Privacy Protection in Dynamic Systems Based on RFID Tags

Jacek Cichoń Marek Klonowski Mirek Kutyłowski

Wrocław University of Technology
DELIS project

PERSEC'2007, White Plains, NY

RFID Technology

Basic Properties

1. RFID device responds to a reader
2. almost no internal logic
3. minimal memory

Potential Applications

1. objects identification
2. movement tracing
3. electronically readable ID's

Advantages

Cheap and uncomplicated to use

Privacy Problems

Big Brother Scenario

1. trace people by tracing their items,

Privacy Problems

Big Brother Scenario

1. trace people by tracing their items,
2. derive consumer preferences, health condition, behavior

Privacy Problems

Big Brother Scenario

1. trace people by tracing their items,
2. derive consumer preferences, health condition, behavior
3. .. new sources of personal data available to anybody

Privacy Problems II

Dangers

1. surveillance: among others for spying, criminal and terrorist purposes

Privacy Problems II

Dangers

1. surveillance: among others for spying, criminal and terrorist purposes
2. unfair competition

Legal Situation

Personal Data Protection Regulations

1. most countries (excluding USA), strict rules in the EU
2. any data concerning **a person that can be identified** is **personal data**
   EU Directive

Legal Situation

Personal Data Protection Regulations

1. most countries (excluding USA), strict rules in the EU
2. any data concerning **a person that can be identified** is **personal data**
   EU Directive
3. personal data protection regarded as condition of freedom of the citizens
4. society becoming sensitive to *personal data protection*,

Legal Situation

Personal Data Protection Regulations

1. most countries (excluding USA), strict rules in the EU
2. any data concerning **a person that can be identified** is **personal data**
   EU Directive
3. personal data protection regarded as condition of freedom of the citizens
4. society becoming sensitive to *personal data protection*,
5. personal data protection obligatory,
   non-respecting is a crime, high penalties

Paradox

Two conflicting demands

1. an RFID tag should show its ID - since it is the main purpose of RFID
2. an RFID tag must restrict showing its ID due to personal data protection

privacy protection - the main usability problem of RFID technology in EU

Countermeasures

Solutions

killing **destroy RFID after use**
but then RFID's not much useful

Countermeasures

Solutions

killing **destroy RFID after use**
but then RFID's not much useful

blocking **block RFID after use**
unblocking by legitimate readers only, but what a problem to capture a reader?

Countermeasures II

Solutions II

hash-lock **re-activation with a key**

additional logic and memory on the RFID, password management

Countermeasures II

Solutions II

hash-lock **re-activation with a key**
additional logic and memory on the RFID, password management

re-encryption **change encoding for untracability**
heavy, asymmetric methods, not really suited for small memory size

Algorithm description

Parameters

1. each ID is a random sequence of n bits,
2. each ID evolve itself, each successful activation changes ID a little bit

Update procedure

1. a subset $B \subseteq \{1, \dots, n\}$ of cardinality l is chosen uniformly at random, say let $B = \{i_1, \dots, i_l\}$,
2. For each $j \leq l$, the bit b_{i_j} of the ID is set uniformly at random: $b_{i_j} \leftarrow b \in_U \{0, 1\}$, independently from the previous value of b_{i_j}.

Update

1	0	0	1	0	1	1	1	0	0	1	1	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

1	0	0		0	1	1	1		0	1	1	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

1	0	0	0	0	1	1	1	0	0	1	1	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

	0	0	0	0	1	1	1	0	0		1	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	0	0	0	1	1	1	0	0	0	1	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	0	0	0	1	1	1	0		0		0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	0	0	0	1	1	1	0	1	0	0	0	1

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0		0	0	1	1	1	0	1	0	0	0	

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0	1	1	1	0	1	0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0		1	1	0		0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0	0	1	1	0	0	0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0	0	1	1	0	0	0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0	0	1	1	0	0	0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Update

0	0	1	0	0	0	1	1	0	0	0	0	0	0

current ID

1	0	0	1	0	1	1	1	0	0	1	1	0	1

original ID

Motivation

Solution idea

1. after a certain time a new ID is so different from the original one that it cannot be linked anymore,
2. loosing control over RFID communication for a certain time results in unlinkability,
3. **even a powerful adversary cannot spy always and everywhere.**

Recognizing evolving ID's by the system

System activities during an activation

1. the system running with own RFID's has a database with records [ID, object description]

Recognizing evolving ID's by the system

System activities during an activation

1. the system running with own RFID's has a database with records [ID, object description]
2. after each activation ID is compared with the database, ID update is recorded by overwriting the old ID

Questions

Problems

1. **How long it takes so that an adversary cannot link an old ID with a new one?**
2. **How frequent are the collisions?**
   collisions require special handling!

Unlinkability

Ad hoc Answers

1. even if two n-bit ID are unrelated, on about 50% of positions they agree
2. would it be better to choose l positions and **switch the bits** there?

Unlinkability

Ad hoc Answers

1. even if two n-bit ID are unrelated, on about 50% of positions they agree
2. would it be better to choose l positions and **switch the bits** there?
3. **No: due to some stochastic peculiarities**

Unlinkability as Stochastic Distance

Markov Process

1. value of a stochastic process - current ID
2. step of the process: random update step (of the algorithm)
3. initial state: starting ID

Unlinkability as Stochastic Distance

Markov Process

1. value of a stochastic process - current ID
2. step of the process: random update step (of the algorithm)
3. initial state: starting ID

4. D_t - probability distribution of ID's after step t
5. D_t should be almost uniform,

Total Variation Distance

Measure of distance between two probability distributions

1. given random variables $\Gamma_1, \Gamma_2 : \Omega \to \mathcal{Y}$
2. total variation distance

$$\mathrm{TVD}(\Gamma_1, \Gamma_2) = \frac{1}{2} \sum_{y \in \mathcal{Y}} |\Pr(\Gamma_1 = y) - \Pr(\Gamma_2 = y)| \ .$$

Convergence

Approaching uniform distribution

Let D_t be the state of the RFID-tag after t-th activation of the tag according to the algorithm, starting with an arbitrary initial ID.
Let

$$\tau(\varepsilon) = \max_s \min_t \{t \in \mathbf{N} \mid \mathrm{TVD}(D_t, \mathrm{uniform}) \leq \varepsilon \wedge D_0 = s\} .$$

Convergence

Approaching uniform distribution

Let D_t be the state of the RFID-tag after t-th activation of the tag according to the algorithm, starting with an arbitrary initial ID.
Let

$$\tau(\varepsilon) = \max_s \min_t \{t \in \mathbf{N} \mid \mathrm{TVD}(D_t, \mathrm{uniform}) \leq \varepsilon \wedge D_0 = s\} .$$

Convergence

For this process (l=number of bits set in one update), for each $k > 1$

$$\tau\left(\frac{1}{n^k}\right) \leq \frac{n \cdot \log n^{k+1}}{l} .$$

Proof Techniques

Rapid Mixing of Markov Chains

1. just a standard use of path coupling technique

Proof Techniques

Rapid Mixing of Markov Chains

1. just a standard use of path coupling technique
2. however: less restrictive divergence measures required, but still keeping it guaranteed safe in a stochastic sense

bounded variation distance?

new proof techniques?

Collision

Collision Event

1. during an update an ID reaches the same value as another ID used (and stored in the database)
2. additional updates in order to escape such a condition

How much time 2 IDs need to reach the same value?

Problem statement

1. is suffices to examine time T for reaching all-zero state

How much time 2 IDs need to reach the same value?

Problem statement

1. is suffices to examine time T for reaching all-zero state
2. what is T, if an ID contains only w ones?

How much time 2 IDs need to reach the same value?

Problem statement

1. is suffices to examine time T for reaching all-zero state
2. what is T, if an ID contains only w ones?

Results

- for $w = \sqrt{n}$, we need $T \approx n^{\sqrt{n}/2}$.
- for $w = n^{1/8}$, we need $T \approx n^{7/8 \cdot n^{1/8}}$.

How much time 2 IDs need to reach the same value?

Problem statement

1. is suffices to examine time T for reaching all-zero state
2. what is T, if an ID contains only w ones?

Results

- for $w = \sqrt{n}$, we need $T \approx n^{\sqrt{n}/2}$.
- for $w = n^{1/8}$, we need $T \approx n^{7/8 \cdot n^{1/8}}$.
- In both cases T is superpolynomial in n, while the time required for reaching almost uniform distribution is only slightly higher than linear.

How much time 2 IDs need to reach the same value?

Problem statement

1. is suffices to examine time T for reaching all-zero state
2. what is T, if an ID contains only w ones?

Results

- for $w = \sqrt{n}$, we need $T \approx n^{\sqrt{n}/2}$.
- for $w = n^{1/8}$, we need $T \approx n^{7/8 \cdot n^{1/8}}$.
- In both cases T is superpolynomial in n, while the time required for reaching almost uniform distribution is only slightly higher than linear.
- It follows that IDs tend to *escape from each other*.

Collision Free Periods

Many IDs and Birthday Paradox

1. what is the probability that with k IDs in M steps of the protocol there are no collisions when we start with a random distribution?

Collision Free Periods

Many IDs and Birthday Paradox

1. what is the probability that with k IDs in M steps of the protocol there are no collisions when we start with a random distribution?
2. expected number of collisions in M steps

$$\simeq M(1 - e^{-\frac{k^2}{2^n}}) \simeq \frac{Mk^2}{2^{n+1}}$$

Collision Free Periods

Many IDs and Birthday Paradox

1. what is the probability that with k IDs in M steps of the protocol there are no collisions when we start with a random distribution?
2. expected number of collisions in M steps

$$\simeq M(1 - e^{-\frac{k^2}{2^n}}) \simeq \frac{Mk^2}{2^{n+1}}$$

3. pbb of collision $\leq$ expected number of collisions so pbb of *no collision within M steps* $\leq \frac{Mk^2}{2^{n+1}}$.

Collision Free Periods

Many IDs and Birthday Paradox

1. what is the probability that with k IDs in M steps of the protocol there are no collisions when we start with a random distribution?
2. expected number of collisions in M steps

$$\simeq M(1 - e^{-\frac{k^2}{2^n}}) \simeq \frac{Mk^2}{2^{n+1}}$$

3. pbb of collision $\leq$ expected number of collisions so pbb of *no collision within M steps* $\leq \frac{Mk^2}{2^{n+1}}$.
4. for $k < \sqrt{\pi 2^{n-1}}$, $p \in [0, 1]$ and $M < \frac{p2^{n+1}}{k^2}$, then the pbb of *at least one collision within M steps* is $< p$.

Minimal Distance

Expected minimal distance for a random set C of ID's:

	$\lvert C\rvert = 2^{10}$	$\lvert C\rvert = 2^{15}$	$\lvert C\rvert = 2^{20}$	$\lvert C\rvert = 2^{25}$
$n = 40$	4.40111	1.66771	0	0
$n = 50$	7.30512	3.96913	1.55622	0
$n = 60$	10.4371	6.60167	3.68943	1.47741
$n = 70$	13.7348	9.46138	6.13539	3.48876
$n = 80$	17.1601	12.4914	8.80223	5.79598

Conclusions

1. stochastic behavior of the process quite well understood
2. stable security conditions
3. .. one has to prevent activation by a non-legitimate reader