# OPTIMIZING CELL TOWERS DISTRIBUTION USING SPATIAL MINING AND GEOGRAPHIC INFORMATION SYSTEM

**Alaa H. AL-Hamami,  *Soukaena H. Hashem**
Amman Arab University for Graduate Studies
*Department of Computer, University of Technology. Baghdad- Iraq.

**Abstract**
   The appearance of wireless communication is dramatically changing our life. Mobile telecommunications emerged as a technological marvel allowing for access to personal and other services, devices, computation and communication, in any place and at any time through effortless plug and play. Setting up wireless mobile networks often requires: Frequency Assignment, Communication Protocol selection, Routing schemes selection, and cells towers distributions. This research aims to optimize the cells towers distribution using spatial mining with Geographic Information System (GIS) as a tool. The proposed optimized towers distribution done by applying the Digital Elevation Model (DEM) on the image of the targeted area should be covered. DEM applied with two levels of hierarchy, first level for getting the cells and second for getting best location in each cell. The research will apply the spatial association rules technique on the second level to select the best square in the cell for placing the antenna. From that the proposal will try to minimize the number of installed towers, makes tower's location feasible, and provides full area coverage.

## توزيع امثل للابراج الخلوية بأستخدام التعدين الجغرافي ونظام المعلومات الجغرافي

**علاء حسين الحمامي، *سكينة حسن هاشم**
جامعة عمان العربية للدراسات العليا
*قسم الحاسبات، الجامعة التكنولوجية. العراق-بغداد.

**الخلاصة**
   ظهور الاتصالات اللاسلكية سببت تغير كبير في حياتنا، الاتصالات الجوالة تعتبر اعجوبة تكنلوجية حيث انها تسمح بالولوج الشخصي وبقية الخدمات والاجهزة والحسابات والاتصالات في اي مكان وفي اي وقت بدون معوقات الربط والتشغيل. تهيئة شبكات الاتصالات اللاسلكية الجوالة تتطلب تحديد الذبذبات واختيار بروتوكولات الاتصالات واختيار قواعد التمرير وتوزيع الابراج. هذا البحث يهدف الى الوصول الى الامثلية بتوزيع الابراج من خلال استخدام التعدين الجغرافي ونظام المعلومات الجغرافي كأدوات اساسية. امثلية التوزيع تتم من خلال تطبيق نموذج الرفع الرقمي على صورة المنطقة التي يجب ان نغطيها بالابراج على مستويين وبشكل هرمي. المستوى الاول هو للحصول على الخلايا والمستوى الثاني تحديد افضل موقع بكل خلية. هنا سنطبق التعدين الجغرافي من خلال تقنية القواعد المترابطة على المستوى الثاني لاختيار افضل مربع في الخلية لوضع الهوائي الممثلة بالبرج. من هذا المقترح سوف نحاول تقليص  عدد الابراج التي سيتم تنصيبها

لجعل المنطقـة مغطـاة بالشبكة وجعـل موقـع الابـراج ذي جدوى واخيـرا جعـل المنطقـة مغطـاة بالشبكة بشـكل

افضل.

## 1. Introduction

In this section we will present the definitions of all concepts and tools would be used in the proposed system.

**Cellular telephony** is the next and perhaps the most representative example of mobile communication systems. The cellular phone system is characterized as a system ensuring bidirectional wireless communication with mobile stations moving even at high speed in a large area covered by a system of base stations. The cellular system can cover whole country. Moreover, a family of systems of the same kind can cover the area of many countries. Initially, the main task of a cellular system was to ensure the connections with vehicles moving within a city and along highways. The power used by cellular mobile stations is higher than that used by the wireless telephony and reaches the values of single watts, for more details see references [1, 2].

**Spatial Databases** Spatial databases are databases that, in addition to usual data, store geographical information like maps, and global or regional positioning. Such spatial databases present new challenges to data mining algorithms. Spatial data mining is a process to discover interesting, potentially useful and high utility patterns embedded in spatial databases. Efficient tools for extracting information from spatial data sets can be of importance to organizations which own, generate and manage large spatial data sets, for more details see references [3, 4].

**Association analysis** is the discovery of what are commonly called *association rules*. The traditional problem is stated and solved in the following references [5, 6].

**Geographical Information System (GIS)**, continuous surface such as terrain surface, meteorological observation (rain fall, temperature, pressure etc.) population density and so on should be modeled. Grid at regular intervals: Bi-linear surface with four points or bi-cubic surface with sixteen points is commonly used. Random points: Triangulated Irregular Network (TIN) is commonly used. Interpolation by weighted polynomials is also used. Contour lines: Interpolation based on proportional distance between adjacent contours is used. TIN is also used. Profile: Profiles are observed perpendicular to an alignment or a curve such as high ways. In case the alignment is a straight line, grid points will be interpolated. In case the alignment is a curve, TIN will be generated [7, 8].

**Digital Elevation Model (DEM)** is a digital representation of topographic surface with the elevation or ground height above any geodetic datum. Followings are widely used DEM in GIS, see figure (1), [9, 10].
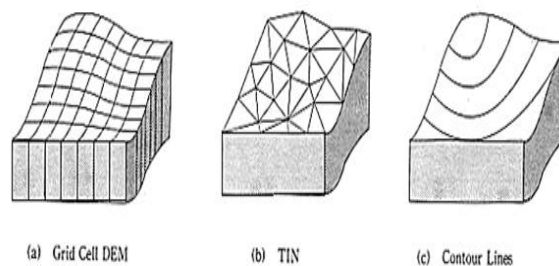


**Figure 1: DEM models.**

**Digital Terrain Model (DTM)** is digital representation of terrain features including elevation, slope, aspect, drainage and other terrain attributes. Usually a DTM is derived from a DEM or elevation data. Several terrain features including the following DTMs. Slope and Aspect, Drainage network, Catchment area, Shading, Shadow and Slope stability, see figure (2), [11, 12].
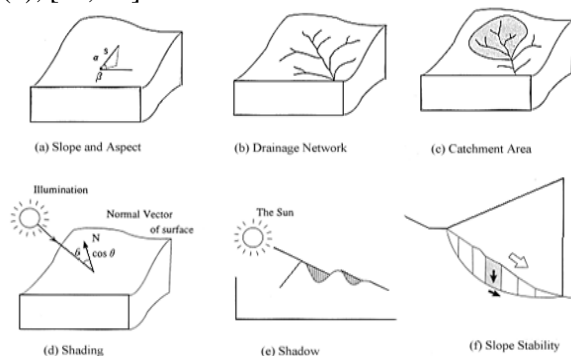


**Figure 2: DTM models.**

## 2. The Problem and Related Works

Presume that you work for a cellular phone company that is interested in expanding the extent of its coverage on the country. You have been tasked with identifying the most suitable locations in the county for the placement of new cellular phone towers. As you might suspect there are many factors that govern the placement

of cellular phone towers. Some factors are based on physical requirements, others on political and economical issues.

Most previous related works [13, 14] suffer from many points related with tower distribution these are:

- This problem is difficult to be modeled mathematically so it is a NP-Complete problem because it has no exact solution.
- It is multi-objective problem, these objectives are Frequency Assignment, Communication Protocol selection and Routing schemes selection.
- Also it is constrained since it depend on the different geographically nature of the area to be covered.

So some sort of approximation is required to minimize number of antennas installed, to have practically feasible location, to provide full area coverage and finally to reduce the problem into a solvable solution.

## 3. The Proposed System

As we mention distributing towers is difficult to model, so some approximation is required. This research introduces a suggestion to optimize the towers distribution; this will be explained in the following steps:

### First step (External Grid)

Take an image for the regions that must be covered with wireless mobile phones. Then applying GIS on the image using DEM for dividing the image to Grid, this grid called *External Grid*, which represent the first level, where each square in that grid will represent a cell, see figure (3), the division done by a specific proposed program depending on requirements of area covering administrators with mobile phones.
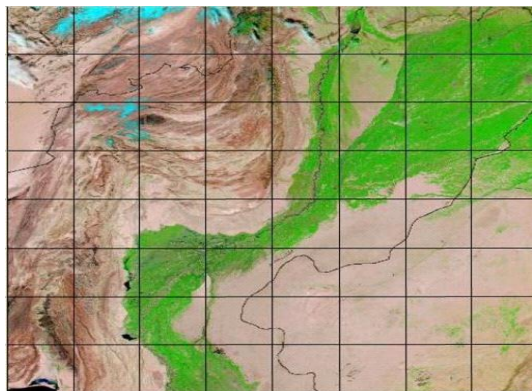


**Figure 3: First level in the proposed program with GIS using DEM on all the area must be covered.**

### Second step (Internal Grid):

Each cell will represent a region that must be covered by placing a tower so each cell in the external grid will also be divided into internal grid of squares by program, see figure (4), to select the most suitable square for placing the antenna. The size of these squares depends on the coverage radius of the antenna used and here we suggest using the omni-directional antenna which covers eight adjacent grid squares around it.
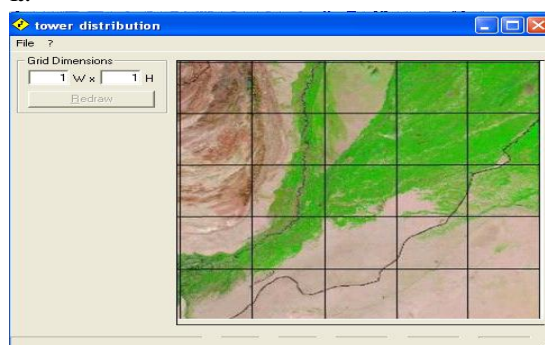


**Figure 4: Second level in the proposed program, where cell divided to squares.**

### Third step (preprocessing the internal grid):

Omni Directional Antennas are used to provide coverage in each cell. The tower must be placed in the most and most suitable location of the cell to make the covered area as good as coverage. To optimize the location of antenna we suggest the following:

### First:

Data mining technique (association rules) will be used since we have the area (each one of cells in the external grid) that will be divided into grid of squares of earth and for each square all it is GIS information will be registered. Both explicit and implicit relations and patterns among spatial objects in all squares will be extracted which are represent the presented area. The spatial and non-spatial attributes for all spatial objects in all the squares in the internal grid will be gotten.

### Second:

The following non spatial attributes (type, size, population, employment rate, etc. …..) and the spatial attributes are presented in the spatial database. This database would constructed as in the following points

- The first attribute will represent the spatial objects (O1, O2, O3… On), and will represent the identification of the transaction in the database.
- The second attribute will represent the type of the spatial object such as (town, road,

….) and will be represented as in the following codes:

***(Code of the type of the spatial object)***
1= town, 2 = road, 3 = river, 4 = sea, 5 = lake, 6 = mine, 7 = forest, 8 = bridge, 9 = highway, 10 = peak, 11= trough.

- The third attribute will represent the size of the spatial object and will have the following codes:

***(Code of the size of the spatial object)***
1= large, 2 = medium, 3=small.

- The fourth attribute will represent the shape of the spatial object and will have the following codes:

***(Code of the shape of the spatial object)***
1 = point, 2 = line, 3=polygon.

- The fifth attribute will represent the directions state: north of, south of, east of, west of, north west of, north east of, south west of, south east of. These spatial attributes will have the following codes:

***(Code of the direction, Code of the related spatial object)***
(A, Oi) = (north of, Oi), (B, Oi) = (south of, Oi), (C, Oi) = (east of, Oi), (D, Oi) = (west of, Oi), (E, Oi) = (north east of, Oi), (F, Oi) = (north west of, Oi), (G, Oi) = (south east of, Oi), (H, Oi) = (south west of, Oi)

- The sixth attribute will represent the Position state: disjoint, overlap, meet, covers, covered by. For more explanation see figure (5). These spatial attributes will have the following codes:

***(Code of the direction, Code of the related spatial object)***
(I, Oi) = (overlap, Oi), (II, Oi) = (meet, Oi), (IIIC, Oi) = (covers, Oi), (IV, Oi) = (covered by, Oi), (V, Oi) = (disjoint, Oi)

- The seventh attribute will represent the distance between spatial objects, see figure (5).
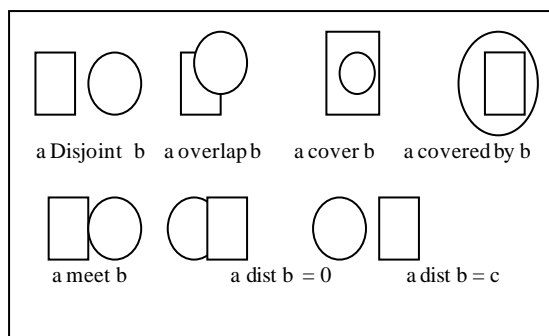


**Figure 5: the position state spatial attribute.**

For each square in the cell the proposed spatial database will be in the following design, Figure (6). Figure (6) present the spatial database for the square numbered as (3, 3).

| Spatial object | Type | Size | Direction | Position | Distance | Population | Employment |
|---|---|---|---|---|---|---|---|
| O1 | 4 | 1 | (a, o2) | (I, o3) | O3 <50 km | High | High |
| O2 | 2 | 1 | (a, o3) | (I, o4) | O3 <50 km | low | Low |
| … | | | | | | | |
| … | | | | | | | |
| … | | | | | | | |
| …. | | | | | | | |

**Figure 6: the proposed spatial database for each square in a cell (sample of the square (3, 3)).**

**Fourth step (mining the internal grid):**

Now after building spatial database for each square in the cell we must mine this database to extract the patterns that will aid to select the best square to be the place of the antenna.

Since the proposed spatial database is usually relied upon alphanumerical and often transaction-based. The problem of discovering association rules is to find relationships between the existence of a spatial object with each others (spatial or non spatial attributes) in a large repetitive collection. Association rules would give the probability that some objects attributes appear with others based on the processed transactions, for example:

(large town and near to water → high population [90%])

That meaning there is a probability of 0.9 that high population is found when the town is large and near water. Essentially, the problem consists of finding objects attributes that frequently appear together, known as frequent or large objects attributes-sets.

We find spatially related rules from all the databases represent the squares. For declaration we display some rules extracted from the database of the square (3, 3). These Association rules describe patterns, which are often in the database, two of these rules are.

*1.If type O1 = 4 (sea) and size O1 = 1 and direction O1 (south, o2 (type = 2)) then position O2 (type = 2) (intersect, O1) (s = 50%, c = 80%).*

2. *If type O1 = 4 (sea) then dist between O1 and (O2 (type =2) < 50 km (s = 50%, c = 90%).*

## Analysis of extracted patterns:

From the analysis we see:

In general the resulted association rules from mining all spatial databases of squares in cell are little and limited since the size of them is limited so usually the numbers of objects also limited.

From analyzing the association rules of the square (3, 3) (two of them presented in the above section) we see that this square presented nearly as a sea. Since the major object in it is the sea.

## Fifth step (classifying the squares of the internal grid):

Now the system will classify the squares into two classes they are: first priority and second priority. The classification will be done according to the position of the square in the cell. The parameter of position taken as basic threshold for classifying since the used antenna was to be Omni Directional Antennas which has covers eight adjacent grid squares around it. So the rule of classification the squares in the cell of n*n squares, where each square has the position (x, y) will be as in the following:

*If square position was (1, any y) or (n, any y) or (any x, n) or (any x, 1)*

*Then the square is **second priority class** else the square is **first priority class***

## Sixth step (selecting the optimal square):

For selecting the optimal square follow the following steps;

1. From the results of spatial association rules on each square analyze these rules. From the analysis we will give the square a ratio of goodness. The ratio of goodness is come from *<u>summation of</u>*:

   - Class type (first priority present 100% and second priority present 50%). For example the square which its position (3, 3) has 100% since it return to the first priority class.

   - With suitability of the square (which come from the analyzing the association rules then the nature of the square will be understood. For example the rules extracted in step four with the square which its position (3, 3) which appear from the rules as a sea with little piece of road

in this cases we think the suitability for placing antenna is nearly 50%.

*So the result for that square is 100% + 50% = 150%*

2. For each square repeat step one then compare which square has highest ratio of goodness will be chosen to be the best square for placing the antenna.

3. If the best ratio of goodness was not good enough then will begin in taken the squares in the class of second priority. Then apply point one and point two above in that sixth step. And take the best two corresponded squares to placing two antennas instead of one since that will guarantee the coverage of cell.

## 4. Discussion and Results

The proposed system aimed to optimize the towers distribution by finding the most suitable location in cell for placing the antenna with considering the antenna type is Omni directional.

The proposal use GIS especially DEM and association rules data mining technique as a basic two tools for solving the problem of towers distribution.

The workflow of the system is sequenced steps begin with taking an image of targeted area then divide it to cells. Each cell divided into squares to select suitable one for placing the antenna in it there are internal sequenced steps, initially build a spatial database for each square this base will store all the spatial objects with all their attributes (spatial and non spatial). Then mining each spatial database separately after mining analyze the extracted association rules that for declare the geographical environment of the square and by it we find the goodness ratio. The last ratio plus the class of square (class of square depend on it is position) result the best square for placing antenna.

The optimization in this research gotten from the summation of goodness ratio and class ratio which give final ratio to each square then the square has higher ratio will be select for placing the antenna in it.

## 5. Conclusions

From the suggestions and their results we get the following conclusions:

1. Distributing and placing towers is a difficult problem to be modeled so the presented work was being as approximation for an optimal solution.

2. Using GIS and DEM especially DTM making the division of area more accurate and presenting the surfaces of the square in more precious.

3. Building spatial database as a flat database will make the spatial mining much more efficient that by reduce the mining to one level only so this will prevent the time and space consuming resulted in the previous work by extending the mining to multilevel.

4. We proposed a novel approach for building a spatial database to accommodate all the necessary requirements for applying Association rules, then extract all the patterns which help in distributing the towers.

5. With the proposed spatial database the extraction of association rule is could done by the traditional apriori algorithm without confusing, that make the proposed approach easy to use and understand by the administrators. Also the analysis step followed by extraction rules is easy because it depends on generalization and normalization determined by the miner.

6. In this research the classifier will be built without the need to measure the entropy of each attribute only it depends on the position of the squares.

## REFERENCES

1. Korhonen, J. **2003**. *Introduction to 3G mobile communication*, second edition, Artech house, INC. pp.202-215.

2. Ganz, A.; Ganz, Z. and Wong, K. **2004**. *Multimedia wireless networks: technologies, standard and QOS*; Prentice Hall PTR, England, pp.400-407.

3. Samet, H. **1990.** *Application of Spatial Data Structures: Computer Graphics, Image Processing and GIS,* Reading, MA: Addison-Wesley.  pp.23-30.

4. Mitra, S. and Ahharya, T. **2003.** *Data mining multimedia, soft computing, and bioinformatics,* John wiley and sons, Germany. Pp.90-98.

5. Kantardzic, M. **2003.** *DM concepts, models, methods and algorithms,* jhon wiley &Sons, Germany, Pp.300-310.

6. Ordonez, C. and Omiecinski, E. **1999.** Discovering association rules based on image content. *Proceedings of the IEEE Advances in Digital Libraries Conference (ADL'99),* pp.57-66.

7. Whiteside, A. **1997.** *Recommended standard image geometry models. Open GIS Consortium*, URL: http://www.opengis. org/ ipt/ 9702tf / UniversalImage / TaskForc.pp, Pp:200-250.

8. Zhou, G., and Li, R. **2000.** Accuracy evaluation of ground points from IKONOS high-resolution satellite imagery, *Photogrammetric Engineering & Remote Sensing Magazine*, **66**(9): 1103-1112.

9. Sun, M.; Xue, Y. and Ma, A. **2002.** 3D Visualization of Large DEM Data Set Based on Grid Division. *Journal of Computer-aided Design & Computer Graphics*, **14**(6):188-193 (in chinese).

10. Baltsavias, E.; Pateraki, M. and Zhang, L. **2001.** Radiometric and geometric evaluation of IKONOS GEO images and their use for 3-D building modeling, Proceedings of ISPRS Joint Workshop High Resolution Mapping from Space, 19-21 September, Hanover, Germany.

11. Toutin, T. **2001**. Geometric Processing of IKONOS Geo Images with DEM, Proceedings of ISPRS Joint Workshop High Resolution Mapping from Space, 19-21 September, Hanover, Germany.

12. Toutin, T. and Cheng, P. **2000.** *Demystification of IKONOS*, *Earth Observation Magazine*, **9**(7): 17-21.

13. Sun, M. **2008.** 3D Visualization of Large DEM Data Set, *Journal of Computer-aided Design & Computer Graphics,* Pp.30-43.

14. Reddy, M. **2009.** Terravision II Visualizing Massive Terrain Databases in VRML, *IEEE Computer Graphics& Applications,* Pp.355-370.