A Pharmaceutical Anti-counterfeiting Method Using Time Controlled Numeric Tokens

Emil Nilsson, Björn Nilsson, and Eric Järpe
CERES and MPE-lab
Halmstad University
Halmstad, Sweden
emil.nilsson@hh.se

Abstract—An anti-counterfeit and authentication method using time controlled numeric tokens enabling a secure logistic chain is presented. Implementation of the method is illustrated with a pharmaceutical anti-counterfeit system. The method uses active RFID technology in combination with product seal. Authenticity is verified by comparing time controlled ID-codes, i.e. numeric tokens, stored in RFID tags and by identical numeric tokens stored in a secure database. The pharmaceutical products are protected from the supplier to the pharmacist, with the possibility to extend the authentication out to the end customer. The ability of the method is analyzed by discussion of several possible scenarios. It is shown that an accuracy of 99.9% telling the customer she has an authentic product is achieved by the use of 11-bit ID-code strings.

Keywords—Active RFID, security, pharmaceuticals

I. INTRODUCTION

Today there is an increasing problem with fake drugs, all over the world, the weak part being the logistic chain from supplier to the pharmacist. To pave the way for security in the logistic chain, RFID could be used. Drugs would be tagged in order to counteract forgery and theft. Reports from the European Union state that counterfeit goods increased by 13% in 2008 with a 50% increase in fake medicines [1].

There are several opportunities for stealing, counterfeiting, and manipulation of drugs in the logistic chain and therefore a growing need to verify the authenticity of goods delivered. To this end, a system based on existing seal technologies in combination with a novel authenticity method, is proposed. The new method is based on the low tech “Calling in the Numeric Token” (CNT) strategy proposed in [2], but with a higher degree of automation and protection against, copying or guessing numeric tokens [3]. The novel method offers unique ID, anti-copy, and tampering evident functionality by use of time controlled numeric tokens (TCNT). TCNT is intended to be implemented in combination with a strong sealing technology. With the proposed new TCNT method, efforts to forge the identity of the shipped goods would have to be extremely strong. Further, any tampering with the goods would easily be detected.

II. ANTI-COUNTERFEIT AND AUTHENTICATION TECHNOLOGY

A. Related research in RFID

RFID has been used in for tracking and verifying of goods in the supply chain in different applications. Although RFID offer new possibilities it is also vulnerable to new types of hostile attacks, such as eavesdropping, denial of service (DoS) and cloning [2,4,5]. Issues concerning security and privacy are mainly handled by the communication protocol. The computational cost is increasing together with the complexity and safety provided by the system. When multiple communication rounds are used between a tag and a reader the energy cost is increased. In this article active RFID is discussed and the energy consumption is of great importance to increase the tag life time. Technologies used today for authentication uses multiple communication rounds between server-reader-tag and would consequently drain the tag energy source (the battery) faster if used in an active RFID system [6][7][8]. Other system relies on reference tags being securely distributed [9].

In this paper the idea of storing a bit string in a hardware serial register rather than having a high performance processor that needs to process secret keys to solve authentication arithmetic is presented.
B. Introduction to TCNT

The TNCT method is based on the low tech CNT strategy. The CNT strategy uses numeric tokens, ID-codes, being associated with the pharmaceutical product. There are three rules for the strategy. 1) The ID-codes should be random and unique for each product container within a given Lot. 2) The ID-code should not be a serial number and it should be unpredictable. 3) There must be at least 1000 times more possible ID-codes than actual product containers for a given Lot. Valid ID-codes are stored in a database controlled by the supplier. When a customer wants to check the validity of their purchased product they call in the numeric token, i.e., ID-code, they find on the container. If the ID-code is a valid number the customer is told that the product is authentic. However, if some earlier caller has called in an identical ID-code the customer is told that their product is a counterfeit.

Obviously this system is simple, but it is also hazardous for the first caller who is not properly warned. The TCNT method is based on the same philosophy as CNT but with a higher grade of automation, and mitigating the risk for the first caller by having time limited validity of the ID-codes. Further, uniqueness of the ID-code is granted over a rolling period rather than for a Lot.

The method is based on a secure database controlled by the supplier or manufacturer of the pharmaceutical product, see Fig. 1. The database is used to store lists of time controlled ID codes which also are imprinted in a volatile electronic memory on each tamper proof active RFID tag. The ID-codes are generated by a pseudorandom code generator when the tag is configured for shipping. Each tag is equipped with a clock which is being synchronized with the clock utilized in the database. A seal contains the tag and the product to be delivered. When the seal is broken, either by someone trying to manipulate the goods or by the customer opening the product container, the tag is triggered. During such an event several tasks are performed by the tag. The ID-code associated with the time being recorded by the internal clock is written to an internal register. The tag is now waiting for an RFID-interrogator with a specific terminal ID, associated with the pharmacist or customer receiving the goods, to read the content in the register. The transmitted time controlled ID-code is compared with the time controlled ID-code pointed out by the database common clock.

III. IMPLEMENTATION

A complete system, see Fig. 2, would consist of the following components, an RFID tag with internal clock, a secure database server with an internal common clock, terminal interrogators (PDA), a network connection (WAN), product seals, and a Pseudo Random (PN) ID-code generator.

At the time of outgoing delivery from the supplier, a reset signal transfers the tag from sleep mode to serial programming, see Fig. 3. A list of ID-codes are generated and written into the volatile memory of the RFID tag. The same list will also be written to the memory of the common database. Each ID-code in the list is valid for a specified time interval $T_{slot}$. Further the ID of the end customer RFID interrogator is written into the tag. During this data transfer the tag is connected to a dedicated programming device by a physical connector. Once the tag is programmed it will be inserted to a sealing together with the goods to be delivered. Several different tamper proofing technologies may be used to detect an event of tampering. Examples are, sensing current conducting wires wound around the goods, tear strips, or pressure sensitive sensors together with gas tight sealing bags. In the latter case the following scenario may be imagined. The atmosphere inside the sealing bag is evacuated, and the resulting pressure drop triggers the tag into its transport mode. The timer is running during the complete delivery, until stopped by someone opening the seal. Manipulation or normal opening at delivery may be monitored by a pressure sensor inside the gas-tight sealing bag. When the seal is broken the under-pressure is neutralized by the surrounding atmospheric pressure. At this tampering event the ID-code in the list pointed out by the timer...
is transferred to the transmission (TX) register. All other codes in the list are being erased. The now selected and stored ID-code is wirelessly transmitted when the RFID tag is contacted by an interrogator with correct ID. Note that this interrogator ID might not necessarily be bound to a specific hardware, but rather being generated from e.g. a customer debit- or credit-card. The RFID technology used may be a commercial available, hosted in a smartphone or similar PDA. After the transmission of the ID-code the RFID tag may return to sleep mode, see Fig. 3. Since the tag now not contains any compromising information it may be recycled, depending on the RFID technology platform being used.

IV. Analysis

The TNCT authenticity method has been analyzed by discussing possible scenarios in a delivery chain that might compromise the functionality. However, before going into this discussion it is important to make an analysis of some vital properties of the system. The system is based on synchronization between the tag clock and the database common clock. This functionality protects the tag from being easily cloned.

A. Clock drift and synchronization

An internal low cost, non crystal, clock is used by the RFID tag for selection of the ID-code stored in slots of the list. A drift value $\epsilon$ is associated with the clock. Precision is limited by the power budget of the electronics and by the will to minimize the number of lumped components in the electronics. The total clock skew is

$$\Delta t = \epsilon T_D$$

(1)

where $T_D$ is the total delivery time. Irrespective of how small the clock drift might be, the system has to accept two possible readout results around the ID-code slot shift occasion. The slot dwell time $T_{\text{slot}}$ must at least be obeying the following criteria if not even larger ID-code slot intervals must be taken into account,

$$T_{\text{slot}} > 2|\Delta t|$$

(2)

Since we have to accept two different ID-codes any given time, the probability $P$ of guessing the next ID-code stored in the list is

$$P = 2/N_{ID}$$

(3)

where $N_{ID}$ is the number of distinct codes. (As an example: someone guesses a code and writes it to each time slot in the list). If we are to address each single slot with a unique code we have to, in each tag, store a number $K$ of codes according to,

$$K = (T_D + \Delta t)/T_{\text{slot}}$$

(4)

If $K$ is larger than $N_{ID}$ a looped slot memory may be used without increasing the risk of correct guesses by any forger.

For each delivery to a customer with a specific terminal ID we have to use unique ID-codes over a specified period and associated with a specific time. The number of lists, $M$, of permuted combinations possible to generate out of $N_{ID}$ ID-codes and $K$ positions is

$$M = N_{ID}/(N_{ID} - K)!.$$  

(5)

However, this number is reduced by the limiting restrictions due to requirement that each slot position corresponding to a specific time has to be unique for each list. Further, in order to minimize the interference of the clock drift the neighboring positions also must not have the same ID-code. Without the clock drift the number of possible lists, $M$, are merely $N_{ID}$. Including the effects of clock drift, limited by (2), the number of code lists shrinks to

$$L \equiv \lfloor N_{ID}/K \rfloor \lfloor K/2 \rfloor,$$  

(6)

where $\lfloor x \rfloor$ denotes the integer part of $x$.

B. Cloning of the RFID tag

The method with time controlled ID-codes is supposed to protect system from attackers trying to clone the RFID tag. An attempt to break the seal could be carried out either by the customer or by an intruder; this would have the consequence that the RFID tag sends the time controlled code. All the other ID-codes stored in the list are erased immediately after writing to the TX register. An attempt to base a cloned RFID on the transmitted ID-coder is useless, unless cloning and the counterfeit delivery is made within the remaining time interval of the slot. This remaining time is not known exactly by the forger, but can as a maximum be $2T_D$. The end customer may mitigate this risk simply by suspending the opening of the delivery enough time to ensure a shift of ID-code. Further, the imposter has to know the RFID-reader-terminal ID to be able to read the RFID tag time slotted code.

A forger who wants to read the complete memory content of the chip must do it when it is in a live and running state. The content of the volatile ID-code memory is destroyed when the power supply is broken. The protecting glue on top of the chip makes it necessary to mechanically drill down to the chip metal wire layers. Bond wires may be designed to cross over the chip embedded in glue top, see Fig. 4. Breaking the wires will break the supply, leading to a memory loss and failure to read out the ID-codes. The battery could be placed on the opposite side of the carrier material for maximum protection. Making an effort to circumvent these precautions requires expensive equipment and personnel, but also a large amount of time. By that time the obtained list would be obsolete.

![Figure 4. Configuration of RF-ASIC, PCB, and battery.](image-url)
The RFID tag electronics is reached via the writing port. This is supposed to be a one-way path isolated from the by data-buffers and thereby protected from eavesdropping. The RFID tag can only be written by mechanically connected writing device. This means that the seal has to be broken before attempting any attack. As discussed before only the currently valid ID-code will be available then.

The geographical position and IP-number in combination with debit card number may be included as part of the authentication process and thereby hedge against the event that someone steals the whole RFID-reader terminal.

Instead of attacking the RFID tag physically a forger may try “phishing” for valid ID-codes by reporting random ID-codes and see if the system accepts them. Again the obtained ID-code will only be valid for a limited time. Further, the attacker does not know what specific delivery the code is validating.

An attack from the inside of the system can be imagined, where list of ID-codes are stolen. However, the attacker must be highly organized since stolen lists of ID-codes are only valid for a specific customer who is expecting a specific brand or type of drug within a specific time. The ID-code lists are only generated at the time of outgoing delivery. The forger would need costly storing capacity for different kinds of pharmaceuticals and must have a very quick response. This kind of attack can largely be avoided by having a good security.

C. Scenarios

Below a number of scenarios are discussed. They illustrate problems and their possible solutions. The scenarios cover hostile attacks by an attacker as well as normal operations that might compromise the functionality of the system. Scenario 1), 2), and 3) cover theft and attempted thefts. Scenario 4) and 5) describe organized attacks against the tag and attacks to the system. In scenario 6) customer integrity is briefly mentioned, and finally scenario 7) discusses how the system handles the resale of drugs.

1) Scenario 1
An attacker breaks the seal and steals the goods. The attacker may try to sell the goods somewhere else, as untouched, diluted or as a different product:

When the seal is broken the RFID tag will lose the information written into the list of ID-codes, and therefore it will become useless for reuse by the intruder. There is a possibility to read out the currently valid ID-code by an interrogator impersonating the ID-code of the end customer. However this information will only be useful for a maximum time period of $2T_{idc}$. A counterfeiter might guess the ID-code, but the odds of guessing a correct code is very small according to (3).

2) Scenario 2
An attacker breaks the seal and fiddle with the content but lets it through to the customer:

Scenario 2) is practically the same as scenario 1). The RFID tag will be useless as in first case. The customer would note the intrusion by not being able to confirm the ID-code of the goods.

3) Scenario 3
An attacker steals the un-tampered package and delivers it somewhere else:

The ID of the interrogator of the final customer will not match the interrogator ID stored in the RFID tag. Further the ID-code read out from the RFID tag will not be matched to any ID-code list stored at the supplier. The transfer of the information is under control of the end customer and the supplier.

4) Scenario 4
Selling counterfeit product. Trying to use cloned or impersonating RFID-tags:

In this scenario the sealing technology would prevent insertion of any impersonating or cloned RFID tag. Further the actual cloning of the RFID tag would be impossible due to the time controlled ID-codes used, see previous discussion.

5) Scenario 5
Trying to break the system by DoS- or spoofing-attacks:

The CNT system has to handle Denial of Service (DoS) attacks, Spoofing attacks, fake call-in sites, and innocent redundancy. DoS-attacks are still a potential problem for TCNT, and have to be mitigated the same way as in CNT. For instance the caller could be forced to ID himself so multiple calls by the same ID can be detected. There could be private systems set up for high volume customers. Spoofing by calling in valid IDs several times is no problem in TCNT since the valid code is changing with time. By contrast to CNT there is no need to raise the detection threshold of callers before reporting fake drugs. Fake call in sites might be a problem and has to be solved similar as with CNT, with information from trusted sources. Innocent redundancy call-in is no problem with TCNT. The system is automated and will only call in ones. The customer could be diverted to fake call in sites set up to always confirm the authenticity. If this turns out to be a problem, the system might be configured to do the comparison of ID-codes locally at the end customer PDA. Thus the customer requests for the valid ID-code from the common database and compares this with the ID-code obtained from the RFID-tag.

6) Scenario 6
Illegitimate tracking of the goods:

This scenario is not applicable here. The TCNT method is not intended for continuous tracking of the tag and the goods, but rather for authentication controlled by the end customer. However the TCNT method works as a complement to existing RFID technologies, which in turn may provide tracking services.

7) Scenario 7
Drug resale and repacking:

Drugs are sometimes resold by an original purchaser and containers might be sub-dived. If each subdivided container is already tagged from the supplier there is no problem to be
handled, the end customer will check the validity as before. However, if the container in order to be subdivided has to be broken together with the seal, the TCNT is effectively transformed to a CNT system. The original purchaser test the validity as described and register the so obtained ID-code. This ID-code may now function as a numeric token that might be called in. The security is lower for the end customer, but still is good.

V. NUMERICAL EXAMPLES

A typical drift value for a good clock is $\varepsilon = \pm 50$ ppm [10], but for a low power integrated RC oscillator used in sleep mode it can be in the range of $\varepsilon = \pm 1000$ ppm [11]. With delivery times $T_D$ of up to two weeks the system has to absorb time skews up to $+/-20$ minutes according to (1). The delay over WAN is typically much smaller than this value. On the other hand the risk for reading an ID-code in a prior or later slot than the database reference is not negligible, and the risk increases the longer the delivery time. Using (3) we find the probability of guessing the correct code to be smaller than 0.1%. The number of codes must be $N_{ID} > 2000$. Further, using (4) with a delivery time of two weeks and time slots of one hour we get $K > 337$. The number of codes possible to generate from a string of 11-bit length is 2048. Equation (6) gives the number of lists, $L$, of permuted combinations possible to generate out of 2048 ID-codes and 337 positions to be 840.

The numbers in our example gives us the possibility to ship more than to 800 deliveries within a rolling period of two weeks to a single customer. If we accept longer strings, say 16-bit codes, we will get a risk of 16 ppm that someone is guessing the correct code and more than 32,000 possible unique deliveries within the time period to a specific customer.

A comparison of the efficiency between CNT and TCNT using 11-bit numeric tokens is found in Table I. The accuracy for CNT is defined as the ability to tell the next caller that he has a counterfeit under the assumption that the same valid bottle ID has already been calle4d in multiple times. There exist one authentic bottle, and C counterfeit bottles, all with the same valid Bottle ID [2]. The accuracy for TCNT is defined using (3), i.e. the possibility of someone guessing the correct numeric token from a pool of $N_{ID}$ codes. Since different accuracy definitions are used in the comparison of the two methods conclusions must be made carefully. TCNT offers improved protection primarily at small numbers of counterfeits and also protection for the first caller. Further it can be seen that the accuracy easily can be made higher in TCNT method with longer ID-code strings. Further it is important to stress that, in contrast to CNT, the TCNT method offers protection also to the first caller.

The possibility to sell a guaranteed authentic medicine will enable responsible pharmacists to defend their position as trusted dealers and thereby giving them a competitive advantage.

TNCT enable the end customer to label the shipment with her credit card number during purchase on e.g. a web-store. The RFID tag will wait for an interrogator with the same ID as the debit- or credit-card number.

We have here assumed that the customer can order directly from a trusted supplier, and that the drug is delivered within a period of two weeks. We further assume is possible to set up web-pages and PDA applications that can communicate the information stored in the RFID-tag.

ACKNOWLEDGMENT

The authors want to thank Sten Wandel, researcher at NGIL at Lund University of Technology, for valuable discussions and viewpoints.

REFERENCES


<table>
<thead>
<tr>
<th>No. of counterfeits made (C)</th>
<th>First caller</th>
<th>Next caller</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CNT accuracy not defined</td>
<td>TNCT accuracy $= 2N_{ID}$</td>
</tr>
<tr>
<td>2</td>
<td>Not def.</td>
<td>99.9%</td>
</tr>
<tr>
<td>10</td>
<td>Not def.</td>
<td>99.9%</td>
</tr>
<tr>
<td>100</td>
<td>Not def.</td>
<td>99.9%</td>
</tr>
<tr>
<td>1000</td>
<td>Not def.</td>
<td>99.9%</td>
</tr>
</tbody>
</table>