

Lightweight AAA for Cellular IP

Hitesh Tewari, Donal O'Mahony

NTRG, Computer Science Department, Trinity College Dublin, Ireland

Ph: + 353 1 608 2896, Fax: + 353 1 677 2204, Email: htewari@cs.tcd.ie

ABSTRACT

As Mobile IP based communications networks become increasingly ubiquitous, current trust relationships employed in these networks will no longer remain adequate. With a large number of networks and many millions of roaming users, there is a need to put into place authentication, authorization and accounting (AAA) procedures for secure updating of router entries and accounting of network usage by mobiles. We present a lightweight AAA scheme for micro mobility based architectures which makes use of hash chain trees for efficient storage of hash values and fast authentication of datagrams. A node attaches the next unused hash value in a chain to each datagram, which can be used by intermediate routing nodes to verify the authenticity of the message. We give an overview of current AAA proposals for Mobile IP and the emerging area of micro mobility protocols, before presenting the design goals of our protocol and details of our solution.

1 AAA IN MOBILE IP

The IETF has specified the Mobile IP [1] protocol for wide area mobility management. It enables a node to move freely from one point of connection on the Internet to another, without disrupting end-to-end TCP connectivity. Mobile nodes are required to securely register a care-of-address (COA) with their home agent (HA) while roaming in a foreign domain. If however security mechanisms are not employed, the network can be compromised through remote redirection attacks by malicious nodes. In addition, mechanisms are needed that allow foreign agents (FA) in the visited domain to verify the identity of mobile nodes and authorize connectivity based on local policies or the ability to pay for network usage. Figure 1 adapted from [2] shows the Mobile IP network model where a correspondent node is communicating with a roaming mobile node via the MN's home agent.

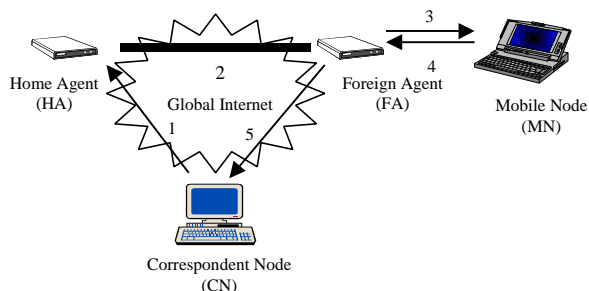


Figure 1: Mobile IP network model

With the availability of low-cost handheld devices such as PDAs and wireless networking hardware, there has been a lot of interest in building Mobile IP based wireless access networks. However for such networks to become commercially viable, authentication and accounting procedures need to be put in place.

Current work in the IETF is focusing on providing an infrastructure for authentication, authorization and accounting (AAA) services to the various entities in the network [3]. Figure 2 depicts the entities and the relationships between them. A foreign domain contains one or more AAA servers (AAAF) and multiple foreign agents. The FAs also known as *attendants* interact with a mobile node to authenticate its credentials. A foreign agent has a security association with its local AAA server, which in turn will have further security associations with other AAA servers. If the AAAF cannot verify the credentials of a mobile node, it will contact the MNs home AAA server (AAAH) with whom it must share a security association. A security association at a very minimum consists of a shared secret between two entities.

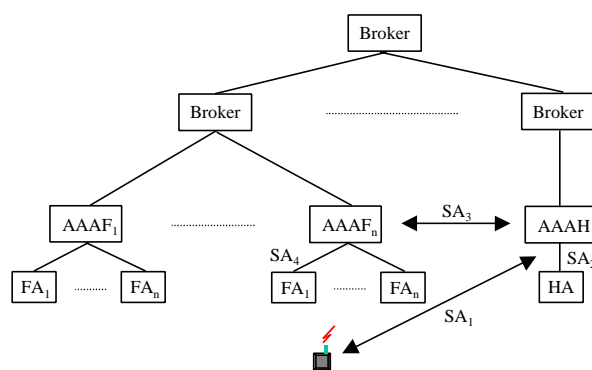


Figure 2: Mobile IP/AAA entities and trust model

In the AAA trust model for Mobile IP, a mobile node shares a security association SA_1 with the AAA server in its home domain. The AAAH in turn shares a security association SA_2 with the home agent. It is also necessary for the AAAH and the AAAF to share a security association SA_3 in order that the AAA server in the foreign domain can verify the credentials of roaming mobile node. Finally the attendant must share a security association SA_4 with the AAAF in order for it to allocate local resources to a mobile node. For scalability reasons the concept of *brokers* (AAAB) is employed which means that a foreign domain does not need to keep security associations with every possible home domain.

Once a mobile node has been authenticated three session keys are generated by the AAAH. Each session

key that is generated by an AAA server will generally be distributed to two entities in the network. The method by which the keys are encoded depends upon the security association between the entities. The Mobile-Home key $K_{MN,HA}$ is shared between the mobile node and the home agent and is encrypted using the security association SA_2 for the HA and using SA_1 for the MN. When a mobile is roaming in a foreign network, this key has to be transported via the AAAF and the serving FA in the foreign network. The Mobile-Foreign key $K_{MN,FA}$ is shared between the MN and the FA. It is encrypted using SA_3 for the FA and SA_1 for the MN. The AAAF forwards the key to the correct FA using the security association SA_4 . Finally the Foreign-Home key $K_{FA,HA}$ is shared between the FA and the HA. It is encrypted using SA_3 for the FA and SA_2 for the HA.

After the initial distribution of the session keys, there is no further requirement to invoke the AAA protocols until the keys expire. During intra-domain handover the new FA will contact the AAAF and obtain the session keys $K_{MN,FA}$ and $K_{FA,HA}$, which were previously assigned to the old FA. Inter-domain handover requires that the MN obtain a new set of sessions keys from the AAAH.

From the above discussion it is clear that the IETF AAA proposal for Mobile IP involves a number of entities each with one or more pre-established security associations. This makes the protocol fairly complex and security analysis difficult. Use of brokers in the system requires that the two administrative domains have security associations with the broker. The broker then becomes privy to all security exchanges between the two domains and has to be trusted.

2 MICRO MOBILITY

Mobile IP is suitable for wide area or macro mobility management when handoffs occur infrequently between base stations. With the rapid growth in the number of portable computers and handheld devices, a new family of mobility protocols has been developed. Micro mobility can be defined as mobility within a domain with frequent handoffs. Local handoffs result in signaling messages that are confined to the local domain and do not result in a change in the care-of-address for the mobile in the home network. Only when a node moves between administrative domains are Mobile IP signaling messages generated, which travel across the core network to the home agent. In general, micro mobility architectures consist of a hierarchy of nodes and base stations connected via a single point of attachment to the Internet.

Cellular IP [4] is a micro mobility protocol, which is based around cellular telephony concepts such as passive connectivity, paging and support for fast intra-domain handover for mobile nodes. A Cellular IP (CIP) network, see figure 3, consists of a gateway router that connects it to the Internet as well as several CIP nodes that are responsible for routing of datagrams within the network. The gateway (GW) embodies both home and foreign agent functionality, and is responsible for filtering out all signaling messages that are specific to the Cellular IP network.

Cellular IP distinguishes between *idle* and *active* mobile hosts and maintains two types of caches to hold hop-by-hop mappings for the same. A *page cache* maintains a mapping for mobile hosts that are not actively transmitting or receiving data but want to be reachable for incoming packets, whereas a *route cache* maintains mappings for only those hosts that are currently receiving or expecting to receive data. Each Cellular IP node maintains a route cache. Paging caches can be placed at strategic points within the network so as to maximize network efficiency.

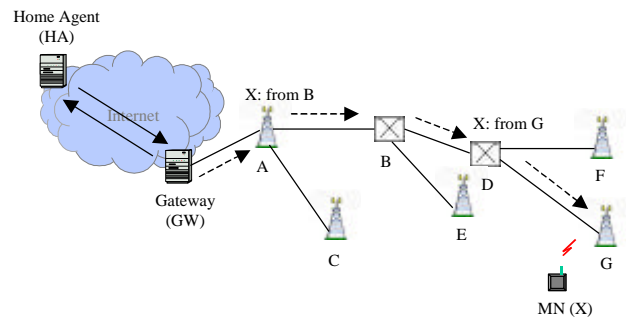


Figure 3: Cellular IP architecture

On power up a mobile node registers itself with the Cellular IP network. The initial registration message is transmitted towards the gateway on a hop-by-hop basis. Each intermediate Cellular IP node creates a mapping for the IP address of the mobile node and the neighbor that forwarded the packet. When the message arrives at the gateway router, it is dropped after the gateway has added an entry for the mobile node in its route cache. Subsequent data packets are used to refresh the existing cache entries, which are valid for a system specific time known as the *route timeout* period. As long as the mobile node has data to send, Cellular IP nodes along the path to the gateway keep an up to date mapping for the mobile node's point of attachment on the network. In cases where a mobile node does not have any data to send but wishes to maintain a valid route cache entry, it periodically sends a *route update* message towards the gateway. Only route update control messages can be used to establish or refresh a route cache entry while data packets can only refresh an existing cache entry. Similarly *page update* control messages can be used to establish page cache entries.

Page and route update messages can be used create entries within Cellular IP caches, which can result in changes to the routing of packets within the network. It is therefore of vital importance that each Cellular IP node, prior to acting on any signaling information authenticates all such messages. Unauthenticated signaling messages can be used to impersonate another node and create denial-of-service attacks. A malicious host could generate false signaling messages and trick the node's home agent into adding a false care-of-address for the node in its routing tables. This would result in packets destined for a node being routed incorrectly by the HA to an unknown destination.

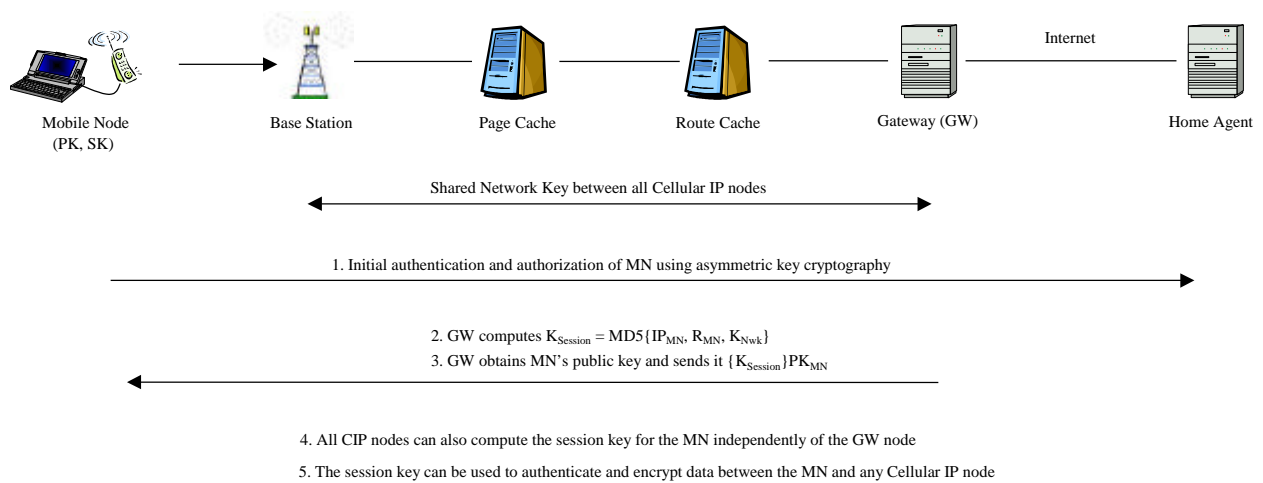


Figure 4: Key management in Cellular IP

2.1 Authentication in Cellular IP

Cellular IP employs a fast session-key management scheme that allows for authentication of control packets by CIP nodes, see figure 4. All CIP nodes in a Cellular IP network have knowledge of a *shared network key*. A session key for a mobile node is calculated by combining the mobile node's IP address (IP_{MN}), a random value assigned to the host (R_{MN}) and the shared network key (K_{Nwk}) as follows:

$$K_{\text{Session}} = \text{MD5}\{\text{IP}_{\text{MN}}, \text{R}_{\text{MN}}, \text{K}_{\text{Nwk}}\}$$

The random number can be assigned by the gateway and is carried in all in all signaling messages generated by the mobile node. All other CIP nodes can independently compute $\text{K}_{\text{Session}}$ as they can obtain the IP address and random number from the signaling messages generated by the mobile. The payload of a signaling message carries a message authentication code (MAC), which comprises of the session key, a timestamp and the packet contents.

There are a number of drawbacks with authentication scheme employed in Cellular IP. Each mobile node in the system requires a public-key pair. For large numbers of nodes this can quickly become a scalability issue. All Cellular IP nodes have knowledge of the shared network key. Repeated use will lead to a decrease in the effective security of the key. Data packets can refresh a cache entry but do not need to be authenticated by a CIP node. This implies that any node can form a malicious data packet and keep cache entries alive to disrupt traffic flows. Finally there are no mechanisms in place for accounting for network usage within a Cellular IP network. Once a node stops transmitting signaling or data packets, the cache entries expire and are subsequently deleted by the CIP node. There will be no further record of the mobile node ever being present on the network.

3 LIGHTWEIGHT AAA PROPOSAL

We propose a new lightweight approach to authentication and accounting for micro mobility in IP networks. We make use of unbalanced one-way binary

trees (UOBT) for generating authentication values to be sent with datagrams. A UOBT [5] is a *hash chain tree* that allows one to generate multiple hash chains dynamically and efficiently, while minimizing storage requirements on the mobile device. Below we identify a number of security related requirements, followed by a short description of UOBT and finally our scheme.

3.1 Protocol Goals

The protocol has been designed to address the issues of authentication and accounting in Cellular IP networks. However it is sufficiently generic to be applied to other micro mobility or hierarchical schemes and has the following goals:

- Authentication of signaling messages. Datagrams that carry signaling or control messages should be accompanied by authentication information. A node must be able to quickly verify the authenticity of such packets prior to creating or updating a routing table entry.
- Accounting of network usage. Each datagram should carry a unique identifier that allows the network operator to associate packets belonging to a specific user and allows him to charge for usage of network resources.
- Minimal use of cryptographic keys. The number of entities in the system that require cryptographic keys should be kept to a minimum. Public keys should only be assigned to long-lived entities such as the home agents and gateway nodes.
- No user digital signatures. End users should not be issued with public key pairs, as this requires the existence of a public key infrastructure (PKI). With large numbers of users maintaining such a PKI becomes a complex task.

3.2 Hash Chains

In order to verify the authenticity of signaling messages efficiently, the number of computationally expensive operations needs to be minimized. Asymmetric key algorithms such as RSA are more computationally intensive than symmetric key cryptographic algorithms

such as DES, which in turn require more computation than hash functions such as MD5.

A user generates a *hash chain* [6, 7] of length N by applying a hash function N times to a random value P_N , the *root* of the hash chain to obtain a final hash P_0 , the *anchor* of the hash chain. The user *commits* to the hash chain by digitally signing the anchor with his private key.

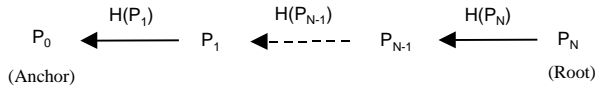


Figure 5: Generating a hash chain

For each authentication value, the user releases the *pre-image* of the last hash to be sent along with a datagram. P_1 is released with the first datagram, P_2 for the second datagram and so on. Since the hash function is one-way, only the user could have generated the hash value.

3.2.1 UOBT Hash Chain Trees

The UOBT scheme is an efficient hash chain scheme, where the root of each chain is derived from another hash chain. The scheme is ideal for devices such as mobile nodes that may have storage limitations, as only the *tree root* value has to be stored on the device to be able to reconstruct the entire UOBT.

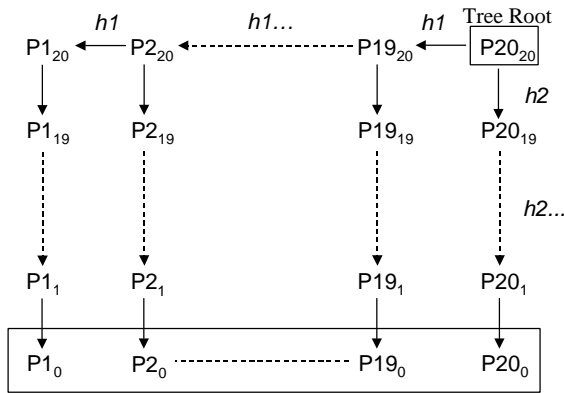


Figure 6: Unbalanced one-way binary tree

An example UOBT where $P_{20_{20}}$ is the tree root is depicted in figure 6. We repeatedly apply a hash function $h1$ to this value to obtain the *backbone hash chain* ($P_{20_{20}}, \dots, P_{1_{20}}$). Each of these hash values is used in turn as the *secret root* value for deriving the individual sub-chains by applying another hash function $h2$. For example, the value $P_{20_{20}}$ is the root of the $P2$ chain, which consists of the values $P_{20_{20}}, P_{20_{19}}$ and so on until the anchor of that chain P_{2_0} is reached. The result is a UOBT with a backbone chain length of 20, with each sub-chain also consisting of 20 hash values and an overall hash chain of 400 hash values.

The whole UOBT has to be used in a single domain, however only parts of the sub-chains can be spent if required. Also there is no need to store the whole UOBT

on the user's device as the individual sub-chains can be generated as and when required.

3.3 Authentication and Route Establishment

A mobile node pre-computes a UOBT prior to registering with the Cellular IP network. On receiving an advertisement from a nearby base station, the mobile node sends a registration request to the gateway. The *RegReq* contains a keyed MAC [8] on the set of anchors of the UOBT with the secret key (K_{MN}) that the mobile node shares with its home agent.

Set of anchors: $\{P_{1_0}, P_{2_0}, P_{3_0}, \dots, P_{18_0}, P_{19_0}, P_{20_0}\}$

The network access identifier (NAI) [9] is used to identify the mobile node's home domain by the gateway and to forward the registration request to the node's HA.

$MN \rightarrow GW \text{ RegReq}\{MN, GW, NAI_{MN}, \{P_{1_0}, \dots, P_{20_0}\}, (NAI_{MN}, GW, \{P_{1_0}, \dots, P_{20_0}\})MAC_{K_{MN}}\}$

The MNs home agent verifies the MAC and adds the GW address as the new care-of-address for the mobile node. The home agent then sends a *RegRepl* message back to the gateway, which contains a signed commitment. The commitment consists of a digital signature on the set of anchors of the sub-chains that comprise the UOBT. The gateway in turn appends its own signature to the message, which binds the mobile node's address to the set of anchors of the UOBT. The gateway broadcasts the message to all Cellular IP nodes in the network and also forwards it to the MN concerned.

$GW \rightarrow MN \text{ RegRepl}\{GW, MN, NAI_{MN}, \{P_{1_0}, \dots, P_{20_0}\}, (NAI_{MN}, GW, \{P_{1_0}, \dots, P_{20_0}\})Sig_{HA}, (MN, \{P_{1_0}, \dots, P_{20_0}\})Sig_{GW}\}$

We employ the use of an *authentication cache* at each CIP node to keep a copy of the commitment. Intermediate CIP nodes on the path between the MN and the GW keep a copy of the current sub-chain anchor and the last hash value received in their route cache, see figure 7.

3.3.1 Releasing Authentication Values

When a mobile node sends a packet or a signaling message on the network it appends the next unused hash value from the sub-chain that it is currently using. Packets originating at a MN contain the next hash value (PX_Y), the hash number (Y), the anchor of the sub-chain (PX_0) and the number of the sub-chain (X).

Datagram: $\{MN, Dest, PX_Y, Y, PX_0, X, Data\}$

This allows for intermediate Cellular IP nodes to quickly verify the anchor by consulting their route cache and in turn to verify the individual hash value by either hashing back to the anchor or the last hash value received.

3.3.2 Fast Handover

During handover a mobile node receives advertisements from other base stations in the vicinity which contains the base station identifier. This triggers the use of the next unused sub-chain in the UOBT within the mobile.

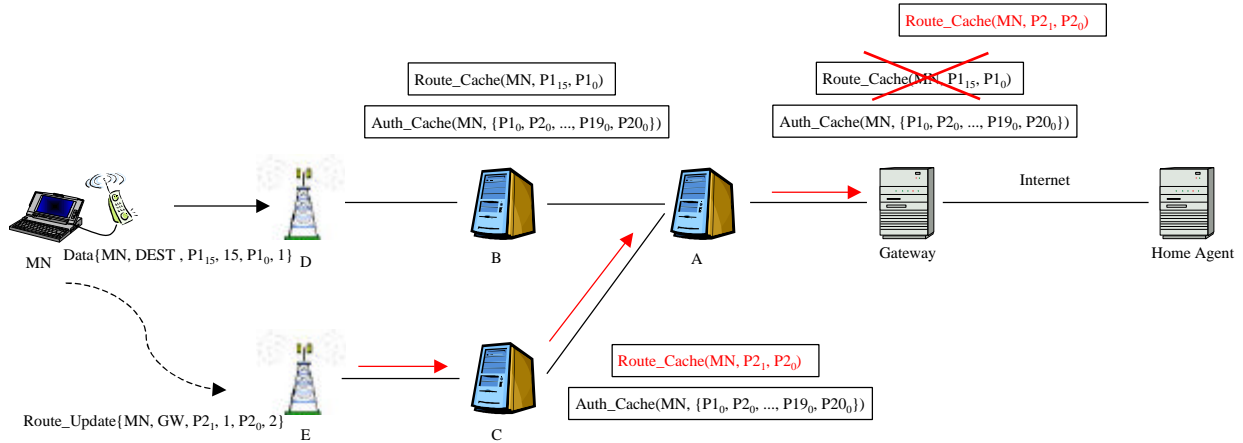


Figure 7: Authentication using UOBT hash values

Figure 7 shows the message exchanges when the MN moves from node D to E. Prior to moving to node E the MN was using the first sub-chain (P1) and the last hash value it released was the fifteenth ($P1_{15}$) in that chain.

When the mobile switches to node E it must transmit a *route update* message towards the gateway node. At this point the MN switches to the next available sub-chain (P2) and transmits the first hash value ($P2_1$). All intermediate CIP nodes in the *new path* to the gateway can verify the accompanying chain anchor ($P2_0$) by consulting their authentication cache. A node only has to perform a *single hash operation* to verify the authenticity of the packet. This aids fast handover between nodes for micro mobility.

3.4 Accounting

The use of hash chain trees allows us to fulfill our requirement of *intra-domain* accounting for network usage. The next unused hash in the UOBT to which the MN has committed itself must accompany each datagram that is transmitted by a mobile node. All packets pass through or are terminated at the gateway node depending on whether they are ordinary data packets or special signaling messages. Thus the gateway is an ideal location to perform accounting procedures. The gateway can keep a record of the *highest* spent hash in each sub-chain belonging to the UOBT to which the MN has committed. Hashing back from this value to the sub-chain anchor will yield the number of hashes spent in a particular sub-chain.

Since the hash chain is generated using a *one-way* function and since the mobile has committed to the UOBT of which the chain is part, we have a means of associating network usage to a mobile node. An account of the actual resources may then be assembled based on the total number of hashes spent. Alternatively there may be an *inter-domain* accounting relationship between the gateway and the home agent in the mobile node's home network, which can be used to settle any charges generated by a visiting mobile.

4 DISCUSSION

We have presented a solution to allow for authentication and accounting procedures to be implemented efficiently

in micro mobility architectures. Authentication of signaling messages is guaranteed without the need for a large scale PKI to be in place. Hash functions and minimal use of cryptographic keys allows the solution to be efficient and scalable. Asymmetric key procedures are only used during the initial registration phase and verification of the signed commitment. Unlike the Cellular IP security architecture where each mobile node requires a public key pair, our system limits public key pairs to entities such as the CIP gateway and the mobile node's HA. Since these are trusted secure entities within the system, there is less of an overhead with regards to lost or compromised keys and maintenance of certification revocation lists (CRLs). This further implies that we require only a minimal PKI, with a top-level certification authority (CA) cross-certifying the public key certificate of the gateway node and home agent. A mobile node shares a symmetric key with its HA and possess the public-key certificate of the HA. The mobile node is required to perform a minimal amount of public key processing in verifying digital signatures generated by the home agent and the gateway nodes.

The main disadvantage of the scheme is that hash chains are of a finite length and there is a possibility that a node may run out of hash values during a session. In the case of a UOBT if there are unused sub-chains, then the user can switch to the next chain by sending a *route update* message. However if there were no further sub-chains available in the UOBT, then this would result in the dropping of a connection or the expiration of a cache entry. This would be particularly true for real-time communications such as voice telephony or video conferencing where the MN may transmit a large number of datagrams during a session. An alternative approach is to negotiate at registration time a value, which corresponds to the number of datagrams that the gateway will allow before expecting the next hash value in the chain to be released by the mobile node. For example each hash value could allow a MN to transmit a hundred datagrams. This could be implemented simply as a counter at the gateway node.

On small devices with limited storage, such as a PDA, it may not be possible to store all the hash values

of a long hash chain. If a UOBT with a backbone chain length equal to the length of each sub-chain is used. It can be shown that the average computational overhead is $n^{1/2} - 1$, where n is the number of values in the UOBT, and $n^{1/2}$ is the square root of n . A 30x30 UOBT requires 29 hashes on average to compute an authentication value. There is an initial overhead in transporting the set of anchors to the home agent for the commitment to be signed. The size of this message directly depends on the length of the backbone chain. However, once the initial registration has taken place, the number of cryptographic operations performed during the communications session are greatly reduced.

In [10] the author has carried out a comprehensive study of micropayment performance. He has made efficiency comparisons of various hash algorithms versus symmetric and asymmetric key algorithms. One observation is with regards to the number of operations performed per second. It was noted that one could perform five times as many MD5 hash computations as DES encryption operations and over five thousand times as many RSA signature generations. In general the author observed that hashing is an order magnitude faster than symmetric encryption, three orders of magnitude faster than signature verification and four orders of magnitude faster than signature generation.

Our scheme makes use of hash algorithms for authentication of signaling messages in the network. This allows for very fast generation and verification of hash values. From our discussions above we are confident that the performance of our scheme will be comparable if not better than the IETF AAA and Cellular IP authentication schemes.

5 CONCLUSION

With the widespread use of PDAs and the availability of cheap wireless networking hardware there has been a mushrooming in the number of wireless access networks. Currently such networks are operated by individual organizations and are usually closed to users who belong to other network operators or organizations. One of the main reasons is that such closed networks do not have any AAA provisioning policies in place, and thus cannot deal with nodes with which they do not have a pre-established security relationship.

We believe that in the future there will be a large number of micro and pico-cellular based All-IP mobile networks, that will provide the next generation of telecommunications services to a very large user population. These networks will require secure and scalable AAA provisioning. Current AAA proposals are too heavyweight and not suitable for large-scale deployment. Our solution aims to provide an efficient means of authenticating signaling messages and accounting of resources for micro mobility.

6 REFERENCES

- [1] C. Perkins ed., "IP Mobility Support", IETF RFC 2002, Oct. 1996.
- [2] C. Perkins, "Mobile IP", *IEEE Communications Magazine*, Vol. 35, No. 5, May 1997, pp. 84-99.

- [3] S. Glass, T. Hiller, S. Jacobs, C. Perkins, "Mobile IP Authentication, Authorization and Accounting Requirements", IETF RFC 2977, Oct. 2000.
- [4] A. Campbell et al., "Design, Implementation, and Evaluation of Cellular IP", *IEEE Personal Communications*, Vol. 7, No. 4, Aug. 2000, pp. 42-49.
- [5] S. Yen, L. Ho, C. Huang, "Internet Micropayment Based on Unbalanced One-way Binary Tree", *Proc. CryptTEC'99*, Hong Kong, July 1999, pp.155-62.
- [6] L. Lamport, "Password Authentication with Insecure Communication", *Communications of the ACM*, Vol. 24, No. 11, Nov. 1981, pp. 770-72.
- [7] D. O'Mahony, M. Peirce, H. Tewari, *Electronic Payment Systems for E-Commerce*, 2nd Edition, Artech House Publishers, Boston/London, 2001.
- [8] P. Metzger, W. Simpson, "IP Authentication using Keyed MD5", IETF RFC 1828, Aug. 1995.
- [9] B. Aboba, "The Network Access Identifier", IETF RFC 2486, Jan. 1999.
- [10] M. Peirce, "Multi-party Micropayments for Mobile Communications", PhD Thesis, Trinity College Dublin, Ireland, Oct. 2000.