On the Goubin-Courtois Attack on TTM

T. Moh*

Sept 23, 2000 revised Sept 29, 2000

Abstract

In the paper [1] L. Goubin and N.T. Courtois propose an attack on TTM cryptosystem through a misidentification of TTM as a special case of TPM (for definition, the reader is referred to § 1). By attacking the TPM system, they produce an interesting formula of complexities, $q^{\lceil \frac{m}{n} \rceil r} \times m^3$, for the TTM cryptosystem, and then wrongfully assume that $r = 2$ for the "non-linear" cases of the TTM cryptosystem. In the present article, we will show that the assumption of $r = 2$ is an oversimplification by giving a concrete example with $r = 8$. Practically, we always have $r > 2$, if it is "non-linear". We further use the analysis of their attack to conclude that the "Learner's Challenge II (+)", which is one possible implementation of TTM, posted by US Data Security Inc (http://www.usdsi.com/), has a complexity of 2^{115}. This is directly opposite to their general conclusion of a complexity of 2^{52}. A complexity of 2^{70} is "strong" for a cryptosystem and 2^{80} is "very strong" for a cryptosystem. Therefore, their attack is ineffective.

1 Introduction

In [1] L. Goubin and N.T. Courtois introduce TPM as

"- n, u, r integers such that $r \leq n$. We also systematically put $m = n + u - r$.

"- $K = GF(q)$ a finite field.

"We first consider a function $\Psi : K^n \to K^{n+u-r}$ such that $(y_1, \cdots, y_{n+u-r}) = \Psi(x_1, \cdots, x_n)$ is defined by the following system of equations:

$$\begin{cases} y_1 = x_1 + g_1(& x_{n-r+1}, \cdots, x_n) \\ y_2 = x_2 + g_2(x_1; & x_{n-r+1}, \cdots, x_n) \\ y_3 = x_3 + g_3(x_1, x_2; & x_{n-r+1}, \cdots, x_n) \\ \cdots & \\ y_{n-r} = x_{n-r} + g_{n-r}(x_1, \cdots, x_{n-r-1}; x_{n-r+1}, \cdots, x_n) & \\ y_{n-r+1} = g_{n-r+1}(x_1, \cdots, x_n) & \\ \cdots & \\ y_{n-r+u} = g_{n-r+u}(x_1, \cdots, x_n) & \end{cases}$$

with each $g_i (1 \leq i \leq n + u - r)$ being a randomly chosen quadratic polynomial."

The TTM cryptosystem (cf [4], [5]) is given by $\prod_i \Phi_i$ from K^n to K^m where $n \leq m$. For simplicity we consider only four maps $\Phi_1, \Phi_2, \Phi_3, \Phi_4$ with the map $\pi = \Phi_4\Phi_3\Phi_2\Phi_1$ where Φ_1

*Math Department, Purdue University, West Lafayette, Indiana 47907-1395. tel: (765)-494-1930, e-mail ttm@math.purdue.edu

is an affine linear map of K^n to K^n, Φ_2, Φ_3 tame maps and Φ_4 an affine linear map of K^m to K^m. The maps Φ_1, Φ_4 are beside our discussion, we shall look at the composition $\Phi_3\Phi_2$ which can be expressed as

$$\Phi_3\Phi_2 = \begin{cases} y_1 = x_1 + P(y_3, \cdots, y_m) = x_1 + P(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n)) \\ y_2 = x_2 + Q(y_3, \cdots, y_m) = x_2 + Q(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n)) \\ y_3 = x_3 + f_3(x_1, x_2) \\ \cdots \\ y_n = x_n + f_n(x_1, \cdots, x_{n-1}) \\ y_{n+1} = 0 + f_{n+1}(x_1, \cdots, x_n) \\ \cdots \\ y_m = 0 + f_m(x_1, \cdots, x_n) \end{cases}$$

There is a technical paper about TTM crytposystem by B. Lucier [7] using Motorola's AltiVec chips with encoding speed 18 Mbits/second and decoding speed more than 50 Mbits/second..

There is an apparent similarity between the above two cryptosystems which causes confusion to casual eyes. However, they are built on **totally different** principles as it will become clear in the discussions of the next section.

In the past, the cryptanalysis of TTM cryptosystem were provided by [4], [5]. An attempt to solve the general problem of overdefined multivariate cryptography (which includes TTM) was made by the article [2] with its inefficiency presented by the article [6]. The article [1] is a new attempt on TTM which will be proved to be ineffective by the present article.

2 Mistaken identity

In their TPM cryptosystem, the number r is the number of useful equations discarded. If $r = n$, then all useful equations are discarded with only "random" type equations left. The legitimate user is on the same footing as the attacker. No one will dream such a nightmare. If r is not really "small" as $0, 1, 2, 3$, the decrypting process as described in § 2.2 of [1] with searching through K^r for the correct values of $\{x_{n-r+1}, \cdots, x_n\}$ will be painfully slow. In the case that $K = GF(2^8)$, if $r \geq 9$, then there are $2^{8\times 9} = 2^{72}$ possibilities for searching, i.e., it will be physically impossible to find the correct plaintext for the legitimate user. Even if $r = 4, 5, 6, 7, 8$, the legitimate user still has to search through $2^{32}, 2^{40}, 2^{48}, 2^{56}, 2^{64}$ possibilities, the decrypting process will be impractical. Assume that we have a computer with speed 10^{10} operations per second, it will take thirty years to search through 2^{64} possibilities, with the improbable assumption that it takes only one operation to verify one possiblity which includes loading data, performing computations and checking the result to see if it is "meaningful". In other words, it will take at least thirty years to decrypt just one block. On the other hand, for $r = 0, 1, 2, 3$, according to their analysis, the complexity is $q^{\lceil \frac{m}{n} \rceil r} \times m^3 \leq 2^{68}$ if $q = 2^8, m \leq 100$ and $\lceil \frac{m}{n} \rceil = 2$. Therefore, their hypothetic TPM cryptosystem is either too slow (slower than one bit per second for decrypting) or insecure in the cases we have discussed.

In our TTM cryptosystem, the decrypting process is straightforward and can easily reach tens of million bits per second. If one wants to use the apparent similarity of those two cryptosystems as in [1] (i.e., one treat TTM as a subcase of TPM), then the number r in the TPM has to be identified. In their TPM system, it is easy to see the number r is the number

of variables $x_{n-r+1}, \cdots, x_n$. If they insist on treating our TTM cryptosystem as their TPM system, then the number r has to stands for the (genuine) number of variables in the expressions of $P(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n))$ and $Q(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n))$ in the formulation of our TTM cryptosystem in §1.. This number could be sizable. Let us give an example to show that this number may easily be 8.

Example: The kernel construction of an implementation of our TTM cryptosystem is to construct two polynomials P and Q which are based on a $\mathbf{Q}_8$ component in [3]. Every $\mathbf{Q}_8$ component will generate a huge number of public keys. Let us define a $\mathbf{Q}_8$ as follows,

$\mathbf{Q}_8$: Let the field $\mathbf{K}$ be of 2^8 elements. Let

$$
\begin{array}{ll}
q_1(z_1, \cdots, z_{14}) = z_2 z_5 + z_7; & q_2(z_1, \cdots, z_{14}) = z_6 z_7 + z_8; \\
q_3(z_1, \cdots, z_{14}) = z_6 z_5 + z_9; & q_4(z_1, \cdots, z_{14}) = z_4 z_2 + z_{10}; \\
q_5(z_1, \cdots, z_{14}) = z_3 z_5 + z_{11}; & q_6(z_1, \cdots, z_{14}) = z_1 z_3 + z_{12}; \\
q_7(z_1, \cdots, z_{14}) = z_3 z_7 + z_{13}; & q_8(z_1, \cdots, z_{14}) = z_8 z_3 + z_{14}; \\
q_9(z_1, \cdots, z_{14}) = z_2 z_{12} + z_3; & q_{10}(z_1, \cdots, z_{14}) = z_{10} z_1 + z_4; \\
q_{11}(z_1, \cdots, z_{14}) = z_2 z_{11} + z_{13}; & q_{12}(z_1, \cdots, z_{14}) = z_4 z_{13} + z_7; \\
q_{13}(z_1, \cdots, z_{14}) = z_4 z_{11} + z_5; & q_{14}(z_1, \cdots, z_{14}) = z_9 z_2 + z_6; \\
q_{15}(z_1, \cdots, z_{14}) = z_9 z_{10} + z_{14}; & q_{16}(z_1, \cdots, z_{14}) = z_6 z_{10} + z_8; \\
q_{17}(z_1, \cdots, z_{14}) = z_6 z_1 + z_9; & q_{18}(z_1, \cdots, z_{14}) = z_9 z_1; \\
q_{19}(z_1, \cdots, z_{14}) = z_9 z_4 + z_6; & q_{20}(z_1, \cdots, z_{14}) = z_6 z_3 + z_9; \\
q_{21}(z_1, \cdots, z_{14}) = z_9 z_7 + z_{14}; & q_{22}(z_1, \cdots, z_{14}) = z_9 z_{13} + z_6; \\
q_{23}(z_1, \cdots, z_{14}) = z_1 z_8 + z_{14}; & q_{24}(z_1, \cdots, z_{14}) = z_{14} z_2 + z_8; \\
q_{25}(z_1, \cdots, z_{14}) = z_{14} z_1; & q_{26}(z_1, \cdots, z_{14}) = z_{14} z_{10}; \\
q_{27}(z_1, \cdots, z_{14}) = z_{10} z_2; & q_{28}(z_1, \cdots, z_{14}) = z_2 z_3; \\
q_{29}(z_1, \cdots, z_{14}) = z_1 z_{11}; & q_{30}(z_1, \cdots, z_{14}) = z_1 z_7 + z_5; \\
q_{31}(z_1, \cdots, z_{14}) = z_1 z_{13} + z_{11}; & q_{32}(z_1, \cdots, z_{14}) = z_1 z_5; \\
q_{33}(z_1, \cdots, z_{14}) = z_{12} z_{11} + z_{13}; & q_{34}(z_1, \cdots, z_{14}) = z_{12} z_7; \\
q_{35}(z_1, \cdots, z_{14}) = z_{12} z_{13}; & q_{36}(z_1, \cdots, z_{14}) = z_{12} z_5 + z_7; \\
q_{37}(z_1, \cdots, z_{14}) = z_{10} z_8; & q_{38}(z_1, \cdots, z_{14}) = z_{14} z_4 + z_8; \\
q_{39}(z_1, \cdots, z_{14}) = z_8 z_{11}; & q_{40}(z_1, \cdots, z_{14}) = z_{14} z_{11}; \\
q_{41}(z_1, \cdots, z_{14}) = z_8 z_5 + z_{14}; & q_{42}(z_1, \cdots, z_{14}) = z_{14} z_{13} + z_8;
\end{array}
$$

Then the following $\mathbf{Q}_8$ is a minimal generating polynomial of $(z_9 z_8 + z_6 z_{14})$ with degree 8 in q_i,

$$
\begin{aligned}
\mathbf{Q}_8 = \; & (q_{14} q_{23} + q_{17} q_{24})(q_{10} q_9 + q_6 q_4)^2 (q_{11} q_{30} + q_1 q_{31}) + (q_9 q_{27} + q_4 q_{28})^2 \\
& (q_{29} q_{30} + q_{31} q_{32})(q_{18} q_{23} + q_{17} q_{25}) + ((q_{33} q_{34} + q_{35} q_{36})(q_{15} q_{37} + q_{16} q_{26}) \\
& + (q_{19} q_8 + q_{20} q_{38})(q_{13} q_7 + q_{12} q_5)) + (q_{21} q_{39} + q_{40} q_2 + q_{22} q_{41} + q_{42} q_3)
\end{aligned}
$$

■

By a trick of H. Hironaka [3] (and later, Patarin) considering the dimension of all partial derivatives, it is easy to show that the expression $(z_9 z_8 + z_6 z_{14})$ needs four variables for any representation. It will be a simple exercise (cf Appendix, it is perhaps better to read the Appendix and skip the next three sentences) to (1) replace the above z_i by some suitable x_j so that (2) it is permissible to replace the above q_i by y_k (where $k \geq 3$) such that the set $\{y_i\}$ defines a *Tame* transformation of $\{x_j\}$. We simply define $P = \mathbf{Q}_8$. We should use the

same $\mathbf{Q}_8$ to construct the polynomial Q for a different set of variables $\{z'_9, z'_8, z'_6, z'_{14}\}$. Then the number r for this TTM cryptosystem (if to be treated as their TPM system) will be 8. It will be shown in § 3 that this system is secure in their analysis. In view of the preceding discussions on their TPM system, if the above TTM cryptosystem is treated as a subcase of TPM system, then it will be impractical to decrypt, while as a TTM cryptosystem itself, the decrypting speed reaches millions of bits per second. Clearly, the above TTM cryptosystem can not be considered as their TPM cryptosystem.

There are many elegant examples of $\mathbf{Q}_8$ for the implementation of TTM for practical purposes. The above may or may not be one of them.

Since their TPM cryptosystem is either slow or insecure in the cases we discussed, while the above implementation of TTM cryptosystem is both fast and secure, the statement "The cryptosystem TTM, proposed by T.T. Moh at CryTec'99 is in spite of an apparent complexity, shown in 2.4 to be a subcase of TPM" in § 1 of [1] is clearly false. The above misunderstanding is partially due to the unfamiliarities with the huge possibilities of polynomials.

3 The inefficiency of the Goubin-Courtois attack on TTM

In most public key systems, there are "weak keys". Usually, the "weak keys" are set aside for "beginner's challenge" or "learner's challenge". In the case of TTM cryptosystem, the weak key happens if the $\mathbf{Q}_8$ component produces a polynomial of the form x_i^2 which is "linear" for the computational purpose. In those weak key cases, we have "$r = 2$" and "linearity", which are precise the cases analyzed by L.Goubin and N.T.Courtois in [1]. Their work is interesting in the sense that they showed in general the weak keys have complexities 2^{52}. The US Data Security Inc posted two "Learner Challenge I & II" based on the weak keys to stimulate people's interests on TTM cryptosystems. The intention is obvious. The "challenge TTM 2.1" mentioned in [1] is officially called "Learner's Challenge II".

The mistake of L. Goubin and N.T. Courtois is that they first mistakenly believe that TTM is a subcase of their TPM and then they oversimplify TTM cryptosystem by assuming all implementations of TTM cryptosystem are "linear" or "$r = 2$". Note that not only we repeatly point out there are other $\mathbf{Q}_8$ components, but also they know the existence of the "non-linear" $\mathbf{Q}_8$ by mentioning "Even if $\mathbf{Q}_8$ is non-linear, and since $r = 2, \cdots \cdots$" (cf § 7 of [1]).Practically, $r = 2$ happens only if $\mathbf{Q}_8$ produces a polynomial of the form cx_i^2, i.e., $\mathbf{Q}_8$ is "linear". How could they insist that $r = 2$ even if $\mathbf{Q}_8$ is "non-linear"? By analyzing their corresponding TPM system, they wrongfully conclude that "Even if $\mathbf{Q}_8$ is non-linear, and since r=2, it still broken in 2^{52} elementary operations for a 512-bits cryptosystem." and "There is very little hope that a secure triangle system will ever be proposed" (cf § 7 of [1]). How wrong could they be!

The non-weak keys are "non-linear", therefore the "linearity attack" of § 4 of [1] can not be applied. For the "kernel attack" of § 5 of [1], after many pages of arguments, they finally produced an interesting Goubin-Courtois formula of complexity for the TTM cryptosystem $q^{\lceil \frac{m}{n} \rceil r} \times m^3$ where n is the number of variables (the length of plaintext) and m is the number of functions (the length of ciphertext). The implementation of TTM cryptosystem based on the example of the preceding section (cf Appendix), will be non-linear with the number $r = 8$ (please recall that $K = GF(2^8)$ and $q = 2^8$), then the complexity will be $2^{8\times 2\times 8} \times m^3 = 2^{128} \times m^3$, which is very impressive. Let us consider a concrete example. For the "Learner's

Challenge II (+)" posted by US Data Security Inc (http://www.usdsi.com/contests.html), the implementation is "non-linear" with the number $r = 6$, and $n = 44,\ m = 80$, then the complexity for the 352 bits cryptosystem according to their formula is $2^{8\times 2\times 6+19} = 2^{115} > 2^{80}$, which shows that the "Learner's Challenge II (+)" is super strong under the "Goubin-Courtois" attack.

Appendix

We will give an implementation of TTM cryptosystem based on the example of § 2. In this implementation, we will illustrate one trick (among other known (unpublished) tricks) to shorten the lengths n and m.

We shall use the notations of $\mathbf{Q}_8$, q_i of § 2. Let $K = GF(2^8)$, $n \geq 30$, $m = n + 68$. As usual, we have four maps, $\Phi_1, \Phi_2, \Phi_3, \Phi_4$ with the map $\pi = \Phi_4\Phi_3\Phi_2\Phi_1$ where Φ_1 is an affine linear map of K^n to K^n, Φ_2, Φ_3 tame maps and Φ_4 an affine linear map of K^m to K^m. we should look at the composition $\Phi_3\Phi_2$ which can be expressed as

$$\Phi_3\Phi_2 = \begin{cases}
y_1 = x_1 + P(y_3, \cdots, y_m) = x_1 + P(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n)) \\
\quad = x_1 + \mathbf{Q}_8(y_{n-7}, \cdots, y_{n+34}) = x_1 + x_{n-5}x_{n-6} + x_{n-8}x_n \\
y_2 = x_2 + Q(y_3, \cdots, y_m) = x_2 + Q(x_3 + f_3(x_1, x_2), \cdots, f_m(x_1, \cdots, x_n)) \\
\quad = x_2 + \mathbf{Q}_8(y_{n-21}, \cdots, y_{n-14}, y_{n+35}, \cdots, y_{n+68}) = x_2 + x_{n-19}x_{n-20} + x_{n-22}x_{n-14} \\
y_3 = x_3 + f_3(x_1, x_2) \\
\cdots \\
y_{n-22} = x_{n-22} + f_{n-22}(x_1, \cdots, x_{n-23}) \\
y_{n-21} = q_1(x_{n-27}, \cdots, x_{n-14}) = x_{n-21} + x_{n-26}x_{n-23} \\
\cdots \\
y_{n-14} = q_8(x_{n-27}, \cdots, x_{n-14}) = x_{n-14} + x_{n-20}x_{n-25} \\
y_{n-13} = x_{n-13} + f_{n-13}(x_1, \cdots, x_{n-14}) \\
\cdots \\
y_{n-8} = x_{n-8} + f_{n-8}(x_1, \cdots, x_{n-9}) \\
y_{n-7} = q_1(x_{n-13}, \cdots, x_n) = x_{n-7} + x_{n-12}x_{n-9} \\
\cdots \\
y_{n+34} = q_{42}(x_{n-13}, \cdots, x_n) = x_{n-6} + x_nx_{n-1} \\
y_{n+35} = q_9(x_{n-27}, \cdots, x_{n-14}) = x_{n-26} + x_{n-16}x_{n-25} \\
\cdots \\
y_{n+68} = q_{42}(x_{n-27}, \cdots, x_{n-14}) = x_{n-20} + x_{n-14}x_{n-15}
\end{cases}$$

where f_i's are quadratic polynomials such that the vector space dimension of all homogeneous degree 2 parts of the above system is $n + 68$, and 8 is the minimal degree of all polynomials of the above system which generate $x_{n-5}x_{n-6} + x_{n-8}x_n$ and $x_{n-19}x_{n-20} + x_{n-22}x_{n-14}$.

As usual we further require that $\pi(0, \cdots, 0) = (0, \cdots, 0)$. Then π is the public key, while $\{\Phi_1^{-1}, \Phi_2^{-1}, \Phi_3^{-1}, \Phi_4^{-1}\}$ is the private key.

References

[1] Goubin, L. and Courtois, N.T. *Cryptanalysis of the TTM Cryptosystem.* Accepted by Asiacrypt 2000, Dec 2000.

[2] Courtois, N. Shamir, A. Patarin, J. and Klimov, A. *Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations.* Eurocrypt'2000

[3] Hironaka, H. *Resolution of singularities of an algebraic variety over a field of characteristic zero.*Ann. Math Vol 79, 1964.

[4] Moh, T. *A Public Key System with Signature and Master Key Functions.* Communications in Algebra, 27(5), 2207-2222 (1999).

[5] Moh, T. *A Fast Public Key System With Signature And Master Key Functions.* CrypTEC'99 (Proc. International Workshop on Cryptographic Techniques & E-commerce), City University of Hong Kong Press, July, 1999.

[6] Moh,T. *On The Method of "XL" And Its Inefficiency to TTM.* http://www.usdsi.com/ttm.html/

[7] Lucier, B *Cryptograpgy, Finite Fields and AltiVec* http://www.AltiVec.org/articles/