ISSN 2313-1527 (PRINT)
ISSN 2313-1535 (ONLINE)

IJEIE

International Journal of Electronics and Information Engineering

Vol. 7, No. 1 (Sept. 2017)

INTERNATIONAL JOURNAL OF ELECTRONICS & INFORMATION ENGINEERING

PUBLISHING OFFICE

Min-Shiang Hwang
Department of Computer Science & Information Engineering, Asia University, Taichung 41354, Taiwan, R.O.C.
Email: mshwang@asia.edu.tw

International Journal of Electronics and Information Engineering is published both in traditional paper form (ISSN 2313-1527) and in Internet (ISSN 2313-1535) at http://ijeie.jalaxy.com.tw

PUBLISHER: Candy C. H. Lin

23-75, P.O. Box, Taichung, Taiwan 40199, R.O.C.

International Journal of Electronics and Information Engineering

Vol. 7, No. 1 (Sept. 1, 2017)

On Bilinear Groups of a Large Composite Order

Lihua Liu[1], Zhengjun Cao[2], Wenping Kong[1], Jinbo Wang[3]
(Corresponding author: Zhengjun Cao)

Department of Mathematics, Shanghai Maritime University, Shanghai, 201306, China[1].
Department of Mathematics, Shanghai University, Shanghai 200444, China[2].
(Email: caozhj@shu.edu.cn)
Science and Technology on Communication Security Laboratory, Chengdu 610041, China[3].
(Received Apr. 2, 2017; revised and accepted May 12, 2017)

Abstract

The algebraic structure of bilinear groups with a large composite order which supports subgroup decision problem, was introduced into cryptography by Boneh et al., in order to design new homomorphic public-key encryption schemes. In this paper, we would like to point out that the structure loses the advantages of elliptic curve cryptography which gained mainly from smaller parameter size. From the practical point of view, this structure is unlikely applicable to cryptographic schemes.

Keywords: Bilinear Groups of Composite Order; Homomorphic Public-key Encryption; Elliptic Curve Cryptography; Pairing-based Cryptography

1 Introduction

The use of elliptic curves in cryptography was suggested independently by Koblitz [11] and Miller [17] in 1985. The advantages of elliptic curve cryptography (ECC) are mainly gained from smaller parameter size. It is generally accepted that a 160-bit elliptic curve key provides the same level of security as a 1024-bit RSA key [9].

Weil pairing plays a key role in elliptic curve cryptography. In 1993, Menezes, Okamoto and Vanstone [16] suggested to use Weil pairing to reduce elliptic curve logarithms to logarithms in a finite field. In 2001, Boneh and Franklin [2] proposed a fully functional identity-based encryption scheme based on Weil pairing. Since then, an abundance of research has been published on the efficient implementation of these pairings, such as modified Weil pairing, Tate pairing.

Except the general restrictions on the domain parameters for an elliptic curve E over a finite field $\mathbb{F}_q$, these bilinear pairings require that the underlying groups should be of prime order ($> 2^{160}$) so that the elliptic curve discrete logarithm problem (ECDL) is resistant to all known attacks, such as Pohlig-Hellman attack, Pollard's rho attack. Thus, it is necessary that the cardinality $\#E(\mathbb{F}_q)$ should be divisible by a sufficiently large prime.

In order to alleviate user's pairing computation burden, Chevallier-Mames et al. [8] explored the problem of secure delegation of elliptic-curve pairing. Other algorithms for outsourcing of bilinear pairings were discussed in [4, 5, 7, 15]. In 2016, Hsien et al. [10] presented a survey of public auditing

for secure data storage in cloud computing. Others discussed the problem of public auditing for shared data storage with user revocation [6, 12, 13].

In 2005, Boneh et al. [3] introduced the subgroup decision problem in bilinear groups of a large composite order n so that it supports a homomorphic public key encryption.

They assume that it is hard to decide if an element in a subgroup, without knowing the factorization of n. Since then, researchers have proposed some cryptographic schemes [14, 21, 22] based on subgroup decision problem. It seems that the algebraic structure (bilinear groups of a large composite order) facilitates the security arguments of these protocols.

In this paper, we would like to remark that bilinear groups of a large composite order (at least 1024 bits) could make group operation very slow. So far, there are no testing reports on this topic. Taking into account that the common elliptic curve cryptosystems have not still displaced the classical public key cryptosystems such as RSA and ElGamal, we stress that the algebraic structure, bilinear groups of a large composite order, is unlikely applicable to cryptographic schemes although it facilitates the security proofs of some complicated cryptographic protocols.

2 Preliminaries

2.1 Weil Pairing

Definition 1. *An elliptic curve E over a finite field $\mathbb{F}_q$ is defined by an equation $E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, where $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$, $\triangle \neq 0$, and $\triangle$ is the discriminant of E and is defined as follows:*

$$\begin{aligned} \triangle &= -d_2^2d_8 - 8d_4^3 - 27d_6^2 + 9d_2d_4d_6, \quad \textit{where} \\ d_2 &= a_1^2 + 4a_2, \\ d_4 &= 2a_4 + a_1a_3, \\ d_6 &= a_3^2 + 4a_6, \\ d_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2. \end{aligned}$$

If L is any extension field of $\mathbb{F}_q$, then the set of L-rational points on E is

$$E(L) = \{(x, y) \in L \times L : y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0\} \cup \{\infty\},$$

where ∞ is the point at infinity.

The number of points in the group $E(\mathbb{F}_q)$, denoted by $\#E(\mathbb{F}_q)$, is called the order of E over $\mathbb{F}_q$. Hasse's theorem provides tighter bounds for $\#E(\mathbb{F}_q)$.

Theorem 1. *(Hasse) Let E be an elliptic curve defined over $\mathbb{F}_q$. Then $q + 1 - 2\sqrt{q} \leq \#E(\mathbb{F}_q) \leq q + 1 + 2\sqrt{q}$.*

Let n be a prime and coprime to the characteristic of $\mathbb{F}_q$. Suppose n divides $\#E(\mathbb{F}_q)$.

Then there exists a n-torsion group [23]:

$$E(\mathbb{F}_{q^k})[n] := \{P \in E(\mathbb{F}_q^k) \,|\, nP = \mathcal{O}\},$$

where the number k is called the *embedding degree* which is the smallest positive integer such that n divides $(q^k - 1)$, and nP denotes the sum of n copies of P. The existence of the n-torsion group is due to Balasubramanian and Koblitz [1].

Theorem 2. *Let E be an elliptic curve over $\mathbb{F}_q$ and let n be a prime dividing $\#E(\mathbb{F}_q)$. Suppose that n does not divide $(q-1)$ and that $\gcd(n,q)=1$. Then the n-torsion group $E[n] \subset E(\mathbb{F}_{q^k})$ if and only if n divides (q^k-1).*

The structure of n-torsion group $E(\mathbb{F}_{q^k})[n]$ is described by the following relation:

$$E(\mathbb{F}_{q^k})[n] \cong (\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}).$$

That means $\#E(\mathbb{F}_{q^k})[n] = n^2$. The Weil pairing is defined on the n-torsion group $E(\mathbb{F}_{q^k})[n]$, not on any n-order group $G \subset E(\mathbb{F}_q)$.

Let μ_n be the group of nth roots of unity. Clearly, $\mu_n \subset \mathbb{F}_{q^k}$, but $\mu_n \not\subset \mathbb{F}_{q^j}$ for $j = 1, \cdots, k-1$.

Definition 2. *The Weil pairing is a map $e_n : E(\mathbb{F}_{q^k})[n] \times E(\mathbb{F}_{q^k})[n] \to \mu_n$ with the following properties:*

1) Linearity: If $P, Q, R \in E(\mathbb{F}_{q^k})[n]$, then

$$\begin{aligned} e_n(P+Q,R) &= e_n(P,R)e_n(Q,R), \\ e_n(P,Q+R) &= e_n(P,Q)e_n(P,R). \end{aligned}$$

2) Alternating: If $P \in E(\mathbb{F}_{q^k})[n]$, then $e_n(P,P) = 1$. This, along with linearity, implies that if $P, Q \in E(\mathbb{F}_{q^k})[n]$, then $e_n(Q,P) = e_n(Q,P)^{-1}$.

3) Non-degeneracy: If $\mathcal{O} \neq P \in E(\mathbb{F}_{q^k})[n]$, there exists $Q \in E(\mathbb{F}_{q^k})[n]$ such that $e_n(P,Q) \neq 1$.

To construct a concrete Weil pairing, we need the following equivalent definition [18].

Definition 3. *Let $n > 1$ be an integer and let $\mathfrak{D}_1, \mathfrak{D}_2$ be divisors on an elliptic curve, E, with disjoint supports, such that $n\mathfrak{D}_1, n\mathfrak{D}_2 \sim 0$. This means that there are functions f_1 and f_2 such that $n\mathfrak{D}_i = div(f_i)$ for $i = 1, 2$. The Weil pairing is defined as $e_n(\mathfrak{D}_1, \mathfrak{D}_2) = \frac{f_1(\mathfrak{D}_2)}{f_2(\mathfrak{D}_1)}$.*

Remark 1. *Most literatures use the notation $E[n]$ to denote the n-torsion group, which does not specify that the representation of points, the computation of functions and the evaluation for those functions should be performed in the extension field $\mathbb{F}_{q^k}$.*

2.2 Miller's Algorithm

In order to calculate Weil pairing, one should evaluate $f(\mathfrak{D})$, where $\mathrm{div}\,(f) = n([P]-[O])$. In 1985, Miller gave an explicit algorithm for calculating Weil pairing. Of course, it can be used to calculate Tate pairing, because it is also defined on the n-torsion group.

Let E be an elliptic curve over the field K and $P, Q \in E(K)$. Let $L_{P,Q}$ be the normalized function, such that $L_{P,Q} = 0$ is the equation of the line passing through P and Q (or the equation of the tangent line to the curve if $P = Q$). Then $\mathrm{div}(L_{P,Q}) = [P] + [Q] + [-(P+Q)] - 3[\mathcal{O}]$. Let $h_{P,Q} := \frac{L_{P,Q}}{L_{P+Q,-(P+Q)}}$. We have $\mathrm{div}\,(h_{P,Q}) = [P] + [Q] - [P+Q] - [\mathcal{O}]$. Let $f_{0,P} = f_{1,P} = 1$. Inductively, for $n > 0$, define $f_{n+1,P} := f_{n,P}\, h_{P,nP}$, we have

$$\mathrm{div}\,(f_{n,P}) = n[P] - (n-1)[\mathcal{O}] - [nP].$$

It is easy to find that $f_{m+n,P} = f_{m,P} f_{n,P} h_{mP,nP}$, $f_{mn,P} = f_{m,P}^n f_{n,mP} = f_{n,P}^m f_{m,nP}$. This means the calculation of $f_{n,P}(Q)$ resembles exponentiation and it can be done in $O(\log n)$ point additions on E/K. Using the constructed functions, we obtain the following formulas [18].

Proposition 1. *Suppose that T is a point in $E(K)$ different from $P, Q, Q-P$, and $\mathcal{O}$. Then $[P]-[\mathcal{O}] \sim [P+T]-[T]$, and the supports of $[Q]-[\mathcal{O}]$ and $[P+T]-[T]$ are disjoint. We have*

$$e_n(P,Q) = \frac{f_{n,Q}(T) f_{n,P}(Q-T)}{f_{n,P}(-T) f_{n,Q}(P+T)}.$$

Proposition 2. *Let E/K be an elliptic curve, $P, Q \in E(K)[n]$, and $P \neq Q$. Then*

$$e_n(P,Q) = (-1)^n \frac{f_{n,P}(Q)}{f_{n,Q}(P)}.$$

Remark 2. *The functions $f_{n,P}, f_{n,Q}$ must be calculated in the group $E(K)$ where the field K satisfies $\mu_n \subset K$. So do the evaluations of $f_{n,P}(Q), f_{n,Q}(P)$.*

3 BGN Homomorphic Encryption Scheme

3.1 Bilinear Groups of Composite Order

Let $\mathcal{G}$ be a group generation algorithm that takes security parameter 1^λ as input and outputs tuple $(p, q, \mathbb{G}, \mathbb{G}_1, e)$ where p and q are distinct primes, $\mathbb{G}$ and $\mathbb{G}_1$ are cyclic groups of order $n = p\,q$, and $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_1$ is a non-degenerate bilinear map, i.e., it satisfies: (i) bilinear: for $\forall\, g_1, g_2 \in \mathbb{G}$ and $\forall\, a, b \in \mathbb{Z}$, $e(g_1^a, g_2^b) = e(g_1, g_2)^{a\,b}$; (ii) non-degenerate: for generator g of $\mathbb{G}$, $e(g, g)$ generates $\mathbb{G}_1$.

Let $\mathbb{G}_p$ and $\mathbb{G}_q$ denote the subgroups of $\mathbb{G}$ of order p and q, respectively. Then $\mathbb{G} = \mathbb{G}_p \times \mathbb{G}_q$. If g is a generator of $\mathbb{G}$, then g^q and g^p are generators of $\mathbb{G}_p$ and $\mathbb{G}_q$, respectively. Let g_p and g_q denote the generators of $\mathbb{G}_p$ and $\mathbb{G}_q$, respectively. For all random elements $h_p \in \mathbb{G}_p$ and $h_q \in \mathbb{G}_q$, We have $e(h_p, h_q) = 1$ because $e(h_p, h_q) = e(g_p^a, g_q^b)$ for some integers a, b, and

$$\begin{aligned} e(g_p^a, g_q^b) &= e(g^{q\,a}, g^{p\,b}) \\ &= e(g, g)^{p\,q\,a\,b} \\ &= 1 \end{aligned}$$

for some generator g in $\mathbb{G}$.

We here want to stress that the subgroup decision assumption over a bilinear group $\mathbb{G}$ *requires a large composite order n so that it is resistant to all known factoring methods.*

3.2 Review of BGN Scheme

In 2005, Boneh, Goh and Nissim [3] introduced the subgroup decision problem over bilinear groups of a large composite order so that it supports a homomorphic public key encryption. We now describe the scheme as follows.

KeyGen: Set up the tuple $(q_1, q_2, \mathbb{G}, \mathbb{G}_1, e)$. Let $n = q_1 q_2$. Pick two random generators $g, u \in \mathbb{G}$ and set $h = u^{q_2}$. Note that the order of h in group $\mathbb{G}$ is q_1. The public key is $PK = (n, \mathbb{G}, \mathbb{G}_1, e, g, h)$. The private key is $SK = q_1$.

Encrypt: Assume the message space consists of integers in the set $\{0, 1, \cdots, T\}$ with $T < q_2$. To encrypt a message m, pick randomly $r \in Z_{n-1}$ and compute $C = g^m h^r \in \mathbb{G}$. Output C as the ciphertext.

Decrypt: Let $\hat{g} = g^{q_1}$. Note that $C^{q_1} = (g^m h^r)^{q_1} = (g^{q_1})^m$. Compute the discrete log of C^{q_1} base $\hat{g}$ to recover m. Since $0 \leq m \leq T$ this takes expected time $\tilde{O}(\sqrt{T})$ using Pollard's lambda method.

Extension: The authors [3] pointed out that anyone can multiply two encrypted messages once using the bilinear map. Set $g_1 = e(g, g)$ and $h_1 = e(g, h)$. Then g_1 is of order n and h_1 is of order q_1. Also, write $h = g^{\alpha q_2}$ for some (unknown) $\alpha \in \mathbb{Z}$. Suppose there are two ciphertexts

$$C_1 = g^{m_1} h^{r_1} \in \mathbb{G}, \quad C_2 = g^{m_2} h^{r_2} \in \mathbb{G}.$$

To build an encryption of the product $m_1 \cdot m_2 \bmod n$ given only C_1 and C_2, do: 1) pick a random $r \in \mathbb{Z}_n$; 2) set $C = e(C_1, C_2) h_1^r \in \mathbb{G}_1$. Then we have

$$C = e(C_1, C_2) h_1^r = e(g^{m_1} h^{r_1}, g^{m_2} h^{r_2}) h_1^r = g_1^{m_1 m_2} h_1^{m_1 r_2 + r_2 m_1 + \alpha q_2 r_1 r_2 + r} = g_1^{m_1 m_2} h_1^{\tilde{r}} \in \mathbb{G}_1$$

where $\tilde{r} = m_1 r_2 + r_2 m_1 + \alpha q_2 r_1 r_2 + r$ is distributed uniformly in $\mathbb{Z}_n$. Thus, C is a uniformly distributed encryption of $m_1 m_2 \bmod n$, but in the group $\mathbb{G}_1$ rather than $\mathbb{G}$. Note that the system is still additively homomorphic in $\mathbb{G}_1$.

4 Analysis of BGN Scheme

4.1 What Group is Used

It claims that BGN encryption [3] resembles Paillier encryption [20] and Okamoto-Uchiyama [19] encryption. We here want to stress that the claim is not convincing because

- Paillier system is constructed *over a multiplicative subgroup of integers modulo* n^2, where $n = p\,q$, p, q are two large primes.
- Okamoto-Uchiyama encryption is constructed *over a multiplicative group of integers modulo* n, where $n = p^2 q$, p, q are two large primes.
- BGN encryption is constructed *over a bilinear group* $\mathbb{G}$ *of a large composite order* $n = q_1 q_2$, where q_1, q_2 are two large primes so as to prevent the adversary from factoring n. The bilinear group is more complex and hard to implement practically.

4.2 Large Working Parameters

It is well-known that ECC schemes are public-key mechanisms that provide the same functionality as RSA schemes. However, a desired security level can be attained with significantly smaller parameters in ECC systems. For example, a 160-bit elliptic curve key provides the same level of security as a 1024-bit RSA parameter. Smaller parameters in ECC systems consequently save much power, bandwidth and storage, and bring a good speed.

The working parameters for an elliptic curve scheme describe an elliptic curve E over $\mathbb{F}_q$, a base point $P \in E(\mathbb{F}_q)$, and its order n. The parameters should be chosen so that elliptic curve discrete log problem (ECDL) is resistant to all known attacks. Usually, we select E so that $\#E(\mathbb{F}_q)$ is prime or almost prime, that is, $\#E(\mathbb{F}_q) = hn$ where n is prime and h is small (e.g., $h = 1, 2, 3$ or 4). That means the size of the working parameter q in an ECC scheme is approximately equal to the size of n. To avoid the Weil and Tate pairing attacks, one should ensure that n does not divide $q^k - 1$ for all $1 \le k \le 20$.

Pairing-based cryptography (PBC) has many elegant properties and interests many researchers. It is generally claimed that PBC can offer a desired security level with smaller parameters as ECC. Suppose that an elliptic curve E is defined over the field $\mathbb{F}_q$. Then ECC is working with elements which are defined over $\mathbb{F}_q$. But PBC is working with the functions and elements defined over $\mathbb{F}_{q^k}$, where k is the *embedding degree*. The security of PBC depends directly on the intractable level of either ECDL in

the group $E(\mathbb{F}_q)$ or discrete log problem (DL) in group $\mathbb{F}^*_{q^k}$. That means PBC protocols have to work in a running environment with parameters of 1024 bits so as to offer 80 bits security level. From the practical point of view, the shortcoming makes PBC lose its competitive advantages significantly. The parameter size comparison in Table 1 is adapted from [9].

Table 1: RSA, DL, EC, PBC parameter sizes for equivalent security levels. Bitlengths are given for DL parameter q and EC parameter n, and RSA modulus n and DL modulus p, respectively.

security level (bits)	80 (SKIPJACK)	112 (Triple-DES)	128 (AES-Small)	192 (AES-Medium)	256 (AES-Large)
EC parameter n EC parameter q	160	224	256	384	512
EC parameter k	>20				
PBC parameter n PBC parameter q^k	160 1024	224 2048	256 3072	384 8192	512 15360
PBC parameter k	≤6				
DL parameter q DL modulus p	160 1024	224 2048	256 3072	384 8192	512 15360
RSA modulus n	1024	2048	3072	8192	15360

It is easy to find that pairing-based cryptography protocols require that:

- The base point $P \in E(K)$ has a sufficiently large prime order n such that ECDL in $E(K)$ is intractable;
- DL in K^* is intractable in order to resist MOV reduction attacks [16];
- It is efficient to compute pairings in $E(K)$.

From the practical point of view, it is annoying for PBC schemes to have to work in extensions of the base fields, even though the inputting parameters are defined over the base field. Taking into account the very long and complicated programming code for PBC systems (see *PBC 0.5.14*, maintained by Lynn, released on Jun 14, 2013), we find that PBC schemes are far slow than its DL counterparts. Smaller inputting parameters in PBC systems can not truly bring them a good speed.

So far, bilinear groups used in cryptographic protocols are derived *only* from elliptic curves. To construct a bilinear group with a large composite order n, it requires a large q by the Hasse theorem. If n is of 1024 bits, then the parameter q should be of more bits.

As mentioned earlier, in classical elliptic curve cryptograph it requires that the parameter q is about of 160 bits. Apparently, a large parameter q (at least 1024 bits) makes group operation very inefficient and loses the advantages of elliptic curve cryptograph gained mainly from smaller parameter size. By the way, the longest parameter, as of 2016, recommended by NIST for elliptic curves has 571 bits.

4.3 Difficulties of Implementing BGN Scheme

To the best of our knowledge, there were few industrial products being integrated with paring-based cryptosystems. The reasons for this situation could be summarized as follows.

1) The pairing computation is somewhat hard to understand for most engineers.

2) In most paring-based cryptosystems, there is a Private Key Generator (PKG) who is responsible for generating private keys for users. The key escrow problem is not compatible with the current PKI system.

3) The heavy group operation of elliptic curve really lowers the advantages that gained from smaller parameter size, let alone the more complicated pairing computation. This is just the reason that common elliptic curve cryptosystems have not still broadly displaced classical public key cryptosystems such as RSA and ElGamal.

For the current status of applied paring-based cryptography, we refer to the following links.

- PBC Library [`http://crypto.stanford.edu/pbc/`]; pbc-0.5.14 (Released on Jun 14, 2013).
- JPBC Library [`http://gas.dia.unisa.it/projects/jpbc/faq.html`]; v2.0.0 (Released on Dec 04, 2013).
- TinyPairing [`http://www.cs.cityu.edu.hk/~ecc/TinyPairing/`]. TinyPairing v0.1 (Released on Oct 09, 2009).
- Jmiracl [`https://dsl-external.bbn.com/tracsvr/openP3S/wiki/jmiracl\`].

5 Conclusion

We show that the structure of bilinear groups of a large composite order is unlikely applicable to cryptographic schemes. We would like to stress that the fundamental reason of introducing such a structure is to facilitate academic security proofs of some complicated cryptographic protocols.

However, "the theorem-proof paradigm of theoretical mathematics is often of limited relevance and frequently leads to papers that are confusing and misleading" [Koblitz and Menezes: another look at "provable security", J. Cryptology 20(1), 2007]. We hope this note will help to right the balance between academic researches of pairing-based cryptography and its practice.

Acknowledgements

We thank the National Natural Science Foundation of China (61303200, 61411146001). We are grateful to the reviewers for their valuable suggestions.

References

[1] R. Balasubramanian and N. Koblitz, "The improbability that an elliptic curve has sub-exponential discrete log problem under the menezes Okamoto vanstone algorithm," *Journal of Cryptology*, no. 11, p. 141–145, 1998.

[2] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Proceedings of Advances in Cryptology (CRYPTO'01)*, pp. 213–229, Santa Barbara, California, USA, Aug. 2001.

[3] D. Boneh, E. Goh, and K. Nissim, "Evaluating 2-DNF formulas on ciphertexts," in *Proceedings of Theory of Cryptography (TCC'05)*, pp. 325–341, Cambridge, MA, USA, Feb. 2005.

[4] S. Canard, J. Devigne, and O. Sanders, "Delegating a pairing can be both secure and efficient," in *Proceedings of Applied Cryptography and Network Security (ACNS'14)*, pp. 549–565, Lausanne, Switzerland, June 2014.

[5] Z. J. Cao, L. H. Liu, and O. Markowitch, "On two kinds of flaws in some server-aided verification schemes," *International Journal of Network Security*, vol. 18, no. 6, pp. 1054–1059, 2016.
[6] T. Y. Chen, C. C. Lee, M. S. Hwang, and J. K. Jan, "Towards secure and efficient user authentication scheme using smart card for multi-server environments," *Journal of Supercomputing*, vol. 66, no. 2, pp. 1008–1032, 2013.
[7] X.F. Chen and et al., "Efficient algorithms for secure outsourcing of bilinear pairings," *Theoretical Computer Science*, no. 562, pp. 112–121, 2015.
[8] B. Chevallier-Mames, et al., "Secure delegation of elliptic-curve pairing," in *Proceedings of 9th IFIP WG 8.8/11.2 International Conference on Smart Card Research and Advanced Application (CARDIS'10)*, pp. 24–35, Passau, Germany, Apr. 2010.
[9] D. Hankerson, A. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*, USA: Springer, 2004.
[10] W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for secure data storage in cloud computing," *International Journal of Network Security*, vol. 18, no. 1, pp. 133–142, 2016.
[11] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, pp. 203–209, 1987.
[12] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing," *International Journal of Network Security*, vol. 18, no. 4, pp. 650–666, 2016.
[13] C.W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage," *International Journal of Network Security*, vol. 18, no. 5, pp. 900–916, 2016.
[14] J. Liu, T. Yune, and J. Zhou, "Forward secure ring signature without random oracles," in *Proceedings of Information and Communications Security - 18th International Conference (ICICS'11)*, pp. 1–14, Beijing, China, Nov. 2011.
[15] L. H. Liu and Z. J. Cao, "A note on 'efficient algorithms for secure outsourcing of bilinear pairings'," *International Journal of of Electronics and Information Engineering*, vol. 5, no. 1, pp. 30–36, 2016.
[16] A. Menezes, T. Okamoto, and S. Vanstone, "Reducing elliptic curve logarithms to logarithms in a finite field," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1639–1646, 1993.
[17] V. Miller, "Use of elliptic curves in cryptography," in *Proceedings of Advances in Cryptology (CRYPTO'85)*, pp. 417–426, Santa Barbara, California, USA, Aug. 1985.
[18] V. Miller, "The weil pairing, and its efficient calculation," *Journal of Cryptology*, no. 17, p. 235–261, 2004.
[19] T. Okamoto and S. Uchiyama, "A new public-key cryptosystem as secure as factoring," in *Proceedings of Advances in Cryptology (Eurocrypt'98)*, pp. 308–318, Espoo, Finland, May 1998.
[20] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proceedings of Advances in Cryptology (Eurocrypt'99)*, pp. 223–238, Prague, Czech Republic, May 1999.
[21] J. Seo, T. Kobayashi, M. Ohkubo, and K. Suzuki, "Anonymous hierarchical identity-based encryption with constant size ciphertexts," in *Proceedings of Public Key Cryptography (PKC'09)*, pp. 215–234, Irvine, CA, USA, Mar. 2009.
[22] H. Shacham and B. Waters, "Efficient ring signatures without random oracles," in *Proceedings of Public Key Cryptography (PKC'07)*, pp. 166–180, Beijing, China, Apr. 2007.
[23] J. Silverman, *The Arithmetic of Elliptic Curves*, USA: Springer, 1986.

Biography

Lihua Liu is an associate professor with Department of Mathematics at Shanghai Maritime University. She received her Ph.D. degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics, cryptography and information security.

Zhengjun Cao is an associate professor with Department of Mathematics at Shanghai University. He received his Ph.D. degree in applied mathematics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He served as a post-doctor in Computer Sciences Department, Université Libre de Bruxelles, from 2008 to 2010. His research interests include cryptography, discrete logarithms and quantum computation.

Wenping Kong is currently pursuing her M.S. degree from Department of Mathematics, Shanghai Maritime university. Her research interests include combinatorics and cryptography.

Jinbo Wang received his Ph.D. degree in applied mathematics from Shanghai University. His research interests include applied cryptography and network security.

Strategic Business Analytics and Alternative Solutions from Decision Making Perspective

Ahmed Hamdi, Samir AbdElrazek, Ahmed AbuElfotoh, Hazem El-Bakry

(Corresponding author: Hazem El-Bakry)

Information Systems Department, Faculty of Computer and Information Sciences
Mansoura University El Gomhouria
St., Mansoura, Egypt
(Email: helbakry5@yahoo.com)

(Received Mar. 6, 2017; revised and accepted May 12 & June 4, 2017)

Abstract

In today's digital world, Entrepreneurs make strategic decisions every day. Changes in technology, changes in organizational priorities, macroeconomic changes, and growth of global information capacity have been influenced substantial changes in decision making process. This paper refines the deep influence of the strategic business analytics factors and the influence mechanism of these factors on the strategic decision making to optimize data gathering, analysis methodologies and enhance the alternative solutions to make decisions based on the optimal solution.

Keywords: Alternative Solutions Analysis; Decision Making; Information Systems Components; Strategic Business Analytics

1 Introduction

The ability to extract and present information in a meaningful way is vital to business success. Regardless of where they sit, Entrepreneurs need the ability to transform data into actionable insight. Executives need aggregated data that captures a complete view of the company and its operations if they are to measure performance and respond proactively to changes in the marketplace and the industry. Managers, teams, and individuals need the ability to find, share, and use information from across all areas of the business to perform tasks effectively and monitor business operations. We focus here on a new and different perspective to analyze sufficient information to give to the decision maker. In this paper we illuminate a new and different perspective to analyze sufficient information to enhance decision making process. We analyze factors like software, hardware, people that have been used to gather and analyze needed information. We have to realize that we are working on more than one level of information, and the rise in the level means less in terms of quantity, and more in terms of importance. So we will work to ensure containment of the top level for each data that must be transformed to become the way to best decision.

2 Factors Analysis Method

In this Method, the components of information systems are the essential factors and will be analyzed to produce information which will be used in making the decision, while the software is used to analyze hardware, data and people. It will be analyzed to determine the validity and the effectiveness of the Results. We cannot ignore the amount of information that can result by using this method, while striving to provide the right amount to make the best decision [1, 2]. This amount of information may result in difficulties in process and extraction and due to the need to organize the information we use two types of data, Date-Time Data and Location Data as a separated factors. The new additional part on this method is considering the need of analysis to the 6 items (see Figure 1):

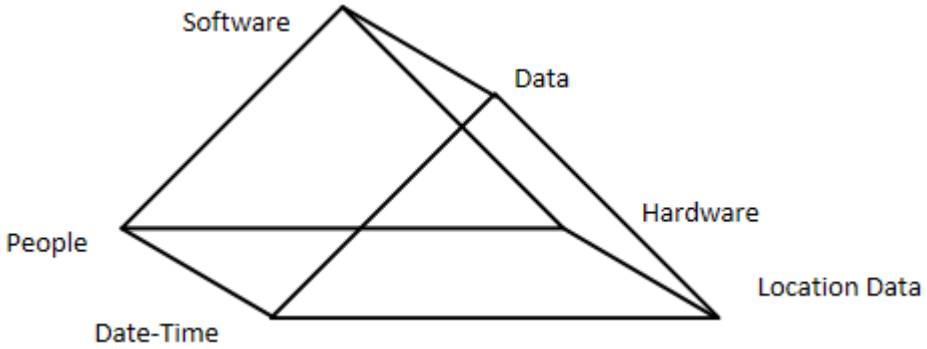

Figure 1: 6 Factors Analysis Method Items Framework

2.1 Software Analysis

Software is the soul by which we use to manage, control and monitoring performance of the hardware. It is used to solve the complex calculations and processes the data to extract valuable information to facilitate our future. What we are discussing here is how the software should be analyzed and provide information about the software and not only used in the analysis, and usually we get this information as a result of questions that will be answered and contribute to the measurement of the amount of the software validity to work.

2.1.1 Software As a Tool

Any Application-based service is using a software, as long as this application is used software will be involved in the process. To clarify this let's assume that you need to book a ticket for a travel from London to Paris. The application that enable you to book it online is a software involved in the process, the application that enable the Airline Employee to trace your ticket and process it is a software, and so on ... there are also application that we use to collect, store and analyze the data to give us insight about the traffic which we further use to make of decisions. So the question here: What is the software we use book a ticket online?

2.1.2 Software As Analysis Item

When we discuss software as a tool, we are not concerning about software we use to book a ticket, to pay online or to receive emails, our concern is about the software we use to analyse software that is mentioned above, software that is used by analysts, managers and decision makers. Here's the difference between both of them. To clarify this let's assume a company will make a decision that allow people to buy a product only online, so what we focus about here is not the software users will use to buy products. But the software that analysts are using to gather, store, process, and retrieve information as reports to give to managers. So the question here: What is the software we use to analyse? (See Figure 2)

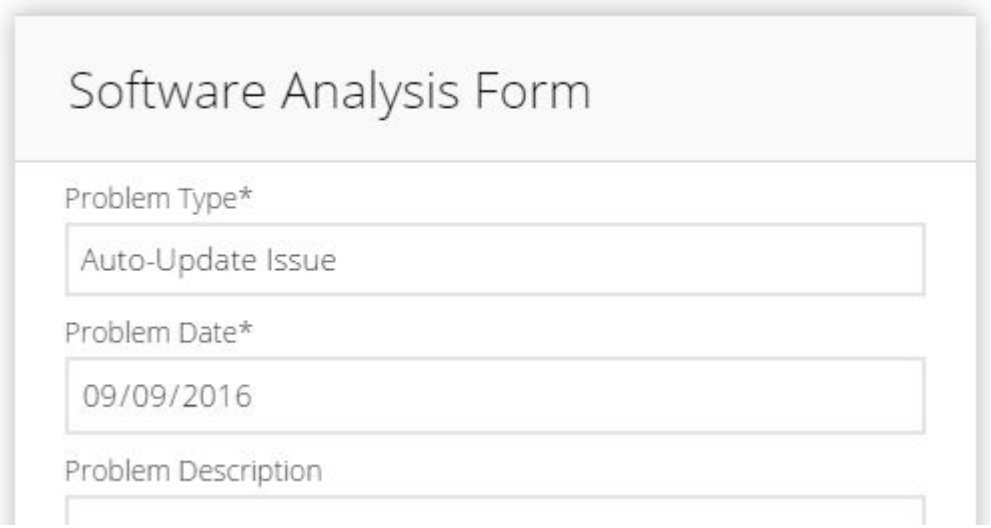

Figure 2: Framework supporting application: Software Analysis Form

2.2 Hardware Analysis

Hardware consists of everything in the physical layer of the information system. For example, hardware can include servers, workstations, networks, telecommunications equipment, fiber-optic cables, mobile devices, scanners, digital capture devices, and other technology-based infrastructure. As new technologies emerge, manufacturers race to market the innovations and reap the rewards [3]. It is the body of the system, which includes the collection of physical elements that are linked together made up the main base of the system. Hardware analysis which includes performance measurement, processing speed and mobility capabilities produces a vital information that helps in entire system stability. However, hardware issues don't involve directly in the analytical considerations.

2.2.1 Hardware As a Physical Layer

Starting from its definition, The Hardware is the base that software is using to be run. We use this term to describe the tangible materials that we use to control, manage and process instructions of programs. To clarify this let's assume that people can only buy a product online, machines will be used to order this product either it is a PC or smart phone, time of servers to respond requests, number of servers in

the network and number of requests could be handled per second, and so on... So the question here: What tools will be used by whom the decision will affect?

2.2.2 Hardware As Analysis Item

On the other hand, The Hardware consists of very important information that should be included into analysis stage about components used to analyze the applications, servers, networks and telecommunications used. By Analyzing, we get information about the environment of gathering and processing data. To clarify this let's assume a company will make a decision that allow people to buy a product only online, so what we focus about here is information about the performance of machines that will be used to gather, store, process and retrieve data rather than information about hardware will be used by people to order this product. So the question here: What I should use to make my decision? (See Figure 3)

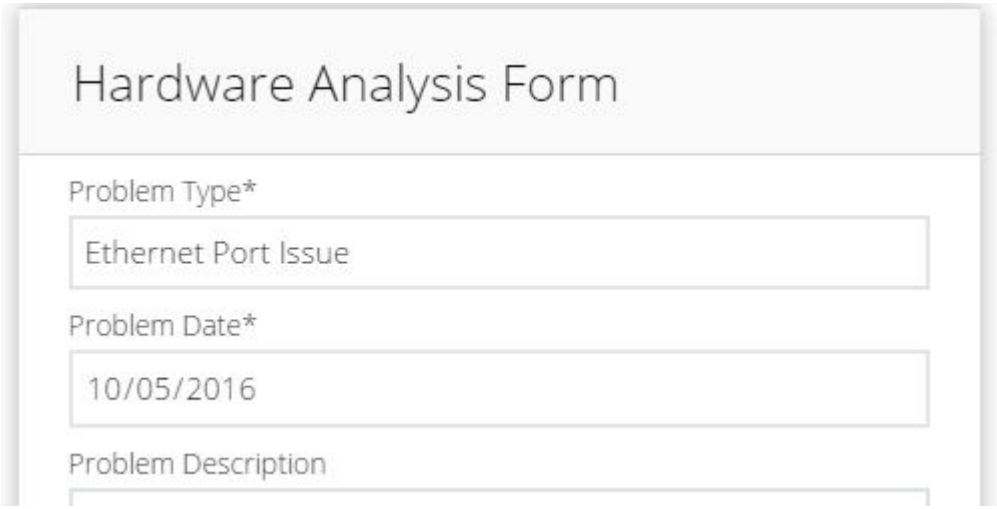

Figure 3: Framework supporting application: Hardware Analysis Form

2.3 People Analysis

Stakeholders, Managers, Analysts and operations users are example of people of various parts of the organization that are involved into the information system, those people are making decisions either uncertainty or regular basis decisions [4]. Involving teams in decision making improves the quality of decisions most of the times [5, 6], being efficient for the organization to generate and evaluate different alternatives of problems solving [7]. Sometimes the democratic decisions are not able to be made because of minority domination or time pressure [8]. The core here is to analyze people who make decisions and apply those decisions.

2.3.1 People As a System Component

There is no doubt that the people is one of the most important pillars of any information system, if not the most important which be relied on to deal with the unexpected and instantaneous changes, and to manage information systems to facilitate our life. People in traditional analysis cycle are the key

that analyze, and make decisions, so who control the system, give orders, store and retrieve data and manage information is one of people who are an information system component. So the question here: Who are the people who book a ticket for us, check the tickets, and so on...?

2.3.2 People As Analysis Item

On the other hand, if we will focus on people as an analysis item that will highlight the required characteristics of anyone who is doing simple task that affect the decision making process. To clarify this let's assume a company will make a decision if they will be able to sell kind of vehicles across a new country, so rather than concern about who will buy them, we will focus on who gathering data, who analyze it. Certifications, qualifications and training he had before making this task either it is a gathering data, analyzing it or making the decisions. So the question here: Who are the people gathering data, analyzing it, reporting information and who also are making the decisions? (See Figure 4)

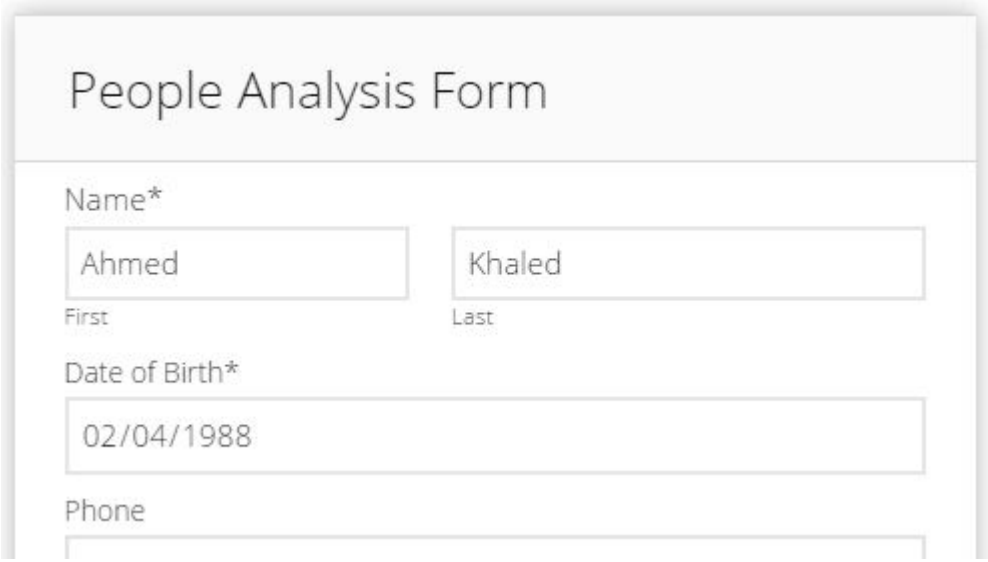

Figure 4: Framework supporting application: People Analysis Form

2.4 Data Analysis

Data is the raw material that an information system transforms into useful information. An information system can store data in various locations, called tables. By linking the tables, the system can extract specific information. Today, a big data is what we are dealing with and According to [9], Big Data should be defined at any point in time as data whose size forces us to look beyond the tried and true methods that are prevalent at that time, whereas for [10] Big Data refers to enormous amounts of unstructured data produced by high-performance applications falling in a wide and heterogeneous family of application scenarios. Data Analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data, but the concentration will be on the analysis of data of the cumulative decisions that are made in the past. We will analyze similar problems of the past in addition to the analysis of solutions that have been put forward at that time. Not only this, but also analyze the way that has been followed to determine the

appropriate solution and also analysis of the solution chosen, then analysis of the applicability of this solution if the same problem emerged of the present or the future. Will it be the appropriate solution to be applied? Or that there will be a need to change the method of selecting the best solution and then change the solution.

2.4.1 Data As Raw Values

Set of raw values are gathered together from different ways and by different tools, data is what we can be called. We cannot make a decision without data although the data might not help directly before processing it, but no data means nothing to be processed and no information will be available to make a decision. To clarify what we mean by data as raw values let's assume that a bank need to classify the most values to be withdrawn, to put them to its ATMs to facility and accelerate the client process, so for each client a report will include the amount of money, number of transactions. So they can determine values should be programmed like 100$, 500$ and 1000$ instead of 100$, 200$ and 300$. The question here: what is the data gathered from resources?

2.4.2 Data As Analysis Item

Our vision of things are different when you have a different vision perspective; as well as data. What we mean here is the data extraction, in addition to the search for data that we collect and analyze it, we should collect and analyze data on the sources of data and resources that collect data. To clarify that let's use the same prior example, a bank need to classify the most values to be withdrawn, to put them to its ATMs to facility and accelerate the client process. So we are looking for data about ATMs itself, how many transactions succeeded, failed or banned and amount of money per each transaction and what the percentage of amount of money withdrawn for the total amount. The question here: What is the data provided about resources? (See Figure 5)

2.5 Date-Time Data Analysis

2.5.1 Date-Time As Data

As all types of data, Time-based data is vital to decision maker as some decision could be valid for 1 days, 1 year, 100 year ... and those kind of data will be analyzed. To clarify this let's assume that a company need to sell a product so the time of advertisements in Paris will be different than in Dubai, simply how long buyers can see the advertisement, will they be at home, in streets, at work, summer or winter and so on ... so the goal here is to include Date-Time data inside the analysis cycle. The question here: When will this decision be happened?

2.5.2 Date-Time As Analysis Item

The difference here is that Date-Time data will not concern about the Time of decision, instead it will be concern about when the analysis will be occurred the day while analysis will be processed. To clarify this let's assume that a company need to sell a product in Paris so what I concern about here when are the decision makers will make their decisions, 10th Feb, 1st May, 30th Dec and so on ... so the goal here is to include the time when the analysis is occurred. The question here: when analysis should be processed? (See Figure 6)

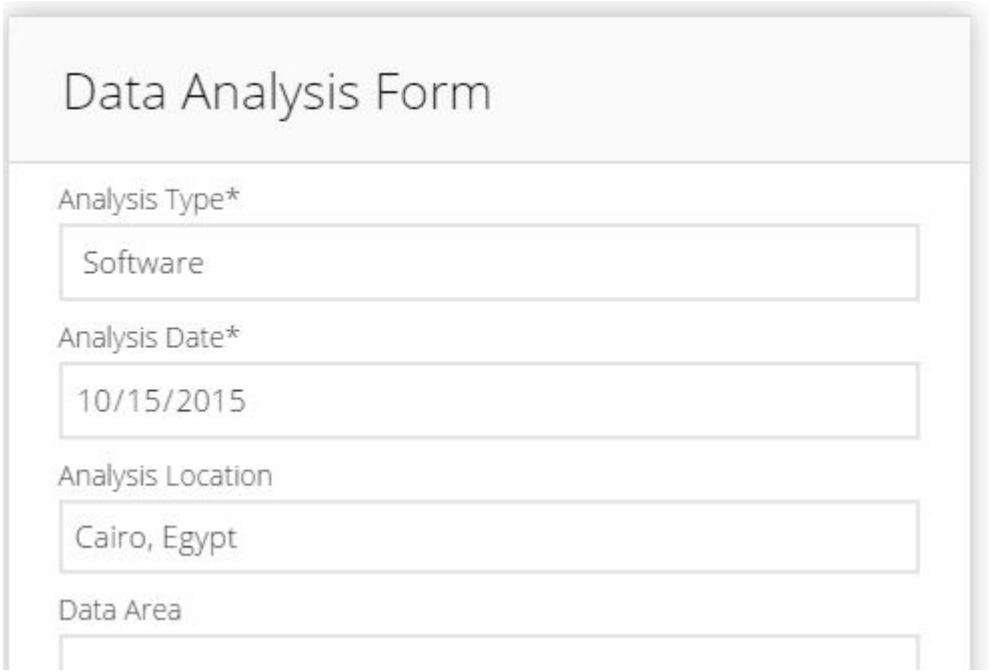

Figure 5: Framework supporting application: Data Analysis Form

2.6 Location Data Analysis

2.6.1 Location As Data

As all types of data, location-based data is vital to decision maker as some decision could be based on specific country, city or distract and those kind of data will be analyzed. To clarify this let's assume that a company need to sell a product so the advertisements in Paris will be different than in Dubai, the language, style, media used [TV, Street, ...], and so on ... so the goal here is to include market or location data inside the analysis cycle. The question here: Where will this decision be happened? (See Figure 7)

2.6.2 Location As Analysis Item

The difference here is that location will not concern about the location of decision, instead it will be concern about where the analysis will be occurred the surrounding environment while analysis will be. To clarify this let's assume that a company need to sell a product in Paris so what I concern about here where are the decision makers while they making decisions, California, London, Moscow and so on ... so the goal here is to include the location where the analysis is occurred. The question here: Where analysis should be processed?

3 A Framework of Application

The proposed framework is meant for designing and developing applications supporting strategic business analytics based on many methods determined by decision maker, it might be a fixed number,

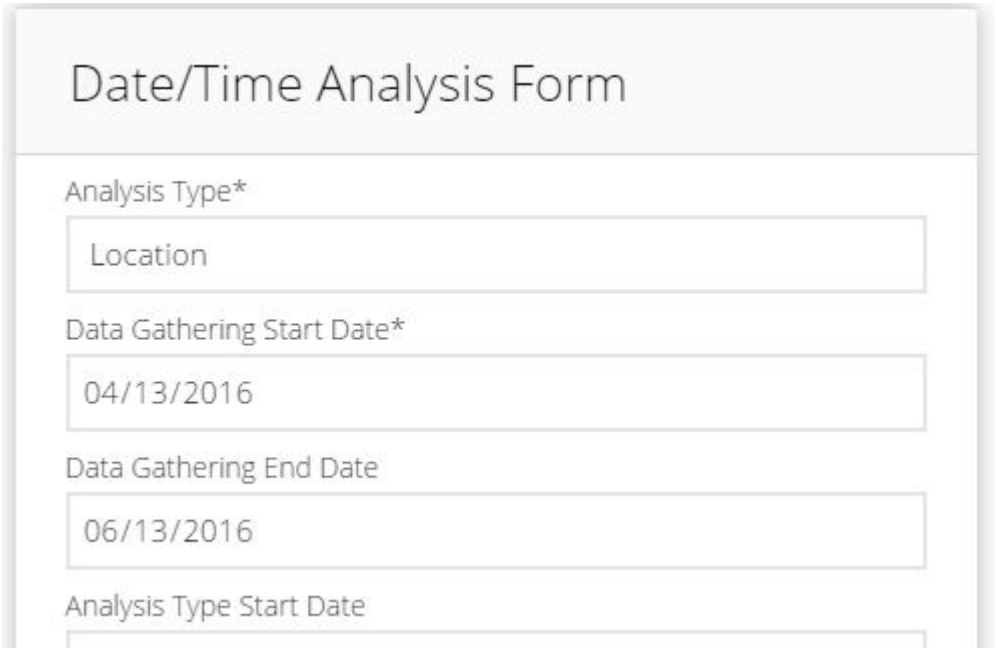

Figure 6: Framework supporting application: Date-Time Analysis Form

random number or using of dynamic fuzzy logic [11].

3.1. Steps to Run Application:

3.1.1. Decision maker will set minimum and maximum value for each factor by default minimum is equal to 0 and maximum is equal to 100.

3.1.2. According to method used, application will determine factor total accuracy.

3.1.3. Application will calculate the decision impact accuracy of all six factors.

3.1.4. Total accuracy risk will formulate "Decision Accuracy Result" chart, providing the decision maker with graph describe the total accuracy of analysis and the risk of making a decision based on gathered and analyzed information.

3.1.5. Accuracy of each factor will be formulated by the equation of Total Factor Accuracy (TFA) divided by the number of rows (FR) have been inserted against this factor. Accuracy of Factor = (TFA / FR) / 100.

3.2. The Class Diagram for the base data model, which should be extended to include data for the particular application being developed with the framework (see Figure 8).

3.3. The accuracy result chart is formulate based on the input data provided by the used methods (see Figures 9, 10, and Table 1).

3.4. For any item, total item impact will be Total Accuracy of Average Factor Result/Total Impact Results (See Figure 11). For Software, Total Accuracy of Average Total Result = 15/Total Impact Results=225=.06666.

So, the percentage of impact will be $0.06666 \times 100 = 6.7\%$.

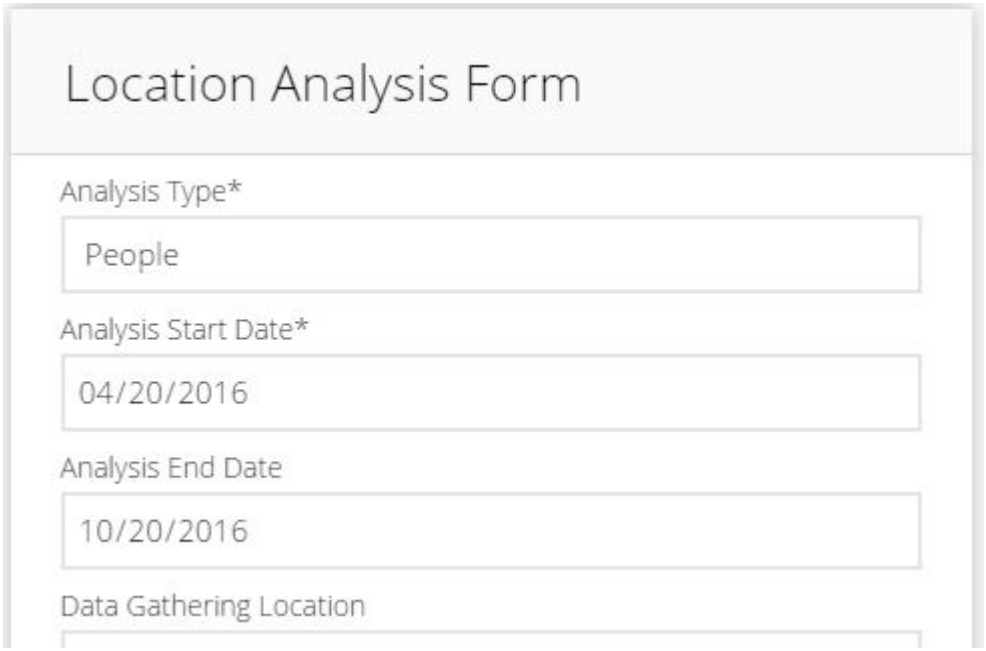

Figure 7: Framework supporting application: Location Analysis Form

Table 1: Framework supporting application, data table

Factor	Rows	Factor Accuracy (FA)	Avg. Factor Accuracy
Software	1	20	(20/1) = 20
	2	10	(30/2) = 15
Factor Accuracy			15 %
Hardware	1	30	(30/1) = 30
	2	30	(60/2) = 30
	3	30	(90/3) = 30
Factor Accuracy			30%
People	1	60	(60/1) = 60
	2	90	(150/2) = 75
Factor Accuracy			75%
Data	1	5	(5/1) = 5
Factor Accuracy			5%
Date-Time	1	0	0
Factor Accuracy			0%
Location	1	100	(100/1) = 100
Factor Accuracy			100%
		Total Impact Results	225

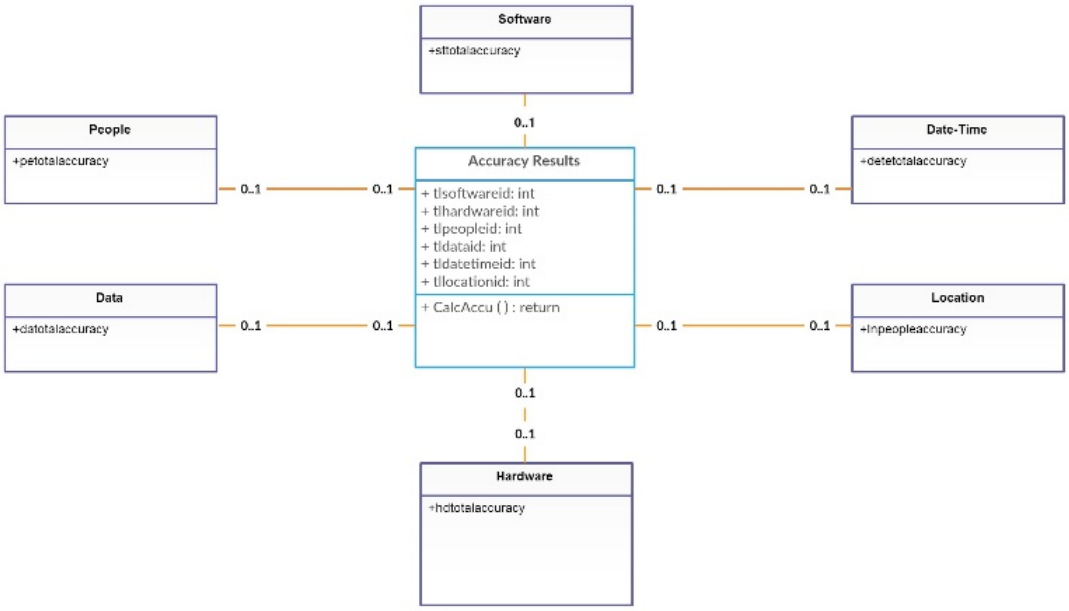

Figure 8: Framework supporting application: Class Diagram

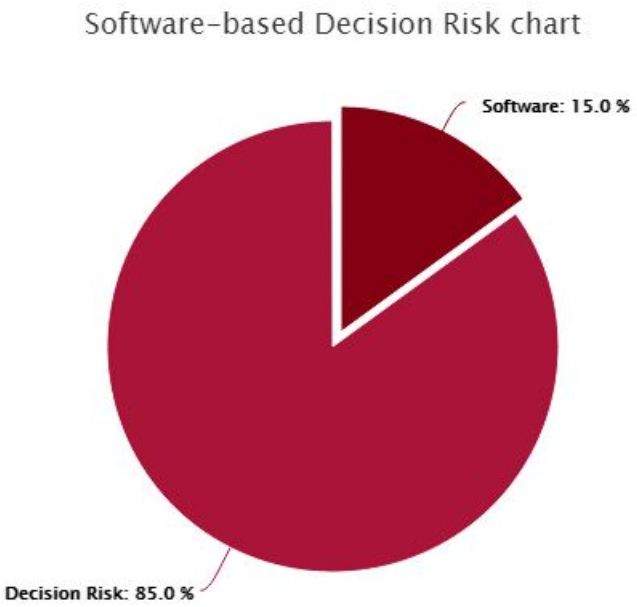

Figure 9: Framework supporting application: Software Analysis Chart

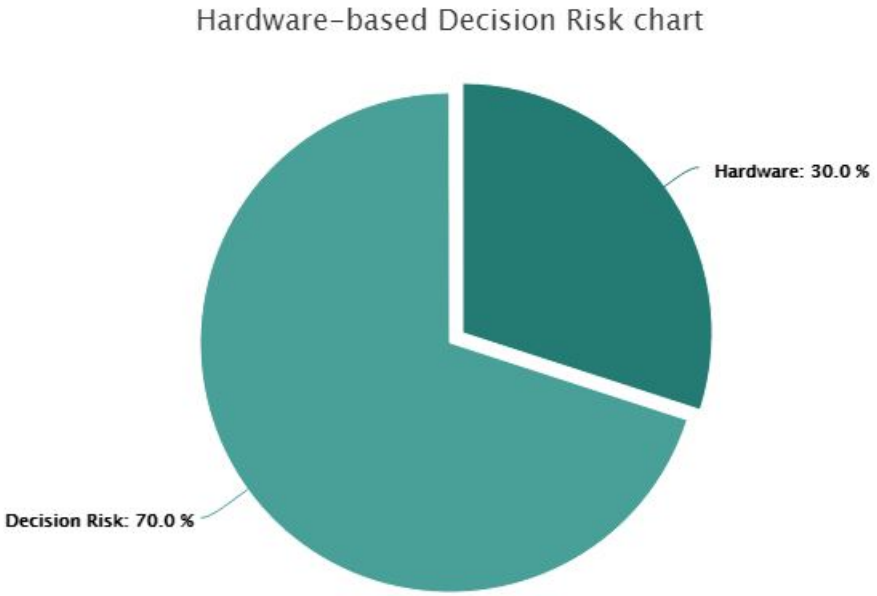

Figure 10: Framework supporting application: Hardware Analysis Chart

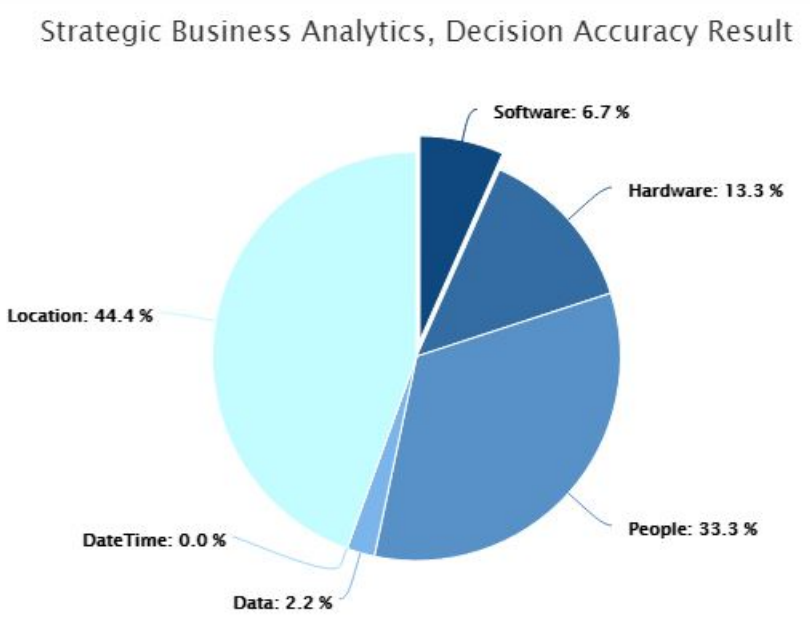

Figure 11: Framework supporting application: Items Impact Chart

4 Conclusions

Current decision making process is focusing on gathered and analyzed data, without considering the information team, analysis team or applications used to do that. In this work, we propose a new and

different perspective to make a better decision, we propose to go one step further by taking advantage of not only the traditional perspective of gathering and analyzing information but also enhanced level of information about software, hardware, data and people that are traditionally used in decision making process as well. Although this method may cause in a huge amount of information about things that decisions will be made for, and information about everything will be used to gather and analyze information to decision makers, but they can determine the reliability of information as well as the risk of making this decision based on the given information.

References

[1] M. Nooraie, "Factors influencing strategic decision-making processes," *International Journal of Academic Research in Business and Social Sciences*, vol. 2, no. 7, pp. 405–429, July 2012.
[2] M. Nooraie, "Organizational slack and decision-making process output," *Journal of Malaysian Management Review*, vol. 42, pp. 105–118, 2007.
[3] H. J. Rosenblatt, *Systems Analysis and Design*, Cengage Learning, 2013.
[4] O. Ytanyi, U. J. F. Ewurum, W. I. Upere, "Evaluation of decision making criteria with special references to quantitative and qualitative paradigms," *African Journal of Business Management*, vol. 44, no. 6, pp. 1110–1117, 2012.
[5] L. McGregor, "Improving the quality and speed of decision making," *Journal of Change Management*, vol. 2, no. 4, pp. 344–356, 2010.
[6] P. F. Drucker, *The Effective Executive*, New York, NY: HarperCollins, pp. 27, 2009.
[7] A. J. DuBrin, *Leadership: Research Findings, Practice and Skills*, Mason, OH, Cengage South-Western, 2012.
[8] J. R. Schermerharm, J. G. Hunt, R. N. Osborn, *Organizational Behaviour (11th ed.)*, New York, NY: Wiley, 2011.
[9] A. Jacobs, "The pathologies of big data," *Communications of the ACM*, vol. 52, no. 8, pp. 36–44, 2009.
[10] A. Cuzzocrea, I. Y. Song, K. C. Davis, "Analytics over Large-Scale Multidimensional Data: The Big Data Revolution!," *Proceedings of the ACM 14th International Workshop on Data Warehousing and OLAP (DOLAP'11)*, New York, USA, pp. 101–104, 2011.
[11] O. Castillo, H. Neyoy, J. Soria, P. Melin, F. Valdez, "A new approach for dynamic fuzzy logic parameter tuning in ant colony optimization and its application in fuzzy control of a mobile robot," *Applied Soft Computing*, vol. 28, pp. 150–159, 2015.

Biography

Ahmed Hamdi graduated from Faculty of Computer and Information Sciences, Mansoura University, Mansoura, Egypt in 2011. Ahmed received his Bachelor degree in Information Systems. Ahmed is a senior quality control engineer and used to be a teaching assistant in Misr University of Science and Technology. Ahmed main research interests are in the areas of strategic business and enhancing decisions making.

Samir Abdelrazek has received his M.Sc degree (2008) in information system, Minufiya University, Egypt. He has completed his PhD degree (2013) in information systems, Mansoura University, Egypt. Currently he is a senior lecture at Information System Department, Mansoura University. He is involved in teaching several courses for both under and postgraduate students. Dr. Abdelrazek research is focusing on Artificial Intelligence, Knowledge Management, Databases, Data Mining, Health Information

Technology, Human-Centered Computing, Software Engineering.

Ahmed AbuElfotoh is full professor at the Faculty of Computer Science and Information Systems - Mansoura University - Egypt. He is the vice dean for higher studies, research, and cultural affairs. His research interests include E-Business, GIS, and Information Systems.

Hazem El-Bakry (Mansoura, EGYPT 20-9-1970) received B.Sc. degree in Electronics Engineering, and M.Sc. in Electrical Communication Engineering from the Faculty of Engineering, Mansoura University - Egypt, in 1992 and 1995 respectively. Dr. El-Bakry received Ph. D degree from University of Aizu-Japan in 2007. Currently, he is associate professor at the Faculty of Computer Science and Information Systems -Mansoura University-Egypt. His research interests include neural networks, pattern recognition, image processing, biometrics, cooperative intelligent systems and electronic circuits. In these areas, he has published many papers in major international journals and refereed international conferences. According to academic measurements, now the total number of citations for his publications is 2909. The H-index of his publications is 28. Dr. El-Bakry has the United States Patent No. 20060098887, 2006. Furthermore, he is associate editor for journal of computer science and network security (IJCSNS) and journal of convergence in information technology (JCIT). In addition, he is a referee for IEEE Transactions on Signal Processing, Journal of Applied Soft Computing, the International Journal of Machine Graphics and Vision, the International Journal of Computer Science and Network Security, WASET Journals, WSEAS Journals and many different international conferences organized by IEEE. Moreover, he has been awarded the Japanese Computer and Communication prize in April 2006 and the best paper prize in two conferences cited by ACM. He has also been awarded Mansoura university prize for scientific publication in 2010 and 2011. Dr. El-Bakry has been selected in who Asia 2006 and BIC 100 educators in Africa 2008.

Computational Error Analysis of Two Schemes for Outsourcing Matrix Computations

Lihua Liu[1], Zhengjun Cao[2], Chong Mao[2], Jinbo Wang[3]

(Corresponding author: Zhengjun Cao)

Department of Mathematics, Shanghai Maritime University, Shanghai 201306, China[1].
Department of Mathematics, Shanghai University, Shanghai 200444, China[2].
Science and Technology on Communication Security Laboratory, Chengdu 610041, China[3].
(Email: caozhj@shu.edu.cn)

(Received June 22, 2017; revised and accepted July 30 & Aug. 3, 2017)

Abstract

Cryptographic schemes are usually computed over finite fields or rings in order to protect the user's privacy, because of computations over finite fields or rings are not order-keeping, which are useful to obscure and dissipate the redundancies in a plaintext message. Recently, Lei et al. proposed two schemes [IEEE Transaction on Cloud Computing, 78-87 (2013), Information Sciences, 205-217 (2014)] for outsourcing matrix inversion computation and matrix multiplication computation to the cloud. All computations in the schemes are computed over the infinite field $\mathbb{R}$. In this paper, we would like to stress that both two schemes are flawed because the computational errors are not considered at all. The proposed verifying equations do not hold. We will analyze the involved rounding errors and revise the original checking mechanisms. We think it is helpful to reinforce the differences between the arithmetic over $\mathbb{R}$ and that over any finite fields or rings.

Keywords: Cloud Computing; Matrix Inversion; Matrix Multiplication; Rounding Error

1 Introduction

It is well known that computational error analysis is a serious topic and plays a key role in numerical computation [12]. However, it is rarely considered in cryptography because most cryptographic schemes are computed over finite fields or rings. The computations over finite fields or rings are not order-keeping which are useful to obscure and dissipate the redundancies in a plaintext message.

With the development of cloud computing, some scientific and engineering applications (data mining, computational financing, and many other computational and data-intensive activities) could be outsourced via the Internet to make use of the massive resources of computing and storage systems. Cloud computing [18] makes it possible to enable customers with limited computational resources to outsource large-scale computational tasks to the cloud, including paring computations, linear equations (LE), linear programming (LP), matrix multiplication computation and matrix inversion computation.

In 2010, Chevallier-Mames et al. [9] considered the problem of outsourcing pairing computations. This work was followed by [1, 4, 8, 17]. In 2011, Dreier and Kerschbaum [10] put forth a method for

secure outsourcing of LP using affine transformation. Wang et al. [22] also presented a scheme for outsourcing of LP based on other transformations. In 2014, Nie et al. [19] proposed another scheme for outsourcing of LP based on the same transformation as that used in Ref.[22]. However, both two LP outsourcing schemes are flawed [3]. Wang et al. [23] have ever proposed a scheme for outsourcing linear equations to cloud. But the scheme [23] fails because the involved homomorphic encryption system [2, 20] is incompatible with the finite field $\mathbb{R}$. In 2014, Chen et al. proposed two computation outsourcing schemes for LE and for LP [6, 7]. But the schemes are insecure because the technique of masking a vector with a diagonal matrix is vulnerable to statistical analysis attacks [5, 11]. Salinas et al. [21] proposed a scheme for outsourcing LE, which makes use of the conjugate gradient method to solve the equivalent quadratic program in the client-server scenario. In 2016, Hsien et al. [13, 16] presented two surveys of public auditing for secure data storage in cloud computing.

Recently, Lei et al. proposed two schemes [15, 14] for outsourcing matrix inversion computation and matrix multiplication computation. All computations in the schemes are computed over the infinite field $\mathbb{R}$. We find both two schemes are flawed because the computational errors over $\mathbb{R}$, especially rounding errors, are not considered at all. The proposed verifying equations do not hold. That means the client cannot check the correctness of the returned values. We will analyze the involved rounding errors and revise the checking mechanisms.

2 Preliminaries

Suppose that a floating-point number system is characterized by four integers [12]: base χ, precision p, exponent range $[L, U]$. Then its accuracy can be characterized by a quantity known as *machine precision*, ϵ. If a given real number x is not exactly representable as a floating-point number, then it must be approximated by some "nearby" floating-point number. The process of choosing $\mathrm{fl}(x)$ to approximate x is called rounding, and the error introduced by such an approximation is called rounding error.

Consider the simple computation $x(y+z)$. In floating-point arithmetic we have

$$\begin{aligned}
\mathrm{fl}(y+z) &= (y+z)(1+\theta_1), \quad \text{with } |\theta_1| \le \epsilon, \\
\mathrm{fl}(x(y+z)) &= (x((y+z)(1+\theta_1)))(1+\theta_2), \text{ with } |\theta_2| \le \epsilon \\
&= x(y+z)(1+\theta_1+\theta_2+\theta_1\theta_2) \\
&\approx x(y+z)(1+\theta_1+\theta_2) \\
&= x(y+z)(1+\theta), \text{ with } |\theta| = |\theta_1+\theta_2| \le 2\epsilon.
\end{aligned}$$

3 Computational error analysis of Lei et al.'s scheme for matrix multiplication computation

3.1 Review of the scheme

Let $\mathbf{X}(i,j), x_{i,j}$ or x_{ij} denote the entry in ith row and jth column in matrix $\mathbf{X}$. Define $\delta_{x,y} = 1$, if $x = y$; $\delta_{x,y} = 0$, if $x \neq y$. Given a matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ and a matrix $\mathbf{Y} \in \mathbb{R}^{n\times s}$, the resource-constrained client wants to securely outsource the computation of $\mathbf{XY}$ to the cloud. The scheme [14] can be described as follows (see Table 1).

The correctness of this procedure can be argued as follows. Set the matrixes $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ as

$$\mathbf{P}_1(i,j) = \alpha_i \delta_{\pi_1(i),j}, \quad \mathbf{P}_2(i,j) = \beta_i \delta_{\pi_2(i),j}, \quad \mathbf{P}_3(i,j) = \gamma_i \delta_{\pi_3(i),j}.$$

Table 1: Lei et al.'s scheme for outsourcing matrix multiplication computation

Client	Server
Setup. Pick three sets of random numbers $\{\alpha_1, \cdots, \alpha_m\}$, $\{\beta_1, \cdots, \beta_n\}$ $\{\gamma_1, \cdots, \gamma_s\}$. Generate three random permutations: π_1 over $\{1, \cdots, m\}$, π_2 over $\{1, \cdots, n\}$, π_3 over $\{1, \cdots, s\}$. Set them as the secret key.	
Input $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{Y} \in \mathbb{R}^{n \times s}$. *Transformation* $\mathbf{X}'(i,j) = (\alpha_i/\beta_j)\mathbf{X}(\pi_1(i), \pi_2(j))$, $\mathbf{Y}'(i,j) = (\beta_i/\gamma_j)\mathbf{Y}(\pi_2(i), \pi_3(j))$. *Outsourcing* Send $\mathbf{X}', \mathbf{Y}'$ to the server.	Compute $\mathbf{Z}' = \mathbf{X}'\mathbf{Y}'$ and return it to the client.
Composition $\mathbf{Z}(i,j) = \left(\gamma_{\pi_3^{-1}(j)}/\alpha_{\pi_1^{-1}(i)}\right)\mathbf{Z}'(\pi_1^{-1}(i), \pi_3^{-1}(j))$. *Verification* Pick an $s \times 1$ random 0/1 vector $\mathbf{r}$. Check that $\mathbf{X}(\mathbf{Y}\mathbf{r}) - \mathbf{Z}\mathbf{r} \stackrel{?}{=} (0, \cdots, 0)^T$. Repeat the process l times. *Output* $\mathbf{Z}$.	

Then

$$\mathbf{P}_1^{-1}(i,j) = (\alpha_j)^{-1}\delta_{\pi_1^{-1}(i),j},\ \mathbf{P}_2^{-1}(i,j) = (\beta_j)^{-1}\delta_{\pi_2^{-1}(i),j},\ \mathbf{P}_3^{-1}(i,j) = (\gamma_j)^{-1}\delta_{\pi_3^{-1}(i),j}.$$

Hence,

$$\mathbf{P}_1\mathbf{X}\mathbf{P}_2^{-1} = (\alpha_i/\beta_j)\mathbf{X}(\pi_1(i), \pi_2(j)) = \mathbf{X}'(i,j),$$
$$\mathbf{P}_2\mathbf{Y}\mathbf{P}_3^{-1} = (\beta_i/\gamma_j)\mathbf{Y}(\pi_2(i), \pi_3(j)) = \mathbf{Y}'(i,j),$$
$$\mathbf{P}_1^{-1}\mathbf{Z}'\mathbf{P}_3 = \left(\gamma_{\pi_3^{-1}(j)}/\alpha_{\pi_1^{-1}(i)}\right)\mathbf{Z}'(\pi_1^{-1}(i), \pi_3^{-1}(j)) = \mathbf{Z}(i,j).$$

Since $\mathbf{X}'\mathbf{Y}' = \mathbf{Z}'$, we have

$$\mathbf{Z} = \mathbf{P}_1^{-1}\mathbf{X}'\mathbf{Y}'\mathbf{P}_3 = \mathbf{P}_1^{-1}\mathbf{P}_1\mathbf{X}\mathbf{P}_2^{-1}\mathbf{P}_2\mathbf{Y}\mathbf{P}_3^{-1}\mathbf{P}_3 = \mathbf{X}\mathbf{Y}.$$

Thus, $\mathbf{X}(\mathbf{Y}\mathbf{r}) = \mathbf{Z}\mathbf{r}$.

Unfortunately, the above reasoning process is true only in some symbolic computing environments. But in a practical floating-point number system, computational errors involved in the above procedure should be considered carefully.

3.2 The checking mechanism fails

Let $\mathbf{X} = (x_{ij})_{m\times n}$, $\mathbf{Y} = (y_{ij})_{n\times s}$, $\mathbf{r} = (r_1, \cdots, r_s)^T$. Then the first entry of $\mathbf{X}(\mathbf{Yr})$ is

$$\begin{aligned}\Sigma_{j=1}^{n} x_{1j}(\Sigma_{i=1}^{s} y_{ji} r_i) &= x_{11}(y_{11}r_1 + y_{12}r_2 + \cdots + y_{1s}r_s)\\ &\quad + x_{12}(y_{21}r_1 + y_{22}r_2 + \cdots + y_{2s}r_s)\\ &\quad + \cdots\\ &\quad + x_{1n}(y_{n1}r_1 + y_{n2}r_2 + \cdots + y_{ns}r_s)\\ &= r_1\Sigma_{i=1}^{n} x_{1i}y_{i1} + \cdots + r_s\Sigma_{i=1}^{n} x_{1i}y_{is}\end{aligned}$$

Since

$$\begin{aligned}\mathbf{Z}(1,1) &= \frac{\gamma_{\pi_3^{-1}(1)}}{\alpha_{\pi_1^{-1}(1)}}\mathbf{Z}'(\pi_1^{-1}(1), \pi_3^{-1}(1)) = \frac{\gamma_{\pi_3^{-1}(1)}}{\alpha_{\pi_1^{-1}(1)}}\left[\Sigma_{k=1}^{n} X'(\pi_1^{-1}(1), k)Y'(k, \pi_3^{-1}(1))\right]\\ &= \frac{\gamma_{\pi_3^{-1}(1)}}{\alpha_{\pi_1^{-1}(1)}}\left[\left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_1}x_{1,\pi_2(1)}\right)\left(\frac{\beta_1}{\gamma_{\pi_3^{-1}(1)}}y_{\pi_2(1),1}\right) + \right.\\ &\quad \left. \cdots + \left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_n}x_{1,\pi_2(n)}\right)\left(\frac{\beta_n}{\gamma_{\pi_3^{-1}(1)}}y_{\pi_2(n),1}\right)\right]\end{aligned}$$

$$\vdots$$

$$\begin{aligned}\mathbf{Z}(1,s) &= \frac{\gamma_{\pi_3^{-1}(s)}}{\alpha_{\pi_1^{-1}(1)}}\mathbf{Z}'(\pi_1^{-1}(1), \pi_3^{-1}(s)) = \frac{\gamma_{\pi_3^{-1}(s)}}{\alpha_{\pi_1^{-1}(1)}}\left[\Sigma_{k=1}^{n} X'(\pi_1^{-1}(1), k)Y'(k, \pi_3^{-1}(s))\right]\\ &= \frac{\gamma_{\pi_3^{-1}(s)}}{\alpha_{\pi_1^{-1}(1)}}\left[\left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_1}x_{1,\pi_2(1)}\right)\left(\frac{\beta_1}{\gamma_{\pi_3^{-1}(s)}}y_{\pi_2(1),1}\right) + \right.\\ &\quad \left. \cdots + \left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_n}x_{1,\pi_2(n)}\right)\left(\frac{\beta_n}{\gamma_{\pi_3^{-1}(s)}}y_{\pi_2(n),1}\right)\right]\end{aligned}$$

the first entry of $\mathbf{Zr}$ is

$$\begin{aligned}&\frac{\gamma_{\pi_3^{-1}(1)}}{\alpha_{\pi_1^{-1}(1)}}\left[\left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_1}x_{1,\pi_2(1)}\right)\left(\frac{\beta_1}{\gamma_{\pi_3^{-1}(1)}}y_{\pi_2(1),1}\right) + \cdots + \left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_n}x_{1,\pi_2(n)}\right)\left(\frac{\beta_n}{\gamma_{\pi_3^{-1}(1)}}y_{\pi_2(n),1}\right)\right] r_1\\ +\ &\cdots\\ +\ &\frac{\gamma_{\pi_3^{-1}(s)}}{\alpha_{\pi_1^{-1}(1)}}\left[\left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_1}x_{1,\pi_2(1)}\right)\left(\frac{\beta_1}{\gamma_{\pi_3^{-1}(s)}}y_{\pi_2(1),1}\right) + \cdots + \left(\frac{\alpha_{\pi_1^{-1}(1)}}{\beta_n}x_{1,\pi_2(n)}\right)\left(\frac{\beta_n}{\gamma_{\pi_3^{-1}(s)}}y_{\pi_2(n),1}\right)\right] r_s.\end{aligned}$$

Since $\{\alpha_1, \cdots, \alpha_m\}$, $\{\beta_1, \cdots, \beta_n\}$ $\{\gamma_1, \cdots, \gamma_s\}$ are randomly chosen in $\mathbb{R}$, the total rounding error in the above equation approximates to $\bar{x}_1\bar{y}_1 ns\epsilon$, where $\bar{x}_1 = \frac{1}{n}\Sigma_{i=1}^{n} x_{1i}$, $\bar{y}_1 = \frac{1}{n}\Sigma_{j=1}^{n} y_{j1}$, and ϵ is the machine precision. Therefore, the practical computational result is

$$\mathbf{X}(\mathbf{Yr}) - \mathbf{Zr} = (\bar{x}_1\bar{y}_1 ns\epsilon, \cdots, \bar{x}_m\bar{y}_m ns\epsilon)^T.$$

Thus, the original checking mechanism in the scheme fails. The authors did not pay more attentions to the differences between the arithmetic over the infinite field $\mathbb{R}$ and that over any finite fields or rings.

3.3 Revision

To fix the scheme, the client has to check that

$$\left|\frac{v_i}{u_i} - 1\right| \leq \lambda n s \epsilon, \quad i = 1, \cdots, m$$

where

$$\mathbf{X}(\mathbf{Yr}) = (u_1, \cdots, u_m)^T, \quad \mathbf{Zr} = (v_1, \cdots, v_m)^T,$$

and λ is a fault-tolerant parameter. If all m inequalities are true, then output $\mathbf{Z}$.

4 Computational error analysis of Lei et al.'s scheme for matrix inversion computation

4.1 Review of the scheme

Given a matrix $\mathbf{X} \in \mathbb{R}^{n \times n}$, the resource-constrained client wants to securely outsource the computation of $\mathbf{X}^{-1}$ to the cloud. The scheme [15] can be described as follows (see Table 2).

Table 2: Lei et al.'s scheme for outsourcing matrix inversion computation

Client	Server
Setup. Pick two sets of random numbers $\{\alpha_1, \cdots, \alpha_n\}$, $\{\beta_1, \cdots, \beta_n\}$. Generate two random permutations: π_1, π_2 over $\{1, \cdots, n\}$, Set them as the secret key.	
Input $\mathbf{X} \in \mathbb{R}^{n \times n}$. *Transformation* $\mathbf{Y}(i,j) = (\alpha_i/\beta_j)\mathbf{X}(\pi_1(i), \pi_2(j))$, *Outsourcing* Send $\mathbf{Y}$ to the server.	
	Compute $\mathbf{R}' = \mathbf{Y}^{-1}$ and return it to the client.
Composition $\mathbf{R}(i,j) = \left(\alpha_{\pi_1^{-1}(j)}/\beta_{\pi_2^{-1}(i)}\right)\mathbf{R}'(\pi_2^{-1}(i), \pi_1^{-1}(j))$. *Verification* Pick an $n \times 1$ random 0/1 vector $\mathbf{r}$. Check that $\mathbf{R}(\mathbf{Xr}) - \mathbf{Ir} \stackrel{?}{=} (0, \cdots, 0)^T$. Repeat the process l times. *Output* $\mathbf{R}$.	

The correctness of the scheme can be easily argued. Set the matrixes $\mathbf{P}_1, \mathbf{P}_2$ as

$$\mathbf{P}_1(i,j) = \alpha_i \delta_{\pi_1(i),j}, \quad \mathbf{P}_2(i,j) = \beta_i \delta_{\pi_2(i),j}.$$

Then

$$\mathbf{P}_1^{-1}(i,j) = (\alpha_j)^{-1} \delta_{\pi_1^{-1}(i),j}, \quad \mathbf{P}_2^{-1}(i,j) = (\beta_j)^{-1} \delta_{\pi_2^{-1}(i),j}.$$

Hence,

$$\mathbf{Y} = \mathbf{P}_1\mathbf{X}\mathbf{P}_2^{-1},\ \mathbf{R} = \mathbf{P}_2^{-1}\mathbf{R}'\mathbf{P}_1 = \mathbf{P}_2^{-1}(\mathbf{P}_1\mathbf{X}\mathbf{P}_2^{-1})^{-1}\mathbf{P}_1 = \mathbf{X}^{-1}.$$

4.2 The checking mechanism fails

Let $\mathbf{X} = (x_{ij})_{n\times n}$, $\mathbf{R} = (\lambda_{ij})_{n\times n}$, $\mathbf{r} = (r_1, \cdots, r_n)^T$. Then the first entry of $\mathbf{R}(\mathbf{X}\mathbf{r})$ is

$$\Sigma_{i=1}^n \lambda_{1i}(\Sigma_{j=1}^n x_{ij}r_j) = r_1\Sigma_{j=1}^n \lambda_{1j}x_{j1} + \cdots + r_n\Sigma_{j=1}^n \lambda_{1j}x_{jn}.$$

Since $\lambda_{ij} = \left(\alpha_{\pi_1^{-1}(j)}/\beta_{\pi_2^{-1}(i)}\right)\mathbf{R}'(\pi_2^{-1}(i), \pi_1^{-1}(j))$, we have

$$\begin{aligned}\Sigma_{i=1}^n \lambda_{1i}(\Sigma_{j=1}^n x_{ij}r_j) \quad = \quad & r_1\Sigma_{j=1}^n \left(\alpha_{\pi_1^{-1}(j)}/\beta_{\pi_2^{-1}(1)}\right)\mathbf{R}'(\pi_2^{-1}(1), \pi_1^{-1}(j))x_{j1} \\ & + \cdots \\ & + r_n\Sigma_{j=1}^n \left(\alpha_{\pi_1^{-1}(j)}/\beta_{\pi_2^{-1}(1)}\right)\mathbf{R}'(\pi_2^{-1}(1), \pi_1^{-1}(j))x_{jn}\end{aligned}$$

Since $\{\alpha_1, \cdots, \alpha_n\}$, $\{\beta_1, \cdots, \beta_n\}$ are randomly chosen in $\mathbb{R}$, the total rounding error in the above equation approximates to $\bar{x}_1\bar{y}_1n^2\epsilon$, where $\bar{x}_1 = \frac{1}{n}\Sigma_{j=1}^n x_{j1}$, $\bar{y}_1 = \frac{1}{n}\Sigma_{j=1}^n \mathbf{R}'(\pi_2^{-1}(1), \pi_1^{-1}(j))$, and ϵ is the machine precision. Therefore, the practical computational result is

$$\mathbf{R}(\mathbf{X}\mathbf{r}) - \mathbf{I}\mathbf{r} = (\bar{x}_1\bar{y}_1n^2\epsilon, \cdots, \bar{x}_n\bar{y}_nn^2\epsilon)^T \neq (0, \cdots, 0)^T.$$

Therefore, the proposed checking mechanism in the scheme fails.

4.3 Revision

To fix the scheme, the client has to check that

$$|u_i - v_i| \leq \psi n^2\epsilon, \quad i = 1, \cdots, n$$

where $\mathbf{R}(\mathbf{X}\mathbf{r}) = (u_1, \cdots, u_n)^T$, $\mathbf{I}\mathbf{r} = (v_1, \cdots, v_n)^T$, and ψ is a fault-tolerant parameter. If all n inequalities are true, then output $\mathbf{R}$.

5 Further discussions

Though the computational errors have been considered in the above revisions, we would like to point out that they are not immune to the following inner attack. The malicious server can cheat the client to accept an inexact result. For example, in the scheme for outsourcing matrix multiplication computation the malicious server computes $\mathbf{Z}' = \mathbf{X}'\mathbf{Y}'$ and returns $\widehat{\mathbf{Z}'}$ to the client, where

$$\widehat{\mathbf{Z}'} = \mathbf{Z}' + \begin{pmatrix} \rho & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

$\rho = (\chi - 1)\epsilon$, χ is the base of the underlying floating-point number system. In such case, we have

$$\mathbf{Z}(\pi_1(1), \pi_3(1)) = (\gamma_1/\alpha_1)\,[\mathbf{Z}'(1,1) + \rho],$$

$$\mathbf{Zr} = (v_1, \cdots, v_{\pi_1(1)} + \frac{\gamma_1}{\alpha_1} r_{\pi_3(1)}\rho, \cdots, v_m)^T$$

Hence,

$$\left| \frac{v_{\pi_1(1)} + \frac{\gamma_1}{\alpha_1} r_{\pi_3(1)}\rho}{u_{\pi_1(1)}} - 1 \right| \leq \left| \frac{v_{\pi_1(1)}}{u_{\pi_1(1)}} - 1 \right| + \left| \frac{\frac{\gamma_1}{\alpha_1}\rho}{u_{\pi_1(1)}} \right|.$$

The probability of the event that

$$\left| \frac{v_{\pi_1(1)}}{u_{\pi_1(1)}} - 1 \right| + \left| \frac{\frac{\gamma_1}{\alpha_1}}{u_{\pi_1(1)}} \right| \cdot (\chi - 1)\epsilon \leq \lambda n s \epsilon$$

approximates to 1, because n, s are supposed to be sufficiently large. Hence, the malicious server can cheat the client to accept an inexact $\mathbf{Z}$, where

$$\mathbf{Z}(i,j) = \left(\gamma_{\pi_3^{-1}(j)} / \alpha_{\pi_1^{-1}(i)}\right) \widehat{\mathbf{Z}'}(\pi_1^{-1}(i), \pi_3^{-1}(j)).$$

Therefore, one has to carefully choose a fault-tolerant parameter which is appropriate for practical processes so that the accumulated errors are kept negligible.

6 Conclusion

We show that Lei et al.'s schemes for outsourcing matrix multiplication and matrix inversion computation are flawed because the computational errors are not considered carefully. We think this paper is helpful to reinforce the differences between the arithmetic over the infinite field of real numbers and that over any finite fields or rings.

Acknowledgements

We thank the National Natural Science Foundation of China (61303200, 61411146001). The authors gratefully acknowledge the reviewers for their valuable suggestions.

References

[1] S. Canard, J. Devigne, and O. Sanders, "Delegating a pairing can be both secure and efficient," in *Proceedings of Applied Cryptography and Network Security - ACNS 2014*, pp. 549–565, Lausanne, Switzerland, June 2014.

[2] Z. J. Cao and L. H. Liu, "Comment on 'harnessing the cloud for securely outsourcing large-scale systems of linear equations'," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 5, pp. 1551–1552, 2016.

[3] Z. J. Cao and L. H. Liu, "A note on two schemes for secure outsourcing of linear programming," *International Journal of Network Security*, vol. 19, no. 2, pp. 323–326, 2017.

[4] Z. J. Cao, L. H. Liu, and O. Markowitch, "On two kinds of flaws in some server-aided verification schemes," *International Journal of Network Security*, vol. 18, no. 6, pp. 1054–1059, 2016.

[5] Z. J. Cao, L. H. Liu, and O. Markowitch, "Comment on 'highly efficient linear regression outsourcing to a cloud'," *IEEE Transactions on Cloud Computing (10.1109/TCC.2017.2709299)*, 2017.

[6] F. Chen, T. Xiang, X. Lei, and J. Chen, "Highly efficient linear regression outsourcing to a cloud," *IEEE Transactions on Cloud Computing*, vol. 2, no. 4, pp. 499–508, 2014.

[7] F. Chen, T. Xiang, and Y. Y. Yang, "Privacy-preserving and verifiable protocols for scientific computation outsourcing to the cloud," *Journal of Parallel and Distributed Computing*, vol. 74, pp. 2141–2151, 2014.

[8] X. F. Chen and et al., "Efficient algorithms for secure outsourcing of bilinear pairings," *Theoretical Computer Science*, no. 562, pp. 112–121, 2015.

[9] B. Chevallier-Mames and et al., "Secure delegation of elliptic-curve pairing," in *Proceedings of Smart Card Research and Advanced Application, 9th IFIP WG 8.8/11.2 International Conference - CARDIS 2010*, pp. 24–35, Passau, Germany, April 2010.

[10] J. Dreier and F. Kerschbaum, "Practical privacy-preserving multiparty linear programming based on problem transformation," in *Proceedings of IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing - PASSAT/SocialCom 2011*, pp. 916–924, Boston, MA, USA, Oct. 2011.

[11] D. Hankerson, A. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*. USA: Springer, 2004.

[12] F. Heath, *Scientific computing, an introductory survey, second edition*. USA: McGraw-Hill Higher Education, 2001.

[13] W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for secure data storage in cloud computing," *International Journal of Network Security*, vol. 18, no. 1, pp. 133–142, 2016.

[14] X. Y. Lei, X. F. Liao, T. W. Huang, and F. Heriniaina, "Achieving security, robust cheating resistance, and high-efficiency for outsourcing large matrix multiplication computation to a malicious cloud," *Information Sciences*, vol. 280, pp. 205–217, 2014.

[15] X. Y. Lei, X. F. Liao, T. W. Huang, H. Q. Li, and C. Q. Hu, "Outsourcing large matrix inversion computation to a public cloud," *IEEE Transactions on Cloud Computing*, vol. 1, no. 1, pp. 78–87, 2013.

[16] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing," *International Journal of Network Security*, vol. 18, no. 4, pp. 650–666, 2016.

[17] L. H. Liu and Z. J. Cao, "A note on 'efficient algorithms for secure outsourcing of bilinear pairings'," *International Journal of of Electronics and Information Engineering*, vol. 5, no. 1, pp. 30–36, 2016.

[18] D. Marinescu, *Cloud Computing Theory and Practice*. USA: Elsevier, 2013.

[19] H. X. Nie, X. F. Chen, J. Li, J. Liu, and W. J. Lou, "Efficient and verifiable algorithm for secure outsourcing of large-scale linear programming," in *Proceedings of 28th IEEE International Conference on Advanced Information Networking and Applications - AINA 2014*, pp. 591–596, Victoria, BC, Canada, May 2014.

[20] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proceedings of Advances in Cryptology - EUROCRYPT 1999*, pp. 223–238, Prague, Czech Republic, May 1999.

[21] S. Salinas, C. Q. Luo, X. H. Chen, and P. Li, "Efficient secure outsourcing of large-scale linear systems of equations," in *Proceedings of 2015 IEEE Conference on Computer Communications - INFOCOM 2015*, pp. 1035–1043, Hong Kong, China, Apr. 2015.

[22] C. Wang, K. Ren, and J. Wang, "Secure and practical outsourcing of linear programming in cloud computing," in *Proceedings of 2011 IEEE Conference on Computer Communications - INFOCOM 2011*, pp. 820–828, Shanghai, China, Apr. 2011.

[23] C. Wang, K. Ren, J. Wang, and Q. Wang, "Harnessing the cloud for securely outsourcing large-scale systems of linear equations," *IEEE Transaction on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1172–1181, 2013.

Biography

Lihua Liu is an associate professor with Department of Mathematics at Shanghai Maritime University. She received her Ph.D. degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics, cryptography and information security.

Zhengjun Cao is an associate professor with Department of Mathematics at Shanghai University. He received his Ph.D. degree in applied mathematics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He served as a post-doctor in Computer Sciences Department, Université Libre de Bruxelles, from 2008 to 2010. His research interests include cryptography, discrete logarithms and quantum computation.

Chong Mao is currently pursuing his M.S. degree from Department of Mathematics, Shanghai University. His research interests include information security and cryptography.

Jinbo Wang received his Ph.D. degree in applied mathematics from Shanghai University. His research interests include applied cryptography and network security.

An Energy Efficient Grid Based Static Node Deployment Strategy for Wireless Sensor Networks

Rajeev Singh and Matendra Singh Manu
(Corresponding author: Rajeev Singh)

G. B. Pant University of Ag. & Technology, Pantnagar, Udham Singh Nagar, Uttarakhand
(Email: rajeevpec@gmail.com)

(Received June 30, 2017; revised and accepted Aug. 5, 2017)

Abstract

Wireless Sensor Network (WSN), an emerging wireless technology, involves deployment of sensing nodes in the surrounding area. A node once deployed in a WSN remains unattended and it therefore becomes critical to design a proper node deployment strategy. A proper node deployment fulfills properties like coverage, connectivity and lifetime. Such WSN deployment strategy meets two objectives: One, it simplifies the complexity in routing, data fusion, communication, etc. Two, it reduces energy consumption and hence extends the lifetime of WSNs. In this paper, we propose a grid based static node deployment strategy for WSN systems that meets these objectives.

Keywords: Coverage Map; Deterministic Node Placement; Energy Efficient Deployment; Grid Based Node Deployment; Pattern Based Deployment; WSN Coverage

1 Introduction

A Wireless Sensor Network (WSN) is a collection of spatially interconnected sensor nodes. A sensor node is a device having capabilities of converting the physical measure of its surroundings into electronic signals. The major components of a sensor node includes: microcontroller, transceiver, external memory, power source and one or more sensors. Each node in WSN depends on a limited source of energy i.e. battery, for its functionalities [6].

An effective WSN is the one which is well deployed in the environment. Positioning of the sensor nodes is done during WSN deployment in such a manner that the design specifications such as coverage, connectivity and energy consumption are met. Coverage ensures that the area of interest is covered by at least one sensor node. Proper connectivity enables direct or indirect connectivity of every sensor node to the sink node. Indirect connection is required when two sensor nodes lie outside the communication range of each other and hence cannot communicate directly. Here, a single node failure may even interrupt the entire system. Due to small capacity of sensor node battery, energy constraint has become a major issue in wireless sensor networks [2, 3, 5, 7].

This paper introduces a deterministic sensor node placement algorithm which provides better coverage and minimizes the energy consumption without affecting the connectivity in Wireless Sensor Networks. Simulation results of the proposed deployment strategy are compared with random and

deterministic node deployment strategies. The paper considers the sensor node deployment in terms of the sensing coverage, communication connectivity and system lifetime.

2 Related Work

Ringwald and Romer [4] refer deployment as setting up an operational sensor network in a real world environment where wireless sensor nodes are installed in a given area of interest. A deployment that gives better coverage and connectivity among sensor nodes to fulfill the needs of the application is considered good.

Some researchers advocate for static adjustment of the node's location which gives uniform coverage at every point within sensor field because of its deterministic architecture. Poe and Schmitt [3] presented three deployments for large-scale WSNs namely, a uniform random, a square grid, and a pattern-based Tri-Hexagon Tiling (THT) node deployment. For each deployment strategy a simple energy model is formulated for studying the energy consumption. For square grid and THT node deployment relative frequency of exactly k-covered points is calculated using the notion of k-coverage map. The k-coverage map is evaluated via basic geometry.

In Uniform Random deployment, each of the sensor nodes is scattered in the area of interest such that each has equal probability of being placed at any point. A uniform random deployment is easy as it can even result from throwing sensor nodes (say, from airplane). Thus, the scattering of nodes is done such that their locations are not known with certainty [8].

In a Square Grid pattern the given area of interest is divided into equal squares and nodes are placed at the corners. It is assumed that the sensing range is equal to the length of a cell. It is shown via a tessellation of the region formed by intersections of sensing ranges that a square grid cell has exact 2-coverage, 3-coverage and 4-coverage regions.

In Tri-Hexagon strategy tiling phenomenon is utilized where the entire plane is covered with figures which do not overlap nor leave any gaps. Tiling also referred as tessellation. A semi-regular tiling having exactly eight different tilings is utilized. In this tiling every vertex uses the same set of regular polygons having same side lengths and interior angles. Tri-Hexagon strategy considers a semi-regular tiling having six equilateral triangles and one regular hexagon where each of the tiling point hosts a node. The same approach is considered for the k-coverage map of a THT cell. The area of each equilateral triangle is fully covered by three nodes, thus having exact 3-coverage. Inside a regular hexagon, there are 3 possible exact k-coverages: 2-, 3-, and 6-coverage. This notion can be utilized further to compute the k-coverage maps for other patterns. A pattern based node deployment involves placing nodes at exact location which may not be suitable for some real application scenarios but the performance enhancements may provide justification in such cases.

Abderrahim [1] proposed Simulated Annealing (SA) based sensor deployment technique for the placement of sensors which minimizes the number of hops between sensor nodes and the sink node while maintaining appropriate coverage. Upon reduction in the number of hops the energy consumption gets reduced resulting in an increase in the lifetime of Wireless Sensor Network.

3 Proposed Node Deployment Strategy

Considering coverage, a grid-based deployment is considered as a stable deployment technique in WSNs. In this work a grid based node deployment strategy for WSNs is proposed namely 'Rhombus Grid' node deployment. Here, a rhombus is considered as a combination of two equilateral triangles. Figure 1 shows the rhombus grid deployment scheme where each crossing point of edges represents the position of sensor

node. Each neighbor node is considered within range of the sensor node. So, the deployment provides a well-connected Wireless Sensor Network.

3.1 Map for Rhombus Grid Node Deployment

The k-coverage map of all possible exactly k-covered points of a rhombus grid cell can easily be inferred via basic geometry as shown in Figure 2. In the rhombus grid cell, nodes are placed at the vertices of the rhombus. A tessellation is formed by intersections of the sensing range of the nodes. Assuming that the sensing range is equal to the length of a cell, a rhombus grid cell has exact 3-coverage and 4-coverage regions.

The white region has exact 4-coverage because it forms the intersection of sensing region of four sensor nodes. The radius of circles is the same for all sensing nodes. Some tessellations are symmetric. A rhombus grid cell has two symmetric gray-regions. The gray regions lie between the border lines of circles and have exact 3-coverage. These white and gray region areas are added for calculating the total area of exact k-coverage of a grid cell.

The Area of equilateral triangle (At) is calculated as per Equation (1), where R_{sense} is the radius of the sensing disk of a sensor node.

$$At = [(\sqrt{3/4}) \times R^2_{sense}]. \tag{1}$$

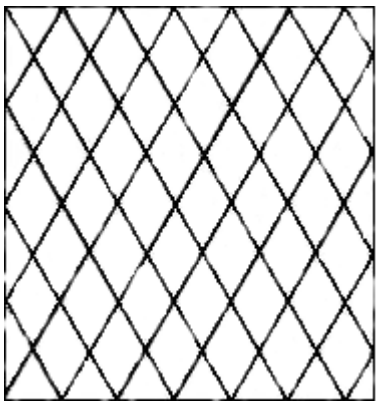

Figure 1: Proposed Rhombus Grid node deployment

The area for segment of circle is formulated as per Equation (2). This area is the difference between 6th part of circle and an equilateral triangle. Hence, gray region area (Equation (3)) is calculated as the area except segments in an equilateral triangle.

$$A1 = [(\prod R^2_{sense}/6) - At] \tag{2}$$

$$A2 = [(\sqrt{3/4}) \times R^2_{sense} - 3A1]. \tag{3}$$

Finally, total 3- and 4-coverage region areas inside a rhombus grid cell are two times A2 and six times A1 which are represented as A3 and A4, respectively.

$$A3 = [2 \times A2] \tag{4}$$

$$A4 = [6 \times A1]. \tag{5}$$

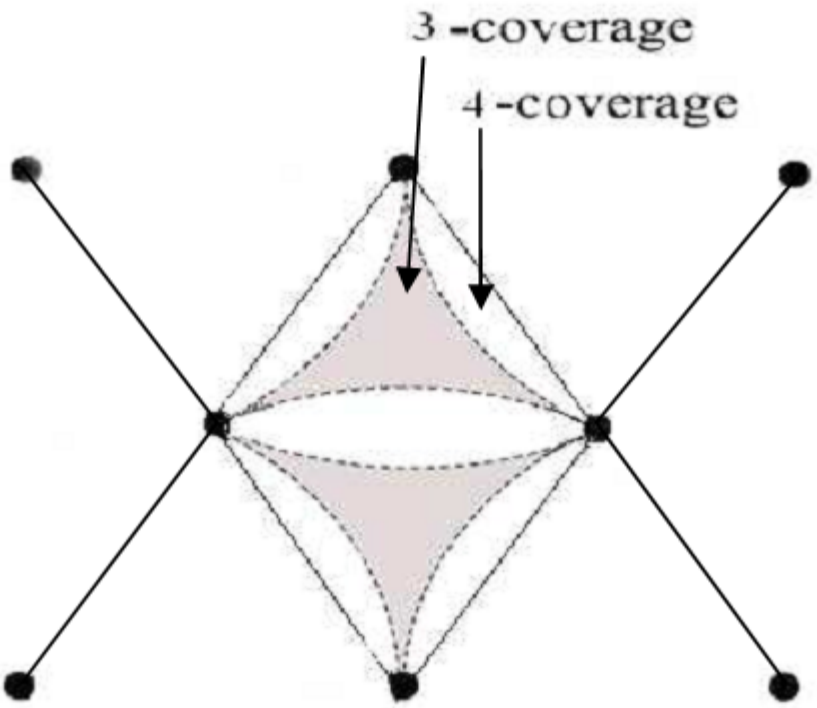

Figure 2: Rhombus Grid k-coverage map

4 Performance Metrics

The performance of all three strategies i.e. Random, Square Grid, Tri-Hexagon [3] and proposed node deployment strategy are evaluated and compared. The coverage and energy consumption metrics are utilized for performance evaluation and comparison among node deployment strategies.

Coverage Area- Coverage in WSN is related to energy saving, connectivity, and network reconfiguration. The coverage gains importance as it is associated with providing high quality of information in the region of interest. Coverage in WSN is either full or partial. In full coverage every point in the area must be covered by at least one sensor without allowing any uncovered points. Node deployment is a crucial parameter in deciding the overall coverage. For specifying conditions on coverage, K-coverage is specified which means that every point in the region or area is covered by at least k sensors.

Energy Model [3] - WSN usually utilizes multi-hop communication as direct communication from every node to sink is very expensive. Thus, transceiver energy consumption is analyzed based on a hop-to-hop transmission. In this work, a simple energy model [3] is taken for the assessment and performance comparison of node deployments in WSNs. It considers the energy consumption by the transceiver unit in calculations. The model does not consider energy consumption for security, routing, and data aggregation. The energy consumption by the processing unit of each node is taken as a constant.

The total energy consumption calculations by the selected model consider energy consumed for

transmitting 1bit data from all the source nodes to their corresponding sinks. Energy consumed for transmitting 1bit data depends on the number of hops in the data transmission path (data flow). Obviously, data flow is considered on the shortest path between a given source node and its nearest sink. The number of data flows is number of nodes per sink. The number of hops in each flow of depends on the node's location. The energy consumption per hop is dependent on its distance from sink. The energy consumption for single hop is calculated by summing energy used by a transmitter, a receiver, and a sensor for 1bit of data. Finally, sum of total energy consumption per sink is added to evaluate the overall energy consumption of the network.

5 Results

The primary factors used in the performance evaluation are: number of nodes and number of sinks. The experiments consider selection of values as per model of MICAz mote running under TinyOS. The sensor nodes utilize Dijkstra's shortest path algorithm as the routing protocol for finding the shortest path (hop distance) from source to sink.

5.1 Performance Evaluation: Coverage

In all the four cases i.e., Random, Square grid, Tri-Beehive and Rhombus deployment pertaining to coverage the x-axis in graph represents the k in k-coverage and y-axis represents percentage of covered area by k-coverage. In Square Grid and in the proposed Rhombus Grid, the area is divided into symmetric square cells. Therefore, single cell analysis is done in both these cases whereas in Tri-Beehive deployment analysis is done on the basis of total number of cells due to the combination of triangle and hexagon. Fig. 3(a)-(d) shows results for all cases related with coverage for 500 node network.

5.1.1 Random Node Deployment Coverage

Both square grid and THT do not consider boundary condition. Hence for uniformity, the boundary condition is not considered in the random deployment. Random deployment requires a systematic sampling over the deployment area. A 100 to 500 nodes sized network is considered for different scenarios (Figure 3(a)) [3].

The average coverage for random deployment turns out to be 3.00. In random deployment, exact 1- to 4- coverage covers most of the area; out of which the exact 3-coverage covers 21.6% to 23% area which is highest in random node deployment (Fig. 3(c)). An important realization is that 5% of the network is not covered by any nodes.

5.1.2 Square Grid Node Deployment Coverage

The average coverage for Square Grid deployment turns out to be 3.14. In Square Grid deployment pattern, about half of the network is covered by 3-coverage. The remaining half is covered by 2- and 4-coverage. The percentage values for the 2-, 3- and 4-coverage are 17.27%, 51.17% and 31.57% respectively (Figure 3(b)).

5.1.3 Tri-Beehive Grid Node Deployment Coverage

Tri-Beehive node deployment pattern has an average 2.81-coverage. In Square Grid deployment pattern, almost three fourth of the network has 3-coverage i.e., covered by three sensor nodes. The remaining network has 2-coverage. Ignoring negligibly small percentage of 6-coverage, the percentage values for the 2- and 3-coverage are 18.53%, 81.47% respectively (figure 3(c)).

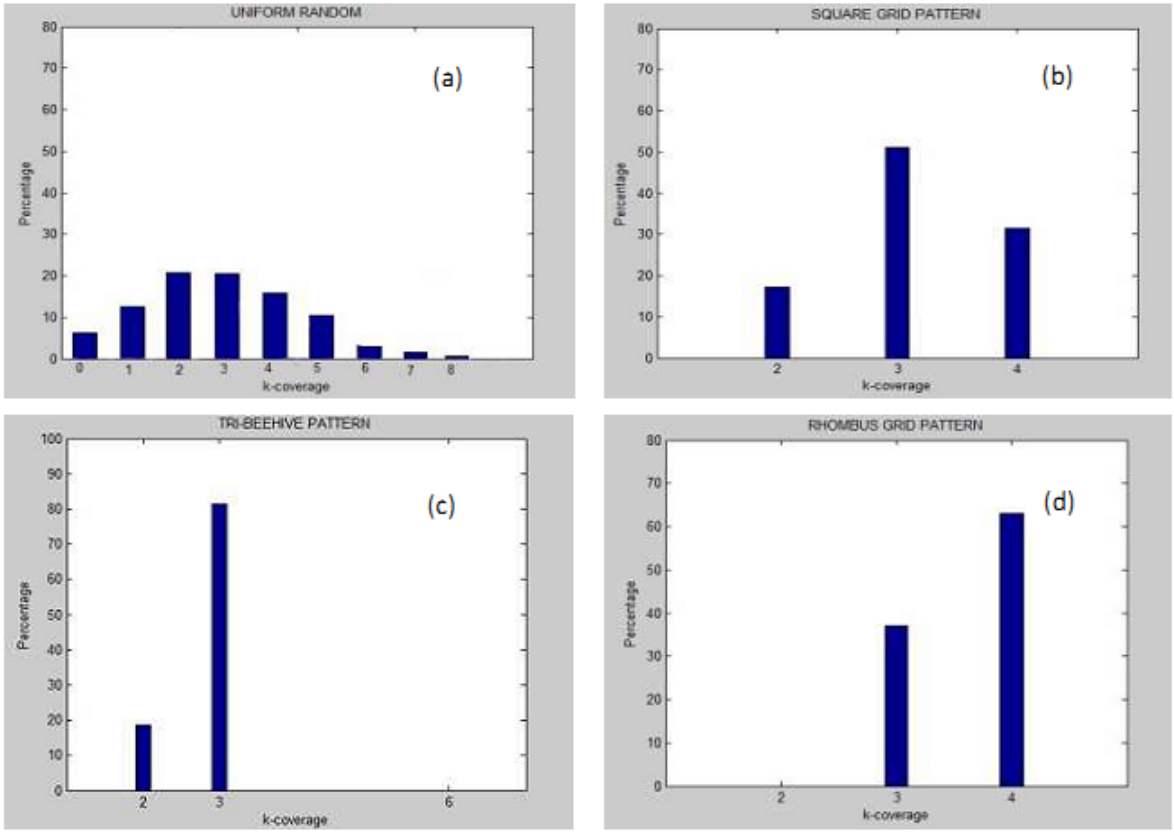

Figure 3: The exact k-coverages of (a) random, (b) square grid, (c) THT and (d) rhombus grid deployments.

5.1.4 Rhombus Grid Node Deployment Coverage

Rhombus grid node deployment pattern has an average 3.62-coverage which is highest regarding the average coverage performance metric. In this deployment, more than half of the network has 4- coverage while the remaining network has 3-coverage. No other network deployment i.e. Random, Square Grid and Tri- Beehive Grid provide such intensive coverage. The percentage values for the 3- and 4-coverage are 37.67% and 62.33% respectively (Figure 3(d)).

5.1.5 Overview of Node Deployment Strategies in Terms of Coverage

Table 1 shows the comparison between all the evaluated node deployment strategies as per 4-coverage and average k-coverage percentage. Table 1 shows that Rhombus Grid deployment strategy gives the highest i.e., 62.33% 4-coverage and highest average k-coverage i.e., 3.62-coverage. Hence, coverage is maximum in Rhombus Grid deployment.

Table 1: k-coverage values

SN	Node Deployment Strategy	4-coverage %	Avg. k-coverage
1	Random	18.23	3.00
2	Square Grid	31.57	3.14
3	Tri-Beehive Grid	-	2.81
4	Rhombus Grid	62.33	3.62

5.2 Performance Evaluation: Energy Consumption

Energy consumption for transmitting 1 bit of data by each node for each deployment strategy is considered for energy evaluation purpose. Each node uses 3V constant voltage for transmitting and receiving bits. In simulations, the transmission and sensing data rates are considered as same. Two scenarios, one with 24 nodes and the other with 100 nodes were simulated for evaluation.

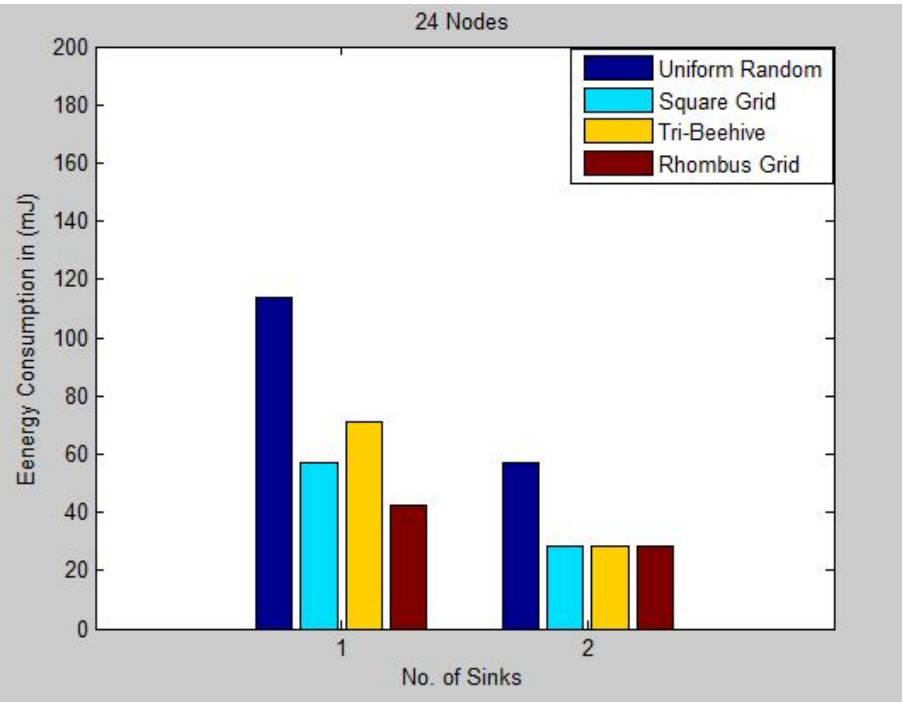

Figure 4: Bar-graph for energy consumption with 24 nodes

5.2.1 Energy Consumption with 24 Nodes Deployed

The energy consumption with 24 nodes is evaluated by taking one and two sink nodes in each node deployment strategy i.e. Uniform Random, Square Grid, Tri-Beehive and proposed Rhombus Grid. Figure 4 shows that when there is a single sink node, energy consumption is different in each node

deployment strategy. The energy consumption is least in the proposed Rhombus Grid method when there is one sink in the area. When two sink nodes are taken, energy consumption is same in all grid base deployment strategies, because the number of hops is same for each node deployment strategy with 24 nodes.

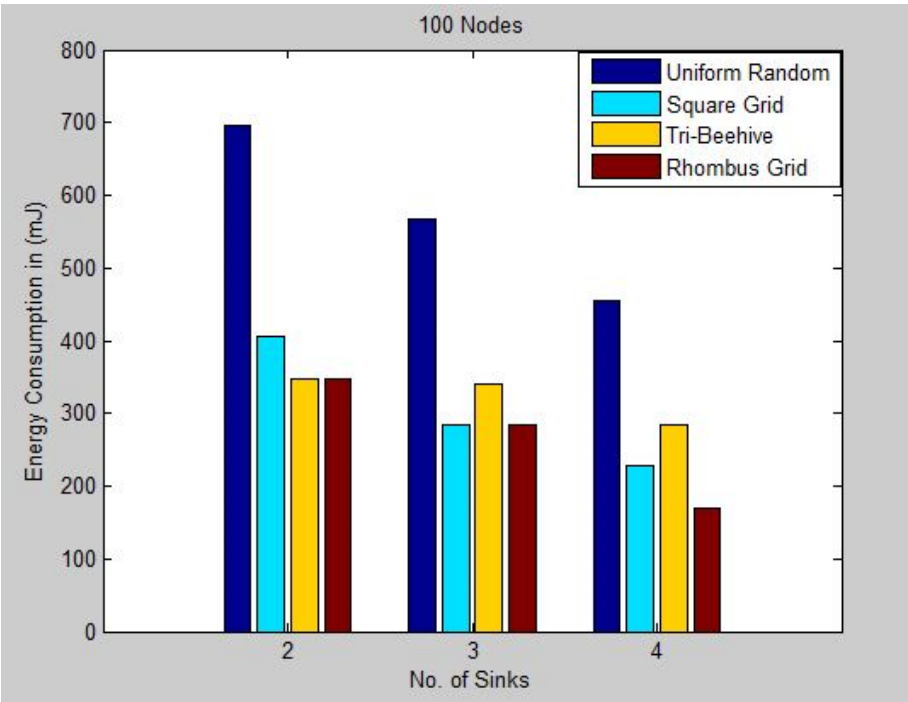

Figure 5: Bar-graph for energy consumption with 100 nodes

5.2.2 Energy Consumption with 100 Nodes Deployed

A simulation result of energy consumption with 100 nodes is shown in the Figure 5. As shown in the Figure 5, energy consumption in random node deployment is more than the grid based node deployment strategies. With increase in number of sink nodes the energy consumption is minimized but the cost of the network is increased and network overheads are also increased. In all the cases the energy consumption for proposed method is least for 100 nodes deployment.

6 Conclusions

Results obtained show that the proposed Rhombus Grid strategy gives better k-coverage in comparison to square grid and Tri-hexagon grid based node deployment strategies. Rhombus grid gives average 3.62-coverage while Square and Tri-beehive gave 3.14-coverage and 2.81-coverage respectively. Energy consumption is less in Rhombus Grid as compared to other strategies, because it takes less number of hops to reach the sink node.

References

[1] T. Abderrahim, E. L. Esteban, V. A. Javier, J. H. Joan, and E. Mohamed, "A novel approach for optimal wireless sensor network deployment," in *Proceedings of the Symposium on Progress in Information & Communication Technology (SPICT'09)*, Kuala Lumpur, Malaysia, vol. 7, no. 8, pp. 40-45, 2009.

[2] N. Meghanathan, "Link expiration time and minimum distance spanning trees based distributed data gathering algorithms for wireless mobile sensor networks," *International Journal of Communication Networks and Information Security*, vol. 4, pp. 196–206, 2011.

[3] W. Y. Poe and J. B. Schmitt, "Node deployment in large wireless sensor networks: Coverage, energy consumption, and worst-case delay," in *Asian Internet Engineering Conference (AINITC'09)*, ACM New York, NY, USA, pp. 77–84, 2009.

[4] M. Ringwald and K. Romer, "Deployment of sensor networks: Problems and passive inspection," in *Fifth IEEE Workshop on Intelligent Solutions in Embedded Systems*, pp. 179–192, 2007.

[5] S. Tang, "Traffic flow analysis of a multi-hop wireless sensor network subject to node failure," *International Journal of Communication Networks and Information Security*, vol. 3, pp. 163–169, 2011.

[6] M. A. M. Vieira, Jr. C. N. Coelho, D. C. da Silva, J. M. da Mata, "Survey on wireless sensor network devices," in *Proceedings of IEEE Conference on Emerging Technologies and Factory Automation (ETFA'03)*, vol. 1, pp. 537–544, 2003.

[7] G. Xing, X. Wang, Y. Zhang, C. Lu, R. Pless, and C. Gill, "Integrated coverage and connectivity configuration for energy conservation in sensor networks," *ACM Transactions on Sensor Networks*, vol. 1, no. 1, pp. 36–72, 2005.

[8] M. Younis and K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 6, no. 4, pp. 621–655, 2008.

Biography

Rajeev Singh received his M. Tech. degree in computer science & engineering from Indian Institute of Technology, Roorkee, India in 2008 and PhD degree from National Institute of Technology, Hamirpur, India in 2014. Currently, he is working as an assistant professor with the Department of Computer Engineering, Govind Ballabh Pant University of Agriculture & Technology, Uttarakhand, India. His research interest includes computer networks and network security.

Matendra Singh Manu received his M.Tech. in Computer Engineering from G.B. Pant University of Agriculture & Technology. His research areas includes computer networks and wireless sensor networks. E-mail - matendra.manu@gmail.com

Impact of Social Networking on Indian Youth - A Survey

Akashdeep Bhardwaj[1], Vinay Avasthi[1], Sam Goundar[2]
(Corresponding author: Akashdeep Bhardwaj)
University of Petroleum & Energy Studies, Dehradun, India[1]
(Email: Bhrdwh@yahoo.com)
Centrum Business School, Lima, Peru[2].
(Received June 2, 2017; revised and accepted Aug. 5, 2017)

Abstract

The extensive use of Social Networking in India has been on the rise among the new generation youths. In today's world, life cannot be imagined without Facebook, YouTube, Instagram, WhatsApp, LinkedIn or Twitter accounts and online handles. The new age social networking culture has been well accepted and has met an enthusiastic response and acceptance. There are reports of cultural changes and in the way traditional interactions and social communications are conducted in India. Research studies on this new age social media impact and usage within India have been limited to specific surveys and theories. The objectives of this study is an attempt to investigate the extent of social networking impact on the Indian youth. The reason for selecting youth as the target audience is because the direction of a country and culture is decided by the direction taken by youths of that country. This paper is an attempt to analyse the pattern of social networking usage and impact in order to determine the social networking addiction.

Keywords: Facebook; Indian Culture; Social Media; Social Networking; Twitter; WhatsApp

1 Introduction

While social networking met with huge enthusiasm among new generation initially, this new social culture seems to have been accepted by all age groups in India. The rise of internet access speeds and smart phones helped social networking even more and days of considering social applications as waste of time is long gone. Initially the use of social networking was limited to corporates and businesses for connecting with peers, customers, clients with twitter handle, Facebook accounts or WhatsApp mentioned on business and visiting cards. Social networking has now branched to include friends, family associates and classmates as well. Social networking offers several opportunities like access to information, videos, extension of social group, ability to express, learning opportunities, seeking and maintaining friends and relatives. A global media survey report on Facebook, Twitter, LinkedIn and Google+ presented statistics as shown in Table 1.

Off the 7.3 billion global population worldwide, social networking has 2.3 billion active users which has seen a rise of 176 million just last year [15]. Social networking advertising earnings are estimated at $8.3 billion in 2015 even as 385 organizations spent over 20% budget on social media channels which has been up by 15% compared to 2015 [5]. The increased use of social networking culture and social networking sites by youth has helped bring friends and family closer for those living in distant locations,

Table 1: Social networking statistics (Source: Brand Watch, October 2016) [17]

Social Media	Active User	Daily Users	Leaders	15-34 Ages	Indian Uses
Facebook	171 billion	113 billion	169	91%	142 million
Twitter	320 million	4.5 million	139	79%	28.5 million
LinkedIn	450 million	100 million	930	56%	37 million
Google+	300 million	1.6 million	89	83%	10 million

low additional cost of connectivity, sharing information, voicing opinions and updating each other on happenings in their lives.

The extensive use of Social Networking however, makes it an interesting study [6] regarding the risks and consequences on the existing youths. Social networking with the ability to effectively vanish boundaries, the anytime anywhere availability has seen impact on privacy as sharing too much, false unnecessary information about themselves or voice opinions, even getting exposed to fraudsters or cyber criminals and most critical of all the increased addiction to Internet and Social applications [13]. These tend to influence the youth for their social, emotional and psychological well-being. Adverse outcomes are seen as increasing exposure to cyber-bullying, unknown persons accessing personal information, online dating, exiting, and sleep deprivation, exposure to unsuitable digital content, outside influences of third-party groups encouraging to transfer money and low social interactions and limited face to face communications.

Examples of popular Social Networking sites are as follows:

- Facebook is currently one of the most famous social networking application site globally, is available in 37 languages and permits registered users to create profiles similar to a 'wall' like a virtual bulletin board, add friends, and send messages, comment, upload and share videos, photographs, web links. This application has several public features like
  - 'Marketplace' to post and respond to classified advertisements online;
  - 'Groups' to publicize events and invite guests and friends for attending that event;
  - 'Pages' to create and promote a personal or business ideas or involve others in a topic;
  - 'Presence Technology' which allows video calls and text chat for those online on the web site o 'Privacy' to block/allow specific or all members from viewing the profile, photos or comments.
- Twitter is a micro blog service which allows registered members to broadcast and follow replies to short posts, better known as 'Tweets' with no approvals required. Other users can subscribe to follow or reply to the tweets which may include hyperlinks to other blogs or posts and receive update messages by adding 'Hashtags' to keyword on the post, this acts like a metatag, expressed as #keyword. The tweets are searchable and available for the public. Twitter works on Ruby-for-Rails which is an open source web framework and its API is available for application developers.
- LinkedIn is designed primarily for corporate business community to promote personal brand online and allows registered members to establish a network of other professionals whom they know and trust as 'connections'. This requires preexisting relationships unlike Facebook or Twitter. Educational and Professional qualifications are the main display items on user pages here. This application is available in 24 languages.

- Google+ provides ability to Google users to post status updates or photographs, available to friends for view and comment in to 'Circles' which is primarily a group for multi-person instant messaging social networking system. Text and Videos are posted on 'Hangouts'.

2 Literature Review

This section presents a review of the research studies related to social networking aspects in order to determine which areas have already been explored and investigation in which area can add value. This help formulate objectives and undertake this research work. These research studies also provided an understanding into the existing efforts for understanding of the complexities of the social network sites.

Isodje, A. [7] presented an overview on the use of Social Media for business promotion, since social media as an online collaborative platform has the power to impact cultures and business. This further infiltrates communities, professional groups, peer bodies, which can be successfully used for promoting ones business.

Mamta et al. [10] tested for affiliation that exists between Higher Education and Social Networking Site. Mining algorithms provided by NASA tools like Like-Analyser, Gephi, Wolfram Alpha and NodeXL to assess presence and participation factor of students and education professionals in social network graphs are utilized in this study and analysis finding related to social network analysis predicted that social networking on Facebook and higher education work in parallel.

In times of traditional print media, there used to be one-way information dissemination which was restricted to geographical limits and presence. The process of information diffusion with arrival of Internet transformed significantly. Purva et al. [12] presented that online social networking like Facebook and Twitter have the fastest means of communication and having gained wide popularity, have revolutionized interpersonal communications by providing a platform to individuals for expressing themselves at a global level, beyond their immediate geography. The authors present the study on diffusion dynamics of specific real world events, discussed on Twitter, with respect to location and time. The events were categorizes into broad categories based temporal (short or long), geographical distribution (local or global), information diffusion (viral or gradual), influence (popular or unpopular) and the cause (natural or planned). It was conclude that the three-dimensional analysis of real-world events by exploring relationships among them.

The number of social networking site users is increasing immensely not only in India but also across the globe. Davmane et al. [3] analysed the factors for the online social networking sites as per users behavior regarding user friends, the peer groups, access patterns, amount of time spend, the effect on personal and professional life. User attitude and behavior is also surveyed for over seven hundred users using a questionnaire consisting of 27 questions which focused on behavior of Indian users in terms of usability, trends and access.

Singh et al. [14] presented the research effort in ensuring awareness about the social networking site concept, merits, demerits and meaning. The research methodology in this paper was based on primary and secondary data regards to grouping of users having similar type of interests, jobs, activities, backgrounds or some other type of real life similarities.

Purti at al. [2] focused on Big Data Management for Social Networking Sites by review and analysis of how Big Data is being managed for social networking sites by Facebook and Twitter. The data size for social networking sites constitutes almost 105 terabytes of data for every thirty minute, which in itself is a huge chunk of the data, unlike other data sources which has structured, limited data to handle. Facebook uses Hive for storing the data on HDFS (Hadoop Distributed File System) while Twitter has implemented a set of solutions storage inside Hadoop to store the data in LZO compressed format.

Kumar et al. [9] propose a sentiment analysis method on the tweets in Cloud environment and utilized Hadoop for intelligent analysis and storage of big data on Facebook and Twitter. The reason is that

handling huge amount of unstructured data is a tedious task to take up. The current Analytics tools and models used that are available in the market are not sufficient to manage big data. Therefore, there is a need to use a Cloud storage for such type of applications. The big data due to rise in social media has gathered huge interest among users and social networking site data is being used for various purposes including prediction, marketing and sentiment analysis.

Mittal et al. [11] analysed the effects of online shared sentiments of emoticons, interjections and comments extracted from posts and status updates. The authors also conducted a survey on the responses on the World Wide Web as an extensive large virtual space with users sharing and expressing views and opinions. Communication with the known and unknown residing anywhere on the globe at any point of time with the consumers being influenced by the social media whether intentionally or unintentionally.

Shang at al. [4] investigated why and how people use location sharing services on social networking platforms in China. To accomplish this the authors conducted research questions and forty three in-depth face-to-face interviews. The results indicated maximum users are concerned about privacy issues15 when using the location sharing services from social networking platforms, even as some indicated that they were not aware of this feature and did not know how to use location sharing services.

Muhammed et al. [1] reviewed research papers from 2010 to 2016 on Sybil attacks regarding use of fake and malicious identities on the online social network. The authors presented ideas for future research and also discussed a new taxonomy for Sybil attacks.

Zhou et al. [18] proposed a unique system called ProGuard for detecting malicious identity accounts in financial institutions dealing online with real and virtual currency. The authors suggested using behaviors, recharging patterns and currency usage by such accounts and even demonstrated experimental results proving their proposed system accomplished 0.3% false positives only.

Kiliroor et al. [8] presented a trust analysis system for online social networks to improve privacy and approval process for authentic social network site users. The authors discussed that real users may not be willing to disclose their identities due to privacy issues on public sites and social networks.

Wang et al. [16] proposed use of a probabilistic model for detecting identity thefts on social networking when using mobiles over unsecure Internet. The authors conducted experiment on real time data sets and displayed their proposed system achieved better performance and response by use of user behavior analysis as a key parameter for identity theft detection.

3 Social Networking Aspects

There are many positive aspects of social networking, but there are equally as many dangers and negative aspects that come with the use of sites such as Facebook, Twitter, LinkedIn, Google+, Pinterest, Tumblr, Instagram, gaming sites, and blogs.

3.1 Positive Aspect

Some of the positives arising from social networking are listed in Table 2.

3.2 Negative Aspects

Some of the negatives arising from social networking are listed in Table 3.

Table 2: Social networking Positive Aspects

Education	- Helps in better collaboration and communication between teachers and students; - Access to online resources helps students to learn better, faster; - Student grades improved along with reduced absenteeism in online sessions; - Educational topics and school assignments being discussed on social sites.
Politics	- Increase in voter participation, seeing their friends voted on Facebook post; - More likely to attend a political meeting and rally seeing others on social sites; - Social movements have easy fast method of mobilizing people and sharing info.
Awareness	- Information dissemination is faster than any media - breaking news spreads fast; - Access to previously inaccessible resources for academic research; - Helps inform and empower individuals to change themselves.
Social Benefits	- Social media allow people to communicate with friends and this increased online communication strengthens those relationships, friendships; - People making new friends - 57% online teens report making new friends online. - Helped find and keep in touch with friends who are geographically far off.
Job Opportunities	- Great for marketing professionals - connect and find business opportunities. - Employers find candidates and unemployed find work faster. - Social media sites have created thousands of ecommerce jobs, new avenues.

4 Social Networking Survey

The authors conducted a survey analyse the impact of Social Networking on Indian youth and culture. The survey involved sending a detailed questionnaire to respondents via Survey Monkey and 532 responses were received, the breakdown and survey analysis is presented as shown in Table 4.

The respondents were asked few questions on the social networking and the responses are illustrated in the below graphs for reference.

Question #1. What is the amount of time you are spending daily on Social Networking sites? Most respondents spent over 1 to 2 hours on the social networking sites each day, which in a country like India is substantial given the closed culture (See Figure 1).

Question #2. What is the extent of the addiction for social networking? The respondents religiously checked their social networking accounts each morning, which shows a trend in the rising interest and addiction to social networking (See Figure 2).

Question #3. What is the main purpose of social networking for your use? The primary reason for using social networking tends to be non-essential, voicing opinions which in Indian culture is limited as speaking out or against is looked down upon (See Figure 3).

Question #4. Mental and physical impact on health? (See Table 5)

Question #5. What are the different ways of accessing social networking applications?

Table 3: Social networking Positive Aspects

Apps access User Data	- Social apps force users to grant access to their apps for list of things; - Access public profile information - user name, profile picture, friend list birthday, favorite movies and books, etc. - Send email - sending direct emails to the user email address; - Access posts in the News feed, Video and Photos posted; - Access family and relationships information; - Post to the wall -Add new message posts on the user's behalf.
Detriment to Work	- Enables copying and cheating when submitting assignments; - Grades improve for light users, while heavy users of social media suffer -; - Students have an average GPA of 3.06 while non-users have 3.82 GPA; - For every 93 minutes over the average 106 minutes spent on Facebook daily, college students' grades dropped. - Students going online while studying scored 20% lower on tests. - Possible negative effects on college admission - 35% of admissions officers scan potential candidate social media blogs and posts which can affects hiring and educational decisions. - Social networking sites harm employees' productivity - 51% of users aged 25-34 checked social media at work. - Harm to employment and prospects - - Job recruiters check a prospective employee's social media accounts, things like profanity, poor spelling, grammar, racism, and sexism, health references to alcohol, drugs, sexual or religious content can count against you.
Lack of Privacy	- Young people often give out personal information when online without reading the fine print privacy policies and unaware about misuse by third parties. - Exposure to corporate and governmental intrusions - Insurance companies use information gleaned from social media. - Online advertising policies are an invasion of privacy. If clicked "like" for a brand, browser cookies give the company information and access about personal information and preferences.
Users Vulnerable to Crime	- Unauthorized sharing of intellectual property can cause loss of potential income. - Cyber-attacks like ransomware, hacking, identity theft and phishing are common problems faced by end users. - Criminals browse social media to know user whereabouts and are known to commit crimes when away on vacation.
Waste of Time	- Constant browsing and replying online posts and blogs, takes the user attention away from core work and often take some time to return to original task.

Social Detriment	- Cyber-bullying or use of electronic communication to bully someone by sending intimidating or threatening messages is commonplace online. This causes emotional trauma and sometimes even leads to suicide. - Excessively being online correlated with personality and brain disorders - poor social skills and narcissistic tendencies or even need for instant pleasure with addictive behaviors and other emotional issues leading to depression, anxiety and loneliness. - Less time for face-to-face interaction with loved ones. - Youngsters are prone to feeling isolated, disconnected from real world and face higher risks of depression, low self-esteem and eating disorders.
Misinformation	Enables the spread of false rumors and unreliable information: - Self-diagnosis of health problems and following amateur medical advice; - Befriending someone to gain information; - Revealing reconnaissance data unknowingly to the public; - Studies have shown that sites such as Facebook influence you, via advertisements, to spend more money.

Table 4: Breakdown of Respondents organization

Organization Category	Respondents	Breakup %
Financial Services	46	9%
Education	173	33%
Information Technology	99	19%
Retail, Ecommerce	65	12%
Internet Service Provider	39	7%
Gaming	22	4%
Media & Travel	31	6%
Pharmacy	57	11%

Accessing social networking applications by users range as follows

- Mobile Devices - 45% (Includes Smart Phones, iPads, Kindles, Tablets);
- Desktop Computers - 22%;
- Laptops - 33%.

5 Conclusion

The Social Networking patterns shown by people in the study are largely consistent with those recorded in previous research studies with respect to impact of popular social media sites on Indian culture and the extent of the use, purposes, mode of access when using these sites. The author also reviewed benefits of the social networking sites in culture development, building self-identity, developing relationships and acquisition of social, communication, and technical skills. For future research, there is a need to increase the sample size and select a better representative sample. This study might also suffer from the disadvantages of judgment sampling viz., researcher's bias and stereotypes that may distort

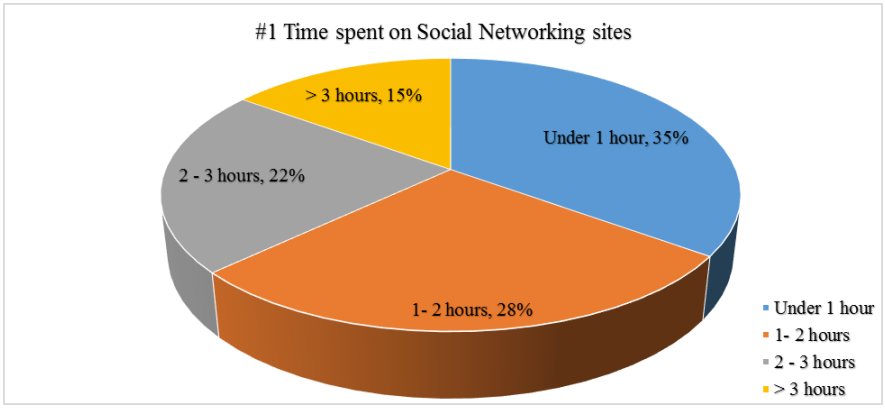

Figure 1: #1 Time per Day Spent On Social Networking Sites

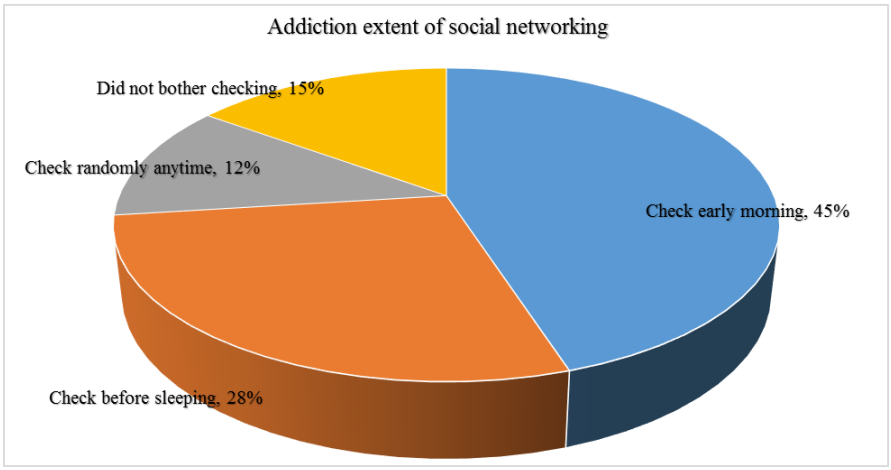

Figure 2: #2 Addiction extent of Social Networking

the results; group selected may not represent all the population and also it might not be possible to accurately identify the sample using this method in case the population is very large. Also, since Social Networking is a global phenomenon, comparative analysis of students from within India and also of various countries can yield interesting findings, implying whether SN addiction exists, also if it does is the pattern of students from different region differs or not.

6 Recommendations

Based on the findings drawn from this study, the researcher has made the following.

- Recommendation to College and University Authorities:
  - Regulation on use of mobile phones during lectures.

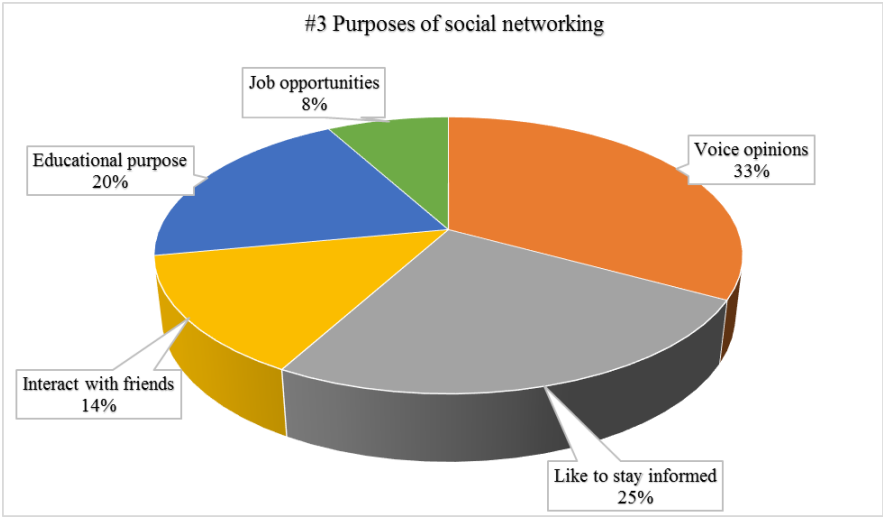

Figure 3: #3 Use of Social Networking

Table 5: Mental and physical impact on Health responses

Social Networking	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
Way of life for youth & old	15%	43%	19%	16%	7%
Is Highly addictive	18%	55%	13%	10%	4%
Compare our lives with others	41%	42%	11%	5%	1%
Making us restless, sleeplessness	23%	53%	13%	7%	4%
Gives rise to Cyber Bullying	18%	43%	22%	13%	4%
Glamourizes Drug & Alcohol	26%	29%	20%	17%	8%
Can make us unhappy	44%	31%	12%	11%	2%
Leads to fear of missing out	39%	28%	26%	3%	4%
Multitasking, loss of concentration	32%	34%	11%	18%	5%
Leads to increased peer pressure	23%	48%	18%	8%	3%

  - Hence the students access the various social networking sites through their mobile phones, it is advisable that university enacts laws, making students' use of phones during lectures an offence which shall attract drastic punitive measures for the culprits.
  - Organize a seminar to enlighten students on the not too-good aspects of using social networking sites as media of interaction. This can be done by exposing students to the importance of face to face communication in the creation of real communication or message sharing. Seminars would be helpful here.
  - Provision of laws on the content of social media: There has to be laws guiding the students' use of the social networking sites and what they disseminate through the media.

- Recommendations to the Ministry of Information Technology:

Since social networking sites fall within the ambit of the Ministry of Communications Technology, it is the duty or responsibility of initiating and coordinating all the policies and programs towards the use and development of information and communication Technologies (ICTs). Social networking is part and parcel of the ICTs, as such from the findings. From this work; these recommendations are made to the ministry:

- The ministry has to mandate all the social service providers to make it mandatory that the condition for one to open an account on any of the social networking sites is having a duly registered GSM SIM card.
- The service providers have to keep the personal details of each of their account owners including their GSM phone numbers; and make the information available to an appropriate government agency if the need arises.
- Enactment of Social Media Use Act: The ministry has to propose to enact a new law that would guide the users of the social network sites with the do's and don'ts. This is quite necessary now, as one of the findings of this study shows that some of the students use the social networking sites to engage in cyber-crimes. Such act shall provide the legal framework that would help Law courts to adjudicate on cyber-crime cases.

References

[1] M. Al-Qurishi, M. Al-Rakhami, A. Alamri, M. Alrubaian, Md M. Rahman, S. Hossain, "Sybil Defense Techniques in Online Social Networks: A Survey," *IEEE Access*, vol. 5, pp. 1200–1219, 2017.

[2] P. Beri, S. Ojha, "Comparative Analysis of Big Data Management for Social Networking Sites," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom'16)*, IEEE, 2016.

[3] M. Devmane, N. Rana, "Usability Trends and Access of Online Social Network by Indian population and its analysis," in *International Conference on Nascent Technologies in the Engineering Field (ICNTE'15)*, IEEE, 2015.

[4] S. Gao and X. Zhang, "Why and how people use location sharing services on social networking platforms in China," in *12th International Joint Conference on e-Business and Telecommunications (ICETE'15)*, IEEE, 2015.

[5] S. Gebauer, *Twitter Statistics 2016. Social Claim Blog*, Nov. 1, 2016. (`https://blog.thesocialms.com/twitter-statistics-you-cant-ignore/`)

[6] D. Houghton, A. Johnson, D. Nigel, M. Caldwell, *Tagger's Delight Disclosure and Liking Behavior in Facebook: The Effects of Sharing Photographs Amongst Multiple Known Social Circles*, Oct. 20, 2016. (`http://epapers.bham.ac.uk/1723/1/2013-03_D_Houghton.pdf`)

[7] A. Isodje, "The use of Social Media for Business Promotion," in *International Conference on Emerging & Sustainable Technologies for Power & ICT in a developing society*, 2014.

[8] C. C. Kiliroor, C. Valliyammai, "Trust analysis on social networks for identifying authenticated users," in *IEEE 8th International Conference on Advanced Computing (ICoAC'17)*, 2017.

[9] M. Kumar and A. Bala, "Analyzing Twitter sentiments through Big Data," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom'16)*, IEEE, 2016.

[10] M. Madan, M. Chopra, M. Dave, "Predictions and recommendations for the higher education institutions from Facebook social networks," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom'16)*, 2016.

[11] S. Mittal, A. Goel, R. Jain, "Sentiment analysis of E-commerce and social networking sites," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom'16)*, IEEE, 2016.

[12] P. Purva, A. Yadav, F. Abbasi, D. Toshniwal, "How Has Twitter Changed the Event Discussion Scenario? A Spatio-temporal Diffusion Analysis," in *International Congress on Big Data (BigData Congress'15)*, IEEE, 2015.
[13] E. Shaw, *Status Update: Facebook Addiction Disorder*, Sept. 15, 2016. (`http://theglenecho.com/2013/01/29/status-update-facebook-addiction-disorder/`)
[14] H. Singh, B. P. Singh, "Social Networking Sites: Issues and Challenges Ahead," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom'16)*, IEEE, 2016.
[15] C. Smith, *Facebook Facts and Statistics*, Oct. 20, 2016. (`http://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats`)
[16] C. Wang, Bo Yang, J. Luo, "Identity Theft Detection in Mobile Social Networks Using Behavioral Semantics," in *IEEE International Conference on Smart Computing (SMARTCOMP'17)*, 2017.
[17] B. Watch, *Social Media 2016*, Nov. 3, 2016. (`https://www.brandwatch.com/blog/96-amazing-social-media-statistics-and-facts-for-2016/`)
[18] Y. Zhou, D. W. Kim, J. Zhang, L. Liu, H. Jin, H. Jin, T. Liu, "ProGuard: Detecting Malicious Accounts in Social-Network-Based Online Promotions," *IEEE Access*, vol. 5, pp. 1990–1999, 2017.

Biography

Akashdeep Bhardwaj, PhD research scholar from University of Petroleum and Energy Studies(UPES), PGDM, B.E (Computer Science) is an Enterprise Risk and Resilience Technology Manager in Information Security and Infrastructure Operations having worked with MNCs for over 20 years and is certified in Ethical Hacking, Cloud, Microsoft, Cisco, and VMware technologies.

Dr. Vinay Avasthi has received PhD in Computer Science and is currently working as Associate Professor with University of Petroleum and Energy Studies, Dehradun for over 8 years and has several research papers in referred international journals.

Dr. Sam Goundar has over 20 years of academic experience in IT, IS, MIS and CS, and researching on cloud & mobile computing, Educational Technology, MOOCs, Smart Cities, Artificial Intelligence, among other domains. He is a Senior Member of IEEE, a member of ACS, a member of the IITP, New Zealand, Certification Administrator of ETA-I, USA and Past President of the South Pacific Computer Society and also serves on IEEE Technical Committee for Internet of Things, Cloud Communication and Networking, Big Data, Green ICT, Cyber security, Business Informatics and Systems, Learning Technology and Smart Cities.

Guide for Authors
International Journal of Electronics and Information Engineering

International Journal of Electronics and Information Engineering (IJEIE) will be committed to the timely publication of very high-quality, peer-reviewed, original papers that advance the state-of-the art and applications of Electronics and Information Engineering. Topics will include, but not be limited to, the following: Algorithms, Bioinformatics, Computation Theory, AI, Fuzzy Systems, Embedded Systems, VLSI/EDA, Computer Networks, Information Security, Database, Data Mining, and Information Retrieval, Digital Content, Digital Life, and Human Computer Interaction, Image Processing, Computer Graphics, and Multimedia Technologies, Information Literacy, e-Learning, and Social Media, Mobile Computing, Wireless Communications, and Vehicular Technologies, Parallel, Peer-to-peer, Distributed, and Cloud Computing, Semiconductor, Software Engineering and Programming Languages, Telecommunication, etc.

1. Submission Procedure

Authors are strongly encouraged to submit their papers electronically by using online manuscript submission at http://ijeie.jalaxy.com.tw/.

2. General

Articles must be written in good English. Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere. It will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

2.1 Length Limitation:

All papers should be concisely written and be no longer than 30 double-spaced pages (12-point font, approximately 26 lines/page) including figures.

2.2 Title page

Title page should contain the article title, author(s) names and affiliations, address, an abstract not exceeding 100 words, and a list of three to five keywords.

2.3 Corresponding author

Clearly indicate who is willing to handle correspondence at all stages of refereeing and publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address.

2.4 References

References should be listed alphabetically, in the same way as follows:
For a paper in a journal: M. S. Hwang, C. C. Chang, and K. F. Hwang, ``An ElGamal-like cryptosystem for enciphering large messages," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 445--446, 2002.

For a book: Dorothy E. R. Denning, *Cryptography and Data Security*. Massachusetts: Addison-Wesley, 1982.

For a paper in a proceeding: M. S. Hwang, C. C. Lee, and Y. L. Tang, ``Two simple batch verifying multiple digital signatures," in *The Third International Conference on Information and Communication Security (ICICS2001)*, pp. 13--16, Xian, China, 2001.

In text, references should be indicated by [number].

2.5 Author benefits

No page charge is made.

Subscription Information

Individual subscriptions to IJEIE are available at the annual rate of US$ 200.00 or NT 6,000 (Taiwan). The rate is US$1000.00 or NT 30,000 (Taiwan) for institutional subscriptions. Price includes surface postage, packing and handling charges worldwide. Please make your payment payable to "Jalaxy Technique Co., LTD." For detailed information, please refer to http://ijeie.jalaxy.com.tw or Email to ijeieoffice@gmail.com.