Multiple-Chi-square Tests and Their Application on Distinguishing Attacks

Ali Vardasbi

vardasbi@ee.sharif.edu

8th International ISC Conference on Information Security and Cryptology

September 2011

Contents

- Introduction
- Preliminaries
- Multiple-chi-square test
- History of ANF monomial tests
- Our modified ANF monomial test
- Trivium
- Experimental results
- Conclusion

Introduction

- Randomness and statistical tests
- Passive tests *vs*. Active tests
- d-monomial test

Preliminaries

- Algebraic Normal Form (ANF)

$$f(x_1,\ldots,x_n) = \oplus_{u \in F_2^n} a_u x^u \quad \leftrightarrow \quad a_u = \underset{v \le u}{\oplus} f(v)$$

- Distribution of ANF coefficients

$$P(a_u = 1) = \frac{1}{2}$$

Preliminaries

- **Pearson's chi-square test**

$$\chi^2 = \sum_{i=1}^{k} \frac{(y_i - u_i)^2}{u_i} \xrightarrow{dist} \chi^2_{k-1}$$

  - y_i : observed frequency
  - u_i : expected frequency
  - χ^2_k : chi-square distribution with k degrees of freedom

Multiple-chi-square test

- **chi-square test for binomial distribution**

$$\chi^2 = \sum_{i=0}^{1} \frac{(y_i - n/2)^2}{n/2} \xrightarrow{dist} \chi^2_1$$

$$y_0 = n - y_1$$

$$\chi^2 = 2 \cdot \frac{(y_1 - n/2)^2}{n/2} = \frac{(y_1 - n/2)^2}{n/4} \xrightarrow{dist} \chi^2_1$$

Multiple-chi-square test

- **Multiple-chi-square test**

$$\chi^2 = \sum_{i=0}^{k} \frac{(O_i - n/2)^2}{n/4} \xrightarrow{dist} \chi_k^2$$

  - **O_i is the number of ones for the i th bit**
  - **A factor of "two" is missing in the old formula**
    - **The impact of this factor "2" is at least "4" bits of reduction in the complexity (in case of Trivium)**

History of ANF monomial tests

- **Filiol's test [1]**
  - **Only used monomials of low degree (e.g. d<4)**
- **Saarinen's test [2]**
  - **Improved Filiol's test by using higher degree monomials**
- **Test by Englund et al [3]**
  - **Improved previous tests by simultaneously using all the coefficients of ANF → so the size of observed samples increased**

Our observation

Lemma. *For* $v \in N$

$$\lim_{v\to\infty} C_{\chi_v^2}(kv) = \begin{cases} 0 & ; k < 1 \\ 1/2 & ; k = 1 \\ 1 & ; k > 1 \end{cases}$$

According to lemma 1, when $\chi^2 = kv > v$, $C_{\chi_v^2}(\chi^2) \to 1$ for large *v*. So there exists χ_0^2 slightly smaller than χ^2 which satisfies $C_{\chi_v^2}(\chi_0^2) > 1-\alpha$. Therefore from $\chi^2 > \chi_0^2$ one can say, with a level of confidence α, that χ^2 is not a sample of a chi-square random variable with degrees of freedom, hence rejecting the null hypothesis.

Our test 1: ANF monomial dist.

- Devide IV into two separate vectors (X_1 , X_2);
- for *P* different values of X_2
  - Construct $Z_0 = f(X_1)$;
  - Compute ANF coefficients from Z_0;
  - Save all the ANF monomials of "*f*"
- let b_i be the number of ones in the i^{th} monomial, during the above procedure.

continue… of test 1: ANF monomial dist.

Compute the chi-square statistic:

$$\chi^2 = \sum_{i=1}^{M}\left(\frac{(b_i - P/2)^2}{P/2} + \frac{(P - b_i - P/2)^2}{P/2}\right) = 2 \cdot \sum_{i=1}^{M}\frac{(b_i - P/2)^2}{P/2}$$

If $\chi^2 > \chi^2_{M,\alpha}$

return **"cipher"**

else

return **"random"**

Our test 2: MD++

- Devide IV into two separate vectors (X_1 , X_2);
- for *P* different values of X_2
  - Construct $Z_0 = f(X_1)$;
  - Compute ANF coefficients from Z_0;
  - Save the ANF monomials of "*f*" with degrees $n-1$ and n
- let b_i be the number of ones in the i^{th} monomial, during the above procedure.

continue... of test 2: MD++

Compute the chi-square statistic:

$$\chi^2 = \sum_{i=1}^{M}\left(\frac{(b_i - P/2)^2}{P/2} + \frac{(P - b_i - P/2)^2}{P/2}\right) = 2\cdot\sum_{i=1}^{M}\frac{(b_i - P/2)^2}{P/2}$$

If $\chi^2 > \chi^2_{M,\alpha}$

return "cipher"

else

return "random"

Trivium

- *Trivium is one of eSTREAM candidates in Profile 2 (hardware)*
- Was designed in 2005 by C. De canni`ere and B. Preneel.
- Became a part of the portfolio for profile.
- Uses three NFSR with a total length of 288 state bits.
- The initialization phase consists of 1152 rounds in which no outputs are generated.

Our result on Trivium – test 1

	ANF monomial test in [1]		Modified ANF monomial test	
rounds	P	n	P	n
672	2^8	18	2^7	14
704	2^6	23	2^7	19
736	-	-	2^8	25

Inconsistency of previous tests

rounds	n	χ^2_{old}	χ^2_{new}	Degrees of freedom
672	12	2041	4084	4096
704	16	32751	65502	65536
736	20	524202	1048404	1048576

Our result on Trivium – test 2

	MD test		MD++ test	
rounds	P	n	P	n
672	2^5	14	2^5	14
704	2^5	19	2^5	18
736	2^6	26	2^6	25

Conclusion

- **Despite the strength of eSTREAM finalists, the distinguishing attacks based on statistical tests can still be useful on them.**
- **chosen IV attacks have been used to find nonrandom properties in recent stream ciphers.**
  - *d*-monomial test: searches for weakness in the distribution of *d*-monomials.

Conclusion

- **There was mistake in computing chi-square statistics in some previous works. Using our notion of multiple-chi-square test, will prevent future possible mistakes.**

References

- [1] E. Filiol, "A New Statistical Testing for Symmetric Ciphers and Hash Functions," In International Conference on Information, Communications and Signal Processing, volume 2119 of Lecture Notes in Computer Science, pages 21–35. Springer-Verlag, 2001.

- [2] M.J. O. Saarinen, "Chosen-IV Statistical Attacks on eSTREAM Stream Ciphers," eSTREAM, ECRYPT Stream Cipher Project, Report 2006/013, 2006. http://www.ecrypt.eu.org/stream.

References

- [3] H. Englund, T. Johansson, and M. S. Turan, “A Framework for Chosen IV Statistical Analysis of Stream Ciphers,” In K. Srinathan, C. Pandu Rangan, and Moti Yung, editors, INDOCRYPT, volume 4859 of LNCS, pages 268–281. Springer, 2007.