# UE INF721, 2011-12
## TD Vulnérabilités

Jean Leneutre

### EXERCICE
### Sécurité du protocole WEP

Les réseaux locaux suivant la norme 802.11 (« Wi-Fi »), ou WLANs, permettent à des terminaux mobiles (ordinateurs portables, assistants personnels, « SmartPhones »...) de communiquer avec d'autres mobiles par onde radio. Un réseau local s'organise autour d'un (ou plusieurs) point(s) d'accès qui dialogue(nt) avec les différents mobiles qui sont dans son (leur) rayon d'émission et de réception[1]. Un point d'accès peut constituer une passerelle IP ou un pont Ethernet pour tous les mobiles qui sont dans son rayon de réception et d'émission.

L'utilisation des ondes radio pour implanter un canal de communication entre une station mobile et un point d'accès facilite certaines attaques : il est facile de brouiller ou de capter ces ondes radio. La connexion au réseau ne nécessitant pas d'accès physique au point d'accès, tout terminal à portée d'un point d'accès peut tenter d'accéder de manière illégitime aux services fournis par celui-ci. La norme 802.11 a initialement introduit le protocole WEP (« *Wired Equivalent Privacy* ») pour essayer de résoudre certains de ces problèmes.

Dans ce qui suit, nous utiliserons les notations suivantes :

- STA identifiant du mobile légitime (STAtion),
- AP identifiant du point d'accès (Access Point),
- $||$ opérateur classique de concaténation de chaînes,
- $\oplus$ loi de groupe XOR (« *Ou Exclusif* »),
- k clef pré-partagée entre STA et AP, de longueur 40 ou 104 bits,
- IV vecteur d'initialisation de longueur 24 bits (Initialization Vector),
- $r_{AP}$ « *challenge* » (valeur pseudo-aléatoire) de 128 octets émis par AP,
- $\{m\}_k$ opération de chiffrement du message m avec la clef k,
- RC4(s) application de l'algorithme de chiffrement par flot RC4 à la graine s (abordé en question 2),
- S séquence pseudo-aléatoire de longueur suffisante (suivant le contexte), générée par RC4(s).
- CRC( M ) code de redondance cyclique de M (abordé en question 3).
- $A \rightarrow B: m$ envoi d'un message m de A vers B.

[1] Ce mode, dont la topologie est étoilée, est appelé « *Mode Infrastructure* », par opposition au « *Mode Ad-Hoc* », dans lequel le protocole opère de façon totalement décentralisée.

1- On suppose que STA possède une clef symétrique k pré-partagée avec AP[2]. Lors de la connexion de STA à un réseau local WiFi avec AP, le protocole d'authentification utilisé peut être abstrait par le protocole suivant :

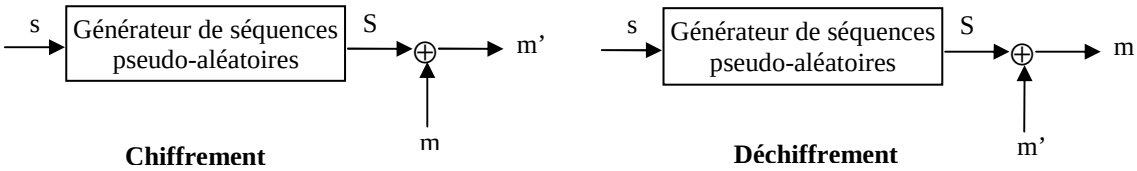
$$STA \rightarrow AP : STA$$
$$AP \rightarrow STA : r_{AP}$$
$$STA \rightarrow AP : \{r_{AP}, AP\}_k$$

Une fois l'authentification réalisée, les messages échangés entre STA et AP sont chiffrés avec k (la clef de chiffrement est la même que celle utilisée pour l'authentification).

Quelle(s) fonction(s) de sécurité ce protocole remplit-il ?
Quels défauts comporte-t'il selon vous ?

2- L'algorithme de chiffrement utilisé est l'algorithme de chiffrement par flot RC4[3]. La plupart des algorithmes de *chiffrement par flot* ou *chiffrement de flux* (« *stream cipher* ») fonctionnent de la manière suivante[4] :

- une séquence pseudo-aléatoire longue, S, (appelée « key stream ») est générée à partir d'une valeur racine secrète (et courte), s, pour chaque chiffrement ;
- le chiffrement d'un message en clair, m, est ensuite réalisé en effectuant un « ou exclusif » (XOR, noté ⊕) entre la séquence fournie par l'algorithme et le message : m'=m⊕ S ;
- le déchiffrement est réalisé en effectuant un « ou exclusif » entre le message chiffré et la séquence S : m'⊕ S =(m⊕ S) ⊕ S = m




Il existe s'autres exemples d'algorithmes de chiffrement par flot : *A5* utilisé dans le GSM, *E0* utilisé dans Bluetooth, … . La spécificité d'un algorithme réside dans la façon de générer une séquence pseudo-aléatoire.

Que peut-il se passer si la séquence pseudo-aléatoire n'est pas renouvelée à chaque envoi de message ?

---

[2] Normalement spécifique à une station, la norme laisse la possibilité que cette clef soit commune à toutes les stations.

[3] Rivest Cipher 4, ou encore *Ron's Code 4*. Ronald L. Rivest est un des auteurs de RSA.

[4] Les algorithmes de chiffrement par flot trouvent clairement leur inspiration dans le chiffre de Vernam et les résultats de la Théorie de l'Information, sans être toutefois en mesure de garantir une confidentialité parfaite, au contraire de la méthode dite du *masque jetable* (« *One Time Pad* »).

3- WEP utilise un vecteur d'initialisation IV, permettant de faire varier la séquence pseudo-aléatoire utilisée pour le chiffrement. Celui-ci est concaténé à la clef pré-partagée k. La séquence S est alors égale à RC4(k || IV). Afin que le récepteur d'un message puisse le déchiffrer, l'émetteur envoie IV en clair dans le message.

a. Lorsque l'on prend en compte la méthode de chiffrement décrite précédemment, le protocole d'authentification d'une STA par un AP, devient :

$$\begin{aligned} &STA \rightarrow AP : \quad STA \\ &AP \rightarrow STA : \quad r_{AP} \\ &STA \rightarrow AP : \quad IV, \ (r_{AP}, AP) \oplus RC4(k\|IV) \end{aligned}$$

On remarquera que dans le dernier message, STA doit maintenant envoyer le vecteur d'initialisation IV, pour que AP puisse déchiffrer.

Montrez qu'un attaquant X peut se faire passer pour STA auprès de AP (sans connaître pour autant la clef k), en explicitant un scénario d'attaque.

b. Le vecteur IV est choisi initialement au hasard, puis incrémenté à l'envoi de chaque nouveau message. La taille d'IV est de 24 bits (environ 17 millions de valeurs possibles). Quand toutes les valeurs d'IV ont été utilisées, IV est réinitialisé à 0[5].

Nous supposons que chaque message fait une taille de 1000 octets, que le débit moyen du réseau est de 8 Mbit par seconde. En combien de temps une collision se produit ?

Déduisez une attaque sur la confidentialité.

c. En 2001, S. Fluhrer, I. Mantin, et A. Shamir observent que pour des valeurs d'IV dites faibles, la séquence générée par RC4 possède un biais sur ses premiers octets[6]. En d'autres termes, les premiers octets du flux utilisé pour le chiffrement ne sont pas aléatoires et peuvent révélées de l'information sur la clef. Pour un message chiffré donné, il est possible de tester si l'on a utilisé un IV faible. Chaque message chiffré avec un IV faible peut permettre de deviner un octet de la clef. En fait, la détermination d'un octet de la clef à partir d'un seul IV faible est statistique et donne une probabilité de succès de 5%. En cas de succès, on parle de « cas résolu », et on peut réitérer afin de deviner l'octet suivant de la clef. L'attaque consiste donc à capturer un maximum de paquets avec des IVs faibles. Cette attaque est connue sous le nom FMS.

Cette attaque théorique a été mise en œuvre par un étudiant américain, A. Stubblefield, associé à J. Ionnadis et A. Rubin. Tout d'abord ils ont observé

---

[5] En fait il existe plusieurs politiques de génération pour l'IV : IV peut être une valeur aléatoire pour chaque message.

[6] Pour plus d'information, consultez :

*S. Fluhrer*, I. Mantin, and A. Shamir, "Weakness in the Key Scheduling Algorithm of RC4", *Proceedings, Workshop in Selected Areas of Cryptography, 2001*.

que quelque soit le protocole utilisé, 802 .11 encapsule tous les paquets avec un entête fixe (le header SNAP de valeur 0xAA).

Suite à cela, des logiciels implémentant l'attaque FMS, comme WEPCRACK ou AIRSNORT, ont rapidement été mis à disposition sur Internet.

Une parade consiste à mettre de côté la première portion de la séquence générée, par exemple les 1024 premiers octets.

Quel est l'avantage de cette dernière attaque sur celle de la question précédente ?

4- En réalité, avant de chiffrer un message on ajoute une protection pour l'intégrité, basée sur le calcul d'un CRC[7]. Un message M est en réalité chiffré de la façon suivante :

*(M || CRC(M)) ⊕ S* (où S est la séquence pseudo-aléatoire calculée par RC4).

a. Montrez qu'un attaquant peut modifier des messages de STA sans que cela soit détecté par AP.

Indice : on s'aidera du fait que la fonction CRC est linéaire par rapport à l'opérateur ⊕, i.e. *CRC(M⊕M') = CRC(M)⊕CRC(M') ).*

b. On suppose que l'AP est relié à internet via une passerelle, et que l'attaquant contrôle un hôte sur le réseau. Par ailleurs, on suppose que l'attaquant capture un paquet IP (chiffré) pour lequel il connaît l'adresse IP de destination. Enfin, on supposera que la passerelle envoie les paquets en clair sur internet.

Proposez alors une attaque permettant à l'attaquant de déchiffrer le paquet précédent en utilisant la question a.

5- Pour chacun des problèmes décelés précédemment, proposez une ou plusieurs améliorations qui permettraient de renforcer la sécurité du protocole WEP.

[7] Désigne un contrôle de redondance cyclique (*Cyclic Redundancy Check*) permettant de détecter les erreurs de transmission par ajout de redondance. Usuellement, le calcul du CRC se base sur une division *modulo 2*.

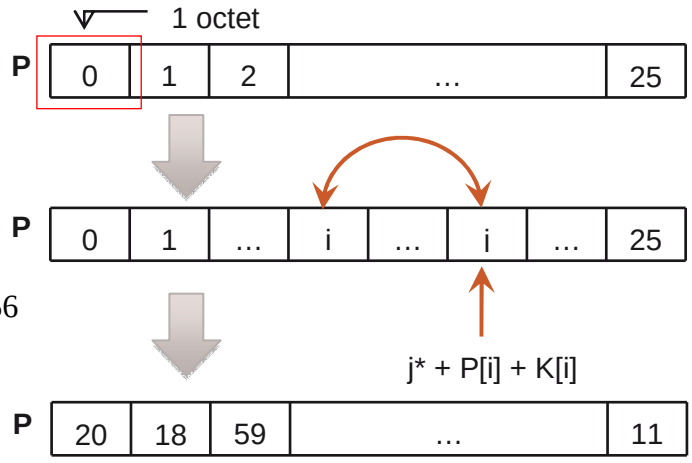
**ANNEXE**
**Description de RC4**

- **Étapes :**
  1. Générer deux tables P et K de taille 256 octets
  2. Initialiser la première table P par les entiers de 0 à 255 (table d'états)
  3. Remplir la deuxième table K avec la clé secrète
  4. Permuter pseudo-aléatoirement la table P en utilisant la clé secrète
  5. Permuter pseudo-aléatoirement la table P avec elle-même
  6. Additionner la séquence ainsi obtenue de la table P avec le flux des données

- **Génération de la permutation :**

pour i de 0 à 255
P[i] := i
finpour

j := 0
pour i de 0 à 255
    j := (j + P[i] + K[i mod |K|]) mod 256
    échanger(P[i], P[j])
finpour




- **Génération de la séquence :**

i := 0
j := 0

tant_que générer une sortie:
    i := (i + 1) mod 256
    j := (j + P[i]) mod 256
    /* permutation de P */
    échanger(P[i], P[j])
    S = P[(P[i] + P[j]) mod 256]
    /* chiffrement du message T */
    C = S ⊕ T
fintant_que