Running head: Identification of missing neighbor cells in GERAN

Title: Identification of missing neighbor cells in GERAN

Authors: M. Toril, V. Wille, R. Barco

Authors' affiliation:

V. Wille (Nokia Networks, Performance Services, Ermine Business Park, Huntingdon, Cambs. PE29 6YJ, UK)

M. Toril, R. Barco (Communications Engineering Dpt., ETSIT, University of Malaga, 29071 Málaga, Spain)

Corresponding author's address:

Volker Wille

Nokia Networks

Ermine Business Park

Huntingdon

Cambs

PE29 6YJ, UK

Tel: +44 7802 690 892

Fax: +44 1480 44 4246

e-mail: volker.wille@nokia.com

Identification of missing neighbor cells in GERAN

M. Toril, V. Wille, R. Barco

Abstract*:* In GSM/EDGE Radio Access Networks (GERAN), the number of neighbor cells defined per cell for handover purposes is limited. Due to the large number of cells and, consequently, neighbor cells in the system, neighbor cell management proves a complex and time-consuming process during network operation. Hence, it is very likely that, on some cells, important neighbor cells are missing. This paper outlines a new approach for determining these missing neighbors, based only on information from the network management system. Thus, the need for propagation data from network planning tools is circumvented, which makes the method especially suitable for the operational stage of an existing network. Results from the application of the method in a live network show its capability to spot such missing neighbors. Although performance benefits from the method are rather limited for the entire network, significant benefits can be obtained on individual cells.

Keywords: mobile communications, network optimization, automation, adjacency management

Abbreviations

BAL	BCCH Allocation list
BCCH	Broadcast Control CHannel
EDGE	Enhanced Data Rates for Global Evolution
GERAN	GSM/EDGE Radio Access Network
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
IM	Interference matrix
KP	Knapsack problem
HO	Handover
MS	Mobile Station
NMS	Network Management System
RP	Relative Position
UTRAN	UMTS Radio Access Network
UR	Utilization Ratio

Introduction

The mobile telecommunication industry is evolving at an extraordinary rate due to the changing nature of services offered by mobile operators, the introduction of new technologies and the increased level of competition. Within this scenario of increasing complexity, efficient network management is a crucial issue to provide high-quality services at a low operational cost. Until a few years ago, most operational tasks in cellular networks were primarily based on manual processes. However, operators are currently showing a growing level of interest in automating most network management activities in order to increase operational efficiency, i.e. reduce costs of running the network. This has stimulated intense research activities in the field of self-regulated networks, which mainly involve automation and optimization [1][2][3][4].

One of the areas where automation can be applied is neighbor cell definition. In GERAN, the handover (HO) process normally ensures that any mobile station (MS) is connected to the cell that offers the best coverage. To perform HO decisions, an MS must provide the base station controller with information of the received signal level from surrounding cells. The frequencies on which an MS performs adjacent cell measurements are given by the Broadcast control channel Allocation List (BAL) [5]. This group of frequencies defines a set of candidate cells to be monitored, which is referred to as *neighbor cell list*. Thus, for handovers (HOs) to occur between a pair of cells, the target cell has to be defined as a neighbor cell of the source cell. Otherwise, a HO cannot take place even if both cells are very close in propagation terms. More importantly, only a limited number of neighbor cells can be defined for each cell in the network. While the maximum number of neighbors per cell is 32 [5], operators try to use as few as possible to improve manageability and reduce the number of frequencies to be measured by mobiles in active mode [6]. This restriction turns the definition of the neighbor cell list into a selection problem.

The problem of defining a neighbor cell list has already been covered in the literature from the perspective of planning procedures. A priori, any method to generate an interference (or inter-cell dependency) matrix can also be used to decide whether an adjacent cell should be added to an existing neighbor cell list. During the past few years, most automatic planning tools have used empirical [7] or semi-empirical [8][9] propagation predictions models. The disadvantage of this approach is that the results provided by such methods are often not accurate enough to determine the correct adjacencies. This observation holds especially true in dense urban areas. In the latter scenario, it is common that antennas are below the rooftops of the surrounding buildings, which causes the near environment of the antenna to have a strong effect on the propagation of radio waves. Likewise, subscribers are normally concentrated in elevated indoor environments, which are inherently difficult to model in propagation predication tools. Even deterministic methods, such as ray-tracing modeling [10], fail to give adequate accuracy due to limited resolution of the available terrain and clutter data as well as uncertainties in the propagation mechanism [11]. In the operational stage, mobile measurements can be used to refine propagation predictions [12][13][14]. Unfortunately, the latter methods rely on complex tools that are not always available for operators. As a consequence of a lack of accurate tools, determining missing adjacencies is not carried out frequently during the operational stage of the network. Hence, it is likely that some important adjacencies might not be defined for certain cells because either they have been missed out at the planning stage, they have been deleted by mistake or the network has just simply evolved during the operational stage.

An inadequate neighbor cell plan can be a major contributor to network quality problems. Obviously, the impact of a missing neighbor cell on network performance depends on the existing coverage available from other cells in the same area. Under certain circumstances, a missing neighbor definition can have a significant performance impact, especially if few other cells provide coverage and the missing adjacency would have been the main HO target.

These facts indicate the need for an automated process that analyzes the current neighbor cells and highlights possible missing ones for any operational cell. In this paper, a new approach for determining these missing adjacencies during the operational stage is proposed. In contrast to the current approach, the main novelty of the proposed method is that it relies on data residing purely in the network management system,. The method is validated by means of field trials in a live GERAN. Thus, the main contribution of this work is to prove the limitations of the current approach and how these can be solved by a simple algorithm, The paper is organized as follows. The next section describes in detail the adjacency-selection problem. The following section presents the novel method to carry out this task. Finally, the subsequent sections discuss the results from the application of this method in a public GERAN.

Outline of the neighbor cell definition problem

In this section, the definition of a neighbor cell list is first formulated as a selection problem. The solution technique currently in use in live networks is then presented. Thus, the limitations of both the approach and the tools currently available for operators are highlighted.

Definitions and notation

In GERAN, each cell has a list of neighboring cells to which a HO is possible. Such a list is referred to as *neighbor cell list* or *adjacency list*. Formally, an *adjacency* (i,j) refers to the definition of cell j as a *neighbor* of cell i. In this work, the terms “neighbor cell” and “adjacency” will be interchangeably used. While the former will refer to the cell that receives HOs from the original cell, the latter will refer to the entity used in network management tools to reflect that HOs are allowed between a pair of cells. Hereafter, lists will be denoted by capital letters and its elements (i.e. either adjacencies or neighbor cells) will be denoted by lower case letters. Likewise, the operator $|\bullet|$ will denote the length (or cardinality) of a list.

It is worth noting that the previous definition of adjacency only reflects HOs in one direction (i.e. from source cell i to target cell j), This type of adjacency is known as *1-way* (or *uni-directional*) adjacency, to differentiate it from adjacencies that comprise HOs in both directions, which are referred to as *2-way* (or *bi-directional*) adjacencies. This work deals with the former type of adjacency, as both HO directions are handled independently in network management tools.

Not all adjacencies defined by the operator are actually used for HO purposes. On the contrary, it is rather frequent that some of them remain unused. Thus, an adjacency is said to be *active* if any HO takes place from the source to the target cell. Likewise, not all active adjacencies of a cell are equally used. Thus, an adjacency $(i.j)$ is said to be more significant than an adjacency $(i.k)$ if the likelihood that a user in i moves from $i \rightarrow j$ is larger than the same indicator from $i \rightarrow k$.

Formulation of the problem

Theoretically, any cell in the system can be defined as a neighbor of any other cell as long as the length of the neighbor cell list is below a certain limit. This constraint converts the definition process into a combinatorial optimization problem known as *selection problem* [17].

In the classical approach, the final purpose of this selection process is to minimize the number of user connections that experience bad quality. As a complementary objective, the management of a minimum number of adjacencies is desired to reduce configuration workload. Hence, the objective function to be minimized should include two different costs, reflecting the cost associated to bad quality connections and overall number of adjacencies. Unfortunately, a linear combination of these costs is not very meaningful, since these costs do not have comparable units. Thus, operators find it difficult to specify the relative weight of each cost. For that reason, the number of adjacencies is usually included in the formulation as a constraint. As a result, the problem of adjacency selection in each cell can be considered as a special case of the *0/1 Knapsack Problem* (KP), which has been extensively covered in the literature (for a survey, see [18][19]). Originally, the conventional KP gets its name from the common situation where a hitch-hiker has to select among various objects to fill up his knapsack, maximizing the profit. The 0/1 prefix emphasizes that a fraction or multiple instances of an object are not selectable.

For each cell in the network, the KP is defined by a finite set $N = \{n_1, n_2, ..., n_k\}$ of cells that are potential candidates for selection as neighbors, and the maximum number of neighbor cells in the list C (i.e. the knapsack capacity). Each candidate cell n_i has a profit p_i and a penalty w_i from its inclusion in the neighbor cell list. The profit p_i reflects the enhancement of network performance in terms of signal-level or signal-quality (e.g. the number of connections that experience bad connection quality) when cell n_i is added to the list. The penalty w_i is simply the number of positions in the list occupied by the new neighbor, i.e. w_i =1. Then, the problem calls for selecting the set of cells with a maximum profit P among those sets that do not exceed the length limit of the neighbor cell list C. The integer linear programming model of the 0-1 KP problem can be formulated as the search of a vector X of binary (decision) variables x_i (i=1, …, k) with the meaning

$$x_i = \begin{cases} 1 & \text{if adjacency is selected,} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

that maximizes the objective function P (profit)

$$P = \sum_{i=1}^{k} p_i x_i \tag{2}$$

subject to the constraint

$$\sum_{i=1}^{k} x_i \leq C \tag{3}$$

From the previous definition, it is clear that the benefit from the ascription of each candidate neighbor cell to a neighbor cell list is not fixed, but might depend on the existence of other candidates in the list. Hence, profit values p_i are dependent on the decision variables x_i, converting P into a non-linear objective function. Thus, the problem can be classified as a 0/1 non-linear KP.

Current Solution Technique

Operators of GERAN networks commonly use propagation predictions from network planning tools to identify which adjacencies should be included in the adjacency list of every new cell added into the network. For the sake of simplicity, these algorithms do not directly optimize any network-quality indicator, but try to maximize coverage probability in every cell in the network, based on the construction of interference matrices (IM) [12][13][14][15][16]. Each element x_{ji} in the IM is an indicator of the average signal level (i.e. potential interference) received by users attached to cell j from cell i. Thus, it is common that the higher x_{ji}, the larger the profit p_i associated to the definition of cell j as an adjacency of cell i. Nonetheless, in situations of strong cell overlapping, once the coverage is assured with an initial set of adjacencies, the benefit from the subsequent addition of new adjacent cells might be limited (or even negligible). Such situations are rather common in dense urban areas, where tens of cells provide coverage on the same location, due to the multi-layered structure of the network and the changing propagation conditions. This dependency of profits from adding an adjacency on the previous state of the list justifies the formulation of the problem as a non-linear KP.

A large number of algorithms have been proposed in the literature for exact and approximate solutions to the KP problem [20][21][22]. Among them, the classical *greedy* algorithm [23] is commonly chosen by operators for the sake of efficiency. In this approach, the adjacencies are ranked and selected based on estimations of received signal-level on the area serviced by the cell under optimization (i.e. IM coefficients). The greediness property refers to the fact that, once a candidate has been included in the new list, its appropriateness is not evaluated again. Apart from its simplicity, the main advantage of the greedy algorithm is its low time complexity, which makes it ideal for large problems as the one considered here. For a set of n candidate cells, the greedy algorithm has polynomial time complexity $O(n)$ in contrast to exponential $O(n2^n)$ and pseudo-polynomial[1] $O(nW)$ time complexities of branch and bound [24] and dynamic programming [25] algorithms, respectively. However, unlike other algorithms, there is no guarantee that the optimal solution is obtained, especially for the non-linear KP. In addition to this shortcoming, the inaccuracy of the propagation prediction models used and the difficulty to account for the spatial distribution of users lead to sub-optimal solutions when applying this approach to live networks.

To compensate for these limitations, operators commonly implement a large number of adjacencies for every newly integrated cell. As a result, some of the adjacencies are infrequently (or never) used and must be removed during the operation of the cell. For this deletion process, it is common not to involve the planning tool, since the Network Management System (NMS) is able to carry out both the identification of under-utilized adjacencies and their automated deletion based on network statistics. Even though the use of network measurements generally enhances the quality of the current solution, it might be the case that deletion decisions proved wrong over time due to network evolution. Changes in propagation conditions (e.g. antenna downtilt, new shadowing obstacle) or user spatial distribution (e.g. opening of shopping mall) are such examples of system evolution. Therefore, an automatic tool is needed to identify possible missing adjacencies network wide. Unfortunately, several reasons prevent operators from using their planning tools to cope with these ongoing changes. First, hardly any operator updates the deleted adjacencies in the planning tool, because the implementation of these changes on the tool is a very time-consuming and tedious task. Similarly, adjacencies that are manually added by network optimization engineers (e.g. to solve localized coverage problems) are generally not reflected in the planning tool either. Consequently, adjacencies on the planning tool do not reflect the actual adjacencies in the network, since this would require synchronizing the planning tool with the NMS, which for most operators is not an automated task. This lack of synchronization, together with the poor accuracy of predictions, causes operators to very seldom use their planning tool for the systematic identification of missing adjacencies on existing cells.

Due to limitations of the current approach, it is desirable to provide operators with a simple NMS-based solution, which does not require any propagation-based information, but is able to rapidly analyze the existing adjacencies in the network and determine if any obvious ones are missing. The need for efficiency is better understood when it is observed that 25 intra-system adjacencies are defined per cell on average in GERAN. Thus, a typical GERAN network with 20.000 cells would have about 500.000 adjacencies. In this context, the speed of the method proves critical, since the higher the time complexity of computation, the less the tool will be used

[1] W can be exponentially large to n, e.g. 2^n

as it increases the load on the NMS. As a result, operators might be forced to limit either the geographical area or the number of times this method can be applied in the network. Ideally, the method to identify obvious missing adjacencies should be run once per week for the entire network to ensure that any possible mistakes in the adjacency definition are spotted fairly quickly and can thus be rectified. For this purpose, the method described in the next section is proposed.

Novel Methodology

This section presents a simple approach to solve the adjacency-selection problem based solely on NMS data. Since it is clear that a simple methodology as the one outlined here will not be able to spot all missing adjacencies in one attempt, some possible refinements are also discussed.

Formulation of Proposed Methodology

The proposed greedy algorithm is based on the comparison of two adjacency lists for a given cell. The first list, referred to as *nominal adjacency list* P, contains the potential adjacencies, ranked in order of significance. The second list, referred to as *actual adjacency list* A, contains the current adjacencies. Subtracting the nominal from the actual list returns the adjacencies in the nominal list that are not in the actual list, referred to as *missing adjacency list* M. In set theory notation, let (P; ≥) be an ordered set of adjacencies where the significance comparison relation ≥ has been defined. The set difference $P \backslash A$, referred as M, is defined as

$$M = P \setminus A = \{x \mid x \in P, x \notin A\} \tag{4}$$

where (M; ≥) is also an ordered set. While A can easily be obtained from the current network configuration in the NMS, the definition of M is the core of the algorithm. For that purpose, geographical information from the network residing in the NMS, such as cell/site coordinates and antenna bearings, is used. Finally, the new adjacency list, referred as *updated adjacency list* U, is built as the combination of A and M, i.e. $U = A \cup M$.

The exact procedure can be summarized in the following steps:

1) *Identification of potential missing adjacencies on a cell basis*: The simplest approach for selecting potential adjacencies for a source cell is to determine the distance to all surrounding cells and rank them by distance in ascending order (i.e. significance of potential adjacencies would be given by proximity between source and target cells). Even though this approach might be suitable for networks deploying mainly omni-directional antenna, in a network deploying sectorized sites it is beneficial to consider the antenna bearing of both source and target cell. Therefore, the maximum angle off the antenna azimuth direction might be limited to a certain value, referred as *catchment angle*, so that the distance ordering criteria is complemented by a directional constraint. Likewise, it is operator strategy to implement only bi-directional

adjacencies (i.e. whenever adjacency ($i.j$) is suggested, adjacency (j,i) is also created). To satisfy this requirement, potential neighbors of a source cell are those target cells that meet the following criteria:

a) The target cell is located within the catchment angle of the source cell

b) The source cell is located within the catchment angle of the target cell

where the catchment angle is dependent on antenna pattern. For omni-directional cells, an angle of 360º is selected. For sectorized cells, an angle of 180º, centered in the maximum radiation direction, is chosen.

To clarify the catchment angle concept, Figure 1 (a)-(b) show the potential neighbors of two cells in a live environment. Both figures show base stations in a network area of 100 km^2. Base-stations with omni-directional antennas are shown by a circle, while those with tri-sectorized antennas are shown by a line in the direction of maximum radiation. In every case, potential neighbor cells are represented by a bold line, while discarded cells are represented by a normal line. The area enclosed in the catchment angle of the original cell is highlighted in grey. Figure 1 (a) shows the potential neighbors of a tri-sectorized cell in the coordinates (10.15, 12.72), whose main direction is 120º azimuth (i.e. measured clockwise from North). From the shaded area, it can be deduced that the cell catchment angle comprises all angles between 30º (i.e. 120-90) and 210º (i.e. 120+90) azimuth. This 180º interval is clearly larger than the theoretical value of 120º (i.e. 360º/3) assigned to a tri-sectorized cell. At the same time, it is observed that all potential neighbors lie within the previous catchment angle. Conversely, it can be checked that the source cell also lie within the catchment angle of every potential neighbor. It is worth noting that, while all omni-directional cells in the shaded area are potential neighbors of the original cell, only those tri-sectorized cells whose main direction is between 210º and 30º azimuth are potential neighbors. More subtle is the fact that cells that share site with the original cell are also potential neighbors, as they fulfill the two catchment angle criteria. The latter is observed in the co-sited neighbors of the original cell with directions 0 and 240º azimuth. Figure 1 (b) shows the potential neighbors of an omni-directional cell in the coordinates (6.25, 13.25). The whole area is now shaded as the catchment angle of the source cell is 360º. From this observation, it can be deduced that all omni-directional cells are potential neighbors of each other.

From the previous figures, it is clear that the ordered set M resulting from these criteria could be arbitrarily large. To restrict the number of new adjacencies, the first k elements of M must be selected, where k is a parameter that must satisfy that $1 \leq k \leq 32\text{-}|A|$.

2) *Discard of potential adjacencies with frequency problems*: Transceivers of neighbor cells should not use the same frequency due to interference reasons. This constraint is critical in the Broadcast Control Channel (BCCH) transceiver. To avoid this issue, potential neighbor cells that share the same BCCH frequency with the source cell are discarded.

3) *Validation of newly added adjacencies*: As in the past, once new adjacencies are configured in the network, their suitability is assessed based on their actual usage, reflected by HO statistics.

It is worth noting that all the information needed for the execution of the proposed algorithm (i.e. cell coordinates and antenna bearings for tri-sectorized cells) and its later assessment (i.e. HO statistics) is available in the NMS. Hence, no additional data from network planning tools is required.

Refinements to Outlined Method

Some benefit might be obtained by repeatedly applying the method within the same geographical area of the network and thereby incorporating a feedback mechanism into the method. For example, once the suggested changes in the adjacency list are implemented, the utilization of the new adjacencies is validated based on network measurements. Thus, whenever newly created adjacencies are not used for HOs within a given time period (e.g. a week), they can easily be removed from the network. To ensure that removed adjacencies are not created again, these are "blacklisted". This blacklist contains those target cells that will never reappear in the nominaladjacency list of a given source cell. With this feedback mechanism, next time the method is re-applied in an area, different adjacencies will be suggested. Hence, over time, a multitude of adjacencies can be tested for their suitability to carry HOs and only the most suitable ones remain as defined adjacencies.

Field Trial Results

The above-described methodology was applied to a public GERAN network in order to test the effectiveness of the algorithm in a real environment. For the sake of clarity, the preliminary conditions are first described, trial results are subsequently discussed and implementation details are finally presented.

Trial Set-Up

The trial area consisted of 9.120 cells providing seamless coverage. On these cells, about 165.000 adjacencies already existed, averaging 18.2 adjacencies per cell. A priori, it was envisaged that this large number of adjacencies in the initial plan would leave little space for improvement of the adjacency lists, i.e. few important adjacencies would be missing. This opinion was reinforced by the preliminary analysis that showed that only 14.7 of those 18.2 adjacencies were used on average in the initial adjacency plan during a 2 day period (i.e. 20% of the adjacencies remained entirely unused).

From the application of the method, potentially missing adjacencies were identified and created in the network. Network statistics were gathered for a period of two working days after the new adjacencies were implemented in the network. Validation of the method was carried out based on these statistics.

The aim of the trial was two-folded: on the one hand, to test the capability of the algorithm to spot missing adjacencies; on the other hand, to prove the benefit of a simple algorithm that relieves the operator from tedious work of identifying missing adjacencies. It is worth noting the independence of both objectives, since even if the performance of the new and current solution coincided (which is not actually the case, as will be shown later), the benefits from a simple, fast and automatic approach would justify its application.

Assessment methodology

A priori, the most intuitive way to quantify the performance gain from the addition of missing adjacencies is through the analysis of signal-level and signal-quality measurements before and after the optimization process. Although this approach proves convenient for simulation studies [15][16], it is not the most suitable one in live networks. In the latter case, statistical noise inherent to these measurements might mask subtle performance changes caused by small extensions of existing adjacency lists. In addition, these measurements prove rather sensitive to other adjustments that might be carried out by optimization engineers during the trial period. In that sense, it is worth noting that due to the large trial area (i.e. the whole network) it was impossible to stop all other network configuration changes in the network during the trial period. Hence, the utilization of the new adjacencies for HO purposes was instead considered as the main figure-of-merit to validate the method, as suggested in [1]. Table 1 presents an example of how the utilization of a newly configured pair of adjacencies is checked based on network measurements. In the example, cell A has 2496 HOs to cell B, which equates to 88% of its entire outgoing HO activity (i.e. 2836 HOs). Similarly, there are 2670 HOs in the opposite direction, which is 13.1% of cell B outgoing HO activity (i.e. 20426 HOs). The previous HO rate will be referred as *utilization ratio* (UR) of the adjacency. From these, it can be concluded that the addition of this adjacency is beneficial for both cells, since the new adjacency pair is well used in both directions.

Another important issue is the fact that HO counters in NMS, used for the calculation of URs, include both inter-cell and intra-cell HOs. While the former type of HO is caused by mobility reasons, the latter type reflects the change of time slot due to interference reasons, which is a common feature supported by several manufacturers. This intra-cell HO mechanism might be viewed as an outgoing adjacency where the target and the source are the same. Since these adjacencies need not be defined in the adjacency list (and do not lead to additional frequency measurements by handsets, as they are already performed as serving-cell measurements), these adjacencies do not compete for a position in the list. Hence, intra-cell HOs must be eliminated from HO statistics. In the following section, the need for this correction is justified by the existence of a non-negligible overall intra-cell HO ratio.

Despite the significance of new adjacencies is reflected by their URs, assessment based on this indicator tends to show limited benefits from the method in a well-planned network. This trend stems from the fact that most HOs are normally carried out by few cells. Thus, the addition of a new adjacency in a list is likely to give negligible performance improvement once a certain list length is reached. In other words, the initial configuration tends to be quite close to the performance optimum with only a few defined adjacencies. Nonetheless, since a missing adjacency would be any new adjacency that had higher significance than the ones in the current list, comparison of the old and new adjacencies is still desirable. For that purpose, the adjacencies after the optimization process are ranked by utilization on a per-cell basis and the position of the new adjacencies is evaluated.

For a fair comparison between adjacencies of different cells, the assessment process must take into account the length of the adjacency list. Thus, it is obvious that an adjacency in the i^{th} position of a list of i adjacencies is less significant than an adjacency in the same position of a list with twice the number of adjacencies. While the former is the least significant of its list, the latter is more significant than half of the adjacencies in its list. Hence, the assessment methodology must consider the relative (rather than the absolute) position in the adjacency list to cope with the variable length of the adjacency list. For that purpose, the *relative position* (RP) is defined on an adjacency basis as

$$RP_{i,j} = \frac{l_j - 1}{N_i - 1} \qquad 1 \le j \le |M_i| \ ; \ 1 \le l_j \le N_i \ ; \ N_i = |U_i| \ ; \tag{5}$$

where i is the source cell to which the adjacency list belongs, j is the new neighbor cell whose significance is assessed, l_j is the position of the adjacency with neighbor cell j, N_i is the number of total adjacencies defined in cell i, and M_i and U_i are the missing and updated adjacency lists, respectively. Intuitively, the previous indicator gives the ratio of adjacencies in the same adjacency list that are more significant than the one considered. Thus, RP is 0 when the new adjacency is the first one of the list (i.e. l_j=1), while 1 when the last one (i.e. l_j= N_i).

Trial Results

The overall results of the trial are presented in Table 2. For the sake of clarity, a division line is used to reflect the situation before, during and after the optimization process. From the application of the method, 9052 new 1-way (4526 2-way) adjacencies were automatically created in the network, resulting in a 5.5% (i.e. 9052/165898*100) increase in the overall number of adjacencies. After such a small relative increase (i.e. 0.99 new adjacencies added to existing 18.2 adjacencies per cell on average), small influence on network quality was expected. The use of utilization ratios as the main figure-of-merit is thus justified.

After the optimization process, the overall intra-cell HO ratio in the network proved small (i.e. 4.6%). Nonetheless, a per-cell analysis showed that this component might be significant on several cells. Figure 2 depicts the histogram of the intra-cell HO ratio on a cell basis, which has been classified into seven bins of variable length. In the Y-axis, the number of cells in each bin and the cumulative relative frequency are presented on the primary and secondary axis, respectively. In the figure, it can be observed that 30% (i.e. 100%-70%) of the cells experienced values of this ratio higher than 5%. As a consequence, this type of handover had to be eliminated from HO statistics.

After the optimization process, the overall share of inter-cell HOs performed through new adjacencies in the network was 1.63%. This low utilization figure was expected, due to the large number of existing adjacencies. For a fair evaluation, the previous indicator should be compared to the mean value of that indicator in the case that all adjacencies (whether old or new) were equally used in the system, i.e. 1/(18.2+0.99)*100=5.2%, From this comparison, it is clear that, in general, the added adjacencies are less used than current ones, but their significance is not negligible. The histogram of the distribution of utilization ratios of new adjacencies is presented in Figure 3. The histogram consists of 8 bins, where each value on the primary axis represents the number of new adjacencies whose UR falls within the bin limits. The cumulative relative frequency of each bin is presented in the secondary axis. For the sake of clarity, the unused adjacencies have been segregated and included in the first bin. Analysis showed that 4486 of the 9052 new 1-way adjacencies remained completely unused. At first glance, it might appear that the accuracy of the proposed method to determine missing adjacencies is not very good, since 49.6% of all added adjacencies were not used. However, it is worth noting that in the initial adjacency lists about 20% of adjacencies were not used, indicating that propagation-based planning tools do not necessarily lead to a perfect adjacency list. Likewise, it can be seen that, on certain cells, the added adjacencies had a significant impact, since they carried a large ratio of the HO traffic from these cells. On other cells, however, the newly added adjacencies do only carry a small percentage of the HO traffic. The fact that a relatively high percentage of new adjacencies carry relatively few HOs (i.e. 80% of new adjacencies carry less than 1% of outgoing HO traffic of their source cells) is mainly due to the large number of adjacencies already in existence.

It is obvious that utilization figures depend on the number of active adjacencies per cell. To prove this relationship, Figure 4 depicts the scatter plot of the UR of the most significant adjacency (i.e. the one with the largest number of HOs) versus the number of adjacencies in the cell. For the sake of clarity, Figure 5 shows how the average maximum utilization relates to the number of adjacencies per cell. A trend line is also superimposed, based on a logarithmic regression model, together with the 95% confidence intervals. Over this trend line, it is observed that the higher the number of adjacencies, the lower UR of the most relevant adjacency, to which all others are compared. Once this dependency is proven, it is shown that the number of active adjacencies is not the same for every cell in the network. Figure 6 represents the histogram of the number of active adjacencies per cell in the trial area after optimization. From the figure, it is clear that the number of active adjacencies per cell greatly varies from cell to cell. This analysis proves the need for breaking down utilization figures by the number of adjacencies, or constructing an indicator that is independent of the number of adjacencies. The RP indicator defined in eq. (5) complies with the latter requirement.

The analysis of the significance of new adjacencies in the network was performed based on the position in the adjacency lists to which they belong. In this comparison, only those old adjacencies that were actively used after optimization were considered. Thus, old adjacencies that were inactive after the optimization (i.e. 20% of the initial adjacencies) were discarded before sorting each adjacency list. Figure 7 depicts the histogram of RP values defined on an adjacency basis. The X-axis has been divided into 11 classes, which reflect the relative position in a list where adjacencies are ranked by significance. The number of cells in each bin and the cumulative relative frequency are presented on the primary and secondary axis, respectively. From the figure, it is observed that only 56% of these new adjacencies displayed RP=1 (last bin in the figure). Thus, 44% (i.e. 100-56%) of these are not the last ones in their list, and can be strictly defined as missing adjacencies. Likewise, a better adjacency list could be defined without any list extension if new adjacencies substituted current ones with less significance. From these results, it can be concluded that a better solution for the adjacency selection problem is easily obtained with the method proposed, regardless of its limited influence on overall network performance.

Finally, the influence of the propagation environment on the results is analyzed. For that purpose, two different scenarios were investigated, taking advantage of the geographical division in the NMS. The first scenario consisted of 1962 cells, which can be broadly classified as dense-urban. The second scenario consisted of 2403 cells, which is a mixture of rural, sub-urban and urban environments (later referred to as rural-urban). As performance results are segregated on a geographical basis (and not purely by propagation scenario), it is expected that only subtle performance differences are observed. In these circumstances, any biasing in the assessment process might spoil the conclusions of the analysis. To cope with this limitation, the confidence intervals of the most relevant figures must be carefully studied.

Table 3 presents results in both scenarios. The average URs of the new adjacencies were 3.13% and 3.08% for the rural-urban and dense-urban scenario, respectively. Likewise, the ratio of HOs performed through news adjacencies was 2.17% and 1.04% for the rural-urban and dense-urban scenarios, respectively. From these two indicators, it might wrongly be concluded that the best results are obtained in the rural-urban scenario, as the utilization of the new adjacencies is higher. On the contrary, the larger value of UR of new adjacencies was due to the smaller average number of adjacencies in the list (i.e. 13.4 vs 17.8), while the higher ratio of HO share through new adjacencies was due to the larger number of new adjacencies added (i.e. 2960 vs 1561). To avoid the influence of these factors, the analysis was focused on the RP indicator. While the average RP of the new adjacencies in the rural-urban scenario was 0.65, the same indicator was 0.61 for the dense-urban scenario. This lower value of RP means the new adjacencies in the latter scenario tend to occupy a more significant position in their adjacency list. This trend is corroborated by Figure 8, where the histogram of RP of the new adjacencies is depicted for the two scenarios. From the relative frequency values, it can be deduced that the relative position of the new adjacencies in the dense-urban scenario is more favorable than in the rural-urban scenario, i.e. a larger number of new adjacencies exhibit a low value of RP, which means a more significant position in the list. This result is more evident in those values of RP corresponding to the first half of the list (i.e. RP≤0.5).

This better performance in dense-urban environments is due to several factors, the foremost of which is the inaccuracy of propagation models used in the planning tool. While accurate propagation prediction is relatively straightforward in rural environments, due to the use of tall isolated masts, dense-urban environments require accurate data to describe terrain topography, which is rarely available. Another factor is the large number cells that cover the same area in dense-urban scenarios. Due to high cell overlapping, cells in these scenarios have many nearby cells that can potentially take a non-negligible share of HOs when added to the neighbor cell list. In contrast, rural cells have few significant adjacencies due to the low level of cell overlap. This statement is backed up by the analysis of the average UR of adjacencies given their position j in the ordered adjacency list, which is presented in Figure 9. For this analysis, only positions up to the average number of active adjacencies (i.e. 14.65) have been considered. While most HOs are concentrated on few adjacencies in rural-urban scenarios, a slightly more equalized HO share is observed in dense-urban environments. Hence, regardless of the number of existing adjacencies, newly added adjacencies always retain a non-negligible utilization in the latter scenario. The last one is the human factor, which is strongly coupled with the previous one. In rural-urban scenarios, the users are confined in localized routes, favoring the concentration of HOs in few adjacencies. Thus, missing adjacencies in these simple scenarios are easily spotted through visual inspection by optimization engineers. By contrast, a missing adjacency in urban scenarios is more difficult to detect due to its subtle effect in network performance and the large number of potential adjacencies. As a consequence, manual inspection is hardly ever performed in these complex scenarios.

Implementation issues

A stand-alone programme was coded to test the validity of the proposed method in a real environment. In this trial scenario, the programme was run on a Windows-based single-processor laptop with a clock frequency of 1GHz and 1GByte of RAM. This computer was able to connect to the commercial NMS used by the operator to collect the required input information for cells under optimization (i.e. defined adjacencies, co-ordinates and azimuth bearings). The execution time for all 9120 cells was about 15 minutes, depending on the hardware used, the same task would take several hours up to a few days to complete on a planning tool. The output created by the software was a configuration file, which was required to automatically implement the missing adjacencies in the network. To reduce download time, the configuration file only reflected changes in the adjacency lists of the network, which can be updated incrementally. After the successful trial, this method will be integrated into the software of the NMS to offer this capability as a standard functionality.

Conclusions

A new method for spotting missing adjacencies in GERAN networks, based only on NMS data, has been outlined. To determine its suitability, this method has been trialed in a public network. Results show that, despite the majority of created adjacencies only carried relatively little HO traffic, on some cells very important missing adjacencies could be found. More formally, 44% of these new adjacencies were not the least significant in their list, and can thus be strictly defined as missing adjacencies. Thus, a better adjacency list could be defined without any list extension if new adjacencies substituted current ones with less significance. The comparison of results in different scenarios has shown that the best performance is achieved in dense-urban environments, where propagation prediction models are less accurate, high cell overlapping exists and user mobility pattern is less deterministic.

It is expected that the impact that the implemented changes have on the overall network performance is small, as the number of existing neighbors is extremely large. Thus, no significant effect could be observed in the signal-quality and signal-level measurements, unless the analysis focuses on network areas where significant adjacencies have been added. Likewise, it is clear that a simple methodology as the one outlined here will not be able to identify all missing adjacencies in one attempt. Hence, methods have also been discussed that could improve the obtained accuracy, such as the refinement based on blacklisting adjacencies. Nonetheless, these limitations are more than compensated for by the benefit of a simple and fast algorithm, which is suitable to be applied regularly for the entire network without the need to involve different tools.

Finally, it is worth noting that, although the method has been conceived and validated for GERAN, the approach is equally valid for other cellular technologies. Since the method does not rely on any system specific information, but only on information about the site location and antenna direction, it is also applicable to UTRAN. For the same reason, it could also be applied to the difficult task of identifying neighbor cells between different technologies (e.g. GERAN-UTRAN, UTRAN-WLAN), where missing neighbors are more likely to exist. In the latter case, the only precaution would be to reserve space in the neighbor-cell list for intra-system neighbors.

Acknowledgements

The authors wish to thank K. Rasmussen, N. Vig, G. Davies and A. Brimer of Nokia for their help, support and advice on this method.

References

[1] T. Halonen, J. Romero, J. Melero, *GSM, GPRS and EDGE Performance*, 2nd ed., John Wiley & Sons, pp 484 – 490, 2003,.

[2] A. Mishra, *Fundamental of Cellular Network Planning and Optimization*, Wiley, West Sussex, 2004.

[3] J.Lempiäinen, M.Manninen, *Radio Interface System Planning for GSM/GPRS/UMTS*, Kluwer Academic Publishers, 2001.

[4] J.Laiho, A.Wacker, T.Novosad, *Radio Network Planning and Optimization for UMTS*, John Wiley & Sons, pp. 424-425, March 2002.

[5] 3GPP TS 04.08 version 7.21.0 Release 1998, ETSI TS 100 940 V7.21.0 (2003-12).

[6] M. Mouly, M.B. Pautet. *The GSM System for Mobile Communications*. Palaiseau, France, 1992.

[7] T. Okumura, E. Ohmori, K Fukuda., "Field Strength and its Variability in VHF and UHF Land Mobile Services", *Rev. Elect. Commun. Lab.*, 16(9–10), 1968, 825–873.

[8] J. Walfish, H. L. Bertoni, "A Theoretical Model of UHF Propagation in Urban Environments", *IEEE Trans. Antenna Propagation*, AP-36, pp. 1788–1796, 1988.

[9] J.E. Berg , "A recursive Method For Street Microcell Path Loss Calculations", in *Proc. 6th IEEE Int. Symp. on Personal, Indoor and Mobile Communications PIMRC'95*, vol 1, pp 140-143, 1995.

[10] P.G. Brown, C.C. Constantinou, "Prediction of Radiowave Propagation in Urban Microcell Environments using Ray Tracing Methods", *ICAP'95*, IEE Conference Publication, No. 407, pp. 373–376, 1995.

[11] P. Daniele, M. Frullone, K. Heiska, G. Riva, C. Carciofi, "Investigation of adaptive 3D microcellular prediction tools starting from real measurements", in *Proc 5th IEEE Int. Conf. on Universal Personal Communications*, pp. 468–472, Sept. 1996.

[12] Y. Timmer, M. Bergenlid, "Estimating the Inter-Cell Dependency Matrix in a GSM Network", *Proc. 50th* IEEE *Vehicular Technology Conference,* vol. 5, pp. 3024-3028, Sept. 1999.

[13] R.Barco, F.J.Cañete, L.Díez, R.Ferrer, V.Wille, "Analysis of mobile-measurement-based interference matrices in GSM networks", *Proc. 54th IEEE Vehicular Technology Conference*, Vol.3, pp.1412-1416, Oct.2001.

[14] A.M.J. Kuurne, "On GSM mobile measurement based interference matrix generation", *Proc. 55th IEEE Vehicular Technology Conference*, vol. 4, pp.1965-1969, May 2002.

[15] H. Olofsson, S. Magnusson, M. Almgren, "A Concept for Dynamic Neighbor Cell List Planning in a Cellular System", *Proc. 7th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. 138-142, Oct. 1996.

[16] H. Olofsson, S. Magnusson, "Dynamic neighbor cell list planning in a microcellular network", *Proc. 6th IEEE Int.. Conf. Universal Personal Communications Record,* vol.1, pp. 223-227, Oct. 1997.

[17] L.A. Wolsey, G.L. Nemhauser, *Integer and Combinatorial Optimization*, *John Wiley & Sons*, July 1999.

[18] S. Martello, P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley & Sons, New York, 1990.

[19] H. Kellerer, U. Pferschy, D. Pisinger , *Knapsack Problems*, Springer Verlag, 2003.

[20] S. Martello and P. Toth. "Algorithms for knapsack problems". In S. Martello, editor, *Surveys in Combinatorial Optimization*, volume 31 of *Annals of Discrete Mathematics*, pp. 213-258. North-Holland, 1987.

[21] G. Carpaneto, S. Martello, P. Toth, "Algorithms and Codes for the Assignment Problem", in B. Simeone, P. Toth, G. Gallo, F. Maffioli, S. Pallottino, editors, FORTRAN Codes for Network Optimization, *Annals of Operations Research* 13, 1988.

[22] S. Martello, D. Pisinger, P. Toth, "New Trends in Exact Algorithms for the 0-1 Knapsack Problem", *European Journal of Operational Research* 123, 325-332, 2000.

[23] G.B. Dantzig, "Discrete variable extremum problems", *Operations Research,* vol. 5, pp. 266-277, 1957.

[24] S. Martello, P. Toth, “An upper bound for the zero-one knapsack problem and a branch and bound algorithm”, *European Journal of Operational Research,* vol. 1, pp. 169-175, 1997.

[25] D. Pisinger, “A minimal algorithm for the 0-1 knapsack problem”, *Methods of Operations Research,* vol. 49, pp. 277-293, 1985.

Authors vitae/bio:

Following under graduate studies in Germany and an internship at Bell Communications Research, USA, **Volker Wille** carried out postgraduate studies in UK leading to a PhD degree. In Nokia, he currently is responsible for development of features that enable self-regulation of cellular networks.

Matías Toril received an M.Sc. degree in telecommunication engineering from the University of Málaga (Spain) in 1995. Since 1997, he is on the staff of the Communications Engineering Department, where he is currently working toward a Ph.D. degree. From 2000 to 2003, he simultaneously joined the Nokia Networks Competence Team in Málaga. His research interests include optimization of mobile radio access networks and radio resource management.

Raquel Barco received her M.Sc. degree in telecommunication engineering from the University of Málaga (Spain) in 1997. From 1998 to 2000 she worked at the European Space Operations Centre (ESOC) of the European Space Agency in Darmstadt (Germany). From 2000 to 2003 she worked part-time for Nokia Networks. Since the end of 1999 she has been with the Communications Engineering Department, University of Málaga. Her research interests include satellite and mobile communications, mainly focused in self-regulation of radio networks.

Figure captions

Figure 1. Potential neighbors cells in a live environment.

Figure 2. Histogram of the intra-cell handover ratio on a cell basis after the optimization process.

Figure 3. Histogram of the handover utilization ratio in newly created adjacencies.

Figure 4. Scatter plot of utilization ratio of the most significant adjacency versus the number of adjacencies in a cell.

Figure 5. Average utilization ratio of the most significant adjacency versus the number of adjacencies in a cell.

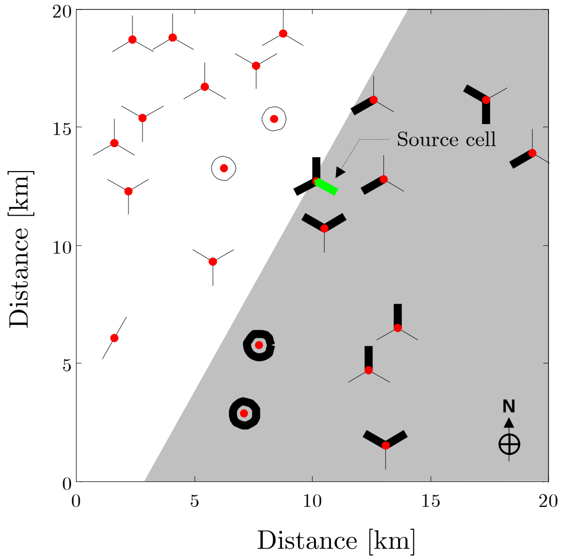
Figure 6. Histogram of the number of active adjacencies per cell in the trial area.

Figure 7. Histogram of the relative-position indicator in new adjacencies.

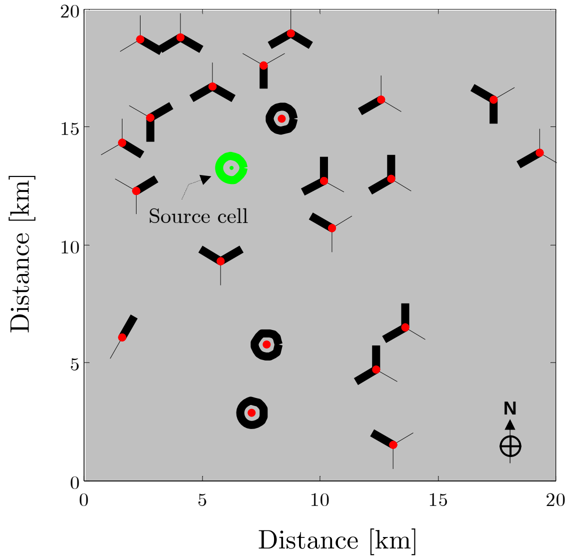
Figure 8. Histogram of the relative-position indicator in new adjacencies for rural-urban and dense-urban scenarios.

Figure 9. Average utilization ratio of adjacencies in terms of their position in the ordered adjacency list.

Figure 1

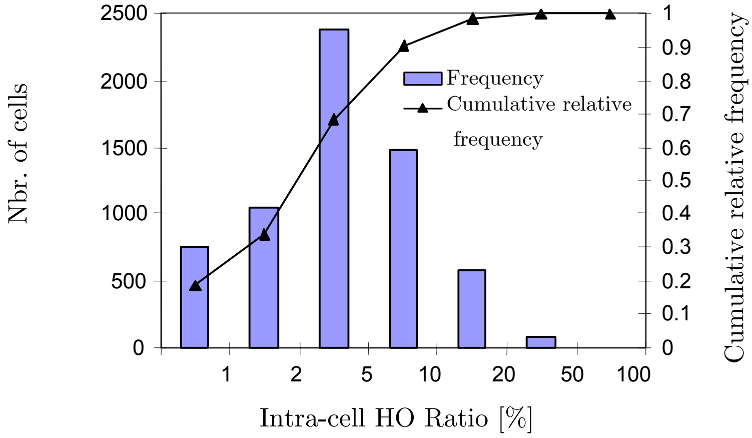


(a) Tri-sectorized



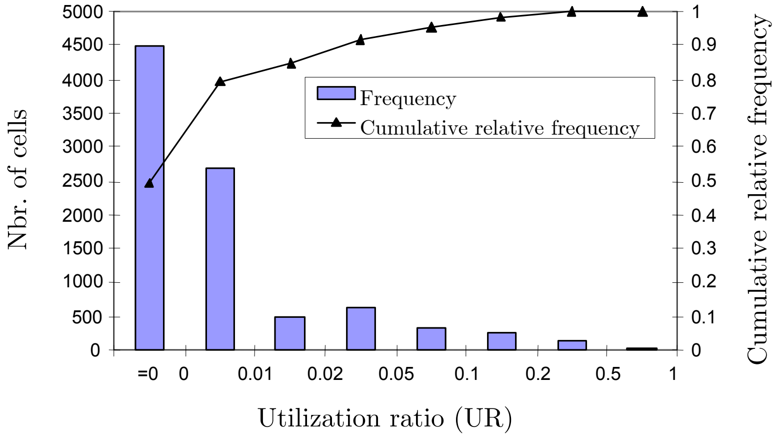
(b) Omni

Figure 2



Nbr. of cells
2500
2000
1500
1000
500
0
1
2
5
10
20
50
100
Intra-cell HO Ratio [%]
Frequency
Cumulative relative frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Cumulative relative frequency

Figure 3



Nbr. of cells
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
=0
0
0.01
0.02
0.05
0.1
0.2
0.5
1
Utilization ratio (UR)
Frequency
Cumulative relative frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Cumulative relative frequency

Figure 4

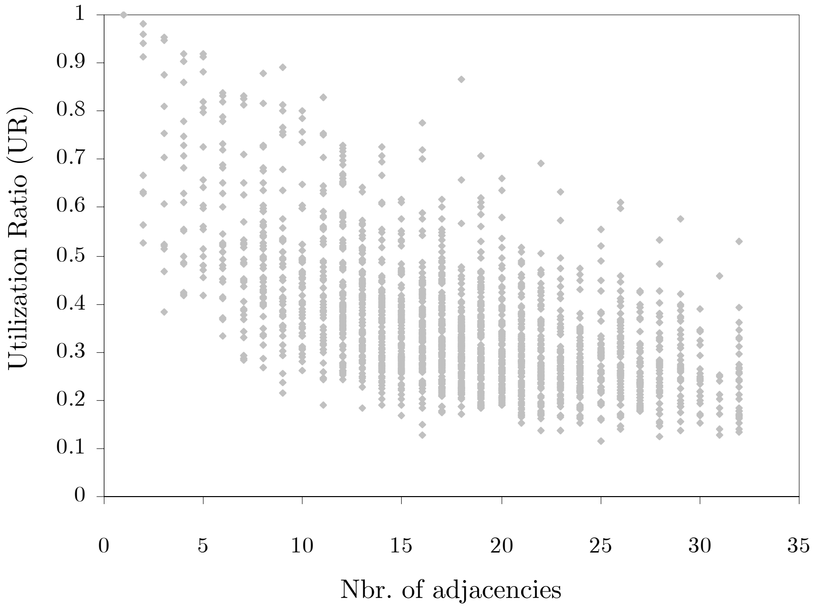
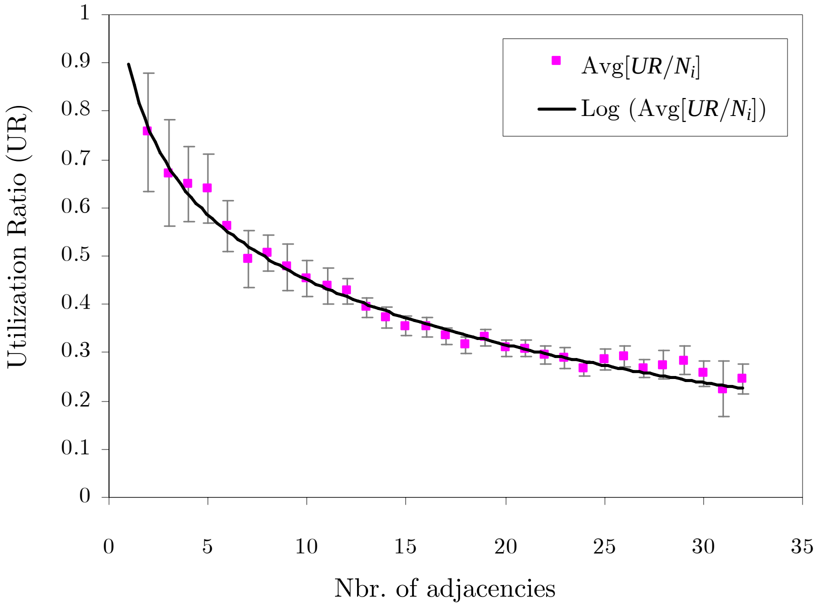
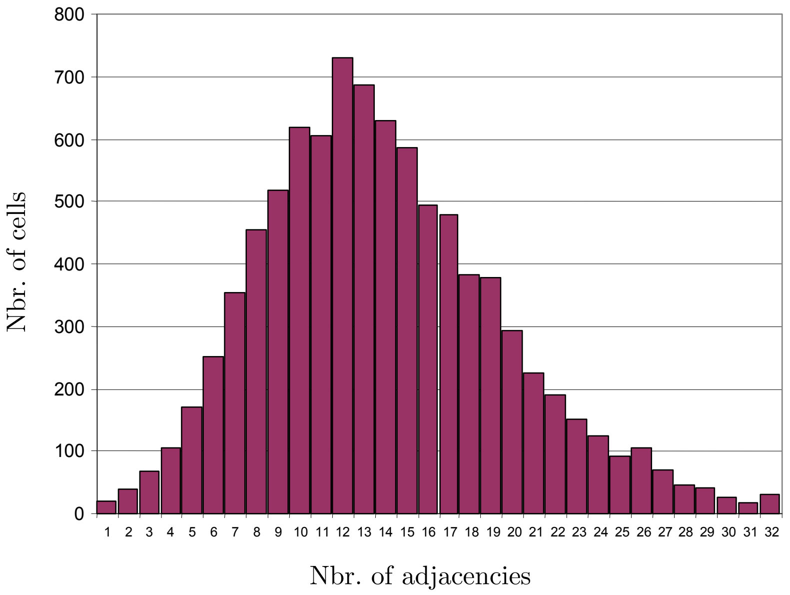

Utilization Ratio (UR)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
30
35
Nbr. of adjacencies

Figure 5



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Utilization Ratio (UR)
Avg[UR/N_i]
Log (Avg[UR/N_i])
0
5
10
15
20
25
30
35
Nbr. of adjacencies

Figure 6



800
700
600
500
400
300
200
100
0
Nbr. of cells
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
Nbr. of adjacencies

Figure 7

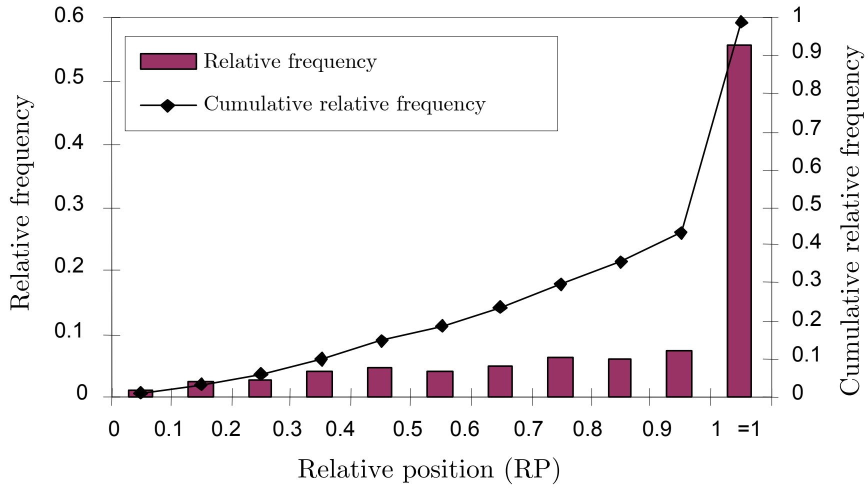
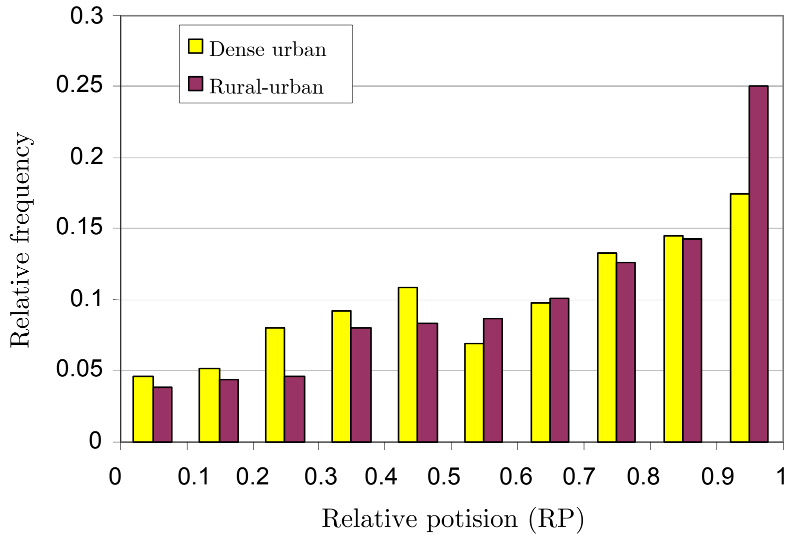
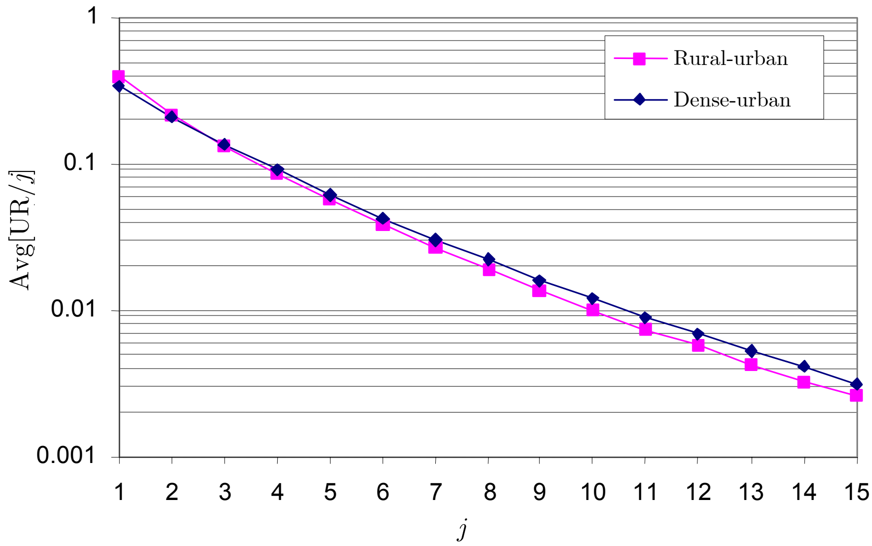

0.6
0.5
0.4
0.3
0.2
0.1
0
Relative frequency
Relative frequency
Cumulative relative frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Cumulative relative frequency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
=1
Relative position (RP)

Figure 8



Dense urban
Rural-urban
Relative frequency
Relative potision (RP)
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 9



Rural-urban
Dense-urban
1
0.1
0.01
0.001
Avg[UR/j]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
j

Table captions

Table 1: Example of outgoing HO activity for a newly created pair of adjacencies

Table 2: Overall results from the optimization process.

Table 3: Comparison of results in dense-urban and rural-urban scenarios.

Table 1

Cell	Total HO #	HO # to cell B	Ratio of HO to cell B	Cell	Total HO #	HO # to cell A	Ratio of HO to cell A
A	2836	2496	88.0%	B	20426	2670	13.1%

Table 2

Nbr. of cells	9120
Nbr. of adjacencies	165898
Avg. nbr. of defined adjacencies per cell	18.2
Avg. nbr. of active adjacencies per cell	14.7
Avg. distance to adjacencies [Km]	4.7
Nbr. of new adjacencies	9052
Avg. nbr of new adjacencies per cell	0.99
Avg. distance to new adjacencies [Km]	3.5
Nbr. of active new adjacencies	4566
Ratio of active adjacency [%]	50.4
Total nbr. of HOs (intra-cell HO excluded)	35105676
Total nbr. of HOs in new adjacencies	573398
Ratio of HO in new adjacencies [%]	1.63
Avg. UR of new adjacencies [%]	3.37
Std. UR of new adjacencies [%]	7.26
Avg. RP of new adjacencies	0.64
Std. RP of new adjacencies	0.28

Table 3

	Rural-urban	Dense-urban	Total
Nbr. of cells	2403	1962	9120
Avg. nbr. of defined adjacencies per cell	16.8	22.5	18.2
Avg. nbr. of active adjacencies per cell	13.4	17.8	14.7
Avg. distance to adjacencies [Km]	6.3	1.3	4.7
Nbr. of new adjacencies	2960	1561	9052
Avg. nbr of new adjacencies per cell	1.23	0.80	0.99
Avg. distance to new adjacencies [Km]	4.5	0.9	3.5
Nbr. of active new adjacencies	1511	741	4566
Ratio of active adjacency [%]	50.9	47.3	50.4
Total nbr. of HOs (intra-cell HO excluded)	9029777	10398071	35105676
Total nbr. of HOs in new adjacencies	195674	108147	573398
Ratio of HO in new adjacencies [%]	2.17	1.04	1.63
Avg. UR of new adjacencies [%]	3.13	3.08	3.37
Std. UR of new adjacencies [%]	7.68	7.00	7.26
95% conf. int. avg. UR of new adjacencies	[3.122,3.143]	[3.060,3.09]	[3.37,3.38]
Avg. RP of new adjacencies	0.65	0.61	0.64
Std. RP of new adjacencies	0.28	0.28	0.28
95% conf. int. avg. RP of new adjacencies	[0.650,0,651]	[0.613,0,615]	[0.641,0,641]