Deckard: A System to Detect Change of RFID Tag Ownership

Luke Mirowski and Jacky Hartnett

University of Tasmania, Hobart, Australia

Summary
Change of tag ownership compromises the security goals of Radio Frequency Identification (RFID). When an attacker clones or steals an authorized subject's tag, they are willingly granted access as RFID assumes the owner of a tag is always the authorized entity. We present Deckard, a new approach to preventing change of tag ownership. Deckard uses the principles of intrusion detection to look for anomalous behavior which may indicate a change of tag ownership has occurred. We have evaluated its performance in detecting synthesized attacks inside a sanitized RFID proximity tag audit log. The results suggest that intrusion detection systems can be used in RFID, although the weaknesses of statistical anomaly detection are also apparent when used on RFID data. We conclude with a call to further research of intrusion detection in RFID systems.

Key words:
RFID, Intrusion Detection, Tag Cloning

1. Introduction

Radio Frequency Identification (RFID) establishes the identity of subjects in the physical world using uniquely numbered electronic tags [1]. When a tag is attached to a subject such as a human, animal or a supply chain product, that subject is identified by its ownership of the tag. The concept is analogous to the way a vehicle license plate identifies a vehicle. However, as RFID assumes that the subject in possession of a tag is the authorized entity; a change of tag ownership, for example, when an attacker clones or steals a tag, can result in the attacker obtaining access to the assets which RFID is being used to protect. Consequently, a change of tag ownership compromises the security goals of an RFID system as attackers cannot be differentiated from authorized subjects.

Change of tag ownership is a serious threat to the security of applications which use RFID. The pharmaceutical industry, for example, has proposed that RFID will be used to track drugs through pharmaceutical supply chains. By attaching RFID tags to drug packaging, they hope that RFID may curtail the USD $40 billion per year counterfeit drug market [2]. Change of tag ownership here may allow counterfeiters to bypass these security checks to introduce inferior counterfeit drugs. Moreover, RFID tags are being used in several countries' electronic passport (ePassport) schemes as a way of preventing passport forgery. Change of tag ownership here may allow terrorists or illegal immigrants to enter country borders undetected. Change of tag ownership is a serious security threat as it compromises the security goals of the RFID applications that use it.

The security necessary to prevent change of tag ownership from occurring is difficult to implement because of the RFID industry's desire to limit tag hardware functionality in order to produce tags that cost around five cents [3, 4]. The trade-off has meant several things. Firstly, on-board the tag; power, storage, processing and gate resources available to low cost tags make it difficult to allow cryptography to be used effectively [5] to prevent tag's from being cloned. Secondly, off-board the tag; passive tags are limited by the amount of power they can obtain from RFID readers. As the transaction time between tags and readers is limited to less than 400 milliseconds in the United States, in order to supply cryptographic components with sufficient power, tags would need to be read from a shorter distance, which would degrade the read-rate of RFID readers [3]. Consequently change of tag ownership is difficult to prevent at the tag layer.

It is therefore not surprising that when the tag layer is the focus of research, as outlined in [6] and [7], that these proposals fail. We believe these solutions fall short of being useful for several reasons. Firstly, the security they propose is situated on the tag. As tags are the weakest link in the chain due to their functional capabilities, an attacker with modest resources can break their security quite easily. Secondly, previous research has primarily focussed on preventing tag cloning from occurring. We believe that the real issue here is whether the subject a tag is attached to is actually the authorized entity, and so preventing an individual tag from being cloned is not a solution. Our proposal improves upon previous research, as it goes beyond static defensive countermeasures, by detecting when change of tag ownership occurs.

We have developed Deckard, a system that uses the principles of intrusion detection to detect change of tag ownership. An intrusion detection system, much like a burglar alarm, monitors the activity occurring within an environment, and responds when it detects suspicious activity. We assume that a subject exhibits their behaviour through their ownership of an RFID tag. We use a tag's

Manuscript received July 5, 2007
Manuscript revised July 25, 2007

audit record's to build a profile of normal behaviour, which can then be used to determine when subject behaviour significantly deviates. We assume that a significant deviation away from normal behaviour is indicative of a change of tag ownership. This system would be useful in detecting when an attacker starts using a cloned or stolen tag. To illustrate, when Mallory steals or clones Alice's tag, his tag usage will be different when compared to Alice's tag usage. Thus, a change in tag ownership is assumed to be visible by a significant deviation in a tag's audit records. Deckard is based on the intrusion detection system proposed by Denning [8] and partly inspired by the character in the 1982 Ridley Scott movie Blade Runner [9]. Our system aims to detect change of tag ownership which to defend against stolen or cloned tags from accessing an RFID system.

The organisation of this paper is as follows. Section 2 will provide the motivation behind our research by briefly providing background information to RFID technology and the problem of change of tag ownership, with particular attention paid to the threat of tag cloning. Section 91 will discuss how our work is different to previous research. Section 92 will present the design and operation of Deckard. Section 93 will discuss our testing methodology. Section 6 will present our results and analysis. In section 7 we summarize our findings. Finally, in section 8 we discuss future work that can be undertaken.

2. Background

Just like the barcode, RFID is regarded as an Automatic Identification and Collection (AIDC) technology [10]. The main difference is that it establishes the identity of subjects in the physical world without requiring line-of-sight, using uniquely identifiable electronic tags. A tag typically contains a unique numeric identifier that can be read from a remote distance via an RFID reader device. Passive tags are powered by the reader device, whereas, active tags have their own on-board power supply. When a reader is in range of a tag, a tag responds with its unique numeric identifier [1]. As RFID assumes that the subject to which a tag was initially attached to has not changed, when a tag is located, the original subject is also located in the physical world. This association between the physical subject and the tag is managed by a middleware database. In essence, this is the concept behind typical RFID systems.

The concept of RFID has existed for a long time. The notion of identifying subjects via unique electronic transponders can be traced back to the 1940's when the British air force developed the Identification Friend or Foe (IFF) system. IFF was a means of identifying British aircraft from enemy aircraft. Nowadays, the term RFID incorporates a number of non-contact integrated circuit technologies for identifying any thing from humans to animals via small electronic tags. These tags operate on the following radio frequencies: < 135 KHz, 13.56 MHz, 862-915 UHF, 2.45 GHz, and 5.8Ghz [11]. Tags which operate on these frequencies include: EPC, VeriChip, proximity tags, payment tokens like SpeedPass, and pet identification chips. The EPC tag is the tag of choice for identifying supply chain products. Wal-Mart uses this type of tag to track and trace products through its supply chain. RFID improves on the efficiency of barcodes by not requiring line of sight contact when automating the identification and collection of data.

Despite the benefits of RFID, it is essentially an insecure technology. Just like the barcode or vehicle license plate, tags can easily be removed from their subjects. To illustrate, in an RFID-enabled liquor store, Mallory can replace an expensive wine bottle's tag with a tag from a cheap bottle of wine. The reader located at the cash register will see the tag belonging to the cheap bottle of wine, which will be the wine that Mallory will end up paying for. This is an example of change of tag ownership via tag stealing. Although quite easy to perform, this form of change of tag ownership has poor scalability as attackers need access to the original tags during the attack phase.

In contrast, tag cloning, when an attacker makes an exact copy of an RFID tag, is a more serious threat to security. As the identifier data on the majority of tags is not kept secret [3], an attacker can simply obtain a tags unique identifier to make their own original tags that are indistinguishable from the originals. Once legitimate tag data has been obtained, attackers can reproduce their cloned tags on a wide scale. Tag cloning is the most widely reported and most serious form change of tag ownership as described below.

There have been a number of well documented examples of tag cloning. Firstly, Indala proximity tags, used by many organizations to control physical access to facilities, have been cloned [12]. With a budget of about USD $100 researchers have built a cloning device that is capable of obtaining an Indala tag identifier and replaying it back to a reader. Cloning an Indala proximity tag can allow an attacker to gain access to a secure building. Secondly, the human implantable VeriChip tag has been cloned [13] using a cloning device called the Prox Mark II. The Mexican government was relying on VeriChip tags to protect access to a secure records room to just eighteen of its workers [14]. However, such attacks now mean an attacker can clone a VeriChip tag and gain access to the facilities that the Mexican government was trying to protect. Thirdly, even though the majority of passive tags cannot support cryptography the Texas Instruments Digital Signal Transponder (TI-DST) is an exception.

Unfortunately, it too is vulnerable to tag cloning. Although the TI-DST is protected by a 40-bit secret cryptographic key, researchers have demonstrated that with modest resources, it is possible to capture enough TI-DST data in a short space of time to crack its encryption key. This is an example of why tag based security does not adequately work. The Exxon-Mobil SpeedPass petrol payment system, which relies on TI-DST tags to authenticate its customers, has been shown to be vulnerable to this attack [15]. This may result in attackers being able to obtain free petrol at a customer's expense. Finally, it has been proposed that RFID systems that assign tag identification numbers sequentially, non-randomly, or using small number spaces, an attacker may simply guess a legitimate tag identification number which can then be replayed to a reader using a simulator device [13], in effect cloning a tag. The EPC tag, a special type of RFID tag used in supply chains, such as the Wal-Mart supply chain, are vulnerable to this type of cloning [16]. A cloning attack on EPC tags may allow counterfeit products to gain access to a supply chain. Tag cloning can allow change of tag ownership to occur.

Change of tag ownership, therefore, is serious threat to the security of RFID systems. By simply removing a tag from a subject, an attacker can obtain access to those assets which RFID is being used to protect. Tag cloning can allow an attacker to achieve this on a much wider scale. Thus, change of tag ownership results in the security goals of an RFID system being compromised.

3. Related Work

Our approach is different to the previous research outlined in [6] and [7] in a several of ways. Firstly, our proposal goes beyond preventing tag cloning by actually determining whether the subject to which a tag is attached has changed. It does not prevent a tag from being cloned, but can be used to detect when a clone or a stolen tag is used by an attacker. We believe this is more effective as it questions the legitimacy of the subject, not simply the tag itself. Secondly, our proposal avoids the difficulties on-board the tags, by situating itself within the middleware. As the reader and middleware components are typically accepted as the expensive components of RFID [17] our proposal is more practical. Finally, our proposal controls the perimeter of an RFID system by moderating access through the readers. When an attack has been detected, it can be configured to inform readers to block a tag's access. Systems that carry out this function for other devices are known as intrusion detection systems.

'Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking resources' [18]. When anomalous behaviour is detected an alert is triggered that informs an administrator of a potential breach of security. There are two established approaches to doing this. Anomaly Detection flags abnormal behaviour, whereas, Signature Detection flags behaviour that is close to some previously defined attack definition. Anomaly detection can detect unknown attacks but the result is typically a high false positive rate. However, signature detection can only detect what it knows about, which means it cannot detect new attacks, but this results in a lower false positive rate. Intrusion detection systems are useful in detecting when an attacker has stolen a user's password, or, when legitimate users abuse their privileges [19].

The Intrusion Detection Expert System (IDES) by Denning [8] has been the inspiration for many intrusion detection systems, and is the inspiration behind Deckard. Although there have been significant advancements in the field of intrusion detection over the years [19], we have chosen to start with a simple approach, as this is the first time intrusion detection has been applied to RFID. We believe that if a simple approach like Deckard works, then future research may consider investigating more sophisticated methods of detecting change of tag ownership in RFID systems.

Briefly, IDES is a general purpose statistical anomaly detection system that monitors a system's audit records to look for abnormal patterns of usage. It observes the standard operations that occur within a target system to detect intrusions, such as logins or file accesses. It uses standard deviation and mean statistics to measure subject behaviour which are categorized into discrete time periods called observations. Observations of subject behaviour that significantly deviate from past observations are regarded as intrusions. The system models each subject's normal behaviour with regard to a profile built from object use. A profile characterizes a subject's past behaviour, based on audit log records. The system then classifies a new observation as either normal, in that it fits the subject's profile, or alternatively as abnormal, in which case it is regarded as an intrusion. The IDES model proposed a profile called Location Frequency that measures the number of times a subject logs into a system at different locations. This profile may be especially useful in detecting attackers that log in from locations that authorized subjects rarely use.

The paradigm of intrusion detection may be useful in detecting change of tag ownership. A system that can model the behaviour of authorized subjects within an RFID system may be capable of detecting when an attacker starts to use a cloned or stolen tag. An RFID intrusion detection system like Deckard may secure an RFID system by preventing attackers from gaining access.

4. System Design and Operation

In this section we discuss the design and operation of Deckard, our intrusion detection system for detecting change of tag ownership. The essential components necessary to perform intrusion detection [18] exist within an RFID system. They are: 1. Target System, 2. Feed, 3. Audit Log, 4. Processing Engine, 5. Knowledgebase, and 6. Activity Records. We now discuss how they exist and how we have adapted them to Deckard.

1. **Target System:** An RFID system that has assets worth protecting from change of tag ownership attacks. For example, the SpeedPass payment system, proximity tag systems, or a product supply chain. Our target system was the University of Tasmania's School of Computing RFID proximity tag system. The target system ultimately determines the type of data that an intrusion detection system can utilize to form profiles. Similarly, the subject to which tags are attached influences the nature of the data. That is, supply chain products would exhibit different behavior when compared to humans. In essence, Deckard may be regarded as a passive host based system, as it monitors RFID tag activity occurring at a reader after the fact and is positioned at the middleware level.
2. **Feed:** The raw data traffic that is produced when a tag is read by a reader. Conceptually, the RFID readers are like sensors, sensing when a tag is in range, reading it, and recording its observations into an audit log. The feed data is encapsulated into an audit record with the following attributes: tag identifier, reader identifier, and timestamp.
3. **Audit Log:** The central repository of a system's audit records produced by the feed. Deckard used a flat text file for this purpose, although it could have easily been a database.
4. **Processing Engine:** The controller responsible for executing all system tasks. It periodically retrieves and updates each tag profile from the knowledgebase. After a profile's model has been updated, it confirms whether or not a profile's thresholds are still in check. Finally, it is responsible for generating activity records which alert a system administrator if a profile's threshold has been exceeded.
5. **Knowledgebase:** A collection of tag profiles which encapsulate the normal behavior of tags, and hence, the behavior of subjects to which the tags are attached. A profile encapsulates subject behavior using the following variables: profile name, RFID operation (read/write), tag identifier number, reader identifier number, time-of-day/timestamp, value (measure of current observation), and threshold (measure of past observations).

   Deckard uses statistical methods to look for anomalies. We developed a single profile called Location Frequency Profile (LFP) to look for behavior that may indicate change of tag ownership. It is necessary to remember, we assume a significant deviation away from normal behavior is indicative of a change in tag ownership. Thus, the LFP defines an administrator controllable threshold called Deviations from Mean (DFM), to determine how "different" a current observation can be from the mean value, before an alert is triggered. The system administrators can trade-off the detection rate against error rate using the DFM threshold. Three DFM threshold levels were specified ($\sigma 1$, $\sigma 2$, $\sigma 3$) in our tests as the data the system was tested on is normally distributed, and accordingly, it uses the statistical Empirical Rule [20]. To illustrate, a threshold of $\sigma 1$ allows a current observation to be ± 1 deviation away from the mean of past observations in the LFP. Finally, the window-size parameter controls the trade-off between the number of audit records used to create an observation, system accuracy, and profile testing and training speed. Using a large number of audit records to create an observation does not necessarily produce a better classification, and so the window-size was used to determine how old the audit records can be. All of this information is encapsulated in a profile which is stored in the knowledgebase for each tag. In essence, Deckard uses statistical anomaly detection to look for change of tag ownership.
6. **Activity Records:** These report the result of a profile update. If a profile's threshold has been exceeded by a current observation, a negative activity record is generated which informs the administrator that it has detected suspicious behavior or an attack. Conversely, positive activity records are simply discarded by the system as no suspicious behavior was detected during the profile update.

With the six components of Deckard in mind, Deckard operates in the following manner. Firstly, the reader records the details of an RFID reader's "read" operation into an audit record. This record is then stored inside an audit log. Secondly, the processing engine periodically performs an update on each tag profile. Each profile is retrieved in turn from the knowledgebase in conjunction with its associated audit records. The window-size determines how old these audit records can be. To determine the LFP, the mean number of times a tag has been used is calculated, and so is the acceptable range that it could normally be used in the current time period using the DFM threshold. The DFM signifies the number of

deviations away from the mean a current observation can occur. Finally, an activity record is produced to report the outcome of a profile update; negative if the profiles threshold has been exceeded by the current profile update or positive if no anomalous behaviour was detected. Thus, Deckard uses a model of tag normal behaviour to detect change of tag ownership attacks within background RFID traffic.

5. Performance Evaluation

The ability of Deckard to detect change of tag ownership was evaluated in two testing phases using the LFP. Phase One determined the systems performance when modelling the normal behaviour of tags by calculating the true error rate of the statistical classifier. Given an audit log that does not contain any attacks; the LFP should not detect any attacks or anomalies in the data. If the LFP can model when subjects are behaving normally, it can use that model to detect when subjects are misbehaving. That is, when an attacker has obtained ownership of a subjects tag and started using it. The ability to model a subject's normal behaviour is an underlying requirement of anomaly detection [18]. Next, Phase Two determined the system's performance at detecting synthesized attacks inside a sanitized audit log. That is, whether the system could detect audit records relating to a change of tag ownership in an RFID system. The results from each testing phase are presented in section 6.

The testing phases used data from four sanitized RFID audit logs that were supplied by the School of Computing in the University of Tasmania. The School operates an RFID proximity tag system that controls subject (student and staff) access to computer laboratories. The system records when a subject uses their tag at a reader. The audit logs represented the activity of 327 proximity tags and their subjects between 25/11/2004 to 02/06/2006. There were some 36,294 useable audit records. Although our tests involved four RFID readers, our system is not dependent on the number of readers within a system, as a profile characterizes the behaviour of a subject at an individual reader. We envisage our system being useful to a wide variety of RFID systems.

The data was sanitized prior to being used to preserve the privacy of individual subjects. The sanitization process changed every tag number to a pseudonym that was in no way related to the original tag number. Although sanitization may sometimes remove the content of the background activity and produce an unrealistic representation of the environment [21], this would not have occurred to these sanitized audit logs as the association between a subject and their audit records was not changed. We now discuss the testing methodology used in each phase.

5.1. Phase One: Performance of Classifying the Normal Behaviour of Tags

The aim of Phase One was to determine the performance of the statistical classifier in modelling the normal behaviour of tags on an attack free data set. The true error rate would indicate the system's ability to do this. A low true error rate is desirable as this would mean that the system can successfully model the normal behaviour of tags. The standard way of predicting the true error rate and future performance of a learning procedure is using stratified ten fold cross validation [22]. As our data was time series dependent, we performed ten fold cross validation without stratification so that the underlying structure of the data would not be altered. The data was then fed into the LFP to calculate the true error rate, in the following manner:

1. The audit logs were partitioned into ten parts of approximately the same number of audit records ordered in time.
2. Using the LFP, the first partition, called the test set, was held out, and then the remaining nine partitions in order of occurrence over time, called the training set, were fed into the LFP. This determined the window-size that produced the lowest error rate, called the optimum window-size.
3. After determining the optimum window-size on the training set, it was then applied to the test set to determine the error rate. This process was repeated ten times, each time with a different test and training set.
4. Finally, the ten error estimates were averaged to produce an estimate of the true error rate. This estimated how well the classifier could model the normal behaviour of tags.

The outcome of this would determine the feasibility of moving on to Phase Two which would determine the systems ability to detect attacks.

5.2. Phase Two: Performance at Detecting Attacks

The aim of Phase Two was to determine the system's ability to detect change of tag ownership. The system was designed on the assumption that change of tag ownership would be indicated by a change in a subject's behaviour exhibited through their tag's audit records. Hence, the

question that needed to be answered here was: how different does an attacker's behaviour need to be from an authorized subject's normal behaviour so they can be detected? To answer this question, we had to simulate attacker behaviour dispersed amongst normal background traffic. As there have not actually been any reported cases of change of tag ownership, for example, tag cloning, we had to synthesize our attack data. Synthesizing attack data is an accepted means of evaluating an intrusion detection system [23] [24] [25].

We simulated an artificial stream of attacker behaviour in the following way. Attacks, in the form of duplicate audit records, were copied and reinserted at known locations throughout the audit log. Their amount, called attack intensity, and their occurrence over time, called attack frequency, were synthesized and then slowly increased to simulate an attacker that was escalating their usage of an authorized subject's tag. The attack frequency was used to ensure attack data was adequately distributed amongst the background traffic. Once attacks had been inserted into the data, the system was allowed to make its classification using the LFP. The LFP's classifications were then verified against the known locations of the attacks. This indicated the point at which an attacker's tag usage could be detected from an authorized subject's tag usage, and hence, whether the system could detect a change of tag ownership.

Although [21] discusses the limitations in synthesizing attacks, we believe that our attacks are a realistic representation of the problem. Unlike a computer system that can look for known traits of computer viruses, the characteristics of a change of tag ownership would only be detected by looking for a deviation in subject behaviour. By manipulating the attack intensity and attack frequency, our aim was to adequately represent a change in behaviour.

We have assumed that our data set is free of pre-existing attacks, and that the only attacks are those which we ourselves have inserted. Although [21] discusses that pre-existing attacks may influence the detection rate, we believe that there would be very few, if any, pre-existing attacks. Any bias within the audit logs, we believe, would have been minimized due to the large sample size of audit records used, and the cross-validation we performed on the test and training data sets. We now present the results from each testing phase.

6. Analysis and Results

The results of each testing phase are now presented in the form of confusion matrices. These summarize the performance at each possible configuration of the system using the LFP. The system's ability to detect attacks was measured as follows. True positives and true negatives are correct classifications when the system correctly classified an observation as containing an attack or not containing an attack. False positives and false negatives are incorrect classifications when the system misclassified an observation as either containing an attack, when in fact it did not, or when it failed to detect the presence of an attack [22]. An ideal intrusion detection system has a high true positive rate and a low false positive value. As is most often the case, when one setting in an intrusion detection system is changed, the results that the system produces will vary. Accordingly, we have used Pearson's correlation coefficient (PCC) [20] to indicate the association between the system's different settings. Together the results from each testing phase form our justification for using an intrusion detection system for RFID data.

6.1. Phase One Results

The results in Table 1 indicate that the system can model the normal behaviour of an authorized subject's tag. The average true error rate using DFM σ1 was 27.33% using 92.5% of the data. Conversely, using a less strict threshold DFM σ3, the average true error rate was 13.29% using 96.88% of the data. Thus, there is a trade-off between the DFM threshold (strictness of what constitutes normal behaviour) and the window-size (amount of data used) to produce an accurate model of normal behaviour.

Table 1. Performance of Deckard using DFM threshold as lower and upper bound on tag usage.

	Average Window-Size %			True Error Rate %		
Reader	σ 1	σ 2	σ 3	σ 1	σ 2	σ 3
1	82.5	90	100	26.87	16.48	12.52
2	90	95	100	29.36	15.13	12.13
3	100	92.5	87.5	24.39	14.76	11.86
4	97.5	97.5	100	28.69	18.71	16.65
Average	92.5	93.75	96.88	27.33	16.27	13.29
	PCC of averages 0.97			PCC of averages -0.95		

Our initial instinct was to use the DFM threshold as a lower and upper bound on tag usage. That is, to constrain the range a tag could allowably be used during an observation period. To illustrate, Alice may have used her tag on average five times in the past. As the DFM represents the number of deviations away from this average, a DFM of σ1, would allow an incoming observation to fall within ±1 deviation from this mean. Thus, increasing the DFM threshold makes the system less strict in its classification of what normal behaviour is, as

behaviour can be further away from the mean value of past observations.

Having used the DFM threshold as a lower and upper bound on tag usage, it can be seen from Table 1, that PCC indicates that as the DFM moves from σ1 to σ3, the true error rate declines. For example, for reader one, the true error rate falls from 26.87% to 12.52%. Thus, the system's ability to model normal behaviour improves as the definition of normal behaviour is made less strict.

However, PCC also indicates that as the DFM threshold moves from σ1 to σ3, becoming less strict in its definition of normal behaviour, the system has to use more audit records to produce a lower true error rate, and hence, a better model of normal behaviour. For example, looking at reader one, the average window-size starts at 82.5% and increases to 100%. Thus, as the definition of normal is made less strict, the system uses more audit records to produce a more accurate classification.

In summary, when the system uses the DFM threshold as a lower and upper bound on tag usage, it performs best at modelling the normal behaviour of subjects when it uses the least strict threshold DFM σ3. Consequently, at DFM σ3, the system must use more audit log data to achieve a better model of normal behaviour.

We refined Phase One in an attempt to improve these results. We changed the DFM threshold so that it acted only as an upper bound on tag usage. That is, when a tags usage exceeded this threshold, a negative activity record would be generated, indicating that an attack had been detected. In a real world RFID application, this may represent the problem more adequately as an increase in tag usage would be more indicative of an attacker's behaviour. To illustrate, when Mallory clones Alice's proximity tag, he may use it more frequently than Alice, or in locations that Alice has rarely used. Furthermore, it may be more feasible to detect subjects whose tags were being misused on a regular basis as they may cause a greater loss to an RFID application.

In Table 2, it can be seen that the association between the DFM threshold, window-size, and true error rate is still present. When compared to using the DFM as a lower and upper bound on tag usage, however, the effect of removing the lower bound on the DFM threshold produces a lower true error rate, and the system does not need to use as much audit log data. For example, at DFM σ1, the average true error rate is 18.52% using 68.75% of the data. A much better performance result when compared to the previous system configuration's result of 27.33% average true error rate using 92.5% of the data.

Table 2. Performance of Deckard using DFM threshold as upper bound on tag usage.

	Average Window-Size %			True Error Rate %		
Reader	σ1	σ2	σ3	σ 1	σ 2	σ3
1	65	87.5	95	17.63	12.53	9.89
2	75	80	92.5	20.38	12.73	9.76
3	50	85	82.5	17.05	11.61	9.21
4	85	97.5	100	19.03	14	11.94
Average	68.75	87.5	92.5	18.52	12.72	10.2
	PCC of averages 0.95			PCC of averages -0.98		

In summary, a lower true error rate, hence, better model of normal behaviour, can be obtained when Deckard uses the DFM threshold as an upper bound on tag usage. The systems performance at modelling the normal behaviour of tags improves when the DFM is relaxed from σ1 to σ3. Consequently, this would actually make it easier for an attacker to evade detection as the system is less strict in what it regards as an anomaly because its definition of normal is not very strict.

6.2. Phase Two Results

As Deckard assumes that a change of tag ownership is indicated by a change in tag behaviour, the system's model of normal behaviour has been applied to a dataset with artificial attacks to see if it can detect them, and hence, detect change of tag ownership. We decided to use the DFM threshold as an upper bound on tag usage as Phase One indicated this produced the best model of normal behaviour. We now present the results from Phase Two testing.

The system detected the greatest number of attacks using DFM σ1, as seen in Table 3. The average true positive rate, which is a tag's true positive rate of detection averaged, was 76.26%, compared to the average true positive rate of 62.97% using DFM σ2, and 46.30% using DFM σ3. At the same time, however, the system incorrectly detected authorized behaviour as attacks. Overall, DFM σ1 made the greatest number of incorrect classifications. Its average false positive rate was 8.40%, compared to 3.69% using DFM σ2 and 2.52% using DFM σ3. Thus, there is a trade-off between the true positive rate and false positive rate. The system detected the greatest number of attacks using DFM σ1, but this setting also produced the greatest number of errors.

Table 3. Performance of Deckard in detecting change of tag ownership attacks.

	True Positive Rate %			False Positive Rate %		
Attack Intensity	σ1	σ2	σ3	σ1	σ2	σ3
Attack Frequency 1						
1	63.64	63.64	63.64	10.77	4.71	2.66
2	81.82	63.64	63.64	9.69	3.49	2.52
3	81.82	81.82	72.73	7.04	3.01	2.39
Average	75.76	69.70	66.67	9.17	3.74	2.52
PCC	0.87	0.87	0.87	-0.97	-0.97	-1.00
Attack Frequency 2						
1	63.64	45.45	36.36	10.09	4.20	2.76
2	81.82	63.64	36.36	7.65	3.61	2.48
3	86.36	81.82	50.00	6.15	3.12	2.17
Average	77.27	63.64	40.91	7.96	3.64	2.47
PCC	0.94	1.00	0.87	-0.99	-1.00	-1.00
Attack Frequency 3						
1	63.64	36.36	24.24	10.29	4.35	2.88
2	78.79	57.58	30.30	7.63	3.49	2.59
3	84.85	72.73	39.39	6.33	3.23	2.27
Average	75.76	55.56	31.31	8.08	3.69	2.58
PCC	0.97	1.00	0.99	-0.98	-0.96	-1.00
Totals						
Overall averages	76.26	62.97	46.30	8.40	3.69	2.52
PCC of averages	0.00	-1.00	-0.97	-0.82	-0.50	0.54

It is equally important to consider the effect that attack intensity and attack frequency has on the system's performance in detecting change of tag ownership. PCC indicates that as attacker behaviour becomes more different from authorized behaviour (attack intensity increases), attacker behaviour becomes easier to detect. The true positive rate of DFM σ1 at an attack frequency of one, increases from 63.64% to 81.82% as attack intensity increases. This association is the same for DFM σ2 and σ3; however, their true positive rates are slightly less than DFM σ1. Thus, Deckard performs better at detecting change of tag ownership as attacker behaviour becomes more different (attack intensity increases). This is a desirable feature of an intrusion detection system.

Another positive aspect of the results, PCC indicates that as attacker behaviour becomes more different from authorized behaviour (attack intensity increases); errors in the system fall. The false positive rate at DFM σ1, at an attack frequency of one, falls from 10.77% to 7.04% as attack intensity increases. Similarly, this association can be seen across DFM σ2 and σ3. Thus, as attacker behaviour becomes more different (attack intensity increases), Deckard produces fewer errors.

The next variation that we tested was the effect of distributing attacks across the dataset. As the number of attack periods increases (attack frequency increases), and attacker behaviour occurs more frequently throughout the dataset, PCC of averages indicates that the behaviour of an attacker is more consistently detected using DFM σ1. Table 3 indicates an average true positive rate starts at 75.76% at an attack frequency of one, and returns to 75.76% at an attack frequency of three. This means that Deckard's ability to detect change of tag ownership remains relatively stable using DFM σ1.

In contrast, PCC of averages indicates the true average true positive rate of DFM σ2 and σ3 decreases as attacker behaviour occurs more frequently. The average true positive rate starts at 69.70% for DFM σ2, and falls to 55.66% at attack frequency three. Thus, these DFM settings are influenced by frequency of attacks, and therefore, they detect fewer attacks as attacks occur more frequently and become incorporated as normal behaviour. This is a problem faced by all systems that aim to detect anomalous behaviour.

At the same time, PCC of averages indicates that the false positive rate falls as attacks occur more frequently. The average false positive rate at DFM σ1 attack frequency one, starts at 9.17% and falls to 8.08% at an attack frequency of three. Thus, the system makes fewer errors as attacks occur more frequently. Although the average false positive rate of DFM σ1 is 8.40%, significantly higher than the other DFM settings, PCC of averages for false positive rate actually indicates that overtime this will perform better than DFM σ3. Using DFM σ3, the false positive rate actually increases from 2.52% to 2.58%. This means that DFM σ1 will produce fewer errors than the other DFM settings when attacks occur more frequently.

In summary, the systems ability to detect change of tag ownership depends on the trade-off between the DFM threshold (strictness of what is classified as normal), attack intensity and attack frequency. Overall, DFM σ1 detected the greatest number of attacks, having produced the highest average true positive rate of 76.26%, but also the highest average false positive rate of 8.40%. In an RFID system with no security to detect change of tag ownership, these results are promising. At the same time, the false positive rate is Deckard's Achilles heel. Any security system that prevents its authorized subjects from carrying out their job is not good enough to be implemented in the real world.

The performance of Deckard is on par with similar statistical anomaly detectors. For example, [26] showed a statistical anomaly detection system based on the Chi-Square statistic and standard deviation produced a detection rate of around 75%. The high false positive rate produced by Deckard is typical of anomaly based data mining intrusion detection systems [23]. Thus, for systems such as this to be useable in real environments the false positive rate needs to be improved.

7. Conclusions

We have shown that change of RFID tag ownership is a serious threat to the security of RFID systems which must be solved. Our approach improves upon previous approaches as we detect when an attacker has obtained ownership of a tag. This is more important than simply determining when a tag has been cloned. In addition, our proposal avoids the limitations of placing security at the tag level, by instead placing security in the middleware.

The system we built, Deckard, can detect synthesized attacks inside a sanitized audit log. The results are promising as there is currently no system, to our knowledge, that can carry out this task for RFID systems. However, the false positive rate dominates the current system's feasibility in the real-world. To illustrate, the SpeedPass petrol payment system (see section 2) has over 7 million customers. A false positive rate of 8.40% in SpeedPass would cause the system to fail by preventing its authorized subjects from paying for petrol. Nevertheless, in applications where there is a potential for more significant loss, like an ePassport system which is controlling access to country borders; the security that an RFID intrusion detection system like Deckard may offer may be better than no security at all.

We believe that our results are sufficiently encouraging to suggest future research that may consider using more sophisticated methods in intrusion detection systems aimed at detecting change of RFID tag ownership.

8. Future Work

The following areas will be considered for further investigation:

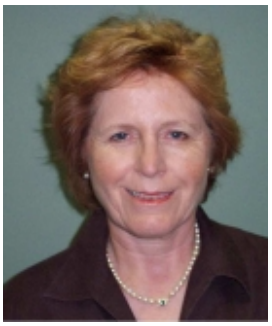
- Replace the statistical detector with a more sophisticated detection technique.
- Incorporate contextual information of an RFID system into the detection process. For example, distance between RFID readers. These may form plausibility checks or signatures to enhance anomaly detection.
- Model the normal behaviour of a specific RFID application and apply it to detecting attacks.

References

1. Garfinkel, S. and H. Holtzman, *Understanding RFID Technology*, in *RFID: Applications, Security, and Privacy*, S. Garfinkel and B. Rosenberg, Editors. 2005, Addison Wesley. p. 15-36.
2. Quirk, R. *E-Pedigree's Evolution*. [Website] 2007 5 March 2007 [cited 2007 27/06/2007]; Available from: http://www.rfidjournal.com/article/articleview/3109/3/82/.
3. Ranasinghe, D.C., D.W. Engels, and P.H. Cole. *Low-Cost RFID Systems: Confronting Security and Privacy*. in *Auto-ID Labs Research Workshop*. 2004. Zurich, Switzerland.
4. Sarma, S., *Towards the 5-cent tag*. 2001, MIT Auto-ID Center.
5. Weis, S.A., S.E. Sarma, R.L. Rivest, and D.W. Engels. *Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems*. in *International Conference on Security in Pervasive Computing*. 2003. Germany: Springer-Verlag.
6. Juels, A., *RFID Security and Privacy: A Research Survey*. 2005, RSA Laboratories.
7. Avoine, G. *Bibliography on Security and Privacy in RFID Systems*. 2007 May 2, 2007 [cited May 3, 2007].
8. Denning, D.E., *An Intrusion-Detection Model.* IEEE Transactions On Software Engineering, 1987. **SE-13, No. 2**: p. 222-232.
9. Chapman, M. *Blade Runner: Frequently Asked Questions*. [Website] 1995 07/08/1995 [cited 2007 23 March]; Available from: http://www.faqs.org/faqs/movies/bladerunner-faq/.
10. Mullen, D. and B. Moore, *Automatic Identification and Data Collection: What the future holds*, in *RFID: Applications, Security, and Privacy*, S. Garfinkel and B. Rosenberg, Editors. 2005, Addison Wesley. p. 3-13.
11. Schuermann, J. *Information technology - Radio frequency identification (RFID) and the world of radio regulations*. ISO Bulletin 2000 May 2000 [cited 22 June 2007]; 3-4]. Available from: http://www.iso.org/iso/en/commcentre/pdf/Radio0005.pdf.
12. Westhues, J., *Hacking the Prox Card*, in *RFID: Applications, Security, and Privacy*, S. Garfinkel and B. Rosenberg, Editors. 2005, Addison Wesley. p. 291-301.
13. Halamka, J., A. Juels, A. Stubblefield, and J. Westhues, *The Security Implications of VeriChip Cloning*. 2006, RSA.
14. Albrecht, K. and L. McIntyre, *Spy Chips*. Second ed. 2005, Nashville,Tennessee: Nelson Current.
15. Bono, S., M. Green, A. Stubblefield, A. Juels, A. Rubin, and M. Szydlo. *Security Analysis of a Cryptographically-Enabled RFID Device*. in *14th USENIX Security Symposium*. 2005.
16. Juels, A. *Strengthening EPC Tags Against Cloning*. in *4th ACM Workshop on Wireless Security*. 2005. New York,USA: ACM Press.
17. IDTechEx. *RFID market to reach $7.26 Bn in 2008*. 2005 April 10 2005 [cited 2006 2 November]; Available from: http://www.idtechex.com/products/en/articles/00000169.asp.
18. Amoroso, E., *Intrusion Detection: An introduction to internet surveillance, correlation, trace back, traps, and response*. First ed. 1999: Intrusion.Net Books.

19. Axelsson, S., *Intrusion Detection Systems: A Survey and Taxonomy*. 2000, Department of Computer Engineering, Chalmers University of Technology: Goteborg, Sweden.
20. Jaisingh, L.R., *Statistics for the utterly confused*. 2000: McGraw Hill.
21. Mell, P., V. Hu, R. Lippmann, J. Haines, and M. Zissman, *An Overview of Issues in Testing Intrusion Detection Systems*. 2003, National Institute of Standards and Technology
Massachuesetts Institute of Technology Lincoln Laboratory.
22. Witten, I.H. and E. Frank, *Data Mining*. Third ed. 2000: Morgan Kaufmann Publishers.
23. Lee, W., S. J.Stolfo, P. K.Chan, E. Eskin, W. Fan, M. Miller, S. Hershkop, and J. Zhang. *Real Time Data Mining-based Intrusion Detection*. in *DARPA Information Survivability Conference and Exposition II, 2001, DISCEX'01*. 2001. Anaheim,CA,USA.
24. P.Lippmann, R., D. J.Fried, I. Graf, J. W.Haines, K. R.Kendall, D. McClung, D. Weber, S. W.Webster, D. Wyschogrod, R. K.Cunningham, and M. A.Zissman. *Evaluating Intrusion Detection Systems: The 1998 DARPA Off-line Intrusion Detection Evaluation*. in *DARPA Information Survivability Conference and Exposition*. 1999.
25. C.Lee, S. and D. V.Heinbuch, *Training a Neural-Network Based Intrusion Detector to Recognize Novel Attacks.* IEEE Transactions on Systems, Man, and Cybernetics, 2001. **31,No.4**(Part A: Systems and Humans).
26. Ye, N. and Q. Chen, *An anomaly detection technique based on chi-square statistic for detecting intrusion into information systems.* Quality and Reliability Eng., 2001. **17**(2): p. 105-112.

Luke Mirowski received his B.Comp. (Hons), from the University of Tasmania, Australia, in 2006. He is currently a PhD candidate. His research interests include RFID and Computer Security.



Jacqueline Hartnett first degree is a B.A. (Hons) in Geography from Exeter University in the UK and a MComp from the University of Tasmania. After graduating, she worked in the then new computing department of the Royal Dutch Shell group and then for IBM both in Australia and the UK. She has taught computer security courses in the University of Tasmania since 1993. Her current research interests are the use of authentication and access control as a means of maintaining privacy and confidentiality of personally identified data and the development of intrusion detection techniques for groups of collaborating network gateways. She is a member of Australian Standard Review Committees; IT-014 04 System & Data Security, Integrity and Privacy subcommittee, and IT-014-06-08 Electronic Communications in Health working groups. She is also currently chair of the Tasmanian branch of the Australian Computer Society.