Quantitative Security of Block Ciphers: Designs and Cryptanalysis Tools

THÈSE N^{O} 4208 (2008)

PRÉSENTÉE LE 14 NOVEMBRE 2008

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS

LABORATOIRE DE SÉCURITÉ ET DE CRYPTOGRAPHIE

PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Thomas BAIGNÈRES

ingénieur en systèmes de communication EPF
de nationalités française et suisse et originaire de Vira (Gambarogno) (TI)

acceptée sur proposition du jury:

Prof. A. Lenstra, président du jury
Prof. S. Vaudenay, directeur de thèse
Dr H. Gilbert, rapporteur
Prof. S. Morgenthaler, rapporteur
Prof. J. Stern, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2008

Contents

Abstract

Block ciphers probably figure in the list of the most important cryptographic primitives. Although they are used for many different purposes, their essential goal is to ensure confidentiality. This thesis is concerned by their *quantitative security*, that is, by *measurable attributes* that reflect their ability to guarantee this confidentiality.

The first part of this thesis deals with well know results. Starting with Shannon's Theory of Secrecy, we move to practical implications for block ciphers, recall the main schemes on which nowadays block ciphers are based, and introduce the Luby-Rackoff security model. We describe distinguishing attacks and key-recovery attacks against block ciphers and show how to turn the firsts into the seconds. As an illustration, we recall linear cryptanalysis which is a classical example of statistical cryptanalysis.

In the second part, we consider the (in)security of block ciphers against statistical cryptanalytic attacks and develop some tools to perform optimal attacks and quantify their efficiency. We start with a simple setting in which the adversary has to distinguish between two sources of randomness and show how an optimal strategy can be derived in certain cases. We proceed with the practical situation where the cardinality of the sample space is too large for the optimal strategy to be implemented and show how this naturally leads to the concept of *projection-based distinguishers*, which reduce the sample space by compressing the samples. Within this setting, we re-consider the particular case of linear distinguishers and generalize them to sets of arbitrary cardinality. We show how these distinguishers between random sources can be turned into distinguishers between random oracles (or block ciphers) and how, in this setting, one can generalize linear cryptanalysis to Abelian groups. As a proof of concept, we show how to break the block cipher TOY100, introduce the block cipher DEAN which encrypts blocks of decimal digits, and apply the theory to the SAFER block cipher family.

In the last part of this thesis, we introduce two new constructions. We start by recalling some essential notions about provable security for block ciphers and about Serge Vaudenay's Decorrelation Theory, and introduce new simple modules for which we prove essential properties that we will later use in our designs. We then present the block cipher C and prove that it is immune against a wide range of cryptanalytic attacks. In particular, we compute the *exact* advantage of the best distinguisher limited to two plaintext/ciphertext samples between C and the perfect cipher and use it to compute the exact value of the maximum expected linear probability (resp. differential probability) of C which is known to be inversely proportional to the number of samples

required by the best possible linear (resp. differential) attack. We then introduce KFC a block cipher which builds upon the same foundations as C but for which we can prove results for higher order adversaries. We conclude both discussions about C and KFC by implementation considerations.

Keywords: Cryptography, block cipher, statistical cryptanalysis, linear cryptanalysis, hypothesis testing, SAFER, Decorrelation Theory

Résumé

Les algorithmes de chiffrement à clef secrète font très certainement partie des primitives cryptographiques les plus importantes. Bien qu'ils soient utilisés à des fins très diverses, leur principale fonction est d'assurer la confidentialité des données. Cette thèse s'intéresse à leur *sécurité quantitative*, c'est-à-dire aux *attributs mesurables* qui reflètent leur habilité à garantir cette confidentialité.

La première partie de cette thèse traite d'un certain nombre de résultats bien connus. En partant de la théorie du secret de Shannon, nous considérons les implications pratiques pour les algorithmes de chiffrement à clef secrète, nous rappelons les schémas élémentaires sur lesquels ces derniers sont conçus, et introduisons le modèle de Luby et Rackoff. Nous décrivons les attaques visant à distinguer une permutation aléatoire d'une autre puis les attaques dont l'objectif est de retrouver la clef secrète pour enfin montrer comment les premières peuvent entraîner les deuxièmes. En guise d'exemple, nous rappelons les concepts de la cryptanalyse linéaire qui est un exemple classique de cryptanalyse statistique.

Dans la deuxième partie, nous considérons l'(in)sécurité des algorithmes de chiffrement à clef secrète face au attaques cryptanalytiques statistiques et développons quelques outils pour exécuter certaines attaques et quantifier leur efficacité. Nous considérons un cadre initial très simple dans lequel un adversaire doit distinguer une source aléatoire d'une autre et montrons que, dans certains cas, une stratégie optimale peut être trouvée. Nous traitons ensuite le cas pratique dans lequel la cardinalité de l'espace échantillon est trop grande pour que la stratégie optimale puisse être utilisée telle quelle, ce qui entraîne naturellement la définition de *distingueurs basés sur des projections* qui réduisent l'espace en compressant chaque échantillon. Dans cette optique, nous reconsidérons le cas des distingueurs linéaires et les généralisons aux ensembles de cardinalité arbitraire. Nous montrons comment ces distingueurs entre des sources aléatoires peuvent être transformés en distingueurs entre des oracles aléatoires et comment, de cette façon, il est possible de généraliser la cryptanalyse linéaire aux groupes Abéliens. En guise de preuve de concept, nous montrons comment casser l'algorithme de chiffrement TOY100, introduisons l'algorithme DEAN qui permet de chiffrer des blocs de chiffres décimaux, et appliquons la théorie à la famille d'algorithmes SAFER.

Dans la dernière partie de cette thèse, nous proposons deux nouvelles constructions. Nous commençons par rappeler quelques notions essentielles concernant la sécurité prouvée des algorithmes de chiffrement à clef secrète et la Théorie de la Décorrélation développée par Serge Vaudenay. Nous introduisons de nouveaux modules

pour lesquels un certain nombre de résultats de sécurité peuvent être prouvés et qui seront au coeur des deux constructions à suivre. Nous présentons ensuite l'algorithme de chiffrement C et prouvons sa sécurité contre une certain nombre d'attaques. En particulier, nous calculons l'avantage *exact* du meilleur distingueur limité à deux paires de textes clairs/chiffrés entre C et l'algorithme de chiffrement parfait et utilisons ce résultat pour calculer la valeur exacte de la valeur moyenne maximum de la probabilité linéaire (ainsi que celle de la valeur moyenne de la probabilité différentielle) de C que l'on sait être inversement proportionnelle au nombre d'échantillons nécessaires pour mener une attaque concluante. Nous introduisons ensuite KFC, un algorithme qui repose sur les mêmes bases que C mais pour lequel nous arrivons à prouver des résultats concernant des adversaires d'ordres plus élevés. Dans les deux cas, nous concluons la discussion par des considérations expérimentales.

Mots-clefs: Cryptographie, algorithme de chiffrement à clef secrète, cryptanalyse statistique, cryptanalyse linéaire, test d'hypothèse, SAFER, Théorie de la Décorrélation