Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

September 2013 **Volume V** **Number 3** **ISSN 2061-2079**

CRYPTOLOGY

Special Issue on Cryptology – GUEST EDITORIAL .... *V. Matyáš, Z. Říha and M. Kumpošt* 1

Protection of Data Groups from Personal Identity Documents
.... *P. Kubiak, M. Kutylowski and W. Wodo* 2

Classes of Garbling Schemes .... *T. Meskanen, V. Niemi and N. Nieminen* 8

On a Key Exchange Protocol Based on Diophantine Equations
.... *N. Hirata-Kohno and A. Pethő* 17

Strongly Secure Password Based Blind Signature for Real World Applications
.... *S. Jose, P. Mathew K. and C. Pandu Rangan* 22

PAPERS FROM OPEN CALL

An Adaptive Load Sharing Algorithm for Heterogeneous Distributed System
.... *P. Neelakantan and A. Rama Mohan Reddy* 30

Energy Efffective Coexistence of LTE-WCDMA Multi-RAT Systems
.... *I. Törős and P. Fazekas* 36

ADDITIONAL

Guidelines for our Authors 43

Technically Co-Sponsored by

IEEE COMMUNICATIONS SOCIETY

hte

IEEE HUNGARY SECTION

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest, Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu
Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Networked Systems and Services
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: FOM Media

www.infocommunications.hu

Special Issue on Cryptology – Guest Editorial

Václav (Vashek) Matyáš, Zdeněk Říha and Marek Kumpošt

***Abstract*—This special issue brings selected papers from the 2013 Central European Conference on Cryptology, held in Telč, June 26-28, 2013.**

This special issue focuses on the area of applied cryptography, bringing up selected papers from the 2013 Central European Conference on Cryptology, covering various aspects of cryptology, including cryptanalysis, cryptographic applications in information security, design of cryptographic systems, general cryptographic protocols, post-quantum cryptography, pseudorandomness, signature schemes, and steganography.

The first paper "Protection of Data Groups from Personal Identity Documents" of Przemysław Kubiak et al. proposes a procedure of presenting a signed face image of the document holder. The aim of this procedure is to authenticate the image by document issuer, but at the same time to prevent misuse of this high quality digital data. The solution reflects the technology challenges related to limits of data storage on a personal identity document chip, and the designed protocols can potentially be used for other than just biometric data.

The second paper "Classes of Garbling Schemes" of Tommi Meskanen et al. extends some results of the work of Bellare et al. from 2012 on garbled circuits from a cryptographic technique to a cryptographic goal, defining several new security notions for garbled circuits. Meskanen et al. provide some new results about the classes of garbling schemes defined by Bellare et al., define new classes of garbling schemes, prove their relation of earlier classes, and also investigate some results concerning the new classes.

The third paper "On a key exchange protocol based on Diophantine equations" of Hirata-Kohno et al. analyzes a key exchange protocol proposed by H. Yosh in 2011, based on the hardness to solve Diophantine equations. The authors analyze the protocol and show that the public key is very large, suggesting also an alternative solution through large families of parameters both in the finite field and in the rational integer cases for which the protocol can be secure.

The last paper "Strongly Secure Password Based Blind Signature for Real World Applications" of Sangeetha Jose et al. password based blind signature that are used in scenarios where a user requires the authentication of the signer without revealing the message to the signer. The authors propose a novel design that ensures the properties unforgeability, blindness and unframeability. Yet for small sizes of passwords, an off-line password guessing attack is of high relevance. The authors propose a strongly secure password based blind short signature that solves the off-line password guessing problem, with the formal proof of the scheme reduced to the computational Diffie-Hellman (CDH) assumption.

VÁCLAV (VASHEK) MATYÁS is a Professor at the Masaryk University, Brno (CZ), and serves as a Vice-Dean for Foreign Affairs and External Relations, Faculty of Informatics. His research interests relate to applied cryptography and security, publishing over a hundred peer-reviewed papers and articles, and co-authoring six books. He was a Fulbright Visiting Scholar with Harvard University, Center for Research on Computation and Society, and also worked with Microsoft Research Cambridge, University College Dublin, Ubilab at UBS AG, and was a Royal Society Postdoctoral Fellow with the Cambridge University Computer Lab. Vashek was one of the Editors-in-Chief of the Identity in the Information Society journal, and he also edited the Computer and Communications Security Reviews, and worked on the development of Common Criteria and with ISO/IEC JTC1 SC27. Vashek is a member of the Editorial Board of the Infocommunications Journal. He received his PhD degree from Masaryk University, Brno and can be contacted at *matyas@fi.muni.cz*.

ZDENEK RÍHA is an Assistant Professor at the Masaryk University, Faculty of Informatics, in Brno (CZ). He received his PhD degree from the Faculty of Informatics, Masaryk University. In 1999 he spent 6 months on an internship at Ubilab, the research lab of the bank UBS, focusing on security and usability aspects of biometric authentication systems. Between 2005 and 2008 he was seconded as a Detached National Expert to the European Commission's Joint Research Centre in Italy, where he worked on various projects related to privacy protection and electronic passports. He was involved in the ePassport interoperability group known as the Brussels Interoperability Group. Zdenek has been working with the WG 5 (Identity management and privacy technologies) of ISO/IEC JTC 1/SC 27. Zdenek's research interests include smartcard security, PKI, security of biometric systems and machine readable travel documents. Zdenek can be contacted at *zriha@fi.muni.cz*.

MAREK KUMPOST is a Research Assistant at the Masaryk University, Faculty of Informatics, in Brno (CZ). He received his PhD in 2009 from the Faculty of Informatics, Masaryk University. The primary area of his doctoral research was oriented on privacy protection, anonymity and user profiling. He was involved in two European-wide project on privacy protection and identity management – FIDIS (Future of Identity in the Information Society) and PICOS (Privacy and Identity in Management for Community Services). He spent 3 months with the LIACC (Laboratory of Artificial Intelligence and Computer Science) group working on user profiling based on information from NetFlow. He is also interested in network security, web application security and cloud security. Marek can be contacted at *kumpost@fi.muni.cz*.

Protection of Data Groups from Personal Identity Documents

Przemysław Kubiak, Mirosław Kutyłowski and Wojciech Wodo

***Abstract*—For personal identity documents, we propose a procedure of presenting a signed face image of the document holder. Our goal is to authenticate the image by document issuer, but at the same time to prevent misuse of this high quality digital data. As the signature is recipient dependent, illegitimate transfer of the signature to third parties is strongly discouraged. Despite that the document issuer is the signatory and that the image recipients are unpredictable in advance, only a very limited amount of information has to be stored on a chip of the personal identity document. Moreover, the solution prevents creating additional signatures by document issuer, as a signature created outside the card leads to a mathematically strong proof of a fraud.**

Although motivation for the protocols presented below was protection of biometric data, the protocols might be used in case of any data.

***Index Terms*—personal identity document, smart card, personal data protection, designated recipient, electronic signature, Merkle tree**

I. Problem Description

A. Background problem

A personal identity document equipped with a cryptographic chip, called *e-ID* for short, offers high level security guarantees against document forgeries: while there is a race between graphical protection techniques and the forgery methods. On the other hand, repeating the same data in electronic form and signing them by the document issuer provide strong and independent security mechanisms at a low price. Advances in cryptanalysis limit the long-term value of these guarantees, nevertheless they are relatively long-lasting.

Electronic layer of e-ID may store a high resolution face image of the document holder – more detailed than the image printed on the document. This enables much more reliable inspection based on e-ID. The strategy applied in particular by biometric passports is to present not only raw data, but also a signature of the document issuer for those data. In this way during an inspection we may become convinced that the image presented originates from the document issuer and has not been replaced even if chip security of e-ID has been broken.

Securing data with a signature of the document issuer is a two-edged sword. Once the signature is created, it can be used by anybody to confirm authenticity of digital data. Therefore this approach leads to privacy threats: once the signed data is shown to a second party, the owner of e-ID has no further control over who has access to it. In particular, this data can be sold to third parties. The signature has a negative influence on the situation, as quality of the data is confirmed by authority issuing the e-ID. This problem has been one of the major factors behind the design of German personal identity card, where the data might be shown without issuer's signature, but via an authenticated and secure channel [1]. The communication and authentication protocols are designed in such a way that even a full transcript of a session together with ephemeral data created during the session on the terminal side cannot be used as a proof against a third party. This is achieved by means of *simultability*. The price is that we have to assume that the chips of the personal identity cards provide full security against all kinds of (practical) attacks.

B. Assumptions about e-ID chips.

We assume that the chip used by e-ID provides certain (limited) security against the issuer of e-ID. Namely, we assume that keys generated privately on the chip can be read by the e-ID issuer as long as the key generation process takes place in environment controlled by the issuer. However, keys generated on the chip when the e-ID is in control of the owner are neither predictable for the e-ID issuer nor they leak from the e-ID.

The assumptions above reflects the setting where the chip vendor does not collude with the authority issuing and personalizing identity documents, but the authority has access to technologies that may break security means on the chip and can access all relevant data on the chip.

C. System goals.

We aim to provide a solution such that:

- Once the face image (or more generally, the data groups containing personal data of the owner) are presented by an e-ID, then a customized signature of the document issuer is attached.
- The signature indicates the recipient of the signature, but the proof is not necessarily unconditional. This means, it should provide traces who is not fulfilling the duties of personal data protection, but on the other hand the signature is not necessarily an undeniable proof of e-ID document presence.
- The authority issuing the e-ID documents cannot create clone documents and customized signatures in order to accuse a certain party for violations of personal data protection.

The simplest solution is to provide a signature $\mathrm{Sign}_K(H(D), R)$, where K is the signing key of the

Authors are with Wroclaw Univeristy of Technology, Institute of Mathematics and Computer Science, e-mail: *firstname.lastname*@pwr.wroc.pl

This research was initiated under support of Foundation for Polish Science.

issuing authority, D denotes the data groups and R is the recipients ID. There are two severe problems with this approach: R must be known in advance and the issuer can create these signatures at any time, distribute them and accuse R of violations of personal data protection.

The first problem can be dealt with by means of proxy signatures [2]: the chip of e-ID receives data that enable it to create signatures on behalf of the document issuer. However, with this approach we solve one hard problem, but create a new harder one. Namely, once an adversary breaks into a chip of e-ID, it can manipulate the e-ID document and in particular replace the face image.

One may also try to use designated verifier schemes - in this case the signature is worthless to anybody, but the verifier determined at signature creation time. The same problems apply as before – the issuing authority has either to create them in advance and store on the chip of e-ID or use a proxy version of it. Moreover, proxy and designated verifier signature schemes are significantly more complicated than the standard signatures, use operations that might be unavailable on the standard chips. Therefore the non-volatile memory requirements for storing program code and data might be quite high regarding limitations for chips on smart cards. Finally, there is nothing so far that would prevent malicious authorities from creating and using the clones of identity documents.

Another option is hiding the signature of the issuing authority by the e-ID. Instead, the chip proves that it holds a signature for given data D (compare [3]). However, such solutions fall into another category as the verifier cannot store the signature for offline verification. Our goal is a real signature - the only difference should be that it has to be customized to show the original recipient.

D. Our contribution

We present two solutions with slightly different properties. The first one is based on hash functions, the second one on asymmetric techniques. In both cases the signature is customized in a way that points to the signature recipient and it is infeasible to change this pointer unless one has access to the secrets stored in the chip of e-ID document.

II. HASH BASED PROTOCOL

Below we sketch the idea of our solution.

A. General settings.

The document issuer holds a conventional pair of keys for creating electronic signatures. Authentication of the public key is achieved again in the standard way (e.g., by publishing or by public key certificates).

For each e-ID document we have k different positions for document verifiers, each verifier is assigned one position. The number k is a system parameter and its value has to be fine tuned depending on system size and trade-off between privacy and detectability of parties misusing personal data. The position of a verifier for each e-ID is determined separately in a pseudorandom way. Namely, for a hash function H, a verifier V for identity document ID_D is assigned position $H(V, D) \bmod k$. In this way, for a given identity document there are good chances that the verifiers the owner of the document visits most frequently have been assigned different positions. On the other hand, without knowledge of the datagroups the position of a given verifier in a given e-ID is completely unpredictable.

As the positions will correspond to leaves of a Merkle tree constructed separately for each e-ID, we assume that k is a power of 2 and throughout the paper log symbol defines the binary logarithm.

B. Document personalization by the e-ID issuer.

For each e-ID document ID_D there is a master secret S_D chosen uniformly at random by the document issuer.

According to standard conventions, we assume that data stored on ID_D consist of data groups D=(D_1,. . . ,D_m), where each D_i, $i = 1, \ldots, k$, is a single data group. As the data might be exposed selectively, the signature is created for $H_D = H(H(D_1), \ldots, H(D_m))$. In this way, for verification of a signature it suffices to present $H(D_1), \ldots, H(D_m)$ as well as the data groups D_j that are to be disclosed.

For the purpose of clone-evidence we need secrets x_D - e.g., x_D might be a signature of the document issuer under the text *"ID_D has been cloned or broken"*.

For ID_D the document issuer creates a Merkle tree [4] of height $\log k + 2$ in the following way:

- for each $i < k$ there are 4 corresponding leaves; they are labelled with the following values: H_D, $x_{D,R,i}$, H_D, $x_{D,L,i}$, where $x_{D,L,i} = R(i||S_D)$, $x_{D,R,i} = x_D - x_{D,L,i}$, and R is a cryptographic pseudorandom generator.
- We construct the labels for higher levels of the tree as always for Merkle trees: if a node A has children nodes with labels h_1 and h_2, then the label of A is $H(h_1, h_2)$.

Let $Root_D$ denote the label of the root of the tree constructed for ID_D. The last step is to create a signature $Sign_D$ of $Root_D$ by the document issuer and to store it on the chip of ID_D.

C. E-ID personalization by the owner.

After delivering ID_D to its owner, it executes a procedure of uploading a random secret X_D to the chip of ID_D. X_D can be kept outside ID_D, but must be unknown for the document issuer.

The purpose of X_D is to determine the leaves used for signing: if $H(i, X_D) \bmod 2 = 0$, then for position i the leaf labeled with $x_{D,L,i}$ is used. If $H(i, X_D) \bmod 2 = 1$, then for position i the leaf labeled with $x_{D,R,i}$ is used. Here we assume that hash function H is cryptographically secure, thus there is no bias for any single bit position.

In particular, the values of $H(i, X_D) \bmod 2$ may be stored in an array A of k bits.

D. Customizing the signature.

Assume that a verifier V is to receive signed data from document ID_D. Apart from $H(D_1), \ldots, H(D_m)$, and chosen data groups D_j the e-ID prepares a signature of the document issuer in the following way:

- compute position i for V as $i = H(V, D) \bmod k$,
- determine a *path* P_i from a leaf holding $x_{D,Z,i}$ to the root, where $Z = L$ if $H(i, X_D) \bmod 2 = 0$, and $Z = R$ otherwise,
- compute a list HP_i of hashes: for each node of P_i, the list HP_i indicates the label of the sibling of the node on P_i. The only exception is the leaf node, for which its label is given and not the label H_D of the sibling node.
- return $H(D_1), \ldots, H(D_m)$, HP_i, $Sign_D$ and the relevant data groups D_j which are to be disclosed.

E. Verification of a signature.

The following steps are necessary to verify a signature $H(D_1), \ldots, H(D_m)$, HP_i, $Sign_D$:

- $H(D_1), \ldots, H(D_m)$ are checked against the data groups disclosed to the verifier,
- the hash values on the path P_i are reconstructed using HP_i, the first value is computed as $H_D = H(H(D_1),\ldots, H(D_m))$,
- the signature $Sign_D$ is verified in the conventional way, against the label $Root_D$ of the root node computed in the previous step.

F. Implementation issues - speeding up signature creation.

Note that the chip does not need to remember the labels of nodes of its hash tree – it can be reconstructed from D_1, $\ldots D_m$ and the secret S_D. Also it is easy to see that auxiliary storage required to compute HP_i is roughly $\log k$ hash values.

If k is relatively small, then computation effort on the chip is acceptable. However, if this is not the case, we can significantly reduce the computational effort by storing the labels of the nodes at height $\frac{1}{2}\log k + 1$ of the tree. In this case the chip has to reconstruct labels for *two* subtrees of depth $\frac{1}{2}\log k + 1$ of total size roughly $6\sqrt{k}$ instead of $\approx 6k$.

G. Clone detection.

As the secret X_D is created *after* the e-ID document is given to the owner, the issuer cannot guess which leaves are used by the chip of e-ID for each position i. A single attempt to create an extra signature on behalf of the document owner leads with probability $\frac{1}{2}$ to disclosure of the secret x_D. An attempt to create, say 20, such signatures will not lead to fraud disclosure with probability $\frac{1}{2^{20}}$, which is the value too low for any authority to dare a fraud.

H. Detection of offenders of personal data protection.

Assume that a verifier V collects data and signatures obtained from e-ID documents. Assume that V has sold n such records to a data bank L which has reached the total size N. Assume that L has been captured by law enforcement authorities.

For each signature found in L we can check if it is possible that it has been obtained from V. If a signature uses the same position in the Merkle tree as it would be used for V, then we say that this is an *accusation* against V. As the positions in the Merkle tree are determined in a pseudorandom way, we may assume that the expected number of accusations against V in L equals

$$(N - n) \cdot \tfrac{1}{k} + n = \tfrac{N}{k} + n(1 - \tfrac{1}{k}).$$

If V is honest, then the expected value equals $\frac{N}{k}$.

Statistical tests indicating dishonest behavior of V can be based on the fact that the Bernoulli distribution is fairly concentrated. For instance, according to Chernoff bounds, probability that there are more than $\frac{2N}{k}$ accusations in case of honest V is bounded by $(e/4)^{N/k}$. For $k = 16$ and $N = 2^{10}$ we get that probability to get more than 128 accusations is $\approx 2^{-35}$, while the expected number of accusations for dishonest V and 70 records sold is higher than 129. This shows that any large scale sale of data is very risky for a verifier. On the other hand, in this kind of business what counts is only large scale sale, as single records have a low price.

Note that higher values of k make detection of dishonest verifiers more reliable. On the other hand, if k is low, then a signature pointing to position i which should be used by V is not an evidence that ID_D has been presented to V. Namely, this position is used by the fraction $\frac{1}{k}$ of all verifiers!

I. Feasibility Issues

We have performed speed tests on Gemalto Java Cards concerning computation of hash values. The results for exemplary parameters are as follows:

SHA-1 (160 bits): 1 hash $\approx$ 5ms, 1280 hashes $\approx$ 4.8s,
SHA-2 (256 bits): 1 hash $\approx$ 9ms, 1280 hashes $\approx$ 10s.

For comparison observe the number of hashes to be computed to create a single signature for tree depth 10 when the hashes of level 5 (32 values) are stored by the chip, is $2 \cdot (32 - 1)$, so the time required is less than 0.5s for SHA-2.

Memory usage for data in case of trees of depth 10 (with intermediate level at depth 5 stored on the chip) equals:

keys:	master secret S_D – 128 bits, user secret X_D – 128 bits, array of hash values on Merkle tree on intermediate level at depth 5 – $32 \cdot 256 = 8192$ bits)
temp. hash values:	at most 6 hashes at a time – 1536 bits.

III. Asymmetric Approach

In this section we sketch a protocol which can be used to create customized signatures by *tagging* a signature of the document issuer. Namely, the chip of e-ID attaches a tag to the data groups and the signature of the issuing authority revealed to a verifier. The point is that without the tag signature verification is infeasible, and that the tag indicates the intended verifier. No prior agreement on the identity of verifiers is necessary.

A. Building Blocks

The main building block for the high-resolution protocol is a solution used to prove equality of two discrete logarithms.

a) System settings.: Let g generate a group of prime order q. Furthermore, assume that Decisional Diffie-Hellman Problem is hard for this group. Let h belong to this group be chosen so that its discrete logarithm is unknown.

We assume that a prover holds a private exponent x. The goal of the prover is to convince that two elements a, b have the form $a = g^x$, $b = h^x$.

b) Schnorr-like proof of equality of discrete logarithms [5].: First the prover performs the following steps:

1) generate r at random,
2) $k := g^r$, $\ell := h^r$,
3) $e := H(k, \ell, g, h, a, b, m)$, where m is some message, for example an empty message or *the name of the addressee of the proof*, i.e. the name of the intended verifier,
4) $s := r + ex \bmod q$,
5) send (e, s) to the verifier.

Then the verifier performs the following steps:

1) $k' := g^s/a^e$,
2) $\ell' := h^s/b^e$,
3) $e' := H(k', \ell', g, h, a, b, m)$,
4) return $\mathtt{ok}$ if $e = e'$.

B. Sketch of the Scheme

The system is supported by a card management system called below CAMS. We refer also to standard protocols for chip authentication (Chip Authentication or ChA) and authenticating terminals (Terminal Authentication or TA) [1].

1) Document personalization.: For each single identity document the following steps are executed by issuing authority:

1) All but two data groups for the e-ID are completed in advance, and are stored in some registry on the side of CAMS.
2) The data groups are copied to the chip of e-ID.
3) The private key and the corresponding public key for ChA are generated by the e-ID chip.
4) The ChA public key is copied to the data groups (i.e., to a copy stored locally on the e-ID chip as well to a copy stored in the registry of CAMS).

The data groups are still not authenticated by the issuing authority. The e-ID is in a state we call "red", which means that all functions of the chip are blocked – only Terminal Authentication and Chip Authentication with terminals of CAMS are allowed.

When the e-ID is in hands of its owner, it must be unblocked. In a private environment the owner connects to a service of CAMS and after mutual authentication via TA and ChA protocols the following steps are executed:

1) The e-ID chip generates its private key $\tilde{x}$ for tagging, and computes $\tilde{a} = g^{\tilde{x}}$, where g is fixed for all users.
2) Key $\tilde{a}$ is written in the remaining empty data group, both in the e-ID chip and in its record in the CAMS registry.
3) The e-ID chip and CAMS each compute $\tilde{h} = H_g(D)$, where H_g is a hash function with the image included in the group generated by g.
4) The e-ID chip computes $\tilde{b} = \tilde{h}^{\tilde{x}}$ and sends $\tilde{b}$ to CAMS.
5) The e-ID chip and CAMS execute zero-knowledge protocol for equality of discrete logarithms for $\tilde{a}, \tilde{b}$ and the corresponding bases $g, \tilde{h}$ (here Schnorr-like protocol described above has to be used, m is chosen to be the string "CAMS").
6) The e-ID chip enters a "yellow" state, which is intermediate between the red one and the "green" one for regular usage. The e-ID chip disconnects from CAMS.

The next phase is generating signature of the issuing authority:

1) User's data groups from CAMS's registry are transferred together with the proof of equality of discrete logarithms to the document issuing authority.
2) The document issuing authority verifiers the proof and if the verification result is positive, then it creates a signature $Sign(\tilde{b})$ under $\tilde{b}$.
3) $Sign(\tilde{b})$ is transferred back to CAMS's registry.

If an e-ID is in the "yellow" state, then any time the e-ID is used it tells the middle-ware to connect to CAMS's service to fetch $Sign(\tilde{b})$. If the signature is available, it is transferred to the chip of e-ID through a secure channel (established by means of TA and ChA protocols). The e-ID verifies the signature, if it is correct, then the e-ID switches from the "yellow" state to the "green" one.

2) Data Group Authentication: To execute this part the e-ID must be in "green" state. After completion of the terminal authentication and the chip authentication procedures the terminal of the verifier and the e-ID chip execute the following protocol (we assume that the terminal is allowed to obtain the whole data D):

1) The e-ID chip sends D and $Sign(\tilde{b})$ to the terminal.
2) The terminal reads $\tilde{a}$ from D and computes $\tilde{h} = H_g(D)$.
3) The e-ID chip computes $\tilde{h} = H_g(D)$ and $\tilde{b} = \tilde{h}^{\tilde{x}}$ and sends $\tilde{b}$ to the terminal (now both sides know the tuple $(\tilde{a}, \tilde{b}, g, \tilde{h})$ and $Sign(\tilde{b})$, but the link between $\tilde{h}$ and $\tilde{b}$ must be proven by the e-ID chip).
4) Both parties execute equality of discrete logarithms protocol for $\tilde{a}, \tilde{b}$ and the corresponding bases $g, \tilde{h}$. Schnorr-like protocol is used for m being a string identifying the verifier.

C. Discussion

As in case of the protocol from Section II the issuing authority cannot create a clone of an e-ID document without breaking into the e-ID chip and reading the secrets installed there by the owner of the document.

Unlike in the previous solution, we are free to make tags as precise as we want: the message m included in the proof of equality of discrete logarithms may fully indicate the verifier's identity. On the other hand, it is also possible to insert restricted information only – as for the protocol from Section II. In the former case the tags are undeniable proofs that an e-ID has issued a customized signature for the verifier indicated in the tag.

Apart from tagging, an e-ID document may check the rights of the terminal to get the data. This can be achieved in a standard way where the terminals are authenticated by certificates and the underlying PKI infrastructure (compare [6]).

Finally, let us remark that despite cryptographic countermeasures and legal restrictions, any party can sell a set of *unauthenticated* personal data. Authentication may be statistical – the party buying the set may confront it with the set of locally stored data. If the records belonging to the intersection set are the same, then the whole set bought is assumed to be correct[1]. In order to prevent such situation, the e-ID could insert steganographic data in images revealed to the verifiers (with such steganographic tags the data would depend from the intended addressee). However, it is a hard challenge to design such protocols: apart from all problems known so far for steganographic security measures we have to deal with the problem of low computational resources on the e-ID chip.

Another option to limit illegal selling of personal data, which may always undergo statistical verification, is to require by law that each record containing personal data should be associated with

- a tag proving that the party that stores the record has obtained it directly from the smart card,
- or a consent signed by the person for selling/revealing her/his data,
- or a pointer to some legal regulations that imposes a duty on the party to process the data (however, the data should still be associated with the tags, indicating whom the data were initially revealed by smart cards).

Then in case of an audit a party that stores the data is safe.

Moreover, each party that sees personal data with the tag issued for another party, and without consent of the citizen for selling/revealing her/his data, should be obliged by law to inform the authorities about the leak (the data seen should be attached to the information). In cases when a party is legally binded to reveal the data to another party it should obtain a signed request for the data, to avoid being accused for data leakage.

IV. Security of the Asymmetric Approach

A. Problem Statement

The exponentiation $\tilde{h}^{\tilde{x}}$, where $\tilde{h} = H_g(D)$, used in the protocol from Section III resembles BLS signature scheme [7]. However, if $\langle g \rangle$ would be a pairing friendly group, no ZKP-EDLP (Zero-Knowledge Proof of Equality of Discrete Logarithms) would be necessary, because equality could immediately be checked with pairing.

[1]See that if the issuing authority creates a duplicate of a document with the same personal data but with different key material, then it could be detected by parties already storing data from the original document. Of course a list of revoked chips should be available online to prevent misuse of cards stolen or lost.

Thus augmenting the exponentiation with ZKP-EDLP we obtain an analog of BLS signature scheme in pairing unfriendly groups. Since D is of the form $(g^{\tilde{x}}, M)$, where M are some data, we obtain a kind of a self-signed certificate of the public key $\tilde{a} = g^{\tilde{x}}$. The document issuing authority makes signature $Sign(\tilde{b})$ under the BLS-like "signature" value $\tilde{b} = \tilde{h}^{\tilde{x}}$.

Problem: *is it feasible to change M and tune $\tilde{x}$ accordingly in such a way that $\tilde{b}$ remains unchanged?* The protocol from Section III assumes negative answer to this question.

B. Argument for Security

We have Schnorr-like dependency here: some randomizer is used inside and outside the hash function: $\tilde{b} = (H_g(g^{\tilde{x}}, M))^{\tilde{x}}$. Hence when we try to change M to M' we search for $x' \in \mathbb{Z}_q^*$ yielding a collision:

$$\tilde{b}^{(x')^{-1}} = H_g(g^{x'}, M').$$

Probability of such an event is not greater than probability of the following collision

$$\tilde{b}^{(x')^{-1}} = H_g(y, M'),$$

where x', y could be independently chosen. But the latter collision occurs no more frequently than the collision

$$\tilde{b}^{(x')^{-1}} = H_g(\tilde{M}), \tag{1}$$

where $\tilde{M}$ could be any bitstring. In the random oracle model for H_g probability of the last event results from the birthday paradox in two rooms setting: Let fix parameter $\gamma \in (0, 1)$. Provided that in each single choice of $(x', \tilde{M})$ an element $\tilde{b}^{(x')^{-1}} \in \mathrm{Im}(H_g)$, the number of choices $(x', \tilde{M})$ yelding collision (1) with probability no smaller than γ is equal to $c_\gamma \cdot \sqrt{|\mathrm{Im}(H_g)|}$, where constant c_γ results from the birthday paradox mentioned above, and is dependent of γ. Since $x', \tilde{M}$ could be chosen independently, the expected number of choices of $(x', \tilde{M})$ to obtain a collision (1) with probability no smaller than γ, equals in the random oracle model for H_g to

$$\frac{c_\gamma \cdot \sqrt{|\mathrm{Im}(H_g)|}}{\Pr\left(\tilde{b}^{(x')^{-1}} \in \mathrm{Im}(H_g)\right)}.$$

V. Conclusions

It turns out that protection of high quality personal data disclosed by personal identity cards is feasible in the model in which there are trust limitations against smart cards manufacturers and authorities issuing the identity documents. Moreover, standard smart cards with cryptographic functions can be used for implementing such a solution.

REFERENCES

[1] BSI, "Advanced Security Mechanisms for Machine Readable Travel Documents 2.10," Technische Richtlinie TR-03110, 2010.

[2] M. L. Das, A. Saxena, and D. B. Phatak, "Algorithms and approaches of proxy signature: A survey," *I. J. Network Security*, vol. 9, no. 3, pp. 264–284, 2009.

[3] J. Monnerat, S. Pasini, and S. Vaudenay, "Efficient deniable authentication for signatures," in *ACNS*, ser. Lecture Notes in Computer Science, M. Abdalla, D. Pointcheval, P.-A. Fouque, and D. Vergnaud, Eds., vol. 5536. Springer, 2009, pp. 272–291.

[4] R. C. Merkle, "Secrecy, authentication, and public key systems." Ph.D. dissertation, Stanford University, Stanford, CA, USA, 1979, aAI8001972.

[5] D. Chaum and T. P. Pedersen, "Wallet databases with observers," in *CRYPTO*, ser. Lecture Notes in Computer Science, E. F. Brickell, Ed., vol. 740. Springer, 1992, pp. 89–105.

[6] BSI, "PKIs for Machine Readable Travel Documents 1.10," Technische Richtlinie BSI TR-03129, 2009.

[7] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the Weil pairing," *J. Cryptology*, vol. 17, no. 4, pp. 297–319, 2004.

Przemysław Kubiak Przemysław Kubiak received master degree in mathematics in 1997 and Phd in 2001. He is an assistant professor for computer science at Wrocław University of Technology.

His research interests are in efficient algorithms for public-key cryptography, and in reducing the need for trust in components of cryptographic protocols.

Mirosław Kutyłowski
Miroslaw Kutylowski is a full professor of computer science at Wroclaw University of Technology, member of Scientific Council of Institute of Computer Science, Polish Academy of Sciences. He received PhD and Habilitation degree from Wroclaw University. Former Alexander von Humboldt-Fellow at Darmstadt Technical University, Hochschuldozent at Paderborn University. He received Mistrz Prize from Foundation for Polish Science in 2009 and IBM Faculty Award in 2012. Member of Steering Committee of ESORICS.

Prof. Kutylowski specializes in ad hoc systems, privacy and security, including in particular protocols for personal identity cards.

Wojciech Wodo
Wojciech Wodo is a PhD candidate of computer science at Wroclaw University of Technology, graduate of special "Top 500 Innovators" program at UC Berkeley focused on science management, technology transfer, commercialization and university-industry collaboration. He received MSc degree from Wroclaw University of Technology in computer science. Mr. Wodo worked as a technology transfer specialist in Wroclaw Research Center EIT+ (2011-2012).

He is a vice-president of MANUS Foundation , responsible for academic entrepreneurship development and consultancy for start-ups.

His field of interest are identifications and authentications methods based on biometric factors as well as cryptographic protocols.

Classes of Garbling Schemes

Tommi Meskanen, Valtteri Niemi, Noora Nieminen

***Abstract*—Bellare, Hoang and Rogaway elevated *garbled circuits from a cryptographic technique to a cryptographic goal* by defining several new security notions for garbled circuits [3]. This paper continues at the same path by extending some of their results and providing new results about the classes of garbling schemes defined in [3]. Furthermore, new classes of garbling schemes are defined and some results concerning them and their relation to earlier classes are proven.**

***Index Terms*—garbled circuits, garbling schemes, secure multiparty computations, privacy**

I. Introduction

The history of garbled circuits traces back to A. Yao, who introduced the technique in [7]. The term *garbled circuit* was introduced by Beaver, Micali and Rogaway [2] where they introduced a way of performing secure multiparty computation with Yao's circuit garbling technique. Since then Yao's garbled circuits have been used for various purposes even though there was no formal definition what is meant by garbling. No proof of security existed either - until Lindell and Pinkas introduced one for a particular garbled circuit using a protocol assuming semi-honest adversaries [5], [6]. After this result, also a proof of security against covert and malicious adversaries has been published [1], [6]. Again, these results are obtained for a specific protocol using garbling schemes rather than considering the security of garbling itself.

The first formal definition of a garbling scheme has recently been proposed by Bellare, Hoang and Rogaway in [3]. A garbling scheme is defined as a five-tuple of functions: the actual garbling procedure $\mathtt{Gb}$, the encryption function $\mathtt{En}$, the decryption function $\mathtt{De}$, the garbled evaluation function $\mathtt{Ev}$ and the original evaluation function $\mathtt{ev}$. The idea behind garbling is the following. Let f be a function which is to be evaluated for different inputs x but in such a way that neither f nor x can be learnt from the evaluation process. Therefore, a garbled version F is created and instead of computing $y = \mathtt{ev}(f, x)$ we compute $Y = \mathtt{Ev}(F, X)$ where X is obtained from x by encryption. After this y is obtained from Y by decryption. Figure 1 illustrates the garbling procedure.

Rogaway et al. define also three security notions for garbling schemes. These notions are expressed via code-based games which are defined in such a way that they capture the intuition behind the different notions: privacy, obliviousness and authenticity are all defined to be reached, if the adversary has only a negligible advantage for winning a particular game. Moreover, these notions have two different models,

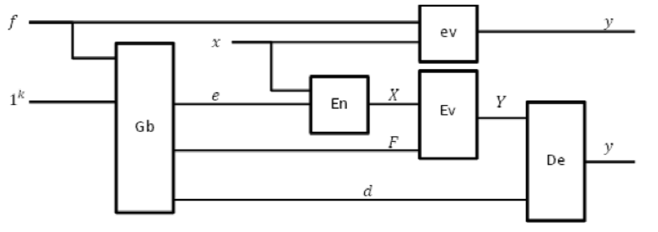


Figure 1: Description of the technique behind garbling. The diagram also illustrates that the result of evaluation with garbling must coincide with the result obtained without garbling.

either based on indistinguishability or simulation. Roughly speaking, indistinguishability means that the adversary cannot distinguish between garblings of two functions. The simulation type means that an adversary is incapable of distinguishing garbling of the function of its own choice from another similar looking function devised by a simulator. Here we refer to the next section for the formal definitions.

Another seminal achievement in [3] is that relations between the different security notions have been proven. Rogaway et al. also provide two concrete garbling schemes, one of which achieves not only privacy but also obliviousness and authenticity. This example assures that the defined security classes are not empty.

This paper consists of three sections. In the first section we define all the necessary concepts, and give an informal description of them so that the idea behind the concept would be more comprehensive to the reader. In the second section we provide new results about the already known classes: some of the results are extensions to the results in [3], some inspired by the results in [3]. The third section provides modified definitions of the games used to define the different security notions. In this manner, we obtain new classes of garbling schemes by minor modifications in the games. Then, we prove some relations not only between the new and existing classes but also among the new classes. We also discuss intuition behind these new classes.

II. Definitions

In this section, we provide the basic definitions and notations. As usual, $\mathbb{N}$ will be the set of positive integers. A *string* is a finite sequence of bits. In addition to the basic strings, there is a special symbol $\perp$. The meaning of this symbol is explained later where the context of usage will be clearer.

Let A be a finite set. Notation $y \twoheadleftarrow A$ means that an element is selected uniformly at random from the set A, and this element is assigned to y. If A denotes an algorithm, then notation $A(x_1, \ldots, x_n)$ means the output of the algorithm A on inputs $x_1, \ldots, x_n$.

T. Meskanen is a researcher at the Department of Mathematics and Statistics, University of Turku, Finland (email: tommes@utu.fi).

V. Niemi is a professor at the Department of Mathematics and Statistics, University of Turku, Finland (email: pevani@utu.fi).

N. Nieminen is a doctoral student at Turku Centre for Computer Science, University of Turku, Finland (email: nmniem@utu.fi).

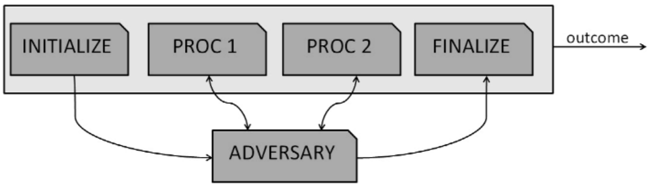


Figure 2: The idea of a code-based game is captured in the above image.

As usual, we say that a function $f : \mathbb{N} \to \mathbb{R}$ is negligible if for every $c > 0$ there is an integer N_c such that $|f(x)| < x^{-c}$ for all $x > N_c$

A. Code-based games

The proofs in this paper are heavily based on *code-based games*. Following the terminology presented in [4], a game is a collection of procedures called *oracles*. This collection may contain three types of procedures: `INITIALIZE`, `FINALIZE` and other named oracles. The word *may* is used, since all the procedures in a game are optional.

The entity playing a game is called *adversary*. When the game is run with an adversary, first the `INITIALIZE` procedure is called. It possibly provides an input to the adversary, who in turn may invoke other procedures before feeding its output to the `FINALIZE` procedure. The `FINALIZE` receives an output of the adversary, and creates a string that tells the outcome from the game typically consisting of one bit of information: whether the adversary has won or not. This description about code-based games is quite informal and gives only the intuition behind the concept. The Figure 2 serves as an illustration. For a more formal description, we refer to [4].

B. Garbling schemes

In this section we provide a formal definition of garbling schemes and their security, and here we follow the guidelines provided in [3].

Formally, a garbling scheme is a 5-tuple $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev})$ of algorithms, from which the first is probabilistic and the rest are deterministic. Let f denote the string that represents the original function. The last component in the 5-tuple is the evaluation function $\mathtt{ev}(f, \cdot) : \{0,1\}^n \to \{0,1\}^m$ which we want to garble. Here, the values $n = f.n$ and $m = f.m$ represent the lengths of the input x and the output $y = ev(f, x)$. They must also be efficiently computable from f. The first component $\mathtt{Gb}$ denotes the garbling algorithm. It takes f and 1^k as its inputs, where $k \in \mathbb{N}$ is a security parameter, and returns (F, e, d) on this input. String e describes the encryption algorithm $\mathtt{En}(e, \cdot)$ which maps an initial input x to a garbled input $X = \mathtt{En}(e, x)$. String F describes the garbled function $\mathtt{Ev}(F, \cdot)$. It returns the garbled output $Y = \mathtt{Ev}(F, X)$. Finally, string d describes the decryption algorithm $\mathtt{De}(d, \cdot)$ which on a garbled input returns the final output $y = \mathtt{De}(d, Y)$. Here we refer to Figure 1 to get an idea of how a garbling scheme works.

`NOTE`: Occasionally, we use a specific evaluation function $\mathtt{ev}_{circ}$ as $\mathtt{ev}$ in the 6-tuple. For it, we first define *a conventional circuit* by a 6-tuple $f = (n, m, q, A, B, G)$. The first component denotes the number of input wires $(n \geq 2)$, the second is the number of output wires $(m \geq 1)$, and the third component represents the number of gates $(q \geq 1)$ in the circuit. The function A identifies the first incoming wire, whereas B identifies the second incoming wire of each gate. The remaining component G is a function identifying the functionality of each gate. For a more specific definition of a circuit, see [3]. Finally, the circuit evaluation function $\mathtt{ev}_{circ}$ is the usual canonical evaluation function:

proc $\mathtt{ev}_{circ}(f, x)$
$(n, m, q, A, B, G) \leftarrow f$
for $g \leftarrow n+1$ **to** $n+q$ **do** $a \leftarrow A(g), b \leftarrow B(g), x_g \leftarrow G_g(x_a, x_b)$
return $x_{n+q-m+1} \cdots x_{n+q}$

There are some additional requirements that garbling schemes must fulfill. These are *length, non-degeneracy* and *correctness* conditions. The length condition means that the lengths of F, e, d may only depend on the security parameter k, the values $f.n, f.m$ and the length of the string f. Non-degeneracy condition means the following: if $f.n = g.n$, $f.m = g.m$, $|f| = |g|$, $(F, e, d) = \mathtt{Gb}(1^k, f; r)$ and $(G, e', d') = \mathtt{Gb}(1^k, g; r)$ where r represents random coins of $\mathtt{Gb}$, then $e = e'$ and $d = d'$. Correctness requires that $\mathtt{De}(d, \mathtt{Ev}(F, \mathtt{En}(e, x)))$ will always give the same result as $\mathtt{ev}(f, x)$.

By the concept of *a side-information function*, we capture the information revealed about f by the garbling process. In the case of circuits and $\mathtt{ev}_{circ}$, this might be the size of the circuit that was garbled, the topology of it or something else - even the whole initial circuit. Formally, a side-information function Φ deterministically maps string f to string $\Phi(f)$. Let $f = (n, m, q, A, B, G)$ be a circuit. Then, we define $\Phi_{size}(f) = (n, m, q)$, which is the side-information function revealing the size of the garbled circuit. Other side-information functions are $\Phi_{circ}(f) = f$ which thus reveals the entire circuit, and Φ_{topo} which reveals the topology of the initial circuit, i.e. $\Phi_{topo} = (n, m, q, A, B)$.

C. The security notions of garbling schemes

There are three types of security: *privacy, obliviousness* and *authenticity*. The first two types also have two distinct models: one based on *indistinguishability* and another based on *simulation*. In all cases, the security is defined through a code-based game consisting of a procedure named `GARBLE` and finalization procedure `FINALIZE`. The procedure `GARBLE` is not to be confused with the garbling function $\mathtt{Gb}$: the garbling function $\mathtt{Gb}$ is a component of a garbling scheme $\mathcal{G}$, whose security the adversary tries to break via the procedure `GARBLE`.

Before the game starts, the garbling scheme $\mathcal{G}$ and the side-information function Φ are fixed in the games based on indistinguishability model. In simulation model, also the simulator $\mathcal{S}$ is fixed although details of it are not assumed to be known to the adversary. The `GARBLE` procedure gives the challenge of the game to the adversary and the `FINALIZE` procedure determines whether the adversary wins the game or not. The adversary is assigned a certain advantage depending on the probability of winning the game. This advantage in turn determines whether the garbling scheme is secure or not.

Table 1 gives the different GARBLE procedures needed in the games to define different security notions. Note that in this first formal description we use the subscripts, but after that, we omit them if they are clear from the context. For example, we will write *PrvSim* game instead of $PrvSim_{\mathcal{G},\Phi,\mathcal{S}}$.

Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev})$ be a garbling scheme, $k \in \mathbb{N}$ a security parameter and Φ a side-information function. The following definitions are informal, and they are mentioned to capture the idea behind the security notions. For a more formal treatment, see [3].

PRIVACY: Privacy has two types of notions, and hence there are two different games with distinct GARBLE procedures, $PrvInd_{\mathcal{G},\Phi}$ and $PrvSim_{\mathcal{G},\Phi,\mathcal{S}}$. The biggest difference between these two is that the latter requires an auxiliary algorithm to be defined, namely the simulator $\mathcal{S}$.

The game $PrvInd$ consists of a GARBLE procedure, which is called by the adversary *exactly once* during one game, and a FINALIZE procedure. Informally the game goes as follows: the adversary calls the GARBLE procedure having two appropriate functions and their inputs as the feed. The procedure returns a garbled version of one of the functions and its input, and the adversary guesses which of the functions got garbled. The FINALIZE procedure takes two inputs, value of parameter b from GARBLE and adversary's guess b', and tells whether the answer given by the adversary was correct or not, and this will then be the outcome of the game.

The game $PrvSim$ has also two procedures, its own GARBLE and FINALIZE, from which the latter has the same functionality as in $PrvInd$ game. The difference in GARBLE procedure is, that now the other function, from which the function f is to be distinguished, is devised by the simulator. The adversary must tell the difference between an actual function and a "fake" function.

We define the advantage of an adversary A in game $PrvInd$ as follows:

$$\mathbf{Adv}_{\mathcal{G}}^{prv.ind,\Phi}(\mathcal{A},k) = 2 \cdot \Pr\left[PrvInd_{\mathcal{G},\Phi}^{\mathcal{A}}(k)\right] - 1.$$

If the advantage function $\mathbf{Adv}_{\mathcal{G}}^{prv.ind,\Phi}(\mathcal{A},\cdot)$ is negligible for all PT adversaries $\mathcal{A}$ then we say that the garbling scheme $\mathcal{G}$ is *prv.ind secure over* Φ. Similarly, we define the advantage of an adversary $\mathcal{B}$ in game $PrvSim$ as $\mathbf{Adv}_{\mathcal{G}}^{prv.sim,\Phi,\mathcal{S}}(\mathcal{B},k) = 2 \cdot \Pr\left[PrvSim_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k)\right] - 1$. Then, we define that a garbling scheme $\mathcal{G}$ is *prv.sim secure over* Φ if for every PT adversary there exists a PT simulator $\mathcal{S}$ such that $\mathbf{Adv}_{\mathcal{G}}^{prv.sim,\Phi,\mathcal{S}}(\mathcal{B},k)$ is negligible.

OBLIVIOUSNESS: At first sight, the games for obliviousness seem similar to the privacy games. The difference is that the decryption algorithm d is not given to the adversary, and hence the adversary cannot compute the final output $y = \mathtt{De}(d, \mathtt{Ev}(F,X))$. Informally, the adversary is asked to distinguish two functions and their inputs from each other without knowing the result of evaluation.

The adversary has an advantage which is calculated as in the privacy model. The obv.ind and the obv.sim security of a garbling scheme $\mathcal{G}$ are defined similarly as in the corresponding Prv-games.

AUTHENTICITY: Here the FINALIZE procedure is a little more complex than in the two cases above. The finalization procedure of a game checks whether the adversary is able to produce a valid garbled output Y different to $\mathtt{Ev}(F,X)$ or not. Also the advantage function is slightly different: $\mathbf{Adv}_{\mathcal{G}}^{aut}(\mathcal{A},k) = \Pr\left[Aut_{\mathcal{G}}^{\mathcal{A}}(k)\right]$. Again, a garbling scheme is aut-secure, if for all polynomial time adversaries $\mathcal{A}$ the advantage function $\mathbf{Adv}_{\mathcal{G}}^{aut}(\mathcal{A},\cdot)$ is negligible.

We denote $\mathtt{GS}(xxx,\Phi)$ to be the set of all garbling schemes that are $xxx-$secure over the side-information function Φ, where xxx denotes the type of security: $prv.ind$, $prv.sim$, $obv.ind$, $obv.sim$, $mod.ind$, $mod.sim$, $mod.ind2$ or $mod.sim2$. The notion $\mathtt{GS}(aut)$ means the set of all $aut-$secure garbling schemes. $\mathtt{GS}(ev)$ means the class of garbling schemes which use the evaluation function $\mathtt{ev}$.

III. RESULTS ABOUT ESTABLISHED CLASSES OF GARBLING SCHEMES

In this section we provide results concerning the security classes *prv.ind, prv.sim, obv.ind, obv.sim* defined in section 2. The first two theorems consider the effect of different side-information functions to the sets of garbling schemes. The following two theorems provide extensions to the existing results in [3] – the non-inclusions are obtained for any side-information function Φ instead of restricting it to Φ_{topo}. Then we continue with two results that provide parallel results to [3]. Finally, the last two theorems in this section provide new results about the established security classes of garbling schemes.

Theorem 1: Suppose that two different side-information functions Φ_a and Φ_b satisfy the condition

$$\Phi_a(f_0) = \Phi_a(f_1) \Rightarrow \Phi_b(f_0) = \Phi_b(f_1). \qquad \textbf{(Condition } (*))$$

Then we have the inclusion $\mathtt{GS}(prv.ind,\Phi_b) \subseteq \mathtt{GS}(prv.ind,\Phi_a)$. If we additionally assume that there exists a polynomial time function g such that $g(\Phi_a(f)) = \Phi_b(f)$ then we also have $\mathtt{GS}(prv.sim,\Phi_b) \subseteq \mathtt{GS}(prv.sim,\Phi_a)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(prv.ind,\Phi_b)$. Suppose now that $\mathcal{A}$ is an arbitrary adversary playing the $PrvInd_{\Phi_a}$ game and let us construct $\mathcal{B}$ as an adversary playing the $PrvInd_{\Phi_b}$ game and using $\mathcal{A}$ as a subroutine. The latter adversary $\mathcal{B}$ tells the first adversary $\mathcal{A}$ to start the game. Adversary $\mathcal{A}$ chooses its input (f_0, f_1, x_0, x_1) which it wants to send to GARBLE procedure, which now in fact the adversary $\mathcal{B}$ pretends to be. Adversary $\mathcal{B}$ forwards the input from $\mathcal{A}$ to GARBLE procedure in $PrvInd_{\Phi_b}$ game. Adversary $\mathcal{B}$ receives an output (F,X,d) or $\perp$ from GARBLE. Now, if $\Phi_b(f_0) \neq \Phi_b(f_1)$, adversary $\mathcal{B}$ sends $\perp$ to $\mathcal{A}$. This is the normal answer: According to our assumption, $\Phi_b(f_0) \neq \Phi_b(f_1) \Rightarrow \Phi_a(f_0) \neq \Phi_a(f_1)$ and hence adversary $\mathcal{A}$ should receive $\perp$ also from its genuine GARBLE procedure. Otherwise, adversary $\mathcal{B}$ forwards the response from

<table>
<tr><td>proc GARBLE(f_0, f_1, x_0, x_1) Game PrvInd$_{\mathcal{G},\Phi}$
$b \twoheadleftarrow \{0,1\}$
if $\Phi(f_0) \neq \Phi(f_1)$ then return $\perp$
if $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ then return$\perp$
if $\mathtt{ev}(f_0, x_0) \neq \mathtt{ev}(f_1, x_1)$ then return $\perp$
$(F, e, d) \leftarrow \mathtt{Gb}(1^k, f_b)$; $X \leftarrow \mathtt{En}(e, x_b)$;
return (F, X, d)</td><td>proc GARBLE(f, x) Game PrvSim$_{\mathcal{G},\Phi,\mathcal{S}}$
$b \twoheadleftarrow \{0,1\}$
if $x \notin \{0,1\}^{f.n}$ then return $\perp$
if $b = 1$ then $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$; $X \leftarrow \mathtt{En}(e, x)$
else $y \leftarrow \mathtt{ev}(f, x)$; $(F, X, d) \leftarrow \mathcal{S}(1^k, y, \Phi(f))$
return (F, X, d)</td></tr>
<tr><td>proc GARBLE(f_0, f_1, x_0, x_1) Game ObvInd$_{\mathcal{G},\Phi}$
$b \twoheadleftarrow \{0,1\}$;
if $\Phi(f_0) \neq \Phi(f_1)$ then return $\perp$
if $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ then return$\perp$
$(F, e, d) \leftarrow \mathtt{Gb}(1^k, f_b)$; $X \leftarrow \mathtt{En}(e, x_b)$;
return (F, X)</td><td>proc GARBLE(f, x) Game ObvSim$_{\mathcal{G},\Phi,\mathcal{S}}$
$b \twoheadleftarrow \{0,1\}$
if $x \notin \{0,1\}^{f.n}$ then return $\perp$
if $b = 1$ then $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$; $X \leftarrow \mathtt{En}(e, x)$
else $(F, X) \leftarrow \mathcal{S}(1^k, \Phi(f))$
return (F, X)</td></tr>
<tr><td colspan="2">proc FINALIZE(b, b') Game PrvInd$_{\mathcal{G},\Phi}$, Game PrvSim$_{\mathcal{G},\Phi,\mathcal{S}}$, Game ObvInd$_{\mathcal{G},\Phi}$, Game ObvSim$_{\mathcal{G},\Phi,\mathcal{S}}$
return $b = b'$</td></tr>
<tr><td>proc GARBLE(f, x) Game Aut$_{\mathcal{G}}$
$(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$; $x \leftarrow \mathtt{En}(e, x)$
return (F, X)</td><td>proc FINALIZE(Y) Game Aut$_{\mathcal{G}}$
return $\mathtt{De}(d, Y) \neq \perp$ and $Y \neq \mathtt{Ev}(F, X)$</td></tr>
</table>

Table I: The games defining the different security notions

its GARBLE to $\mathcal{A}$, who then sends its answer b' to $\mathcal{B}$. The adversary $\mathcal{B}$ answers the same b' in its $PrvInd_{\Phi_b}$ game.

Let us now consider the winning probabilities and advantages of both adversaries in their games. The behavior of adversaries $\mathcal{A}$ and $\mathcal{B}$ are the same at every step of the game: the inputs are the same, and the answers are the same. Therefore the probability of the answer b' being the correct one must be the same in both games. Hence the advantages of both adversaries are also equal. Because $\mathcal{G} \in \mathtt{GS}(prv.ind, \Phi_b)$ the advantage of $\mathcal{B}$ in $PrvInd_{\Phi_b}$ game is negligible. Thus, the advantage of $\mathcal{A}$ is also negligible, and $\mathcal{G} \in \mathtt{GS}(prv.ind, \Phi_a)$, which proves the claim.

For the second part, let us assume that there exists an efficient conversion g from the side-information function Φ_a into Φ_b. Our objective is to prove under these assumptions that $\mathtt{GS}(prv.sim, \Phi_b) \subseteq \mathtt{GS}(prv.sim, \Phi_a)$.

To do this, assume that $\mathcal{G} \in \mathtt{GS}(prv.sim, \Phi_b)$. This means that for every polynomial time adversary $\mathcal{A}'$ there exists a simulator $\mathcal{S}$ such that the advantage of $\mathcal{A}'$ is negligible in $PrvSim_{\mathcal{G},\Phi_b,\mathcal{S}}$ game.

Let $\mathcal{A}$ be an arbitrary adversary playing $PrvSim_{\mathcal{G},\Phi_a,\mathcal{S}}$ games. Similarly to the first part of the proof, let $\mathcal{B}$ be an adversary who plays $PrvSim_{\mathcal{G},\Phi_b,\mathcal{S}}$ games by emulating $\mathcal{A}$, i.e. behaving just like $\mathcal{A}$ would behave in corresponding $PrvSim_{\mathcal{G},\Phi_a,\mathcal{S}}$ games. More precisely, by emulation of $\mathcal{A}$ we mean the following. First, adversary $\mathcal{B}$ tells $\mathcal{A}$ to start its game. Adversary $\mathcal{B}$ receives the GARBLE input (f, x) from $\mathcal{A}$, after which $\mathcal{B}$ forwards this input to its own GARBLE. This procedure returns (F, X, d) or $\perp$ to $\mathcal{B}$, who now consults adversary $\mathcal{A}$ by giving this output to him. Now, $\mathcal{A}$ returns b' to $\mathcal{B}$, who chooses the same b' as its own return value.

The assumption $\mathcal{G} \in \mathtt{GS}(prv.sim, \Phi_b)$ implies that there exists a simulator $\mathcal{S}_{hard}$ such that the advantage of $\mathcal{B}$ is negligible in $PrvSim_{\mathcal{G},\Phi_b,\mathcal{S}_{hard}}$ game. Now, we define another simulator $\mathcal{S}'_{hard}$ by $\mathcal{S}'_{hard}(1^k, y, \Phi_a(f)) = \mathcal{S}_{hard}(1^k, y, g(\Phi_a(f)))$. First of all, $\mathcal{S}'_{hard}$ is polynomial time, because the conversion g is efficient and $\mathcal{S}_{hard}$ is a polynomial time simulator. Secondly, the win probability of $\mathcal{B}$ in its own $PrvSim_{\mathcal{G},\Phi_b,\mathcal{S}_{hard}}$ game is the same as the win probability that $\mathcal{A}$ has in the $PrvSim_{\mathcal{G},\Phi_a,\mathcal{S}'_{hard}}$ game, which implies equal advantages. By assumption, the advantage of $\mathcal{B}$ was negligible, and so is the advantage of $\mathcal{A}$ by the above argument. Now we have found a simulator against which $\mathcal{A}$ has a negligible advantage. □

NOTE: For example, $\Phi_a = \Phi_{topo}$ and $\Phi_b = \Phi_{size}$ satisfy the condition $(*)$.

Theorem 2: Let Φ_a and Φ_b be two different side-information functions satisfying the above condition $(*)$. Then the following inclusion holds: $\mathtt{GS}(obv.ind, \Phi_b) \subseteq \mathtt{GS}(obv.ind, \Phi_a)$. If we additionally assume that there exists a polynomial time function g such that $g(\Phi_a(f)) = \Phi_b(f)$ then we have also $\mathtt{GS}(obv.sim, \Phi_b) \subseteq \mathtt{GS}(obv.sim, \Phi_a)$.

Proof: The proof is similar to that of the previous theorem. □

The next four theorems consider non-inclusions of the form $A \not\subseteq B$ between sets of garbling schemes. In all cases we make an assumption that the set A is non-empty. The following two propositions provide a generalization to Propositions 5 and 7 in paper [3].

Theorem 3: For all Φ and for $\mathtt{ev} = \mathtt{ev}_{circ}$, we have $\mathtt{GS}(obv.sim, \Phi) \bigcap \mathtt{GS}(\mathtt{ev}) \not\subseteq \mathtt{GS}(prv.ind, \Phi)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(obv.sim, \Phi) \bigcap \mathtt{GS}(\mathtt{ev})$. Let us construct another garbling scheme $\mathcal{G}' = (\mathtt{Gb}', \mathtt{En}, \mathtt{De}', \mathtt{Ev}, \mathtt{ev})$ such that $\mathcal{G}' \in \mathtt{GS}(obv.sim, \Phi) \bigcap \mathtt{GS}(\mathtt{ev})$ but $\mathcal{G}' \notin \mathtt{GS}(prv.ind, \Phi)$. The construction is as follows: The function $\mathtt{Gb}'(1^k, f)$ picks $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$ and returns $(F, e, d||e)$. Let $\mathtt{De}'(d||e, Y) = \mathtt{De}(d, Y)$. Including e in the description of the decoding function does not harm $obv.sim$ security, because the adversary is given only (F, X) by the GARBLE procedure in the $obv.sim$ game. Thus $\mathcal{G}'$ inherits the $obv.sim$ security from $\mathcal{G}$.

On the other hand, $\mathcal{G}'$ is not $prv.ind$ secure. Adversary $\mathcal{A}$ makes a query (f_0, f_1, x_0, x_1), where $f_0 = f_1 = \mathtt{AND}$ and $x_0 = 00, x_1 = 01$. This choice is fine for the PrvInd game, since $\mathtt{ev}(f_0, x_0) = 0 = \mathtt{ev}(f_1, x_1)$. Now, the adversary computes $X_0 = \mathtt{En}(e, x_0)$ and $X_1 = \mathtt{En}(e, x_1)$, which must be different because of the non-degeneracy condition (see Section 2). Then he/she compares these two with the garbled input X received from GARBLE. This comparison now reveals which of the inputs, x_0 or x_1, was used. $\square$

Theorem 4: For all Φ and for $\mathtt{ev} = \mathtt{ev}_{circ}$, we have $\mathtt{GS}(aut) \bigcap \mathtt{GS}(\mathtt{ev}) \nsubseteq \mathtt{GS}(prv.ind, \Phi) \bigcup \mathtt{GS}(obv.ind, \Phi)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(aut) \bigcap \mathtt{GS}(\mathtt{ev})$. Let us construct a garbling scheme $\mathcal{G}' = (\mathtt{Gb}, \mathtt{En}', \mathtt{De}, \mathtt{Ev}', \mathtt{ev})$ such that $\mathcal{G}' \in \mathtt{GS}(aut, \Phi) \bigcap \mathtt{GS}(\mathtt{ev})$ but $\mathcal{G}' \notin \mathtt{GS}(prv.ind, \Phi) \bigcup \mathtt{GS}(obv.ind, \Phi)$. The construction is as follows: We define that $\mathtt{Ev}'(F, X||x)) = \mathtt{Ev}(F, X)$, $\mathtt{En}'(e, x) = \mathtt{En}(e, x)||x = X||x$.

The new encoding function $\mathtt{En}'$ and evaluation function $\mathtt{Ev}'$ do not harm $aut-$security, since the adversary has chosen the function f and its input x. On the other hand, appending x to the encoding harms both obliviousness and privacy: In both games the adversary chooses the function f in such a way that $ev(f, \cdot)$ is not injective. This is possible because it is assumed that $\mathtt{ev} = \mathtt{ev}_{circ}$.

In both PrvInd and ObvInd game the adversary chooses inputs x_0, x_1 such that $x_0 \neq x_1$ and $\mathtt{ev}(f, x_0) = \mathtt{ev}(f, x_1)$. Now, the encoding $X||x_b$ reveals which of the inputs was used. $\square$

The following two results provide parallel results compared to Propositions 8 and 9 in [3].

Theorem 5: Let P be a one-way permutation in the set of all functions f. Then, for $\Phi_P(f) = P(f)$ and for any $\mathtt{ev}$, $\mathtt{GS}(obv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev}) \nsubseteq \mathtt{GS}(obv.sim, \Phi_P)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(obv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev})$. We construct a new garbling scheme $\mathcal{G}' = (\mathtt{Gb}', \mathtt{En}, \mathtt{De}, \mathtt{Ev}', \mathtt{ev})$ such that $\mathcal{G}' \in \mathtt{GS}(obv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev})$ but $\mathcal{G}' \notin \mathtt{GS}(obv.sim, \Phi_P)$.

The construction is the following. The algorithm $\mathtt{Gb}'(1^k, f)$ picks $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$ and returns $(F||f, e, d)$. Let $\mathtt{Ev}'(F||f, X)$ return $\mathtt{Ev}(F, X)$. First of all, we claim that the constructed garbling scheme is $obv.ind$ secure over Φ_P. The reasoning goes as follows. The adversary $\mathcal{A}$ sends (f_0, f_1, x_0, x_1) to its GARBLE. For the answer not being $\perp$ it must be that $\Phi_P(f_0) = \Phi_P(f_1)$, and hence $P(f_0) = P(f_1)$ by the definition of Φ_P. Since P is a one-way permutation, $f_0 = f_1$ must hold. Thus prepending f to the description of F does not harm $obv.ind$ security.

However, $\mathcal{G}'$ is not $obv.sim$ secure over Φ_P. We introduce an adversary $\mathcal{B}$ that breaks the $obv.sim$ security with respect to any PT simulator. The adversary chooses (f, x) to be sent to the GARBLE procedure in ObvSim game. Now, if the challenge bit b in the game is 0, the simulator $\mathcal{S}$ is called to produce $(F||f, X)$ from $(1^k, \Phi_P(f))$. However, the PT simulator manages to produce exactly the right function f with negligible probability, because $\Phi_P = P$ is a one-way permutation. In other words, this means that the adversary $\mathcal{B}$ will almost always detect from the parameter $F||f$ whether the simulator was used or not. $\square$

Theorem 6: Let P be a one-way permutation in the set of all functions f and let $\Phi_P(f) = P(f)$ while $\mathtt{ev}$ is arbitrary. Assume that there exist x and y for which $\Phi_P(f) = P(f)$ is one-way even when restricted to functions f such that $y = \mathtt{ev}(f, x)$. Then $\mathtt{GS}(prv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev}) \nsubseteq \mathtt{GS}(prv.sim, \Phi_P)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(prv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev})$. We construct a new garbling scheme $\mathcal{G}' = (\mathtt{Gb}', \mathtt{En}, \mathtt{De}, \mathtt{Ev}', \mathtt{ev})$ such that $\mathcal{G}' \in \mathtt{GS}(prv.ind, \Phi_P) \bigcap \mathtt{GS}(\mathtt{ev})$ but $\mathcal{G}' \notin \mathtt{GS}(prv.sim, \Phi_P)$.

The construction is similar to that of the previous proof. The algorithm $\mathtt{Gb}'(1^k, f)$ picks $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f)$ and returns $(F||f, e, d)$. Let $\mathtt{Ev}'(F||f, X)$ return $\mathtt{Ev}(F, X)$. First of all, the constructed garbling scheme is $prv.ind$ secure over Φ_P by exactly the same reasoning as in the previous proof.

However, $\mathcal{G}'$ is not $prv.sim$ secure over Φ_P. We prove this by introducing an adversary $\mathcal{B}$ having a non-negligible advantage in the $PrvSim_{\Phi_P}$ game. By the assumption, there exist x and y such that $\Phi_P(f)$ is still one-way, when restricted to f such that $y = \mathtt{ev}(f, x)$. Thus the adversary $\mathcal{B}$ can choose (f, x) satisfying $y = \mathtt{ev}(f, x)$ to be sent to the GARBLE procedure. Now, if the challenge bit b in the game is 0, the simulator $\mathcal{S}$ is called to produce $(F||f, X, d)$ from $(1^k, y, \Phi_P(f))$, where $y = \mathtt{ev}(f, x)$. However, the polynomial time simulator manages to produce exactly the right function f with negligible probability, because $\Phi_P = P$ is an injective one-way function. In other words, this means that the adversary $\mathcal{B}$ will almost always detect from $F||f$ whether the simulator was used or not. $\square$

The following two propositions provide new results for garbling scheme classes in [3].

Theorem 7: If the function $h : (f, x) \mapsto (\Phi(f), \mathtt{ev}(f, x))$ is injective, then $\mathtt{GS}(\mathtt{ev}) \subseteq \mathtt{GS}(prv.ind, \Phi)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev})$ be an arbitrary garbling scheme over side-information function Φ. Let $\mathcal{B}$ be an adversary playing the $PrvInd_\Phi$ game. The adversary sends (f_0, f_1, x_0, x_1) to the GARBLE procedure of this game. For the output not being $\perp$ it must be that

$$\Phi(f_0) = \Phi(f_1), \mathtt{ev}(f_0, x_0) = \mathtt{ev}(f_1, x_1).$$

But by injectivity of h this implies

$$\begin{aligned} h(f_0, x_0) &= (\Phi(f_0), \mathtt{ev}(f_0, x_0)) \\ &= (\Phi(f_1), \mathtt{ev}(f_1, x_1)) = h(f_1, x_1) \\ &\Rightarrow (f_0, x_0) = (f_1, x_1). \end{aligned}$$

This in turn is equivalent to $f_0 = f_1$ and $x_0 = x_1$, meaning that the advantage of the adversary $\mathcal{B}$ in this game will be equal to 0. This completes the proof. $\square$

Theorem 8: If the function $\mathtt{ev}$ is injective and efficiently invertible (i.e. given $y = \mathtt{ev}(f', x')$, f and x such that $\mathtt{ev}(f, x) = y$ can be found in polynomial time), then $\mathtt{GS}(\mathtt{ev}) \subseteq \mathtt{GS}(prv.sim, \Phi)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev})$ be an arbitrary garbling scheme over side-information function Φ. Let $\mathcal{B}$ be an adversary playing the $PrvSim_\Phi$ game. The adversary sends (f, x) to the GARBLE procedure of this game. But now, if the challenge bit $b = 0$, the simulator can always find the right f and x to be garbled because $y = \mathtt{ev}(f, x)$ can be inverted efficiently and $\mathtt{ev}(f,x) = \mathtt{ev}(f',x')$ guarantees $f = f', x = x'$. This means that no matter what the challenge bit was, in both cases, $b = 0$ or $b = 1$, the pair (f,x) becomes garbled correctly because the simulator that knows f and x is able to use the normal garbling method. This means that the advantage of the adversary $\mathcal{B}$ in this game equals 0, proving the inclusion $\mathtt{GS}(\mathtt{ev}) \subseteq \mathtt{GS}(prv.sim, \Phi) \bigcap \mathtt{GS}(\mathtt{ev})$. □

IV. NEW CLASSES OF GARBLING SCHEMES

In [3], the definitions and relations between different security types were, at least to some extent, based on intuition about what is meant by a garbling scheme that achieves privacy, obliviousness or authenticity, and the intuition was modeled as a game. In this section we consider the games defined in paper [3] from another point of view; we consider them purely as games, and try to achieve new results by modifying the existing game definitions in certain ways explained later.

The first modification we make is that in the indistinguishability model, the PrvInd game will be modified to the direction of ObvInd game by removing the decryption key d from the return value (F, X, d). The same end result can be obtained by tightening the ObvInd game by adding the evaluation test $\mathtt{ev}(f_0, x_0) \stackrel{?}{=} \mathtt{ev}(f_1, x_1)$ in it. In the absence of a better name we call the new class *ModInd*. The second modification concerns the PrvSim game, in which we again ease the requirements by removing the decryption key d from the return value (F, X, d). In ObvSim game, adding y to the input of the simulator $\mathcal{S}$ will lead to the same intermediate form as above. The new class shall be named *ModSim*.

Another modification is obtained by relaxing the PrvInd game by removing the evaluation test. This can also be achieved by adding d to the output (F, X) in ObvInd game. A similar modification in Sim side is to leave y out from the input of the simulator in PrvSim game, or add d to (F, X) in ObvSim game. The former modification is called *ModInd2* and the latter is called *ModSim2*.

The finalization procedure is not modified in any of these games.

A. Applications

Before proceeding to the descriptions of our modifications, it is convenient to discuss the possible applications that could utilize garbling schemes and more specifically, our modified security models. One typical example is outsourcing of a complex computation to a service in the cloud. In many cases the input data or the algorithm (or both of them) is privacy-sensitive data and should not be revealed to the party running the cloud service. With garbling schemes achieving different types of security, we can hide different amount of this information. In order to have an idea which type of security is most appropriate in different situations, let us take a closer look at which kind of information is revealed by a garbling scheme belonging to a specific security class.

Let the function f represent the algorithm, x represents the privacy-sensitive input data and $f(x) = y$ represents the output of the algorithm. These are all garbled with some garbling scheme, and the garbled function and garbled input are given to the server, which computes the garbled output. It depends on the garbling scheme how much the server is allowed to know about f, x and $f(x)$. It is worth noting that whatever the model of security is, the original function f is not known for the server, only the side-information function $\Phi(f)$ is. The following list provides the central differences between the models.

- **obv.sim**: Garbling does not reveal x, f or $f(x)$ to the server.
- **prv.sim**: The server is allowed to get $f(x)$ but not x or f.
- **mod.sim**: When computing $y_1 = f(x_1)$ and $y_2 = f(x_2)$, the server is allowed to find out whether $y_1 = y_2$ or not.

There are situations in which the output data is not sensitive and can be revealed to the party maintaining the cloud service. According to the list above, a prv.sim secure garbling scheme is then appropriate. Also garbling schemes of the two other types may fit the situation except if the server needs the output in further computations. The issue is that the output will remain garbled in the cloud. Of course, further computations could also be garbled but this arrangement would significantly and unnecessarily add the total complexity of computation.

If the output is sensitive data, an obv.sim secure scheme suits. Our modified model mod.sim is suitable as well except in some cases where the number and/or distribution of different output values may reveal too much information. On the other hand, mod.sim can actually be modified to apply to these cases as well. Instead of considering inputs x, the modified scheme would take inputs $x||i$ where i is for example an ever-increasing counter. The procedure then returns $\mathtt{ev}(f,x)||i$ as the output. The counter at the end of the evaluation result will now make sure that each output appears only once. According to the previous discussion, mod.sim secure garbling schemes can be used in the same applications as prv.sim or obv.sim secure schemes. In the following section we will prove that it is at least as easy to find a mod.sim secure scheme as it is to find an obv.sim or a prv.sim secure scheme. In conclusion, the modified security model mod.sim covers almost all applications except some esoteric cases.

B. Definitions and results

Next we give the formal definition of ModInd and ModSim games. Then we continue by proving some results concerning the new classes of garbling schemes that are secure with respect to these games.

The following proposition shows that mod.ind security is at least as easy to reach as prv.ind security or obv.ind security.

Theorem 9: $\mathtt{GS}(prv.ind, \Phi) \bigcup \mathtt{GS}(obv.ind, \Phi) \subseteq \mathtt{GS}(mod.ind, \Phi)$.

Proof: First suppose that $\mathcal{G}$ is a prv.ind secure garbling scheme. Dropping the decryption key d out of the output of GARBLE procedure does not increase the winning chances of any adversary.

proc $\texttt{GARBLE}(f_0, f_1, x_0, x_1)$ $\text{ModInd}_{\mathcal{G},\Phi}$	proc $\texttt{GARBLE}(f, x)$ $\text{ModSim}_{\mathcal{G},\Phi,\mathcal{S}}$
$b \twoheadleftarrow \{0,1\}$	$b \twoheadleftarrow \{0,1\}$
if $\Phi(f_0) \neq \Phi(f_1)$ **then return** $\perp$	**if** $x \notin \{0,1\}^{f.n}$ **then return** $\perp$
if $\{x_0, x_1\} \not\subset \{0,1\}^{f_0.n}$ **then return** $\perp$	**if** $b = 1$ **then** $(F, e, d) \leftarrow \texttt{Gb}(1^k, f); X \leftarrow \texttt{En}(e, x)$
if $\texttt{ev}(f_0, x_0) \neq \texttt{ev}(f_1, x_1)$ **then return** $\perp$	**else** $y \leftarrow \texttt{ev}(f, x);\ (F, X) \leftarrow \mathcal{S}(1^k, y, \Phi(f))$
$(F, e, d) \leftarrow \texttt{Gb}(1^k, f_b);\ X \leftarrow \texttt{En}(e, x_b)$	**return** (F, X)
return (F, X)	

Table II: The modified GARBLE procedures in Ind and Sim games

Secondly, suppose that $\mathcal{G}$ is an obv.ind secure scheme. Now, following the specification of ModInd GARBLE procedure, the adversary receives $\perp$ on all inputs whose evaluations $\texttt{ev}(f_0, x_0)$ and $\texttt{ev}(f_1, x_1)$ are not equal. However, this evaluation equality test is not a part of ObvInd game. Hence, even though GARBLE procedure in ObvInd game returns an output different from $\perp$, the corresponding procedure in ModInd game might return $\perp$. Otherwise the games are identical. Adversaries of both games are able to find out beforehand whether the GARBLE procedure returns $\perp$ and therefore the adversary in ModInd game does not receive . Therefore, the advantage of adversary playing the ModInd game cannot be better than the advantage of a corresponding adversary in ObvInd game. According to the assumption, the advantage in the ObvInd game is negligible, and thus the advantage in ModInd game must also be negligible. □

Theorem 10: $\texttt{GS}(prv.sim, \Phi) \bigcup \texttt{GS}(obv.sim, \Phi) \subseteq \texttt{GS}(mod.sim, \Phi)$.

Proof: First, suppose that garbling scheme $\mathcal{G}$ is prv.sim secure. As in the PrvInd case, omitting the decryption key d from (F, X, d) does not increase the winning probability of an adversary playing the modified game.

Secondly, suppose that the garbling scheme $\mathcal{G}$ belongs to the set of ObvSim secure schemes. In the ModSim game, the simulator's additional input y cannot make its work of producing a good output (F, X) more difficult. Let us explain in more details why this is the case.

Let $\mathcal{A}$ be an arbitrary adversary playing the ModSim game. Let $\mathcal{A}'$ be the corresponding adversary playing the ObvSim game: adversary $\mathcal{A}'$ behaves in ObvSim game exactly in the same way as $\mathcal{A}$ behaves in ModSim game. According to our assumption, there is a simulator $\mathcal{S}'$ such that the advantage of $\mathcal{A}'$ is negligible. Now, we construct a simulator $\mathcal{S}$ for the ModSim game. The simulator $\mathcal{S}$ will totally omit the additional input y and call simulator $\mathcal{S}'$ to produce an output to the adversary $\mathcal{A}$. Now, this simulator makes the advantage of adversary $\mathcal{A}$ negligible, because the adversary $\mathcal{A}$ behaves just like $\mathcal{A}'$ and also the simulators in both games behave identically. This completes the proof. □

As mentioned in the introductory part of this section, we have created four modifications to the prv.ind and prv.sim models in total, of which we have now covered two. In the rest of this section, we first give the descriptions of the two other modified games and provide some results concerning them. Finally, we give a diagram including the new models and their relations.

After these two definitions, we will now provide a result about mod.ind2 and mod.sim2.

Theorem 11: Assume that the following condition holds:

$$(\forall f_0, f_1)\ (\forall x_0, x_1): \qquad \textbf{(Condition } (**)\textbf{)}$$
$$\Phi(f_0) = \Phi(f_1) \Rightarrow \texttt{ev}(f_0, x_0) = \texttt{ev}(f_1, x_1).$$

Then $\texttt{GS}(mod.ind2, \Phi) = \texttt{GS}(prv.ind, \Phi)$. Otherwise $\texttt{GS}(mod.ind2, \Phi) = \emptyset$.

Proof: Suppose first that $(**)$ does not hold. Then the adversary can choose f_0, f_1, x_0, x_1 such that $\texttt{ev}(f_0, x_0) \neq \texttt{ev}(f_1, x_1)$ but still $\Phi(f_0) = \Phi(f_1)$ holds. In this case, the adversary will always win the game, because $b = 0$ if and only if $\texttt{ev}(f_0, x_0) = \texttt{De}(d, \texttt{Ev}(F, X))$, and thus the advantage would not satisfy the negligibility condition, and no garbling scheme is secure.

Now assume that $(**)$ holds. Then the the adversary in the ModInd2 game has no choice other than choosing f_0, f_1, x_0, x_1 such that $\Phi(f_0) = \Phi(f_1)$ for not receiving $\perp$ which now implies that $\texttt{ev}(f_0, x_0) = \texttt{ev}(f_1, x_1)$ must hold. It follows that the sets of PrvInd secure garbling schemes and ModInd2 secure garbling schemes must be equal. This completes the proof. □

Theorem 12: For any Φ, the inclusion $\texttt{GS}(mod.sim2, \Phi) \subseteq \texttt{GS}(prv.sim, \Phi)$ holds. If $(**)$ holds, and Φ is efficiently invertible (i.e. given $\phi = \Phi(f')$, a function f can be found in polynomial time such that $\Phi(f) = \phi$), then the equality $\texttt{GS}(mod.sim2, \Phi) = \texttt{GS}(prv.sim, \Phi)$ holds. Finally, if $(**)$ does not hold, then $\texttt{GS}(mod.sim2, \Phi) = \emptyset$.

Proof: The difference between the ModSim2 and PrvSim games is, that in PrvSim game the simulator gets $y = \texttt{ev}(f, x)$ as input, whereas the simulator in ModSim2 game does not. This means that simulator's task of creating a good output (F, X, d) in PrvSim game is not harder than the task of the simulator in the other game. Therefore, the advantage of an adversary in PrvSim game cannot be better than in ModSim2 game. This proves the first claim.

For the second part, suppose that $(**)$ holds and Φ is efficiently invertible. Even though y is not provided to the simulator, it still is able to produce (F', X', d') such that the adversary has no better chances than guessing to win the ModSim2 game. Namely, the simulator creates from $\Phi(f)$ such a function f' that $\Phi(f) = \Phi(f')$, and it then creates any suitable input x' to the function f'. Now, because of the condition $(**)$, the equality $\texttt{ev}(f, x) = \texttt{ev}(f', x')$ must hold and hence the simulator always learns the right y. This means that the setting in this new, modified game actually is exactly the same as in PrvSim game.

Finally suppose that $(**)$ does not hold. In the modified game, the adversary can choose f and x such that there

proc GARBLE(f_0, f_1, x_0, x_1) ModInd2$_{\mathcal{G},\Phi}$	proc GARBLE(f, x) ModSim2$_{\mathcal{G},\Phi,\mathcal{S}}$
$b \twoheadleftarrow \{0,1\}$	$b \twoheadleftarrow \{0,1\}$
if $\Phi(f_0) \neq \Phi(f_1)$ **then return** $\perp$	**if** $x \notin \{0,1\}^{f.n}$ **then return** $\perp$
if $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ **then return** $\perp$	**if** $b = 1$ **then** $(F, e, d) \leftarrow \mathtt{Gb}(1^k, f); X \leftarrow \mathtt{En}(e, x)$
$(F, e, d) \leftarrow \mathtt{Gb}(1^k, f_b)$; $X \leftarrow \mathtt{En}(e, x_b)$	**else**$(F, X, d) \leftarrow \mathcal{S}(1^k, \Phi(f))$
return (F, X, d)	**return** (F, X, d)

Table III: Another modification of GARBLE procedure in Ind and Sim games

exists a function f' satisfying $f' \neq f$, $\Phi(f) = \Phi(f')$ and $\mathtt{ev}(f, x) \neq \mathtt{ev}(f', x')$ for some x'. Now the simulator has at most 50% chance to guess the correct f. If the guess was incorrect, distinguishing the simulated version from the actual garbled output is easy since the adversary is able to check if $\mathtt{ev}(f, x) = \mathtt{De}(d, \mathtt{Ev}(F, X))$. Thus, no garbling scheme is ModSim2 secure. □

Corollary 1: The following inclusion holds: $\mathtt{GS}(mod.sim2, \Phi) \subseteq \mathtt{GS}(mod.ind2, \Phi)$.

Proof: The claim follows from Theorem 11 and Theorem 12 and Proposition 2 in [3]. □

NOTE: In practice, condition $(**)$ does not usually hold. Therefore, it is hard to imagine an application in which our second modification would have practical significance because of the above result.

The next theorem provides a relation between the modified simulation type and the modified indistinguishability type garbling schemes under our first modification.

Theorem 13: The following inclusion holds: $\mathtt{GS}(mod.sim, \Phi) \subseteq \mathtt{GS}(mod.ind, \Phi)$.

Proof: Let $\mathcal{G} = (\mathtt{Gb}, \mathtt{En}, \mathtt{De}, \mathtt{Ev}, \mathtt{ev}) \in \mathtt{GS}(mod.sim, \Phi)$. We need to prove that $\mathcal{G} \in \mathtt{GS}(mod.ind, \Phi)$. Let $\mathcal{A}$ be the PT adversary playing the ModInd game. We construct a PT ModSim adversary $\mathcal{B}$ as follows. Let $\mathcal{B}$ run $\mathcal{A}$ as a subroutine. The latter makes its query f_0, f_1, x_0, x_1. Adversary $\mathcal{B}$ returns $\perp$ to $\mathcal{A}$ if $\Phi(f_0) \neq \Phi(f_1)$ or $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ or $\mathtt{ev}(f_0, x_0) \neq \mathtt{ev}(f_1, x_1)$.

Regardless of whether $\mathcal{B}$ returned $\perp$ to $\mathcal{A}$ or not, adversary $\mathcal{B}$ picks $c \in \{0,1\}$ at random and makes its query to GARBLE with input f_c, x_c getting back (F, X) which is sent to adversary $\mathcal{A}$ in case $\perp$ was not sent earlier. In any case, adversary $\mathcal{A}$ returns a bit b' to adversary $\mathcal{B}$. The latter adversary now returns 1 if $\Phi(f_0) = \Phi(f_1)$, $\{x_0, x_1\} \subseteq \{0,1\}^{f_0.n}$, $\mathtt{ev}(f_0, x_0) = \mathtt{ev}(f_1, x_1)$ and $b' = c$ and 0 otherwise. Let $\mathcal{S}$ be any PT algorithm representing the simulator. Then there are two possible outcomes of the game:

1) If $\Phi(f_0) = \Phi(f_1)$, $\{x_0, x_1\} \subseteq \{0,1\}^{f_0.n}$ and $\mathtt{ev}(f_0, x_0) = \mathtt{ev}(f_1, x_1)$, then the input to the simulator $\mathcal{S}$ is the same regardless of c, or
2) $\Phi(f_0) \neq \Phi(f_1)$, $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ or $\mathtt{ev}(f_0, x_0) \neq \mathtt{ev}(f_1, x_1)$ then adversary $\mathcal{B}$ always answers 0 regardless of b' received from adversary $\mathcal{A}$.

Let's analyze the win probabilities of both adversaries. First consider the case 2. Adversary $\mathcal{B}$ always answers 0, and there is 50% chance of it being the right answer, and hence the win probability of $\mathcal{B}$ is one half. The win probability of adversary $\mathcal{A}$ is the same: $\mathcal{A}$ does not get any information linked to the challenge bit, and thus its answer is as good as guessing but there is always 50% chance of answering right.

Next consider case 1. Now, there are two possibilities for challenge bit b. Suppose first that $b = 1$. In this case, adversary $\mathcal{B}$ wins if and only if $\mathcal{A}$ wins. On the other hand, if the challenge bit b equals 0, adversary $\mathcal{A}$ does not have any information because it is getting the same input regardless of c, so its answer is no better than a guess. Thus the win probability equals $\frac{1}{2}$. Furthermore, the adversary $\mathcal{A}$ wins if and only if adversary $\mathcal{B}$ loses, therefore $\Pr[\mathcal{B}\ \ wins] = \Pr[\mathcal{A}\ \ loses] = \frac{1}{2}$.

This case analysis above shows that in all cases $\Pr[\mathcal{B}\ \ wins] = \Pr[\mathcal{A}\ \ wins]$. Now continuing with $\Pr[\mathcal{A}\ \ wins]$ we obtain

$$\begin{aligned}\Pr[\mathcal{A}\ \ wins] &= \frac{1}{2}\Pr[\mathcal{A}\ \ wins|b=1] + \frac{1}{2}\Pr[\mathcal{A}\ \ wins|b=0] \\ &= \frac{1}{2} \cdot \left(\frac{1}{2} + \frac{1}{2} \cdot \mathbf{Adv}_{\mathcal{A}}\right) + \frac{1}{2} \cdot \frac{1}{2} \\ &= \frac{1}{2} + \frac{1}{4} \cdot \mathbf{Adv}_{\mathcal{A}}.\end{aligned}$$

By the definition of advantage of adversary $\mathcal{B}$ we have $\Pr[\mathcal{B}\ \ wins] = \frac{1}{2} \cdot \mathbf{Adv}_{\mathcal{B}} + \frac{1}{2}$ and therefore we obtain $\mathbf{Adv}_{\mathcal{A}} = 2 \cdot \mathbf{Adv}_{\mathcal{B}}$. Now, since the $\mathbf{Adv}_{\mathcal{A}}$ is negligible according to the assumption, $\mathbf{Adv}_{\mathcal{B}}$ is also negligible. □

For our last theorem, we introduce a new condition:

The decryption key d can be efficiently computed from the tuple (F, X). (**Condition** $(***)$)

Theorem 14: The following inclusions hold: $\mathtt{GS}(mod.ind2, \Phi) \subseteq \mathtt{GS}(obv.ind, \Phi)$ and $\mathtt{GS}(mod.sim2, \Phi) \subseteq \mathtt{GS}(obv.sim, \Phi)$. If condition $(***)$ holds, then the classes are equal.

Proof: The difference between $ModInd2$ and $ObvInd$ (respectively $ModSim2$ and $ObvSim$) is that in ModInd2 game (in ModSim2 respectively) the adversary receives the decryption key d as an output from GARBLE together with F and X. This auxiliary output does not make the advantage smaller to the adversary in the modified games. The claim follows from this observation. □

These results complete the considerations about the possible relations between the new and old classes of garbling schemes. Results are collected into Figure 3.

If in addition to $(***)$ we require that $(\Phi, \mathtt{ev})$ is efficiently invertible (i.e. given $y = \mathtt{ev}(f', x')$ and $\phi = \Phi(f')$, f and x such that $y = \mathtt{ev}(f, x)$ and $\Phi(f) = \phi$ can be found in polynomial time) and condition $(**)$ also holds, then all the

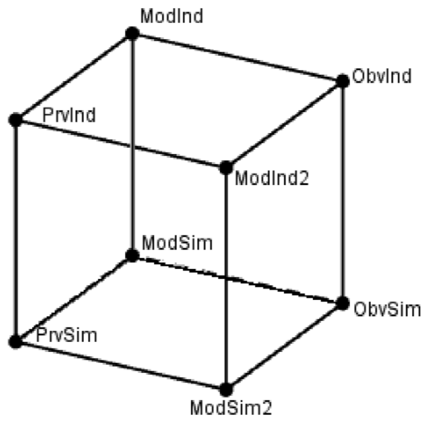


Figure 3: Inclusions between classes of garbling schemes

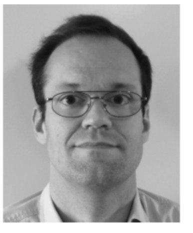
sets in the diagram collapse into one point: a garbling scheme that belongs to one security class will be secure also with respect to any other security model.

V. CONCLUSIONS

In this paper, we have considered different security classes of garbling schemes. Some of our results are obtained for the classes defined by Bellare, Hoang and Rogaway in [3]. We have also introduced new security classes and described their relation to the earlier classes. From these new classes, we see that the new classes $\texttt{GS}(mod.ind, \Phi)$ and $\texttt{GS}(mod.sim, \Phi)$ would be promising targets for future research - at least, it seems that these classes would have practical applications. Namely, our results show that all garbling schemes in the old obv-classes belong also to the new mod-classes, and therefore it is at least as easy to find a garbling scheme that is mod-secure. Moreover, it seems to be harder to find an application which would require obv-security but where mod-security would not suffice. The second new class sets too hard requirements for a secure garbling scheme and this class is practically always empty.

REFERENCES

[1] Y. Aumann and Y. Lindell. Security against covert adversaries: Efficient protocols for realistic adversaries. *Proc. of TCC 2007*, 4392 of LNCS:137–156, 2007.
[2] D. Beaver, S. Micali, and P. Rogaway. The round complexity of secure protocols. *Proc. of the* 22^{nd} *STOC*, pages 503–513, 1990.
[3] M. Bellare, V. T. Hoang, and P. Rogaway. Foundations of garbled circuits. *Proc. of ACM Computer and Communications Security (CCS'12)*, 2012.
[4] M. Bellare and P. Rogaway. Code-based game-playing proofs and the security of triple encryption. *Advances in Cryptology, Proc. of Eurocrypt 2006*, 4004 of LNCS:409–426, 2006.
[5] Y. Lindell and B. Pinkas. A proof of security of Yao's protocol for secure two-party computation. *Electronic Colloquium on Computational Complexity*, TR04-063, 2004.
[6] Y. Lindell and B. Pinkas. A proof of security of Yao's protocol for secure two-party computation. *Journal of Cryptology*, 22(2):161–188, 2009.
[7] A. Yao. How to generate and exchange secrets. *Proc. of* 27^{th} *FOCS, 1986.*, pages 162–167, 1986.



Tommi Meskanen had his PhD in 2005. Since that he has been working as a researcher and lecturer at University of Turku. His main research interests are cryptography and public choice theory. His email address is tommes@utu.fi

Valtteri Niemi is a Professor of Mathematics at the University of Turku, Finland. Between 1997 and 2012 he was with Nokia Research Center in various positions, based in Finland and Switzerland. Niemi was also the chairman of the security standardization group of 3GPP during 2003-2009. His research interests include cryptography and mobile security. Valtteri can be contacted at valtteri.niemi@utu.fi.

Noora Nieminen is a doctoral student at Turku Centre for Computer Science, Department of Mathematics and Statistics at the University of Turku. Her research interests include cryptography and its applications. Contact her at nmniem@utu.fi.

On a key exchange protocol based on Diophantine equations

Noriko Hirata-Kohno, Attila Pethő

***Abstract*—We analyze a recent key exchange protocol proposed by H. Yosh, which is based on the hardness to solve Diophantine equations. In this article, we analyze the protocol and show that the public key is very large. We suggest large families of parameters both in the finite field and in the rational integer cases for which the protocol can be secure.**

I. Introduction

The notion of public key cryptography started with a key exchange protocol [12]. Various protocols have been developed for this purpose, see for example [8], [14]. Hard computational problems lie under these protocols, e.g., factorization into primes of large integers, computation of discrete logarithm, determination of the shortest vector in lattices and decoding of error correcting codes.

D. Hilbert asked in his famous lecture at the second International Congress of Mathematicians in 1900 whether there exists a general procedure which determines the solvability of Diophantine equations. The question was answered 70 years later by Y. Matijasevič, who proved that such an algorithm does not exist [11]. However, the impossibility of a general algorithm does not mean that we cannot solve special equations. There are large classes of Diophantine equations which are algorithmically and numerically solvable, see e.g. [1], [20].

Despite many efforts, finding the solutions to Diophantine equations is usually a hard task. Based on this observation, Lin, Chang and Lee [13] suggested a new public key protocol in 1995. A bit later Cusick showed that this protocol is insecure and it can be broken in polynomial time without solving any Diophantine equations [9]. Although such observations, especially in the case of (non-linear) Diophantine equations of high degree, Yosh [22] proposed a key exchange protocol whose security relies on the hardness to find the solutions to the equations.

N. Hirata-Kohno is a professor at the Department of Mathematics, College of Science and Technology, Nihon University, Suruga-dai, Kanda, Chiyoda, Tokyo 101-8308, JAPAN (email:hirata@math.cst.nihon-u.ac.jp).

A. Pethő is a professor at the Department of Computer Science, University of Debrecen, H-4010 Debrecen, P.O. Box 12, HUNGARY (email:Petho.Attila@inf.unideb.hu).

The first author was partially supported by the NEXT Program, GR 087, JSPS for 2010–2013. The second author was partially supported by the JSPS, Grant in aid 21540010 and Invitation Fellowship Program FY2011, L-11514, by the OTKA grant No. 104208 and by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project.

We present here a more detailed analysis of the protocol. We show that it can be secure both over finite fields and in the original setting, i.e. over the ring of rational integers. In any case there is a big efficiency bottleneck and indeed the size of the public key is enormous.

It might be true that the theory of cryptography does not profit enough from the theory of Diophantine equation of high degree and vice versa. This is the reason to write these notes.

After the celebrated theorem of Shor [19] that factorization and discrete logarithm can be done with quantum algorithms in polynomial time, there is a big demand to develop new public key protocols. These should be based on problems, which cannot be solved by quantum computers in polynomial time, or at least we should have some evidence. A good overview on such efforts is presented in [3]. We hope that these notes might give a small step toward this direction.

II. The protocol of Harry Yosh

In this section, we describe with minor modifications and generalizations, the key exchange protocol proposed by H. Yosh [22]. Let R be a commutative ring with unity 1. Fix $a \in R$ and $b \in \mathbb{N}$ and for $x \in R$, consider the function

$$T_{a,b}(x) = (x+a)^b.$$

Obviously $T_{a,b}$ is a polynomial map from R to R. Assume that b is chosen such that $T_{a,b}$ is injective, i.e. invertible. Let $f(x_1,\dots,x_m)$, $g(x_1,\dots,x_m) \in R[x_1,\dots,x_m]$.

To exchange a secret key, Alice and Bob perform the following steps:

(i) Alice chooses a polynomial $f(x_1,\dots,x_m) \in R[x_1,\dots,x_m]$ and compute a solution $(r_1,\dots,r_m) \in R^m$ to the Diophantine equation

$$f(x_1,\dots,x_m) = 0.$$

She keeps $(r_1,\dots,r_m)$ secret, but makes f public.

(ii) Bob chooses a polynomial $g(x_1,\dots,x_m) \in R[x_1,\dots,x_m]$ and parameters $a_1,\dots,a_n \in R$ as well as $b_1,\dots,b_n \in \mathbb{N}$ such that T_{a_j,b_j} are invertible for $j = 1,\dots,n$. He computes

$$H(x_1,\dots,x_m) =$$

$$= T_{a_n,b_n}(\dots(T_{a_1,b_1}(g(x_1,\dots,x_m)))\dots)$$

and takes an element $h \in H + fR[x_1,\dots,x_n]$.
He keeps $a_1,\dots,a_n,b_1,\dots,b_n$ secret and makes g,h public.

(iii) Knowing g,h Alice computes $s = g(r_1,\dots,r_m)$ and $u = h(r_1,\dots,r_m)$ and sends u to Bob.

(iv) Bob computes $T^{-1}_{a_1,b_1}(\dots(T^{-1}_{a_n,b_n}(u))\dots)$, which is s, the common secret key of Alice and Bob.

For completeness we prove

Proposition 1. *The protocol is correct.*

Proof: Alice can compute s because she knows g and $r_1,\dots,r_m$.
As $f(r_1,\dots,r_m) = 0$ we have

$$u = h(r_1,\dots,r_m) = H(r_1,\dots,r_m).$$

Thus

$$s = H^{-1}(u) = T^{-1}_{a_1,b_1}(\dots(T^{-1}_{a_n,b_n}(u))\dots)$$

and Bob can compute s because he knows $a_1,\dots,a_n,b_1,\dots,b_n$ and $T_{a_j,b_j}, j = 1,\dots,n$ are invertible. ■

In Yosh' analysis, it was only considered one possible attack. The secret key can be computed from common solutions to the system of public equations $f = 0, h = u$. Yosh pointed out that one can choose these equations such that the determination via Gröbner bases technique of the common solution still remains a hard task. Unfortunately only few examples were given in the article.

Here, we present a more detailed cryptoanalysis of the protocol of Yosh. In Yosh's original version, only the case $R = \mathbb{Z}$ was investigated and the finite field case was just mentioned. We investigate two cases, when $R = \mathbb{Z}$ and R is a finite field.

Another difference is that Yosh dealt with the map in three parameters $\hat{T}_{a,b,c}(x) = (x+a)^b + c$, with $a,c \in R$ and $b \in \mathbb{N}$. By the obvious identity

$$\hat{T}_{\hat{a}_n,\hat{b}_n,\hat{c}_n}(\dots(\hat{T}_{\hat{a}_1,\hat{b}_1,\hat{c}_1}(x))\dots) =$$
$$= T_{a_{n+1},b_{n+1}}(T_{a_n,b_n}(\dots(T_{a_1,b_1}(x))\dots)),$$

where $a_1 = \hat{a}_1, a_j = \hat{a}_j + \hat{c}_{j-1}, j = 2\dots,n, a_{n+1} = \hat{c}_n$, $b_j = \hat{b}_j, j = 1,\dots,n$ and $b_{n+1} = 1$ it is enough to work with our map in two parameters.

We point out that the most serious bottleneck is the size of the public key, especially the size of h. To keep this parameter in an acceptable size, we have to use low degree polynomials, in particular $b_1,\dots,b_n$ have to be small.

Another important observation is that the equation $f = 0$ has to be hard to solve. We show in both cases that this can be achieved with large families of polynomials. In the case of $\mathbb{Z}$ we present a concrete example for which the protocol seems to be secure and the public key can be computed within some seconds.

A nice feature of the above algorithm is that the parties are coequal during the key generation, both have own secret, which are not known even by the partner. In this respect it is similar to the celebrated Diffie-Hellmann key exchange protocol [12].

III. PRELIMINARY OBSERVATIONS

Remark that in [22] there is no hints for the secure choice of the parameters, only an example and remarks about possible attacks are given. In these notes we concentrate on the possibility of such a choice of the parameters, which is computationally feasible, but seems secure enough. In this part we collected observations, which are independent from the ground ring R.

To break the system, i.e. to compute the common key, the enemy has to find the secret parameters $r_1,\dots,r_m$ or $a_1,\dots,a_n,b_1,\dots,b_n$. The only public information about the former is that $(r_1,\dots,r_m)$ is a solution to the system of equations

$$f(x_1,\dots,x_m) = 0 \quad (1)$$
$$h(x_1,\dots,x_m) = u. \quad (2)$$

To solve such equations one can use Gröbner bases technique [5], [6], [8] or elimination theory. The latter means that choosing one of the unknowns, say x_m, one computes the resultant $Res_{x_m}(f, h-u)$, which has unknowns one less than those of f or h. Moreover the first $m-1$ coordinates of solutions to (1) and (2) are zeroes of the resultant. Thus m has to be at least three because otherwise after the elimination one of the variables in (1) and (2), we would obtain an equation in a univariate polynomial, which is simple to solve.

Key exchange protocols are used several times with the same parameters. In our case f and $(r_1,\dots,r_m)$ can be fixed. After each running the enemy learn a new h and the corresponding u. After ℓ turns he collects $\ell+1$ public equations for $(r_1,\dots,r_m)$. If $\ell \geq m-2$ then the enemy can easily compute $(r_1,\dots,r_m)$.

Proposition 2. *The protocol can be used with the same polynomial f only at most $m-3$-times.*

A further observation of similar manner is the following.

Proposition 3. *If the adversary can compute many solutions, not necessarily $(r_1,\dots,r_m)$, of (1), then he can compute the element s and break the protocol.*

Proof: Indeed, assume that $(\alpha_1,\dots,\alpha_m) \in R^m$ is a solution to (1) and put $\beta = g(\alpha_1,\dots,\alpha_m)$. As

$$h = H + fV$$

for some $V \in R[x_1,\dots,x_m]$, we have $h(\alpha_1,\dots,\alpha_m) = H(\alpha_1,\dots,\alpha_m)$. Thus we get the equation

$$(((\beta+a_1)^{b_1}+a_2)^{b_2}+\ldots+a_n)^{b_n} = h(\alpha_1,\dots,\alpha_m). \tag{3}$$

for $a_1,\dots,a_n,b_1,\dots,b_n$. Knowing about $2n$ solutions of (1) we obtain about $2n$ equations of form (3), which determine usually the $2n$ unknowns. ∎

Now we investigate the possible choice of $a_1,\dots,a_n,b_1,\dots,b_n$. Let

$$\begin{aligned} t(x) &= t_{a_1,\dots,a_n,b_1,\dots,b_n}(x) = \\ &= T_{a_n,b_n}(\ldots(T_{a_1,b_1}(x))\ldots) = \\ &= (((x+a_1)^{b_1}+a_2)^{b_2}+\ldots+a_n)^{b_n}. \end{aligned}$$

It is clear that the degree of $t(x)$ is $b_1\cdots b_n$. On the other hand its value at each point can be computed by n additions and by at most $O(\log b_1+\ldots+\log b_n)$ multiplications. Furthermore, it can be stored on at most $n(A+B)$ bits, where A and B denote the maximal bit length of the representations of a_i and $b_i, i=1,\dots,n$ respectively. This means that t admits a very sparse representation. Since polynomials in sparse representations are rare, we cannot expect that h has a similar simple representation. We have to expect that the representation of h is dense, i.e. most of its coefficients are non-zero.

Put $d_i = \deg_{x_i} g, i=1,\dots,m$. Then it is clear that

$$\deg_{x_i} H = b_1\cdots b_n\cdot d_i$$

holds for $i=1,\dots,m$. Thus H has at most $(1+o(1))d_1\cdots d_m(b_1\cdots b_n)^m$ terms. We obtain h in Step (ii) by adding a suitable multiple of f to H. Hence we can control the degree of one of the variables. We may assume that it is x_m. By the argument above, we expect that a big portion of the coefficients of the terms of h is non-zero, i.e. we have to store about

$$O(d_1\cdots d_{m-1}(b_1\cdots b_n)^{m-1}) \tag{4}$$

non-zero elements of R. This means that $n,m,b_1,\dots,b_n$ have to be small. To be more specific $b_1,\dots,b_n\le\mathcal{B}$ and $n,m\le N$, where $\mathcal{B},N$ are small positive integers.

IV. The protocol over finite fields

Yosh mentioned in [22] that the protocol works over finite fields too, but no detail is given. We analyze this case in the present section. Set $R=\mathbb{F}_q$, where q is a prime power. In practice q is either a large prime or a large power of 2. It is a classical fact that $x\mapsto x^b$ is bijective on $\mathbb{F}_q^* = \mathbb{F}_q\setminus\{0\}$ iff $\gcd(q-1,b)=1$. Combining this fact with the general remarks of Section III we must have $1\le b_i\le\mathcal{B}$ and $\gcd(q-1,b_i)=1, i=1,\dots,n$.

By Proposition 3 the equation $f(x_1,\dots,x_m)=0$ has to be hard to solve. The next theorem, which is the combination of Theorem 2.1. and Corollary 2.2. The argument by Bérczes, Folláth and Pethő in [4], enables us to define a large class of $f\in\mathbb{F}_q[x_1,\dots,x_m]$ such that if q is large then this holds with high probability.

Theorem 1. *Let*

$$\begin{aligned} F(x_1,\dots,x_m) &:= B(x_1,\dots,x_m)+A(x_1,\dots,x_m) \\ &\in \mathbb{F}_q[x_1,\dots,x_m] \end{aligned}$$

with homogeneous polynomials A,B satisfying $\deg A < \deg B = D$, $\deg_{x_i} B = D$ for each $1\le i\le m$. Further, suppose that there exist indices $1\le j_1<j_2\le n$ such that the binary form

$$B(0,\dots,0,x_{j_1},0,\dots,0,x_{j_2},0,\dots,0) \tag{5}$$

has no multiple zero.
Denote by $P_{coll}(F,\gamma)$ the probability that $F(\mathbf{x})$ assumes the value $\gamma\in\mathbb{F}_q^*$, when $\mathbf{x}$ runs uniformly through the elements of $\mathbb{F}_q^m$. If $q>5\cdot D^{13/3}$, then

$$P_{coll}(F,\gamma)\le\frac{3}{q}.$$

The following construction of f is based on the consequence of Theorem 1.

- Set $q=2^{127}$, which ensures that $\gcd(q-1,p)=1$ for $p=3,5,7$.
- Choose homogenous polynomials $A,B\in\mathbb{F}_q[x_1,\dots,x_m]$ subject to the condition (5) and such that $\deg A<\deg B\sim b_1\cdots b_n/3$.
- Pick randomly $r_1,\dots,r_m\in\mathbb{F}_q$ and set $\gamma = B(r_1,\dots,r_m)+A(r_1,\dots,r_m)$. If $\gamma=0$ then choose $r_1,\dots,r_m$ again, otherwise set $f=B+A-\gamma$.

Then $(r_1,\dots,r_m)$ is a solution of $f=0$. As $D\sim b_1\cdots b_n/3\sim 7$ the condition $q>5\cdot D^{13/3}$ holds too. By Theorem 1 the chance to find $(r_1,\dots,r_m)$ or a different solution of $f=0$ is extremely low.

Remark that in the first step q can be replaced by a larger power of 2 or by an odd prime of similar size. We have to be care to the condition $\gcd(q-1,p)=1$ for all primes $p\le\mathcal{B}$. In [4] it was proved that there exists a large class of polynomials, which satisfy the assumptions of step 2.

We suggest that Bob chooses $a_1,\dots,a_n\in\mathbb{F}_q^*$ randomly. This is appropriate because in Step (iii) of the algorithm Alice makes public the value $u=h(r_1,\dots,r_m)$. Thus the equation

$$(((s+a_1)^{b_1}+a_2)^{b_2}+\ldots+a_n)^{b_n}=u$$

is known for everybody, but the element s is not known. We may assume without loss of generality $b_n=1$ because one can compute small degree roots in finite fields or in $\mathbb{Z}$ in probabilistic polynomial time. Thus our equation has the form

$$x^b+y=c,$$

where c and b are known, but x, y are unknown elements of $\mathbb{F}_q$. Thus the adversary has no chance to find the hidden solution s.

To hide H we suggest to choose $V \in \mathbb{F}_q[x_1, \ldots, x_m]$ randomly of low degree, and put $h = H + fV$.

Proposition 4. *With the above choice the key exchange protocol of Yosh over finite fields is secure.*

V. The case $R = \mathbb{Z}$

The map $T_{a,b}$ is injective if and only if b is odd.

In Step (iii) of the algorithm, Alice make public the value $u = h(r_1, \ldots, r_m)$. Thus the equation

$$(((s + a_1)^{b_1} + a_2)^{b_2} + \ldots + a_n)^{b_n} = u. \quad (6)$$

is known for everybody, but s is not known. We pointed out in the finite field case that $b_n = 1$ can be assumed without loss of generality. Thus our equation has the form

$$x^b + y = c,$$

where c is a known integer, b may be assumed to have some small values and x, y are unknown integers. Let y_0 be the nearest integer to $c^{1/b}$ and compute the two sided sequence $(y_0 \pm k)^b, k = 0, 1, \ldots$ until c appears. If the equation has a small solution in y, say $|y| \leq 10^7$, then with the above procedure, it will be quickly found.

Proposition 5. *We may assume $b_n = 1$. The parameters $a_1, \ldots, a_n$ should be sufficiently large, say $|a_i| \geq 10^8, i = 1, \ldots, n$.*

Let $a = \max\{|a_1|, \ldots, |a_n|\}$. We have to expect that the absolute value most of the coefficients of $t(x)$ hence of H, h are as large as $a^{b_1 \cdots b_{n-1}}$, which is 10^{72} even for the smallest possible parameter values $n = 4, b_1 = b_2 = b_3 = 3$. By (4), we have to store and transmit $3^9 \cdot d_1 \ldots d_{m-1}$ integers. In the simplest case, namely choosing g to be linear, we have to transmit about 10^4 coefficients of size 10^{72}. This is a very large amount of data. Below we give a concrete example showing this fact.

Now we come back to the choice of f. By Proposition 3 f has to be such that the equation $f = 0$ is hard to solve. We suggest to choose f a diagonal polynomial, i.e. of form $c_1 x_1^{d_1} + \ldots + c_m x_m^{d_m} - c_{m+1}$ with $d_1, \ldots, d_m \geq 2$. First of all these polynomials are very simple. It is an important aspect to compute h and one solution of the equation $f = 0$.

On the other hand diagonal polynomials are complicated enough, i.e. by careful choice of $c_1, \ldots, c_{m+1}, d_1, \ldots, d_m$ the adversary can hardly find a solution of the diophantine equation $c_1 x_1^{d_1} + \ldots + c_m x_m^{d_m} - c_{m+1} = 0$. Indeed, it is well known that if at most one exponent is equal to two and we fix the values of $m-2$ variables, then the resulting single equation in two-variables has only finitely many solutions. Moreover it is usually hard to find a solution provided the coefficients are large. If two exponents are equal to 2 then we may get equations of form $x^2 - dy^2 = m$ with infinitely many integer solutions, but the computation of the fundamental solutions is hard. For example, it is well known that finding a solution of $x^2 - y^2 = n$ such that $x - y \neq \pm 1, \pm n$ is equivalent to finding a non-trivial factor of n, see e.g. [17].

Choose $d_1 \leq \ldots \leq d_m$ according to the last paragraph and such that they are small, say $d_i \leq 7, i = 1, \ldots, m$. Let v be a positive integer, which we specify later. After fixing $d_1, \ldots, d_m$ it is not wise to choose $c_1, \ldots, c_m$ and c_{m+1}, because the success probability for the solution of a given equation is the same for everybody. Alice has to carry out in a different manner. She chooses a solution and after this she searches for an equation with the prescribed solution. To be more specific, she chooses $r_1, \ldots, r_m, c_{m+1} \in \mathbb{Z}$ randomly subject to the conditions $|r_i|^{d_i} \leq 2^v, i = 1, \ldots, m, |c_{m+1}| \leq 2^v$ and such that $\gcd(r_1, \ldots, r_m) = 1$. The number of possibilities is about $2^{v\left(1 + \frac{1}{d_1} + \ldots + \frac{1}{d_m}\right)}$. Then she computes $c_1, \ldots, c_m$ by solving the linear Diophantine equation

$$c_{m+1} = c_1 r_1^{d_1} + \ldots + c_m r_m^{d_m}.$$

The assumptions are such that this equation is solvable and that it has infinitely many solutions. From this infinite collection we suggest to choose $c_1, \ldots, c_m$ such that they have similar size. Performing this process Alice has the polynomial f and knows a solution to (1). On the other hand, finding a solution for other peoples (or finding another solution for Alice) is hopeless.

It remains to specify v. It must be so large that a brute force attack is hopeless. This means that the number of choices of the parameters must be large, at least 2^{128}. This implies the inequality

$$v\left(1 + \frac{1}{d_1} + \ldots + \frac{1}{d_m}\right) \geq 128.$$

We suggest to choose g randomly among the quadratic or linear polynomials.

There is no canonical choice for $h \in H + f\mathbb{Z}[x_1, \ldots, x_m]$, provided $m > 1$. One can fix a variable, say x_m, and consider H, f as polynomials in x_m with coefficients in the ring $\mathbb{Z}[x_1, \ldots, x_{m-1}]$. Then one can compute the remainder of H modulo f. The choice of the variable considerably influences the size of h. We give an example below. Another possibility for the choice of h is that we pick a polynomial $V \in \mathbb{Z}[x_1, \ldots, x_m]$ randomly and put $h = H + fV$.

Finally we present a concrete example, which might satisfy the security requirements and the size of the public key is beyond the possibilities.[1] Set $m = 4, n = 3$ and choose the polynomials as follows.

$$
\begin{aligned}
f &= c_1x_1^2 + c_2x_2^5 + c_3x_3^3 + c_4x_4^7 + c_5;\\
c_1 &= 1004439616068996251566977588899652\\
&\ 58647,\\
c_2 &= -3498105123011851201811794864519944\\
&\ 7959092\\
c_3 &= 3637968625340525244277529707911599 9\\
&\ 387383647170627044441713963619543 64,\\
c_4 &= -7075412456027395462040210714939958\\
&\ 108817512020742239926498242401,\\
c_5 &= -9876543234567898765432165432056789\\
&\ 6543210567,\\
g &= 3x_1 + 5x_2^2 + 7x_1x_2 + 93x_3^3 + 753x_4,\\
H &= ((g + 734367)^3 + 537769)^5 + 56478587.
\end{aligned}
$$

A solution of $f = 0$ is

$$
\begin{aligned}
x_1 &= 235452462352353121512,\ x_2 = 43689743,\\
x_3 &= 43216789765432,\ x_4 = 4567973.
\end{aligned}
$$

We left to the readers to find a different solution. With these parameters the computation of h took some seconds. It has 2107 terms and the internal representation in MAPLE has length 800327.

Noriko Hirata-Kohno is a professor at the Department of Mathematics, College of Science and Technology, Nihon University, Suruga-dai, Kanda, Chiyoda, Tokyo 101-8308, JAPAN. Her research interests include Diophantine problems and applications to cryptography. She has a PhD from University of Paris 6. Her email address is hirata@math.cst.nihon-u.ac.jp.

Attila Pethő is a professor and the head of the Department of Computer Science, Faculty of Informatics, University of Debrecen, Hungary. He got PhD degree in mathematics from the Lajos Kossuth University. He is a corresponding member of the Hungarian Academy of Sciences. His research interest are number theory and cryptography. You can contact him: petho.attila@inf.unideb.hu.

[1]It is not at all practical.

REFERENCES

[1] A. BAKER, *Transcendental Number Theory*, Cambridge Univ. Press, Cambridge, (1975).

[2] TH. BECKER and V. WEISPFENNING in cooperation with H. KREDEL, *Gröbner Bases: a Computational Approach to Commutative Algebra,* Graduate Texts in Mathematics, Vol. 141, Springer Verlag, New York, (1993).

[3] D.J. BERNSTEIN, J. BUCHMANN and E. DAHMEN (Editors), *Post-Quantum Cryptography*, Springer Verlag, Berlin, Heidelberg, (2009).

[4] A. BÉRCZES, J. FOLLÁTH and A. PETHŐ, *On a family of collision-free functions*, Tatra Mountains Math. Publ. **47** (2010), 1–13.

[5] B. BUCHBERGER, *Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems,* Aequ. Math. **4** (1970), 374–383.

[6] B. BUCHBERGER, G. E. COLLINS and R. LOOS eds., in cooperation with R. ALBRECHT, *Computer Algebra Symbolic and Algebraic Computation*, Springer, (1982).

[7] J. BUCHMANN, *Introduction to cryptography*, Second edition. Undergraduate Texts in Mathematics. Springer, (2004).

[8] J. BUCHMANN and H.C. WILLIAMS, *A key-exchange system based on imaginary quadratic fields*. J. Cryptology **1**, (1988), 107–118.

[9] T.W. CUSICK, *Cryptoanalysis of a public key system based on diophantine equations*, Inform. Processing Letters, **56** (1995), 73–75.

[10] J. H. DAVENPORT, Y. SIRET and E. TOURNIER, *Computer Algebra Systems and Algorithms for Algebraic Computation,* Academic Press, (1988).

[11] M. DAVIS, Y. MATIJASEVIC and J. ROBINSON, *Hilbert's tenth problem, Diophantine equations: positive aspects of a negative solution,* in: *Mathematical Developments Arising from Hilbert Problems,* Ed.: F.E. Browder, Symp. in Pure Math., (1974), AMS, Providence, RI., (1976), pp. 323–378.

[12] W. DIFFIE and M. HELLMAN, *New direction in cryptography*, IEEE Trans. onInformation Theory, 22 (1976), 644–654.

[13] C. H. LIN, C. C. CHANG and R. C. T. LEE, *A New Public-Key Cipher System Based Upon the Diophantine Equations*, IEEE Transactions on Computers, Volume 44, Issue 1, (1995).

[14] A.J. MENEZES, P.C. VAN OORSCHOT and S.A. VANSTONE, *Handbook of applied cryptography*, CRC, (1996).

[15] M. MIGNOTTE, *Mathematics for Computer Algebra,* Springer Verlag, Berlin, 1992.

[16] R. RIVEST, A. SHAMIR and L. ADLEMAN, *A method for obtaining digital signature and public-key cryptosystems*, Communications of the ACM, 21 (1978), 120–126.

[17] H. RIESEL, *Prime numbers and computer methods for factorization*, Second edition. Progress in Math., 126. Birkhäuser, MA, (1994), xvi+464 pp.

[18] B. SCHNEIER, *Applied cryptography: protocols, algorithms and source code in C*, (1996).

[19] P. SHOR, *Algorithms for Quantum Computation: Discrete Logarithm and Factoring,* Proc. 35th Annual Symposium on Foundations of Computer Science (1994) 124–134 and SIAM J. Comput. 26 (1997), 1484–1509.

[20] N.P. SMART, *The algorithmic resolution of Diophantine equations*, London Mathematical Society Student Texts, 41. Cambridge University Press, Cambridge, 1998.

[21] F. WINKLER, *Polynomial Algorithms in Computer Algebra,* Texts and Monographs in Symbolic Computation, Springer, (1996).

[22] H. YOSII, *The key exchange cryptosystem used with higher order Diophantine equations*, International Journal of Network Security & Its Applications **3** (2011), 43–50.

Strongly Secure Password Based Blind Signature for Real World Applications

Sangeetha Jose, Preetha Mathew K. , C. Pandu Rangan

Abstract—**Digital signature is the cryptographic primitive that ensures authentication and nonrepudiation. A password based blind signature can be used in the scenarios, where the participation of both the signer and the user are required. The user requires the authentication of the signer without revealing the message to the signer. This requirement is needed for real world applications such as client server applications in the banking scenario. As per our knowledge, the first password based blind short signature was constructed by Sangeetha et al. in CECC 2013 which ensures the properties unforgeability, blindness and unframeability. But if the password size is very small, it may be susceptible to off-line password guessing attack. In this paper we propose a strongly secure password based blind short signature which solves the off-line password guessing attack. The formal proof of the scheme is reduced to computational Diffie-Hellman(CDH) assumption.**

Index Terms—**Blind Signature Scheme, Password Based Blind Signature Scheme, Unforgeability, Blindness, Unframeability.**

I. INTRODUCTION

Eventhough there are tremendous growth in technologies in this twenty first century, secure data transmission is still appear to be a big hurdle and a lot of security issues need to be solved. Encryption schemes provide confidentiality where as digital signatures provide unforgeability. Digital signature scheme allows to sign documents in such a way that anyone can verify the authenticity of the signature. Diffie and Hellman [6] coined the notion of public key cryptosystem and Rivest et al. [7] proposed the first known digital signature called RSA(Rivest, Shamir and Adleman) signature scheme. The definition of security requirements for signature scheme was given by Goldwasser et al. [11] and the security proof for signature scheme in random oracle model was proposed by Pointchevel et al. [13]. The cryptosystems which is proved to be secure with random oracle uses cryptographic hash functions(preimage and collision resistant) and in the proof of security we assume that the output of hash functions follows a uniform distribution. Bellare et al. also had given the security proof of a RSA based digital signature in their classical work [1].

The idea of blind signature was put forward by David Chaum [5]. The applications like e-voting, digital cash etc require signatures which conceal the original message. The blind signature allows the user to get a signature without giving any information about the message to the signer and the signer cannot tell which session of the signing protocol corresponds to which message [8]. The properties of blind signatures are blindness and unforgeability. The provable secure design for blind signature is proposed by Pointchevel et al. [12] in which they defined the security for blind signatures with an application to electronic cash. Security arguments for blind signatures are proposed in papers([10],[14],[8]).

Gjosteen et al. [9] presented password based signature schemes based on RSA(Rivest, Shamir and Adleman) assumption and LRSW (Lysyanskaya, Rivest, Sahai and Wolf) assumption in which password is used as a random seed for the digital signature's key generation algorithm. Since passwords are short compared to key size, the key storage constraints can be solved. But these kind of schemes may be susceptible to online and off-line password guessing attacks for the low entropy passwords. In cryptography, Shannon([2],[3]) coined the term "entropy" which has been used as a measure of the difficulty in guessing or finding a password or a key. According to the NIST(National Institute of Standards and Technology) recommendations [4], 80 bits entropy are required for secure passwords. But passwords should be randomly selected passwords. Then the minimum threshold level of entropy can be obtained by using minimum 13 characters for the password from a 94 printable characters (Entropy, $H = log_2(b^l) \approx 85$ bits, where $b = 94$ and $l = 13$) which ensures the secrecy of the passwords. In different banking applications like e-locker facility, the secret information of the customer and the bank are together needed for transaction. For this purpose it is essential to generate signature mutually by using both secret key of customer and banker's secret key. For signature generation if customer is using certain threshold passwords along with banker's secret key it will increase the security as well as the efficiency of the system because customers can remember comparatively smaller passwords rather than using a large secret key. This insight motivates the construction of the password based blind signature(PBBS) scheme described in [21] in which both user's password and server's secret key are simultaneously used for signature generation.

Related Work: Gjosteen et al. [9] proposed password based signatures which prevents dictionary attacks. They introduced two password based signature schemes based on RSA [7] and CL(Camenisch and Lysyanskaya) [15] signatures. First scheme is easy to implement, but it does not achieve the security requirements. Second scheme is less practical, but it achieves stronger security. Password based signatures have a lot of applications in the banking scenario. Hence a password based

Sangeetha Jose is a Ph D scholar from Theoretical Computer Science Lab at Indian Institute of Technology Madras. (email: sangeethajosem@gmail.com).
Preetha Mathew K. is a Ph D scholar from Theoretical Computer Science Lab at Indian Institute of Technology Madras. (email: preetha.mathew.k@gmail.com).
C. Pandu Rangan is Professor in Department of Computer Science and Engineering at Indian Institute of Technology Madras(email: prangan55@gmail.com).

blind signature(PBBS) scheme is proposed in [21] by making use of blind version of BLS(Boneh, Lynn and Shacham) short signature scheme([17],[18]). In this paper we modified [21] to obtain a strongly secure password based blind signature(ss-PBBS).

Motivation: In all client server environment applications, if we use client's password as well as server's secret key for signing a document so as to ensure high efficiency and security. That is, client(user) and server(signer) can sign the document only by mutual agreement, so that user cannot generate signature without secret key of the signer(unforgeability) and the signer is not able to sign on behalf of the user without users password(unframeability). If the server's signature can be obtained by the client without revealing the message to the server, it is called blindness. To achieve the goals of unforgeability, unframeability and blindness, a password based blind signature construction is required which uses secret of both the client and server. This stimulates for the construction of a new password based blind signature(PBBS) scheme [21] in which message is being signed using both client's password as well as server's secret key. The password based blind signature scheme is based on blind version of BLS short signature scheme, which significantly reduces the signature size to 170 bits compared to Gjosteen et al.s password based signatures with 1024 bits and 2κ bits where κ is security parameter which is considered to be large. This scheme can be effectively used in banking applications. The key construction of the scheme is similar to Gjosteen et al.[9], but the rest of the construction is entirely different as shown in Table 1. Security proof of the scheme is elaborately given which is based on computational Diffie-Hellman(CDH) assumption in random oracle model. But there is a constraint in the size of password. In order to overcome this drawback we designed a strongly secure password based blind signature.

A. Organization of the Paper

Section 2 explains the preliminary concepts of bilinear pairing and the hardness assumptions which helps to prove the security of schemes. Section 3 gives the definitions of password based blind signatures and its security. Section 4 explains the password based blind signature scheme in [21]. Section 5 discusses strongly secure password based blind signature scheme, the proof of security and its advantages. The paper concludes in section 6.

II. Preliminary Concepts

A. Bilinear Pairing

Let $\mathbb{G}_1$ be a multiplicative cyclic prime order group q with generator g and $\mathbb{G}_2$ also be a multiplicative cyclic group of the same prime order q. A map $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ is said to be a bilinear pairing if the following properties hold.

1. $Bilinearity$: For all $g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*$, $e(g^a, g^b) = e(g, g)^{ab}$.
2. $Non\text{-}degeneracy$: For all $g \in \mathbb{G}_1$, $e(g, g) \neq I_{\mathbb{G}_2}$ where $I_{\mathbb{G}_2}$ is the identity element of $\mathbb{G}_2$.
3. $Computability$: e is efficiently computable.

B. Computational Diffie-Hellman (CDH) Assumption

Security proof of scheme is based on CDH assumption. CDH problem states that given (g, g^a, g^b), compute g^{ab}, where $g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*$.

Definition 1: **(CDH Assumption):** The advantage of any probabilistic polynomial time algorithm $\mathcal{A}$ in solving the CDH problem in $\mathbb{G}_1$ is defined as
$Adv_{\mathcal{A}}^{CDH} = Prob[g^{ab} \leftarrow \mathcal{A}(g, g^a, g^b) \mid g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*]$
The Computational Diffie-Hellman(CDH) assumption is that, for any probabilistic polynomial time algorithm $\mathcal{A}$, the advantage $Adv_{\mathcal{A}}^{CDH}$ is negligibly small(ϵ).

C. Conference-key Sharing Scheme (CONF)

CONF states that given (g, g^a, g^{ab}), compute g^b, where $g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*$.

Definition 2: **(CONF Assumption):** The advantage of any probabilistic polynomial time algorithm $\mathcal{A}$ in solving the CONF problem in $\mathbb{G}_1$ is defined as
$Adv_{\mathcal{A}}^{CONF} = Prob[g^b \leftarrow \mathcal{A}(g, g^a, g^{ab}) \mid g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*]$

III. Definition of Password Based Blind Signatures

Password based blind signature consists of different algorithms which is defined as follows [9].

Definition 3: **(Password Based Blind Signatures):** A password based blind signature scheme mainly consists of the following six algorithms.

- Setup(1^κ): A trusted third party outputs the public parameters by accepting the security parameter κ as input. It includes group parameters, message space, password space, hash functions, mappings etc. The parameters have public access by all the algorithms.
- KeyGen: These are interactive algorithms run by user and server. This algorithm inputs user password pw and outputs the values needed for obtaining signing key(sk_{PBBS}) of the server. It also generates secret and public keys(sk and pk) of both the user and server.
- Request(m, pk, pw): User runs this algorithm on message m and outputs the signature request L and the state information.
- Issue(L, pk, sk_{PBBS}): Server runs this algorithm in which signature request L is the input and the output is blind signature σ'.
- Unblind($\sigma', pk, state$): This algorithm is also run by the user. This makes use of blind signature σ', public keys and $state$ from $Request$ algorithm and outputs signature σ. But when the check fails, algorithm outputs $\perp$.
- Verify(m, σ, pk): Anyone can verify that whether σ is a valid signature on m under publicly available information like pk by $Verify$ algorithm. If it is a valid signature algorithm outputs 1, otherwise outputs 0.

User has a secret password with a minimum level of entropy which is chosen randomly.

A. Security Definitions of Password Based Blind Signatures

The security of the password based blind signatures can be convinced by proving its different properties which are as follows, *unforgeability, blindness and unframeability*. The additional property which is present in PBBS is that of $unframeability$. The other two, viz $unforgeability$ and $blindness$ are the properties of blind signature. The formal definition of the said properties are detailed below.

*1) **Unforgeability**:* In the formal definition of unforgeability, where the adversary $\mathcal{A}$ plays the role of user and the simulator will have the role of server. This game is based on random oracle and the challenger has to provide hash oracle and sign oracle($Issue$) and $\mathcal{A}$ tries to get "one-more" signature.

Definition 4: **(Unforgeability) [10]:** A password based blind signature scheme PBBS is said to be unforgeable, if the probability that $\mathcal{A}$ wins the following game is negligible.

- Step 1 (Setup Phase): $(pk, sk) \leftarrow KeyGen(1^\kappa)$.
- Step 2 (Training Phase): $\mathcal{A}$ engages in polynomially many(in κ) adaptive interactive protocols (hash and Issue oracles) with polynomially many copies of server(pk, sk). Let 'l' be the number of executions in which server outputs valid message-signature pair at the end of step 2.
- Step 3 (Forgery Phase): $\mathcal{A}$ outputs a set of $\{(m_1, \sigma_1), ..., (m_j, \sigma_j)\}$ where (m_i, σ_i) for $1 \le i \le j$ are all accepted by $Verify(m_i, pk, \sigma_i)$ for distinct m_i.

We can say that $\mathcal{A}$ wins the game when $j > l$. That is, $\mathcal{A}$ outputs more valid tuples (m, σ) than he/she received during the training phase.

*2) **Blindness**:* It ensures that server cannot distinguish between two messages m_0, m_1 which has already signed by him with the interaction of the user. For proving blindness, server plays the role of adversary $\mathcal{A}$ and challenger $\mathcal{C}$ will be the user.

Definition 5: **(Blindness) [10]:** A password based blind signature scheme PBBS satisfies the property of blindness, if the probability that $\mathcal{A}$ wins the following game is negligible.

- Step 1: $(pk, sk) \leftarrow KeyGen(1^\kappa)$
- Step 2: $\mathcal{A}$ produces two messages $\{m_0, m_1\}$ polynomial in 1^κ where $\{m_0, m_1\}$ are by convention lexicographically ordered and give to the $\mathcal{C}$.
- Step 3: $\{m_b, m_{1-b}\}$ are the same messages $\{m_0, m_1\}$ ordered by $\mathcal{C}$ according to the value of bit $b \in \{0, 1\}$ which is hidden from $\mathcal{A}$. $\mathcal{A}$ has given access to two interactive protocols with user $\mathcal{U}$, first with $\mathcal{U}(params, pk, m_b)$ and second with $\mathcal{U}(params, pk, m_{1-b})$.
- Step 4: Initially if the user protocol's output is σ_b(that is, does not output fail) and the next time user protocol's output is σ_{1-b},(that is, does not output fail) then only $\mathcal{A}$ gets σ_b, σ_{1-b} ordered according to the corresponding (m_0, m_1).
- Step 5: $\mathcal{A}$ outputs a bit $b^{'}$.

We can see that $\mathcal{A}$ can predict $b' = b$ only with a guessing probability. Therefore, we can define adversary $\mathcal{A}$'s advantage in the game as $|Pr[b' = b] - 1/2|$. That is, the server is not able to distinguish the messages that he/she signs in the previous sessions.

*3) **Unframeability**:* This is an additional property which is required for proving the security of the password based blind signature schemes. This property ensures that the server is not able to sign on behalf of the user without user's knowledge. Otherwise server has to find out user's password. Thus server will be the adversary $\mathcal{A}$ and tries to construct password based signature without the user intervention of the user. The formal definition of unframeability is as follows.

Definition 6: **(Unframeability) [9]:** A password based blind signature scheme PBBS is unframeable, if the probability that $\mathcal{A}$ wins the following game is negligible.

- Step 1 (Setup Phase): $(pk, sk) \leftarrow KeyGen(1^\kappa)$
- Step 2 (Training Phase): $\mathcal{A}$ engages in polynomially many(in κ) adaptive interactive protocols (hash, Request and Unblind oracles) with polynomially many copies of user(pk, sk). $\mathcal{A}$ can ask any number of queries to this oracles and decides in an adaptive fashion when to stop.
- Step 3 (Frameability Phase): $\mathcal{A}$ outputs a (m^*, σ^*) which has to be verified by $Verify(m^*, pk, \sigma^*)$ algorithm for a distinct m^*.

We say that $\mathcal{A}$ wins the game when Verify$(m^*, pk, \sigma^*) = 1$. That is, $\mathcal{A}$ outputs a valid tuple (m^*, σ^*) other than he/she received during the training phase without the help of the user.

IV. Password Based Blind Signature(PBBS)

Password based blind signature(PBBS) in [21] is shown in Fig. 1 which is an interaction between a user and a server(signer). The authentication protocol should be resistant to eavesdropping attacks, so that the protocol should not be attacked by an adversary to carry out offline attack. Here anyone can have a feel that if we expose $y = g^{H_2(pw)}$ as public key, it is susceptible to offline guessing attacks. But since the password is randomly selected, we can ensure the security by using 13 character passwords. According to the NIST(National Institute of Standards and Technology) recommendations [4], 80 bits entropy are required for secure passwords. The minimum threshold level of entropy can be obtained by using minimum 13 characters for a randomly selected password from a 94 printable characters (Entropy, $H = log_2(b^l) \approx 85$ bits, where b = 94 and l = 13) which ensures the security of the passwords. That is, it is quite infeasible for an attacker to do offline guessing in polynomial time.

Verification algorithm(Verify(m, σ, y_2, y)) helps to verify the validity of the message-signature pair.

if $e(\sigma, g) \stackrel{?}{=} e(H_1(m), y_2 y)$
  $return$ 1
else $return$ 0

To show the **correctness** of verification algorithm(Verify(m, σ, y_2, y)), the equation can be expanded as follows.

$$\text{Note that } \sigma = \frac{\sigma' L^{H_2(pw)} H_1(m)^\eta}{(y_1 y_2)^k} = \frac{L^{x_2 - \eta} L^{H_2(pw)} H_1(m)^{H_2(pw) - x_1}}{(g^{x_1} g^{x_2})^k}$$

Setup(1^κ): Select a pairing $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ where $\mathbb{G}_1$ and $\mathbb{G}_2$ are cyclic prime order group in q and a generator $g \in \mathbb{G}_1$. Select hash functions, $H_1 : \{0,1\}^* \rightarrow \mathbb{G}_1$ and $H_2 : \{0,1\}^* \rightarrow \mathbb{Z}_q{}^*$ and return public parameters $params \leftarrow (e, q, \mathbb{G}_1, \mathbb{G}_2, g, H_1, H_2)$.

USER		**SIGNER**
KeyGen$_{\mathcal{U}}(pw)$:		**KeyGen**$_{\mathcal{S}}(\eta)$:
$x_1 \leftarrow_R \mathbb{Z}_q^*$		$x_2 \leftarrow_R \mathbb{Z}_q^*$
$y_1 \leftarrow g^{x_1}$		$y_2 \leftarrow g^{x_2}$
$y \leftarrow g^{H_2(pw)}$		
$\eta \leftarrow H_2(pw) - x_1$	$\xrightarrow{\eta}$	Signing Key, $sk_{PBBS} = x_2 - \eta$
$return(x_1, y_1, y, \eta)$		$return(x_2, sk_{PBBS}, y_2)$

Secret Keys(sk): $sk_{\mathcal{U}} = x_1, sk_{\mathcal{S}} = x_2,$ Public Keys(pk): $pk_{\mathcal{U}} = y_1, pk_{\mathcal{S}} = y_2,\ y = g^{H_2(pw)}$

Request$_{\mathcal{U}}(m, pk, pw)$:		
$k \leftarrow_R \mathbb{Z}_q^*$		
$L = H_1(m)g^k$	$\xrightarrow{L}$	**Issue**$_{\mathcal{S}}(L, pk, sk_{PBBS})$:
$state \leftarrow (m, k, pw)$		$\sigma' = (L)^{sk_{PBBS}}$
	$\xleftarrow{\sigma'}$	

Unblind$_{\mathcal{U}}(\sigma', pk, state)$:
if $(e(L, y_2 g^{-\eta}) \stackrel{?}{=} e(\sigma', g))$
then
$$\sigma = \frac{\sigma' L^{H_2(pw)} H_1(m)^\eta}{(y_1 y_2)^k}$$
if $Verify(m, \sigma, y_2, y) = 1$
then $return(\sigma)$
$return\ (\perp)$

Fig. 1. Password based blind signature scheme(PBBS) in [21]

$$= \frac{L^{x_2+H_2(pw)-\eta} H_1(m)^{H_2(pw)-x_1}}{(g^{x_1} g^{x_2})^k}$$
$$= \frac{L^{x_1+x_2} H_1(m)^{H_2(pw)-x_1}}{g^{k(x_1+x_2)}}$$
$$= \frac{(H_1(m)g^k)^{x_1+x_2} H_1(m)^{H_2(pw)-x_1}}{g^{k(x_1+x_2)}}$$
$$= H_1(m)^{x_1+x_2+H_2(pw)-x_1}$$
$$= H_1(m)^{x_2+H_2(pw)}$$

Therefore,
$$e(\sigma, g) = e(H_1(m)^{x_2+H_2(pw)}, g)$$
$$= e(H_1(m), g^{x_2} g^{H_2(pw)})$$
$$= e(H_1(m), y_2 y)$$

V. STRONGLY SECURE PASSWORD BASED BLIND SIGNATURE SCHEME(SS-PBBS)

The strongly secure scheme is as in Fig. 2. This is made strongly secure by setting $y = g^{rH_2(pw)}$ where $r \in \mathbb{Z}_q^*$ which made public for verification of signature. Conference-key sharing (CONF) [23] assumption states that given (g, g^a, g^{ab}), compute g^b, where $g \in \mathbb{G}_1$ and $a, b \in_R \mathbb{Z}_q^*$, is hard to achieve [22]. Thus given g, g^r, $g^{rH_2(pw)}$, getting $g^{H_2(pw)}$ is hard. In ss-PBBS, g^r is not public and only g, $g^{rH_2(pw)}$ are public and hence the hardness of solving this is more than CONF. Eventhough $g^{rH_2(pw)}$ is public, offline password guessing attacks will not be effective because it is not possible to distinguish r and $H_2(pw)$ from $rH_2(pw)$. Since $H_2(pw)$ cannot be obtained by enumerating the values of $rH_2(pw)$ and thus finding pw is hard.

Verification algorithm(Verify(m, σ, y_2, y)) helps to verify the validity of the message-signature pair.

if $e(\sigma, g) \stackrel{?}{=} e(H_1(m), y_2 y)$
$return\ 1$
else $return\ 0$

To show the **correctness** of verification algorithm(Verify(m, σ, y_2, y)), the equation can be expanded as follows.

Note that $\sigma = \dfrac{\sigma' L^{rH_2(pw)} H_1(m)^\eta}{(y_1 y_2)^k}$
$$= \frac{L^{x_2-\eta} L^{rH_2(pw)} H_1(m)^{rH_2(pw)-x_1}}{(g^{x_1} g^{x_2})^k}$$
$$= \frac{L^{x_2+rH_2(pw)-\eta} H_1(m)^{rH_2(pw)-x_1}}{(g^{x_1} g^{x_2})^k}$$
$$= \frac{L^{x_1+x_2} H_1(m)^{rH_2(pw)-x_1}}{g^{k(x_1+x_2)}}$$
$$= \frac{(H_1(m)g^k)^{x_1+x_2} H_1(m)^{rH_2(pw)-x_1}}{g^{k(x_1+x_2)}}$$
$$= H_1(m)^{x_1+x_2+rH_2(pw)-x_1}$$
$$= H_1(m)^{x_2+rH_2(pw)}$$

Setup(1^κ): Select a pairing $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ where $\mathbb{G}_1$ and $\mathbb{G}_2$ are cyclic prime order group in q and a generator $g \in \mathbb{G}_1$. Select hash functions, $H_1 : \{0,1\}^* \rightarrow \mathbb{G}_1$ and $H_2 : \{0,1\}^* \rightarrow \mathbb{Z}_q{}^*$ and return public parameters $params \leftarrow (e, q, \mathbb{G}_1, \mathbb{G}_2, g, H_1, H_2)$.

USER		SIGNER
KeyGen$_{\mathcal{U}}(pw)$:		**KeyGen**$_{\mathcal{S}}(\eta)$:
$x_1 \leftarrow_R \mathbb{Z}_q^*$, $r \in_R \mathbb{Z}_q^*$		$x_2 \leftarrow_R \mathbb{Z}_q^*$
$y_1 \leftarrow g^{x_1}$		$y_2 \leftarrow g^{x_2}$
$y \leftarrow g^{rH_2(pw)}$		
$\eta \leftarrow rH_2(pw) - x_1$	$\xrightarrow{\eta}$	Signing Key, $sk_{PBBS} = x_2 - \eta$
$return(x_1, y_1, y, \eta)$		$return(x_2, sk_{PBBS}, y_2)$

Secret Keys(sk): $sk_{\mathcal{U}} = x_1, sk_{\mathcal{S}} = x_2$, Public Keys(pk): $pk_{\mathcal{U}} = y_1, pk_{\mathcal{S}} = y_2,\ y = g^{rH_2(pw)}$

Request$_{\mathcal{U}}(m, pk, pw)$:		
$k \leftarrow_R \mathbb{Z}_q^*$		
$L = H_1(m)g^k$	$\xrightarrow{L}$	**Issue**$_{\mathcal{S}}(L, pk, sk_{PBBS})$:
$state \leftarrow (m, k, pw)$		$\sigma' = (L)^{sk_{PBBS}}$
	$\xleftarrow{\sigma'}$	

Unblind$_{\mathcal{U}}(\sigma', pk, state)$:
if $(e(L, y_2 g^{-\eta}) \stackrel{?}{=} e(\sigma', g))$
then
$$\sigma = \frac{\sigma' L^{rH_2(pw)} H_1(m)^\eta}{(y_1 y_2)^k}$$
if $Verify(m, \sigma, y_2, y) = 1$
then $return(\sigma)$
$return\ (\perp)$

Fig. 2. Strongly secure password based blind signature scheme(ss-PBBS)

Therefore,

$$\begin{aligned} e(\sigma, g) &= e(H_1(m)^{x_2 + rH_2(pw)}, g) \\ &= e(H_1(m), g^{x_2} g^{rH_2(pw)}) \\ &= e(H_1(m), y_2 y) \end{aligned}$$

A. Proof of Security

The security of ss-PBBS scheme can be proved in consideration with the properties of unforgeability, blindness and unframeability. The following theorems show that proposed ss-PBBS scheme is perfectly unforgeable, blind and unframeable in the random oracle under computational Diffie Hellman(CDH) assumption.

Theorem 1: **The strongly secure password based blind signature is existentially unforgeable against adaptive chosen message attack(EUF-CMA) under CDH assumption with an advantage of challenger at least $\epsilon/e(1 + q_I)$.**

Proof:- In this simulation game adversary($\mathcal{A}$) plays the role of user($\mathcal{U}$) and the challenger($\mathcal{C}$) as that of the signer($\mathcal{S}$). The approach to security proof is similar to [16] and is as follows. If there exists an adversary $\mathcal{A}$ who can break the scheme, then there will be a challenger $\mathcal{C}$ who can make use of $\mathcal{A}$ to solve the CDH which is considered to be a hard problem.

- **Setup Phase**: Challenger chooses public system parameters $(e, q, \mathbb{G}_1, \mathbb{G}_2, g, H_1, H_2)$ in which H_1 and H_2 are *cryptographic hash functions* which behave as random oracle. $\mathcal{C}$ sets $y_2 = g^a$ which is considered to be the public key of the signer($pk_{\mathcal{S}}$) and sends public parameters and y_2 to $\mathcal{A}$.
- **Training Phase**: During this phase $\mathcal{A}$ is permitted to access the following oracles.
  - H_1**-Oracle:** H_1-Oracle works in the following way. An adversary can be able to make q_{H_1} queries with m_i and the challenger should be able to respond back to these queries with h_i. $\mathcal{C}$ maintains H_1-list and this will be empty initially. When $\mathcal{A}$ queries the oracle with m_i, $\mathcal{C}$ responds as follows.
    If the query comes with m_i, it checks whether it is in the H_1-list. If it is present in the H_1-list as a tuple $(hcoin_i, m_i, h_i, u_i)$, then $\mathcal{C}$ replies with h_i from the list. Otherwise, $\mathcal{C}$ flips a coin randomly where $hcoin \in \{0, 1\}$, which gives 1 with probability α and 0 with probability $1 - \alpha$. $\mathcal{C}$ also randomly chooses $u_i \in_R \mathbb{Z}_q^*$ and makes the H_1-list tuple as follows.
    1. If $hcoin = 0$, $\mathcal{C}$ sets $h_i = H_1(m_i) = g^{u_i}$ and insert the tuple $(hcoin_i, m_i, h_i, u_i)$ in to the H_1-list. Give h_i to $\mathcal{A}$.
    2. Else, sets $h_i = H_1(m_i) = g^{u_i} g^b$ and insert the tuple $(hcoin_i, m_i, h_i, u_i)$ in to the H_1-list. Respond this h_i as answer to the query by $\mathcal{A}$.

– H_2**-Oracle:** An adversary can be able to make q_{H_2} queries with pw_j and the challenger should be able to respond back to these queries. It is done by maintaining a H_2-list which is initially empty. When $\mathcal{A}$ queries the oracle with pw_j, $\mathcal{C}$ randomly take $w_j \in_R \mathbb{Z}_q^*$ and give $H_2(pw_j) = w_j$. Challenger also randomly selects $r_j \in_R \mathbb{Z}_q^*$ and stores (pw_j, w_j, r_j) in the H_2-list and later if the query appears with pw_j in the H_2-list, then gives the same w_j from the tuple (pw_j, w_j, r_j) to the adversary.

– **Issue Oracle:** In the unforgeability game, adversary $\mathcal{A}$ can access the $Issue$ oracle also. The signature is forgeable if the user is able to sign the message without the participation of the signer. Therefore, signer's privacy should be maintain in the unforgeability game rather than the privacy of the user. $\mathcal{A}$ chooses m_i and pw_j and requests the challenger $\mathcal{C}$ for the signature on message m_i with password pw_j. r_j and $w_j = H_2(pw_j)$ will be obtain from H_2-list, if it is already queried. Otherwise, $\mathcal{C}$ randomly take $w_j \in_R \mathbb{Z}_q^*$, $r_j \in_R \mathbb{Z}_q^*$ and give $H_2(pw_j) = w_j$ and store it as the tuple(pw_j, w_j, r_j) in the H_2-list. Then $\mathcal{C}$ checks that whether m_i is queried or not.
1. If m_i is queried, $\mathcal{C}$ retrieve the corresponding tuple $(hcoin_i, m_i,\ h_i, u_i)$ from the H_1-list. If $hcoin_i = 0$, $\mathcal{C}$ calculates and outputs $\sigma_i = y_2^{u_i}(h_i)^{r_j w_j}$. If $hcoin_i = 1$ then $\mathcal{C}$ aborts and reports failure.
2. If m_i is not queried, $\mathcal{C}$ runs the H_1-$Oracle$ to get the h_i, $hcoin_i$ and u_i values and insert these values in H_1-list. Then by using these values, produce the signature according to step 1 in $Issue\ Oracle$.

- **Forgery Phase**: On getting sufficient training, $\mathcal{A}$ produces a message-signature pair (m^*, σ^*) for a specific pw_j such that m^* is not queried to $Issue\ Oracle$ and σ^* is valid. But m^* should be queried to H_1-$Oracle$ and $\mathcal{C}$ obtains the tuple $(hcoin^*, m^*,\ h^*, u^*)$ from H_1-list. If m^* is not queried to H_1-$Oracle$ abort. From H_2-$Oracle$ $\mathcal{C}$ obtains r_j since the tuple consists of (pw_j, w_j, r_j). If m^* is queried, then in some cases $\mathcal{C}$ can solve hard problem(here CDH) as follows.
If $hcoin^* = 0$, $\mathcal{C}$ cannot do much and responds as simulation failure. But, if $hcoin^* = 1$, $\mathcal{C}$ can solve the CDH problem as follows. First $\mathcal{C}$ returns h^* and u^* from H_1-list and then compute g^{ab} as follows.
$$\begin{aligned}\frac{\sigma^*}{y_2^{u^*}\ (h^*)^{r_j w_j}} &= \frac{(h^*)^{a+r_j H_2(pw_j)}}{(g^a)^{u^*}\ (h^*)^{r_j w_j}} \\ &= \frac{(g^{u^*} g^b)^a\ (h^*)^{r_j H_2(pw_j)}}{(g^a)^{u^*}\ (h^*)^{r_j H_2(pw_j)}} \\ &= g^{ab}\end{aligned}$$
This solves CDH problem which is a contradiction to CDH assumption. This indicates that $\mathcal{A}$ cannot produce a valid signature σ^* for the message m^*. Thus, we can say that there is no forgery possible in polynomial time with non negligible advantage.

Probability Analysis: In the proof of Theorem 1, challenger needs to abort the game in certain situations. The requirement is that the probability of aborting is to be negligible. Suppose adversary makes a total of q_I issue queries. As mentioned earlier, let $hcoin \in \{0, 1\}$,be 1 with probability α and 0 with probability $1 - \alpha$. During simulation $hcoin = 1$ is the abort condition in training phase and $hcoin = 0$ is the abort condition in challenge phase. Therefore, the probability that challenger does not abort in training phase is $(1 - \alpha)^{q_I}$. The probability that challenger does not abort in forgery phase is α. Let challenger does not abort during training phase is E_1 and challenger does not abort during forgery phase is E_2

Pr(Challenger does not abort during simulation)=$Pr(E_1) \wedge Pr(E_2)$
Therefore,
Pr(Challenger does not abort during simulation)=$\alpha(1-\alpha)^{q_I}$. By maximizing this value at $\alpha_{opt} = 1-1/(q_I+1)$, probability that challenger does not abort during simulation is at least $1/e(1+q_I)$ which is non negligible, where q_I is the number of issue queries. Therefore, we can conclude that the advantage of challenger is at least $\epsilon/e(1 + q_I)$ as required. This probability analysis technique is similar to [19], where the authors use an approach similar to Coron's analysis [20] of the full domain hash signature scheme.

Theorem 2: **The strongly secure password based blind signature satisfies blindness such that it is infeasible for a malicious signer to distinguish between the two messages m_0 and m_1 has been signed first in two executions with the honest user.**

Proof:- In this game the role of adversary $\mathcal{A}$ and challenger $\mathcal{C}$ is interchanged from the above game. $\mathcal{A}$ provides public parameters($params$) and two messages $m_0, m_1 \in \mathbb{M}$ and sends to $\mathcal{C}$. A random bit $b \in \{0, 1\}$ is chosen by the $\mathcal{C}$ and order the messages as m_b and m_{1-b} based on the value of the selected bit 'b'. The random bit 'b' is hidden from $\mathcal{A}$. $\mathcal{A}$ has given black box access to two oracles $\mathcal{U}(params, pk, m_b)$ and $\mathcal{U}(params, pk, m_{1-b})$. This $\mathcal{U}$ algorithms perform PBBS protocol and produce the outputs σ_b and σ_{1-b} corresponds to m_b and m_{1-b}. If $\sigma_b \neq \perp$ and $\sigma_{1-b} \neq \perp$ then only $\mathcal{A}$ receives (σ_0, σ_1). If $\sigma_b = \perp$ and $\sigma_{1-b} \neq \perp$ then $\mathcal{A}$ receives $(\perp, \epsilon)$. If $\sigma_b \neq \perp$ and $\sigma_{1-b} = \perp$ then $\mathcal{A}$ receives $(\epsilon, \perp)$. If $\sigma_b = \perp$ and $\sigma_{1-b} = \perp$ then $\mathcal{A}$ receives $(\perp, \perp)$. After accessing the black boxes $\mathcal{A}$ tries to predict 'b' and we prove that $\mathcal{A}$ can do this with negligible advantage. That is, there is only guessing probability.
Challenger selects k randomly from $\mathbb{Z}_q^*$ and sends L to $\mathcal{A}$ where $L = H_1(m_b)g^k$ which is uniformly distributed in $\mathbb{G}_1$. $\mathcal{A}$ returns back $\sigma' \in \mathbb{G}_1$ to the first oracle($\mathcal{U}(params, pk, m_b)$) and chooses the value using any strategy he/she wants. At this point $\mathcal{A}$ fixes on the value and he/she is able to predict the output σ_i of the oracle $\mathcal{U}(params, pk, m_b)$ with negligible advantage as follows.
Step 1: $\mathcal{A}$ checks if $e(L, y_2 g^{-\eta}) = e(\sigma', g)$ holds. If the check fails, record σ_b as $\perp$. Otherwise record the value as σ_b.
Step 2: Similar to above $\mathcal{A}$ chooses any value $\sigma' \in \mathbb{G}_1$ for the second oracle and do the similar check. If the check fails, record σ_{1-b} as $\perp$. Otherwise record the value as σ_{1-b}.
Step 3: If $\sigma_b = \perp$ and $\sigma_{1-b} \neq \perp$ then output $(\perp, \epsilon)$. If $\sigma_b \neq \perp$ and $\sigma_{1-b} = \perp$ output $(\epsilon, \perp)$. If both checks fails then output $(\perp, \perp)$. If anyone of these three cases occurs, abort.
Step 4: Finally the adversary, $\mathcal{A}$ could predicts (σ_b, σ_{1-b}) only

if $\sigma_b \neq \perp$ and $\sigma_{1-b} \neq \perp$. That is, if both check succeeds then $\mathcal{A}$ initiates PBBS protocol on m_b and m_{1-b} and outputs σ_b, σ_{1-b} respectively. If either protocol run fails, abort.
This prediction is true because $\mathcal{A}$ performs the same check as that of honest user. If $\mathcal{A}$ is able to predict the final output of its oracles accurately, then $\mathcal{A}$'s advantage in distinguishing $\mathcal{U}(params, pk, m_b)$ and $\mathcal{U}(params, pk, m_{1-b})$ is the same without this final output. Therefore, all of $\mathcal{A}$'s advantage to distinguish between these signatures must come from distinguishing the earlier message of the oracles(L). These oracles send only uniformly random values and hence $\mathcal{A}$ cannot distinguish between them with non-negligible probability. Therefore we can define adversary $\mathcal{A}$'s advantage in the game as $|Pr[b' = b] - 1/2|$.

Theorem 3: **If CDH assumption holds, the strongly secure password based blind signature provides unframeability under random oracle.**

Proof:- To prove the unframeability, signer should not be able to create a signature on behalf of the user without finding user's password. We can prove the security of the scheme under CDH assumption. In this simulation game signer plays as adversary and user as challenger.

- **Setup Phase**: Challenger $\mathcal{C}$ sets $y = g^a$ where $a = rH_2(pw)$. $\mathcal{C}$ sends public parameters and y to $\mathcal{A}$.
- **Training Phase**: During this phase $\mathcal{A}$ has access to $Request$ and $Unblind$ oracles along with H_1-$Oracle$.
  - H_1**-Oracle:** This hash oracle is similar to that of H_1-oracle in the security proof of $Theorem$ 1 with only difference is that it is provided by the user.
  - **Request Oracle:** In this phase $\mathcal{A}$ selects m_i and queries for signature request,L from the $\mathcal{C}$. It can be simulated as follows.
    1. If m_i is queried, $\mathcal{C}$ retrieve the tuple $(hcoin_i, m_i, h_i, u_i)$ corresponds to m_i from the H_1-list. $\mathcal{C}$ randomly selects $k \in_R \mathbb{Z}_q^*$ and computes $L = h_i g^k$ where $h_i = H_1(m_i)$. $\mathcal{A}$ gets L as output from the $Request\ Oracle$.
    2. If m_i is not queried, run the H_1-$Oracle$ and gets h_i corresponds to m_i and do the similar step as above.
    Here the $Request\ Oracle$ is similar to the normal $Request$ algorithm. $Unblind\ Oracle$ can be simulated as follows.
  - **Unblind Oracle:** $\mathcal{A}$ queries this oracle with a message,m_i.
    1. If m_i is queried, $\mathcal{C}$ retrieve the tuple $(hcoin_i, m_i, h_i, u_i)$ corresponds to m_i from the H_1-list. Then, if $hcoin_i = 0$, $\mathcal{C}$ calculates and outputs $\sigma_i = (y\ y_2)^{u_i}$. If $hcoin_i = 1$ then $\mathcal{C}$ aborts and reports failure.
    2. If m_i is not queried, run the H_1-$Oracle$ and insert the tuple $(hcoin_i, m_i, h_i, u_i)$ in to the H_1-list. Then produce the signature according to step 1 in $Unblind\ Oracle$.
- **Frameability Phase**: After getting sufficient training, $\mathcal{A}$ produces a message-signature pair (m^*, σ^*) such that such that m^* is not queried to $Request$ and $Unblind\ Oracle$ and σ^* is valid. But m^* should be queried to H_1-$Oracle$ and $\mathcal{C}$ obtains the tuple $(hcoin^*, m^*, h^*, u^*)$ from H_1-list. If m^* is not queried to H_1-$Oracle$ abort. If m^* is queried, then in some cases challenger $\mathcal{C}$ can solve hard problem(here again CDH) as follows.
  If $hcoin^* = 0$, $\mathcal{C}$ cannot do much and responds as simulation failure. But, if $hcoin^* = 1$, $\mathcal{C}$ can solve the CDH problem as follows. First $\mathcal{C}$ returns u^* from H_1-list and then compute g^{ab} and g^{bx_2} as follows.
  $$\frac{\sigma^*}{(y\ y_2)^{u^*}} = \frac{(h^*)^{a+x_2}}{y^{u^*}\ g^{x_2 u^*}} = \frac{(g^{u^*} g^b)^a\ (g^{u^*} g^b)^{x_2}}{(g^a)^{u^*}\ (g^{x_2})^{u^*}} = g^{ab}\ g^{bx_2}$$
  $\mathcal{C}$ knows (g, g^a, g^b, g^{x_2}) only and compute g^{ab} and g^{bx_2} is known to be CDH problem which is considered to be hard problem. Till today, there is no polynomial time algorithm exists for solving CDH problem. This indicates that $\mathcal{A}$ cannot produce valid signature σ^*. Thus, we can say that there is no frameability possible in polynomial time with non negligible advantage or the scheme is unframeable.

The probability analysis of Theorem 3 is similar to Theorem 1.

Scheme	Underlying Signature	Hardness Assumption	Signature Size
Gjosteen et al. Scheme 1 [9]	RSA	RSA Inversion	1024 bits
Gjosteen et al. Scheme 2 [9]	CL	LRSW	$2\kappa^*$ bits
PBBS Scheme [21]	BLS	CDH	170 bits(constraint in password size)
ss-PBBS Scheme	BLS	CDH	170 bits(no constraint in password size)

$^*\kappa$ is security parameter

TABLE I
COMPARISON WITH EXISTING SCHEMES

B. Advantages

Since the scheme(ss-PBBS) is using both signer's secret key and user's password, it provides more stronger security and it has more efficiency than the existing schemes [9] as shown in Table 1. There is no constraint for the password size and the scheme is not susceptible to offline-password guessing attacks. Thus ss-PBBS scheme is more suitable for client server applications especially for banking applications where both customer and bank secret information are needed for transaction without any password guessing attack.

VI. CONCLUSION

ss-PBBS scheme is strongly secure scheme and is not susceptible to off-line password guessing attack even if the password size is small. Security proof for this scheme in standard model is an open problem. The scheme can also be made to a honest-user unforgeable password based blind signature scheme using the generic transformation given in [8].

REFERENCES

[1] Mihir Bellare and Phillip Rogaway. "The Exact Security of Digital Signatures - How to Sign with RSA and Rabin". In *EUROCRYPT*, pages 399-416, 1996.

[2] Claude E. Shannon. "A mathematical Theory of Communication". In Bell System Technical Journal, Vol. 27, pp. 379423, October, 1948.

[3] Claude E. Shannon. "Prediction and Entropy of Printed English". In Bell System Technical Journal, Vol. 30, n. 1, pp. 50-64, 1951.

[4] William E. Burr, Donna F. Dodson and W. Timothy Polk. "Information Security". Electronic Authentication Guideline, Recommendations of the National Institute of Standards and Technology(NIST), Special Publication 800-63, Version 1.0.2, April, 2006.

[5] David Chaum. "Blind Signatures for Untraceable Payments". In *Advances in Cryptology - Crypto '82*, Lecture Notes in Computer Science, pages 199-203, Springer, 1983.

[6] Whitfield Diffie and Martin E. Hellman. New Directions in Cryptography. In *IEEE Transactions on Information Theory*, Volume 22(6), pages 644-654, 1976.

[7] Ronald L. Rivest, Adi Shamir and Leonard M. Adleman. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. In *Communications of the ACM*, Volume 21(2), pages 120-126, 1978.

[8] Dominique Schroder and Dominique Unruh. "Security of Blind Signatures Revisited". In *Public Key Cryptography*, Lecture Notes in Computer Science, pages 662-679, Springer, 2012.

[9] Kristian Gjosteen and Oystein Thuen. "Password-Based Signatures". In *EuroPKI*, Lecture Notes in Computer Science, pages 17-33, Springer, 2011.

[10] Ari Juels, Michael Luby and Rafail Ostrovsky. "Security of Blind Digital Signatures (Extended Abstract)". In *CRYPTO*, Lecture Notes in Computer Science, pages 150-164, Springer, 1997.

[11] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. "A digital Signature Scheme Secure Against Adaptive Chosen-message attacks". *SIAM J. Comput.*, 17(2):281-308, April 1988.

[12] David Pointcheval and Jacques Stern. "Provably Secure Blind Signature Schemes". In *ASIACRYPT*, pages 252-265, 1996.

[13] David Pointcheval and Jacques Stern. "Security Proofs for Signature Schemes". In *EUROCRYPT*, pages 387-398, 1996.

[14] David Pointcheval and Jacques Stern. "Security Arguments for Digital Signatures and Blind Signatures". In *J. Cryptology*, 13(3):361-396, 2000.

[15] Jan Camenisch and Anna Lysyanskaya. "Signature Schemes and Anonymous Credentials from Bilinear Maps". In *CRYPTO 2004*, pages 56-72, 2004.

[16] Jianhong Zhang and Xiuna Su. "Another Efficient Blind Signature Scheme based on Bilinear Map". In *6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1-4, 2010.

[17] Dan Boneh, Ben Lynn and Hovav Shacham. "Short Signatures from the Weil Pairing". In *ASIACRYPT*, Lecture Notes in Computer Science, pages 514-532, Springer, 2001.

[18] Alexandra Boldyreva. "Threshold Signatures, Multisignatures and Blind Signatures Based on the Gap-Diffie-Hellman-Group Signature Scheme". In *Public Key Cryptography*, Lecture Notes in Computer Science, pages 31-46, Springer, 2003.

[19] Dan Boneh and Matthew Franklin. Identity-Based Encryption from the Weil Pairing. *SIAM Journal of Computing*, Volume 32(3), pages 586615, 2003.

[20] Jean-Sébastien Coron. "On the Exact Security of Full Domain Hash". In *CRYPTO 2000*, Lecture Notes in Computer Science, pages 229-235, Springer, 2000.

[21] Sangeetha Jose, Preetha Mathew K. and C. Pandu Rangan. "Password Based Blind Short Signature for Real Time Applications". In *Central European Conference on Cryptography, CECC'13*, Telc, Czech Republic, June 2013.

[22] Kouichi Sakurai and Hiroki Shizuya. "Relationships Among the Computational Powers of Breaking Discrete Log Cryptosystems". In *EUROCRYPT*, Lecture Notes in Computer Science, pages 341-355, Springer, 1995.

[23] Tatsuaki Okamoto. "Encryption and Authentication Schemes Based on Public-key Systems". In *Ph.D. Thesis*, The University of Tokyo, 1988.

Sangeetha Jose is a Ph D scholar from Theoretical Computer Science Lab at Indian Institute of Technology (IIT) Madras, Chennai, India. She is working under the guidance of Prof. C. Pandu Rangan. Her research interests are in provable security mainly focus on the design and analysis of public key encryption and digital signatures and the security of cloud computing. Contact her at sangeethajosem@gmail.com.

Preetha Mathew K. is a Ph D scholar in Indian Institute of Technology (IIT) Madras, Chennai, India. She is working under the guidance of Prof C. Pandu Rangan. Her areas of interest focus on provably secure post quantum cryptosystems, especially in code-based cryptosystem and the security issues in cloud computing. Preetha Mathew K. can be contacted at preetha.mathew.k@gmail.com.

C. Pandu Rangan is a Professor in the department of computer science and engineering of Indian Institute of Technology (IIT) Madras, Chennai, India. Theoretical Computer Science Lab in IIT Madras is headed by him. His areas of research interests mainly focus on cryptography, security issues in cloud computing, algorithms and data structures, game theory, graph theory and distributed computing. C. Pandu Rangan can be contacted at prangan55@gmail.com.

An Adaptive Load Sharing Algorithm for Heterogeneous Distributed System

P.Neelakantan, A.Rama Mohan Reddy

Abstract— Due to the restriction of designing faster and faster computers, one has to find the ways to maximize the performance of the available hardware. A distributed system consists of several autonomous nodes, where some nodes are busy with processing, while some nodes are idle without any processing. To make better utilization of the hardware, the tasks or load of the overloaded node will be sent to the under loaded node that has less processing weight to minimize the response time of the tasks. Load balancing is a tool used effectively for balancing the load among the systems. Dynamic load balancing takes into account of the current system state for migration of the tasks from heavily loaded nodes to the lightly loaded nodes. In this paper, we devised an adaptive load-sharing algorithm to balance the load by taking into consideration of connectivity among the nodes, processing capacity of each node and link capacity.
Keywords: Load balancing, Distributed System, heterogeneous, response time .

I. INTRODUCTION

An important attribute in a dynamic load balancing policy is to initiate the load balancing activity that specifies which node is responsible for detecting imbalance of the load among the nodes [9]. A load-balancing algorithm is invoked when load imbalance among the nodes is detected. The initiation of load balancing activity will have a higher impact on complexity, overhead and scalability. The load balancing algorithm is designed in such a way to make the overloaded node to transfer its excess load to the underloaded node which is called sender – initiated and when underloaded node requests the load from the overloaded node then it is called receiver-initiated [6][8].

Domain balancing is used to decentralize the balancing process by minimizing its scope and decreasing the time complexity of the load-balancing algorithm. A domain is defined as subset of nodes in a system, such that a load balancing algorithm can be applied for this subset of nodes in a single step. Domain balancing is used in load balancing algorithms to decentralize the balancing. The balancing domains are further divided into two types: The first type is overlapped domains, which consists of node initiating the balancing activity and balances its load by migrating the tasks or load units with the set of surrounding nodes. [3]

Global balancing is achieved by balancing every domain and by diffusing the excess load throughout the overlapped domains in a distributed system. Another important attribute in load balancing algorithm is the degree of information. The degree of information plays an important role in making the load balancing decisions. To achieve global load balancing in a few steps, the load balancing should get absolute information instead of getting the obsolete information from the nodes. In general, the collection of information by a node is restricted to the domain or nearest neighboring nodes (which are directly connected to a node)[4].

Although collecting information from all the nodes in a distributed system gives the exact knowledge of the system, it introduces large communication delay, so from this perspective, it will have a negative impact on the load balancing algorithm. In such cases, it has been observed, that the average response time is kept minimum without load balancing instead of doing the load balancing which induces overhead in migrating the load from one node to another node in the system [5].

In this section, an abstract view of the software details is presented for load balancing. The distributed system consists of several nodes and the same load balancing software is installed to run on all the nodes in the distributed system. By installing the same software in all the nodes, the load balancing decision is taken by a node locally (decentralized) by collecting the information from the neighboring nodes as opposed to the centralized load balancing policy [14].

The program must use a multi-threaded concept to implement load balancing in a distributed system. Two communication ports are available: TCP and UDP. UDP is preferable as it incurs less communication overhead. In general the architecture provides three layers: Communication layer, Load balancing process and application layer [14][10]. For storing information two data structures were used.

The communication link is responsible for four phases: node status information phase, node status reception, tasks reception and task migration. The node status information is responsible for disseminating the load

Manuscript submitted May 3, 2013, revised Aug 13, 2013. P. Neelakantan is with the Department of CSE, SVU College of Engineering, Tirupati-517501 (e-mail:pneelakantan@rediffmail.com). Dr.A.Rama Mohan Reddy is with the Department of CSE, SVU College of Engineering, Tirupati, (e-mail:

information to the node that has requested it. The exchange of the information has a profound effect on the load balancing decision; it has to be done according to the predefined intervals of time specified on each node[7][14].

The status reception is responsible for receiving the status information from the other nodes and it will be updated in the local node list which is running the status reception phase. Here it is possible to distinguish the old information from the new information. The technique that is used to find is to associate the timestamp for the information that it has received from some node (say $TS_j^i(Inf)$, the time stamp attached to the information received from j to i). The local node say i maintaining the status about the node j is kept in the memory. If any estimate regarding node j exists in the node i memory, it will be compared to the received time stamp message and drops the old time stamp and the new timestamp message has been saved in the memory as the old time stamp has the obsolete information [11][1][2].

Once a node collects the above information, it knows whether it is overloaded or underloaded. In case if it is overloaded node, it transmits the excess tasks (loads) to the underloaded nodes in a "tasks transmission" phase. The next initiation of load balancing activity will be done only when the current migration of load units to the underloaded nodes is completed.

The "task reception" is responsible for listening to the requests and accepts the tasks sent from the other nodes. As we can observe from the above situations, the minimum time to initiate the new load balancing activity takes three time instants. One instant for receiving the status of all the nodes and second time instant for determining the underloaded nodes and computing the excess load and third time instant for transferring the excess load to the underloaded nodes which has been determined in the second time instant. So, the new load balancing activity takes place only at the fourth time instant [12] [14].

In a few papers [3] [9] [10], it is assumed that the nodes will not fail. The problem arises when the nodes fail which is common in the distributed systems. Sometimes a communication link will also fail, so the node will be unreachable. These two aspects i.e., failure of a node and the communication link will affect greatly the load balancing algorithms. Let us assume the following scenario. The overloaded node has collected the load information from the neighboring nodes and found some of the nodes are low loaded as discussed earlier. Now at the given time instant when the node tries to send its excess load to the overloaded node, it will not succeed because of the failure of the node. The node may fail after sending the status information. If this happens, an alternative must be chosen to avoid a failure of the load-balancing algorithm.

II. NOTATIONS & ASSUMPTIONS

N: Number of nodes
V= {1, 2… N} a set of nodes in a system
q_i: Number of tasks in the queue of node i
$w_i(t)$: Expected waiting time experienced by a task inserted into the queue at the i^{th} node in time t
$A_i(t)$: rate of generation of waiting time on i^{th} node caused by the addition of tasks in time t.
$S_i(t)$: rate of reduction in waiting time caused by the service of the tasks at the i^{th} node in time t.
$r_i(t)$: rate of removal(transfer) of the tasks from node j to node i at time t by the load balancing algorithm at node j.
ts_i: Average completion time of the task at node i.
b_i: Average size of the task in bytes at node i when it is transferred
d_{ij}: Transfer rate in bytes/sec between node i and node j
$\bar{q}_i(t)$: Average size of the queue calculated by node i based on its domain information at time t.
D_i: Neighboring nodes to i which is defined as $D_i = \{j | j \in V \text{ and } (i,j) \in E\}$ where V= {1, 2…N}
$E_i(t)$:Excess number of tasks at node i at time t.
f_{ij}: Portion of the excess tasks of node i to be transferred to node j decided by the load balancing algorithm.

The following assumptions were made in this paper:

1. It is assumed that a distributed system consists of N heterogeneous nodes interconnected by an underlying arbitrary communication network. Each node i in a system has a processing weight P_i >0 and processing capacity S_i>0. The load is defined to be $L_i = P_i/S_i$. In homogenous case the value of $L_i=P_i$.
2. It has been assumed that tasks arrive at node i according to Poisson process with rate $\lambda_i(t)$. A task arrived at node i may be processed locally or migrated through the network to another node j for remote processing. Once the task is migrated it remains there until its completion.
3. It is assumed that there is a communication delay incurred when task is transferred from one node to another before the task can be processed in the system. The communication delays are different for each link.

Each node contains an independent queue where arrived tasks are added to the queue, which results in accumulation of waiting time. Load balancing must be done repeatedly to maintain load balance in the system. Each node runs the load-balancing algorithm individually and hence the proposed algorithm is distributed in nature.

The second level of the system is a load-balancing layer, which consists of load balancing algorithms. The load balancing process is initiated by using predefined or randomly generated time instants, kept in a file. The algorithm determines the portion of the excess load to be sent to the underloaded node based on the current state of the node and availability of the nodes in the network. The load balancing algorithm must consider the communication delay while migrating the tasks to the other nodes. The algorithm selects the tasks to migrate to other nodes by setting their status as inactive to avoid execution of the tasks by current node application during the transition period. After completion of the task transmission activity, the status of the tasks is set to active when they are not transmitted to any node. When the tasks are transmitted to other nodes during the task transmission phase then those tasks are removed from the task queue of the current node.

Application layer consists of two threads of control: Task input and task execution threads. The task input creates a number of tasks defined in the initialization file and inserts them in the task queue. This task input is also responsible for adding the new tasks to the task queue either from the current node or from other nodes in the system. The task execution thread is responsible for execution of the tasks and updating the QSize variable by removing the task from the task queue.

The load balancing policy must take into account of processing capacity of the node while migrating the tasks to it. The selected node may become a candidate for one or more overloaded node in a given time instant because of the decentralized policy. Another issue to be considered is variable task completion times. Taking these issues a priori is not possible so a load balancing strategy must be adaptive to the dynamic state changes in the system and act accordingly to transfer the tasks. Even this can result in task shuttle between the nodes, so a migration limit for a task should be set to avoid task thrashing.

Another issue to be considered while migrating the tasks from one node to another node in a system is communication overhead. Large communication delays will have a negative impact on the load balancing policy, so, the transfer delays must be taken into account while migrating the task. When the completion of the task time in current node is greater than the completion time on task in another node inclusive of communication overhead, then only a task is considered for migration.

III. MATHEMATICAL MODEL

The mathematical model for load balancing in a given node i is given by [1] [2]

$$\frac{dw_i(t)}{dt}=A_i - S_i + r_i(t) - \sum_{j=1}^{\neq N_i} f_{ij}\frac{ts_i}{ts_j} r_j(t-\tau_{ij}) \quad (1)$$

$$E_i(t)= q_i(t)- \bar{q}_i(t)$$

$$r_i(t) = G_i(E_i(t))$$

$$f_{ij} \geq 0, f_{ii}=0,\ \sum_{j=1}^{\neq N_i} f_{ij} = 1$$

$$E_i(t) = \begin{cases} E & if \quad y \geq 0 \\ 0 & if \quad y < 0 \end{cases}$$

When a task is inserted into the task queue of node i, then it experiences the expected waiting time which is denoted by $w_i(t)$.

Let the number of tasks in i^th node is denoted by $q_i(t)$.

Let the average time needed to service the task at node i ts_i .

The expected (average) waiting time is given by at node i is given by $w_i(t) = q_i(t)ts_i$.

Note that $w_i(t)/ts_i = q_i$ is the number of tasks in the node i queue.

Similarly $w_k(t)/ts_k = q_k$ is the queue length of some node k. If tasks on node i were transferred to some node k, then the waiting time transferred is $q_i ts_k = \frac{w_i(t)ts_k}{ts_i}$, so that the fraction ts_k/ts_i converts waiting time on node i to waiting time on node k.

A_i : Waiting time generated by adding the task in the i^th node.

S_i : Rate of reduction in waiting time caused by the service of tasks at the i^th node is given by $S_i = (1 * tp_i)/tp_i=1$ for all $w_i(t) > 0$.

$r_i(t)$: The rate of removal (transfer) of the tasks from node i at time t by the load balancing algorithm at node i. f_{ij} is the fraction of i^th node tasks to be sent out to the j^th node. In more detail $f_{ij}r_i(t)$ is the rate at which node i sends waiting time (tasks) to node i at time t where f_{ii}>=0 and f_{ii}=0.That is, the transfer from node i of expected waiting time (tasks) $\int_{t_1}^{t_2} E_i(t)dt$ in the interval of time $[t_1, t_2]$ to the other nodes is carried out with the j^{th} node receiving the fraction $p_{ij}(t_{p_j}/t_{p_i})\int_{t_1}^{t_2} u_i(t)dt$ where the ratio t_{p_j}/t_{p_i} converts the task from waiting time on node i to waiting time on node j. As $\sum_{i=1}^{n}(f_{ij}\int_{t_1}^{t_2} E_i(t)dt) = \int_{t_1}^{t_2} E_i(t)dt$, this results in removing all of the waiting time $\int_{t_1}^{t_2} E_i(t)dt$ from node i.The quantity $f_{ij}E_i(t-\tau_{ij})$ is the rate of increase (rate of transfer) of the expected waiting time (tasks) at time t from node i by (to) node j where $\tau_{ij}(\tau_{ii} = 0)$ is the time delay for the task transfer from node i to node j.

In this model, all rates are in units of the rate of change of expected waiting time, or time/time which is dimensionless. As $E_i(t) \geq 0$, node i can only send tasks to other nodes and cannot initiate transfers from another node to itself. A delay is experienced by transmitted tasks before they are received at the other node. The control law $E_i(t) = G_i * E_i(t)$ states that if the i^{th} node output $w_i(t)$ is above the domain average $(\sum_{j=1}^{n} q_j(t-\tau_{ij}))$/n, then it sends data to the other nodes, while if it is less than the domain average nothing is sent. The j^{th} node receives the fraction $\int_{t_1}^{t_2} F_{ij}(t_{p_i}/t_{p_j})\, u_i(t)dt$ of transferred waiting time $\int_{t_1}^{t_2} E_i(t)dt$ delayed by the time τ_{ij}.The model described in (1) is the basic model for load balancing, but an important feature is to determine f_{ij} for each underloaded node j. One approach is to distribute the excess load equally to all the underloaded neighbors.

$$f_{ij} = \frac{1}{n-1} \text{ for } i \neq j.$$

Another approach is to use the load information collected from the neighbors to determine the deficit load of the neighbors. The deficit load of the neighbours shall be determined by node i by using the formula (2)

$$q_j(t\text{-}\tau_{ij}) - \bar{q}_i \quad (2)$$

The above formula is used by node i to compute the deficiency waiting times in the queue of node j with respect to the domain load average of node i.

If node j queue is above the domain average waiting time, then node i do not send tasks to it. Therefore $(\bar{q}_i - q_j(t$-

τ_{ij})) is a measure by node i as how much node j is behind the domain average waiting time. Node i performs this computation for all the other nodes which are directly connected to it and then portions out its tasks among the other nodes that fall below the domain queue average of node i.

$$f_{ij} = \frac{(\bar{q}_i - q_j(t-\tau_{ij}))}{\sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij})} \qquad (3)$$

If the denominator $\sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij})$=0 then fij are defined to be zero then no waiting times are transferred. If the denominator $\sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij})$=0, then$(\bar{q}_i - q_j(t - \tau_{ij}) \leq 0 \forall j \in N_i$. However by definition of the average $\sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij})+\bar{q}_i - q_i(t) = \sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij}))=0$ which implies $\bar{q}_i - q_j(t) = \sum_{j=1}^{N_i}(\bar{q}_i - q_j(t-\tau_{ij})) > 0$

That is, if the denominator is zero, the node j is not greater than its domain queue average, so $E_i(t) = G_iE_i(t))=0$, where G is Gain Factor.f_{ij} :Portion of the excess tasks of node i to be transferred to node j decided by the load balancing algorithm. Except the last three parameters remaining information is known at the time of load balancing process. Before the instance of load balancing activity, every variable is updated.

IV. PROPOSED ALGORITHM

Algorithm ALS

The current node i, performs the followings:

a. Calculate the average queue size ($\bar{q}_i$)based on the information received from the neighbouring nodes.

$$\bar{q}_i = \frac{1}{\neq N_i + 1}\sum_{j=1}^{\neq N_i}\left(q_i + q_j\frac{ts_j}{ts_i}\right)$$

if ($q_i > \bar{q}_i$)then $E_i = (q_i - \bar{q}_i) * G$
else Exit.

b. Determine the participant nodes in load sharing process.

Participants= {j| $q_j < \bar{q}_i$, ∀j∈N_i}

c. Calculate the fraction of the load (f_{ij}') to be sent to the participants

$$f_{ij}' = \frac{\bar{q}_i - (\frac{ts_j}{ts_i})q_j}{\sum_{j=1}^{N_i}(\bar{q}_i - (\frac{c_j}{c_i})q_j}$$

d. Calculate maximum portion of the excess load (f_{ij}'')

$$f_{ij}'' = \frac{(q_i - E_i)\, ts_i \, \mathrm{dij}}{E_i b_i}$$

e. fij = Min (f_{ij}', f_{ij}'')

a. Announce to node j about its willingness to send $T_{ij} = fij * E_i$ tasks;
b. nowReceived = call procedure acceptanceFromNodej()
c. if(nowReceived >0)
   i. Transfer NowReceived to j
   ii. $T_{ij}=T_{ij}$- NowReceived

End if

g. Repeat steps from (a) to (f).

```
Procedure acceptanceFromNodej()
  if ((q_j+ T_ij)< q̄_j nowSend=q̄_j − q_j;
  else nowSend=-1;
return now Send;
end acceptanceFromNodej
```

In general it is assumed that keeping the Gain factor G=1 will give the good performance. But in a distributed system with largest delays and the nodes that have domain queue average outdated gives poor result. This phenomenon was first observed by the load balancing group at the University of New Mexico [7]. So the G values are set in the way that yields an optimal result. Another step that is added in the above algorithm is to test the node availability. It checks both node availability as well as the amount of waiting times it can receive. The node executing the ALS is permitted to send the tasks to the neighbors after receiving the acknowledgement specifying the amount of the load they can be able to process.. The time complexity of the proposed algorithm is O(d) shown in table 1.

Table 1: ALS Operations

Sno	Actions	Operation	Quantity, (d is the number of neighbors)
1	Compute average queue size	Addition Division Multiplication	d+1 d d
2	Compute Ei	Subtraction Multiplication	1 1
3	Determine the participant nodes	Comparison	d
4	Compute f_{ij}'	Subtraction Division Multiplication	d+1 d+1 d+1
5	Compute f_{ij}''	Subtraction Division Multiplication	1 1 3
6	Compute T_{ij}	Multiplication	d
7	Message to node	Transfer	d

8	Compute nowReceived	Addition Comparison Message Transfer	d d d

V. SIMULATION

To test the performance of the newly proposed load-balancing policy, a Java program is developed to test the performance of the existing and proposed algorithms. The existing algorithms ELISA and DOLB are used to compare with the proposed algorithm ALS. The DOLB is very much related to the above problem. The initial settings and parameters are shown in Table 2. The average network transfer rates between each node are represented by the cost adjacency matrix.

The proposed algorithm ALS is tested with DOLB & ELISA for the gain values *G* between 0.3 and 1 with 0.1 incremental steps. The α parameter introduced in the previous section was set to 0.05 by running several experiments and observing the behavior of the *tsi* parameter. Note that, the first time the load-balancing process was triggered after 40s from the start of the system and then the strategy executed regularly at 20s interval.

Table 2: Simulation Parameters

Number of nodes	16,32,64
Initial task distribution	[100…1000] tasks distributed randomly at each node
Average task processing time(ts in ms)	Processing time is randomly distributed in a range [300…800]
Size of task(in KB)	100
Load balancing instance	First time the load balancing was triggered at 5s then for every 10s the load balancing is initiated
Bandwidth distribution (d_{ij})	A cost adjacency matrix denotes the transfer rate between the nodes.It is uniformly distributed in the range [1..5] Mbps

The above constraint ensures that the *ts* parameter had enough time to adapt and reflect the current computational power of each node before the occurrence of any task migration between the nodes. Note that the ratio $\frac{ts_i}{ts_j}$ are fixed over time. The proposed and rival methods were evaluated by conducting 10 runs for each value of *G between* 0.3 and 1 with 0.1 incremental step.

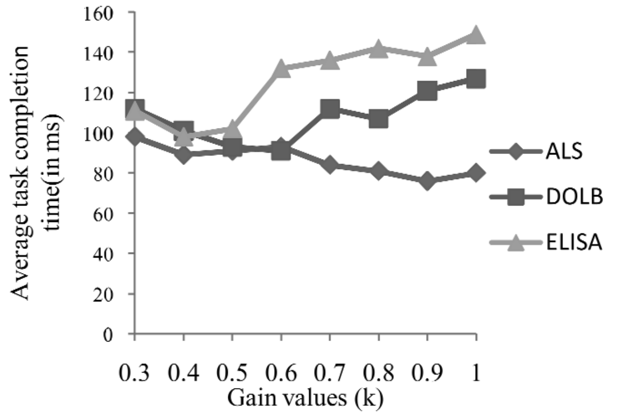


Figure 1: Completion time averaged over 5 runs vs different gain values K. The graphs shows the results of three policies for system size=64.

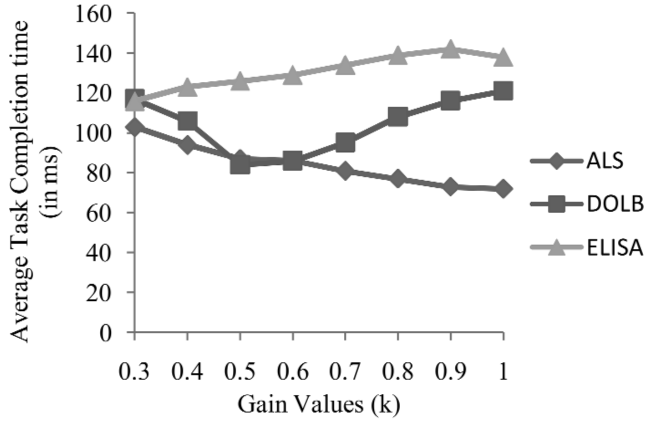


Figure 2: Completion time averaged over 5 runs vs. different gain values K. The graphs shows the results of three policies for system size=32.

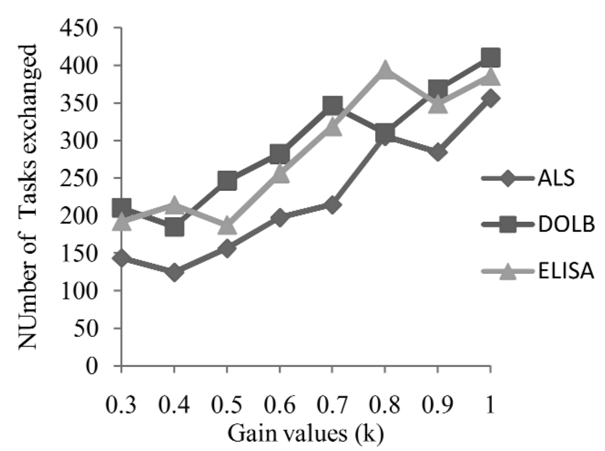


Figure 3: Total number of tasks exchanged averaged over 5 runs Vs different Gain values K. The graphs shows the performance of the three policies for system size=16.

VI. CONCLUSION

The proposed algorithm is better when compared to the existing algorithms in the literature. In simulation, we assumed the tasks with no precedence and with no deadlines. However, in heterogeneous systems, load balancing technique must take into account of OS scheduling policies like round robin, priority scheduling and to consider the deadline of the task, In this paper, these factors are not considered while designing the proposed algorithm. As a future work, these factors must be taken into account in designing a load-balancing algorithm.

P.Neelakantan received B.E(CSE) from Kuvempu University and M.E (CSE) from Madurai Kamaraj University in the year 2002.Presently he is pursuing Ph.d at SV University, Tirupati.

Dr. A. Rama Mohan Reddy received the B. Tech. from JNT University, Hyderabad in 1986, M. Tech degree in Computer Science from National Institute of Technology in 2000 Warangal and Ph. D in Computer Science and Engineering in 2008 from Sri Venkateswara University, Tirupathi, Andhra Pradesh, India. He worked as Assistant Professor, Associate Professor of Computer Science and Engineering, Sri Venkateswara University College of Engineering during the period 1992 and 2005. Presently, working as Professor of Computer Science and Engineering, Sri Venkateswara University College of Engineering. He has 28 years of Industry and Teaching experience. He is life member of ISTE and IE. Research interests include Software Architecture, Software Engineering and Data Mining. He has 10 international publications and 14 international conference Publications at International and National level.

References

[1] M. M. Hayat, S . Dhakal, C. T. Abdallah I 'Dynamic time delay models for load balancing. Pan 11: Stochastic analysis of the effect of delay uncertainty, CNRS-NSF Workshop: Advances in Control of tirne-delay Systems, Paris France, January 2003.

[2] J. Ghanem, C. T. Abdallah, M. M. Hayat, S. Dhakal, J.D Birdwell, J. Chiasson, and Z. Tang. Implementation of load balancing algorithms over a local area network and the internet. 43rd IEEE Conference on Decision and Control, Submitted, Bahamas, 2004.

[3] L. Anand, D. Ghose, and V. Mani, "ELISA: An Estimated Load Information Scheduling Algorithm for Distributed Computing Systems," Int'l J. Computers and Math With Applications, vol. 37, no. 8, pp. 57-85, Apr. 1999.

[4] WEI Wen-hong, XIANG Fei, WANG Wen-feng, et al. Load Balancing Algorithm in Structure P2P Systems[J],Computer Science, 2010, 37(4):82-85.

[5] Khalifa, A.S.; Fergany, T.A.; Ammar, R.A.; Tolba, M.F," Dynamic online Allocation of Independent tasks onto heterogeneous computing systems to maximize load balancing," IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2008,Pages:418 – 425.

[6] Andras Veres and Miklos Boda. The chaotic nature of TCP congestion control. In Proceedings of the IEEE Infocom, pages 1715-1723, 2000.

[7] J. Chiasson, J. D. Birdwell, Z. Tang, and C.T. Abdallah. The effect of time delays in the stability of load balancing algorithms for parallel computations.IEEE CDC, Maui, Hawaii, 2003.

[8] Ming wu and Xian-He sun, A General Self Adaptive Task Scheduling System for Non Dedicated Heterogeneous Computing, In Proceedings of IEEE International Conference on Cluster Computing, PP 354-361, Dec 2003.

[9] Z. Zeng and B. Veeravalli, "Design and Performance Evaluation of Queue-and-Rate-Adjustment Dynamic Load Balancing Policies for Distributed Networks", presented at IEEE Trans. Computers, 2006, pp.1410-1422.

[10] K. Lu, R. Subrata, and A. Y. Zomaya, Towards decentralized load balancing in a computational grid environment, in: Proceedings of the first International Conference on Grid and Pervasive Computing, 2006, Vol. 3947, pp. 466-477, Springer-Verlag Press.

[11] Acker, D., Kulkarni, S. 2007. A Dynamic Load Dispersion Algorithm for Load Balancing in a Heterogeneous Grid System. IEEE Sarnoff Symposium, 1- 5.

[12] M. Luczak and C. McDiarmid. On the maximum queue length in the supermarket model. The Annals of Probability, 34(2):493–527, 2006.

[13] Zhou, S. (1987). An Experimental Assessment of Resource Queue Lengths as Load Indices. Proc. Winter USENIX Conf., p.73-82.

[14] J. Ghanem, C. T. Abdallah, M. M. Hayat, S. Dhakal, J.D Birdwell, J. Chiasson, and Z. Tang. Implementation of load balancing algorithms over a local area network and the internet. 43rd IEEE Conference on Decision and Control, Submitted, Bahamas, 2004.

Energy effective coexistence of LTE-WCDMA multi-RAT systems

István Törős, *Member, IEEE,* Péter Fazekas, *Member, IEEE*

***Abstract*—As the amount of today's mobile traffic, including internet data and voice calls, highly increases, more effective technologies have to be integrated into the cellular wireless networks to serve the new demands. Actually the "green" networks conception is highly promoted, so the coexistence of radio technologies is very important in terms of energy consumption. By energy effective radio network planning procedure, this paper presents the energy consumption of multi-RAT (Radio Access Technology) structure. During analyses the traffic distribution among RATs is changed representing the user's traffic transition. The primary purpose is to examine the energy consumption in the phases of transition between telecommunication technologies demonstrating the energy efficiency of the multi-RAT systems.**

I. Introduction

The mobile telecommunication is one of the most dynamically developing services in the world. The traffic via mobile networks has exploded in the last few years, so the investments in more effective telecommunication technologies and equipments have become more important to serve the increased size of data. As the occupied bandwidth used by a telecommunication technology is limited and the data transfer conditions over this bandwidth are defined, to follow the increasing traffic the providers have to install more and more equipments in the radio access networks. The total number of mobile subscriptions in the world has passed 5 billion by the end of 2010, more than 70 % of the population of the planet. The number of worldwide base station sites is circa 5.5 million and the total global RAN (Radio Access Network) power consumption is 70 TWh, which equals to the total annual electricity consumption of the countries of Ireland and Portugal together.

The service providers and the largest mobile telecommunications equipment vendors collaborate to research more and more innovative solutions, by which the modern mobile telecommunication systems can be improved. One of the most important criteria is the energy efficiency. Taking the EARTH project for example, which aims to improve the energy efficiency of mobile communication systems, from components over protocols up to the system level. The main target is an average 50 % reduction of electricity consumption of wireless networks [1].

Numerous cellular network planning algorithms are presented in the literature [2], [3], [4], [5], [12], [13], and these can be classified into three major groups. One class uses exact algorithms as core mechanisms. Although exact algorithms are able to find optimal solution, they are often too computationally intensive and time consuming to be applied even to a relatively small data set. The other, more popular class includes the heuristic algorithms, for example simulated annealing, clustering methods, or any others. The disadvantages of these are the long running time, the hard verification as well as the chance of stopping in a local optimum. Our multi-RAT method is the member of this group. Finally the last group is the genetic algorithms, which transform the optimization problem to a simplified representation.

The radio network planning algorithms are the members of location-allocation problems. The target is to find the locations considering to be optimal depending on the pursued objectives, such as minimal transportation costs or maximal accessibility, which are reflected in the location-allocation models used.

The Facility Location Problem (FLP) is a classical question in computer science and one of the NP-complete problems. The capacitated version of FLP (CFLP) contains the capacities of subsets, which is called supplies. The energy efficient cellular network planning can be identified with facility location problem, where the supplies change dynamically taking the signal propagation and the used radio resource management into account.

$$minP_{in} = \sum_{j=1}^{k} P_0(j) + \sum_{j=1}^{k} \Delta * P_{out}(j). \qquad (1)$$

where k is the number of sectors, $P_0(j)$ is the static power consumption and $\Delta * P_{out}(j)$ is the dynamic power consumption. In the case of LTE (Long Term Evolution), the P_{out} depends on the used resources near linearly, and $P_0(j)$ is a technology specific value.

Actually the cellular wireless networks are made up of multiple access technologies. This multi-RAT topology is a heterogeneous network including the mixture of different generation standards starting with 2G, 3G and 3.5G technologies. This solution increases the capacity of system, because the different standards use different carrier frequencies avoiding the interferences between technologies. Furthermore, the multi-RAT system represent many generations of mobile technologies, so this heterogeneous wireless network is available for more subscribers. As the traffic increases the data are shared among RANs. The high demands, like internet multimedia service, are served by the highest capacity RAN. The other, low demand services are served by other technologies. The density of stations of actually highest capacity RAN increases more and more following the traffic explosion. The coexistence of multi-RAT systems is an interesting question. The daily

I. Törős is with Dept. of Networked Systems and Services, Budapest University of Technology and Economics, Magyar tudósok körútja 2., 1117 Budapest, Hungary Email: toros@hit.bme.hu

P. Fazekas is with Dept. of Networked Systems and Services, Budapest University of Technology and Economics, Magyar tudósok körútja 2., 1117 Budapest, Hungary Email: fazekasp@hit.bme.hu

energy consumption of mobile systems can be reduced by effective base station cooperation [14], [15]. The electricity consumption of an access network can be predicted. This analysis requires a cellular network planning procedure, which determines the positions of necessary stations of every RAT to serve the predefined demand generations. Under demand generation can be understood 2G,3G or 4G subscribers with traffic data.

Our work deals with the multi-RAT energy consumption mentioned above. The analyses are based on a feasible cellular network planning algorithm, which focuses on the energy efficiency. It determines the topologies of radio access networks one by one optimizing the energy consumption of multi-RAT system. The dimensioning phase of planning is not necessary, the algorithm can start with an empty environment placing and configuring the stations of the different RAT layers. When the algorithm plans a radio access network, it is assumed, that the topologies of earlier planned standards (reference system) have already known. So first the reference topology has to be determined by planning algorithm symbolizing the starting state, when only one type of telecommunication technology was installed.

Furthermore, the network planning algorithm determines an effective coexistence of the analyzed technologies. The subscriber attraction by new generation standard affects the other RATs reducing their total traffic, hence these older topologies can be changed by shutting off stations, reducing transmitter power, orientating antenna main lobes, etc..

The rest of this paper is organized as follows. In Section II the models used in this study are presented, and we describe the multi-RAT planning and transmitter power reduction methods, which are used in the analyses. In Section III the results of algorithms are provided, and the conclusion is given in Section IV.

II. System model and used algorithms

This section introduces the system model and the submethods of planning algorithms used for investigation of energy effective coexistence of LTE-WCDMA multi-RAT systems. The examined scenario can be simply described by the set of applicable coordinates over the area and the given traffic amount per generations of technologies (GSM-Global System for Mobile Communications,WCDMA-Wideband Code Division Multiple Access,LTE) over the area, assigned to any subset of the coordinates on the terrain. We suppose that the amount of traffic demands is given by a set of discrete coordinates (denoted as Demand Positions, DPs), along with the amount of traffic generated at that position. This approach is flexible to describe any kind of traffic distribution (continuous, if every point of the area is a DP, discrete service areas if there are much smaller number of DPs). The set of DPs is denoted by:

$$\mathcal{DP}^s = \{\cup_{i=0}^{m} DP_i^s\}; \tag{2}$$

where m denotes the number of DP_i^ss in the traffic environment of s_{th} demand generation. These points are represented by (x_i,y_i,dem_i), where x_i, y_i are the coordinates and dem_i is the traffic demand of DP_i^s, expressed in kbps. DP is an input parameter.

We assume that a base station (BS) operates three cells through three sectorized antennas. The resources are given to the radio access networks by these equipments to serve the users. Some equipment can be shared by different access networks to reduce the installation and energy consumption costs.

The stations are represented by

$$\mathcal{BS}^s = \{\cup_{j=0}^{t} BS_j^s\}; \quad \mathcal{BS} = \{\cup_{s=0}^{n} BS^s\} \tag{3}$$

where t is the number of BS_j^ss in the traffic environment of s_{th} demand generation.

We suppose that base stations cannot be placed arbitrarily, but to given possible (e.g. in an urban environment to rooftops) candidate positions (CP):

$$\mathcal{CP}^s = \{\cup_{j=0}^{r} CP_j^s\}; \quad \mathcal{CP} = \{\cup_{s=0}^{n} CP^s\} \tag{4}$$

where r is the number of CP_j^ss in the traffic environment of s_{th} demand generation.

The stations of other RATs (GSM,WCDMA...) were placed also to any candidate positions.

$$\mathcal{CPE} \subseteq \mathcal{CP}; \tag{5}$$

where CPE denotes the candidate positions of the earlier placed stations (reference topology).

We use COST 231 Okumura-Hata path loss model for big city environment in our simulations. This has the advantage that it can be implemented easily without expensive geographical database, yet it is accurate enough, captures major properties of propagation and used widely in cellular network planning. A sector is defined as the set of DPs that are covered by a given transmitter. The "best server" policy is followed within the network, namely a demand is served by the sector whose signal strength is the highest in the position of DP_i [6].

The resources of network can be managed by frequency adaptation and power management. Our planning procedure uses the properties of 3GPP LTE radio resource management (RRM). The relationship between SINR (Signal to Interference plus Noise Ratio) and spectral efficiency is given by the so called Alpha-Shannon Formula which is suggested to be used for LTE networks in [7].

The RRM of LTE is modelled in our case by a semi dynamic frequency allocation strategy. It is the so called C/I scheduler. The sectors allocate Physical Resource Blocks (PRBs) to the demands in the order of decreasing SINRs. The frequency allocation simultaneously deals the PRBs one by one in every sector. Note that the amount of traffic a PRB can carry is determined from the SINR by the alpha-Shannon formula. If a sector is ready (serves all DP^s sets) then it won't transmit on the remaining PRBs (hence the SINR on these PRBs will be better for the neighbours). This method is very fast and reasonably high SINR values can be achieved by cell borders as well. It has to be emphasized, that any RRM algorithm can be supposed for our planning mechanism, RRM function is actually an input to the planning (and thus affects final results). In practice LTE base stations are transmitting with constant power spectral density (regardless the number of PRBs actually used), hence using less PRBs require proportionally less transmit power, as described below in (6).

The transmitter output power, P_{out} can be described by

$$P_{out} = \frac{usedPRB}{allPRB} * P_{max} \tag{6}$$

where P_{max} is the maximum top of cabinet output power of transmitter, $usedPRB$ and $allPRB$ are the number of actually used PRBs and all PRBs respectively. This latter depends on the configured bandwidth of the system, that is also a parameter of the deployment method. Namely, as a PRB is a 180 kHz wide chunk of the channel, in a 1 ms subframe, e.g. a 20 MHz bandwidth configuration typically means 100 PRBs in every 1 ms subframe.

In practice LTE base stations are transmitting with constant power spectral density (regardless the number of PRBs actually used), hence using less PRBs require proportionally less transmit power. Furthermore, it is assumed, that the P_{out} depends on the allocated resources also linearly in the cases of the other standards (GSM,WCDMA).

The power consumption of the base station follows the linear model:

$$P_{Cons} = P_0 + \Delta * P_{out} \tag{7}$$

where the first part (P_0) describes the static power consumption. Depending on the load situation, a dynamic power consumption ($\Delta * P_{out}$) part adds to the static power. The factor Δ is mainly due to the power amplifier inefficiency and feeder loss.

A. Base station placement and multi-RAT planning methods

This subsection deals with the base station placement and multi-RAT planning methods. To analyze the energy consumption of multi-RAT system, first the network topology has to be planned. These methods determine the quasi optimal station position and configuration for every RAN and reduce the number of applied equipments. The explaining of these algorithms are necessary to understand the numerical results.

The base station placement method can be configured for given coverage (in terms of percentage of the area covered by at least a minimum signal strength) and service (in terms of percentage of total traffic requirements served) criteria. The default is 100% for both. The input parameters are the used bandwidth, maximum transmit power parameters of transmitters as well as the DP scenario of every demand generation. The geographical area is fixed. The output data are the base station topology (BS) [9].

1) Base Station Placement Algorithm (BSPA): This algorithm determines a base station topology, which guarantees the serving and coverage criteria on the given demand scenario.

The BSPA is based on K-means dynamic clustering method. The clusters are the sites of stations including the covered subscribers. The criterion function of K-means, which has to be minimized, is the sum of squared Euclidean distances between the locations of demands and the position of serving base station.

Algorithm 1: K-means

Input: K is the number of centroids/clusters.
M is the number of objects.
$O = \{\cup_{i=0}^{M} o_i(i,x,y)\}$ is the set of position of objects.
I is the number of iterations.
map is the $max_x * max_y$ scenario.
Output: C is the set of positions of centroids after clusterization.

```
begin
    Initialization Step:
    C = {∪_{j=0}^{K} c_j^0}
    ∀c_j^0.x ← random(max_x)
    ∀c_j^0.y ← random(max_y)
    S = {∪_{k=0}^{K} S_k}
    ∀S_k ← ∅
    Iteration Step:
    for t ← 0 to I do
        Reassignment Step:
        ∀S_k ← ∅
        for i ← 0 to M do
            min ← ∞
            id ← −1
            for j ← 0 to K do
                if distance(o_i, c_j) < min then
                    min ← distance(o_i, c_j)
                    id ← j
                end
            end
            o_i joins to the S_id
        end
        Update Step:
        for k ← 0 to K do
            c_k^(t+1) ← 1/#S_k^t ∑_{i⊂S_k^t} o_i
        end
    end
end
```

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. It is a dynamic clustering method which attempts to directly decompose the data into disjoint clusters. The number of clusters (K) is fixed a priori. The different located centroids of clusters cause different results, so the algorithm has to be started with different initial states and run as much as possible.

Briefly overview the K-means, it can be composed of the following steps:

1. Place K (parameter) points into the space represented by the objects that are being clustered.

2. Assign each object to the group (cluster) that has the closest centroid. (Reassignment step)

3. When all objects have been assigned, recalculate the properties of the K centroids. (Update step)

4. Repeat Steps 2 and 3 until the centroids no longer move or the counter of iteration expire.

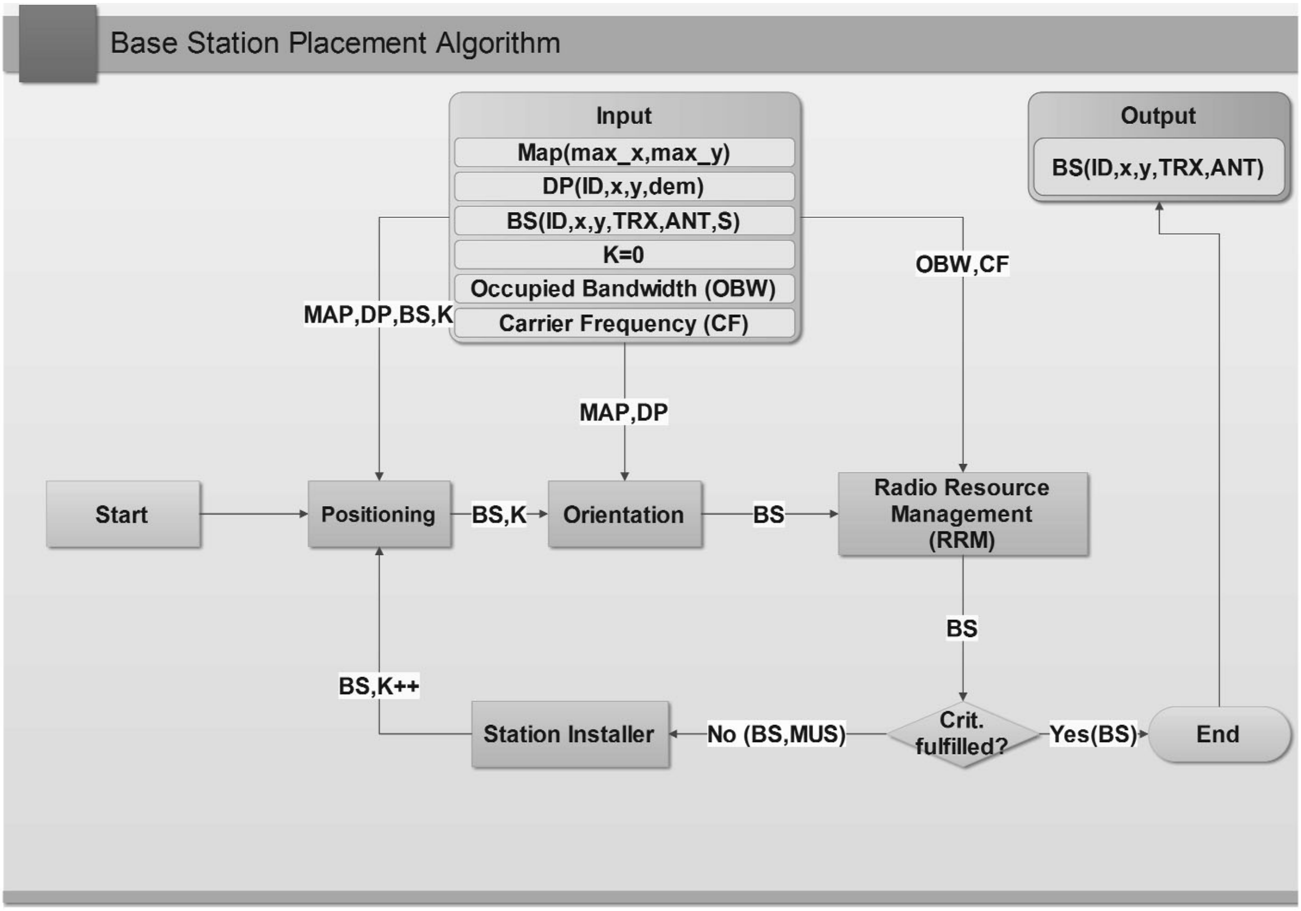


Fig. 1. State chart diagram of Base Station Placement Algorithm

The objective function is the total energy consumption of access network.

$$min\ P(\mathcal{BS}) = min \sum_{j=1}^{k}(P_0(j) + \Delta * P_{out}(j)). \qquad (8)$$

The objective function has four changeable parameters to reduce the total power consumption. Δ and P_0 depend on the type of BSs, so these parameters are independent from BSPA, because the algorithm places only one type of stations. The k is the number of sites in the wireless network. As it is pointed out in the related work section, the minimization of installed stations (sites) is the first priority target. $P_{out}(j)$ is the output power of j_{th} site. This parameter is the function of allocated resources depending on the network topology. So the k and $P_{out}(j)$ parameters can be reduced by BSPA.

The BSPA, including station positioning, antenna beam orientation, RRM and station installer, can be realized as a closed loop (Figure 2). Starting with an empty environment (K is 0), it places the stations (K=1,2,3,4...) iteratively until the mentioned criteria are fulfilled.

Positioning:

The BS positioning algorithm is based on the mentioned K-means procedure. The centroids of clusters are the BSs and the assignment step is the procedure of sector creation. One cluster is made up of three sectors of BSs. In the update step the position of covered DP (x_i) is weighted by the demands of DP (dem_i) determining the positions of stations. So the modified objective function of K-means is

$$min\ Z = \sum_{j=1}^{k} \sum_{DP_i \in S_j} dem_i * ||DP_i^{(j)} - BS_j||^2 \qquad (9)$$

where dem_i is the demands of i_{th} subscriber, and $||DP_i^{(j)} - BS_j||$ is the Euclidean distance between subscriber (Demands positions) and the serving station. $S_j = \cup_{h=1}^{N} S_{j,h}$, where N is the number of sectors per BS [9].

Orientation:

The antenna beam orientation is also based on K-means clustering. Our aim that the directions of covered DPs with higher demand are subtended smaller angle with the main direction of serving antenna. The assignment step is also the procedure of sector creation. In the update step, x_i is the included angle between the direction of covered DP_i within the sector and the main direction of serving transmitter weighted by the dem_i. This mechanism determines the beam directions of antennas[9].

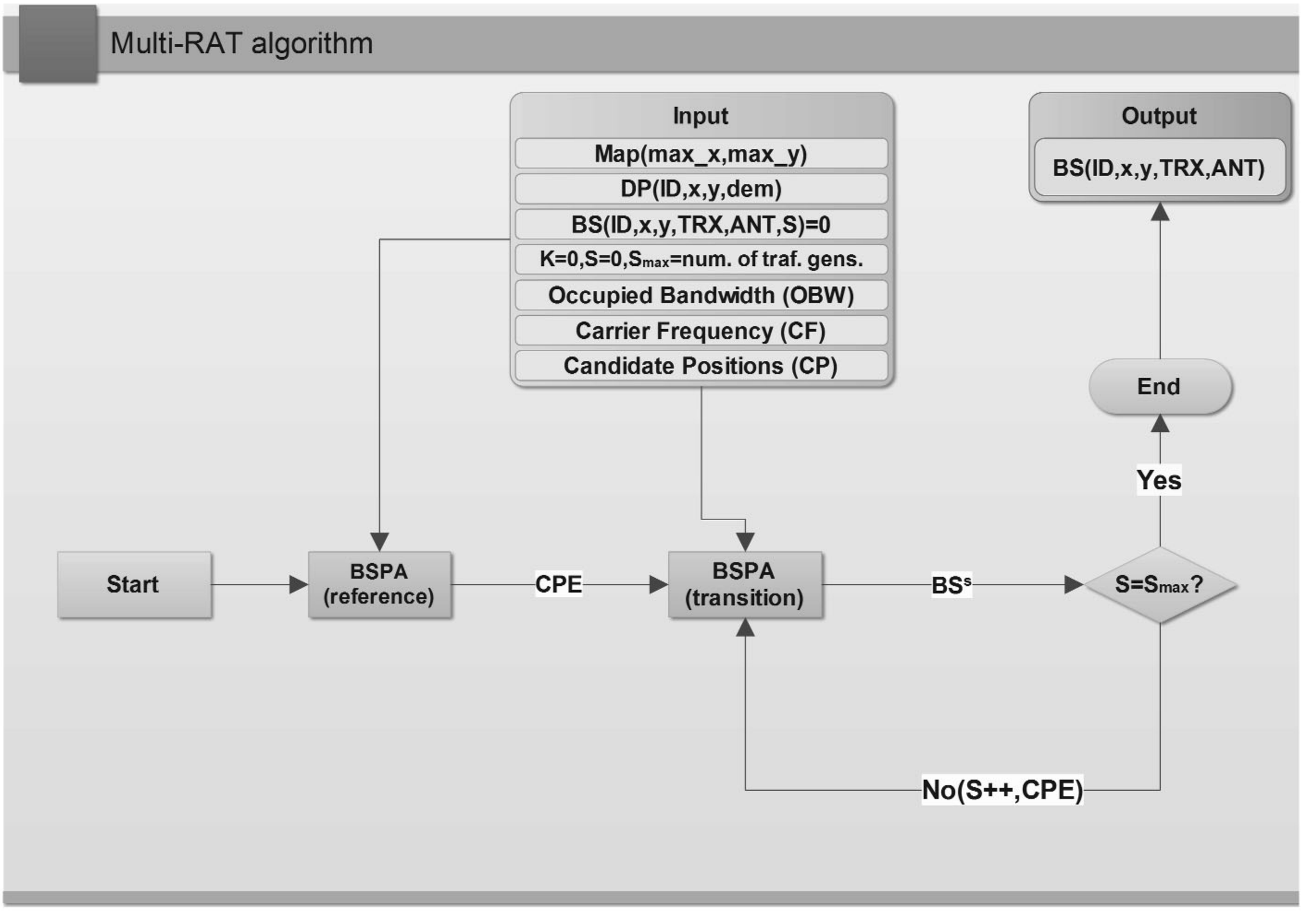


Fig. 2. State chart diagram of Multi-RAT planning

Radio Resource Management:

The radio resource management is described in model section as a parameter. In our investigations, the RRM is a max C/I scheduler. It is executed after base station positioning and antenna beam orientation to analyze the loads of sectors. The radio resource management is an input parameter of BSPA. The target of this method is to determine the required/used number of PRBs per sector and to give these informations to the station installer as results.

Station Installer:

After RRM the most unserved sector (MUS) has to be found, which is the sector with the highest total unserved traffic (DPs with not enough PRBs allocated to) under its coverage. If the number of required PRBs is less than the number of available PRBs within all sectors then there is no MUS and the algorithm stops. Otherwise the algorithm locates a new base station near the serving antenna of MUS in the main direction and runs the positioning, rotation and RRM mechanisms again. So our clustering algorithm is an increasing number of K-means (X-mean) [11].

Figure 3 shows the mentioned station movement, as the algorithm runs iteratively. Actually the black sector is the MUS, so the station installer places the new station near the serving antenna of this.

The complexities of BSPA is $O(IK^2NM)$, where I is the fix number of iterations, N is the number of DPs, K is the number of BSs in the final state, and $O(M)$ is the operation cost (signal propagation).

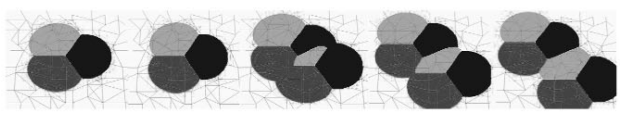


Fig. 3. Station movement within placement algorithm

2) Multi-RAT planning method: This method uses the BSPA to plan an energy effective multi-RAT topology. First it plans a reference network topology using an older telecommunication standard (WCDMA,GSM). As the subscribers are attracted by new standard, the stations of reference RAN can be shut off, because the reduced overall demands can be served by fewer capacities. Furthermore, the high demands, like internet multimedia service, connect with the highest capacity RAN (LTE).

The reference system contains the stations of older topology determining the candidate positions of transition phases (CPE). The second procedure is the planning of transition cases. The traffic scenarios contains the demand generations ($DP_i{}^s$) starting with the reduced number of subscribers of older generation and ending with the new generation demands.

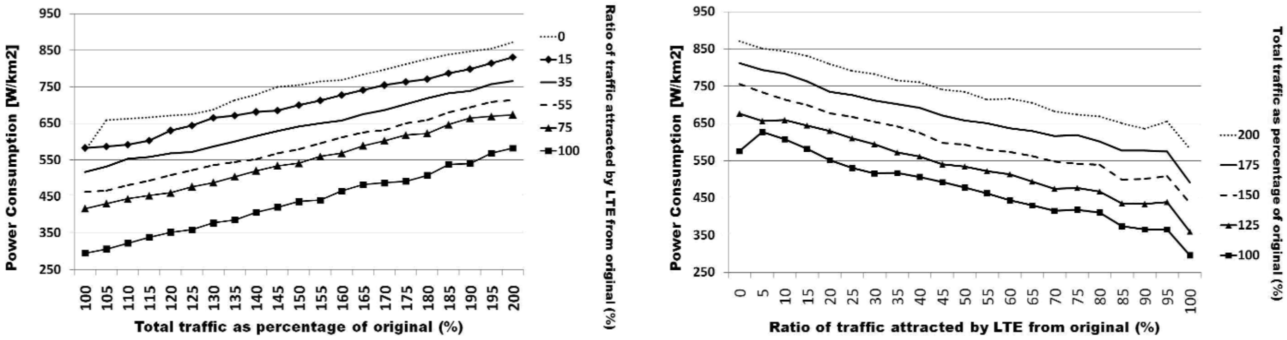


Fig. 4. The power consumption of multi-RAT systems as a function of the increase in the traffic demands and demands migration between standards (WCDMA,LTE).

The BSPA installs the stations only to the reference candidate positions (CPE), which denote base stations of reference topology, so the new multi-RAT structure reuses the elements of older topology. If a CPE is empty on every scenario, then the station can be removed. If the number of CPE is not enough in the case of new generation demands, then the set of CPE need to be complemented with the rest of candidate positions ($CP \setminus CPE$).

III. Numerical results

The analyses discussed below use the multi-RAT planning algorithm. The geographical topology is constant, the sizes of total demands and the ratio of traffic attracted by LTE from WCDMA (reference) are changed illustrating the phases of transition between telecommunication technologies. The multi-RAT planning algorithm gets the WCDMA and LTE demand scenarios as input parameters and gives back a WCDMA-LTE multi-RAT topology. Table III shows the main input parameters of algorithm derived from [16].

Input parameters

Carrier frequency	2 GHz
Occupied bandwidth of WCDMA	5 MHz
Occupied bandwidth of LTE	10 MHz
Frequency reuse factor	1
Static power of stations	300 W
Max top of cabinet output power of tx	30 W
Inefficiency of power amplifier	3
Size of environments	$9km^2$
Default traffic	85 Mbps/$9km^2$

The analyses were run with same parameters on the studied scenario and the results were averaged.

Figure 4 shows total power consumption of multi-RAT networks as a function of size of traffic demands (left) and a function of the ratio of traffic attracted by LTE from WCDMA (right). The new demands always connect with the LTE system. The different lines of the figures represent the horizontal axis of other one, and vice versa. The curves can not intersect each other, because more data traffic requires more stations increasing the power consumption of system. The reasons of high steps (left figure dotted line 100 % and right figure at the end of lines) are caused by the establishment of new technology and the complete removing of the other one. In the establishment phase the service providers have to place many new transmitters to guarantee the coverage criterion of new telecommunication technology. In the complete removing phase the transmitters of WCDMA system can be switched off totally, reducing the energy consumption. These simulation results show that the LTE system is more effective than the WCDMA (wider bandwidth) one, so the service providers can save the budget of energy consumption if the users change over from 3G to 4G.

IV. Conclusion

In this paper we examined the energy consumptions of multi-RAT network topologies focussing on WCDMA-LTE coexistence. In the analyzed cases it was assumed, that the future demands would connect with the new LTE network, furthermore, some percents of 3G users would change technology. The results showed that the energy consumption of cellular system could be reduced by LTE technology. Assuming same overall demands, the energy efficiency of network increased as the LTE gains ground.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement no INFSO-ICT-247733 EARTH Project. The support of the Hungarian Government through the TMOP-4.2.1/B-09/1/KMR-2010-0002 project at the Budapest University of Technology and Economics is also acknowledged.

References

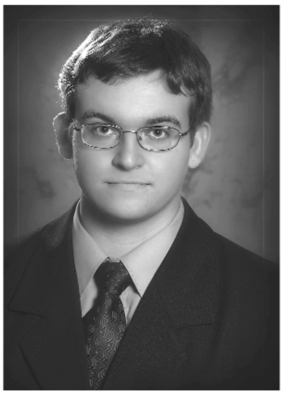
[1] Gergely Biczók, Jens Malmodin, Albrecht Fehske, "INFSO-ICT-247733 EARTH Deliverable D2.1", Economic and Ecological Impact of ICT (2011)

[2] Omar H. Karam, Lamia Fattouh, Nourhan Youssef, Ahmad E. Abdelazim, "Employing Clustering Techniques in Planning Wireless Local Loop Communication Systems: PlanAir",11th International Conference On Artificial Intelligence Applications Cairo, Egypt, February 23-26, (2005)

[3] Harish Ramamurthy, Abhay Karandikar: "B-Hive: A cell planning tool for urban wireless networks", 9th National Conference on Communications, (2003)

[4] Kurt Tutschku, "Demand-based Radio Network Planning of Cellular Mobile Communication Systems", INFOCOM, pp. 1054–1061 (1998)
[5] Hurley, S., "Planning effective cellular mobile radio networks." IEEE Trans. Vehicular Technol. v51 i2. 243-253. (2002)
[6] Les Barclay, "Propagation of Radiowaves, p. 194, The Institution of Electrical Engineers", London (2003)
[7] Abdul Basit, "Dimensioning of LTE Network, Description of Models and Tool, Coverage and Capacity Estimation of 3GPP Long Term Evolution radio interface" (2009)
[8] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks" in Proc. of 19th Future Network & MobileSummit 2010, Florence, Italy, (June 2010)
[9] István Törős, Péter Fazekas, "Automatic Base Station Deployment Algorithm in Next Generation Cellular Networks", Accessnet 2010 Budapest (2010)
[10] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, 1:281-297 (1967)
[11] Dan Pelleg, Andrew Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters", Proceedings of the 17th International Conf. on Machine Learning 2000
[12] Gonzlez-Brevis P., Gondzio J., Fan Y., Poor H.V., Thompson J.S., Krikidis I., Chung P., "Base Station Location Optimization for Minimal Energy Consumption in Wireless Networks.", In VTC Spring(2011)1-5
[13] Z. Zheng, S. He, L. X. Cai, X. Shen, "Constrained Green Base Station Deployment with Resource Allocation in Wireless Networks", Handbook on Green Information and Communication Systems, Editors M. S. Obaidat, A. Anpalagan, and I. Woungang, John Wiley & Sons, Inc., 2012.
[14] F. Han, Z. Safar, W.S. Lin, Y. Chen, and K.J.R. Liu, "Energy-efficient cellular network operation via base station cooperation", ;in Proc. ICC, 2012, pp.4374-4378.
[15] István Törős, Péter Fazekas, "Planning and network management for energy efficiency in wireless systems", In Future Network & Mobile Summit (FutureNetw), 2011
[16] Gunther Auer (DOCOMO), Oliver Blume (ALUD), Vito Giannini (IMEC), Istvan Godor (ETH), Muhammad Ali Imran (UNIS), Ylva Jading (EAB), Efstathios Katranaras (UNIS), Magnus Olsson (EAB), Dario Sabella (TI), Per Skillermark (EAB), Wieslawa Wajda (ALUD), "INFSO-ICT-247733 EARTH Deliverable D2.3", Energy efficiency analysis of the reference systems, areas of improvements and target breakdown (2012)

Author Biographies



István Törős was born in Pecs, Hungary in 1985. He received his Ing. (MSc.) degree in 2009 at the Department of Telecommunications, Budapest University of Technology and Economics. His recent research interests are wireless network planning based on energy consumption optimization and repeaters in telecommunication systems.

Péter Fazekas received an MSc degree in electrical engineering from the Technical University of Budapest (now Budapest University of Technology and Economics) in 1998. Currently he is with the Department of Networked Systems and Services where he received his PhD in 2013. His research area includes the performance analysis of cellular networks, mobility modeling, and packet scheduling disciplines in wireless environment.

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on
http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2,
"Template and Instructions on How to Create Your Paper".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *index terms (Keywords)*. For the final version of accepted papers, please send the *short cvs* and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984.

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

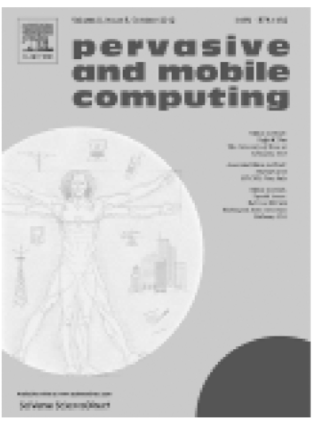
Authors are requested to send their manuscripts via electronic mail or on an electronic medium such as a CD by mail to the Editor-in-Chief:

Csaba A. Szabo
Department of Networked Systems and Services
Budapest University of Technology and Economics
2 Magyar Tudosok krt.
Budapest, 1117 Hungary
szabo@hit.bme.hu

IEEE PerCom in Budapest in 2014!

The **IEEE Pervasive Computing and Communication** (PerCom) conference is the worldwide premier scholarly venue in the areas of pervasive computing and communications. Since 2003, the conference has grown significantly in terms of quality and variety of the technical programs – it is recognized as a top tier conference by most universities and organizations across the world.

PerCom provides a high profile, leading edge forum for researchers, engineers, and practitioners to present state-of-the-art research in the respective fields of pervasive computing and communications. The conference features a diverse mixture of presentation forums including core technical sessions, keynote talks, panel discussions from worldwide experts, demonstrations, a PhD forum, and work in progress posters. The conference also hosts a number of workshops that have themselves become well recognized in the community as forums for specialized topics within the field.



The conference also provides formats to honor excellence in the field. The Mark Weiser Best Paper award, sponsored by Elsevier, is given to authors of the PerCom's best paper. In addition, the highest quality papers from the conference are published in a special issue of the Pervasive and Mobile Computing Journal.

The IEEE PerCom Steering Committee has recently decided to accept the joint application of the Budapest University of Technology and Economics (BME), Department of Telecommunications and the Scientific Association for Infocommunications (HTE) to organize PerCom 2014 in Hungary. Thus, after the 2013 edition in San Diego, Budapest will host this prestigious conference in 2014. It will be a 5 day-event, with associated workshops, and will hopefully attract several hundreds of participants. More information can be found on the conference's website:

www.percom.org

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its more than 1300 individual members, the Scientific Association for Infocommunications (in Hungarian: Hírközlési és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society. HTE is corporate member of International Telecommunications Society (ITS).

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **DR. GÁBOR HUSZTY** • *ghuszty@entel.hu*

Secretary-General: **DR. ISTVÁN BARTOLITS** • *bartolits@nmhh.hu*

Managing Director, Deputy Secretary-General: **PÉTER NAGY** • *nagy.peter@hte.hu*

International Affairs: **ROLLAND VIDA, PhD** • *vida@tmit.bme.hu*

Addresses

Office: H-1055 Budapest, V. Kossuth Lajos square 6-8, Room: 422.

Mail Address: 1372 Budapest, Pf. 451., Hungary

Phone: +36 1 353 1027, Fax: +36 1 353 0451

E-mail: *info@hte.hu*, Web: *www.hte.hu*