

CRYPTOGRAPHIE

SMI-S6

A. EL HIBAOUI

Faculté des Sciences de Tétouan – Université Abdelmalek Essaâdi
Département Informatique
2024-2025

aelhibaoui@uae.ac.ma

Cours de Cryptographie

Introduction

Depuis fort longtemps, les hommes ont tenté de rendre sécuritaires leurs communications confidentielles. Différentes techniques ont été utilisées.

Au début, il s'agissait seulement de cacher l'existence du message. Cette technique s'appelle la **stéganographie**.

Puis, des techniques de plus en plus sophistiquées furent utilisées pour rendre les messages compréhensibles seulement par leurs destinataires légitimes.

Tout au cours de l'histoire, une difficile bataille eut lieu entre les constructeurs de code (**cryptographes**) et ceux qui essayaient de les briser (**les cryptanalystes**). Il n'est toujours pas clair, même aujourd'hui, qui sera le vainqueur.

Stéganographie

Stéganographie

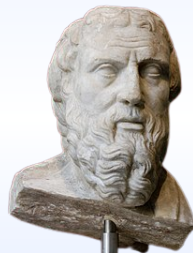
la stéganographie consiste à cacher des informations dans une image, tout comme la cryptographie qui est une manière de dissimuler le contenu d'un message,

Le plus ancien exemple de stéganographie a été rapporté par Hérodote. C'était lors du conflit entre la Grèce et la Perse au 5^{ème} siècle av. J.-C.

Les Perses voulaient conquérir la Grèce et avaient préparé pendant 5 années une imposante armée. Heureusement pour les Grecques, Damaratus, un Grec exilé en Perse eu vent de ce projet.

Il inscrivit son message sur des tablettes de bois et les recouvrit de cire. Les tablettes avaient donc l'air vierges. Elles n'attirèrent pas l'attention des gardes tout au long du parcours.

Les Grecques, une fois mis au courant de l'attaque perse à venir, eurent le temps de se préparer et lors de l'attaque, ils mirent l'armée perse en déroute.



Stéganographie

Hérodote rapporte aussi l'histoire d'Histaïæus qui, pour transmettre un message, rasa la tête de son messenger et inscrivit le message sur son crane. Une fois les cheveux repoussaient, le messenger peut circuler sans attirer l'attention.

Durant la Deuxième Guerre mondiale, les Allemands utilisaient la technique du micropoint. Il s'agit de photographier avec un microfilm le document à transmettre. La taille du microfilm était de moins d'un millimètre de diamètre. On plaçait le micropoint à la place du point final d'une lettre apparemment anodine.

En 1941, le FBI repéra le premier micropoint. De nombreux messages furent par la suite interceptés.

La cryptographie

La cryptographie

La cryptographie est l'étude des méthodes permettant de transmettre des données de manière confidentielle. Afin de protéger un message, on lui applique une transformation qui le rend incompréhensible ; c'est ce qu'on appelle le chiffrement, qui à partir d'un texte en clair donne un texte chiffré ou cryptogramme. Inversement, le déchiffrement est l'action qui permet de reconstruire le texte en clair à partir du texte chiffré.

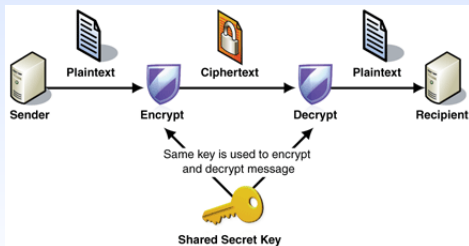
Dans la cryptographie moderne, les transformations en question sont des fonctions mathématiques, appelées algorithmes cryptographiques, qui dépendent d'un paramètre appelé clé. Deux types de chiffrement existent :

- Chiffrement symétrique ou à clé secrète
- Chiffrement asymétrique ou à clé publique

Message en clair

cipher message

Types de chiffrement



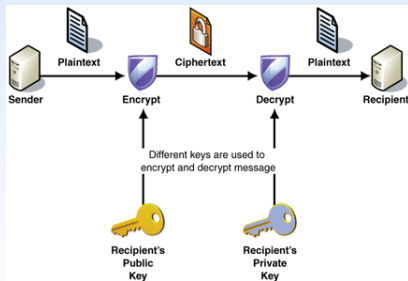
Chiffrement symétrique ou à clé secrète

Les algorithmes de chiffrement symétrique se fondent sur une même clé pour chiffrer et déchiffrer un message. L'émetteur utilise une clé pour chiffrer le message et le destinataire utilise la même clé (le même algorithme mais en sens inverse) pour déchiffrer le message.

Cette technique est rapide et efficace mais son problème est que la clé, qui doit rester totalement confidentielle, doit être transmise au correspondant de façon sûre, en plus de quelques difficultés de gestion des clés (surtout sur des réseaux ouverts).

Quelques algorithmes de chiffrement symétrique très utilisés : DES (Data encryption standard), AES (Advanced encryption standard)

Types de chiffrement



Chiffrement asymétrique ou à clé publique

Un message chiffré avec une clé publique donnée ne peut être déchiffré qu'avec la clé privée correspondante. Par exemple si A souhaite envoyer un message chiffré à B, il le chiffrera en utilisant la clé publique de B (qui peut être publié dans l'annuaire). La seule personne qui déchiffre le message est le détenteur de la clé privée de B. Principaux algorithmes : RSA (Rivest Shamir Adleman), DSA (Digital Signature Algorithm). Le principal inconvénient de RSA et des autres algorithmes à clés publiques est leur grande lenteur par rapport aux algorithmes à clés secrètes.

Types de chiffrement

Chiffrement mixte

Le chiffrement mixte se définit par l'utilisation conjointe d'algorithmes symétriques et asymétriques pour chiffrer des données.

Pourquoi procède-t-on ainsi ?

- Premièrement parce que les algorithmes symétriques sont plus rapides que les algorithmes asymétriques. De plus, dans le processus de chiffrement mixte, l'algorithme asymétrique ne chiffre qu'une clé symétrique... ce qui représente peu de bits.
- Deuxièmement, cette méthode permet de chiffrer un même document pour plusieurs destinataires sans doubler à chaque fois la taille des données chiffrées. Si on ne chiffrait qu'avec les clés asymétriques, il faudrait en effet rechiffrer les données pour chaque nouveau destinataire.

Chiffrement de César

Chiffrement de César



Jules César, 46 av. J.-C.



Cette technique simple de chiffrement effectuant un décalage est appelée chiffrement de César.

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

Exemple

avec un décalage de trois $p = 3$, mon nom devient

EL_HIBAUI = **HOCKLEDRXL**

(On décale aussi les espaces ...)

Cette technique de chiffrement est-elle sécuritaire ?

Chiffrement de César

On intercepte le message

FAGEMYPREMPURZV_EMZR_R_FMNMDAZR

Essayons différents décalages pour le décoder :

$p = 1$

E_FDLXQDLOTQYUZDLYQZQZELMLC_YQ

$p = 2$

DZECKWPCKNSPXTYCKXPYPYDKLKBZXP

$p \in \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

....

$p = 13$

TOUS_LES_CHEMINS_MENENT_A_ROME

Clairement, le chiffrement de César n'est pas sécuritaire.

Chiffrement de César

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

Question

Comment connaître le décalage utilisé sans lister toutes les possibilités ?

Chiffrement de César

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

Question

Comment connaître le décalage utilisé sans lister toutes les possibilités ?

Réponse

Utiliser une analyse fréquentielle. Le caractère le plus fréquent permet d'avoir le décalage.

FAGEMYREMPURZV_EMZR_R_FMNMDAZR

_	A	D	E	F	G	M	N	P	R	Z	U	V	Y
3	2	1	3	2	1	5	1	1	5	3	1	1	1

Le caractère le plus fréquent est *M* ou *R*.

$$M = _ + p \rightarrow p = 13$$

$$R = _ + p \rightarrow p = 18$$

- $p = 13$: TOUS_LES_CHEMINS_MENENT_A_ROME
- $p = 18$: OJPNVG_NVYC_HDINVH_I_IOVWVMJH_

Substitution alphabétique

Substitution mono-alphabétique

Essayons autre chose.

␣	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
R	D	O	H	X	A	M	T	C	␣	B	K	P	E	Z	Q	I	W	N	J	F	L	G	V	Y	U	S

TOUS_LES_CHEMINS_MENENT_A_ROME

devient

FQLJRP AJRHC AE_ZJREAZAFRDRNQE A

Le décodage devrait être plus difficile. Peut-on essayer tous les décodages possibles ?

Substitution mono-alphabétique

Essayons autre chose.

␣	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
R	D	O	H	X	A	M	T	C	␣	B	K	P	E	Z	Q	I	W	N	J	F	L	G	V	Y	U	S

TOUS_LES_CHEMINS_MENENT_A_ROME

devient

FQLJRP AJRHC AE_ZJREAZAZFRDRNQEA

Le décodage devrait être plus difficile. Peut-on essayer tous les décodages possibles ?

Il y a $27! = 10\,888\,869\,450\,418\,352\,160\,768\,000\,000$ possibilités ...

Substitution mono-alphabétique

- La substitution mono-alphabétique apparaît déjà dans le kàma-sùtra qui fut écrit au 5^{ème} siècle mais qui est basé sur des écrits datant du 4^{ème} siècle av. J.-C.
- Le premier usage révélé de chiffrement par substitution dans un usage militaire est rapporté par Jules César dans La guerre des Gaules. César utilisait fréquemment le chiffrement et en particulier le décalage de trois caractères.
- La substitution mono-alphabétique fut la technique de chiffrement la plus utilisée durant le premier millénaire. Nombreux savants de l'antiquité tenaient cette technique pour inviolable.
- Ceux sont les Arabes qui réussirent à briser ce code et qui inventèrent la cryptanalyse au 9^{ème} siècle.

Analyse fréquentielle



Abu Yusuf Ya'qub ibn Ishaq al-Kindi
(801 à Koufa-873 à Bagdad)

Al-Kindi

- Al-Kindi au IX^e siècle fait la plus ancienne description de l'analyse fréquentielle, méthode de cryptanalyse probablement utilisée pour décrypter les documents administratifs et économiques de la dynastie abasside mais aussi pour reconstituer la chronologie des révélations du Coran.
- Il expose les fondements de cette méthode de cryptanalyse dans son traité intitulé Manuscrit sur le déchiffrement des messages cryptographiques. Il montre qu'un message chiffré conserve la trace du message clair original en gardant les fréquences d'apparitions de certaines lettres.

Substitution mono-alphabétique

Exemple. Comment déchiffrer le message ci-dessous ?

BQPSNRSJXJNJXLDPCLDLPQBE_QRKJXHNKPKSJPJIKSPUN
BDKIQRBKPQPBQPZITEJQDQBTSPKELNIUNPHNKPBPCKSS
QWKPSLXJPSNVVXSQCCKDJPBLDWPXBPSPNVVXJPGKPJKDXI
PZLCEJKPGKSPSJQJXSJXHNKSPGPLZZNIIKDZKPGKSPGXV
VKIKDJKSPBKJJIKS

Chaque lettre est chiffrée de la même façon ...
Certaines lettres sont utilisées plus souvent.

Occurrence des lettres

On peut constater que selon la langue, un texte comportera une répartition particulière des fréquences de lettres. Par exemple en français, les lettres les plus fréquentes, c'est-à-dire les lettres que l'on retrouve le plus souvent, sont le E, suivi du A, du I et du S ...

Les fréquences d'apparition des différentes lettres en langue française sont données par le tableau suivant

En Francais

␣	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
19.3	6.7	0.6	2.4	2.9	13.9	0.9	0.8	0.8	6.1	0.3	0	4.7	2.1	5.6	4.1	2.5	1.3	5.3	6.3	6.3	5.2	1.3	0	0.4	0.3	0.1

Dans le cryptogramme

␣	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.5	0	5.1	2.6	4.6	2	0	2.6	1.5	4.6	9.2	12.8	4.1	0	5.6	0	14.3	5.6	1.5	9.2	1	1	3.1	1	5.6	0	2.6

Remplaçons P par _ et K par E

BQ_SNRSJXJNXLD_CLDL_QBE_QREJXHNE_ESJ_JIES_UN
 BDEIQRBE_Q_BQ_ZITEJQDQBTSE_ELNUN_HNE_BE_CESS
 QWE_SLXJ_SNVVXSQCCEDJ_BLDW_XB_SNVVXJ_GE_JEDXI
 _ZLCEJE_GES_SJQJXSJXHNE_S_G_LZZNIIEDZE_GES_GXV
 VEIEDJES_BEJJIES

Remplaçons Q par A et B par L

LA_SNRSJXJNXLD_CLDL_ALE_AREJXHNE_ESJ_JIES_UN
 LDEIARLE_A_LA_ZITEJADALTSE_ELNUN_HNE_LE_CESS
 AWE_SLXJ_SNVVXSACCEDJ_LLDW_XL_SNVVXJ_GE_JEDXI
 _ZLCEJE_GES_SJAJXSJXHNE_S_G_LZZNIIEDZE_GES_GXV
 VEIEDJES_LEJJIES

Remplaçons S par S et G par D

LA_SNRSJXJNXLD_CLDL_ALE_AREJXHNE_ESJ_JIES_UN
 LDEIARLE_A_LA_ZITEJADALTSE_ELNUN_HNE_LE_CESS
 AWE_SLXJ_SNVVXSACCEDJ_LLDW_XL_SNVVXJ_DE_JEDXI
 _ZLCEJE_DES_SJAJXSJXHNE_S_D_LZZNIIEDZE_DES_DXV
 VEIEDJES_LEJJIES

Remplaçons J par T et I par R

LA_SNRSTXTNTXLD_CLDL_ALE_ARETXHNE_EST_TRES_UN
 LDERARLE_A_LA_ZRTETADALTSE_ELNRUN_HNE_LE_CESS
 AWE_SLXT_SNVVXSACCEDT_LLDW_XL_SNVVXT_DE_TEDXR
 _ZLCETE_DES_STATXSTXHNES_D_LZZNRREDZE_DES_DXV
 VEREDTES_LETTRES

Remplaçons X par I, H par Q et N par U

LA_SURSTITUTILD_CLDL_ALE_ARETIQUE_EST_TRES_UU
 LDERARLE_A_LA_ZRTETADALTSE_ELURUU_QUE_LE_CESS
 AWE_SLIT_SUVVISACCEDT_LLDW_IL_SUVVIT_DE_TEDIR
 _ZLCETE_DES_STATISTIQUES_D_LZZURREDZE_DES_DIV
 VEREDTES_LETTRES

Remplaçons V par F et D par N

LA_SURSTITUTILN_CLNL_ALE_ARETIQUE_EST_TRES_UU
 LNERARLE_A_LA_ZRTETANALTSE_ELURUU_QUE_LE_CESS
 AWE_SLIT_SUFFISACCENT_LLNW_IL_SUFFIT_DE_TENIR
 _ZLCETE_DES_STATISTIQUES_D_LZZURRENZE_DES_DIF
 FERENTES_LETTRES

Remplaçons R par B et L par O

LA_SUBSTITUTION_MONO_ALPHABETIQUE_EST_TRES_VULNERABLE_A_LA_CRYPTANALYSE_POURVU_QUE_LE_MESSAGE_SOIT_SUFFISAMMENT_LONG_IL_SUFFIT_DE_TENIR_COMPTES_DES_STATISTIQUES_D_OCCURRENCE_DES_DIFFERENTES_LETTRES

Remplaçons C par M et L par O

LA_SUBSTITUTION_MONO_ALPHABETIQUE_EST_TRES_VULNERABLE_A_LA_CRYPTANALYSE_POURVU_QUE_LE_MESSAGE_SOIT_SUFFISAMMENT_LONG_IL_SUFFIT_DE_TENIR_COMPTES_DES_STATISTIQUES_D_OCCURRENCE_DES_DIFFERENTES_LETTRES

Remplaçons E par P, W par G, _ par H, Z par C, U par V, et T par Y

LA_SUBSTITUTION_MONO_ALPHABETIQUE_EST_TRES_VULNERABLE_A_LA_CRYPTANALYSE_POURVU_QUE_LE_MESSAGE_SOIT_SUFFISAMMENT_LONG_IL_SUFFIT_DE_TENIR_COMPTES_DES_STATISTIQUES_D_OCCURRENCE_DES_DIFFERENTES_LETTRES

Substitution+

- Au lieu de faire la substitution mono-alphabétique, on peut rendre le code plus difficile à briser en faisant une substitution de mots. Chaque mot est remplacé par un nombre, d'où la nécessité d'un dictionnaire. On peut utiliser des synonymes.
- Cette technique n'est pas vraiment pratique. La construction du dictionnaire est fastidieuse. Il faut se déplacer avec le dictionnaire qui pourrait être intercepté. Il est difficile de changer le code.

Substitution++

Substitution++

- Différentes techniques peuvent être utilisées pour rendre le chiffrement par substitution plus sécuritaire tout en gardant une clef de **taille raisonnable**.
- Premièrement, on peut utiliser des synonymes. Par exemple, la lettre E se retrouve 14% du temps et on pourrait utiliser 14 symboles différents pour représenter E et ainsi de suite pour les autres symboles. On obtient un code de 100 symboles.
- On peut aussi utiliser des blancs (symbole sans signification).
- On peut coder certains mots courants par un seul symbole.
- etc ...

Code de Marie Stuart



Marie Stuart

14 décembre 1542 – 24 juillet 1567

Marie Stuart

- En 1586, Marie Stuart, reine d'Écosse fut jugée en Angleterre.
- Elle était accusée d'avoir comploté pour assassiner la reine Elizabeth.
- Le complot eut lieu durant son emprisonnement en Angleterre mais Marie utilisait le chiffrement lors de ses communications avec ses complices.
- La Reine était réticente à exécuter Marie car elle était sa cousine. Le déchiffrement des lettres rendrait la preuve accablante et ne laisserait aucune chance à Marie.

Code de Marie Stuart


a b c d e f g h i k l m n o p q r s t u x y z
 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 10

Nulles ff. r. u. d.

Dowbleth 6

and for with that if but where as of the from by
 2 3 4 4 4 3 7 11 13 8 10

so not when there this in wich is what say me my wyrt
 ꝥ x tt ꝥ ꝥ x t ꝥ m n m m d

send lre receive bearer I pray you Mte your name myne


Code de Marie Stuart

- Gifford transmettait secrètement les lettres de Marie mais c'était en fait un agent double et il transmettait aussi les lettres au services de renseignement de la Reine qui réussirent à briser le code utilisé par Marie.
- En plus de lire toute sa correspondance et d'apprendre le contenu, ils ont falsifié un message demandant explicitement la liste des personnes impliquées.
- Ils furent tous exécutés, incluant Marie. La preuve était accablante.

Exercice

Décrypter le message de Marie Stuart suivant :

6-1-1-8-5-2-7-0-5-0-8-2-5-4-3-1-1

Chiffre de Vigenère/Le chiffre indéchiffrable

Chiffre de Vigenère/Le chiffre indéchiffrable

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

Code de Blaise de Vigenère

Au 16^{ième} siècle, on brisait les codes de façon routinière. La balle était dans le camp des cryptographes. Vigenère inventa un code simple et subtile. Il s'agit d'une amélioration du chiffre par décalage. On choisit un mot de code par exemple COVID et on l'utilise pour chiffrer.

COVID = 3,15,22,9,4



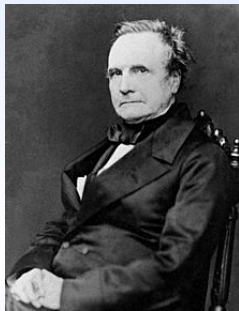
Blaise de Vigenère

COVIDCOVIDCOVIDCOVIDCOVIDCOVIDCOVIDCOVIDC
 ⊕ LE_CODE_DE_VIGENERE_EST_IL_INDECHIFFRABLE
 OTVLSGTVMICJDPIQTMNDHGOIMOODWHHRCRJIFWKPH

Clairement, une attaque statistique simple ne fonctionnera pas. Si le mot de code est suffisamment long (une phrase), essayer toutes les clefs est aussi impossible.

Le chiffre de Vigenère est-il indéchiffrable ?

Chiffre de Vigenère/Le chiffre indéchiffrable



Charles Babbage
1791 – 1871

Déchiffrement du Code de Vigenère

Le nombre de possibilités = $\sum_{i=1}^m (27)^i$

où m est la taille du message.

- Les cryptanalystes furent déjoués pendant près de 3 siècles par le chiffre de Vigenère.
- Au 19^{ième} siècle, Charles Babbage réussit à le briser.
- La technique est relativement simple.

Exemple

OTDHRHSIEGTD_LVISHFIESPVFLHDUOIWEGXJKLRMQHOEEEFMXHFDVXTDQ
 DOWZEGXNWIXNRBDRRSED_TMDQIYLEYJCXPEIIXEEFMXHOTFUOFFEQEL
 HOYSHOJTLGDQDOPTQVYJXFDIHOPFCRPJIOVJWFSZYTIEOTDEGTGTGPKJ
 CTWYCIREMTYEGOISIPRYCTRHRFIEXBIEICMXC_IEPTWXDVIEGCMYCTXWH
 OEXVTCEOCRL

Occurence des lettres

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	3	3	11	14	24	12	9	11	19	7	2	7	7	2	15	8	6	11	7	17	2	7	7	14	9	2
	1.27	1.27	4.66	5.93	10.17	5.08	3.81	4.66	8.05	2.97	0.85	2.97	2.97	0.85	6.36	3.39	2.54	4.66	2.97	7.20	0.85	2.97	2.97	5.93	3.81	0.85

Il est clair qu'une attaque basée sur la fréquence des caractères ne pourra pas réussir.

- Dans ce type de codage, un caractère peut être dans le message peut être codé par plusieurs caractères.
- Connaissant la taille de la clé, on découpera le chiffré en bloc de même taille. Ensuite, une position donnée, on calcule les fréquence des caractères qui se trouvent dans une position donnée.

OTDH RSIE GTD_ LVIS HFIE SPVF LH DU OIWE GXJK LRMQ HOEE EFMX HFDV
 XTDQ DOWZ EGXN WIXN RBDR RBSE D_TM DQIY LEYJ CXPE IIXE EFMX HOTF
 UOFF EQEL HOYS HOJT LGDQ DOPT QVYJ XFDI HOPF CRPJ IOVJ WFSZ YTIE
 OTDE GTGT GPKJ CTWY CIRE MTYE GOIS IPRY CTRH RFIE XBIE ICMX C_JE
 PTWX DVIE GCMY CTXW HOEX VTCE OCRL

- Les caractères à la position 0 des blocs sont :
 ORGLHSLOGLHEHXDEWRRDDLCIEHUEHHLDDQXHC IWYOGGCCMGICRXICPDGCHVO
- Les caractères à la position 1 des blocs sont :
 TSTVFP HIXROFFTOGIBB_QEXIFOOQOOGOVFOROFTTTPTITOPTFBC_TVCTOTC
- Les caractères à la position 2 des blocs sont :
 DIDIIVDWJMEMDDWXXDSTIYPXMTFEYJDPYDPPVSIDGKWRYIRRIIMI WIMXECR
- Les caractères à la position 3 des blocs sont :
 HE_SEFUEKQEXVQZNNREMYJEEXFFLSTQTJIFJJZEETJYEESYHEEXEXEYWXEL

Occurrence des caractères

ORGLHSLOGLHEHXDEWRRDDLCIEHUEHHLDDQXHCIWYOGGCCMGICRXICPDGCHVO

C	D	E	G	H	I	L	M	O	P	Q	R	S	U	V	W	X	Y
7	5	4	6	8	4	5	1	4	1	1	4	1	1	1	2	3	1
11.86	8.47	6.78	10.17	13.56	6.78	8.47	1.69	6.78	1.69	1.69	6.78	1.69	1.69	1.69	3.39	5.08	1.69

Le caractère le plus fréquent ici est H. En considérant que les caractères apparaissant le plus souvent sont soit $_$ ou E, on peut essayer différentes possibilités pour trouver le décalage utilisé.

- $H = _ + p_0$, le décalage $p_0 = 8 \rightarrow H$
- $H = E + p_0$, le décalage $p_0 = 3 \rightarrow C$

TSTVFPHIXROFFTOGIBB_QEXIFOOQOOGOVFOROFTTTPTITOPTFBC_TVCTOTC

$_$	B	C	E	F	G	H	I	O	P	Q	R	S	T	V	X
2	3	3	1	7	2	1	4	11	3	2	2	1	12	3	2
3.39	5.08	5.08	1.69	11.86	3.39	1.69	6.78	18.64	5.08	3.39	3.39	1.69	20.34	5.08	3.39

Le caractère le plus fréquent ici est T. En considérant que les caractères apparaissant le plus souvent sont soit $_$ ou E, on peut essayer différentes possibilités pour trouver le décalage utilisé.

- $T = _ + p_1$, le décalage $p_1 = 20 \rightarrow T$
- $T = E + p_1$, le décalage $p_1 = 15 \rightarrow O$

DIDIIVDWJMEMDDWXXDSTIYPXMTFEYJDPYDPPVSDGKWRYIRRIIMIWMXECR

C	D	E	F	G	I	J	K	M	P	R	S	T	V	W	X	Y
1	9	3	1	1	10	2	1	5	4	4	2	2	2	4	4	4
1.69	15.25	5.08	1.69	1.69	16.95	3.39	1.69	8.47	6.78	6.78	3.39	3.39	3.39	6.78	6.78	6.78

Le caractère le plus fréquent ici est I. En considérant que les caractères apparaissant le plus souvent sont soit $_$ ou E, on peut essayer différentes possibilités pour trouver le décalage utilisé.

- $I = _ + p_2$, le décalage $p_2 = 9 \rightarrow I$
- $I = E + p_2$, le décalage $p_2 = 4 \rightarrow D$

HE_SEFUEKQEXVQZNNREMYJEEXFFLSTQTJIFJJZEETJYEESYHEEXEXEYWXEL

$_$	E	F	H	I	J	K	L	M	N	Q	R	S	T	U	V	W	X	Y	Z
1	16	4	2	1	5	1	2	1	2	3	1	3	3	1	1	1	5	4	2
1.69	27.12	6.78	3.39	1.69	8.47	1.69	3.39	1.69	3.39	5.08	1.69	5.08	5.08	1.69	1.69	1.69	8.47	6.78	3.39

Le caractère le plus fréquent ici est E. En considérant que les caractères apparaissant le plus souvent sont soit $_$ ou E, on peut essayer différentes possibilités pour trouver le décalage utilisé.

- $E = _ + p_3$, le décalage $p_3 = 5 \rightarrow E$
- $E = E + p_3$, le décalage $p_3 = 0 \rightarrow _$

A partir des différents décalages trouvés $p_0.p_1.p_2.p_3$, on déduit que la clé utilisée pour chiffrer le message est le mot **CODE**. Ainsi, on peut déchiffrer le message.

En utilisant la clé CODE

LE CODE DE VIGENERE PARAÎT PLUS DIFFICILE À BRISER QUE LA SUBSTITUTION MONO ALPHABÉTIQUE IL FUT BRISÉ PAR BABBARGE UNE FOIS LA LONGUEUR DE LA CLÉ RETROUVÉE LE DÉCODAGE EST UN JEU D'ENFANT ENCORE UNE FOIS LE MESSAGE DOIT ÊTRE ASSEZ LONG

Et si on connaît pas la taille de la clé, comment faire alors ?

Une idée, c'est de calculer pour chaque longueur l'indice de coïncidence, proposé par William Friedman en 1920.

$$IC = \sum_{q=_}^{q=Z} \frac{n_q(n_q - 1)}{n(n - 1)}$$

et prendre la longueur dont l'indice est le plus grand^a. Probablement, c'est la bonne taille de clé.

a. Le calcul de IC ne concerne que l'ensemble des premiers caractères des blocs issus de découpage.

Vidéo liée : Vigenère Codage et Décodage 

Déterminer la taille de la clé

Algorithme

Pour tout k de 1 à n faire

- Découper le message crypté initial en blocs de taille k ;
- Construire un texte à partir des premiers caractères des blocs obtenus ;
- Calculer les fréquences des caractères présents dans le texte construit.
- Calculer

$$IC_k = \sum_{q=\sqcup}^{q=Z} \frac{n_q(n_q - 1)}{n(n - 1)}$$

Déterminer la taille de la clé

Exemple de calcul de l'indice de coïncidence pour **une clé de taille 4**.

ORGLHSLOGLHEHXDEWRRDDLCIEHUEHHLDQXHCIWYOGGCCMGICRXICPDGCHVO

C	D	E	G	H	I	L	M	O	P	Q	R	S	U	V	W	X	Y
7	5	4	6	8	4	5	1	4	1	1	4	1	1	1	2	3	1
11.86	8.47	6.78	10.17	13.56	6.78	8.47	1.69	6.78	1.69	1.69	6.78	1.69	1.69	1.69	3.39	5.08	1.69

$$\begin{aligned}
 IC_4 &= \sum_{q=_}^{q=Z} \frac{n_q(n_q - 1)}{n(n - 1)} \\
 &= \frac{(7 * 6 + 5 * 4 + 4 * 3 + 6 * 5 + 8 * 7 + 4 * 3 + 5 * 4 + 1 * 0 + 4 * 3 + 1 * 0 + 1 * 0 + 4 * 3 + 1 * 0 + 1 * 0 + 1 * 0 + 2 * 1 + 3 * 1)}{59(59 - 1)} \\
 &= \frac{112}{1711} \\
 &= 0.0654587960257159555815312682641...
 \end{aligned}$$

On calcule les indices de coïncidences pour différentes tailles de clé (IC_1, IC_2, \dots, IC_n).
L'indice de coïncidence le plus élevé étant celui qui correspond à la longueur cherchée.

Masque jetable

Masque jetable

- Peut-on avoir un cryptosystème ayant une confidentialité absolue et qui soit impossible à briser ?
- Qu'arrive-t-il si on utilise le chiffre de Vigenère avec une clef aussi longue que le message ?
Avec une clef aléatoire, on obtient le masque jetable.
- Pour être inconditionnellement sécuritaire, la clef doit être choisie aléatoirement et être utilisée une seule fois.

Sécurité du masque jetable



Claude Elwood Shannon.

1916–2001

- Si la clef est : 12,17,21,7,1,10,24,23,7,15
ABDELA AZIZ devient MSYLMKYVPN
- Pour toute interprétation du message, il existe une clef la justifiant.
Avec la clef : 11,4,11,2,25,22,20,22,16,14
BONJOUR _ _ _ devient MSYLMKYVPN
- C'est Shannon en 1949^a qui a démontré formellement que le masque jetable est inconditionnellement

Cryptosystème à clef courte

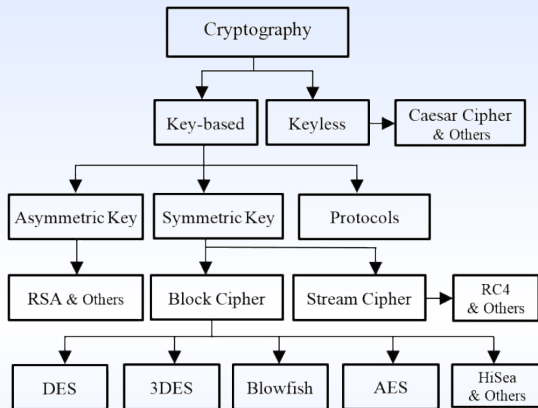


Auguste Kerckhoffs.

1835–1903

Principe de Kerckhoff (La cryptographie militaire 1883)

- La sécurité d'un système de cryptographie **ne doit pas dépendre** de la préservation du secret de l'algorithme.
- La sécurité ne repose que sur le secret de la clef.
- Le masque jetable n'est pas pratique.
- Peut-on chiffrer avec une clef courte de façon sécuritaire ?



Vue d'ensemble des algorithmes de chiffrement cryptographique.

Références

Livres & Polycopiés

- Cours de Cryptographie, Alain Tapp, IFT2105-H2008



Exo7





Neso Academy


- Chiffrement de César
- Chiffrement de Vigenère
- Neso Academy
- Machine Émigre et les clés secrètes
- Cryptographie à clé publique
- Cryptographie à clé publique
- Arithmétique pour RSA
- Chiffrement RSA

Références (suite)

Mathématiques & Informatique

- Vigenère Codage et Décodage 
- AES Rijndael Encryption Cipher Overview 

Autres références

- Public Keys Part 2 - RSA Encryption and Decryptions 
- Crypto Museum, www.cryptomuseum.com
- Free Password Hash Cracker