Passive Cryptanalysis of an Ultralightweight Authentication Protocol of RFIDsec'10 Asia

Authors:
P. Peris-Lopez (U. Delft, Netherlands)
J. C. Hernandez-Castro (U. Portsmouth, UK)
R. Phan (U. Loughborough, UK)
Juan E. Tapiador (U. York, UK)
T. Li (I^2R, Singapore)

Workshop on RFID Security, June 7 - 9, 2010, Istanbul

Abstract

At RFIDSec'10 Asia, Yeh, Lo and Winata proposed a process-oriented ultralightweight RFID authentication protocol. This protocol is claimed to provide strong security and robust privacy protection, while at the same time the usage of resources on tags is optimized. Nevertheless, in this paper we show how the protocol does not achieve any of its intended security objectives; the main result is that the most valuable information stored on the tag, that is, the static identifier ID, is easily recovered even by a completely passive attacker in a number of ways. More precisely, we start by presenting a traceability attack on the protocol that allows tags to be traced. This essentially exploits the fact that the protocol messages leak out at least one bit of the static identifier. We then present a passive attack (named Norwegian attack) that discloses $\lfloor \log_2 L \rfloor$ bits of the ID, after observing roughly $O(L)$ authentication sessions. Although this attack may seem less feasible in retrieving the full 96-bits of the ID due to the large number of eavesdropped sessions involved, it is already powerful enough to serve as a basis for a very effective traceability attack. Finally, we use a recent cryptanalysis technique (called Tango attack) which allows for an extremely efficient full disclosure attack, capable of revealing the value of the whole ID after eavesdropping only a very small number of sessions.

Keywords: RFID, Cryptanalysis, Ultralightweight, Authentication

Yeh-Lo-Winata Protocol (I)

Step 1 *Reader* $\rightarrow$ *Tag*: *Hello*

Step 2 *Tag* $\rightarrow$ *Reader*: IDS_t

Step 3 *Reader* $\rightarrow$ *Tag*: $A \parallel B \parallel C \parallel flag$

If $(IDS_t = IDS_{tr_{new}})$: $flag = 0$ and $K = K_t$.

Else: $flag = 1$ and $K = ID$.

$$\begin{aligned} A &= (IDS \oplus K) \oplus n_1 \\ B &= (IDS \vee K) \oplus n_2 \\ C &= (\widehat{K} \oplus n_1) + n_2 \qquad \widehat{K} = Rot(K \oplus n_2, n_1) \end{aligned}$$

Step 4 *Tag* extracts $\{n_1, n_2\}$, computes $\widehat{K}$ and verifies C.

Then *Tag* $\rightarrow$ *Reader*: D

$$D = (\widehat{K'} \oplus n_2) + n_1 \qquad \widehat{K'} = Rot(K \oplus n_1, n_2)$$

Yeh-Lo-Winata Protocol (II)

Step 5 *Reader* computes $\widehat{K'}$ and verifies D. If OK, it updates the secrets:

$$\begin{aligned} IDS_{tr_{old}} &= IDS \\ IDS_{tr_{new}} &= (IDS + (ID \oplus \widehat{K'})) \oplus n_1 \oplus n_2 \\ K_{tr} &= \widehat{K} \end{aligned}$$

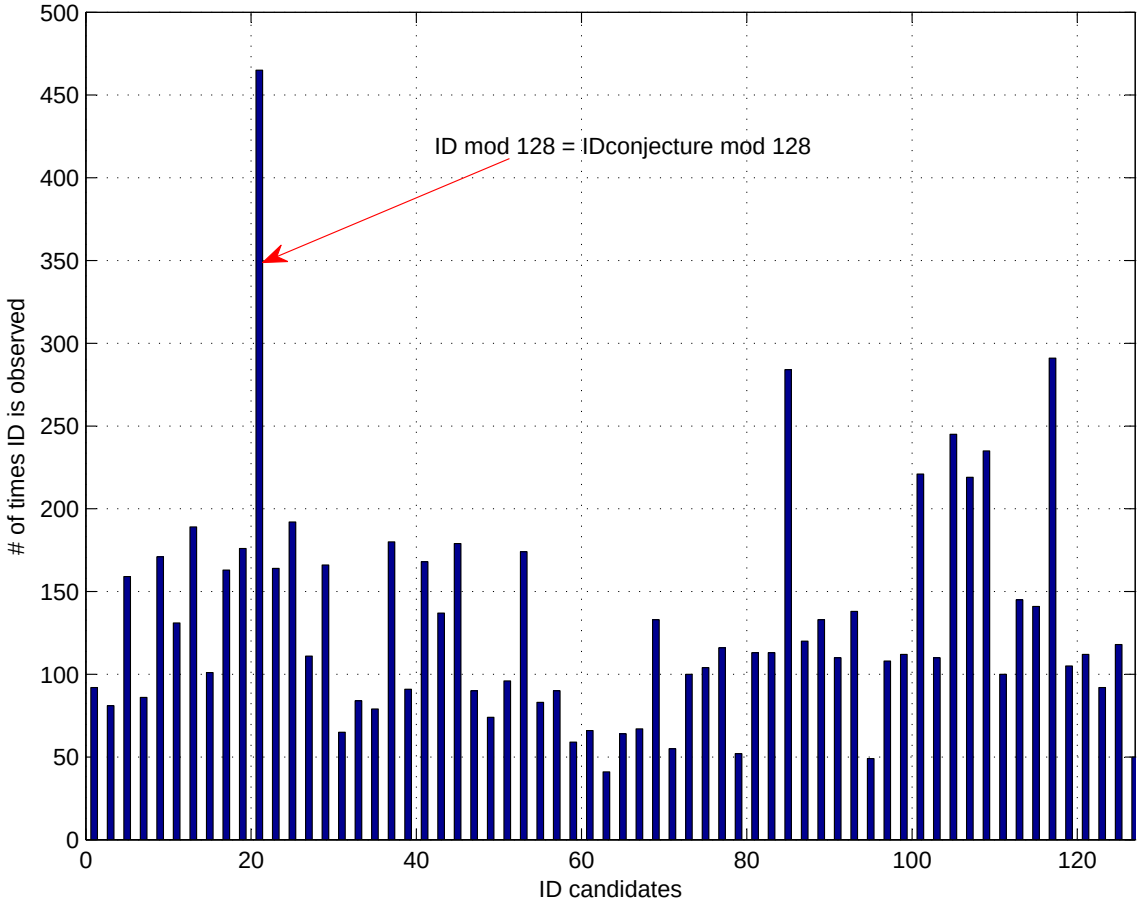
Reader $\rightarrow$ *Tag*: *Update command*

Step 6 *Tag* updates IDS and K

Full Disclosure Norwegian Attack (I)

1. For $i = 0$ to L
2. $\quad Observations[i] = 0$
3. Repeat a sufficiently high number of times N the following steps:
4. $\quad$ Observe an authentication session and get IDS, A, B, C and D
5. $\quad$ Check if for these values it holds that $C \mod L = D \mod L$
6. $\quad$ If this is not the case, go to step 4.
7. $\quad$ Perform the following tasks:
8. $\quad\quad$ Wait for the authentication session to finish.
9. $\quad\quad$ Send to the tag a "Hello" message to obtain $IDS_{tr_{new}}$.
10. $\quad\quad$ Compute $ID_{estimated} \mod L = (IDS_{tr_{new}} - IDS) \oplus D \mod L$
11. $\quad\quad$ Increment $Observations[ID_{estimated}]$
12. Filter: find $ID_{conjecture}$, the maximum of the values in $Observations[i]$.
13. Guess that $ID_{conjecture} = ID \mod L$.

Full Disclosure Norwegian Attack (II)



Histogram of ID candidates ($L = 128,\ N = 2^{18}$)

Full Disclosure Tango Attack

- Can we do it better? Here's the idea:
  - *How much information about the secrets is leaked out by the public messages exchanged during one session?*
- Let's consider only very simple combinations of public messages after session i:

$$L^k = a_0 IDS^k \oplus a_1 A^i \oplus a_2 B^i \oplus a_3 C^i \oplus a_4 D^i \oplus a_5 IDS^{k+1} \quad a_i \in \{0,1\}$$

  and then see whether there's any correlation between L^k and ID

- One simple measure: bias w.r.t. optimal Hamming distance

$$\epsilon = \left| d_H(L^k, ID) - \frac{m}{2} \right|$$

A Scaled-down Example

$ID(base10) = 85$ $\qquad$ $ID = [0, 1, 0, 1, 0, 1, 0, 1]$

- **Session k:**
  Eavesdropping of vectors $\{IDS^k, A^k, B^k, C^k, D^k, IDS^{k+1}\}$
  Computing of an approximation: i.e. $ID_{approx}(1) =$ [0 1 0 1 1 1 1 1]

- **Session $k+1$:**
  Eavesdropping of vectors $\{IDS^{k+1}, A^{k+1}, B^{k+1}, C^{k+1}, D^{k+1}, IDS^{k+2}\}$
  Computing of an approximation: i.e. $ID_{approx}(2) =$ [0 1 0 1 0 1 0 0]

- **Session $k+2$:**
  Eavesdropping of vectors $\{IDS^{k+2}, A^{k+2}, B^{k+2}, C^{k+2}, D^{k+2}, IDS^{k+3}\}$
  Computing of an approximation: i.e. $ID_{approx}(3) =$ [0 1 1 0 0 1 0 1]

- **Conjecture ID:**
  Sum of the vectors:

$$ID_{approx} = \begin{array}{l} [0\ 1\ 0\ 1\ 1\ 1\ 1\ 1] \\ [0\ 1\ 0\ 1\ 0\ 1\ 0\ 0] \\ + \ [0\ 1\ 1\ 0\ 0\ 1\ 0\ 1] \\ \hline [0\ 3\ 1\ 2\ 1\ 3\ 1\ 2] \end{array}$$

  Average value:

$$\begin{cases} \text{if } (id_i^{approx} \geq \gamma) & id_i^{conjecture} = 1 \\ \text{if } (id_i^{approx} < \gamma) & id_i^{conjecture} = 0 \end{cases}$$

  i.e. If $\gamma = 1.5$ $\qquad$ $ID_{conjecture} = [0, 1, 0, 1, 0, 1, 0, 1]$

  Conjecture: $\qquad$ $ID_{conjecture}(base10) = 85$

No Trivial Approximations

L	$d_H(L, ID)$
A	47.94 ± 4.9481
B	48.02 ± 4.9290
C	47.91 ± 4.9111
D	47.89 ± 4.8949
$IDS_{tr_{old}}$	48.01 ± 4.94
$IDS_{tr_{new}}$	48.01 ± 4.94
$A \oplus B$	47.96 ± 4.85
$A \oplus C$	48.00 ± 4.95
$A \oplus D$	48.05 ± 4.87
$A \oplus IDS_{tr_{old}}$	47.98 ± 4.98
$A \oplus IDS_{tr_{new}}$	48.01 ± 4.95
$B \oplus C$	47.99 ± 4.88
$B \oplus D$	48.00 ± 4.91
$B \oplus IDS_{tr_{old}}$	48.00 ± 4.87
$B \oplus IDS_{tr_{new}}$	47.96 ± 4.94
$C \oplus D$	47.95 ± 4.94

L	$d_H(L, ID)$
$C \oplus IDS_{tr_{old}}$	47.97 ± 4.86
$C \oplus IDS_{tr_{new}}$	47.93 ± 4.87
$D \oplus IDS_{tr_{old}}$	47.93 ± 4.92
$D \oplus IDS_{tr_{new}}$	48.01 ± 4.84
$IDS_{tr_{old}} \oplus IDS_{tr_{new}}$	47.99 ± 4.95
$A \oplus B \oplus C$	48.07 ± 4.89
$A \oplus B \oplus D$	48.08 ± 4.94
$A \oplus B \oplus IDS_{tr_{old}}$	47.94 ± 4.85
$A \oplus B \oplus IDS_{tr_{new}}$	48.05 ± 4.90
$A \oplus C \oplus D$	48.03 ± 4.93
$A \oplus C \oplus IDS_{tr_{old}}$	47.95 ± 4.91
$A \oplus C \oplus IDS_{tr_{new}}$	47.97 ± 4.86
$A \oplus D \oplus IDS_{tr_{old}}$	47.95 ± 4.89
$A \oplus D \oplus IDS_{tr_{new}}$	48.05 ± 4.90
$A \oplus IDS_{tr_{old}} \oplus IDS_{tr_{new}}$	47.99 ± 4.91
$B \oplus C \oplus D$	47.96 ± 4.86

... and many more with similar results

Some Manipulations...

- From the ID update equation

$$IDS_{tr_{new}} = (IDS + (ID \oplus \widehat{K'})) \oplus n_1 \oplus n_2$$

we can work out ID

$$ID = ((IDS_{tr_{new}} \oplus n_1 \oplus n_2) - IDS) \oplus \widehat{K'}$$

- Now, $n_1 \oplus n_2$ can be approximated by $A \oplus B$

$$A = (IDS \oplus K) \oplus n_1$$
$$B = (IDS \vee K) \oplus n_2$$

with probability $\frac{3}{4}$.

Some Manipulations...

- From message $D = (\widehat{K'} \oplus n_2) + n_1$ we have

$$\widehat{K'} = (D - n_1) \oplus n_2$$

- Searching for an approximation to this value we found

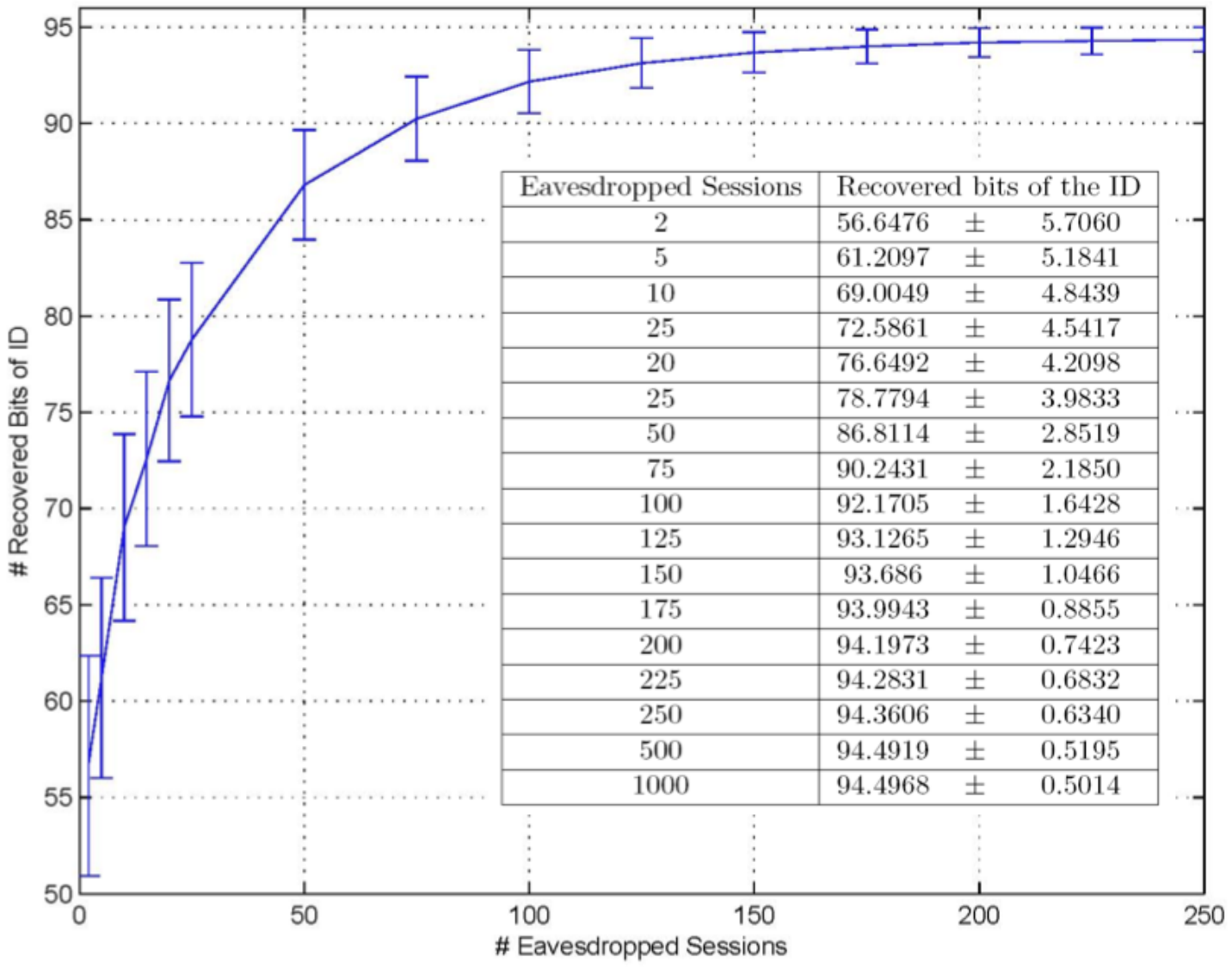
$$\widehat{K'} \approx \overline{D + (A \oplus B)}$$

which has $d_H(\widehat{K'}, \overline{D + (A \oplus B)}) = 40.4185 \pm 5.2096$

- So finally we have

$$ID \approx ((IDS_{tr_{new}} \oplus A \oplus B) - IDS) \oplus \overline{D + (A \oplus B)}$$

YLW Protocol: Recovering the ID



Eavesdropped Sessions	Recovered bits of the ID
2	56.6476 ± 5.7060
5	61.2097 ± 5.1841
10	69.0049 ± 4.8439
25	72.5861 ± 4.5417
20	76.6492 ± 4.2098
25	78.7794 ± 3.9833
50	86.8114 ± 2.8519
75	90.2431 ± 2.1850
100	92.1705 ± 1.6428
125	93.1265 ± 1.2946
150	93.686 ± 1.0466
175	93.9943 ± 0.8855
200	94.1973 ± 0.7423
225	94.2831 ± 0.6832
250	94.3606 ± 0.6340
500	94.4919 ± 0.5195
1000	94.4968 ± 0.5014

Conclusions

- Informal claims about security can be simultaneously very intuitive and very incorrect.
- Dangers of linearity are well understood. In this sort of protocols one should guarantee that there are no significant leakages due to (direct or indirect) linear relations.
- Furthermore, (non)linearity is not the only property that matters. *Any* algorithm that learns something about the secret using only public information will do.
- Techniques used in algebraic cryptanalysis are useful here too.