何大可　黄月江　编

密码学进展

——ChinaCrypt'2007

中国密码学会2007年会论文集

西南交通大学出版社
Http://press.swjtu.edu.cn

PROGRESS ON CRYPTOGRAPHY

密码学进展
——ChinaCrypt'2007
中国密码学会2007年会论文集

何大可 黄月江 编

西南交通大学出版社
Http://press.swjtu.edu.cn

内 容 简 介

本书是2007年10月在成都召开的中国密码学会2007年会论文集。书中收录了涉及密码学若干分支的研究论文54篇。主要内容包括：序列密码与分组密码、公钥密码、Hash函数与数字签名、密码协议、量子密码、密码实现与应用等。

本书可供从事密码学、信息安全、通信与信息系统、计算机应用技术等专业的科技人员和高等院校师生参考。

图书在版编目（CIP）数据

密码学进展：中国密码学会2007年会论文集 / 何大可，黄月江编. —成都：西南交通大学出版社，2007.10
ISBN 978-7-81104-742-4

Ⅰ. 密… Ⅱ. ①何…②黄… Ⅲ. 密码－理论－文集
Ⅳ. TN918.1-53

中国版本图书馆CIP数据核字（2007）第152334号

密码学进展——中国密码学会2007年会论文集

何大可　黄月江　编

*

责任编辑　万　方
特邀编辑　于　河
封面设计　本格设计
西南交通大学出版社出版发行
（成都二环路北一段111号　邮政编码：610031　发行部电话：028-87600564）
http: //press.swjtu.edu.cn
E-mail: cbsxx@swjtu.edu.cn
四川森林印务有限责任公司印刷

*

成品尺寸：185 mm × 260 mm　　印张：22.75
字数：655千字
2007年10月第1版　　2007年10月第1次印刷
ISBN 978-7-81104-742-4
定价：88.00元

中国密码学会2007年会程序委员会

主　席：何大可（西南交通大学）

副主席：黄月江（中国电子科技集团公司第三十研究所）

委　员：（按姓氏笔画及汉语拼音排序）

马建峰（西安电子科技大学）
王小云（山东大学）
冯克勤（清华大学）
冯登国（中国科学院软件所）
刘木兰（中国科学院数学与系统科学研究院）
朱　洪（复旦大学）
李　宝（中国科学院研究生院）
李　祥（贵州大学）
杨义先（北京邮电大学）
杨伟成（中国船舶重工集团公司第七二二研究所）
张焕国（武汉大学）
秦志光（电子科技大学）
徐茂智（北京大学）
曹珍富（上海交通大学）
符方伟（南开大学）
彭国华（四川大学）
裴定一（广州大学）

序　言

由中国密码学会主办、西南交通大学承办的中国密码学会 2007 年会（China Crypt'2007）于 2007 年 10 月 19 日至 22 日在中国成都西南交通大学召开。

本次年会共收到投稿论文 116 篇，每篇论文至少由两位专家评审。程序委员会认真讨论了评审结果，并且征询拟录用论文作者本人意见，最后确定录用论文 54 篇，其中 37 篇为全文录用，17 篇为短文录用。

本论文集收录的这 54 篇论文，内容涉及序列密码与分组密码、公钥密码、Hash 函数与数字签名、密码协议、量子密码、密码实现与应用等研究方向。这些论文部分地反映了我国密码学学术界当前的研究动态和学术水平。

本次年会，无意间创造了 3 个月征文、3 个月论文成集的新记录。为此，我们首先要感谢所有向本次年会投稿的作者，感谢他们对本次年会征文的迅速响应，这是对中国密码学会及本次年会最大的支持。其次，要感谢所有参与稿件评审的专家，他们为了从众多的稿件中遴选出最具代表性的论文参加年会交流付出了辛勤的劳动。我们还要感谢西南交通大学信息安全与国家计算网格实验室的老师和研究生们以及西南交通大学出版社，没有他们的帮助，不可能在如此短的时间内完成论文集的稿件处理、编辑校对和印刷出版。

本次年会和论文集的出版得到中国密码学会和学会主管单位的大力支持，在此一并致谢！

中国密码学会 2007 年会程序委员会

2007 年 10 月

目 录

序列密码与分组密码

公钥密码

杂凑函数与数字签名

密码协议

量子密码

密码实现与应用

短 文

附　件

序列密碼與分組密碼

Algebraic Immunity Hierarchy of Boolean Functions*

Ziran Tu Yingpu Deng
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, P.R.China
tuzr007@126.com, dengyp@amss.ac.cn

Abstract: Algebraic immunity of Boolean functions is a very important concept in recently introduced algebraic attacks on stream cipher. For a n-variable Boolean function f, the algebraic immunity $AI_n(f)$ takes values in $\left\{0,1,\ldots,\left\lceil\frac{n}{2}\right\rceil\right\}$. For every k in this range, denote $B_{n,k}$ the set of all n-variable Boolean functions with algebraic immunity k, and we know that $B_{n,k}$ is always non-empty. According to the algebraic immunity, we can form a hierarchy of Boolean functions. Trivially, $|B_{n,0}|=2$ In general, about this integer sequence $|B_{n,k}|, k=1,\ldots,\left\lceil\frac{n}{2}\right\rceil$, very few results are known. In this paper, we show an explicit formula for $|B_{n,1}|$. That is, we obtain an exact formula for the number of Boolean functions with algebraic immunity one. This is the first exact formula for the terms in the above integer sequence. We also give a tight upper bound about non-linearity of Boolean functions with algebraic immunity one.
Key words: Boolean functions; algebraic attack; algebraic immunity; non-linearity; stream cipher

1 Introduction

Boolean functions are very important in stream ciphers, of which there are two models: the combiner model and the filter model. They have been proved to be theoretically equivalent, but the attacks do not work quite similarly on each model. What they have in common is that both the combining function and the filtering function should be balanced, have high algebraic degree, high non-linearity and high correlation immunity.

Recently, a new attack [1] [2] [3] upon stream cipher, the so-called algebraic attack, brings a completely new criterion for the design of secure stream cipher systems, known as algebraic immunity.

A Boolean function on n-variables is a mapping from F_2^n into F_2, which is the finite field with two elements. We denote B_n the set of all n-variable Boolean functions. Any Boolean function f in B_n has a unique representation as multivariate polynomials over F_2, which is called the algebraic normal form (ANF)

* The work was supported by NNSF of China (No. 10501049) and 973 Project (No. 2004CB318000).

$$f(x_1,x_2,\dots,x_n)=\sum_{I\subseteq\{1,\dots,n\}} a_I\prod_{i\in I} x_i$$

where the a_I's are in F_2. The algebraic degree $\deg(f)$ of f equals the maximum degree of those monomials with nonzero coefficients in its algebraic normal form. A Boolean function f is called affine, if $\deg(f)\leqslant 1$. The support of f is defined as $Supp(f)=\{x\in F_2^n : f(x)=1\}$, and the $wt(f)$ is the number of vectors which lies in $Supp(f)$.

Definition 1.1[6] The algebraic immunity $AI_n(f)$ of an n-variable Boolean function f is defined to be the lowest degree of nonzero functions g such that $fg=0$ or $(f+1)g=0$.

It is known that for an arbitrary n-variable Boolean function f, we have $AI_n(f)\leqslant\left\lceil\frac{n}{2}\right\rceil$.

Let $B_{n,k}=\{f\in B_n : AI_n(f)=k\}$ where $k=0,1,\dots,\left\lceil\frac{n}{2}\right\rceil$. From [5], we know that $B_{n,k}$ is always non-empty. Thus we have an integer sequence $|B_{n,k}|, k=0,1,\dots,\left\lceil\frac{n}{2}\right\rceil$. Trivially, $|B_{n,0}|=2$. We are interested in what kinds of Boolean functions in $B_{n,k}$, and their cardinals. If we know this, we can successfully form a hierarchy of Boolean functions according to their algebraic immunities, but unfortunately, for a general k, it seems rather difficult to determine completely the number $|B_{n,k}|$,so far as we know, there is little results about this. For example, the references [7] [4] give some lower bound for $|B_{n,\left\lceil\frac{n}{2}\right\rceil}|$.

In this paper, we have a try to understand more about this problem, we can give an explicit formula to count the number of Boolean functions in $B_{n,1}$, this is the first nontrivial exact formula for the terms in the above integer sequence, and we also give a tight upper bound on non-linearity for those functions.

2 Main Results

In this section, we give our main results and their proofs. Let us start with a simple fact.

Lemma 2.1 Let $f\in B_n$ be a non-constant Boolean function, then $AI_n(f)=1$ if and only if there exists a hyper-plane (i.e. $(n-1)$-dimensional subspace of F_2^n) H in F_2^n such that $Supp(f)\subseteq H$ or $Supp(f)\supseteq H$ or $Supp(f)\subseteq\overline{H}$ or $Supp(f)\supseteq\overline{H}$, where $\overline{H}=F_2^n\setminus H$.

Proof. $AI_n(f)=1$ means there exists a degree-1 function g such that $fg=0$ or $(f+1)g=0$, the support of g is a hyper-plane or its complement, then it's easy to derive the lemma.

Lemma 2.2 We choose m distinct vectors from F_2^n to form a $m\times n$ matrix over F_2 with rank r,denote the total number of this kind of matrices by $f_n(m,r)$,then

$$f_n(m,r)=\begin{cases}0 & r>m\\ f_n(m-1,r)\cdot(2^r-m)+f_n(m-1,r-1)\cdot(2^n-2^{r-1}) & otherwise\end{cases}$$

Proof. Suppose we've already had a matrix composed by $m-1$ distinct non-zero vectors

$\alpha_1,\alpha_2,...,\alpha_{m-1}$ in F_2^n, we need to choose α_m such that $rank\{\alpha_1,\alpha_2,...,\alpha_m\}=r$, there are two cases to be considered: first, if $rank\{\alpha_1,\alpha_2,...,\alpha_{m-1}\}=r$, then we should choose α_m in the subspace spanned by $\alpha_1,\alpha_2,...,\alpha_{m-1}$, there are 2^r-m choices for α_m; second, if $rank\{\alpha_1,\alpha_2,...,\alpha_{m-1}\}=r-1$, we should choose α_m not in the subspace spanned by $\alpha_1,\alpha_2,...,\alpha_{m-1}$, there are 2^n-2^{r-1} possibilities, then we obtain our recursive relation.

When $m=r$, from [8] we know

$$f_n(r,r)=(2^n-1)\cdot(2^n-2)\cdot\ldots\cdot(2^n-2^{r-1})$$

and by Lemma 2.2 we can obtain iteratively all $f_n(m,r)$.

Lemma 2.3 We denote $F_n(m,r)$ the number of possibilities to choose m distinct non-zero vectors from F_2^n whose rank is r, then $F_n(m,r)=f_n(m,r)/m!$

Proof. It is obvious.

Now, we can deduce our formula to count the number of n-variable Boolean functions with algebraic immunity one, this is the following theorem.

Theorem 2.4 We have $|B_{n,1}|=2-2^{n+1}+\sum_{m=1}^{2^n-1}\sum_{r=1}^{n}F_n(m,r)\cdot 2^{r+1}\cdot(2^{2^{n-r}}-1)\cdot(-1)^{m+1}$.

Proof. By Lemma 2.1, we only need to consider the following set

$A=\{X\subseteq F_2^n: X\neq\varnothing, X\neq F_2^n, \exists$ a hyper-plane H such that $X\subseteq H$ or $X\subseteq\overline{H}$ or $X\supseteq H$ or $X\supseteq\overline{H}\}$, and $|A|$ is what we want, because $|A|=|B_{n,1}|$.

Let us give an order on all 2^n-1 non-zero vectors in F_2^n, and let α_i be the i-th vector and H_i be the hyper-plane which is $\{x\in F_2^n: <x,\alpha_i>=0\}$, where $<x,\alpha_i>$ denotes the inner-product of x and α_i, $i=1,2,...,2^n-1$.

We denote $A_i=\{X\subseteq F_2^n: X\neq\varnothing, X\neq F_2^n, X\neq H_i, X\neq\overline{H_i}$ and $X\subseteq H_i$ or $X\subseteq\overline{H_i}$ or $X\supseteq H_i$ or $X\supseteq\overline{H_i}\}$, we have $|A|=|\bigcup_{i=1}^{2^n-1}A_i|+2^{n+1}-2$, in which $2^{n+1}-2$ is the number of non-constant affine functions. By the Inclusion and Exclusion-Principle, then

$$|\bigcup_{i=1}^{2^n-1}A_i|=\sum_i|A_i|-\sum_{i,j}|A_i\cap A_j|+\ldots+(-1)^{m+1}\sum_{i_1,i_2,\ldots,i_m}|\bigcap_{j=1}^{m}A_{i_j}|+\ldots+|\bigcap_{i=1}^{2^n-1}A_i|.$$

We need to compute $|\bigcap_{j=1}^{m}A_{i_j}|$.

If $m=1$, it is easy to compute that $|A_i|=2^2\cdot(2^{2^{n-1}}-2)$. Now suppose $m>1$, we can divide $\bigcap_{j=1}^{m}A_{i_j}$ into two parts

$$\bigcap_{j=1}^{m}A_{i_j}=\bigcup_{S_{i_j}=H_{i_j}orS_{i_j}=\overline{H_{i_j}}}\{X\subseteq F_2^n: X\neq\varnothing, X\subseteq\bigcap_{j=1}^{m}S_{i_j}\}\cup\bigcup_{S_{i_j}=H_{i_j}orS_{i_j}=\overline{H_{i_j}}}\{X\subseteq F_2^n: X\neq F_2^n, X\supseteq\bigcup_{j=1}^{m}S_{i_j}\}$$

Since $\{X\subseteq F_2^n: X\neq\varnothing, X\subseteq\bigcap_{j=1}^{m}S_{i_j}\}$ and $\{X\subseteq F_2^n: X\neq F_2^n, X\supseteq\bigcup_{j=1}^{m}S_{i_j}\}$ are symmetric, these two parts have the same cardinal, so we can only consider the first part. If $rank\{\alpha_{i_1},\alpha_{i_2},\dots,\alpha_{i_m}\}=r$, then $\bigcap_{j=1}^{m} H_{i_j}$ is a $n-r$-dimensional subspace, and then $\bigcap_{j=1}^{m} S_{i_j}$ is either $\varnothing$ or a $n-r$-dimensional flat, and note that the components of the first part are disjoint, in other words, there are 2^r disjoint flats with dimension $n-r$, we get

$$|\bigcup_{S_{ij}=H_{i_j} or S_{i_j}=\overline{H_{i_j}}} \{X \subseteq F_2^n : X \subseteq \bigcap_{j=1}^{m} S_{i_j}\}| = 2^r \cdot (2^{2^{n-r}}-1)$$

then
$$|\bigcap_{j=1}^{m} A_{i_j}| = 2^{r+1} \cdot (2^{2^{n-r}}-1).$$

When we choose randomly m non-zero vectors from F_2^n, its rank may distribute from 1 to $Min\{m,n\}$, by lemma 2.3, there are $F_n(m,r)$ possibilities that the rank of this group of vectors is r. We have

$$\sum_{i_1,i_2,\dots,i_m} |\bigcap_{j=1}^{m} A_{i_j}| = \sum_{r=1}^{n} F_n(m,r)\cdot 2^{r+1}\cdot(2^{2^{n-r}}-1).$$

Finally, $$|A| = 2^{n+1} - 2 + \sum_{m=2}^{2^n-1}\sum_{r=1}^{n} F_n(m,r)\cdot 2^{r+1}\cdot(2^{2^{n-r}}-1)\cdot(-1)^{m+1} + F_n(1,1)\cdot 2^2\cdot(2^{2^{n-1}}-2).$$
$$= 2 - 2^{n+1} + \sum_{m=1}^{2^n-1}\sum_{r=1}^{n} F_n(m,r)\cdot 2^{r+1}\cdot(2^{2^{n-r}}-1)\cdot(-1)^{m+1}$$

This proves our theorem.

Remark: From our formula, we have the following table.

n	$\lvert B_{n,1}\rvert$	$\lvert B_{n,1}\rvert/\lvert B_n\rvert$
1	2	0.5
2	14	0.875
3	198	0.7734375
4	10582	0.161468505859
5	7666550	0.00178500777110457420349121093750
6	1081682871734	0.000000058638145973718101833238591780
7	9370945806264076577334	2.75387346428130707474160629154766355497062e-17

We can see from the above table, that $B_{n,1}$ constitutes only a very small part of B_n, and as n grows up, the proportion of $B_{n,1}$ in B_n approaches 0.

It is well known that for any $\alpha \in F_2^n$, the value

$$W_f(\alpha) = \sum_{x\in F_2^n} (-1)^{f(x)+<x,\alpha>}$$

is called the Walsh coefficient of f at α. The non-linearity of Boolean function f can be expressed via its Walsh coefficients by

$$nl(f)=2^{n-1}-\frac{1}{2}Max_{u\in F_2^n}|W_f(u)|.$$

We also derive a tight upper bound on the non-linearity of Boolean functions with algebraic immunity one.

Theorem 2.5 Let f be in B_n with $AI_n(f)=1$, then $nl(f)\leqslant 2^{n-2}$, and this bound is tight.

Proof. Suppose f and g in B_n, it's easy to verify that

$$2\cdot(-1)^{f\cdot g}=1+(-1)^f+(-1)^g-(-1)^{f+g}.$$

By the definition of Walsh coefficient, we have

$$2\cdot W_{f\cdot g}(\alpha)=W_0(\alpha)+W_f(\alpha)+W_g(\alpha)-W_{f+g}(\alpha).$$

if $f\cdot g=0$, then

$$2^n\cdot\delta_{\alpha,0}+W_{f+g}(\alpha)=W_f(\alpha)+W_g(\alpha)$$

Since $AI_n(f)=1$, we assume $g(x)=<\beta,x>+a_0$, in which β is nonzero in F_2^n and a_0 is in F_2. Let $\alpha=(0,0,...,0)$, we get

$$2^n+(-1)^{a_0}W_f(\beta)=W_f(0).$$

Then $$2^n\leqslant|W_f(0)|+|W_f(\beta)|\leqslant 2\cdot Max_{u\in F_2^n}|W_f(u)|$$

Finally $$nl(f)=2^{n-1}-\frac{1}{2}Max_{u\in F_2^n}|W_f(u)|\leq 2^{n-1}-2^{n-2}=2^{n-2}.$$

Note that the upper bound we obtained above is also tight. For $n=1$, the above bound gives

$nl(f)\leqslant\frac{1}{2}$, that is, $nl(f)=0$. Suppose $n\geqslant 2$. Consider $f(x_1,x_2,...,x_n)=x_1x_2$ in B_n, clearly $AI_n(f)=1$, because $x_1x_2(x_1+x_2)=0$. The Walsh coefficient of f at $(a_1,a_2,...,a_n)\in F_2^n$ is

$$W_f(a_1,a_2,...,a_n)=\sum_{x\in F_2^n}(-1)^{x_1x_2+a_1x_1+a_2x_2+...+a_nx_n}=\sum_{x_1,x_2\in F_2}(-1)^{x_1x_2+a_1x_1+a_2x_2}\prod_{i=3}^{n}\sum_{x_i\in F_2}(-1)^{a_ix_i}$$

If $(a_3,a_4,...,a_n)\neq(0,0,...,0)$, then $W_f(a_1,a_2,...,a_n)=0$. By $W_f(0,0,...,0)=2^{n-1}$, $W_f(0,1,0,...,0)=W_f(1,0,0,...,0)=2^{n-1}$ and $W_f(1,1,0,...,0)=-2^{n-1}$, we get $nl(f)=2^{n-2}$.

3 Conclusion

According to the algebraic immunity, we can form a hierarchy of Boolean functions. It is very difficult to determine the number of Boolean functions with a specified algebraic immunity. In this paper, we obtain the first complete answer to this problem, that is, we give the exact formula for the number of any n-variable Boolean functions with algebraic immunity one, and we also give a tight upper bound of non-linearity for those functions.

References

[1] F. Armknecht, Improving fast algebraic attacks, FSE 2004, Springer LNCS vol. 3017, pp. 65-82

[2] N. T. Courtois, Fast algebraic attacks on stream ciphers with linear feedback, Crypto 2003, Springer LNCS vol. 2729, pp. 176-194

[3] N. T. Courtois, W. Meier, Algebraic attacks on stream ciphers with linear feedback, Eurocrypt 2003, Springer LNCS vol. 2656, pp. 345-359

[4] N. Li, W. F. Qi, Construction and count of Boolean functions of an odd number of variables with maximum algebraic immunity. preprint

[5] M. Lobanov, Tight bound between nonlinearity and algebraic immunity. Cryptology ePrint Archive 2005/441

[6] W. Meier, E. Pasalic, C. Carlet, Algebraic attacks and decomposition of Boolean functions, Eurocrypt 2004, Springer LNCS vol. 3027 pp. 474-491

[7] L. Qu, G. Feng, C. Li, On the Boolean functions with maximum possible algebraic immunity: construction and a lower bound of the count. preprint

[8] Z. Wan, Geometry of Classical Groups over Finite Fields, Second Edition, Science Press, Beijing, 2002

布尔函数的代数免疫度分层

涂自然　邓映蒲

中国科学院数学与系统科学研究院系统科学所　北京　100080　中国

tuzr007@126.com, dengyp@amss.ac.cn

摘　要：布尔函数的代数免疫度是在对流密码的代数攻击中产生的新概念，对任意 n 元布尔函数，其代数免疫度 $AI_n(f)$ 可取值 $\left\{0,1,\ldots,\left\lceil \frac{n}{2} \right\rceil\right\}$，对其中任意 k，记 $B_{n,k}$ 为代数免疫度为 k 的布尔函数全体，我们知道 $B_{n,k}$ 总是非空，根据代数免疫度我们可以对布尔函数进行分层。$|B_{n,0}|=2$ 是平凡的，但一般地，关于整数序列 $|B_{n,k}|, k=1,\ldots,\left\lceil \frac{n}{2} \right\rceil$ 结果不多。本文给出了 $|B_{n,1}|$ 的明确公式，这是关于该序列的第一个精确公式，并且我们得到了一个关于代数免疫度为1的布尔函数的非线性度的紧的上界。

关键词：布尔函数　代数攻击　代数免疫度　非线性度　流密码

Joint Linear Complexity of Multiple Linear Recurring Sequences

Fangwei Fu[1] Harald Niederreiter[2] Ferruh Özbudak[3]
[1]Chern Institute of Mathematics, Nankai University, Tianjin 300071, P.R.China
[2]Department of Mathematics, National University of Singapore, Singapore 117543, Republic of Singapore
[3]Department of Mathematics, Middle East Technical University, Ankara 06531, Turkey
fwfu@nankai.edu.cn, nied@math.nus.edu.sg, ozbudak@metu.edu.tr

Abstract: In this paper, we study the joint linear complexity of multisequences consisting of linear recurring sequences. The expectation and variance of the joint linear complexity of random multisequences consisting of linear recurring sequences are determined. Then we enumerate the multisequences consisting of linear recurring sequences with fixed joint linear complexity. A general formula for the appropriate counting function is derived.
Key words: Linear recurring sequences; joint linear complexity; expectation; variance; counting function

Extended Abstract

The linear complexity of sequences is one of the important security measures for stream cipher systems [1,5,18,22,23]. The linear complexity of a finite or periodic sequence is the length of the shortest linear feedback shift register that can generate it. When a sequence is used in stream ciphers as a keystream, it must have high linear complexity to resist an attack by the Berlekamp-Massey algorithm, since one can use the Berlekamp-Massey algorithm to generate the whole sequence from some initial terms. It is well known that a stream cipher system is completely secure if the keystream is a ``truly random" sequence that is uniformly distributed.

A fundamental research problem in stream ciphers is to determine the expectation and variance of the linear complexity of random sequences that are uniformly distributed. Recently, in the study of vectorized stream cipher systems [4,12] the joint linear complexity of multisequences has been investigated [1,2,5,8-10,16-21,26,27]. The multisequence shift-register synthesis with applications to decoding cyclic codes has been studied in [6,7,24,25].

Rueppel [22,23] determined the expectation and variance of the linear complexity of random finite sequences over the binary field. Gustavson [11] derived the general formula for the counting function of the linear complexity of finite sequences over a finite field. Dai, Imamura, and Yang [2], Feng and Dai [8], Niederreiter [18,19], and Niederreiter and Wang [20,21,26] studied the expectation and variance and counting function of the joint linear complexity of finite multisequences over a finite field.

For periodic sequences, Rueppel [22,23] determined the expectation of the linear complexity of random periodic sequences over the binary field for some special values of the period. Blahut and