A NEW (t,n) MULTI-SECRET SHARING SCHEME BASED ON LINEAR ALGEBRA

Seyed Hamed Hassani
Department of Electrical Engineering
Department of Mathematical Science
Sharif University of Technology
P.O.Box 11365-9363, Tehran Iran

Mohammad Reza Aref
Department of Electrical Engineering
Sharif University of Technology
P.O.Box 11365-9363, Tehran Iran

Keywords: Threshold scheme, Secret Sharing, Multi-secret sharing, Linear algebra, Cryptography.

Abstract: In this paper, a new multi-secret threshold scheme based on linear algebra and matrices is proposed. Unlike many recently proposed methods, this method lets the use of conventional cryptographic algorithms in sharing multiple secrets. Our scheme is a multi-use scheme, which in some cases, the amount of computations is considerably reduced. Also, in this paper bounds on the maximum number of participants, for a given threshold value, are obtained.

1 INTRODUCTION

Secret sharing has been a subject of study for the last three decades and is a useful tool in modern cryptography. It plays an important role in protecting information from getting lost, stolen, or destroyed and has been more applicable in recent years. Secret sharing schemes either can be identified as threshold schemes or generalized group-oriented cryptosystems. Various approaches have been proposed for the general problem. In 1979, the first (t,n) threshold scheme was proposed by Blakley (Blakley , 1979) and Shamir (Shamir, 1979) independently, where, Blakley's scheme is based on linear projective geometry and Shamir's scheme is based on the Lagrange interpolating polynomial. In a (t,n) threshold scheme, a secret can be shared among n participants and at least t participants are required to reconstruct the secret, while $(t-1)$ or fewer participants can obtain no information about the secret. In a multi-secret sharing scheme, there are multiple secrets to be distributed during a secret sharing process but only one share is kept by each participant and many secrets can be shared without refreshing the share .

As mentioned in (Jackson et al, 1994), multi-secret sharing schemes may be classified into two groups: One time-use and multi-time use schemes. In a one time-use scheme, the secret holder redistributes new shares to each participant once a particular secret is reconstructed. The schemes in (Blakley , 1979), (Shamir, 1979) and (Karnin et al., 1983) are of this type. On the other hand, in a multi-time use scheme, the shadows owned by any participant remain still secret to others, after the reconstruction of multiple secrets. Therefore, there is no need to redistribute new shares to each participant, which is a costly process in both time and resources ((Jackson et al, 1994)). The schemes in (Deng et al., 1995) (Chien et al., 2000) (Bertilsson et al.,, 1992) and (Pang, 2005) are of this type.

A special class of secret sharing schemes are the linear threshold schemes which are based on vector spaces and systematic linear block codes ((Deng et al., 1995) (Chien et al., 2000) (Bertilsson

Hamed Hassani S. and Reza Aref M. (2006).
A NEW (t , n) MULTI-SECRET SHARING SCHEME BASED ON LINEAR ALGEBRA.
In *Proceedings of the International Conference on Security and Cryptography*, pages 443-449
DOI: 10.5220/0002098704430449

et al.,, 1992) (Karnin et al., 1983)).The proposed scheme is a linear threshold scheme in which the secret S is stored in a matrix (or in a linear subspace). This matrix (or linear subspace) has the property of being constructed by special (authorized) set of vectors and of course, a non-special (non-authorized) set would give no information about it. The scheme is applicable in both one-secret and multi-secret sharing systems and is a multi-time use scheme.

One major problem in many of the previously proposed linear threshold schemes such as Deng, et al. scheme (Deng et al., 1995) , Chien, et al. scheme (Chien et al., 2000), is the dependency of secret reconstruction process on the number of participants (n). For example, in Chien's scheme, in order to reconstruct p secrets, solving $n+p-t$ simultaneous linear equations is required. This dependency, especially when n is a large number with respect to t, would make the scheme somehow inefficient. In the proposed scheme, the secret decomposition scheme is totally independent of the number of participants.

In most of the previous schemes the computations are performed in $GF(q^m)$ where, q^m is larger than all the used numbers. In fact, all the used numbers are elements in this field. But in our scheme the computations could be done in a field of any size, hence, reducing the amount of computations.

The paper is organized as follows. In section 2, we shall briefly review Chien's scheme as an example of a recently proposed linear threshold scheme. In section 3, we shall present the proposed scheme and make some security analysis. In section 4, a comparison is given between our scheme and other schemes such as Chien's. Finally, in section 5, conclusions are presented.

2 REVIEW OF CHIEN'S SCHEME

Before presenting Chien's scheme (Chien et al., 2000), we give a definition of a one-way function $f(x,y)$ with two variables x and y. One-way function has been used in Chien's scheme, and will be used in our scheme.

Definition 1 If function $f(x,y)$ denotes a two-variable one-way function that maps any x and y to a bit string $f(x,y)$ of fixed length ((He et al., 1995)), the function has the following properties:

a) Given x and y, it is easy to compute $f(x,y)$.

b) Given x and $f(x,y)$, it is hard to compute y.

c) Given y and $f(x,y)$, it is hard to compute x.

d) Having no knowledge of y it is hard to compute $f(x,y)$ for any x.

e) Given y, it is hard to find two different values x_1 and x_2 such that $f(x_1,y)=f(x_2,y)$.

f) Given pairs of x_i and $f(x_i,y)$, it is hard to compute $f(x',y)$ for $x'\neq x_i$.

The proof of existence, as well as few examples on construction of such one-way functions is given in ((He et al., 1995)). Chien's scheme is as follows:

Step 1 Let $GF(2^m)$ be a large finite field such that all the used numbers are its elements and let g be a primitive element ; Let $G(n+p,2(n+p)-t)$ denote a special type of systematic block codes generator matrix $[P\ \ I]_{(n+p)\times 2(n+p)-t}$ where I is an identity matrix of order $(n+p)\times(n+p)$ and P is a $(n+p)\times(n+p-t)$ matrix $\left[g^{(i-1)(j-1)}\right]$ for $i=1\sim n+p$ and $j=1\sim n+p-t$ ((Lin, 2004)); The secret holder randomly selects $s_1,...,s_n$ as participants' secret shadows.

Step 2 The secret holder randomly selects r and computes $f(r,s_i)$ for each participant.

Step 3 Assuming $P_1,P_2,...,P_p$ are the p secrets to be shared, let $D=[P_1,P_2,...,P_p,f(r,s_1),f(r,s_2),...,f(r,s_n)]$ be the vector of information symbols. The secret holder computes the corresponding code word $V=DG$ as follows:

$$V=[c_1,c_2,...,c_{n+p-t},P_1,P_2,...,P_p,f(r,s_1),f(r,s_2),...,f(r,s_n)]^T \quad (1)$$

where:

$$c_j=\sum_{i=1}^{p} g^{(i-1)(j-1)}P_j+\sum_{i=p+1}^{n+p} g^{(i-1)(j-1)}f(r,s_{i-p}) \quad (2)$$

Step 4 Publish $(r,c_1,c_2,...,c_{n+p-t})$ in an authenticated manner.

The secret reconstruction process is very simple and straightforward. In order to reconstruct the p secrets, at least t participants poll their pseudo shadows $f(r,s_i)$. Thus, the $(n+p-t)$ equations in (2) are obtained with only $(n+p-t)$ unknown symbols. Therefore, the p secrets can be obtained by solving the simultaneous $(n+p-t)$ linear equations in (2). The important point is that the secret shadows s_i will not be revealed even though all of the symbols $f(r,s_i)$ are exposed to the participants. Therefore, redistribution of secret shadows among the participants, after the secret-reconstruction process, is not required. This is due to the properties of the one-way two-variable function $f(r,s)$. The secret holder only has to choose and publish another random integer r. The number of public values in Chien's scheme is $(n+p-t+1)$ according to Step 4.

3 THE PROPOSED SCHEME

3.1 Description of the Basic Idea

Suppose that V is a finite m-dimensional vector space over a field F and E is a t- dimensional subspace of V. Obviously E is spanned by any linearly independent set $A=\{\alpha_1,\alpha_2,...,\alpha_t\}$ of its vectors ($(\alpha_l\in E)$ for $l=1\sim t$). The idea is first to find a special and unique set of linearly independent vectors $T=\{T_1,T_2,...,T_t\}$ of E such that, given any t-linearly independent set $A=\{\alpha_1,\alpha_2,...,\alpha_t\}$ of vectors in E, the set T can easily be obtained from A. We call the set T as the characteristic set of E. For example, suppose that $m=4,t=3$, having any three vectors $\alpha_1,\alpha_2,\alpha_3\in E$ which do not lie in a plane, E is totally known with respect to V. But the idea is to find 3 special vectors $T=\{T_1,T_2,T_3\}$ in which T could easily and uniquely be obtained from any arbitrary and linearly independent set of 3 vectors in E.

By the method of finding the row-equivalent matrix (Hoffman et al., 1971), the characteristic set of any subspace can easily be found.

Lemma 1: Every matrix $Z_{m\times n}$ has a unique row-equivalent matrix $T_{m\times n}$.

Proof: (Hoffman et al., 1971).

Lemma 2: The row space of a matrix is the linear space spanned by its rows as vectors. The row space of a matrix and its row-equivalent matrix are the same. Also, the row spaces of two matrices are the same if and only if their row equivalent matrices are the same. As a result, the row equivalent matrix of all the matrices with the same row spaces is the same.

Proof: (Hoffman et al., 1971).

So, representing the vectors of V in the standard coordinates, the characteristic set of any t-dimensional subspace E can be found as follows:

Step 1 Choose an arbitrary base $(\alpha_1,\alpha_2,...,\alpha_t)$ for E.

Step 2 Generate the matrix $Z_{t\times m}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_t\end{bmatrix}$ in which the i^{th} row of $Z_{t\times m}$ is α_i $(i=1\sim t)$.

Step 3 Compute the row-equivalent matrix $(T_{t\times m})$ of $Z_{t\times m}$.

Step 4 According to Lemma 2, the rows of $T_{t\times m}$ span E and $T_{t\times m}$ is uniquely found. We call $T_{t\times m}$ the characteristic matrix of E. So the collection of rows of the characteristic matrix of E is also its characteristic set.

So, if by some means the secret S is fitted into a matrix $T_{t\times m}$ which is the row equivalent of itself, then by choosing each participant's share as a vector in E ($T_{t\times m}$ is the characteristic matrix of E),with the property that every t shares are linearly independent, we are done.

3.2 Description of our Scheme

First, we describe the algorithm for sharing one secret, and then the multi-secret sharing algorithm is described. We assume that the secret S is a binary data of size $|S|$.Also, the computations are performed in $GF(2^m)$ (there is no limitation on m but of course for security purposes m is chosen large enough).

Step 1 By adding sufficient number of zeros at the end of S, S is divided into t sub-secrets S_i $(i = 1 \sim t)$. The size of each sub-secret is $|S_i| = (\left\lceil \frac{|S|}{t} \right\rceil + 1)$.

Step 2 Sufficient number of zeros are added to the end of each S_i. Then, each S_i is divided into blocks $S_{i,j}$ $(j = 1 \sim c = (\left\lceil \frac{|S_i|}{(m-1)} \right\rceil + 1))$ of length $m-1$. So each S_i $(i = 1 \sim t)$ is a vector as follows:

$$S_i = (S_{i,1}, S_{i,2}, ..., S_{i,c}) \quad (3)$$

Step 3 Let T be a matrix $T_{t\times(t+c)} = [I \; S]$, where I is an identity $t \times t$ matrix and $S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_t \end{bmatrix}$ is a $t \times c$ matrix in which the i^{th} row is S_i. Obviously, T is row-equivalent of itself so the rank of T is t, and T is the characteristic matrix of the space spanned by its rows (E).

Step 4 The share for the i^{th} participant ($i = 1 \sim n$, where n is the number of participants) is a vector $\alpha_i \in E$. The necessary property for the set $A = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ is that any arbitrary collection B of t vectors in A is a linearly independent set and as mentioned in section 3.1, T (therefore E) can be easily computed from B. To generate the α_i s , one method is to generate the matrix $A_{n\times(t+c)} = P \times T$, where $P_{n\times t} = \lfloor g^{(i-1)(j-1)} \rfloor$ $(i = 1 \sim t, j = 1 \sim n)$ (g is a primitive element in $GF(2^m)$) and let α_i be the i^{th} row of A. It is proved in section 3.3 that the α_i s generated in this method have the above mentioned property (when $n \le 2^m - 1$).

Step 5 The secret holder randomly selects $s_1, ..., s_n$ as participants' secret shadows. Also the secret holder randomly selects r and computes $f(r, s_i)$ for each participant.

Step 6 Each participant's share (α_i) is encrypted with $f(r, s_i)$ as the key(for example by those in (Elgamal, 1985),(Rivest et al., 1978)) and each participant's public share (β_i) is generated.

Step 7 Publish $(r, \beta_1, \beta_2, ..., \beta_n)$ in an authenticated manner.

Secret reconstruction: In order to reconstruct the secret, t participants poll their pseudo shadow $f(r, s_i)$ s. Using the pseudo shadows, the respective public shares (β_i) are deciphered and α_i s are generated. By the use of t vectors (α_i) and according to section 3.1 the characteristic matrix T, and as a result, the secret S are generated.

According to step 1 the secret is first divided into t parts. So, our one-secret threshold scheme is already a multi-secret threshold scheme with t secrets. For having the general multi-secret threshold scheme, we first perform steps 1 and 2 for each of the p secrets individually. But in step 3, for generating the i^{th} row of matrix S, the i^{th} sub-secrets of each p secret are put together in a vector which is the i^{th} row of matrix S. The other steps are the

same. The reconstruction of the secrets is also similar. Therefore, when matrix T is found, the matrix S and each of the p secrets are easily reconstructed.

3.3 Bounds on the Maximum Value of n

As mentioned in section 3.2 The public share for the i^{th} participant $(i = 1 \sim n)$ is a vector $\alpha_i \in E$. A necessary property for the set $A = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ is that any arbitrary collection B of t vectors in A is a linearly independent set. Therefore, the problem of finding maximum number of the users is equivalent to finding the maximum possible cardinality of set A. The members of the characteristic set of E form a basis for it. For each set $B = \{z_1, z_2, ..., z_t\}$ of vectors in, E we have

$$z_i = \sum_{j=1}^{t} \gamma_{i,j} T_j \qquad \text{where} \quad \gamma_{i,j} \in GF(2^m) \tag{4}$$

Definition 2 for each $z_i \in E$ the vector $z_i^{relative} = (\gamma_{i,1}, \gamma_{i,2}, ..., \gamma_{i,t})$ is called the relative vector of z_i and for each set B, the Matrix $Z_{t\times t} = \lfloor \gamma_{i,j} \rfloor$ $(i = 1 \sim t, j = 1 \sim t)$ is called the relative matrix of set B.

Obviously B is a linearly independent set if and only if its relative matrix is invertible. Let $Q_{n\times t}$ be the matrix in which the i^{th} row is the relative vector of α_i (each participant's share vector), i.e.

$$Q_{n\times t} = \begin{bmatrix} \alpha_1^{relative} \\ \alpha_2^{relative} \\ .... \\ \alpha_n^{relative} \end{bmatrix}$$

The problem is equivalent to finding the maximum number of rows for $Q_{n\times t}$,or similarly the maximum cardinality of a set A.

Lemma 3: If $f(t)$ is the maximum cardinality of set A over a t dimensional space, then $f(t) \geq 2^m + 1$

Proof. Let Q be a $(2^m + 1) \times t$ matrix defined as $Q = \begin{bmatrix} R \\ P \end{bmatrix}_{(2^m+1)\times t}$ where R is a $2 \times t$ matrix $\begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 0 & ... & 1 \end{bmatrix}_{2\times t}$ and $P = \lfloor g^{(i-1)(j-1)} \rfloor$ $(i = 1 \sim 2^m - 1, j = 1 \sim t)$ where g is a primitive element in $GF(2^m)$. So for $0 \leq i, j \leq 2^m - 1$ and $(i \neq j)$ we have: $g^i \neq g^j$. Therefore, according to the inevitability of Vandermonde matrix, every set of t arbitrary and distinct rows are linearly independent. Let $Q_t = \{q_1, q_2, ..., q_{2^m+1}\}$ be the collection of all rows of Q respectively. So

$f(t) \geq 2^m + 1$

Lemma 4: The intersection of any two non-overlapping hyper planes in a t dimensional space is a $t-2$ dimensional space.

Proof: (Hoffman et al., 1971).

Lemma 5: $f(2) = 2^m + 1$.

Proof. According to Lemma 3, $f(2) \geq 2^m + 1$. Any vector in $(GF(2^m))^2$ (the space of all ordered pairs (a,b) which $a,b \in GF(2^m)$) is a multiple of a vector in the set Q_2, so no new vector can be added to Q_2. Also for any other set of vectors, with the property of linearly independence of any two vectors, there is a one to one correspondence between the set and a subset of Q_2 (relation: being multiple of each other). So the proof is complete.

Theorem 1: $2^m + 1 \leq f(t) \leq 2^m + t - 1$

Proof. Let $\Pi = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ be a maximal set with the properties mentioned above. Let Ω be the linear subspace spanned by $\lambda_1, \lambda_2, ..., \lambda_{t-2}$. Because of linearly independence of these vectors, there are two independent vectors, e.g., θ_1, θ_2 which the spanned plane by θ_1, θ_2 is perpendicular to the Ω .Let Γ be the family of $t-1$ dimensional hyper planes passing Ω. Obviously, each hyper plane in Γ can at most include one of the vectors in Π other than $\lambda_1, \lambda_2, ..., \lambda_{t-2}$. Let Ψ be the family of projection of such vectors on the plane passing through θ_1, θ_2. Trivially, none of the members of Ψ are multiple of each other and also all the members are non-zero (Lemma 4). So there is a one to one correspondence between Γ and Ψ and according to the proof of lemma 5, there is a one to one correspondence between Ψ and a subset of Q_2. So:

$|\Psi| = |\Gamma| \leq f(2)$

Therefore, we have:

$$|\Pi| \le f(2) + t - 2$$

(5)

From (5) and lemma 5:

$$f(t) \le |\Pi| \le 2^m + t - 1$$

(6)

Finally, from Lemma 3 and (6) we get the result.

3.4 Cheater Identification and Security Analysis

It is important for any secret sharing scheme to detect cheating and to identify the cheater. There are numerous works on this issue such as in Hwang et al., 1999), (Tan et al., 1999). Therefore, we will not address this issue here. However, in order to keep the participants' shadows secret after the cheater identification process, there are some points, which should be mentioned: The secret holder should use the pseudo shadows $f(r, s_i)$ instead of the real shadows in generating the public values for cheater identification. At the same time, with no knowledge of the real shadows, the verifier should be able to use the pseudo shadows in order to determine the existence of a cheater. Hence, in the cheater identification process, each participant polls his pseudo shadow. Therefore, the real shadows will not be disclosed by the properties of the one-way two-variable function.

The security of our scheme can be analyzed from the following different views:

a) Having $t' \le t-1$ pseudo shadows, only t' linearly independent vectors from E is found. Since the characteristic set of E is determined by computing the row equivalent matrix T, and also in order to find T ($T_{t\times(t+c)} = [I\ S]$), using t' vectors would give no information about matrix S in which the secrets are stored. Because every new vector from the remaining $t-t'$ vectors has a direct impact on each element of S.

b) Our scheme will not disclose the participant's real shadows s_i even after multiple secret reconstruction. Even though pseudo shadows $f(r, s_i)$ have been exposed among many co-operating participants, the real shadows are well protected by the properties of the one-way two-variable function. Therefore in order to share next p secrets, the secret holder only needs to randomly choose a new integer r without redistributing every participant's secret shadow s_i.

c) Given the public values $(r, \beta_1, \beta_2, ..., \beta_n)$, an adversary has no way of determining α_i's, without having the pseudo shadows $f(r, s_i)$ (as the keys). Furthermore, encryption of less than t of the β_i s, according to part a, shall give no information about the secrets.

4 PERFORMANCE COMPARISON

In this section, we will shortly compare the performance of our scheme with other schemes, especially Chien's scheme.

In Chien's scheme, the secret reconstruction costs solving the simultaneous $(n+p-t)$ linear equations, while in our scheme, the dependence of reconstruction process on n is totally omitted. Therefore, in the cases where n is a relatively large number, our scheme would be more efficient.

In our scheme the computations are performed in $GF(2^m)$ where 2^m could be smaller than the secrets (if their decimal representation is assumed). This would reduce the number of computational bit operations in some cases. For example, according to (Koblitz, 1998) in a finite field $GF(p^m)$, two elements can be divided or multiplied in $O(\ln^2 q)$ ($q = p^m$) bit operations, and one element can be raised to the N^{th} power in $O((\ln N)(\ln^2 q))$ bit operations. So, obviously, when $\ln(q)$ is reduced by the factor k, the operations are reduced by the factor k^2. In our scheme, (for simplicity, suppose one secret threshold scheme) we approximately (in step 2) reduced the binary size of the field order ($\log_2(q)$) by $\frac{|S|}{tm}$ and instead there are $\frac{|S|}{m}$ parallel processes (for each block in each secret). So since the reduction in each multiplication or division is by the factor k^2 (as defined above), a total reduction in the computations is expected.

5 CONCLUSION

In this paper, a new (t, n) threshold multi-secret sharing scheme was proposed. The scheme is mainly based on vector spaces and matrices and one-way functions. In this scheme, the computational complexity is independent of the number of users .Furthermore, the scheme is a multi-use scheme which lets the use of conventional cryptographic methods in the secret sharing problem.

REFERENCES

Blakley G.R.: Safeguarding Cryptographic Keys. AFIPS Conference Proceedings. Vol.48 (1979).pp. 313-317.

Shamir A.: How to Share a Secret. Communications of ACM .Vol.24 (1979).pp. 612-613.

Deng R.H., Gong L., Lazar A. A., and Guo W.: Authenticated Key Distribution and Secure Broadcast Using No Conventional Encryption: A Unified Approch Based on Block Codes. In Proceedings of the IEEE GLOBE Telecommunication Conference (1995). pp. 1193-1197.

Chien H.-Y., Jan J.-K., and Tseng Y.-M.: An Efficient Multi-Secret Sharing Scheme. IEICE Transactions on Fundamentals of Electronic Communications and Computer Sciences. Vol.E83-A (2000). No. 12. pp. 2762-2765.

Bertilsson M. and Ingemarsson I.: A Construction of Practical Secret Sharing Schemes Using Linear Block Codes. In Advances in Cryptology-Auscrypt' 92. Springer-Verlag(1992). pp. 2-21.

Karnin E.D., Greene J. W. and Hellman M. E.: On Secret Sharing Systems. IEEE Transactions on Information Theory, Vol.29 (1983). pp. 35-41.

Hoffman K., Kunze R.: Linear Algebra. Second Edition. Prentice-Hall. Englewood Cliffs.NJ. (1971)

Pang L., Wang Y.: A New (t, n) Multi-Secret Sharing Scheme Based on Shamir's Secret Sharing. Applied Mathematics and Computation, Vol.167, Nr.2, (2005). pp. 840-848.

Jackson W. A., Martin K. M. and O'Keefe C. M.: On Sharing Many Secrets. In Advances in Cryptology-Asiacrypt'94, Springer-Verlag (1994). pp. 42-54.

He J. and Dawson E.: Multi-Secret Sharing Scheme Based on One-Way Function. Electronics Letters, Vol.31. No. 2 (1995) .pp. 93-94.

Elgamal T.: A Public-key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. IEEE Transactions on Information Theory Vol.31 (1985). pp. 469-472.

Rivest R.L, Shamir A., Adleman L.A: A Method for Obtaining Digital Signatures and Public-key cryptosystems. Communications of the ACM, Vol.21.Nr.2 (1978). pp.120-126.

Hwang R. J., Lee W. B., and Chang C. C.: A Concept of Designing Cheater Identification Methods for Secret Sharing. The Journal of Systems and Software. Vol.46(1999). pp.7-11.

Tan K. J., Zhu H. W. and Gu S. J.: Cheater Identification in (t, n) Threshold Scheme. Computer Communications. Vol.22(1999). pp.762-765.

Lin S., Costello D.J.: Error Control Coding. Second Edition. Prentice Hall.Inc.(2004)

Koblitz N.: Algebraic Aspects of Cryptography. Algorithms and Computation in Mathematics.Vol.3 . NY. Springer-Verlag.(1998)