Attaque par clair-connu sur le chiffrement de Hill

Chiffrement de Hill

Le chiffrement de Hill est un algorithme de cryptographie symétrique (c'est à dire qu'il utilise la même clé pour chiffrer et déchiffrer).
L'algorithme permet seulement de chiffrer les caractères présents dans la table de conversion suivante[1] :

0	1	2	3	4	5	6	7	8	9	10	11	12	13	
A	B	C	D	E	F	G	H	I	J	K	L	M	N	
14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
O	P	Q	R	S	T	U	V	W	X	Y	Z		.	=

Les étapes de l'algorithme de chiffrement sont les suivantes:

1. **Transformation de la chaîne d'entrée de longueur ℓ en une liste** $m = (m_0, ..m_{\ell-1})$, correspondant aux **indices des caractères dans la table de conversion** (donc $m_i \in \{0, .., 28\}, i = 0..\ell - 1$)
2. **Transformation de** m **en** v, en regroupant par **paquets de** n **élements** qui seront vu comme des **vecteurs colonnes en dimension** n (note: on s'arrangera pour que n divise ℓ en rajoutant si besoin des '=', du padding, à la fin des messages)
3. Pour chaque vecteur v_i, on **calcul** $c_i \equiv Mv_i \mod 29$, où M est une **matrice** $n \times n$ **inversible modulo 29**, la clé de chiffrement
4. On recolle les c_i entre eux, et on reconvertit le tout en une chaîne de caractères, qui sera notre texte chiffré

Déchiffrement

Il s'agit des même opérations que pour le chiffrement, sauf que l'on utilise M^{-1} à la place de M dans l'algorithme.

Un petit exemple pour la root

Pour que cela soit plus clair, voilà un petit exemple. On choisira $n = 2$, pour chiffrer la chaîne "SECURIMAG". On utilisera la clé de chiffrement $M = \begin{pmatrix} 22 & 14 \\ 21 & 23 \end{pmatrix}$

1. On commence par remarquer que la longueur de la chaîne "SECURIMAG" est 9, qui n'est pas divisible par n. On rajoute donc un '=' de padding derrière, on a alors $\ell = 10$.

[1] Le vrai algorithme utilise simplement l'alphabet et travaille modulo 26, mais pour des raisons de simplicité et de lisibilité, on utilisera un alphabet personnalisé et on travaillera modulo 29 (qui est premier ! *wink* *wink*)

2. On calcule ensuite m, qui vaut $(18, 4, 2, 20, 17, 8, 12, 0, 6, 28)$
3. On regroupe nos indices par paquets de 2, donc $v = (\begin{bmatrix}18\\4\end{bmatrix}, \begin{bmatrix}2\\20\end{bmatrix}, \begin{bmatrix}17\\8\end{bmatrix}, \begin{bmatrix}12\\0\end{bmatrix}, \begin{bmatrix}6\\28\end{bmatrix})$
4. On calcule chaque c_i :
   - $c_0 \equiv Mv_0 \equiv \begin{pmatrix}22 & 14\\21 & 23\end{pmatrix}\begin{bmatrix}18\\4\end{bmatrix} \equiv \begin{bmatrix}17\\6\end{bmatrix} \mod 29$
   - $c_1 \equiv Mv_1 \equiv \begin{pmatrix}22 & 14\\21 & 23\end{pmatrix}\begin{bmatrix}2\\20\end{bmatrix} \equiv \begin{bmatrix}5\\9\end{bmatrix} \mod 29$
   - $c_2 \equiv Mv_2 \equiv \begin{pmatrix}22 & 14\\21 & 23\end{pmatrix}\begin{bmatrix}17\\8\end{bmatrix} \equiv \begin{bmatrix}22\\19\end{bmatrix} \mod 29$
   - $c_3 \equiv Mv_3 \equiv \begin{pmatrix}22 & 14\\21 & 23\end{pmatrix}\begin{bmatrix}12\\0\end{bmatrix} \equiv \begin{bmatrix}3\\20\end{bmatrix} \mod 29$
   - $c_4 \equiv Mv_4 \equiv \begin{pmatrix}22 & 14\\21 & 23\end{pmatrix}\begin{bmatrix}6\\28\end{bmatrix} \equiv \begin{bmatrix}2\\16\end{bmatrix} \mod 29$
5. On obtient alors la liste $c = (17, 6, 5, 9, 22, 19, 3, 20, 2, 16)$, qui une fois convertie en lettre donne le texte chiffré "RGFJWTDUCQ".

Le challenge

Vous arrivez sur le canal IRC principal de Securimag, #securimag. Tout le monde a décidé de chiffrer ses messages avec le chiffrement de Hill, en utilisant pour clé une matrice donnée à l'amphi de rentrée. Pas de bol, vous l'avez manqué, cet amphi.
Vous pouvez néanmoins voir le message de topic du canal, qui n'est pas chiffré, lui : "MERCI DE CHIFFRER VOS MESSAGES COMME CONVENU A L'AMPHI DE RENTREE, AINSI QUE D'INCLURE VOTRE PSEUDO AU DÉBUT DE VOS MESSAGES POUR QUE L'ON SOIT SUR QUE VOUS EN ÊTES L'AUTEUR. TOUT VOS MESSAGES DOIVENT DONC ÊTRE DE LA FORME SUIVANTE UNE FOIS DÉCHIFFRÉ : 'PSEUDO CORPS DU MESSAGE.'"

Ni une, ni deux, vous voyez une faille potentielle qui pourrait vous permettre d'accéder aux messages du canal : l'attaque par clair connu. En effet, vous savez que chaque message chiffré envoyé dans la discussion commence par le pseudo de son auteur suivi d'un espace.
Il ne vous reste plus qu'à utiliser cette information à votre avantage, afin de retrouver la clé de chiffrement et de vous faire passer pour un des leurs ! A vous de jouer !