SECURITY PROTOCOLS FOR 2G AND 3G WIRELESS COMMUNICATIONS

T. Newe & T. Coffey
Data Communications Security Group,
Department of Electronic & Computer Engineering,
University of Limerick, Ireland.
E-Mail: (Thomas.Newe@UL.ie) (Tom.Coffey@UL.ie)
Fax: 353-61-338176

ABSTRACT

Wireless communications are being driven by the need for providing network access to mobile or nomadic computing devices. The need for wireless access to a network is evident in current work environments. A number of new protocols have been recently published with the goal of providing both privacy of data and authentication of users for mobile systems. Such protocols can employ private-key and/or public key cryptographic algorithms. Public-key algorithms hold the promise of simplifying the network infrastructure required to provide security services such as: privacy, authentication and non-repudiation, while symmetric algorithms require less processing power on the mobile device.

In this paper a selection of protocols are reviewed and they are broadly divided into two categories: second generation and third generation protocols. A summary of the capabilities and services provided by each protocol is then provided.

Keywords: 2G, 3G, mobile, end-to-end, radio-link, security.

1. INTRODUCTION

In addition to the security requirements for fixed network such as; identity authentication, data confidentiality and non-repudiation, mobile systems have a number of additional requirements due to the nature of the mobile environment. These include:

- Communications path: The communications path contains many parts, one of which, the radio link is particularly vulnerable to attack.
- Location Privacy: Since mobiles roam freely it may be advantageous to keep its location a secret.
- Computational constraints: Most mobile devices are computationally and/or power limited, therefore it is necessary to limit the computational complexity of any security algorithm used.

It is also necessary to consider what type of security service to supply, is it end-to-end or radio link security? Radio link security is generally used with second generation mobile systems where the main security concern is protection of data (generally voice data) over the radio link. Third generation systems can involve the use of the Internet (WAP enabled devices) and services such as: non-repudiation, data confidentiality and end-to-end security can be required. End-to-end security protocols provides for integrity over the entire link, both mobile and fixed.

A generic set of security requirements for protocols in mobile devices, against which the protocols in this paper are evaluated, is as follows:

1. Mutual Authentication of user and network
2. Exchange of certified public keys
3. Session key agreement
4. Joint control of session key
5. Mutual implicit key authentication
6. Mutual key confirmation
7. Mutual assurance of key freshness
8. Confidentiality
9. Initialisation of payment mechanism
10. Non-repudiation of origin

2. 2G MOBILE PROTOCOLS

Second generation protocols generally rely on shared long-term keys between mobile devices and home network servers in order to allow the generation of session keys. This means that once a shared session key is established between two mobiles across a network, symmetric encryption algorithms can protect digitally encoded speech and control information. Symmetric key algorithms are generally used due to the computational limitations of second-generation mobile devices.

The following notation is used to describe the security protocols in this section:
B - Base Station, *MD* - Mobile device, PK_X - Public key of entity *X*. PK_X^{-1} - Secret key of entity *X*. *Cert(X)* -Certificate of entity *X* obtained from a central authority. $\{X\}_K$ - Denotes encryption of *X* with a key *K*, where *K* can be either a symmetric key or a public key depending on the encryption algorithm used. In this case the algorithm is a public key algorithm.

2.1 Beller-Chang-Yacobi and Carlsen's BCY protocols

The original BCY [3] protocol was designed to provide session key agreement and radio link security services. This protocol uses a combination of asymmetric and symmetric cryptographic algorithms. It combines a Diffie-Hellman (DH) key exchange with the modular square root (MSR) encryption technique. The MSR technique reduces the computational burden placed on the mobile device, however the computational requirements of the base station remain high. This computational requirement limits the usefulness of the protocol for anything other than link security.

1. B → MD : $B, N_B, PK_B, Cert(B)$
2. MD → B : $\{x\}_{PK_B}$
3. MD → B : $\{N_B, MD, PK_{MD}, Cert(M)\}_x$
4. MD ↔ B : $\{known\ messages\}_{sk}$

Figure 1 Original BCY Protocol

The original BCY protocol had some shortcomings with regards to hiding the identity of the initiating mobile and the identity of the responding mobile station. Carlsen identified these shortcomings and suggested improvements, see figure 2.

1. B → MD : $B, N_B, PK_B, Cert(B)$
2. MD → B : $\{x\}_{PK_B}, \{N_B, MD, PK_{MD}, Cert(M)\}_{PK_B}$
3. MD ↔ B : $\{known\ messages\}_{sk}$

Figure 2 Carlsen's BCY Protocol

Both the base *B*, and the mobile device *MD*, generate a shared secret η using the DH technique and each others public key. The session key *sk* is then generated by encrypting *x* with η using a symmetric algorithm. To complete the protocol *MD* and *B* exchange a pre-agreed set of messages encrypted under the session key. If these messages decrypt properly then the two parties are authenticated.

2.2 Beller-Yacobi and Boyd-Mathuria BY protocols

The Beller and Yacobi (BY)[6] protocol was designed to provide low-cost, two-way public-key authentication and key agreement. The protocol also supplies location/identity privacy. The computational load on the mobile side is two orders of magnitude lower than that required by the RSA [10] public-key algorithm. However, a full computational load is required on the base unit side, making it suitable for link security only. This protocol is only suitable for link security given its computational requirements on the base side.

1.	B → MD :	B, PK_B, $Cert(B)$
2.	MD → B :	$\{x\}_{PK_B}$
3.	B → MD :	$\{N_B\}_x$
4.	MD → B :	$\{MD, PK_{MD}, Cert(MD), \{N_B\}_{PK_{MD}^{-1}}\}_x$

Figure 3 Beller-Yacobi Protocol

In [5] Boyd *et al.* presented a potential attack on the BY protocol. This attack made some strong assumptions but it indicated a flaw in the protocol design which left the protocol susceptible to the *man in the middle* attack after step 3. The solution to this attack, as presented by Boyd *et al.*, is to sign the session key in step 2. This essentially reduces the protocol to a three step protocol as shown in figure 4.

1.	B → MD :	B, PK_B, $Cert(B)$, N_B
2.	MD → B :	$\{x\}_{PK_B}$, $\{MD, PK_{MD}, Cert(MD)\}_x$, $\{h(B, MD, N_B, x)\}_{PK_{MD}^{-1}}$
3.	B → MD :	$\{N_B\}_x$

Figure 4 Boyd-Mathuria BY Protocol

In step 1 *B* now also sends the challenge N_B to *MD*. In step 2 *MD* now sends the encrypted session key *x* to *B* together with the signed challenge N_B to guarantee the freshness of *x*. The one-way hash function *h* is used to protect the confidentiality of the session key *x*. *x* is now authenticated and *B* verifies its arrival in step 3.

2.3 Summary

In second-generation systems the mobile's home domain server needs to be on-line at the time of the call setup in order to help establish a session key between mobiles. This means that the home server needs to be ultra reliable. A solution to this problem is to use public key algorithms between mobiles. However, this solution was not feasible in second generation systems due to the computational limitations of mobile devices. Public key protocols can be more readily used in 3G systems due to their improved computational abilities.

3. 3G MOBILE PROTOCOLS

Third generation systems are more widely distributed than 2G and have additional security requirements. The services required by 3G systems include the ten requirements outlined in section 1.These were identified as part of the European ASPeCT project [7].

The following notation is used to describe the protocols in this section:

h_1, h_2, h_3	Three hash functions (Can all be based on the same function).
$Sig_U\{X\}$	U's signature transformation on message *X*.
UCert	A's certificate which contains A's public signature verification information.
VCert	Certified public key of the VASP, (g^v).
chd	Charging data.
pay	Payment data.
T^V	Timestamp issued by the VASP *V*.

K_{UV}	Secret session key between Mobile *U* and the VASP *V*.
id_{CA}	Identity of central authority whose signatures the mobile can verify.
g^x	Generation of public key using Elliptic curve generator *g* and secret number *x*.
COUNT	Used to detect cloning fraud in mobile handsets.
$Enc_V\{X\}$	Denotes encryption of the field *X* using the public key of *V*.

3.1 ASPeCT (Variant B) Protocol

The ASPeCT protocol was developed by the ACTS collaborative research project ASPeCT [7]. The ASPeCT Authentication and Initialisation of Payment (AIP) protocol was developed for authentication between a mobile user and a valued added service provider (VASP) in Universal Mobile Telecommunications System (UMTS) environments. One of the basic properties of the ASPeCT protocol is the establishment of a secret session key, which can be used to encrypt subsequent communications between two entities. The ASPeCT protocol is also responsible for the secure transfer of payment and charging information between the mobile and the VASP. Two basic models were designed, classed as B and C variants. The main difference is that in the C variant an on-line Trusted Third Party (TTP) exists which serves to certify the identity of the mobile. Variant B is discussed below.

With this variant it is assumed that the mobile is in possession of a valid certificate on the VASP's public key and that the VASP has a valid certificate on the public key of the mobile.

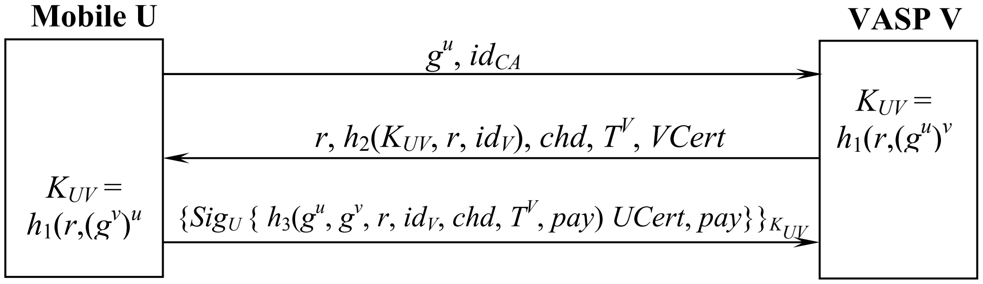
Mobile U
VASP V
g^u, id_{CA}
$K_{UV} = h_1(r,(g^u)^v$
r, $h_2(K_{UV}, r, id_V)$, chd, T^V, $VCert$
$K_{UV} = h_1(r,(g^v)^u$
$\{Sig_U\{\ h_3(g^u, g^v, r, id_V, chd, T^V, pay)\ UCert, pay\}\}_{K_{UV}}$

Figure 5 Variant B of the ASPeCT Protocol

At the start of the protocol *U* generates a random number (temporary secret key) *u* and then computes g^u which is sent together with the identity of the authority whose certificates *U* can verify, id_{CA}. On receipt of the first message the VASP *V* does not know with whom he is communicating. *V* generates a random number *r* and computes $(g^u)^v$, where *v* is *V's* private key, and then a session key $K_{UV} = h_1(r,(g^u)^v)$ is computed. *V* then sends *U* the random number *r*, the hash value $h_2(K_{UV}, r, id_V)$, and its certificate *VCert*, together with the charging data *chd* and a timestamp T^V.

On receipt of the second message the mobile *U* computes the key $K_{UV} = h_1(r,(g^v)^u)$ and compares the hashed value $h_2(K_{UV}, r, id_V)$ against the one received from the VASP *V*. If the check is successful then *U* knows that *V* generated the session key and *U* replies with the signature $\{Sig_U\ \{\ h_3(g^u, g^v, r, id_V, chd, T^V, pay)\ UCert, pay\}\}_{KUV}$ encrypted with the secret session key and a symmetric algorithm. *V* now deciphers the message using K_{UV} and retrieves the certificate *UCert*. After this certificate is verified *V* can verify *U's* signature using the public key of *U*. *V* then stores the signature and the corresponding message for later use in the payment scheme.

Arguably the most obvious weakness in the ASPeCT protocol is the delay in the identification of the mobile unit until the end of the last step. This leads to a possible attack described by Boyd-Park in [8]. The authors published a new authentication and key establishment protocol to correct this weakness.

3.2 Boyd-Park and NC Boyd-Park Key Agreement Protocols

The Boyd-Park protocol [8] was published in an attempt to correct the mobile identification delay weakness highlighted during an analysis of the ASPeCT protocol.

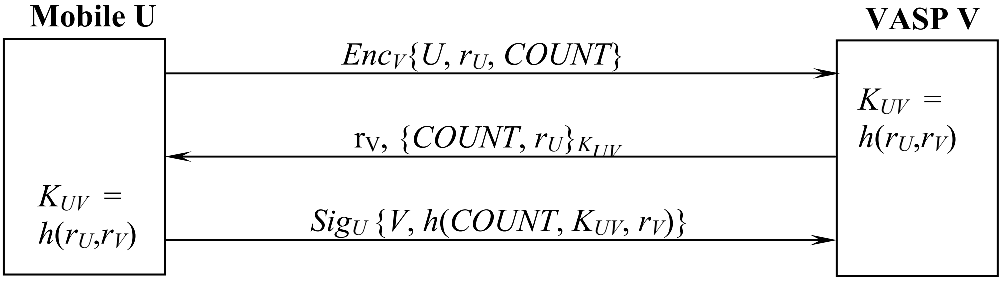

Figure 6 Boyd-Park Key Agreement Protocol

In step 1 of the protocol, the mobile *U* sends a message to the VASP *V* encrypted under the public key of *V* containing: a random nonce r_U, the mobile's identity *U* and the *COUNT* field. In step 2, V calculates the secret session key K_{UV} using a hash of random nonce's r_U and r_V. Then $r_{V,}$ and an encrypted message (using the session key K_{UV}) containing *COUNT* and r_U is sent to *U*. In step 3, *U* calculates the session key K_{UV} and uses it to reply with a signed message to *V*. From this it can be seen that key agreement between *U* and *V* has taken place prior to completion of authentication of the mobile *U* by *V*.

A modified protocol presented in [9] corrects this and enables the authentication of *U* by *V* prior to key agreement. This protocol is outlined in Figure 7.

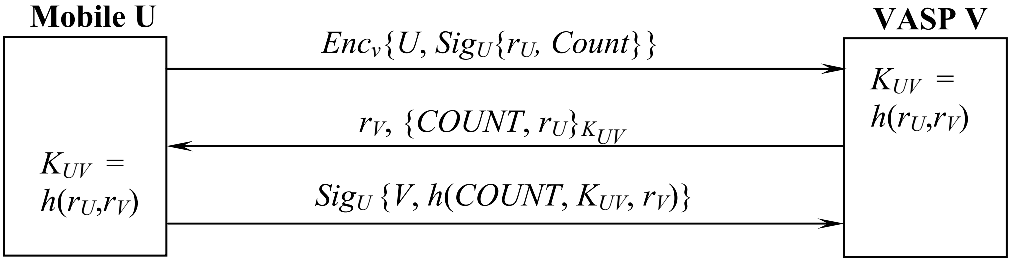
Mobile U
VASP V
$Enc_V\{U, Sig_U\{r_U, Count\}\}$
r_V, $\{COUNT, r_U\}_{K_{UV}}$
$Sig_U\{V, h(COUNT, K_{UV}, r_V)\}$
$K_{UV} = h(r_U,r_V)$
$K_{UV} = h(r_U,r_V)$

Figure 7 Newe-Coffey Improved Boyd-Park Key Agreement Protocol

The first message supplies the identity of *U* and signed information $Sig_U\{r_U, Count\}$ to enable authentication of *U* by *V*. User anonymity is achieved by encrypting the message with the public key of *V*. The second message supplies confirmation of the arrival of r_U to *U*, and a random nonce for generating the session key K_{UV} and for use with step 3. The third message provides key confirmation and key freshness for *V*.

4. SUMMARY

In this paper an introduction to wireless communication security protocols was given and both second-generation and third-generation protocols were discussed. The primary difference between second generation protocols and third generation protocols is the requirement of the mobile device/phone to handle information transfer to and from internet sites as WEB active components. While second generation systems can supply the necessary security for data transfer, the ability to setup initialisation of payment for valued added services, or supply non-repudiation services is not readily provided. These services are an essential requirement for third generation mobile systems where e-commerce applications will take a central role.

Protocol	**Requirements List**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Improved BYC - MSR + DH	✓	✓	✓	x	✓	✓	x	✓	*✓	*✓
Improved Beller-Yacobi	✓	✓	x	x	✓	✓	x	✓	*✓	*✓
ASPeCT	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Improved Boyd-Park	✓	*✓	✓	✓	✓	✓	✓	✓	*✓	✓

*✓ - Indicates that the protocol can be modified to provide this service.

Table 1 Requirements achieved by mobile protocols

From table 1 it can be seen that the ASPeCT and Improved Boyd-Park protocols meet all of the security requirements listed in section 1. The protocols which were developed for second generation systems meet only some of the requirements. 2G protocols were primarily designed to provide link security to protect against passive attacks on the air interface. In 3G systems end-to-end security is the primary concern.

REFERENCES

[1] Beller M.J., Chang L.-F. and Yacobi Y., "Privacy and Authentication on a Portable Communications System", Proceedings of GLOBECOM '91, pp. 1922-1927, 1991.

[2] Beller M.J., Chang L.-F. and Yacobi Y., "Security for Personal Communication Services: Public-Key vs Private Key Approaches", Proceedings of Third IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '92), pp. 26-312, 1992.

[3] Beller M.J., Chang L.-F. and Yacobi Y., "Privacy and Authentication on a Portable Communications System", IEEE Journal on Selected Areas in Communications, vol. 11, pp. 821-829, 1993.

[4] Carlsen U., "Optimal Privacy and Authentication on a Portable Communications System", ACM Operating Systems Review, vol. 28, No. 3, pp. 16-23, 1994.

[5] Boyd C. and Mathuria A., " Key establishment protocols for secure mobile communications: A selective survey", Information Security and Privacy '98, vol. 1438, pp. 344-355, 1998.

[6] Beller M.J. and Yacobi Y., "Fully-Fledged two-way Public Key Authentication and Key Agreement for Low Cost Terminals", Electronic Letters, vol. 29, No. 11, pp. 999-1001, 1993.

[7] Advanced Security for Personal Communications Technologies. http://www.esat.kuleuven.ac.be/cosic/aspect/index.html

[8] Boyd C. and Park D.G., "Public Key Protocols for Wireless Communications", ICISC '98, pp. 47-57, 1998.

[9] Newe, T., Coffey, T., 2003. “Formal Verification Logic for Hybrid Security Protocols”. International Journal of Computer Systems Science and Engineering, Vol. 18 no 1, January 2003, pp 17-25. ISSN 0267 6192

[10] Rivest R.L., Shamir A. and Adleman L.M., "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". Communications of the ACM, Vol. 21, No. 2, Feb 1978, pp. 120-126.