Sécurité des communications informatiques
TP — Attaque active en chiffré seul contre le mode CTR

2024-02-23/03-08

On considère le contexte pseudo-réaliste suivant : un capteur embarqué communique avec un contrôleur en échangeant des messages $m = m_0 || \cdots || m_7$ possédant la structure suivante :

- $m_0, \ldots, m_6$ sont des chaînes de 8 bits représentant des caractères ASCII 7 bit (ceci implique en particulier que le bit le plus significatif de chacune de ces chaînes vaut toujours zéro).
- m_7 représente la somme modulo 256 de $m_0, \ldots, m_6$ interprétés comme des entiers non signés (entre 0 et 127).

Ces messages sont chiffrés avec un mode CTR instancié avec un chiffre par bloc E qui est supposé être une « bonne PRP ». Les chiffrés (toujours d'un bloc) envoyés ont la forme $c || E(k, c) \oplus m$, où c est un compteur public. On suppose que les compteurs utilisés sont bien choisis, et ne se répètent pas.

Lors de la réception d'un message chiffré, celui-ci est déchiffré comme $m'_0 || \cdots || m'_7$, et il est vérifié que m'_7 est bien égal à la somme modulo 256 de $m'_0, \ldots, m'_6$; si ce n'est pas le cas, un message d'erreur (non chiffré) « `SENDAGN` » est émis, demandant de renvoyer le message.

Préparation : *différences signées*

Soit deux chaînes binaires a, $b \in \{0,1\}^n$, on dit que la paire *ordonnée* (a, b) a une *différence signée* (de un bit) $+2^i$ ($i \in [\![0, n-1]\!]$) si $a \boxplus 2^i = b$ (où $\boxplus$ définit l'addition modulo 2^n de ses opérandes interprétées comme des entiers dans $[\![0, 2^n - 1]\!]$). De façon équivalente, cela veut dire que a et b diffèrent exactement sur leur $i^{\text{ème}}$ bit et que : ce bit vaut zéro dans a ; ou $i = n - 1$. Symétriquement, (a, b) a une différence signée -2^i si $a = b \boxplus 2^i$ (ou $a \boxminus 2^i = 1$, pour l'opérateur de soustraction modulaire $\boxminus$)*.

Q.1 : Soit a, b, $c = a \boxplus b$ des chaînes binaires de 8 bits, et $a' = a \oplus 1$, $b' = b \oplus 1$, $c' = a' \boxplus b'$ (où 1 désigne ici la chaîne de 8 bits dont l'unique bit non nul est le bit le moins significatif (d'index 0)).

1. Montrez que $c = c'$ ssi. les différences signées (a, a') et (b, b') ont des signes différents.
2. Expliquez comment un adversaire en mesure de vérifier si cette égalité tient peut déduire « un bit » d'information sur a et b.
3. Montrez que ceci n'est plus vrai si $a' = a \oplus 2^7$ et $b' = b \oplus 2^7$ (où 2^7 désigne ici la chaîne de 8 bits dont l'unique bit non nul est le bit le plus significatif (d'index 7)).

*. Ces définitions s'étendent naturellement à des différences sur plus d'un bit, mais dans ce cas plusieurs différences distinctes peuvent exister pour une même paire (a, b), car la représentation signée des chaînes binaires est redondante.

Attaque active en chiffré seul

On considère maintenant un adversaire actif pour le contexte décrit ci-dessus : cet adversaire peut lire (et éventuellement bloquer) tout message échangé entre le capteur et le contrôleur, ainsi qu'injecter un message arbitraire.

Q.2 :

1. Développez une attaque qui étant donné un chiffré $\mathrm{c} \| E(\mathrm{k}, \mathrm{c}) \oplus \mathrm{m}$ injecte 7×7 chiffrés, intercepte au plus 7×7 messages d'erreur, et renvoie 2^7 messages candidats possibles pour m.

   INDICE : Utilisez des différences signées et le mécanisme de messages d'erreur pour déterminer (pour chaque j de 0 à 6) la valeur des bits d'indice 1 à 6 de m_j relativement à la valeur du bit d'indice 0 de ce même m_j.

⋆. Expliquez pourquoi cette attaque ne fonctionnerait pas si m_7 était calculé comme le XOR bit à bit des $\mathrm{m}_0, \ldots, \mathrm{m}_6$ (ou plus généralement comme n'importe quelle application linéaire sur $\mathbb{F}_2 \simeq \mathbb{Z}/2\mathbb{Z}$).

Q.3 :

1. Implémentez une preuve de concept de cette attaque, en utilisant comme base le fichier https://membres-ljk.imag.fr/Pierre.Karpman/ctr_active_attack_poc.c (à compiler avec l'option *-march=native*).