Revisión de protocolos para la identificación escalable, segura y privada en sistemas RFID

Rolando Trujillo-Rasúa, Antoni Martínez-Ballesté y Agusti Solanas
Universitat Rovira i Virgili
Departament d'Enginyeria Informàtica i Matemàtiques
Av. Països Catalans 26
E-43007 Tarragona, Catalunya
{rolando.trujillo,antoni.martinez,agusti.solanas}@urv.cat

***Abstract*—Los sistemas RFID permiten la identificación rápida y automática de etiquetas a través de un canal de comunicación inalámbrico. Información tal como el nombre, el modelo o la fecha de expiración de un producto, puede ser almacenada en dichas etiquetas y luego leída sin la necesidad de un contacto visual entre el lector y la etiqueta. Para que el uso de la tecnología RFID se generalice con éxito, es conveniente cumplir con varios objetivos: eficiencia, seguridad y protección de la privacidad.**

Actualmente no existe ningún protocolo de identificación para sistemas RFID de propósito general, que proporcione seguridad, privacidad y que, a la vez, sea escalable.

En este artículo presentamos un breve estudio de los protocolos de identificación seguros más conocidos para sistemas RFID. Nuestro objetivo es mostrar que, hasta el momento, sólo aquellos protocolos diseñados para ciertas aplicaciones concretas son los que pueden garantizar los tres objetivos anteriores. Planteamos, así, la necesidad de un estudio profundo en esta línea de investigación.

I. Introducción

La identificación por radio frecuencia (RFID) es una tecnología diseñada para la identificación eficiente de productos. Este tipo de sistemas es muy atractivo para cadenas de suministros, procesos de manufactura, etc. ya que esta identificación se realiza de forma inalámbrica y, además, no se requiere contacto visual entre el lector y la etiqueta. Estas características de los sistemas de identificación por radiofrecuencia permiten reducir costes y agilizar procesos relacionados con la distribución y gestión de productos entre múltiples organizaciones y empresas. Es por ello que los sistemas RFID están llamados a reemplazar los códigos de barras y cada vez son más las empresas que hacen uso de esta tecnología.

En términos generales, los componentes básicos de un sistema RFID son etiquetas y lectores. Los lectores son los componentes activos, que inician el proceso de identificación y proveen de energía a las etiquetas (*etiquetas pasivas*). Aunque existen etiquetas que pueden tener su propia fuente de energía (*etiquetas activas*), éstas suelen ser bastante más caras que las pasivas y su penetración en el mercado es menor. Así, a pesar de sus limitados recursos computacionales, las etiquetas pasivas son las más usadas actualmente.

En muchas aplicaciones orientadas a la logística y distribución de productos, un lector debe ser capaz de identificar una etiqueta entre millones de posibilidades. En este contexto, la escalabilidad del *proceso de identificación* es una propiedad irrenunciable[1]. Por otro lado, además de la escalabilidad existen otras dos propiedades importantes para el proceso de identificación:

- *Seguridad.* La propiedad de seguridad tiene dos vertientes: por una parte, sólo los lectores

[1]En el escenario que consideramos, en el que es necesario garantizar la seguridad y la privacidad en la identificación, la escalabilidad tiene relación con el número de operaciones criptográficas realizadas por un lector para identificar una etiqueta y depende de forma directa del número de posibles etiquetas a identificar.

autorizados deben poder acceder a los datos guardados en las etiquetas; y por otra parte, la suplantación de etiquetas debe tener una probabilidad de éxito matemáticamente despreciable.
- *Privacidad.* Consiste en que la operación de identificación debe evitar que potenciales lectores no autorizados puedan tracear una etiqueta.

La privacidad suele ser un concepto que depende de las aplicaciones, los tipos de adversarios, las restricciones físicas, etc. Por ello, dar una única definición de privacidad que satisfaga las necesidades de todas las aplicaciones suele ser difícil. En la literatura existen varias definiciones de privacidad ([6], [11]). En este artículo entendemos el concepto de privacidad como la incapacidad de un lector no autorizado de tracear una etiqueta [1].

Definición 1 (Intrazabilidad existencial)**.** Se dice que un protocolo cumple con la propiedad de *Intrazabilidad existencial* si dos respuestas de una misma etiqueta no pueden ser correlacionadas por un adversario.

Las propiedades de seguridad y privacidad están, en cierta forma, entrelazadas y, en este sentido, algunas de las soluciones destinadas a proporcionar seguridad tienen incidencia en la privacidad proporcionada y viceversa. Por ejemplo, para lograr seguridad en un proceso de identificación, se pueden usar funciones criptográficas de bajo coste computacional, tales como las funciones de resumen o los cifrados simétricos: la etiqueta necesita cifrar su identidad usando su clave secreta (que evidentemente será conocida sólo por los lectores autorizados). Para lograr privacidad es necesario usar, además, generadores de numeros aleatorios que diversifiquen las respuestas de las etiquetas (y, por consiguiente, se dificulte la traceabilidad). Así pues, la etiqueta debe enmascarar su identidad mediante un número aleatorio antes de hacer el cifrado. De este modo, todos los mensajes de identificación de la etiqueta conformarán, ante los ojos de un lector no autorizado, una secuencia de números aleatorios sin correlación.

En este escenario, para que el lector pueda obtener la identidad de una etiqueta, necesita conocer la clave con la que se cifró el mensaje: el lector debe buscar de forma exhaustiva en su base de datos una clave que pueda descifrar dicho mensaje. Dado que por motivos de seguridad cada etiqueta tiene una clave distinta, los lectores deben revisar todas las claves, lo cual resulta costoso computacionalmente. Este problema de escalabilidad representa un gran obstáculo para una tecnología que pretende ser un pilar fundamental de la nueva computación ubícua.

En el presente artículo realizamos un breve estudio de protocolos de identificación privada, segura y escalable para sistemas RFID. El resto del artículo se organiza de la siguiente forma: la Sección II presenta una revisión de los protocolos de propósito general. Los dividimos en cuanto a escalabilidad se refiere (teniendo en cuenta su complejidad computacional). La Sección III introduce algunos de los protocolos que mejoran la escalabilidad pero requieren una arquitectura o disposición específica de los lectores. Finalmente la Sección IV concluye el artículo.

II. Protocolos de propósito general

En esta sección revisamos los protocolos de identificación RFID de propósito general (en un escenario simple de lector y etiquetas) que contemplan seguridad y privacidad, en relación a la escalabilidad.

A. Protocolos con complejidad computacional lineal

El esquema *Improved Randomized Hash-lock* [5] es posiblemente el protocolo de identificación seguro para sistemas RFID más popular. Éste logra altos niveles de seguridad y privacidad a muy bajo coste computacional (en el lado de la etiqueta) dado que sólo hace uso de una función pseudoaleatoria y una función de resumen para generar las respuestas de las etiquetas. El protocolo funciona de la siguiente forma:

- El lector envía un número aleatorio r_1 a la etiqueta.
- Cuando ésta lo recibe, genera otro número aleatorio r_2 y computa la respuesta $r =$

Lector (R)		**Etiqueta (T)**
Genera r_1	$\xrightarrow{r_1}$	
		Genera r_2
	$\xleftarrow{r}$	$r = (r_2, h(r_1 \|\| r_2 \|\| ID_T))$
Computa r^{T_i}		
$r^{T_i} = (r_2, h(r_1 \|\| r_2 \|\| ID_{T_i}))$		
Identifica T como T_i si $r^{T_i} = r$		

Fig. 1. Improved Randomized Hash-lock

$h(r_1, r_2, ID)$ donde ID es su identificador y $h(...)$ es una función de resumen.

- Finalmente, el lector recibe r y r_2 como mensaje de identificación, con los cuales realiza una búsqueda exahustiva de un identificador ID_i tal que $r = h(r_1, r_2, ID_i)$.

La Figura 1 muestra una descripción esquemática de este protocolo. Aunque el esquema es seguro y proporciona privacidad, tiene un grave problema de escalabilidad, ya que el lector debe analizar la respuesta de la etiqueta mediante todas las posibles claves[2].

B. Protocolos con complejidad computacional logarítmica

Desde el punto de vista de la escalabilidad, los protocolos basados en árboles son posiblemente los más destacados [7]. Estos logran una complejidad computacional logarítmica ($O(b \log_b^N)$) donde N representa la cantidad de etiquetas en la base de datos y b el factor de ramificación. La idea es que cada etiqueta sea representada por una y sólo una de las ramas del árbol, es decir, por el conjunto de nodos que se sigue desde la raíz hasta una hoja. De este modo, cada etiqueta T_i contiene un secreto único k_i y, además, contiene las claves de los nodos que conforman su rama (véase la Figura 2).

Sin duda, el secreto único de la etiqueta es suficiente para identificarla como se hace en el esquema *Improved Randomized Hash-lock*. Sin embargo, los protocolos basados en árboles usan la rama de las etiquetas para descartar grandes conjuntos de identificadores posibles a medida que van avanzando por una única rama del árbol.

[2]Nótese que puede tratarse de cientos de miles o millones de posibilidades.

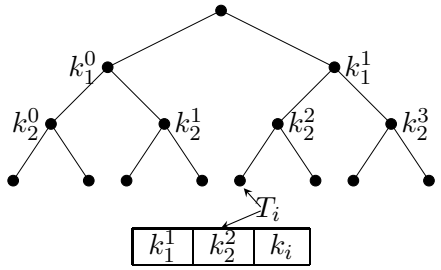


Fig. 2. Árbol de profundidad tres y factor de ramificación dos. La etiqueta T_i es representada por la rama k_1^1, k_2^2 y por su secreto k_i.

A pesar de que los protocolos basados en árboles suponen un significativo avance en cuanto a escalabilidad, presentan déficits de seguridad y privacidad importantes. SU problema radica en que las etiquetas comparten subramas y, por tanto, también comparten claves. Es por ello que comprometiendo[3] una sóla etiqueta otras etiquetas del sistema pueden ser rastreadas [3].

C. Protocolos con complejidad computacional sublineal

Una idea parecida a la de los protocolos basados en árboles es la de los protocolos basados en grupos [2]. En estos protocolos cada etiqueta tiene dos claves: una clave de grupo y una clave individual. Con el fin de reducir el tiempo de identificación, el lector primero identifica el grupo al que pertenece la etiqueta, y posteriormente sólo tiene que buscar entre las claves de las etiquetas que pertenecen a este grupo para poder descifrar el mensaje de identificación. De este modo, la complejidad

[3]Entendemos por comprometer una etiqueta la acción de conseguir sus claves secretas de forma ilícita.

computacional de este tipo de protocolos es sublineal, $O(\frac{N}{k})$, donde k es el número de grupos. No obstante, los protocolos basados en grupos tienen un problema similar al de los protocolos basados en árboles, es decir, si una etiqueta T es comprometida por un atacante, toda etiqueta T' que pertenezca al mismo grupo que T puede ser rastreada usando la clave de grupo.

III. Protocolos para disposiciones específicas

Existen otros protocolos que asumen topologías y distribuciones de lectores específicas en las que éstos colaboran para identificar etiquetas en un entorno distribuido. Estos protocolos están diseñados para reducir la complejidad computacional a $O(\frac{N}{k})$ donde k es el número de lectores en el sistema. El primer trabajo en esta dirección se debe a Solanas et al. [8]. Su propuesta implica que los lectores deben cubrir el sistema completo con sus campos de interrogación. Bajo esta suposición, dos lectores se definen como vecinos si y sólo si sus campos de interrogación están solapados. A partir de esta relación de vecindad entre los lectores se define un grafo de comunicaciones (véase la Figura 3). La idea se basa en que cada etiqueta, por definición, sólo podrá moverse entre lectores vecinos. De este modo, un lector solamente necesita almacenar la información de las etiquetas que están en su campo de interrogación y en el de los lectores vecinos. Así, cada lector maneja su propia base de datos *(caché)*, y ésta se va actualizando según los movimientos de las etiquetas.

A pesar de los beneficios en términos de escalabilidad, este protocolo es difícilmente aplicable en la práctica debido a que los lectores deben computar la distancia exacta entre su posición y la de las etiquetas.

Para resolver este problema práctico, se propuso un nuevo protocolo en [9]. Este protocolo, basado en el de Solanas et al., tiene diferencias significativas con respecto al original. Se introduce un nuevo parámetro $p \in [0, 1]$ que permite una reducción considerable de los tamaños de las cachés. Teóricamente, cuando $p = 1$ el tamaño de las caché es aproximadamente el mismo que en el protocolo

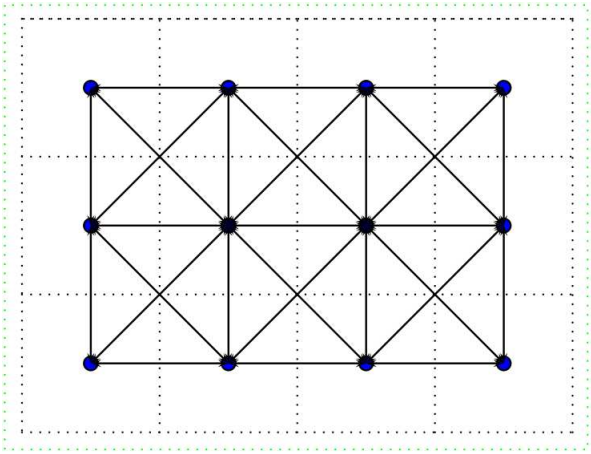


Fig. 3. Topología de comunicaciones entre lectores colaborativos

de Solanas et al., mientras que si $p = 0$ se alcanza el tamaño mínimo. Sin embargo la reducción del tamaño de cache implica un augmento en el número de mesajes enviados entre los distintos lectores colaborativos. Además el protocolo relaja las restricciones del original en cuanto a la superficie de interrogación de los lectores.

De forma parecida, Fouladgar y Afifi [4] apuntan que, en muchas aplicaciones, las etiquetas son leídas por un mismo conjunto de lectores. Por tanto, proponen agrupar las etiquetas según los lectores que más las leen. Esta idea mejora las propuestas de los protocolos basados en grupos [2] dado que las etiquetas no son asignadas a grupos de forma aleatoria, sino que son inteligentemente agrupadas según sus localizaciones más frecuentes. El problema en esta propuesta es que cuando las etiquetas tienen una largo ciclo de vida y se mueven por diferentes partes del sistema, la complejidad computacional se degrada de sublineal a lineal.

Finalmente, un nuevo protocolo basado en lectores colaborativos fue propuesto en [10]. Esta propuesta es capaz de mejorar la escalabilidad de propuestas anteriores tales como [8] y [4] en hasta un 50%. La mejora se debe al uso de una heurística que estima las localizaciones de las etiquetas. Además de las mejoras en cuanto a escalabilidad, esta propuesta mejora la flexibilidad y aplicabilidad de propuestas anteriores tales como [9], [8].

Lo más destacable de los protocolos basados en lectores colaborativos, es que no degradan ni la privacidad ni la seguridad a la vez que consiguen mejorar la escalabilidad del sistema de identificación.

IV. Conclusiones

En este artículo, hemos revisado protocolos de identificación para sistemas RFID en relación a la escalabilidad. Hemos mostrado que identificar etiquetas usando una complejidad computacional lineal es posible de forma segura y privada. Sin embargo, esta complejidad computacional sólo ha sido reducida por protocolos que asumen propiedades especiales de las aplicaciones. Otros protocolos genéricos que han intentado alcanzar complejidades computacionales logarítmicas o sublineales no logran ser privados y seguros como los de coste lineal.

Como trabajo futuro, proponemos hacer un estudio de los mejores protocolos de identificación para diferentes aplicaciones. Haciendo esto, se podría mostrar que no existe un protocolo de identificación genérico que es el más adecuado para todas las aplicaciones, sino que muchas veces es necesario usar características típicas de las aplicaciones para mejorar las prestaciones de los protocolos de identificación para sistemas RFID.

Referencias

[1] Gildas Avoine. Adversary Model for Radio Frequency Identification. Technical Report LASEC-REPORT-2005-001, Swiss Federal Institute of Technology (EPFL), Security and Cryptography Laboratory (LASEC), Lausanne, Switzerland, September 2005.

[2] Gildas Avoine, Levente Buttyant, Tamas Holczer, and Istvan Vajda. Group-based private authentication. In *World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a*, pages 1 –6, 2007.

[3] Gildas Avoine, Etienne Dysli, and Philippe Oechslin. Reducing Time Complexity in RFID Systems. In Bart Preneel and Stafford Tavares, editors, *Selected Areas in Cryptography – SAC 2005*, volume 3897 of *Lecture Notes in Computer Science*, pages 291–306, Kingston, Canada, August 2005. Springer.

[4] Sepideh Fouladgar and Hossam Afifi. Scalable privacy protecting scheme through distributed rfid tag identification. In *Proceedings of the workshop on Applications of private and anonymous communications*, AIPACa '08, pages 3:1–3:8, New York, NY, USA, 2008. ACM.

[5] Ari Juels and Stephen Weis. Defining Strong Privacy for RFID. In *International Conference on Pervasive Computing and Communications – PerCom 2007*, pages 342–347, New York City, New York, USA, March 2007. IEEE, IEEE Computer Society.

[6] Ari Juels and Stephen A. Weis. Defining strong privacy for rfid. Technical report, 2006.

[7] David Molnar and David Wagner. Privacy and Security in Library RFID: Issues, Practices, and Architectures. In Vijayalakshmi Atluri, Birgit Pfitzmann, and Patrick Drew McDaniel, editors, *Conference on Computer and Communications Security – ACM CCS'04*, pages 210–219, Washington, DC, USA, October 2004. ACM, ACM Press.

[8] Agusti Solanas, Josep Domingo-Ferrer, Antoni Martínez-Ballesté, and Vanesa Daza. A Distributed Architecture for Scalable Private RFID Tag Identification. *Computer Networks, Elsevier*, 51(9), January 2007.

[9] Rolando Trujillo-Rasua and Agusti Solanas. Efficient probabilistic communication protocol for the private identification of rfid tags by means of collaborative readers. *Computer Networks*, In Press, Corrected Proof:–, 2011.

[10] Rolando Trujillo-Rasua and Agusti Solanas. Scalable trajectory-based protocol for RFID tags identification. In *IEEE International Conference on RFID-Technology and Applications (RFID-TA 2011)*. IEEE, (to appear) 2011. Available at: http://crises2-deim.urv.cat/docs/publications/conferences/633.pdf.

[11] Serge Vaudenay. On privacy models for rfid. In *Proceedings of the Advances in Crypotology 13th international conference on Theory and application of cryptology and information security*, ASIACRYPT'07, pages 68–87, Berlin, Heidelberg, 2007. Springer-Verlag.