Chapter 2
Anti-Counterfeiting and Supply Chain Security

Thorsten Staake[1], Florian Michahelles[1], Elgar Fleisch[1], John R. Williams[2], Hao Min[3], Peter H. Cole[4], Sang-Gug Lee[5], Duncan McFarlane[6], and Jun Murai[7]

[1]Auto-ID Lab St.Gallen, University of St.Gallen and ETH Zurich, Switzerland
{thorsten.staake,elgar.fleisch}@unisg.ch, fmichahelles@ethz.ch
[2]Auto-ID Lab MIT, Massachusetts Institute of Technology,
Department of Civil and Environmental Engineering, Cambridge, MA, USA. jrw@mit.edu
[3]Auto-ID Lab Fudan, Fudan University, Shanghai, China. hmin@fudan.edu.cn
[4]Auto-ID Lab Adelaide, School of Electrical & Electronic Engineering,
The University of Adelaide, South Australia 5005, Australia. cole@eleceng.adelaide.edu.au
[5]Auto-ID Lab ICU, Information and Communications University,
Daejon, Republic of Korea sglee@icu.ac.kr
[6]Auto-ID Lab Cambridge, Centre for Distributed Automation and Control,
Institute for Manufacturing, Department of Engineering,
University of Cambridge, Cambridge UK. dcm@eng.cam.ac.uk
[7]Auto-ID Lab Keio, Keio University SFC Research Institute, Tokyo, Japan
jun@sfc.keiu.ac.jp

Abstract: Counterfeit trade developed into a severe problem for many industries. While established security features such as holograms, micro printings or chemical markers do not seem to efficiently avert trade in illicit imitation products, RFID technology, with its potential to automate product authentications, may become a powerful tool to enhance brand and product protection. The following contribution contains an overview on the implication of product counterfeiting on affected companies, provides a starting point for a structured requirements definition for RFID-based anti-counterfeiting systems, and outlines several principal solution approaches that are discussed in greater detail in the subsequent chapters.

1 Counterfeit trade and implications for affected enterprises

Intangible assets constitute a considerable share of many companies' equity. They are often the result of extensive investments in research and development, careful brand management, and a consistent pledge to high quality and exclusiveness. However, the growth of markets in Asia where these intangible assets are difficult to protect, the trend in favour of dismantling border controls to ease the flow of international trade, and the increasing interaction of organizations in disparate locations require new measures to protect these assets and safeguard companies from unfair competition. Especially product counterfeiting, the unauthorized manufacturing of articles which mimic certain characteristics of genuine goods and which may thus pass off as products of licit companies, have developed into threats to consumers and brand owners alike.

Counterfeit trade appears to affect a wide range of industries. Alongside the traditionally forged items such as designer clothing, branded sportswear, fashion accessories, tobacco products, and digital media, customs statistics show a considerable growth of fakes among consumer products as well as among semi-finished and industrial goods including foodstuff, pharmaceuticals, fast moving consumer goods, electrical equipment, mechanical spare parts, and electronic components (e.g. TAXUD 2004). The implications are numerous and wide ranging. Counterfeiting undermines the beneficial effects of Intellectual Property Rights (IPR) and the concept of brands as it affects the return on investment in research, development, and company goodwill. Producers of reputable products are deterred from investing within a national economy as long as their intellectual property is at risk. National tax income is reduced since fake goods are largely manufactured by unregistered organizations. Social implications result from the abovementioned costs: the society pays for the distorted competition, eventually leading to fewer innovative products and a less secure environment as earnings from counterfeiting are often used to finance other illegal activities (ICC 2005). However, for selected emerging markets, the phenomenon also constitutes a significant source of income and an important element of their industrial learning and knowledge transfer strategy. As a consequence, not all governments determinedly prosecute counterfeiters, which often renders legal measures to eradicate the source of illicit goods ineffective.

For companies, counterfeit trade can lead to a direct loss of revenue since counterfeit products, at least partly, replace genuine articles, a reduction of the companies' goodwill as the presence of imitation products can diminish the exclusiveness of affected brands and the perceived quality of a product, and to a negative impact on the return on investment for research and development expenditures which can result in a competitive disadvantage to those enterprises which benefit from free-ride effects. Moreover, counterfeit trade can result in an increasing number of liability claims due to defective imitation products, and may facilitate the emergence of future competitors as it can help illicit actors to gather know-how which may enable them to become lawful enterprises in the future. These implications explain the vivid interest in organizational and technical protection measures – especially since established security features have apparently not been able to prevent the increase of counterfeit occurrences. RFID technology has the potential to overcome the shortcomings of the established technologies and may become a powerful tool for product and brand protection. However, the wide range of affected products and industries, the large number of stake holders, and last but not least the considerable reengineering capabilities of many illicit actors require a thoughtful solution design and thus a careful requirement definition.

2 Requirements for Auto-ID based anti-counterfeiting solutions

The specification of Auto-ID-based anti-counterfeiting technologies is strongly influenced by the security related requirements as well as by the design parameters which stem from an integration in the desired production and inspection settings. The security related requirements can be deduced from an attack model, whereas the additional practical requirements stem from interviews with industry experts at various workshops (e.g. from Special Interest Group Anti-Counterfeiting). Both aspects are discussed below, including a description of the potential capabilities of an attacker and the security-related constraints of RFID-based systems.

2.1 Attack model

A critical design parameter of anti-counterfeiting technologies is their desired level of security which can be defined as the cost and effort that is required to compromise or bypass the system.

Since the level of security strongly influences the cost of the solution, it should be carefully adjusted to the risk (i.e. the damage and probability of occurrence) imposed by counterfeit goods. Formal attack models allow for structuring the requirement analysis. In cryptography, such models usually take the form of an "experiment," a program that intermediates communications between a fictional adversary, and a runtime environment containing the system components (often referred to as oracles) (c.f. Juels 2006). Security models have to accurately reflect real-world threats (i.e. the capabilities of illicit actors) as well as the actual system characteristics. With respect to RFID, appropriate models should not only address the top-layer protocols, but also include the basic characteristics of RFID transponders down to the bit-level. The latter may lead to a less formal description but is necessary in order to capture potentially challenging threat scenarios like power analyses (and other side channel attacks) or destructive reengineering tests; in fact, while purely algorithmic models may help to evaluate cryptographic primitives and communication protocols, they do not sufficiently capture less standardized hardware attacks which impose realistic threats to RFID systems. Therefore, the attack model that is outlined below consists of a non-formal description of the system characteristics, the capabilities of the illicit actors as well as the identification and evaluation of the potential attack scenarios.

System Capabilities

Low-cost RFID transponders are limited with respect to their maximum transistor count (as the chip size influences the transponder cost), the available energy (due to restrictions of the transmitting power of readers, the size of the antenna, and the often required considerable distance between tag and reader devices), and the frequency spectrum. This ultimately results in limited computational power, confines the memory size and communication bandwidth, and hampers the integration of sophisticated pseudo random number generators or sensors against hardware attacks.

Since cryptographic operations usually rely on computationally intense primitives, complex encryption procedures are difficult to realize in low-cost transponders. Even promising proposals that outline a lean integration of established security standards in RFID devices (e.g. Feldhofer et al. 2004) would dramatically increase the energy consumption and the required communication bandwidth, and thereby lead to lower read ranges and reduced bulk reading capabilities.[1]

More complex – and thus more expensive – transponders allow for more sophisticated cryptographic measures. In principle, the complexity of the design can escalate up to those of battery powered smart cards (i.e. active tags) with public-key crypto systems. When evaluating a potential migration path towards more secure systems, it should be considered that the silicon chip only constitutes one cost factor of the device (besides the cost for packing and the antenna) and that doubling the gate count does not necessarily double the price of the of the transponder.

Another relevant characteristic of RFID results from the radio connection between tag and reader. Connectivity is connectionless and communication is provided over an unreliable channel. This allows illicit actors to listen to the data exchange and, for example, detect existing identification numbers. Moreover, conflicts have to be considered when sharing the channel. Due to the limited power of the readers and computational constraints among tags, a more powerful sender can easily jam legitimate readers (Walters et al. 2006). An intentional violation of the tag-to-reader communication protocol, e.g. by continually transmitting messages in an attempt to generate collisions, can also disable a meaningful data exchange, which gives rise to several potential attacks.

Capabilities of Illicit Actors

The computational power and hardware complexity of low-cost RFID transponders is rather limited compared to the potential capabilities of illicit actors. Moreover, the unattended and distributed deployment of RFID transponders makes the devices highly susceptible to physical attacks. In fact, the access of the adversary to the system is a critical parameter of the attack model. Most cryptographic security analyses base on the assumption that illicit actors are able to experiment extensively with the elements of the system (e.g. Bellare et al. 1998), and thus are able to submit a large number of "oracle" queries to expose weaknesses of the design or to "guess" secret information. In this context, the limitations of RFID systems also restrict the capabilities of the attackers; illicit actors may have unlimited access only to selected transponders (e.g. after purchasing original articles with the security feature still in place), but limited access to arbitrary components. The latter is the case since attackers can only read tags which are in close proximity to their reader devices, or listen to tag reader communications which are within eavesdropping range (see Juels (2006) for a definition of various read ranges). However, in most supply chain related applications, the vast majority of transponders are hidden to other parties most of the time.

[1] In some scenarios, however, the application of such systems is nevertheless meaningful. Integration in existing standards is discussed later in this section.

With respect to transponders that are in the possession of the attackers, a wide variety of tools is available. Potential steps include power analyses and the exact measurement of response times, the application of different clock speeds or the elimination of the air interface in order to increase the frequency of queries, and hardware attacks (e.g. opening the packaged IC) in the attempt to directly read out key registers on the circuit or to reverse engineer the underlying algorithms. Therefore, when designing RFID-based anti-counterfeiting features, care must be taken that compromising accessible transponders does not affect the security of the remaining system. The protection should base on secret keys which are different and non-related among the tags rather than on secret algorithms that a large number of transponders may have in common.

Attack Scenarios

A simple attack model for low-cost RFID devices is provided by Juels (2004), who mainly addresses threats to data security, authentication, and privacy. With respect to anti-counterfeiting features, however, the focus of potential attacks is shifted to an extended set of threats. Interviews with brand protection experts conducted during this research revealed the relevance of the following issues: tag cloning which is strongly related to tag authentication, obfuscation and deception, tag omission, removal-reapplication, and, new in the context of product security features but frequently discussed in computer security, denial-of-service attacks. Each issue is addressed below.

Cloning refers to the duplication of security features such that they are likely to pass of as authentic during inspection. With respect to RFID, tag cloning may be defined as the replication of a transponder with the duplicate being able to emulate the original tag's behavior. In a system with cloned entities, investigators (or reading devices) can no longer ensure that the distinguishing mark they observe originates from the correct source; moreover, without taking the existence of duplicate features into account, observers would even falsely certify the authenticity of bogus components. Large scale tag cloning attacks can severely compromises anti-counterfeiting solutions and therefore should be addressed during the system design.

Obfuscation connotes the use of misleading protection technologies. In practice, licit companies frequently change security features to prevent counterfeiters from copying or cloning their protection technology. While following this paradigm of "creating a moving target", the licit parties unintentionally complicate the inspection process. Especially third parties can be overwhelmed by the coexistence of different, mostly visual security features. Consequently, counterfeit producers can often rely on the lack of knowledge (and the lack of time and motivation to acquire it) during inspection processes. A very common attack stems from the application of security mechanisms which are not related to the genuine product, such as the use of holograms instead of micro printings or flip colors instead of complicated packaging designs. However, the need to change anti-counterfeiting primitives when they become ineffective as well as their user-friendliness given the limited resources during inspection translates into the requirement of a flexible security system with a static user interface.

In anti-counterfeiting systems that rely on more than one component, threats may not only originate in bogus product security features but also in malicious backend systems. When a barcode, a micro printing, or an RFID transponder references a database containing track and trace information or advanced shipment notices, the authenticity of the relevant source has to be verified.

Tag Omission, i.e. the abdication of the security features by counterfeit producers even if the corresponding genuine articles are equipped with protective measures, relies on low inspection rates among many categories of goods. The phenomenon shows the need of large scale and consequently low-cost inspections. Preferably, inspections can be automated even in loosely guided processes as given in many warehouses, at customs, or at retail stores.

Removal-Reapplication attacks refer to the application of genuine security features from (mostly discarded) genuine products to counterfeit articles. This constitute a potential threat for tagging technologies where security features are being attached to an object (like holograms or RFID transponders) rather than being an inherent part of it (such as chemical markers). The consideration of this attack is of importance especially when protecting high value goods like aviation spare parts which are, when out of service, often still accessible to illicit actors. When relying on tagging technologies, a defense is to tightly couple the security feature to the object, e.g. by tamper-proofing its physical package or by establishing a logical link between object and tag.

Denial-of-Service Attacks may be defined as "any event that diminishes or eliminates a network's capacity to perform its expected function" (Wood and Stankovic 2002). Since established anti-counterfeiting technologies usually do not rely on network resources, this attack is new to the brand and product protection domain. However, when authentication processes involve entities in disparate locations, the access to these resources may be disturbed. With respect to RFID devices, attacks can cut off the connection between individual transponders and reading devices. When illicit actors target major distribution centers or customs, e.g. at harbors or airports, denial-of-service attacks may severely slow down inspection processes and thus interfere with the unobstructed flow of goods.

Eliminating any possibility of such attacks is difficult on a technical level given the limited functionality of low-cost transponders. However, providing tools for detecting attacks and localizing the illicit device is not a major issue. In actual systems, the operator would have to physically remove or deactivate the attack device.

2.2 Practical Requirements

The attack model led to a set of security related requirements. They include measures to avert a duplication of security features; the design of a stable, easy to use interface; the necessity to efficient inspection processes at low cost even in loosely guided processes; a tight coupling of the security feature to the object; and measures against denial of service attacks. In addition to this set, a number of – partly interrelated – conditions stem from the practical requirements on anti-counterfeiting and supply chain security solutions which are not directly related to breaches of security:[2]

- *Different levels of security:* The desired level of security has a major impact on the fixed and variable costs of the solution. It can be determined i) by the risk or cost resulting from a compromised system, and ii) by the lifetime of the object which is to be protected. Risk or cost can be classified in terms of the potential health and safety hazards for consumers, or the incremental financial losses of licit manufacturers and brand owners. Depending on the probability of individual occurrence, health and safety hazards may require highly secure systems. In the context of RFID, these can be realized by the application of complex cryptographic primitives (e.g. public-key-based authentication mechanisms implemented on certified RFID transponders); for critical spare parts in the aviation industry, for example, the cost of RFID transponders may be as high as 10 EUR or above. However, if illicit products primarily cause incremental financial losses (e.g. due to dissatisfied consumers and substitution effects), a detailed cost-benefit analysis is helpful in order to select an appropriate protection mechanism.
- *Migration path:* Anti-counterfeiting technologies often constitute a barrier for illicit actors only for a limited, unknown period of time. Holograms, for example, have been considered highly secure features when introduced and are now widely available on the market. Consequently, it is desirable to have the opportunity to change the underlying security primitive at low cost, i.e. without the need to alter the technical infrastructure or to require the user to get accustomed to new checking procedures. RFID technology, if properly designed, allows for separating user interfaces and underlying technologies, and may therefore constitute a sustainable solution.
- *Manufacturing requirments:* Existing manufacturing settings are often highly optimized with respect to throughput and down times. The addition of supplementary process steps can severely impact the key performance measures of the production facilities. This is especially the case in high volume production environments e.g. in the pharmaceutical or fast moving consumer goods industry, where the required line speeds severely limit the technology choice. Process steps that are necessary to integrate security features have to be as non-intrusive as possible.
- *Product specific requirements:* Prouct related characteristics can impose a number of additional constraints on the technology choice. Restrictions may result from the available size for such features, the object's material, and operating conditions such as temperature, electrical discharge, abrasion etc. When the security features are to be deployed at an early stage of the production process, aggravated conditions may apply. The product specific requirements should be analyzed on a case-by-case basis at an early stage of the design process.

[2] The practical requirements result from a group work undertaken during the second Special Interest Group workshop in Hamburg at July 1, 2005.

- *Invariance of the product design:* I order to enhance the level of security, it is desirable to integrate the security features in the product and not to rely on tagging its packing. However, companies seem to be rarely willing to subordinate product design to anti-counterfeiting measures. This limitation may further complicate the tag-in-product integration.
- *Technology specific requirements:* Idividual security technologies may be chosen due to the specific advantages they exhibit such as the possibility to automate checking processes, which may have to be defined in greater detail. With respect to RFID, bulk reading (i.e. the number of tags which can be read quasi-simultaneously; read rates (i.e. the share of transponders which is actually detected during a bulk read); read ranges (i.e. the maximum distance between tag and reader during the inspection process); data standards, etc. are to be considered.
- *Confidentiality:* Last but not least, securiy features shall not reveal confidential information of the manufacturer (e.g. on production output) nor infringe the privacy of the user or consumer.

Depending on the actual application, several solutions concepts are applicable which are outlined below.

3 Solution concepts

RFID technology comes at various levels of complexity and offers several functionalities which make it applicable as anti-counterfeiting measures in various application scenarios. The following section discusses in greater detail the usability of unique serial numbers, a technique to avert removal-reapply-attacks, and the usability of tags with authentication capabilities in a standard reader environment.

3.1 Using unique serial numbers

Marking objects with unique identifiers, i.e. on item-level rather than for individual project categories only, helps to monitor the flow of goods and thus to detect illicit trade activities. If designed carefully, a numbering system can significantly reduce counterfeit trade. The latter is possible if an approach is chosen which is difficult to apply for illicit actors, but whose identifiers are easy to check for supply chain partners or end-users. In an ideal scenario,

- the number is assigned in a random way, with the numbering space significantly larger than the number of items to be identified, so that illicit actors are unlikely to simply guess valid IDs,
- the validity of the number can be easily checked by the supply chain partner or, if desired, by the consumer,
- the number can be read automatically by authorized persons, allowing for large scale searches for invalid or duplicate identifiers, thus increasing the chance to seize illicit goods,

- a duplication of the number carrier is unreasonably expensive, and
- the number carrier cannot be removed nor can illicit actors overwrite the number which would allow them to disguise the identity of the object.

The basic operating principle of a unique ID system is quite simple: The manufacturer generates a random number, writes it to the data carrier and stores it in a database. When the product ID is checked e.g. in a store or at customs, a reader device retrieves the product ID, sends it to a service offered by the manufacturer (or an IT provider) which looks up the number in the database and returns the result to the reader device. An operational implementation, however, should provide additional features such as a system for user access management that prevents illicit actors from discovering licit numbers or competitors from monitoring the flow of goods. When the system is applied by a larger number of vendors, it becomes impractical to store the access information of individual service providers on the reading devices; therefore, the system should contain a lookup system which allows the readers to retrieve the corresponding addresses from a known online source.

3.2 Plausibility checks based on track and trace

A track and trace system is a potentially powerful tool as it can provide an enormous degree of supply chain visibility. In principle, information on an object's location and the corresponding time, possibly together with data on the owner, its status, the object's operating conditions, etc., is recorded and stored for further processing. If such measurements are repeated over time, they allow for plausibility checks of the recorded products history. Heuristics can be applied as for example done by credit card companies which routinely freeze cards if they inhabit a suspicious transaction history.

Track and trace systems rely on the ability to uniquely identify individual articles. In order to facilitate meaningful analyses, numerous data points have to be collected, which requires an efficient way to capture supply chain events. In this regard, RFID can be seen as an enabling technology. Though the operating principles of track and trace systems may appear simple, an actual implementation of the infrastructure is a considerable challenge. From a technical perspective, especially access management in non-predetermined supply chains constitutes a major hurdle. However, even bigger barriers seem to be organizational issues on the ownership of the data, the distribution of system costs, and the lack of interest among some industries to provide a higher degree of supply chain visibility for their customers. In fact, the solution requires numerous stakeholders – often with conflicting interests – to work together, which renders it impractical in many application scenarios. However, track and trace is likely to become the dominant solution in highly regulated industries where powerful stakeholders can enforce an adoption (e.g. the Food and Drug Administration with respect to pharmaceuticals).

3.3 Object specific security

Security solutions that are based on tagging technologies have a system specific drawback: when checking an object, it is still the tag (e.g. a hologram or a simple RFID transponder) which is authenticated and not the object or document the tag is attached to. The link between tag and object is often not strong enough. In theory – and also in practice if the solution is not designed properly – a tag can be removed from an original article and attached to another object, thereby compromising the security system.

In contrast to most other tagging technologies, RFID can overcome this shortcoming. Even low-cost RFID tags with a certain amount of memory can store data that binds a tag to a given product, as a picture in a passport binds a passport to its holder. As a result, illicit actors are detained from simply removing a tag from a legitimate product and reapplying it to a counterfeit article in a way that the fake is not detected during product validation. An in-depth description of this technique is given in the chapter 13, "Product Specific Security Features Based on RFID Technology".

3.4 Secure authentication

Concepts that allow for a proof of identity (or authentication) are common in computer systems. However, establishing efficient means of authentication in RFID infrastructures constitutes a major challenge. The lack of cryptographic functionalities of basic RFID transponders is a big impediment to current designs. Since serial numbers are usually not read protected, an attacker can obtain an identifier from a tag and program it into another transponder, or emulate the tag using other wireless device. If done at a larger scale, duplicate devices render track and trace or anti-counterfeiting solutions ineffective.

Challenge-response protocols can avert tag cloning as they allow for a comparison of secret keys at disparate locations without transferring them over a possibly insecure channel. In a carefully designed system, third parties are prevented from reconstructing the secret, even if it is used numerous times. These properties qualify challenge-response protocols for an application in an RFID environment, where the channel must be regarded as insecure.

Critics of this approach frequently mention the increasing tag costs which may result from the integration of the required cryptographic unit in RFID transponders. However, Feldhofer et al. (2005) showed an implementation of an 128 bit version of the Rijndael cipher (Daemen and Rijmen 2002) using less than 4,000 gates, which, given an approximate gate count of current EPC Gen2 tags of 15,000, would only lead to a small increase in tag cost.[3] This motivates further research on the actual integration of authentication protocols in RFID systems and thus is a major topic in the remainder of this book.

[3] This is especially true since the cost of the actual chip is only one component of transponders which is also made up by packing, the antenna, assembly, etc.

References

1 Daemen, J. and Rijmen, V. 2002. The design of Rijndael: AES – the Advanced Encryption Standard. Berlin, Germany: Springer

2 Feldhofer M., Dominikus S., and Wolkerstorfer J. (2004). Strong authentication for RFID systems using the AES algorithm, Proceedings of the Workshop on Cryptographic Hardware and Embedded Systems (CHES '04), pp. 357–370

3 Feldhofer, M., Wolkerstorfer, J., and Rijmen, V. 2005. AES implementation on a grain of sand. In Proceedings, Information Security, 152(1): 13–20

4 ICC International Chamber of Commerce (2005). Current and emerging intellectual property issues for business – A roadmap for business and policy makers. Document no 450/911 Rev. 6 (Paris, France: ICC, March 2005), p. 2. www.insme.info/documenti/Roadmap-2005-FINAL.pdf.

5 Juels A. (2004). Minimalist cryptography for low-cost RFID tags, RSA Working Paper, www.rsasecurity.com/rsalabs/staff/bios/ajuels/publications/minimalist/Minimalist.pdf.

6 Juels, A. (2006). RFID security and privacy: a research survey, IEEE Journal on Selected Areas in Communications, 24(2): 381–394

7 TAXUD European Taxation and Customs Union (2004). Breakdown of the number of cases registered and the number of articles seized by product type: EU – 2004, http://ec.europa.eu/taxation_customs/resources/documents/customs/customs_controls/counterfeit_piracy/statistics/counterf_comm_2004_en.pdf.

8 Walters, J. P., Liang, Z., Shi, W., and Chaudhary, V. 2006. Wireless sensor network security: A survey. In Xiao, Y. (Ed.), Security in Distributed, Grid, and Pervasive Computing: Chapter 17. Sound Parkway, NW: CRC Press

9 Wood, A. D. and Stankovic, J. A. 2002. Denial of service in sensor networks. Computer, 35(10): 54–62