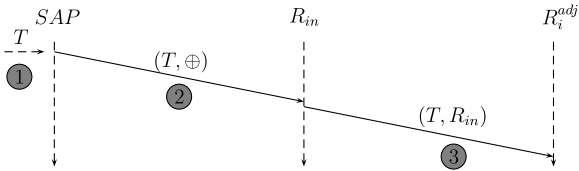


Fig. 4. Messages generated upon the arrival of an authenticated tag into the system.

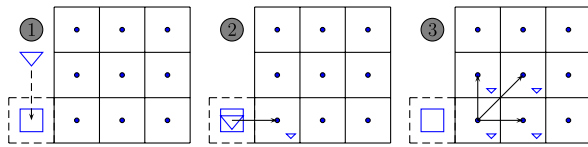


Fig. 5. Graphical scheme of the arrival of an authenticated tag into the system.

The roaming protocol is launched when any tag  $T$  moves from its current cell to another (adjacent) cell. The protocol works as follows:

**Protocol 2 (Roaming).**

1. A tag  $T$  is detected by a reader  $R_i$  other than the owner of the tag (see Step 1 of Figs. 7 and 6), where we denote by owner of  $T$  the last reader that informed the rest of the readers that  $T$  was in its cell. Due to the spatial distribution of the readers,  $T$  must come from one of the adjacent readers to  $R_i$ , so  $R_i$  has in its cache the ID information of  $T$  and is able to identify it.
2. After identification,  $R_i$  sends a message to its adjacent readers  $R_i^{adj}$  in order to inform them that the new owner of  $T$  is  $R_i$  (see Step 2 of Figs. 7 and 6).

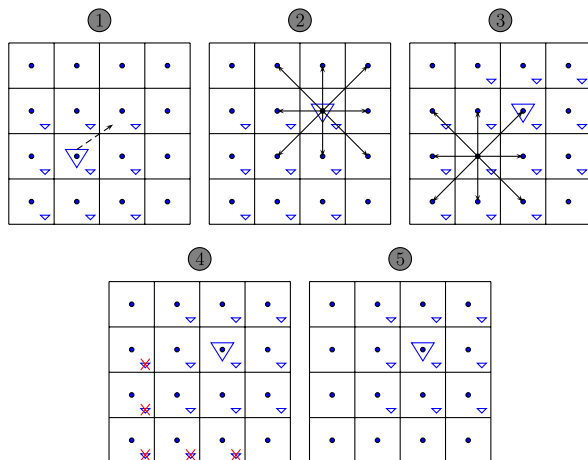


Fig. 6. Graphical scheme of an authenticated tag roaming about the system.

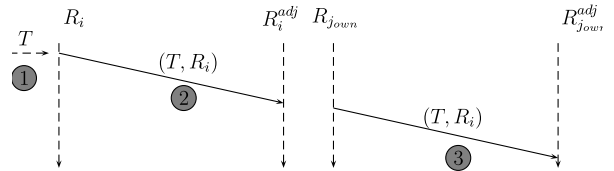


Fig. 7. Messages generated during the roaming protocol.

3. Upon message reception, the adjacent readers behave differently depending on their current cache information:

- (a) If an adjacent reader  $R_j \in R_i^{adj}$  has no information about  $T$  in its cache then it simply appends  $T$  and its owner information (see Step 3 of Fig. 6).
  - (b) If an adjacent reader  $R_j \in R_i^{adj}$  has information about  $T$  in its cache but was not the previous owner of  $T$ , then it only needs to update the name of the owner of  $T$  (i.e., in this case  $R_i$ ).
  - (c) The adjacent reader  $R_{owen}$  which was the previous owner of  $T$  must communicate to its adjacent readers  $R_{owen}^{adj}$  that the new owner of  $T$  is  $R_i$ . To do so,  $R_{owen}$  propagates the message from  $R_i$  to its adjacent readers (see Step 3 of Fig. 7 and 6).
4. When the adjacent readers of  $R_{owen}$  receive the message, they behave differently depending on their adjacency relations:
    - (a) If a reader  $R_k \in R_{owen}^{adj}$  is adjacent to  $R_i$  then it does nothing.
    - (b) If a reader  $R_k \in R_{owen}^{adj}$  is not adjacent to  $R_i$  then it removes the information on  $T$  from its cache (see Step 4 of Fig. 6).
  5. At the end of the protocol, only readers adjacent to the current owner keep information on  $T$  in their cache (see Step 5 of Fig. 6).

**3.3. Departure protocol**

In almost any RFID application, tags which have entered a controlled system must leave it. In a supermarket, grocery, warehouse, etc., tags travel from the shelves to the checkout. Similarly, the tags in a production line leave the system when the parts they correspond to have been completely assembled.

In order to control the departure of tags from the system, the System Exit Points (SEP)s mentioned above are used. A SEP is an area covered by a

reader from which no tag can go back into the system (e.g., a checkout in a supermarket).

The departure protocol works as follows:

**Protocol 3 (Departure).**

1. The protocol starts when a tag  $T$  is detected by a SEP (see Step 1 of Figs. 8 and 9).
2. The SEP informs its adjacent readers  $R_{SEP}^{adj}$  that  $T$  must be removed from their caches because there is no chance for  $T$  to go back (see Step 2 of Figs. 8 and 9). The adjacent readers of the SEP, including the previous owner  $R_{own}$  of  $T$ , erase the information on  $T$  from their caches (see Step 2 in Fig. 9).
3. The previous owner  $R_{own}$  of  $T$  propagates the removal message to its adjacent readers  $R_{own}^{adj}$  (see Step 3 of Figs. 8 and 9).
4. The readers  $R_{own}^{adj}$  remove any information on  $T$  from their caches, and nothing remains in the system about the departed tag (see Step 4 of Fig. 9).

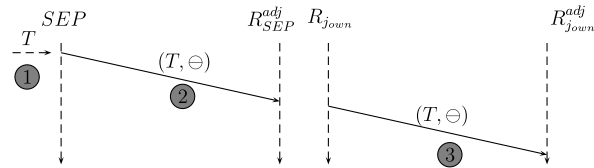


Fig. 8. Messages generated during the departure protocol.

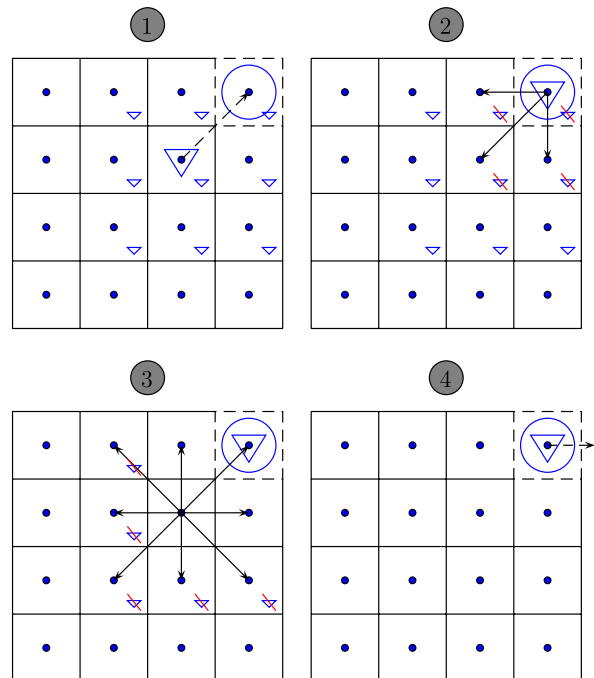


Fig. 9. Graphical scheme of a tag departing through a SEP.

**3.4. Comparison with the centralized approach**

To the best of our knowledge, ours is the first contribution to scalable private tag identification. Thus, the only possible comparison is with the (non-scalable) centralized approach to private tag identification, whose shortcomings are at least:

*Dependability:* Any failure in the back-end server of a centralized private tag identification system results in a complete system crash.

*Message explosion:* The number of messages generated by the tags grows with their number, and it is obvious that a single back-end will eventually become swamped with incoming traffic as the number of tags grows.

*Computational explosion:* In a single back-end approach the computations for private tag identification cannot be distributed, so that the back-end server will be unable to privately identify all tags if their number keeps growing.

*Delay:* A centralized approach is supposed to work with a FIFO structure to queue the messages of the tags waiting to be identified. Even without reaching back-end collapse, buffering a large number of tag messages may lead to delays incompatible with the roaming pattern of tags: if a tag moves to a new cell before the previous messages have been processed by the back-end, the system fails in its purpose of keeping control of the tag movements.

The above shortcomings are obviously mitigated by the distributed scheme proposed in this paper.

**4. Experimental results**

The simulation of a complex environment is not straightforward because of the quantity and variety of the parameters involved. The aim of the presented simulation is to show that the size of the cache used at the readers quickly stabilizes and tends to become uniform over time, regardless of the number and the roaming pattern of tags. This is an indication of the scalability of the proposed system.

Three different simulation scenarios depicted in Fig. 10 have been considered:

*Empty:* In an empty scenario there are no obstacles, so the tags that enter the system can roam

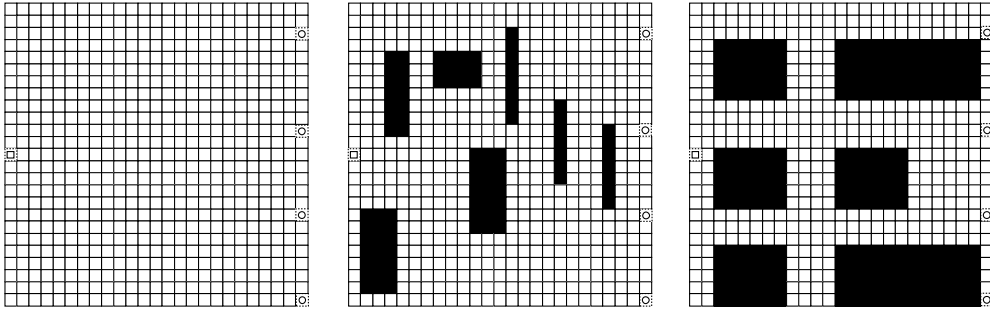


Fig. 10. The three test scenarios. From left to right “Empty”, “Unstructured” and “Corridor”.

freely. This kind of environment is not very common but it is useful to demonstrate how the system works.

*Unstructured*: This scenario includes randomly distributed obstacles which hinder free tag roam-

ing. It is suitable for analyzing the behavior of the system in the presence of bottlenecks.

*Corridor*: Most real-life scenarios can be viewed as corridors (e.g., supermarkets, offices, libraries, etc.).

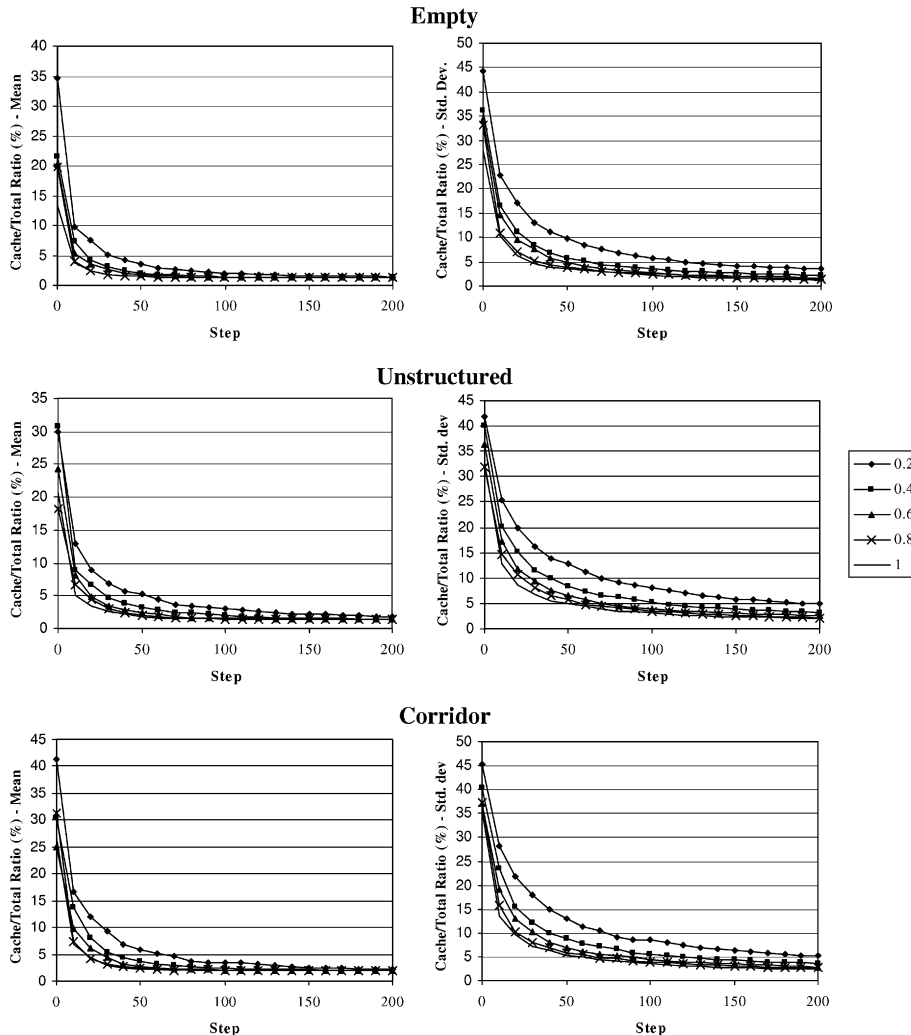


Fig. 11. Evolution of the Cache/Total Ratio mean and deviation for different values of the roaming probability.



Looking at Fig. 10, each scenario consists of a SAP located on the left, 4 SEPs on the right and a grid of 25 by 25 readers evenly distributed. We do not need to specify the real spatial dimensions of the environment nor the cover range of the readers, because our protocols are invariant to these parameters, as they only make use of the adjacency relations between the readers.

In order to test the scalability of the system, we have taken as a measure the ratio between the cache size at the readers and the total number of tags. The Cached/Total Ratio (CTR) of a reader  $R$  at simulation step  $s$  is defined as follows:

$$CTR_{(R,s)} = 100 \times \frac{I_{(R,s)}}{E_s}$$

where  $I_{(R,s)}$  is the number of tag IDs stored in the cache of reader  $R$  and  $E_s$  is the total number of tags in the system.

The means and the standard deviations of the CTR are computed by taking into account only the readers with a non-empty cache.

For each scenario, 100,000 tags are allowed to enter into the system through the SAP during the simulation.

In order to test our proposal under several conditions, we have performed tests using the three aforementioned scenarios and taking into account two parameters:

- *Roaming probability of a tag*: This is the probability that a tag moves to another cell (and thus to a different reader).
- *Number of new tags per simulation step*: This is the number of tags that enter the system (through the SAP) at each simulation step.

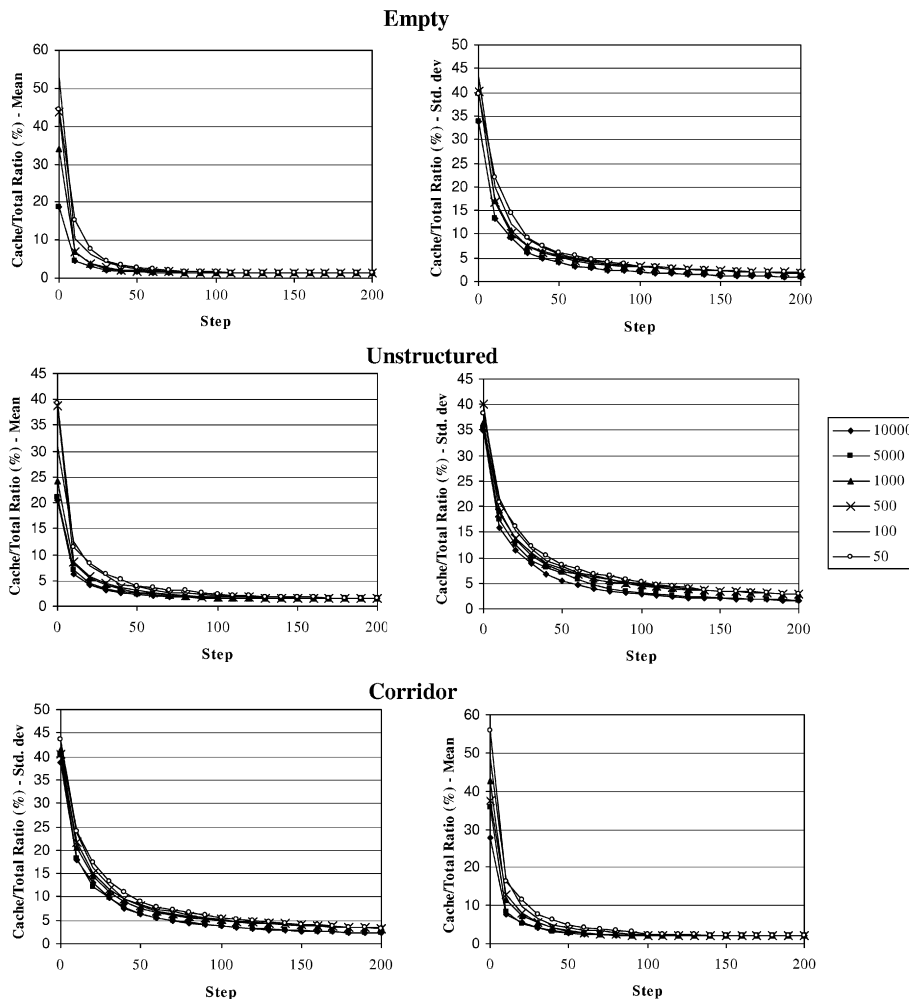


Fig. 12. Evolution of the Cache/Total Ratio mean and deviation for different values of the number of entering tags per step.

For roaming probabilities ranging from 0.2 to 1.0 and 1000 new tags per simulation step, Fig. 11 shows the evolution of the mean and standard deviation of the CTR for readers with non-empty cache. Regardless of the roaming probability, the load of the reader caches tends to become uniform.

For a number of new tags per step ranging from 50 (slow entrance) to 10,000 (fast entrance) and a roaming probability 0.5, Fig. 12 shows the evolution of the mean and standard deviation of CTR for readers with non-empty cache. Regardless of the tag entrance rate, the load of the reader caches tends to become uniform.

After analyzing these results, we can conclude that neither tag mobility nor tag entrance rate significantly affect the overall behavior of the system, in the sense that the cache load at the readers becomes evenly distributed over time.

For all three scenarios, the following can be observed:

- During the first simulation steps, the mean of the CTR is high. This happens because all tags enter

the system through the SAP and the only readers that store tag identifiers are the ones close to the SAP.

- After a few steps, the mean and the standard deviation of the CTR sharply decrease. As tags move around, the proposed protocol ensures that their identifiers become evenly distributed among the readers and the average size of the non-empty caches decreases. Consequently, it can be concluded that the proposed information sharing protocol scales properly regardless of the scenario.

Fig. 13 shows the number of identifiers for each cache at two different simulation steps ( $s = 20$  and  $s = 300$ ) for the “Empty” scenario, using 0.5 roaming probability and with a rate of 1000 new tags per simulation step. It can be observed that, for  $s = 20$  the caches which are close to the SAP have a higher number of stored identifiers. Figs. 14 and 15 show the same information for the “Unstructured” and “Corridor” scenarios, respectively. It can be seen that, thanks to the proposed

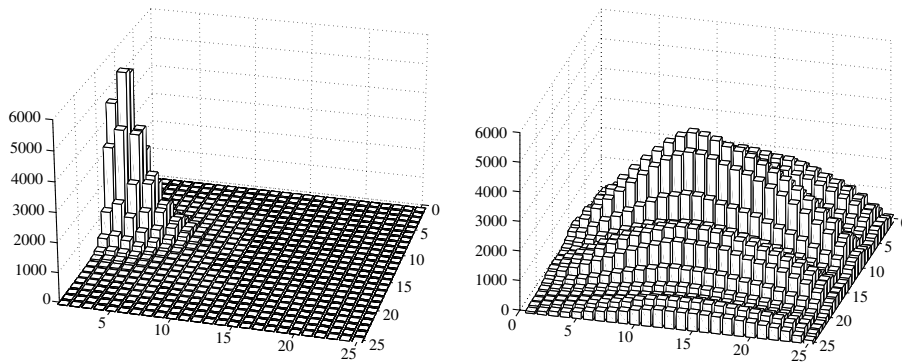


Fig. 13. Number of tags in each cache for the “Empty” scenario at steps 20 (left) and 300 (right).

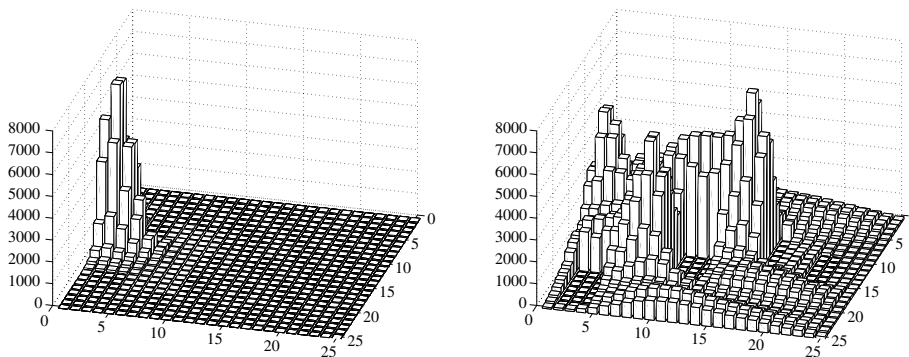


Fig. 14. Number of tags in each cache for the “Unstructured” scenario at steps 20 (left) and 300 (right).

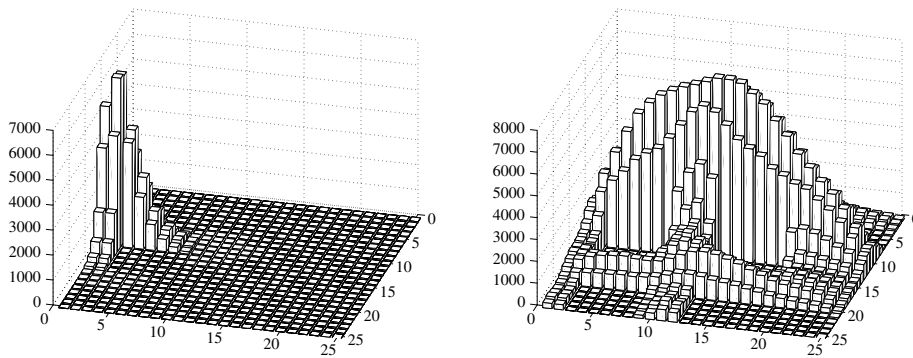


Fig. 15. Number of tags in each cache for the “Corridor” scenario at steps 20 (left) and 300 (right).

protocol, the number of identifiers stored in the caches is much smaller than the total number of identifiers in the system.

From the above results, it becomes clear that the proposed distributed architecture and protocol are useful to reduce the cache size required at the readers, which makes the RFID system scalable while still allowing private tag identification.

## 5. Conclusions and future work

RFID technology has already become a powerful reality. However, there is still a long way to go in order to evolve RFID into a ripe technology that can be widely deployed without compromising with privacy and security matters.

In this article we have proposed a cell-based architecture which provides scalability without renouncing private tag identification. The above architecture is engineered by an information sharing protocol suite whose basic idea is to distribute among the readers the computations needed for private tag identification.

To the best of our knowledge, this is the first contribution in the literature that tackles scalability of private RFID tag identification.

The experimental results provided clearly demonstrate the usefulness of our approach and pave the way to future research and development. In particular, we envision at least the following future research topics:

- Study and optimize the number of messages generated by the information sharing protocol suite.

- Generalize the proposed system to allow the use of heterogeneous readers (i.e., with different cover ranges and capabilities).
- Test the system in real scenarios.

## Acknowledgements

This work was partly supported by the Spanish Ministry of Education through project SEG2004-04352-C04-01 “PROPRIETAS” and by the Government of Catalonia under grant 2005 SGR 00446.

## References

- [1] K. Albrecht, Supermarket cards: the tip of the retail surveillance iceberg, *Denver University Law Review* 79 (2002) 534–565. <http://www.spsychips.com/documents/Albrecht-Denver-Law.pdf>.
- [2] K. Albrecht, L. McIntyre, RFID: the big brother bar code, *American Legislative Exchange Council Forum* 6 (2004) 49–54. <http://www.spsychips.com/alec-big-brother-barcode-article.html>.
- [3] G. Avoine, P. Oechslin, A scalable and provably secure hash-based RFID protocol, in: *Third IEEE International Conference on Pervasive Computing and Communications Workshops 2005*, March, pp. 110–114.
- [4] E. Batista. What your clothes say about you, *Wired News* <http://www.wired.com>, 2003. <http://www.wired.com/news/wireless/0,1382,58006,00.html>.
- [5] Boycott Benetton, <http://www.boycottbenetton.com>.
- [6] M. Burmester, T. van Le, B. de Medeiros, Provably secure ubiquitous systems: Universally composable RFID authentication protocols, 2006. Referenced 2006 at <http://eprint.iacr.org/2006/131.pdf>.
- [7] M. Burmester, T. van Le, A. Yasinsac, Adaptive gossip protocols: managing security and redundancy in dense ad hoc networks, *Journal of Ad hoc Networks* 4 (3) (2006) 504–515.

- [8] C.A.S.P.I.A.N. <http://www.nocards.org>.
- [9] S.J. Engberg, M.B. Harning, C.D. Jensen, Zero-knowledge device authentication: privacy & security enhanced RFID preserving business value and consumer convenience, in: Second Annual Conference on Privacy, Security and Trust 2004. Referenced 2006 at [http://www.obivision.com/Papers/PST2004\\_RFID\\_ed.pdf](http://www.obivision.com/Papers/PST2004_RFID_ed.pdf).
- [10] C.C. Haley, Are you ready for RFID? Internetnews.com, November 2003. <http://www.internetnews.com/wireless/article.php/3109501>.
- [11] J.E. Henning, Preserving privacy in RFID deployment, Technical Report RVS-Occ-04-01, RVS, Faculty of Technology, University of Bielefeld, March 2004.
- [12] A. Juels, RFID security and privacy: a research survey, IEEE Journal on Selected Areas in Communications 24 (2) (2006) 381–394.
- [13] A. Juels, S.A. Weis, Defining strong privacy for RFID, 2006. <http://www.rsasecurity.com/rsalabs/node.asp?id=3046>, submitted for publication.
- [14] Y. Nohara, S. Inoue, K. Baba, H. Yasuura, Quantitative evaluation of unlinkable ID matching schemes, in: Workshop on Privacy in the Electronic Society (WPES), 2005.
- [15] M. Ohkubo, K. Suzuki, S. Kinoshita, Efficient hash-chain based RFID privacy protection scheme, in: International Conference on Ubiquitous Computing – Ubicomp, Workshop Privacy, Current Status and Future Directions, 2004.
- [16] Boycott Tesco, <http://news.zdnet.co.uk/0,39020330,39185481,00.htm>.
- [17] G. Tsudik, YA-TRAP: yet another trivial RFID authentication protocol, in: PerCom Workshops 2006, IEEE Computer Society, Pisa, Italy, 2006, pp. 640–643.
- [18] S.A. Weis, S.E. Sarma, R.L. Rivest, D.W. Engels, Security and privacy aspects of low-cost radio frequency identification systems, in: Security in Pervasive Computing, vol. 2802 of Lecture Notes in Computer Science 2004, pp. 201–212.



**Agusti Solanas** (Tarragona, Catalonia, Spain, 1980) is a predoctoral researcher at the CRISES Research Group in the department of Computer Science and Maths at Rovira i Virgili University (URV) of Tarragona, Catalonia. He received his B.Sc. and M.Sc. degrees in Computer Engineering from URV in 2002 and 2004, respectively, the latter with honours (Outstanding Graduation Award). He received a Master in Tele-

matics Engineering from the Technical University of Catalonia (UPC) in 2006. He was selected to take part into the Young Researchers Award 2005 organized by the Spanish Ministry of Education, Culture and Sport. His fields of activity are data privacy, data security, watermarking, neural networks and evolutionary computation. He participated in several Spanish-funded and Catalan-funded research projects. He has authored over 20 publications and he has delivered several talks. He has served as program committee member and reviewer in several conferences and journals.



**Josep Domingo-Ferrer** is a Full Professor of Computer Science at Rovira i Virgili University of Tarragona, Catalonia. He received with honors his M.Sc. and Ph.D. degrees in Computer Science from the Autonomous University of Barcelona in 1988 and 1991 (Outstanding Graduation Award). He also holds a M.Sc. in Mathematics. His fields of activity are data privacy, data security and cryptographic protocols. In 2003, he was a co-recipient

of a research prize from the Association of Telecom Engineers of Catalonia. In 2004, he got the TOYPS'2004 Award from the Junior Chambers of Catalonia. He has authored 3 patents and over 170 publications, one of which became an ISI highly-cited paper in early 2005. He was the co-ordinator of EU FP5 project CO-ORTHOGONAL and of several Spanish funded and US funded research projects. He has chaired or co-chaired 8 international conferences (including 2 CARDIS conferences) and has served in the program committee of over 42 conferences on privacy and security. He is an Associate Editor of three international journals and has been a Guest Editor of Computer Networks, IJUFKS and Data Mining and Knowledge Discovery. In 2004, he was a Visiting Fellow at Princeton University. He is the Secretary of IFIP WG 8.8 on Smart Cards and he is the founder and chairholder of the forthcoming UNESCO Chair in Data Privacy.



**Antoni Martínez-Ballesté** (Tarragona, Catalonia, Spain, 1976) is a Junior Lecturer at Rovira i Virgili University of Tarragona (URV). He is currently a member of the CRISES research group. His research interests are related to security in computer networks and secure applications. In 1999, he received his B.Sc. in Computer Systems Engineering from the URV. In 2002, he received his M.Sc. in Computer Engineering also from URV. In 2004, he received with honors a Ph.D. in Telematics Engineering from the Technical University of Catalonia (UPC). In 2004, he was a visiting researcher at LAAS-CNRS, Toulouse, France. He has authored more than 25 articles in journals and conferences. He is also currently working on innovative learning techniques in engineering and the European Higher Education Space.



**Vanesa Daza** graduated in Mathematics in University of Barcelona in 1999 and received her Ph.D. in Mathematics in Technical University of Catalonia in 2004. Afterwards, she joined a crypto-based security company as a Senior Researcher. Currently she holds a post-doc position in the CRISES research group at Rovira i Virgili University. Her research interests are mainly related with distributed cryptography and its applica-

tions. She has co-authored over 20 articles in journals and conferences. She has also co-authored 2 international patents. In 2001, she was a visiting researcher at BRICS (Basic Research in Computer Science) at Aarhus University (Denmark), supported by a EU Marie Curie Fellowship.