2009

Taxonomy development in information systems: Developing a taxonomy of mobile applications

Nina Oertel
SAP Research CEC Karlsruhe, nina.oertel@sap.com

TRACKING BASED PRODUCT AUTHENTICATION: CATCHING INTRUDERS IN THE SUPPLY CHAIN

Oertel, Nina, SAP Research CEC Karlsruhe, Vincenz-Priessnitz Strasse 1, 76131 Karlsruhe, Germany, nina.oertel@sap.com

Abstract

Product counterfeiting is a growing problem worldwide, threatening the health of consumers and reducing company profits. By detecting and intercepting counterfeits before they reach the customer, the problem can be mitigated. In this paper, an approach to detect counterfeit items based on their claimed history is presented. The necessary data is provided by tracking infrastructures that enable the recording and retrieval of movements of individual items in the supply chain based on unique identifiers assigned to products. If the movement history of an item deviates from the movements of genuine items that have been learned before, a warning about a potential counterfeit is issued. Counterfeiter activities that are possible in a tracking enabled environment are modelled and the capability of the proposed approach to detect these strategies is assessed.

Keywords: Anti-Counterfeiting, Authentication, Supply Chain, RFID

1 INTRODUCTION

Counterfeiting is growing in terms of industries and countries affected, as well as the number of fake goods seized by customs each year. The OECD (2007) estimates that the international trade in counterfeit products was up to 200 billion $ in 2005. Products subject to counterfeiting range from luxury handbags, sneakers and liquor to plane spare parts and pharmaceuticals. A counterfeit is any product that infringes intellectual property rights such as trademarks, patents and designs. This includes items produced in factory overruns - by the same manufacturers that produce genuine products during normal working hours - although they do not differ in appearance from genuine products.

Each counterfeit that is sold to a customer affects the rights holder negatively. Potential consequences range from reducing the sales of genuine products and affecting brand value to disappointed or even physically harmed customers (OECD 2007, Staake et al. 2008). Rights holders attempt to minimize the negative consequences by implementing anti-counterfeiting measures that aim at cutting off counterfeit supply and increasing the number of detected counterfeits. Product authentication approaches, i.e., methods that aim at deciding whether a product is counterfeited or genuine, enable stakeholders such as customers, customs and supply chain partners to check products and contribute thus directly to an increased counterfeit detection rate. Moreover, the higher the detection rate, the less profitable is the business for the counterfeiter as revenues decrease with each detected counterfeit, while counterfeit production costs persist. For each counterfeiter organization there is a detection rate that reduces the profits to zero and drives the counterfeiter out of business. Additionally, as profits decline, counterfeiters might target products that are less well protected and thus promise larger profits. This contributes to decreasing the supply of counterfeits.

Traditional authentication approaches, e.g., the check of holograms, base authentication on something an item *has*. Authentication can also be based on something an object *did* in the past (Stajano 2002). With the advent of item serialization – facilitated by radio-frequency identification (RFID) technology - and tracking infrastructures, history information about individual items is available. Serialization means that each item is assigned a unique identifier (UID) that can be stored on an RFID tag attached to the item. Suitable readers are deployed in the supply chain. When an item passes, an event is created that contains UID, time, location and further context information. Tracking infrastructures then enable the storing and sharing of events between companies. To illustrate how this information can be used for authentication, assume that product #123 was seen in Paris at 12:00 h, followed by a sighting in Tokyo at 12:05 h. As item #123 cannot physically have moved over this distance in such a short period of time, it can be concluded that the item is carrying a duplicated serial number and is therefore probably a counterfeit.

While the potential of tracking for counterfeit detection has been recognized (Koh et al. 2003; Staake et al. 2005; Inaba 2008), few concrete methods have been suggested so far. The general difficulty is to specify rules operating on the item history that accurately distinguish between counterfeit and genuine products, while keeping the specification effort manageable. The research question is thus how to design a counterfeit detection approach based on tracking data that addresses these requirements. The approach proposed in this paper learns rules governing the behaviour of genuine items from available tracking data, thus eliminating the need for manual specification. The warning about potential counterfeits is based on anomaly detection. The research thus contributes to the general aim of developing methods and systems for analysing tracking data to support business decisions (Asif & Mandviwalla 2005; Curtin et al. 2007).

The research method pursued is reflected in the structure of the paper: In the next chapter, related work is reviewed and shortcomings are identified. Chapter 3 gives an overview of item tracking and the proposed authentication approach while chapter 4 presents the methods for learning and counterfeit detection in detail. To evaluate the approach, a model of counterfeiter behaviour in a

tracking enabled world is developed and utilized to assess the detection capabilities of the approach, before concluding and suggesting areas for future work.

2 RELATED WORK

A large share of authentication approaches in use today are based on adding a hard to imitate, artificial feature to the product, e.g. holograms, printing in colour-changing ink or chemical markers (OECD 2007). Authentication with these approaches requires special equipment, expertise or can only be performed one product at a time. As the counterfeit detection rate depends not only on the detection capability of an approach, but also on the check rate, i.e. the fraction of products that are actually checked, approaches that do not allow for bulk authentication in the supply chain will usually result in lower detection rates.

With the advent of RFID technology, cryptographic authentication approaches that make use of the computational capabilities of RFID tags have been proposed (Juels 2006). Cryptographic authentication cannot be performed with low-cost RFID tags as they have only very limited processing power. Once a method has been devised to perfectly imitate a security feature – for instance by hacking cryptographic RFID tags - the approach is no longer secure and counterfeits become undetectable.

Some proposed authentication approaches utilize item serialization, i.e. the fact that each genuine item was assigned a UID during manufacturing. It was proposed to check the validity of the serial number of products against a database of all serial numbers assigned during manufacturing (Lei et al. 2005; Johnston 2005). In the context of tracking infrastructures, this approach was suggested to be used for pharmaceuticals (Koh et al. 2003). Drawbacks of this approach include the low protection against UID copies and the possibility of counterfeiters to skim the database for valid numbers.

When more tracking data is available, the full product history can be made accessible and checked by the prospective buyer for plausibility (Staake et al. 2005). Some types of plausibility rules are suggested by Illic et al. (2009). For specifying individual rules, background knowledge and manual input is required, e.g. the maximum allowed item stay time per location must be configured. The drawback of these approaches is the effort imposed on the user for rule specification and decision.

Few approaches have been proposed that learn rules from available data. It has been suggested to learn which product transitions between any two locations are plausible (Illic et al. 2009, Lehtonen et al. 2007). By considering only two locations, these rules are not precise enough to detect some counterfeit types. The approaches use a probabilistic decision function that classifies items as counterfeit if the observed movements are infrequent. This poses a problem if genuine items perform unusual movements. The Hidden-Markov-Model based approach by Lehtonen et al. (2007) detects counterfeits based on similarity with previously detected counterfeits, but does not adapt well to new counterfeiter strategies.

The counterfeit detection approach proposed in here extends on previous work by learning rules from tracking data, but adds precision by considering the full context information contained in events, event timing, and the full sequence of events instead of two consecutive events. The decision will be based on the discovery of anomalies in event sequences. The concept of anomaly detection has been applied in fraud and computer intrusion detection. In both domains, anomalies caused by fraudster activity are sought among sequences of events, e.g. credit card transactions, phone calls, or system calls (Bolton & Hand 2002). Fraud or intrusions are detected by identifying deviations from previously learned profiles (Forrest 1997; Hilas & Sahalos 2005). A main difference to counterfeit detection is that these systems operate on continuous event streams, while a product has a lifecycle and the corresponding event sequence has a beginning, end, and a regular structure determined by the underlying business processes.

3 OVERVIEW OF TRACKING-BASED COUNTERFEIT DETECTION

3.1 Serialization and Tracking

Serialization and tracking infrastructures are necessary prerequisites for tracking-based counterfeit detection. The UID may be stored on an RFID tag attached to a product, facilitating the simultaneous identification of multiple items without line of sight. However, tracking is not limited to RFID tagged products, and in principle any technology able to provide UIDs can be employed, e.g., printed 2D-barcodes. To enable tracking, suitable reading devices need to be distributed in the supply chain that capture events when items pass. The attributes to be contained in an event are usually determined by the tracking infrastructure. For example, the attributes *(UID, time, location, business step)* specify a simple event format. An event instance contains values in the domain of each attribute, e.g., *(123, 12:00 h, Paris Shop A, Sale)*. In a business setting, further useful attributes include the organizational unit, associated business transactions, lifecycle information, and aggregation with other products. The captured events are submitted to a local database – the repository - for future reference. A tracking infrastructure provides methods for finding and assembling the events pertaining to a specific item, which are usually distributed over multiple repositories and even multiple companies.

The trace of an item is defined as the sequence of all events containing the item UID ordered by their timestamp. The trace may not only contain information about the physical movements from location to location, but also about the underlying business processes and the actions performed. This is illustrated by the traces of two sample products in table 1. Both items are produced in Lyon and then delivered to a distribution centre in Paris. While item #123 does not pass the quality control and is therefore disposed of, item #126 is shipped to a retailer and subsequently sold to a customer.

UID	Event	Time	Location	Step
123	1	12.08.2008;12:00 h	Lyon Plant A	Production
123	2	22.08.2008;12:00 h	Paris Distribution Centre A	Receiving
123	3	22.08.2008;14:00 h	Paris Distribution Centre A	Quality Control
123	4	22.08.2008;15:00 h	Paris Distribution Centre A	Disposal
126	1	23.08.2008;11:00 h	Lyon Plant A	Production
126	2	25.08.2008;12:00 h	Paris Distribution Centre A	Receiving
126	3	26.08.2008;12:00 h	Paris Distribution Centre A	Shipping
126	4	26.08.2008;14:00 h	Paris Shop B	Receiving
126	5	27.08.2008; 11:00 h	Paris Ship B	Sale

Table 1. Sample traces of two items represented in an event format that contains the two context attributes location (represented by city and facility) and process step.

Serialization and tracking are not yet in widespread use, although trials of limited scope are performed. However, some product types, e.g., luxury watches or plane parts, already carry serial numbers and proprietary systems are used to track them over their lifecycle. Standardization of tracking infrastructures is under way, most notably in the context of the EPCglobal framework, a set of standards targeted at enabling an interorganisational, distributed tracking infrastructure based on globally unique identifiers (EPCglobal 2007b). Part of the framework is also a specification of an event format suitable for a business context (EPCglobal 2007a). Government mandates requiring serialization and tracking for high risk products such as pharmaceuticals might further foster the adoption of tracking based systems (e.g. FDA 2004).

The event format presupposed for the proposed counterfeit detection approach is generic in so far as only a UID, timestamp and at least one recurring attribute is required. The approach is thus independent from a specific standard such as the EPCglobal framework and can be adapted to the event format utilized by any tracking system that fulfils these basic requirements.

3.2 Components of the Authentication Approach

An overview of the components of the proposed anti-counterfeiting approach and the data flows between them is given in figure 1. The model learner takes as input a set of sample traces from the tracking repository and infers from them the rules according to which items move. The output is a model of the supply chain containing all rules. The learning step can be repeated periodically to update the model with current traces. For every product type, a separate model should be constructed, so that the trace of a handbag is not compared to traces of car brakes. The Electronic Product Code (EPC) standard contains the product type identifier as prefix in the UID, followed by the actual serial number of the item (EPCglobal 2007b). For all products identified by EPCs and UIDs constructed in a similar manner, the product type is directly available from the events.

The quality of the learned model depends on the quality of the traces contained in the sample. The sample must fulfil the following requirements:

- The traces should cover all potential product movements, including legitimate exceptions in product routing or process execution, as well as seasonal variations.
- All traces must belong to genuine items, otherwise the movements exhibited by counterfeit products will later be accepted as normal

To fulfil the first requirement, the sample size must be chosen large enough and cover a sufficiently long period of time. Furthermore, commonly used data cleansing techniques aimed at filtering out outliers and exceptions based on low observation frequencies should not be applied. The problem of separating events triggered by legitimate users from events generated by intruders is well known from the intrusion detection domain. But unlike intrusion detection, which operates in the virtual space, counterfeit detection deals with actual physical objects. This allows for authenticating items with non-tracking based approaches, e.g., relying on security features or expert judgement. It is also possible to send designated test items from the manufacturer through the supply chain, authenticating them at each read point, and basing the sample on their traces. Only traces of items confirmed to be genuine should be admitted to the sample. Another possibility to ensure that only traces of genuine items are contained is to visualize the model with a suitable representation (e.g. Illic et al. 2009, Hoffman et al. 2008) and have it inspected by an expert.

The *decider* takes the trace of the product to be authenticated, the learned model and possibly also the position at which the authentication is performed as input. Position information is required because items might also be authenticated at intermediate points in the supply chain. In fact, every capturing of an event may be coupled with authentication, thus leading to a high check rate. This allows to intercept counterfeits early, before they reach the customer. The position information can be supplied by the user, the reading device or by an authentication service. The decider compares the trace reported by the item to be authenticated with the supply chain model (up to the authentication position) and computes the authentication result. If a mismatch is found, the item is a potential counterfeit and a warning is issued. If no anomaly is detected, the trace is reported to be plausible. These cautious results are justified because even for plausible traces, it cannot be guaranteed that the item is genuine. Likewise, deviating traces might be caused by errors, e.g., wrong item routing and need not necessarily be caused by counterfeiting activity. The detailed algorithms used by the decider and the model learner are presented in the next chapter.

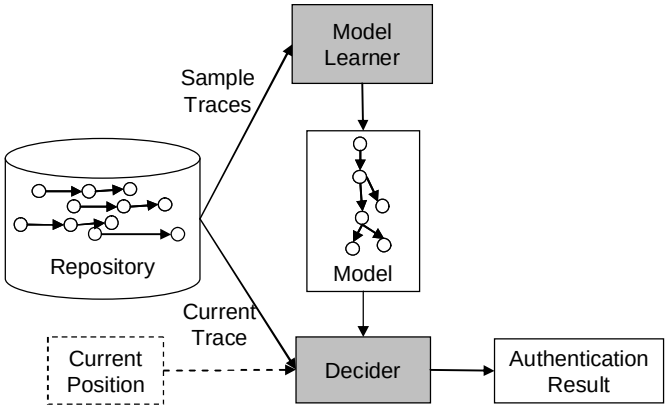


Figure 1. Overview of the components and data flow of the proposed authentication approach.

4 AUTHENTICATION APPROACH

4.1 Representing Item Movements

The supply chain model shall in a general yet precise way describe the movements of genuine items. The movements as represented in tracking data are characterized by

- the events contained in the trace,
- the sequence of events, and
- the timing of events in the trace relative to each other.

Before describing how the model is constructed from sample traces, it will be described how each of these three concepts is represented.

The attributes specified by an event format can be divided into recurring and non-recurring attributes. The values of recurring attributes can be expected to be observed for further items, while the values of non-recurring attributes have a limited lifespan. The UID – bound by the lifetime of an item - and the event time – being monotonically increasing – are non-recurring. In the above example (table 1), the attributes location and process step are recurring. As their values are determined by the underlying business processes and the design of the supply chain, the values can be expected to be observed also for future items. The set of recurring attributes can be determined by the user (as this is necessary only once per event format) or automatically by scanning the sample traces for attribute values recurring over the sample period. A unique combination of values of recurring attributes will be called a *station*. Examples of stations are *(Lyon Plant A, Production)* or *(Paris Distribution Centre A, Receiving)*. Each event maps to exactly one station, while a station represents many different events.

To accurately represent an event, its position in the trace and prior events must be considered in addition to the attribute values contained. Take for example a process in which items are shipped from a central distribution centre to shops, and, in case they are not sold, sent back to the distribution centre and shipped to outlet centres. Some items will pass the shipping ramp - the station *(Paris Distribution Centre A, Shipping)*- twice, but depending on the position in the trace, the legitimate next event will vary, being either the reception at a shop or at an outlet centre. Therefore, the concept of a *node* is introduced to represent an event at a specific position in the trace. Each node references a station that indicates which attribute values are contained in the event represented by the node. Multiple nodes (representing events at different positions in the trace) may reference the same station.

The ordering of events in a trace is represented by *transitions* between nodes. A transition is a directed arc from a start node to a successor node. The timing of events relative to each other is given by the difference in the timestamps of two consecutive events in a trace. The time is therefore a property of

the transition between event representations. When considering multiple items that share the same transition in their traces, the average, minimum and maximum *transition times* are computed for the transition.

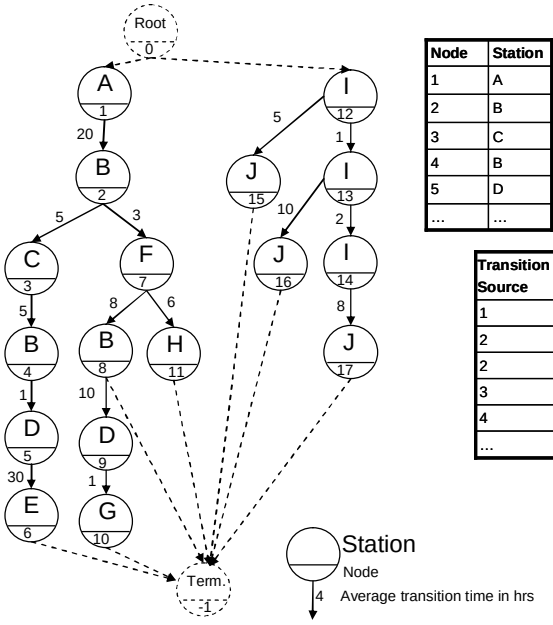
4.2 Constructing the supply chain model

Given stations, nodes, transitions and transition times as building blocks, a model of the legitimate behaviour of items in the supply chain can be constructed from a set of traces of genuine items. Put shortly, the model is built up by transforming each trace into a sequence of nodes connected by transitions and then merging sequences sharing the same prefix together. The resulting model is a tree shaped directed acyclic graph containing nodes and transitions, and an additional look-up table mapping the nodes to stations. The model represents all prior observed sequences of movements of genuine items.

The graph is initialized with a root node and a terminal node. The root node is a common predecessor of all nodes representing potential first events in traces. The root node connects for example the traces of items produced in Lyon and items produced in Tokyo. The terminal node serves as a common end node for all traces. It is needed to mark the potential last events so that the completeness of traces can be checked.

Each trace in the input set is processed in the following way: If the station corresponding to the current event is not yet contained in the model, the station is created and added. Starting with the root node of the graph as the current node, it is tested whether the station corresponding to the current event is referenced by one of the successor nodes. Successor nodes are all nodes to which a transition exists. If such a node is found, the average, minimum and maximum time of the transition between the current and the successor node are updated with the difference between the timestamps of the last and current event from the trace. If no suitable successor node is found, a new successor node referencing the station corresponding to the current event is created. A transition between the current and the successor node is added and the transition times are initialized. As long as there are more events in the trace, the successor node then becomes the new current node and the next event from the trace is processed. If the last event of a trace is reached, a transition to the terminal node is added to indicate that the trace ends at this point.

Figure 2 presents a sample supply chain model learned by this algorithm. Stations are labelled with characters, while nodes are labelled with consecutive integers. The model has the following properties: All direct successors of the root node represent valid start events of traces, the predecessors of the terminal node represent valid last events. The transition from node 8 to the terminal node indicates that (A, B, F, B) is a valid trace. Although the node pairs 4;5 and 8;9 reference the same stations, the preceding nodes 3 respectively 7 influence not only the transition times between stations B and D, but also the set of valid successor events (represented by E respectively G). It is important to note that events may not only have implications for the directly following event, but also on events that happen at any later point in time. These higher level dependencies (spanning multiple events) are commonly found in supply chains. For example, an unpacking event is only valid if items have been packed earlier. Nodes 12-14 showcase a potential repetition of events. By representing events as nodes it is possible to determine how many repetitions are allowed (maximum 3 in the example) and how transition times may change with the number of repetitions.



Node	Station
1	A
2	B
3	C
4	B
5	D
...	...

Station	Location	Step
A	Lyon Plant A	Production
B	Paris DC A	Shipping
C	Paris Shop A	Receiving
D	London Outlet A	Receiving
E	London Outlet A	Sale
...	...	...

Transition Source	Transition Target	Min. Time	Max. Time	Average Time
1	2	5	30	20
2	3	1	20	5
2	7	2	10	3
3	4	2	7	5
4	5	0,5	5	1
...	...	...	...	...

Figure 2. Graphical and partial data representation of a supply chain model.

4.3 Item Authentication

The decider classifies a trace as plausible if its trace conforms to the supply chain model. If this is the case, a similar trace has been observed for a genuine item, rendering the item in question unsuspicious. For a trace to conform to the model, the following conditions must be fulfilled:

- All events in the trace must be valid, i.e., for each event the corresponding station must be contained in the model
- The events in the trace must form a valid sequence, i.e., there exists a path in the graph on which all stations corresponding to events in the trace are connected by transitions
- The trace must be complete, i.e., the path starts at the root node and ends at the terminal node
- The timing of events in the trace is plausible, i.e. the time difference between any two events is within reasonable boundaries of the time span observed for the corresponding transition

If any of these conditions is violated, the item is classified as a potential counterfeit. The four types of mismatches are illustrated in figure 3. The completeness condition needs to be relaxed if an item is authenticated at an intermediate point in the supply chain. For example, if all traces represented by the model end at a shop, but an item is authenticated at a distribution centre, it will unjustifiably be classified as counterfeit. Therefore, if no information about the authentication context is available, the test for a trace ending with the terminal node may be skipped. In case some information about the authentication context is available, transitions from all stations matching with this context to the terminal node should be added temporarily to the supply chain model. The context information may correspond directly to a station, e.g., authentication is performed at the reader located at *(Paris Distribution Centre A, Shipping)* or only to some attribute values contained in a station, e.g., *(?, Sale)*.

The decider processes the current trace as follows: For each event, the corresponding station is computed and searched in the look-up table. If the station is not found, the item is classified as potential counterfeit. Starting with the root node in the model, a node referencing the current station is searched among the successor nodes. If no such node exists, the item is again a potential counterfeit. If a valid successor node is found, the difference of event times of the last and current event is compared to the corresponding transition time. We propose to check whether the event time difference is smaller than the minimum or greater than the maximum transition time, each extended by a small margin to account for minor variations in event timing. The extension may, e.g., be set to 10% of the difference

between average time and minimum time for the lower bound and to 10% of the difference between maximum and average time for the upper bound. If all tests are passed, the successor node becomes the new current node and the next event is processed. If the last event in the trace is reached, it is additionally tested whether a transition from the current node to the terminal node exists.

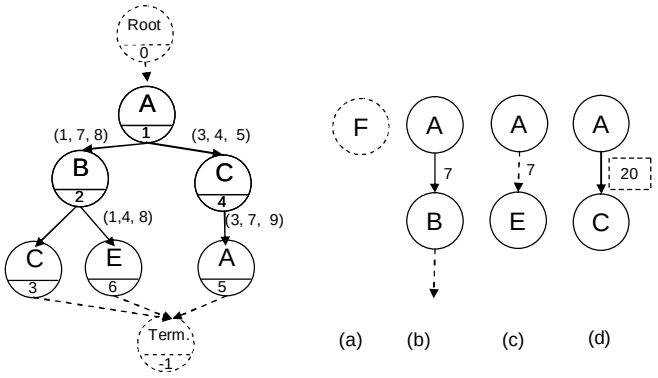


Figure 3. Mismatch between supply chain model and trace: a) invalid event, b) incomplete trace, c) invalid transition, d) invalid transition time

5 EVALUATION

The information about counterfeiter strategies and the routes on which counterfeited items are put on the market is limited due to the clandestine nature of the operations. To assess the detection capabilities of the proposed approach, it is moreover necessary to systematically explore the possibilities of counterfeiters in tracking enabled supply chains, for which the available data is not suitable. A model of the potential counterfeiter activities is therefore suggested, the effect of the behaviour on item traces is analyzed and used to evaluate the detection capabilities of the proposed approach. In future work, it is planned to computationally verify the results of the analysis by simulating the modelled counterfeiter strategies and normal supply chain operations and assessing the detection capability.

Suppose all items of a certain type are equipped with UIDs and a tracking infrastructure covering the supply chain is in place. The first challenge for the counterfeiter is to obtain UIDs for the counterfeited goods. One option is to completely omit the UID, but this strategy should be easily detectable during authentication. For actually obtaining a UID, the options include guessing identifiers, transferring the UIDs of genuine products to counterfeits or copying the UIDs found on genuine products. If identifiers are guessed randomly, the resulting UID may in rare cases be a duplicate of a valid UID, but most likely it will be invalid, i.e., no genuine item carrying the same UID exists. To transfer UIDs, counterfeiters may remove RFID tags from genuine items and reapply them to counterfeits, or they may seek access to UIDs of disposed products and reuse them. Any UID found on a counterfeit will have at least one of these properties: it is invalid, has been transferred or is duplicated.

The counterfeited items must be distributed, either by injecting them in licit channels or by distributing them outside of the legitimate supply chain, for example by smuggling, selling on flea markets or through online shops. For the resulting item trace, it is important whether there are read points on the path of the item or not. Illicit distribution channels are likely to contain no read points and thus be invisible, while licit channels can be assumed to be visible. The combined effects of UID and distribution strategies on traces will now be analyzed. The various scenarios are illustrated in figure 4.

An item with an *invalid UID* that is traded completely through invisible channels has no trace. This can be detected immediately. Items with invalid UIDs that are injected in the licit supply chains will appear suddenly after the injection point (figure 4a). Some events at the beginning of their trace are missing and the trace will fail the completeness check as the first event in the trace is not a successor

of the root node in the model. However, if a counterfeiter manages to inject counterfeits in the licit channel before the first event is usually captured – which is most likely in the production facilities of genuine items – and it is ensured that this manipulation does not distort the further routing of the item, the resulting trace will fully correspond to the model and the counterfeit will not be detected. If distortions surface later in the trace (not only in terms of location but also in terms of wrong business steps or missing transaction information) the counterfeit will also be detected (figure 4g). An additional check for UID validity offered by the rights holder and strict number management can mitigate this threat

In case a counterfeit carries a *transferred UID* (valid and unique), the events in the trace were triggered by the movements of two items: Up to the point in time when the UID was removed, the events were created for a genuine product and the beginning of the trace will conform to the model. After that, the counterfeit carries the stolen UID. Depending on the distribution of the counterfeit, the following possibilities for the continuation of the trace exist: If the counterfeit is distributed outside visible channels, the trace ends early (figure 4b). Unless the last captured event was a valid terminal event, the completeness check will fail. If the counterfeit is injected in the licit supply chain, the sequence of events will only be valid if the counterfeit directly replaces a genuine item, the time needed for the product exchange does not lead to a time threshold violation and the manipulation does not lead to distortions on the future path of the counterfeit. If the injection is not a direct replacement but takes place further upstream (figure 4c), downstream (figure 4d) or in another branch of the supply chain (figure 4e), the model will contain no transition allowing this sequence of events. If the replacement leads to an exceeded transition time, invalid transitions or events later in the trace (figure 4g), it will also be detected.

If counterfeits with *duplicate UIDs* are distributed through invisible channels, they can be detected as long as the trace of the genuine item carrying the UID is not complete (figure 4a). If the duplicates are injected in the licit supply chain, the trace that is retrieved for the UID is a mix of all sub-traces created by the multiple items carrying the same UID (figure 4f). As soon as the first item with a copied UID is injected in the supply chain, this will result in an invalid trace as a transition between the last event captured for the genuine item (from which the UID was copied) and the first event triggered by the counterfeit is missing. However, if the counterfeiter manages to inject the counterfeit at exactly the station where the next event for the genuine item is expected, the injection will remain undetected, but only as long as the genuine item does not arrive at this station. As soon as the genuine item triggers the next event, the trace becomes invalid as a transition is missing and the counterfeit can be detected at the next sighting. Note that in this specific case, a genuine item will be classified as counterfeit. As counterfeiters will probably put the same copied UID on many items, it may be acceptable to misclassify one genuine item if this allows detecting mass copied items.

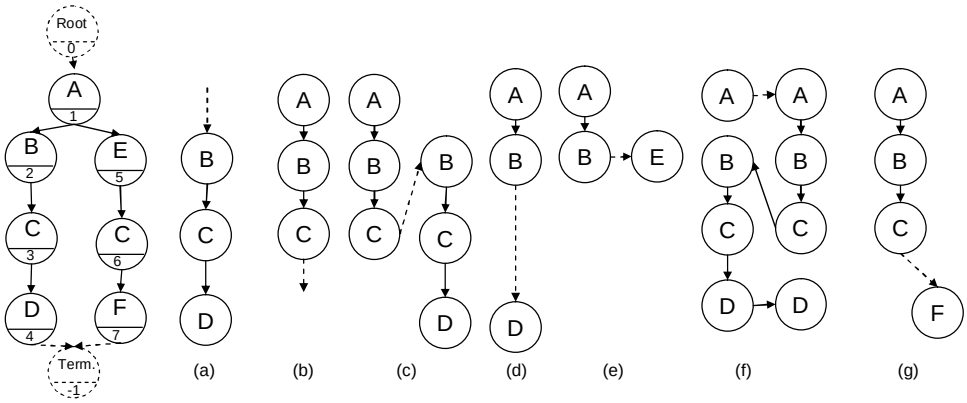


Figure 4. Consequences of counterfeiter strategies on trace data

From this analysis follows that a wide range of potential counterfeiter activities result in traces that deviate from the traces of genuine items and can thus be detected by the proposed approach. Only few options are open to counterfeiters to “construct” traces for counterfeits that arte indistinguishable from the traces of originals. Moreover, each of these remaining options requires considerable effort in terms of knowledge and access to the licit supply chain.

Tracking based authentication enables a high check rate compared to traditional authentication, as automated event capturing can be combined with authentication and, in case of RFID, this can be done for many items simultaneously. The possibility to check items by their UID might also be extended to end customers, further increasing the check rate. The detection capability of the approach for low effort strategies such as identifier guessing without access to the licit supply chain is given, as well as the capability to detect the strategy most prevalent today - the distribution through illicit channels. Additional safeguards such as random inspection of inconspicuous goods, investigation of the sources of detected counterfeits facilitated by tracking data and UID validation may help in limiting the number of counterfeits that cannot be spotted with the proposed approach. In combination with a high check rate, this may result in detection rates that are high enough to drive counterfeiters out of business or make them target more vulnerable products. In any case, tracking based counterfeit detection will raise the bar for counterfeiters and each detected counterfeit decreases the overall negative consequences of counterfeiting.

6 CONCLUSION AND OUTLOOK

In this paper, a counterfeit detection approach based on item serialization and tracking was proposed. Rules governing the behaviour of genuine items are learned from available tracking data. The traces of items to be authenticated are compared to these rules and a counterfeit warning is issued if no match with a rule is found and the trace is thus distorted. By modelling the potential behaviour of counterfeiters in a tracking enabled environment and analyzing the consequences of these actions on item traces, it could be shown that the approach is able to detect a large share of the modelled strategies. It was also analyzed which strategies are undetectable by the suggested approach as they result in traces that are indistinguishable from the traces of originals. These strategies have in common that they require prerequisites from the counterfeiter such as access to the licit supply chain and knowledge about the current location of items. It was argued that the combination of the detection capability and the attainable check rate may lead to favourable counterfeit detection rates, particularly in comparison with traditional authentication approaches.

As some organization are already tracking items or are preparing to do so in the future, they may wish to consider the suggested approach to discourage counterfeiting of their products. Organizations that are affected severely by counterfeiting might assess the suitability of the approach for their specific context. Of particular importance is the trust that companies have in their supply chain partners, their expected willingness to participate in such an approach, as well as the current and expected future strategies of counterfeiters.

It was demonstrated that counterfeit detection is a problem that is in principle amenable to information technology support, but before tracking based counterfeit detection can be applied in practice, many questions remain to be answered. A major concern that could not be addressed in this paper is data sharing in the supply chain. Organizations are generally reluctant to reveal fine-grained and potentially sensitive business data as contained in tracking events. Data access is a prerequisite for model learning and authentication, so a challenge is to enable these functionalities with minimal or no divulgence of critical information. To address this issue, concepts from the area of privacy preserving data mining and secure multi party computation may be assessed. Furthermore, mechanisms for avoiding high false positive rates are required, as a large share of genuine items for which a counterfeit warning is issued might hamper the practical applicability of the approach. Given current rates of read errors in RFID based systems, issuing a warning for any deviation might be a too harsh decision criteria in real world operations and the decision logic may be further refined to account for this requirement.

Operator models need to developed, e.g., a service hosted by trusted third party, the tracking infrastructure provider, or a distributed model learning by every supply chain participant. The enforceability, cost and benefit distribution in various supply chain settings might also be explored. How the approach can be extended to end customers, the security implications of this step and user acceptance of a mobile or web based authentication system are further areas of interest.

The proposed tracking based authentication approach might prove to be a useful complement to traditional authentication approaches. Given item serialization and tracking are in place as prerequisites, the proposed approach may help to increase security in the supply chain and to lessen the negative consequences on health and finances caused by counterfeiting.

References

Asif, Z. and Mandviwalla, M. (2005). Integrating the Supply Chain with RFID: A Technical and Business Analysis. Communications of the Association for Information Systems, 15, 393 – 427.

Bolton, R. and Hand, D. (2002). Statistical fraud detection: A review. Statistical Science, 17(3), 235–255.

Curtin, J., Kauffman, R. and Riggins, F. (2007). Making the 'MOST' out of RFID technology: a research agenda for the study of the adoption, usage and impact of RFID, Information Management and Technology, 2007, 8(2), 87 – 110.

EPCglobal (2007a). EPCglobal Information Services (EPCIS) Standard Version 1.0.1. http://www.epcglobalinc.org/standards/epcis/epcis_1_0_1-standard-20070921.pdf

EPCglobal (2007b). EPCglobal Architecture Framework Version 1.2. http://www.epcglobalinc.org/standards/architecture/architecture_1_2-framework-20070910.pdf

FDA (2004). Combating Counterfeit drugs: A report of the food and drug administration, http://www.fda.gov/oc/initiatives/counterfeit/report02_04.pdf

Forrest, S., Hofmeyr, S. and Somayaji, A. (1997). Computer immunology. Communications of the ACM, 40(10), 88–96.

Johnston, R. (2005). An anti-counterfeiting strategy using numeric tokens. International Journal of Pharmaceutical Medicine, 19(3), 163–171.

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

Hilas, C. and Sahalos, J. (2005). User profiling for fraud detection in telecommunication networks. In 5th International Conference on Technology and Automation ICTA'05.

Hoffmann, F., Oertel, N. and Vogt, H. (2008). Visualization of Supply Chain Events for Anti-Counterfeiting. In 4th European Workshop on RFID Systems and Technologies, 2008.

Illic, A., Andersen, T. and Michahelles, F. (2009). Increased Supply Chain Visibility through Rule-Based Analysis of RFID Data. IEEE Internet Computing, 13(1).

Inaba, T. (2008). EPC System for a safe and secure supply chain and how it is applied In Cole, P. and Ranasinghe, D. (Eds.) Networked RFID Systems and Lightweight Cryptography: Raising Barriers to Product Counterfeiting, Springer, Berlin, p. 191-210.

Koh, R., Schuster, E., Chackrabarty, I. and Bellman, A. (2003). Securing the pharmaceutical supply chain. Auto-ID Labs Technical report.

Lei, P., Claret-Tournier, F., Chatwin, C. and Young, R. (2005). A secure mobile track and trace system for anti-counterfeiting. In Proceedings of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service, p. 686–689.

OECD (2007). The economic impact of counterfeiting and piracy. Report. http://www.oecd.org/dataoecd/11/11/2090589.pdf

Staake, T. and Fleisch, E. (2008). Countering Counterfeit Trade. Springer, Berlin.

Staake, T., Thiesse, F., and Fleisch, E. (2005). Extending the EPC network: The potential of RFID in anti-counterfeiting. In Proceedings of the 2005 ACM symposium on applied computing, p. 1607–1612, New York, NY, USA.

Stajano, F. (2002). Security for Ubiquitous Computing. Wiley & Sons.