Optimal Routing: A compilation of *Cellular Networking Perspectives* articles.

Cellular Networking Perspectives

Editor: David Crowe • Phone +1-403-289-6609 • Email crowed@cnp-wireless.com

Optimal Routing, Part I

Routing calls to a roaming mobile can be very inefficient in utilization of trunks and switching resources. In the worst case, a Brazilian calling a Canadian mobile roaming in Brazil could invoke two international long distance calls (from Brazil to Canada and back), just to call across Rio.

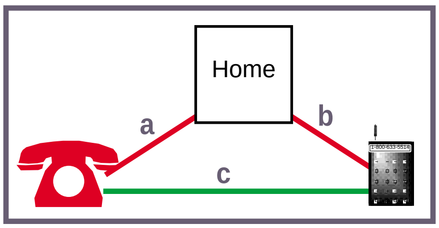
Rectifying this situation involves 'thinking before doing'— determining the whereabouts of the destination mobile before routing a call blindly towards its home system, where the mobile may very well not be. However, solutions stumble because it is difficult to examine a phone number and determine whether it is a mobile or not, especially if it is homed in another country. And, it is even more difficult to determine whether the phone is accessible using the same Mobile Application Part (e.g. GSM MAP or TIA/EIA-41).

Basic Concept

Optimal Routing allows a call to be routed directly from the Originating Switch to the MSC currently serving a mobile (leg 'c' in Figure 1), which replaces both the leg from the Originating Switch to the home MSC (leg 'a') and the leg from the home MSC to the Serving MSC (leg 'b'). Optimal Routing is not applicable (or at least, has no benefit) when a mobile is within its home system (scenario i in Figure 4). It is most useful when the roaming mobile being called is within the local calling area of the calling mobile. Currently, such a situation may require two long distance calls even if the two parties are within spitting distance (see Figure 2). Optimal Routing can reduce this to a single local call (see Figure 3).)

> ***Huh?***
>
> If there are any acronyms or terms you are unfamiliar with, check our website glossary; you will probably find them there:
>
> www.cnp-wireless.com/glossary.html

Figure 1: Optimal Routing



Using a MAP

The *Mobile Application Protocol* or *MAP* allows cellular or PCS systems to interconnect at a high level of intelligence. The two most commonly used are GSM MAP and TIA/EIA-41, although in the near future, 3G systems will likely develop new MAPs. Both GSM MAP and TIA/EIA-41 have the ability to perform optimal routing once the dialed digits have been identified as a compatible number (i.e. served by a network conforming to the same MAP). However, there are several major reasons why this is rarely done:

- Billing Complications
- Recognizing Digits
- Number Portability Concerns)

Billing Complications.

Billing for a call to a roamer without optimal routing is quite simple. The calling party pays for the leg to the home system and the called mobile pays (if it is roaming) for the leg to the current serving system. Airtime charges may be paid by the calling party (Calling Party Pays) or the called mobile (Terminating Party Pays).

With Optimal Routing, new billing scenarios are possible, as shown in Figure 4. When the mobile is at home (scenario i), optimal routing has no effect, so there are no billing complications.

When the mobile is roaming near the calling party (scenario ii), two long distance calls may be avoided, so the toll charges for both parties can be reduced or eliminated.

When the mobile is roaming further away from the calling party than the home system (scenario iii), charges must be split because the charge from the calling party to the mobile's actual location (c) may be greater than either the leg to the home system (a) or the leg from the home system to the current serving system (b), but the charge will almost certainly be less than the sum of charges for legs (a) and (b).

There are several solutions to the billing problems raised by optimal routing. One

Figure 2: **Routing to Roamer: Worst Case Routing**

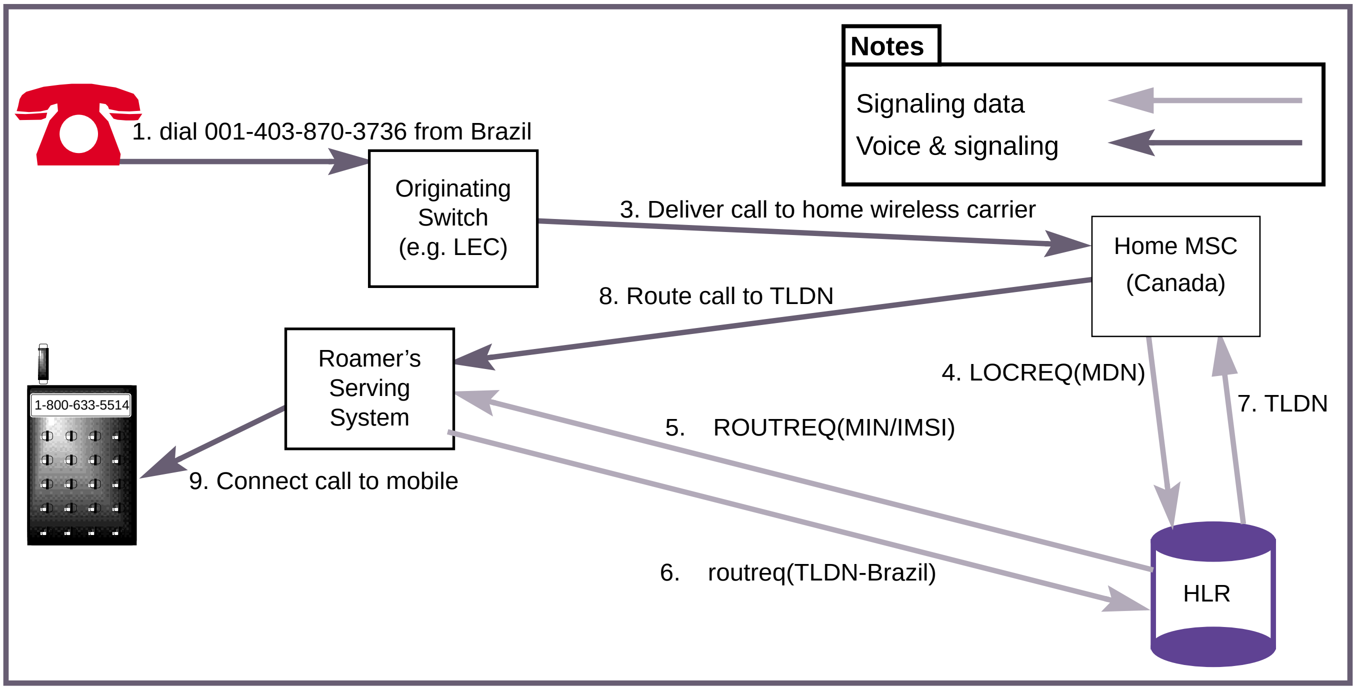
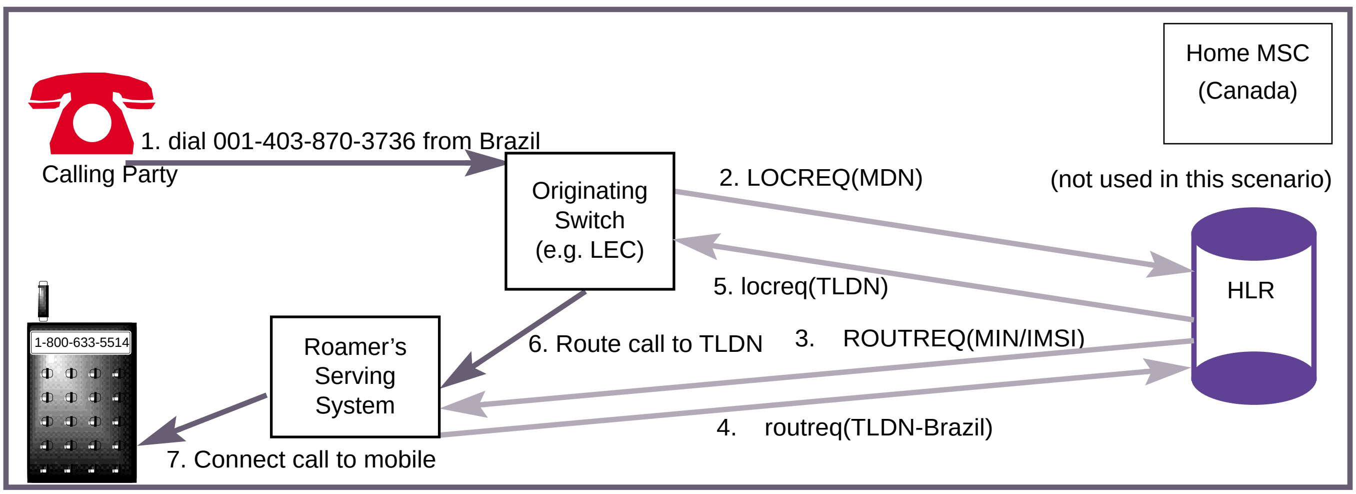


Figure 3: **Optimal Routing using MAP (e.g. TIA/EIA-41)**



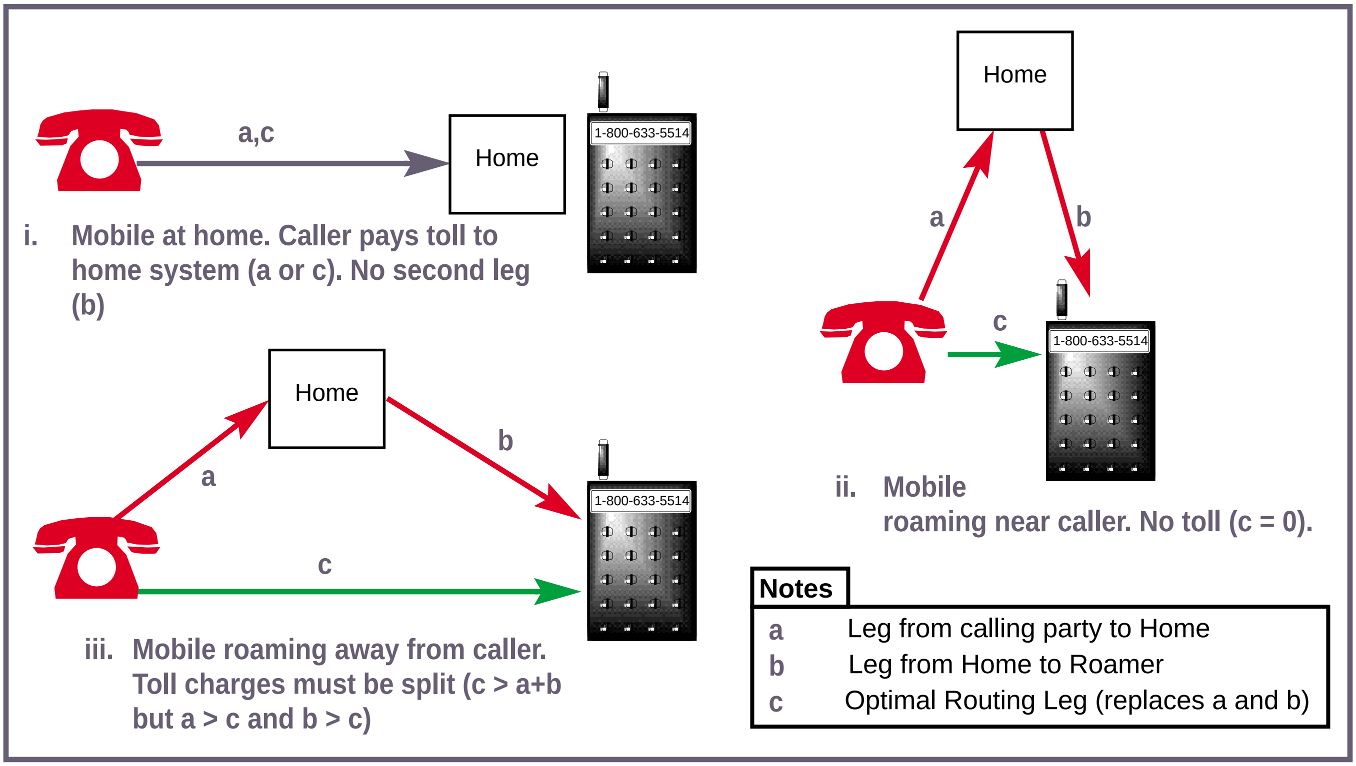
is to leave the charges paid by the calling party unchanged. This means they will sometimes pay toll charges for a call with no toll leg (which is not likely to upset carriers, but this could possibly upset regulators and consumers). Another solution is to apportion the charges, ensuring neither party ever pays more than they would in a non-optimal routing case. This may require wireless–landline billing record exchange, something that is avoided today.

With the trend towards fixed rate calling plans from both wireless and long distance companies, a simple solution is probably the best. The calling party can be charged their standard long distance calling rate, and the called mobile can be charged their standard combined airtime/long distance rate, with the carrier pocketing the savings.

Recognizing Digits.

In many countries, wireless carriers are assigned unique numbering prefixes, so it is possible to recognize that a call is to a wireless carrier by analyzing only the first 2 to 3 digits of the number (requiring a table of 100 or 1,000 entries, at most). Some wireless carriers make this easily recognizable number an essential part of their marketing image (e.g. Shinsegi 017 in Korea). However, it is difficult to extend this ability to international dialing because it would require maintaining much larger tables; it would also require tracking the numbering plan changes in other countries. In North America, the problem is even worse because wireless carriers are assigned blocks of 10,000 numbers (and sometimes even smaller!). A table to examine the first 6 digits of every dialed number would have 1,000,000 entries, and because it would change frequently, it would be virtually impossible to manage.

Figure 4: Optimal Routing Billing Cases



Number Portability Concerns

Number portability makes it risky to identify a carrier based on analysis of dialed digits. A switch may determine that dialed digits are for a mobile, only to discover later they have been ported to a wireless carrier using a different MAP or, even worse, they have been ported to a landline carrier. In the opposite direction, a lesser problem occurs when a carrier determines that dialed digits are not for a wireless carrier, and does not initiate optimal routing when it is actually possible.

Solutions to this problem depend on the method being used, and these will be discussed separately with each approach.

When is it Feasible?

Optimal Routing is feasible under a number of situations:

- Within a single carrier network, digit translations can route intra-carrier calls as mobile-to-mobile calls, initiating the TIA/EIA-41 LocationRequest INVOKE (LOCREQ), for example.
- In countries where wireless carriers are assigned entire region or area codes, all carriers can maintain lists of the directory number prefixes for other carriers.
- Number Portability is not only a potential problem for optimal routing, but also a potential solution. If it is extended nationwide (and that is a big if), queries will need to be moved to the originating switch and the LNP database could include information about the type of MAP supported, facilitating optimal routing.
- Enhanced Roamer Agreement Tables. These existing tables can be enhanced to support optimal routing for mobile-to-mobile calls.
- ISUP *Release to Pivot*. This SS7 ISUP protocol has promise for allowing optimal routing for all types of calls — without number portability problems — although current protocols are not yet flexible enough.

Optimal Routing, Part II: Optional Routes

The purpose of Optimal Routing, as described in Part I, is to route calls from their originating point directly to the terminating mobile without passing (through) Home. There are several options for the route taken to accomplish this; these vary in important aspects, such as their reliance on new network capabilities, their ability to handle calls not originated by a mobile and their compatibility with existing systems.

Using MAP

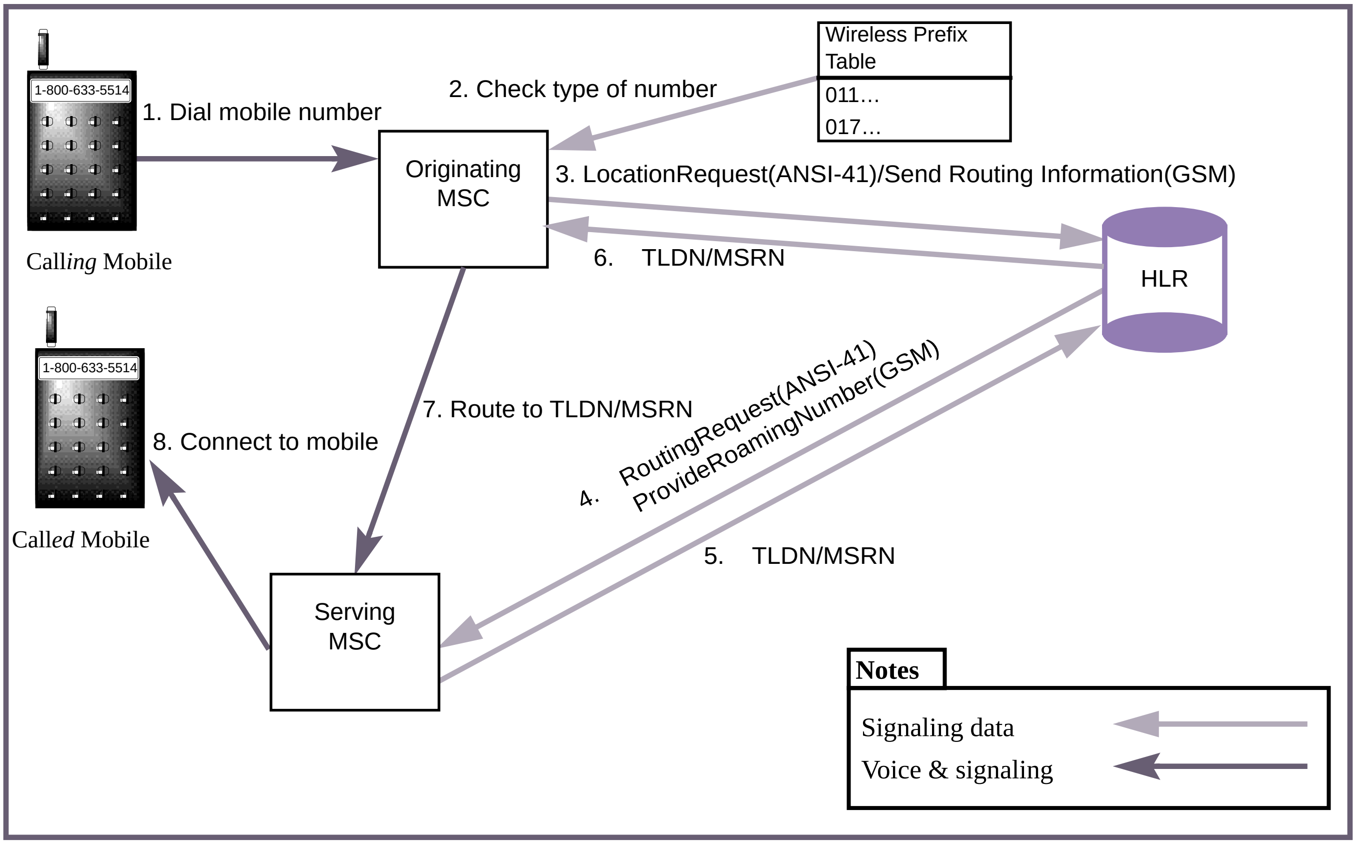
Optimal routing using MAP (GSM or ANSI-41) has long been an option, but it is not often implemented because it requires the system from which a call is being made to be able to recognize the dialed digits as mobile and because it can only optimize mobile-to-mobile calls. In countries where wireless systems are assigned a unique dialing prefix digit, recognition is easy to implement, and it may also be possible within a single carrier network in countries where wireless numbers are not so easy to separate out (such as the United States).

Figure 5 illustrates the use of either GSM or ANSI-41 MAP for optimal routing. The originating MSC has to be able to recognize mobile numbers (step 2), and it must support standard MAP call delivery transactions (steps 3-6).

When a routing number is obtained (known as a TLDN (Temporary Local Directory Number) in ANSI-41 and as an MSRN (MS Roaming Number) in GSM), the call can then be routed directly to the current serving MSC (step 7) and to the mobile (step 8).

This solution is only applicable to mobile originating calls (or the currently unlikely case of a landline switch that supports a subset of MAP), and it assumes wireless numbers cannot be ported (whether to another wireless carrier or to a landline carrier).

Figure 5: Optimal Routing using GSM or ANSI-41 MAP



Number Portability

Number portability already requires the ability to query a database (NPDB), to obtain the LRN (Location Routing Number) for ported phone numbers (see our May, June and July 1999 issues for a discussion of number portability). A minor modification to this database would enable the type of number (e.g. ANSI-41 MAP, GSM MAP, other) to be included in the response to the query, whether the number was ported or not.

Apart from requiring modifications to the number portability database, it would also require modifications to the protocols that query it (e.g. the ANSI-41 NumberPortabilityRequest message), it would require queries outside the local portability region, and it would require all blocks of numbers to be included in the database, even those that contain no ported numbers.

Figure 6 shows how optimal routing could use number portability queries. The major difference from the standard MAP method (highlighted in the shaded oval area) is that an external database (NPDB) is queried instead of an internal table.

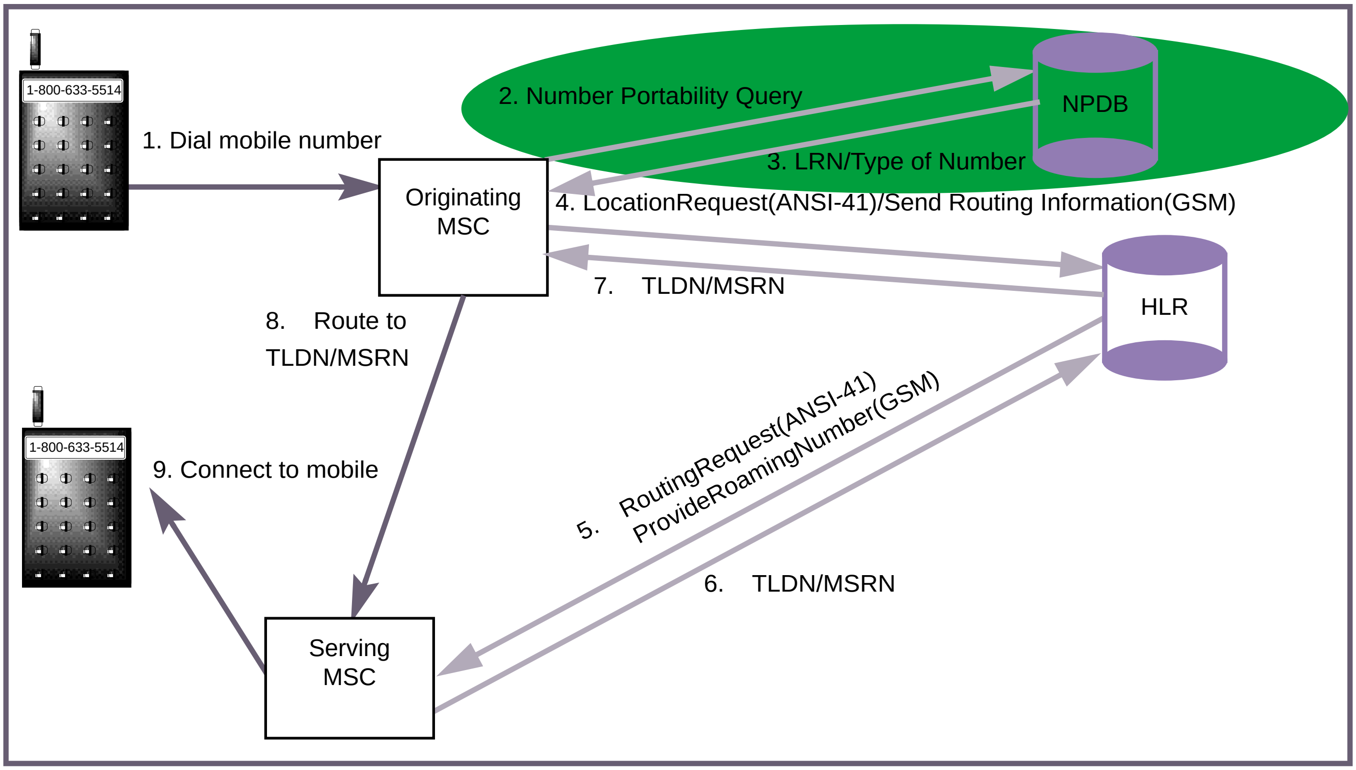
A big reason why this solution is a long shot is because US wireless carriers are still hoping to avoid having the number portability mandate applied to them. Also, the portability infrastructure is so unwieldy, any non-essential change is likely to be rejected.

Even if implemented, this method of optimal routing would only work when both the originating and terminating switches supported number portability, and, in particular, it would only work when this new database field is supported. In addition, it would only work when the originating switch is able to determine the network address of the NPDB for the terminating number.

Enhanced Roamer Agreement Tables

When a roaming mobile makes a call in ANSI-41 cellular or PCS systems, it provides its MIN or IMSI to the serving system (or it provides a TMSI, which can be mapped into a MIN or IMSI). This identifier allows the HLR to be

Figure 6: Optimal Routing using Number Portability



addressed, which then allows the presence of the mobile to be reported through a RegistrationNotification operation, which allows future call delivery.

These Roamer Agreement Tables are already close to providing the information needed to determine whether or not mobile originated calls are to another mobile on a compatible network. When the MIN of a mobile is the same as its directory number (MDN), the corresponding record from the Roamer Agreement Table could be used directly to obtain the network address of the HLR, to which a LocationRequest could be sent.

In general, the MIN of a mobile is not the same as the MDN, and even when a simple mapping between the two identifiers exists, it is not the same in every country. Consequently each entry in this table would (should) be updated with the corresponding MDN prefix for call delivery as well as the MIN prefix for originations. The MDN prefix identifies a group of mobiles sharing a contiguous block of numbers (e.g. the prefix 403-870 could be used to represent the ten thousand mobiles with numbers from 403-870-0000 through 403-870-9999).

Figure 7 illustrates how enhanced roamer agreement tables could facilitate optimal roaming. This method is similar to the first method described, but it is easier to support in countries with complex mobile numbering plans — such as the United States — because it works by enhancing an existing data table, rather than by creating another large table, with its corresponding huge data management burden.

The critical difference between the first two methods (highlighted by the shaded oval area) is the Roamer Agreement Table queried at step 2. This table does not need to store the type of number, since all numbers in this table are mobile (ignoring the impact of number portability). The absence of an entry identifies a number to which this method of optimal routing does not apply.

This method only works with mobile originated calls, it is not fully compatible with number portability and it is not compatible with GSM systems, which do not usually keep a Roamer Agreement Table.

Release-to-Pivot (RTP)

The most general purpose method for implementing optimal routing takes a completely different approach. Because of this, it can apply to land-to-mobile calls and it is not affected by number portability.

Release-to-Pivot (RTP) is a general concept allowing a call routed to one switch to be redirected to another,

allowing the trunking to be re-originated from the beginning point of the call. This can optimize trunking for call forwarding, for directory assistance with call completion, and for wireless optimal routing.

Figure 8 shows how Release-to-Pivot can be used to implement optimal routing. The originating switch indicates its ability to support RTP in an SS7 ISUP IAM call setup message. The Home MSC performs the normal ANSI-41/GSM process to obtain a temporary routing number (steps 3-6). However, instead of directly routing the call to the Serving MSC, the Home MSC releases the incoming trunk (at step 7) – with a caveat – the call must be reoriginated to the number provided (the temporary

Figure 7: Optimal Routing using Enhanced Roamer Agreement Tables

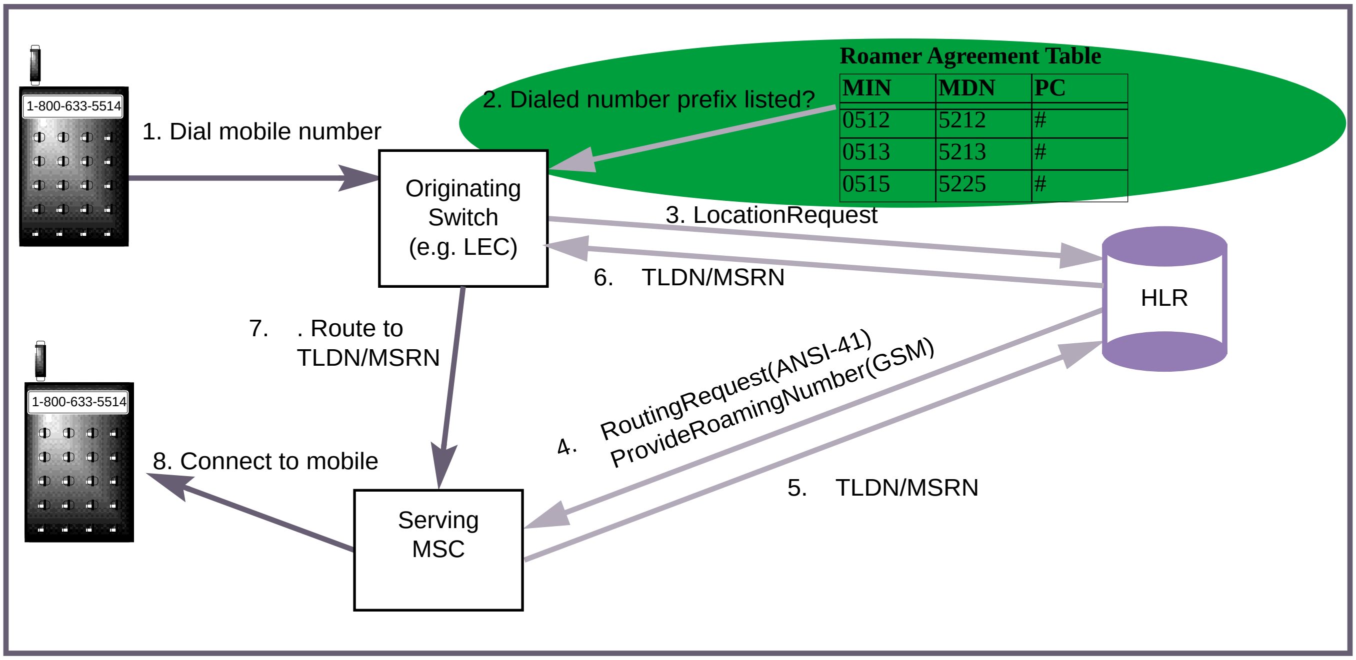
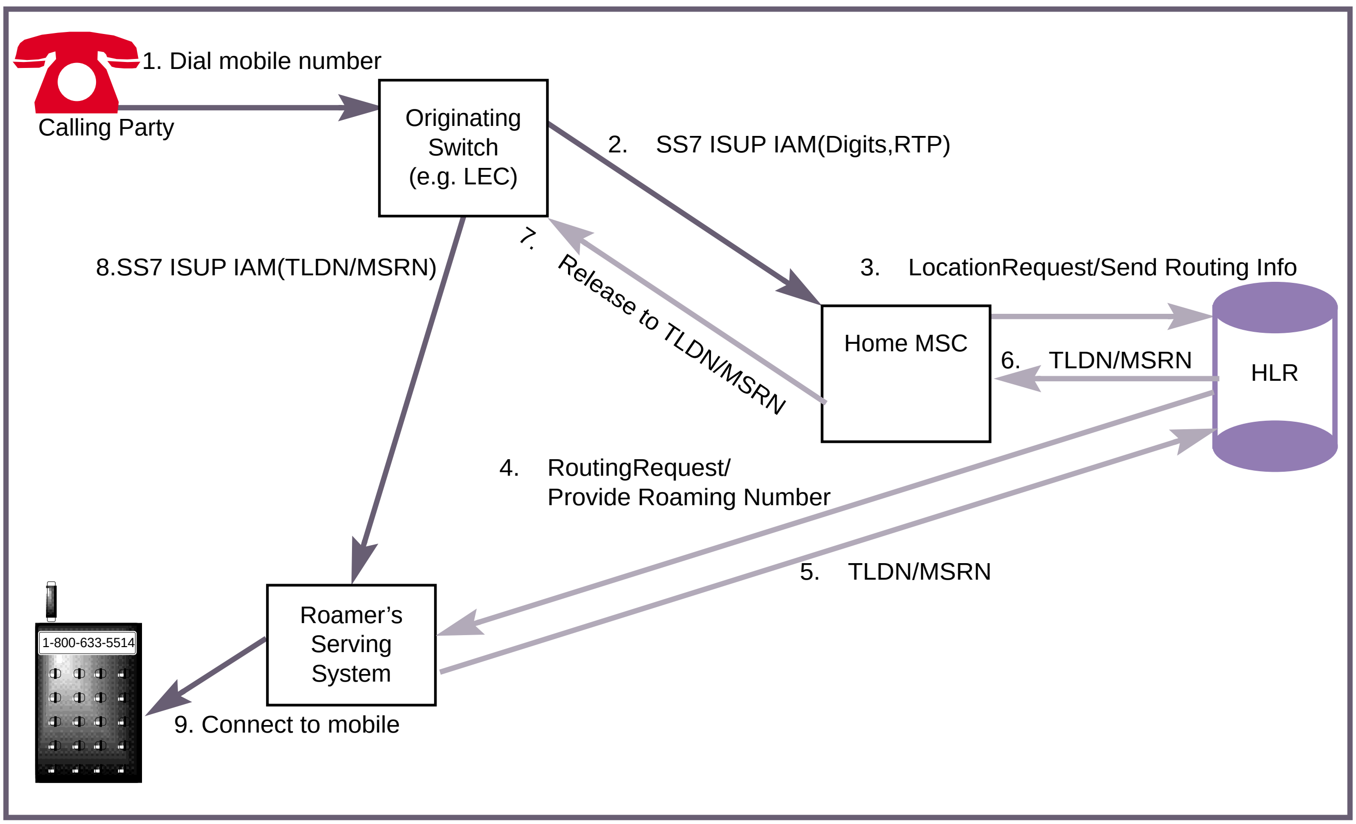


Figure 8: Optimal Routing using Release-to-Pivot



routing number, TLDN or MSRN). The originating switch can then extend the call to the Serving MSC. If, for any reason, it does not want to perform this function (e.g. lack of a billing agreement), it can reject the release, and the Home MSC will have to extend the call normally, although without optimal routing.

Release-to-Pivot has one serious limitation – it fails to work with some types of call forwarding. If a call is forwarded to a serving MSC through release-to-pivot, but the mobile does not answer, does not respond to a page or is busy, it is now impossible to route the call back to the home system where a new route can be selected (e.g. call forward no answer number). To overcome this problem, an enhanced form of Release-to-Pivot will be required, which we call Conditional Release-to-Pivot (cRTP).

Optimal Routing, Part III: The Ultimate Solution

The first two parts of this article on Optimal Routing described some partial solutions, but they did not describe all that is needed to make it work even when the caller is using a landline phone, when the mobile is ported and when call forwarding is considered.

The closest solution is based on SS7 ISUP Release-to-Pivot (RTP). It is, however, deficient in certain call forwarding cases because the connection to the home MSC is released prematurely. A slight modification, which we call Conditional Release-to-Pivot (cRTP), maintains the trunk to the Home MSC until either the roaming mobile answers or until call forwarding occurs. It is important to note, this solution is not yet under development by any standards organizations.

Conditional Release-to-Pivot (cRTP)

There is a period of time during a call to a roaming mobile when it is not known whether the call will be completed to the mobile or forwarded under the control of the home system. Figure 9 illustrates how cRTP can provide Optimal Routing while allowing for this period of uncertainty.

Steps 1 and 2 of this figure show normal call establishment to the Home MSC. At Step 3, however, while a route is established to the Serving MSC, the route to the Home MSC is still maintained (although it is not part of the voice path). This requires a new message which we have named cRTP.

Once a three-way call path is established, two different possibilities open up. Step 4a illustrates that, if the mobile answers (resulting in an ISUP ANM to the Originating Switch), the trunk to the Home MSC can be released normally and the Serving MSC will continue with the call for the roamer.

If, however, the mobile does not respond to a page, if it is busy, or if it does not answer after being paged, the mobile may be redirected to a call-forward busy/no-answer number, as shown in Step 4b. The Serving MSC sends an ISUP REL message to the Originating Switch at about the same time it sends an ANSI-41 RedirectionRequest to the Home MSC, which will then query the HLR for the forward-to number.

This type of call processing is not yet possible with GSM systems; forwarding would have to occur from the Serving MSC back to the Home MSC, and this would negate the benefits of optimal routing.

Figure 9: Optimal Routing using Conditional Release-to-Pivot (cRTP)