Investigating Enhanced Route Optimization for Mobile IPv6

Shilei Kuang, Robert Elz, and Sinchai Kamolphiwong
Centre for Network Research (CNR), Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University, Thailand
kuang_shilei@hotmail.com, kre@munnari.OZ.AU, ksinchai@coe.psu.ac.th

Abstract

Enhanced Route Optimization for Mobile IPv6 applies Cryptographically Generated Addresses (CGAs) to improve security and reduce handover delays. However, the use of CGAs requires computationally expensive algorithms. This may be an issue for small mobile devices with low processing power. It is a problem for correspondent nodes that simultaneously communicate with a large number of mobile nodes, such as publicly accessible servers. This paper investigates Enhanced Route Optimization for Mobile IPv6, concentrating upon the costs particularly for the Correspondent Node. The costs of implementing Enhanced Route Optimization for Mobile IPv6 are not negligible. The Correspondent Node will have to protect itself against potential denial-of-service attempts from attackers by limiting the amount of resources it spends on CGA verification.

1. Introduction

Mobile IPv6 is designed to provide mobility support on top of the existing IPv4 infrastructure, without requiring any modifications to routers or applications. The key benefit of Mobile IPv6 [2] is that even though the mobile node changes locations and addresses, existing connections are maintained.

An advantage of Mobile IPv6 over Mobile IPv4 is that data packets can be directly relayed between the mobile node and its correspondent node. This mode is called route optimization and is possible as IPv6 correspondent nodes are assumed to generally understand the Mobile IP procedures. Route Optimization provides the Mobile Node the opportunity to eliminate inefficient triangle routing and bidirectional tunneling. A Binding Update (BU) message from the mobile node to each active correspondent node enables this Route Optimization.

However, an unauthenticated or malicious Binding Update message would provide an intruder with an easy means to launch various types of attack [9]. Therefore, Route Optimization uses an authentication mechanism called the Return Routability (RR) procedure to protect Binding Update messages (see Figure 1). The Return Routability procedure uses HoT/HoT and CoT/CoT messages respectively to test the Mobile Node’s home address and care-of address reachability. The CoT and HoT message each carry a token from the correspondent node that the mobile node must use to authenticate its BU to the correspondent node. When the mobile node demonstrates knowledge of both token the correspondent node assumes that both the Home address (HoA) and Care-of address (CoA) belong to the mobile node, as packets sent to both those addresses reached the mobile node. The return routability procedure was designed with the objective to provide a level of security that compares to that of today’s non-mobile Internet [9].

Several enhancements to the RR procedure have been proposed [5]. Vogt and Doll measured performance with and without those enhancements [12,13] and concluded that TCP performance, in particular, was much improved when the Binding Update messages could be performed more rapidly. The IETF created Enhanced Route Optimization for Mobile IPv6 [5] which provides lower handoff delays, increased security, and reduced signaling overhead. Our purpose is to investigate the costs of implementing this new procedure.

This paper is organized as follows: Section 2 presents background information for CGA method in Mobile IPv6 and Enhanced route optimization. Section 3 gives our evaluation of this protocol and proposes Permanent Home Keygen token process before handoff. Sections 4 and 5 respectively, present our

Figure 1. Route Optimization Procedure

2. Background Information

2.1 Cryptographically Generated Addresses

A Cryptographically Generated Address (CGA) [3] is a method for securely associating a cryptographic public key with an IPv6 address in the Secure Neighbor Discovery (SEND) protocol [4]. The basic idea is to generate the interface identifier (i.e., the rightmost 64 bits) of the IPv6 address by computing a cryptographic hash of the public key. The resulting IPv6 address is called a cryptographically generated address. The corresponding private key can then be used to sign messages sent from the address.

To verify the association between the address and the public key, the verifier needs to know the address itself, the public key, and the values of the auxiliary parameters. The verifier can then go on to verify messages signed by the owner of the public key (i.e., the address owner). No additional security infrastructure, such as a public key infrastructure (PKI), certification authorities, or other trusted servers, is needed. The specification [3] requires the use of an RSA public/private key pair [17] in CGA implementation. Note that an attacker can always create its own CGA address but he will not be able to spoof someone else's address since he needs to sign the message with the corresponding private key which is assumed to be known only by the real owner.

2.2 The Use of CGAs in Enhanced Route Optimization

The purpose of HoTI/HoT is ensuring packets can only be redirected by the legitimate recipient. The legitimate recipient is identified through the home address, and only the legitimate recipient is expected to receive the Home Keygen Token sent to the home address.

CGA [3] can provide the same functionality without sending a packet to the home address. A node that uses a CGA at a certain time can prove at a later time that it is still the same node when it uses this CGA again. But instead of relying on a routing property, as with the home-address test, this proof can be drawn from the CGA’s special interface identifier. The CGA owner signs important packets with its private key and includes its public key along with the auxiliary data in these packets. Since it is computationally hard to produce another public/private-key pair that hashes to the same CGA, the recipient of the signed message can verify, by recomputing the hash and comparing it with the CGA’s interface identifier, that the sender must be the legitimate owner of this CGA.

Enhanced Route Optimization (ERO) use a CGA rather than use HoTI/HoT to prove the ownership of Mobile Node’s home address. A mobile node uses a CGA as its home address, and it signs BUs with its private key. The correspondent node can thus verify that the BU is from the same mobile node that used this home address before. Unfortunately, HoTI/HoT are still required to prove the association with the prefix in the home address because CGA does not ensure that a mobile node can indeed receive packets at the home address it claims to own. This property could be misused for a flooding attack against the home network.

ERO attends to these problems by combining CGAs with HoTI/HoT and CoTI/CoT. A HoT is performed at first contact between a Mobile Node and a Correspondent Node. This test verifies that the mobile node is the legitimate owner of the home address. Since the home address is cryptographically generated, the correspondent node will recognize the mobile node as the owner of this home address during subsequent registrations without having to do the home address test again. But CGA-based authentication involves public-key cryptography and is hence computationally much less efficient than authentication.
through a shared secret key. [5] proposes use an initial CGA-based authentication to securely exchange a secret Permanent Home Keygen Token (PHKT) between a mobile node and correspondent node.

An issue with CGAs in general is that they involve computationally expensive algorithms. This may be an issue for small mobile devices with low processing power. It is an issue for correspondent nodes that simultaneously communicate with a large number of mobile nodes, such as publicly accessible servers. Let along the computational overhead required for legitimate mobile nodes, a correspondent node will have to protect itself from potential denial-of-service attempts from attackers by limiting the amount of resources it spends on CGA verification.

2.3.3 Enhanced Route Optimization

In the standard RR procedure, the BU is delayed until both the HoTt/HoT and CoTt/CoT tests complete. The HoTt/HoT may itself be delayed until the Home Agent has been informed (using a BU protected and authenticated using IPSec [2]) of the new location of the Mobile Node. An aggressive Mobile Node however can send the HoT packet immediately after the Home Agent BU is transmitted, trusting the network to deliver the packets in order, and the Home Agent to accept and process the BU ready for the following HoT. Arkko, Vogt and Haddad [5] claim that Mobile Node should not be aggressive this way, the justification being the possible loss of the HoTt/HoT messages. This however seems to be a small price to pay if we assume that usually the BU at the Home Agent is successfully processed and the gain when it is is significant. Regardless, the HoTt/ HoT combination will almost always limit the timing of the BU to the Correspondent Node. Both HoTt/HoT and CoTt/CoT involve a round trip between Mobile Node and Correspondent node, but the HoTt/HoT takes a detour via the Home Agent, so it should normally be slower than the CoTt/CoT.

The ERO process is illustrated in Fig.2. It uses a CGA to ensure mobile node’s home address. However, the HoTt/HoT may be performed before the Mobile Node changes location. The BA or EBA that follows the CGA protected BU carries the PHKT, that is used to authenticated future Bus from this Mobile node to the same correspondent node. After removing that delay source, ERO permits deferred CoTt/CoT—actually performing the CoT as an option in the first BU which is called Early Binding Update (EBU) [5] and with the CoT merged with the Early Binding Acknowledgement (EBA). To avoid attacks upon an unauthenticated victim care-of address, a Correspondent Node processing a deferred CoT limits the amount of data it will send to the Mobile Node’s care-of address. This is known as Credit Based Authentication (CBA) and means that the Correspondent node does no more harm to a victim at the care-of address than the Mobile node could have done by sending its packets directly at the care-of address.

![Figure 2. Enhanced Route Optimization Procedure](image)

3. Permanent Home Keygen Token Process Before Handoff

It is clear that for the first BU using ERO to a particular Correspondent Node (before the permanent token is obtained) to gain any performance enhancement over RR authenticate BUs, the HoTt/HoT sequence must have completed before movement. Otherwise the delay waiting for this process (particularly if the aggressive process is not adopted) would allow the CoTt/CoT to occur, and the regular RR procedure would be just as quick, without either CGA Public Key computations or credit calculations.

ERO will benefit only if either HoTt/HoT can be performed before movement, or if we can expect the association between Mobile Node and Correspondent Node remain active for more than one movement event, so that subsequent movement Binding Updates reap the benefits established during the first. To send proactive HoTt/HoT the mobile node could commence
this procedure as soon as an association with a Correspondent Node is established, assuming that the mobile node will move before that association terminates, or in the words of [12] "The Mobile Node can invoke proactive home-address test on a just-in-time basis, if its link layer provides a trigger announcing imminent handoff." While certainly plausible with certain link level technologies, imminent handoff trigger events are not something we have encountered in practice.

In practice a Mobile Node that does not anticipate long continuous associations (through multiple movements) with most Correspondent Nodes, and that desires the speedy ERO handovers will need to perform early HoTI/HoT exchanges with relevant Correspondent Nodes. Further, since Mobile IP processing is typically (by design) well separated from the applications, there generally is no knowledge of which Correspondent Nodes need speedy handover, or which may remain associated for lengthy periods and thus expect to remain existing when any motion occurs. Thus, in practice, the mobile node is likely to simply perform the HoTI/HoT tests whenever a new Correspondent Node association is formed, and then repeat that before the expiry of the lifetime of the returned token, or perform an early no-motion BU to obtain a "permanent token".

![Figure 3. Permanent Home Keygen Token process before handoff](image)

There is also no way to distinguish a Mobile Node from any other node which is "mobile" in the Mobile IP sense which merely implies that the node's address has altered. Often this will be because the node has physically moved, but there are other causes. All nodes should be able to benefit from the Mobile IP procedures, not just the portable ones. The effect of thousands or millions of nodes performing HoTI/HoT exchange with the more popular internet servers, even if those servers refuse to participate needs further study.

To determine the effectiveness of ERO we need to determine the practicality of performing the HoTI/HoT sequence before the Mobile Node moves, the effectiveness of the CBA scheme, and evaluate the cost of using CGA with the associated RSA Public Key algorithms. Normally a Permanent Home Keygen Token is allocated and returned in response to the first BU (CGA authenticated), after the mobile node moves the first times. Until then short lifetime Temporary Home Keygen Tokens are used, and renewed as required. While the specification does not specifically suggest it, we have implemented a mode of early BU (a de-registration without previous registration) that allows the "Permanent" token to be obtained before the mobile node moves (see Figure 3). This allows the periodic HoTI/HoT sequence to be halted whereas normally the lack of security in those messages requires repetition and obtaining a new token to lessen the effects of packet snooping attacks.

4. Testbed Network Components

The experimental Testbed consists of three FreeBSD [14] nodes playing the roles of the Mobile Node, Home Agent and Correspondent Node. The testbed network shown in Figure. 4 consists of one PC-based rooter (HA), one notebook PC-based Mobile Node and one PC-based Correspondent Node.

![Figure 4. Testbed topology](image)

The Mobile Node may attach to its Home Agent or
to either of the access routers in the visited domains. The exterior interfaces of all routers and the Correspondent Node connect to the “Internet-IPv6 network”. The Mobile IPv6 implementation is that from KAME-SHISA [8], we generate CGAs using software from DoCoMo labs [10] which was primarily intended to test the Secure Neighbor Discovery (SEND) [4] protocols. The RSA algorithms used are those provided by the OpenSSL project [11].

5. Problem Statement and Experiment Scenarios

To determine the burden of implementing and using ERO we must measure its costs. In particular, one of the requirements of [5] is:

*Attackers should not be able to cause denial of service against mobile or correspondent nodes through exploiting expensive computations involved in the mobility protocol.*

We seek determine whether or not ERO meets this requirement.

We expect the costs of ERO to be dominated by the costs of the Public Key computations (RSA algorithms). That cost is controlled by the size chosen for the modulus parameter from which the keys are generated. Thus we vary that over a selection of values ranging from 512 to 3072 bits. Values smaller than 512 bits (probably including 512 bits) are too small to be secure. Values larger than 1575 bits cause the generated public keys, encoded as specified in [5, 6, 7], with the other CGA parameters [3] to exceed the capacity of a single CGA parameter option. [5] allows for this, permitting several options to be used to carry the complete CGA parameter block. At key modulus sizes greater than 2040 bits, the packet signature no longer fits in a single option, nor does the returned permanent home keygen token. [5] does not provide for those option size limits to be exceeded; to enable measurements to be made with longer keys, we copied the solution for CGA parameter blocks, and permit multiple options to be combined to carry a single large parameter value.

If attackers send lots of useless EBU packets to the target it will require determination of failure to verify the signatures, requiring much computation. If an attacker were to send many fraudulent EBU packets, each CGA authenticated to a target node that node would need to determine the failure of the RSA signature verification, consuming CPU in so doing.

We measured and report that cost. For comparison we also measured the CPU cost to a target of processing several after packet flooding attacks, in comparison to the network resources require to launch it. Those results are reported below.

The cost of the standard Route Optimization Return Routability Procedure authentication is lightweight, so the cost of ERO for our proposes, amounts to the elapsed CPU time to perform the necessary CGA and signature calculations. To avoid the reported measurements depending upon the actual CPU used in our experiments1 and as we are most interested in the relationship between the various measurements, we have chosen a time unit we call N, all times reported will be given as multiples of N.

![Figure 5. Process Cost time for CGA-based authentication](image)

6. Results and Discussion

For each selected key length Figure 5 shows the time required to generate the RSA signature used to authenticate the CGA-based EBU (CGA-Sign), the time taken to verify that signature from a received packet (CGA-Verify), the time taken to encrypt a Permanent home keygen token to return (PHKT-Encrypt), and the time taken to decrypt the Permanent home token at the Mobile Node (PHKT-Decrypt). It is immediately obvious that the most expensive operations are generating the signature (CGA-Sign) and decrypting the Permanent home keygen token (PHKT-Decrypt); operations using the much larger

1 In our experiment the measurements were obtained using an Intel Pentium4 model Prescott running at a clock frequency of 2.80 GHZ. With this environment, one N represents approximately 7 microseconds.
Private Key. Both of these functions are performed by the Mobile Node, at its option, and hence are not a significant concern here.

![Graph showing the comparison of Process BU with Process Receive BU](image)

**Figure 6. Compare Process BU with Process Receive BU**

While the other operations, those performed by the Correspondent Node, are considerably cheaper, the costs are not negligible, and clearly grow as the key size increases. To obtain a better feel for the Correspondent Node costs, we also measured the total time to process an enhanced EBU. Those results are provided in Fig 6. We also include there the cost of processing a received EBU where the signature verification fails as it might in a Denial of Service attack upon the Correspondent Node. Clearly the time to process the Standard BU does not depend upon the key length (there is no key involved).

However it can be seen that as the key size increases the cost to the Correspondent Node increases rapidly. With greater key length than we were able to test the cost of even discarding a BU with an invalid signature may become intolerable. On the processor we used, with a key modulus of 3072 bits, the Correspondent Node would exhaust its processor time merely validating, and rejecting, approximately 500 invalid BU packets a second, and can respond to only half that number of valid BUs carrying CGA parameters requesting a permanent home keygen token be received.

That is, a Correspondent Node server of this power could handle connections from at most 250 distinct clients per second, if each was to prepare itself for possible later rapid mobile IP handover using ERO.

For comparison, we also measured the network bandwidth an attacker would need to consume to achieve a 100%CPU utilization of a victim node using several other possible packet types. Figure 7 shows those results. It can be seen that EBU consuming less than 5 MB/Sec would saturate the CPU, whereas IPv4, ICMP Echo (ping) packets 65KB big more than 3 GB/sec to achieve the same result. This makes EBU a much more effective denial of service (by CPU exhaustion) attack than ICMP echo, or any of the after methods we tested.

![Graph showing the 100%CPU Cost of EBU, Flooding IPv6/IPv6 and UDPIPv6](image)

**Figure 7. 100%CPU Cost of EBU, Flooding IPv6/IPv6 and UDPIPv6**

7. Conclusion and Outlook

Enhanced Route Optimization applies CGA to increase its security. But attackers may use denial of service against correspondent nodes through exploiting expensive computations involved in the mobility protocol. In this paper we have shown the costs to the Correspondent Node implementing Enhanced Route Optimization for Mobile IPv6 are not negligible, we suspect that with large keys the costs may be considerably greater than reasonable for a busy correspondent node. While CGA use currently mandates the RSA algorithms, it would be interesting to discover whether use of Elliptic Curve Cryptography (ECC) [15, 16] with its smaller key for equivalent security would reduce the computational cost.

We suggest that implementation of correspondent nodes at the very least rejects very large keys without processing. An easy way to achieve this is to reject packets containing multiple CGA parameter options. The maximum key length then will be about 1575 bits, which seems as if it may be manageable. The security