33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Improvement of Fuzzy Vault Scheme for Securing Key Distribution in Body Sensor Network

Cun-Zhang Cao, Chen-Guang He, Shu-Di Bao, *Member, IEEE,* and Ye Li, *Member, IEEE*

***Abstract*—The security of Body Sensor Network (BSN) has become a vital concern, as the massive development of BSN applications in healthcare. A family of biometrics based security methods has been proposed in the last several years, where the bio-information derived from physiological signals is used as entity identifiers (EIs) for multiple security purposes, including node recognition and keying material protection. Among them, a method named as Physiological Signal based Key Agreement (PSKA) was proposed to use frequency-domain information of physiological signals together with Fuzzy Vault scheme to secure key distribution in BSN. In this study, the PSKA scheme was firstly analyzed and evaluated for its practical usage in terms of fuzzy performance, the result of which indicates that the scheme is not as good as claimed. An improved scheme with the deployment of Fuzzy Vault and error correcting coding was then proposed, followed by simulation analysis. The results indicate that the improved scheme is able to improve the performance of Fuzzy Vault and thus the success rate of authentication or key distribution between genuine nodes of a BSN.**

I. Introduction

As there have been more and more medical applications of Body Sensor Network (BSN), the security of BSN has become a vital concern. A family of biometrics based security methods has been proposed in the last several years, where the bio-information derived from physiological signals, such as electrocardiogram (ECG) and photoplethysmogram (PPG) is used to generate entity identifiers (EIs) [1] for multiple security purposes, including node recognition and keying material protection. As one of the biometrics methods, Physiological Signal based Key Agreement (PSKA) proposed in [2] uses the frequency-domain information of physiological signals to generate EIs [3][4], which is used together with Fuzzy Vault scheme [5] to deploy key distribution within BSN.

Because of the fuzzy characteristic of human body, the identifying information based on physiological signals captured by different nodes within the same BSN can not be matched exactly. Thus, in order to achieve the cryptographic key distribution, a fuzzy identification algorithm, called Fuzzy Vault which has been very popular for fingerprints identification [6] and other research fields [7], was adopted in [4]. The Fuzzy Vault Scheme (FVS) was first introduced by Juels and Sudan [5], springing from the Fuzzy Commitment Scheme (FCS) [8] where error-correcting codes were applied for the approximate matching. FVS takes two deficiencies in FCS into account: intolerance of substantial symbol reordering, and security over non-uniform distributions [9]. However, In order to reliably and efficiently deploy the key distribution in BSN, which is depicted in Fig. 1, PSKA and FVS still face some problems. To address the problem of high False Rejection Rate (FRR) in PSKA, Miao *et al* [10] proposed a modified scheme which was claimed to have potential ability for reduction of the error rate, at the cost of decreased security level, which actually is not acceptable.

In this study, the procedures of PSKA and parameters of FVS in terms of feasibility and security will be analyzed in details. Based on the analysis results, an improved scheme with double fuzziness will be proposed to increase the success rate of keying material recovery, while keeping the same level of security compared to the original one.

The remainder of this paper is organized as follows. In Section II, FVS and how it is used in PSKA is introduced. In Section III, the analysis of FVS and the improved scheme are presented in details, followed by the simulational analysis and discussions in Section IV. Finally, conclusions and future work are given in Section V.

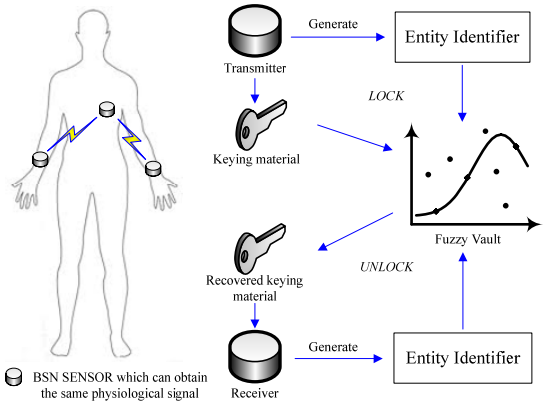


Fig. 1. Simplified Illustration of Biometrics based Key Distribution in BSN

II. Fuzzy Vault Scheme and Its Use in PSKA

The original scheme of Fuzzy Vault includes two algorithms: *LOCK* and *UNLOCK* algorithms [5]. Supposing the secret K should be secured, select a polynomial p in a

This work was supported in part by National Basic Research Program 973 (No. 2010CB732606), Key Basic Research Program of Shenzhen China (No. JC201005270257A), and Guangdong Innovation Research Team Fund for Low-cost Healthcare Technologies.

The authors are with the Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, and Key Lab for Health Informatics, Chinese Academy of Sciences (email: cz.cao@siat.ac.cn, cg.he@siat.ac.cn, sd.bao@siat.ac.cn, ye.li@siat.ac.cn).

Please direct correspondence to Dr. Bao (phone: +86-755-86392200; fax: +86-755-86392299; email: sd.bao@siat.ac.cn).

single variable *x,* and take the information symbols in *K* to be the coefficients of the polynomial.

LOCK algorithm can be depicted as follows: if we aim to lock the secret *K* under a particular set *A*. Treating the elements of *A* as distinct *x*-coordinate value, compute evaluations of *p* on the elements of *A*. And then we can construct a new set $P=\{x, p(x)\}$, where *x* is the element of *A*, and $p(x)$ represents the calculated value of *p* on the element *x*. In order to secure the elements of set *P*, a much larger chaff points set $C=\{c, d\}$ should be randomly chosen, where *c* is not the element of *A*, and *d* is unequal to the calculated value of *p* on the element *c*. Compounding the elements of *P* and *C*, a vault *R* is constructed.

UNLOCK algorithm aims to unlock the vault *R* with the set *B* which should have a substantial number of values in common with set *A*. Firstly, we should look for the matched points both in the set *A* and *B*. And compute the set $Q=\{y, p(y)\}$, where *y* is the matched point, and $p(y)$ represents the calculated value of *p* on the element *y*. If set *A* and *B* are approximately same, the number of matched points will be big enough to reconstruct the polynomial *p* by means of mathematics method and the secret *K* can be unveiled correctly. It can be seen that there are two difficulties in the *UNLOCK* algorithm, one is search of enough matched points, and the other is polynomial reconstruction.

In PSKA [4], FVS was the core for securing keying material transmission, however, the parameters which would be important for the deployment of FVS were neglected in the scheme. Furthermore, there was no detailed information about the reconstruction of polynomial, which actually is a crucial problem to the efficiency of FVS.

III. Analysis of Existing Scheme and the Improved Scheme

The hardness of cracking PSKA is based on the security of FVS. Therefore, how to maintain the security of FVS is the key problem. In this study, a simulation scheme is designed to analyze the system performance in terms of security and feasibility, where many parameters and procedures will be considered, including the length of keying material, the construction of polynomial, the choice of feature points, the addition of chaff points, the algorithm of reconstruction, and the recovery of keying material. The parameters of FVS that will be considered in the simulation study are shown in Table I.

A. Analysis of Parameters

There are some relationships among the aforementioned parameters. For example, the length of keying material is related to the construction of polynomial. As is explained in Section II, the keying material should be separated according to the chosen order of polynomial, and it's no doubt that the two aspects, i.e. keying material and polynomial, are both important to the security of FVS.

Besides, on the process of construction of vault to lock the keying material, another security consideration is the choice of feature and chaff points, i.e. both of the value range and the resolution of vault, which will be also discussed in details. Other parameters will not be analyzed in details because of their less impact on the performance and the space limitation of the paper.

TABLE I
Parameters Considered in the Simulation Study

Symbol	Parameter
key	Keying material
k_l	Bit length of keying material
c_l	Averaged bit length of coefficients
v_{th}	Order of polynomial
N	Number of symbols that has been RS encoded
K	Number of symbols to be RS encoded
q	q parameter of Galois field
f_{max}	Maximum of feature range
f_{min}	Minimum of feature range
$c_{\max}$	Maximum of chaff range
c_{min}	Minimum of chaff range
σ_f	Resolution of feature
σ_c	Resolution of chaff
n_f	Number of feature points
n_c	Number of chaff points

B. Analysis of Procedures

Firstly, the reconstruction algorithm of polynomial is a crucial problem. It should be noted that the recovery of keying material is much up to the reconstruction algorithm. In MATLAB environment, we had attempted several different algorithms including Inverse Matrix approach, Elementary Row Transformation approach and Polynomial Fitting approach. It was found that Polynomial Fitting approach took advantages over the others on the aspects of calculation accuracy and maximum order executable. The success rate of reconstruction, refers to Table II, is above 97% when the order lower than 21; however, more errors occurred when the order of polynomial becomes larger, and the success rate decreases significantly, e.g. 84.7% at the 25th order, and 27.3% at the 31st order. In order to achieve better performance of reconstruction, obviously seeking for a better algorithm of reconstruction could be the first consideration. However, better algorithm in terms of computational accuracy might also mean more cost of calculation. In the resource-limited BSN, it might not be the best plan.

Secondly, In PSKA the requirement of the number of matched points seems to be too strict to be achieved. Thus, this applicative restriction is another key problem. In Section II, the process of unlock algorithm has been introduced in details. The receiver node of BSN must find enough matched feature points based on the EIs generated by itself. For example, providing that v_{th} was the order of polynomial

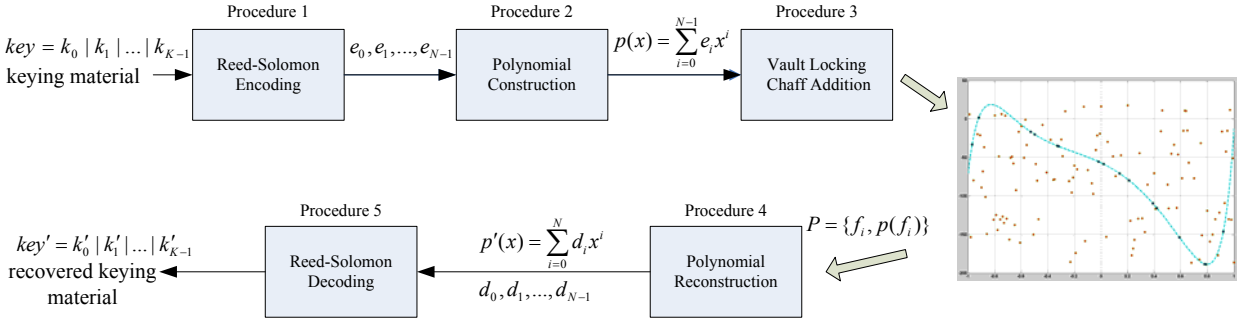


Fig. 2. Diagram of the proposed scheme

chosen by the transmitter, there are at least $v_{th}+1$ matched feature points should be found out to do reconstruction. But in practice, to successfully find out $v_{th}+1$ matched feature points might be unrealistic because that the inherent differences between physiological signals captured by different node even within the same BSN.

At last but not least, with regards of security and feasibility, many factors should be taken into account. For instance, if the keying material *key* is k_l bits, and v_{th} denotes the order of chosen polynomial, an equation can be obtained, i.e.

$$k_l / (v_{th}+1) = c_l \tag{1}$$

where c_l represents the averaged bit length of coefficients of polynomial. Generally speaking, in the actual deployment of BSN the length of keying material k_l is confirmed, therefore, v_{th} and c_l are inversely correlated. However in view of security, v_{th} and c_l are both expected to be large enough to keep resistance to brute force attacks, so it is a trade-off need to be considered.

The setting of value range and resolution of feature or chaff points are also important. Given the bit length of feature point, the value range and resolution is also inversely correlated. Here each feature point is actually a segmentation of the EI. For example, if the bit length of EI is 128, there are a number of 16 feature points, and the value range is set to be [-127, 128], the resolution of feature points will be 1, which means the possible minimum distance between two feature points. If the value range of feature points is [$f_{\min}$, f_{max}], the number of feature points $n_f = (f_{max} - f_{\min})/\sigma_f$, where σ_f denotes the resolution of feature points. Correspondingly, $n_c = (c_{max} - c_{\min})/\sigma_c$ is the number of chaff points. The number n_c should be much larger than n_f so as to meet the security requirement in FVS. To summarize, to efficiently utilize FVS in keying material protection of BSN still faces many challenges in terms of feasibility and security.

A. Improved Scheme with Employment of ECC

Based on the original FVS an improved scheme is proposed in this study, which utilizes the fault-tolerant capability of error-correcting code (ECC). As is demonstrated in Fig. 1, our proposed scheme employs Reed-Solomon (RS) code which is a widely used error-correcting code.

*1) **Procedure 1**: Before the polynomial is created, the keying material is encoded into RS codewords.*

In fact, any other ECC scheme can be applied, although we only choose RS code to employ in our simulation, due to its ease-of-use in *q*-ary Galois field, and highly-efficient capability of error correcting. Firstly, the keying material is segmented as $key = k_0 \mid k_1 \mid ... \mid k_{K-1}$, where $\mid$ represents concatenation. Supposing k_l is the length of keying material, and (N, K, q) are the basic parameters of Reed-Solomon Code, which denote that K symbols will be encoded to codeword with the symbol length N in *q*-ary Galois field, this equation can be obtained,

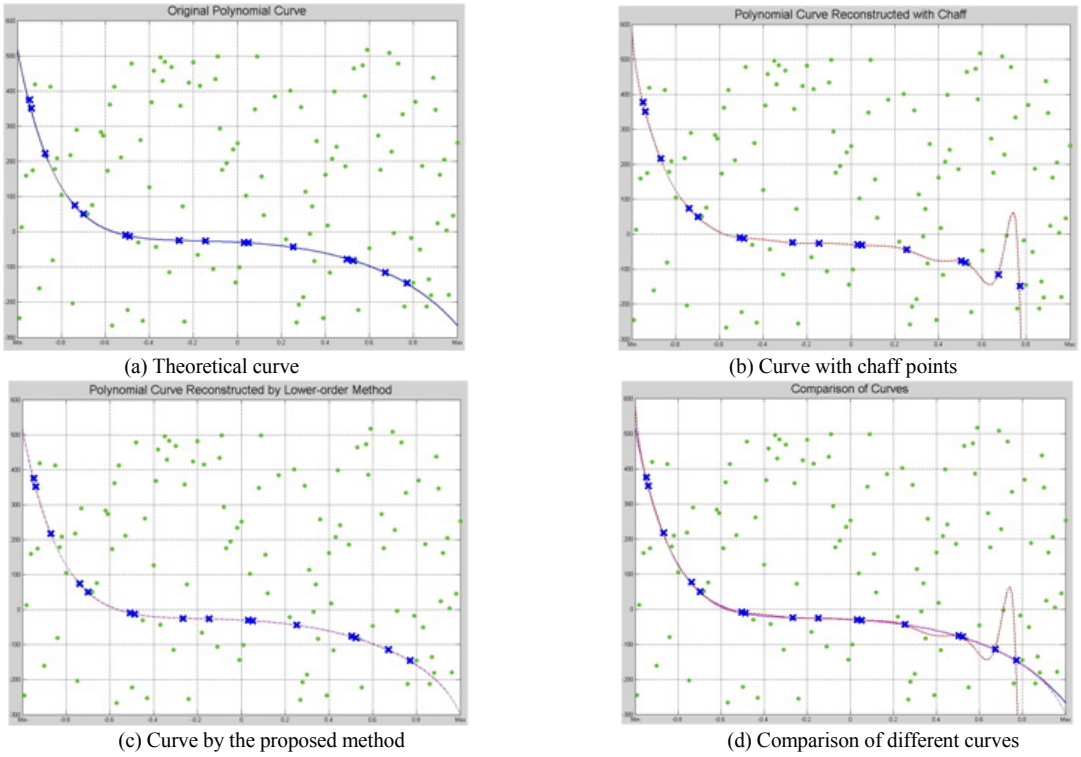
$$q * K = k_l \tag{2}$$

*2) **Procedure 2**: The RS codeword is employed as the coefficients of the chosen polynomial.*

The keying material is encoded to the codewords, i.e. $e_0, e_1, ..., e_{N-1}$, and the polynomial will be constructed as $p(x) = \sum_{i=0}^{N-1} e_i x^i$. It is known that construct a v_{th} -order polynomial, $v_{th}+1$ coefficients are needed at least. Therefore, the parameters of RS code should be chosen in particular,

$$v_{th} + 1 = N \tag{3}$$

Furthermore, before or after encoding the bit lengths of symbol in *q*-ary Galois field are identical. In other words, the length of polynomial coefficients c_l is equal to q. To some extent, the employment of RS code solve the trade-off problem mentioned before. The length of coefficients, i.e. $c_l = k_l / K$, is no longer inversely related to the order of polynomial $v_{th} = N - 1$, and they can increase simultaneously if N and K are appropriately selected.



(a) Theoretical curve (b) Curve with chaff points

(c) Curve by the proposed method (d) Comparison of different curves

Fig. 3. Comparison of curves reconstructed by different methods (✖ represents the genuine points in vault, and ● represents the chaff points. X-axis represents the value of feature point and Y-axis represents the calculated value of polynomial.)

1) ***Procedure 3****: Construct the vault to lock the keying material.*

This procedure in our scheme is similar to the *UNLOCK* algorithm of original one, which has been introduced in section II. However we will do some discussions on this procedure. The relationship of feature and chaff points is very important. In terms of security, it might be considered that the range of chaff points shall be much larger than that of feature points, at the same time, the resolution shall be smaller. But, when take the enormous costs in construction and reconstruction of vault into account, the enhanced security level by this way is questionable. On the consideration of efficiency, we propose to use the same parameters of value range and resolution, i.e. $f_{max} = c_{\max}$, $f_{\min} = c_{\min}$, and $\sigma_f = \sigma_c$.

2) ***Procedure 4****: Reconstruct the polynomial with fewer feature points.*

This procedure aims to answer the question that how to recover the keying material with fewer matched feature points, which also is the key to breakthrough the constraint of applications. The set $P = \{f_i, p(f_i)\}$ obtained in Procedure 3, is used for the polynomial reconstruction. In the original scheme, there are at least $v_{th}+1$ matched feature points needed to reconstruct a v_{th} -order polynomial, but in our proposed scheme the matched points can be less than $v_{th}+1$. In case that there is no enough matching points found out at the receiving node, one can do reconstruction to use the matched feature points together with a few chaff points. This is just like the attempt of attackers to crack the locked vault. But according to the simulation results, it's impossible to successfully reconstruct the polynomial with any chaff point. In most of same network cases, the EI generated by a genuine receiver node is very similar to that of the transmitter, and thus the differences between mismatching feature points are always small. Based on this assumption, in order to reduce FRR and improve success rate of recovery of keying material, a new reconstruction method, named Lower-Order Twice Reconstruction (LOTR), is proposed as follows. Firstly, the matched feature points, which are not enough in most of cases, are used to construct a lower-order polynomial. Secondly, estimate the remaining feature points by making use of the lower-order polynomial. Finally, reconstruct the polynomial with the matched and estimated points. A comparison of curves constructed by different methods is illustrated in Fig. 3, where LOTR is demonstrated to be effective in recovery of keying materials. Detailed discussion will be given in Section IV.

3) ***Procedure 5****: Decode the reconstructed polynomial coefficients in order to obtain the original keying material.*

This procedure is to recover the keying material that is locked in vault. After Procedure 4, the reconstructed polynomial is given as

$$p'(x) = \sum_{i=0}^{N} d_i x^i$$

where coefficients is the RS codewords ready to be decoded. After the decoding process, the original keying material is

recovered, i.e. $key' = k'_0 \mid k'_1 \mid ... \mid k'_{K-1}$. Though the LOTR method is able to reduce the errors of coefficients recovery, it is still not enough for a receiver node to obtain the exactly correct key. That is why the coefficients are error-correcting encoded in the proposed scheme. Furthermore, the error-correcting code is also helpful to improve the overall efficiency of the system, where the reconstruction procedure can potentially introduce another kind of errors because of the computational deviation.

IV. Simulation Results and Discussions

We simulated the proposed scheme using Matlab software to evaluate its key recovery performance. Assume that the bit length of keying material is 64. According to Equation (2), the parameters of RS code is selected as (N,8,8) in the simulation. Table II shows results of success rates with different orders of polynomial, which also determines the parameter N as shown in Equation (3). In addition, the column of Threshold indicates the capability of ECC. For example, RS(20,8,8) with N=20 can correct up to 6 symbol errors.

Table II shows the results in case there is one mismatched feature point, where the sub-column *With chaff points* means the success rate of reconstruction using one chaff point instead of a genuine point. For instance, if the order is 27 with one mismatched point, just as the 5th row in Table II shows, the success rate of the LOTR method is 94.9% while the method with a chaff point is 2.2% only. It is noted that while increasing the order of polynomial, the success rate without ECC decreases dramatically; on the contrary, the rate with ECC increases. Table III shows a decreasing trend of success rates with a different number of mismatched feature points while the order of polynomial is set to 27.

V. Conclusion

This study focuses on the applicability analysis of FVS for its usage in biometrics based security solution for BSN. It is found that there are many problems unsolved to make effective use of the original FVS. A simulation study is carried out to understand the relationships among different parameters in the protocol. Based on the simulation results, an improved scheme with double fuzziness is proposed, where the polynomial coefficients are Error-Correcting encoded, and a new reconstruction method, called the lower-order twice reconstruction, is employed to increase the success rate of *UNLOCK* process in FVS. More follow-up works need to be carried out in order to reach an actually feasible solution, where frequency-domain information of physiological signals can be fully utilized for security purposes.

TABLE II. Performance with One Mismatched Feature Point

Order	Threshold	Success rate of reconstruction without ECC	Success rate of recovery for keying material with one mismatched points	
			With chaff points	*Lower-order twice reconstruction*
19	6	98.1%	2.1%	55.3%
21	7	97.4%	2.1%	70.5%
23	8	92.0%	1.9%	82.0%
25	9	84.7%	3.7%	84.8%
27	10	74.4%	2.2%	94.9%
31	12	27.3%	3.0%	98.2%

TABLE III. Performance with More Mismatched Feature Points

Number of mismatched points for reconstruction with RS(28,8,8)	Success rate of recovery for keying material	
	With chaff points	*Lower-order twice reconstruction*
1	2.2%	94.9%
2	0.7%	87.9%
3	0.2%	70.1%
4	0.4%	61.1%
5	0.6%	48.9%

References

[1] Shu-Di Bao, Carmen C. Y. Poon, Yuan-Ting Zhang, and Lian-Feng Shen, "Using the Timing Information of Heartbeats as an Entity Identifier to Secure Body Sensor Network," IEEE Transactions on Information Techonology in Biomedicine, 12(6): 772-779, 2008.

[2] Krishna K. Venkatasubramanian, Ayan Banerjee, and Sandeep Kumar S. Gupta, "PSKA: Usable and Secure Key Agreement Scheme for Body Area Networks," IEEE Transactions on Information Technology in Biomedicine, 14(1): 60-68, 2010.

[3] Krishna K. Venkatasubramanian, Ayan Banerjee and Sandeep K. S. Gupta, "Plethysmogram-based secure inter-sensor communication in Body Area Networks," Military Communications Conference, MILCOM 2008. IEEE, pp. 1-7, 2008.

[4] Krishna Kumar Venkatasubramanian, Ayan Banerjee, and Sandeep K. S. Gupta, "EKG-based Key Agreement in Body Sensor Networks," INFOCOM Workshops 2008, IEEE, pp. 1-6, 2008.

[5] Ari Juels, and Madhu Sudan, "A Fuzzy Vault Scheme," Designs, Codes and Cryptography, 38(2): 237–257, 2006.

[6] Karthik Nandakumar, Anil K. Jain, and Sharath Pankanti, "Fingerprint-Based Fuzzy Vault:Implementation and Performance," IEEE Transaction on Information Forensics and Security, 2(4): 744-757, 2007.

[7] Youn Joo Lee, Kang Ryoung Park, Sung Joo Lee, Kwanghyuk Bae, and Jaihie Kim, "A New Method for Generating an Invariant Iris Private Key Based on the Fuzzy Vault System," IEEE Transactions on Systems, Man, and Cybernetics—part B: Cybernetics, 38(5): 1302-1313, 2008.

[8] A. Juels and M.Wattenberg, "A fuzzy commitment scheme," in Proc. 6th ACM Conf. Comput. Commum. Sec., pp. 28–36, 1999.

[9] Walter J. Scheirer, Terrance E. Boult, "Cracking Fuzzy Vaults and Biometric Encryption," Biometrics Symposium, Baltimore, MD, pp.1-6, 2007.

[10] Fen Miao, Shu-Di Bao, Ye Li, "A Modified Fuzzy Vault Scheme for Biometrics-Based Body Sensor Networks Security," IEEE Global Telecommunications Conference, Miami, FL, 2010.