Wireless Indoor Resource Optimisation meeting
Prescribed User Capacity Requirements

Eamonn Kenny, Eamonn O Nuallain

Abstract—The optimal locations of transmitters in a building, that best meet the bandwidth requirements of all users, was studied in the past by a small number of authors. They used empirical path-loss models to constrain an optimisation cost function. This paper concentrates on using a signal to interference ratio cost function with realistic field parameters obtained using a fast ray-tracing algorithm. The solution obtained is shown to perform well in realistic scenarios, meeting a large proportion of user bandwidth requirements. However, it does not perform well in an oversaturated environment.

Index Terms—ray-tracing, convex space, indoor, optimisation.

I. INTRODUCTION

THE term resource optimisation mentioned in this context refers to the optimisation of the resources that make up a communication system. The system provider may attempt to save financially by providing cheaper services while maintaining good coverage. Examples of where the provider may incur large costs are in:

• antenna costs - antennas are costed according to type, ranging from very expensive smart antennas down to omni-directional antennas.
• location costs - the position in a building where the antenna needs to be mounted may be expensive or cheap.
• capacity requirements of the user - may be low in the case of voice links or high in the case of data links. The higher the bit rate required by the user equipment (UE) the better the planning of the system should be to give each user adequate service.
• environmental safety issues - planning the antennas in an area that will not affect nature or humans.

This paper will concern itself with defining a resource optimisation algorithm that uses at its core a ray-tracing engine to provide propagation coverage information. The optimisation problem will be defined as follows:

Generate the least number of base stations at optimal locations giving adequate coverage to a set of users with a predetermined capacity requirement.

When a simulation is performed to obtain the optimal positions of base stations while providing the largest number of possible users wanting a connection to the system, this is usually performed in two stages known as downlink and uplink optimisation. In a telecommunications system the downlink is the connection from many servers in a network to many user equipments (UEs). User equipment (UE) is a general term given to laptops, tablets, desktops, mobile phones, e-readers etc, that use wireless technologies. The uplink is the connection from the UE back to a receiving server. Servers are also called base transceiver stations (BTS) with the ability to both send and receive information. They can contain one or more antennas in a working system. For our system we will assume that the downlink optimisation requires a number of user equipments with a particular capacity requirement assigned to each, assumed already to be connected to the system.

The capacity requirements of the user will depend on what is known as the signal to interference ratio (SIR). This is a measure of the signal power received by a user in the presence of possible interfering base stations. The other base stations which are possibly emitting at the one time can drown out the signal from the best server. This needs to be avoided where possible, and this is why an optimally positioned set of base stations is necessary. The uplink optimisation determines how many users can connect to the system assuming that each UE is possibly interfering with one another and that they are randomly distributed in the building.

The optimisation algorithm depends heavily on propagation effects which must be computed to give the signal to interference ratios (SIR). Without accurate values for the signal power it cannot be assumed that the simulation of the system is reliable. The decision was made to use asymptotic methods to obtain the electric field strength values.

The optimisation algorithms described by Rappaport [11] are very involved but excellent as a basis on which to build a better algorithm. They require field strength information but do not contain signal to interference ratio information. In other words it does not consider the capacity problem. The basic ideas within the algorithm of Rappaport are used as a framework in this paper to build a far more sophisticated algorithm that accounts for SIR values. Many unexplained variables will be replaced with realistic parameters based on statistics made available by Laiho [6].

The optimisation of a system depends on many issues. The nature of the function to be optimised is one issue and the type of constraints is another. Much of the background on convex problems or linear programming problems is covered by Boyd [2] but some of the techniques can be applied to non-linear programmes. The solution to this type of programming problem is usually sought by a path-following method as described in Gonzaga [3] using a good restart method as...
Both the downlink (section III-F) and the uplink (section III-H) algorithms are described and results are presented for test buildings given the parameters of Laiho [6] as test statistics to obtain physically meaningful results (section IV).

A tighter constraint problem is obtained by expressing the field strength summation as a mean field strength value. This mean value is obtained by a random phase generator that is usually computationally time consuming (described in subsection III-F7). A fast numerical technique is presented that speeds up the computation considerably. The results from initial tests regarding the new optimisation technique are presented and conclusions are drawn regarding the effectiveness of the technique.

II. STATE OF THE ART

Accurate field strength is virtually impossible to attain using a direct three dimensional integral equation formulation of the full wave solution over large surface areas. The computational complexity is far too great due to the dense matrix systems arising during the calculation of the current density. Asymptotic methods however, have the advantage of being accurate at very high frequencies and are known to be computationally fast to calculate.

The asymptotic solutions most widely used are ray optic techniques calculating direct line of site (LOS), reflections and diffractions from a server to a UE. The ray-tracing method we consider uses the method of images to calculate reflected and transmitted rays (see Kenny [5]). Two papers in the literature use very similar electromagnetic wave equations to the ones mentioned in this paper. The first is a planning tool called CINDOOR by Torres [15] containing a full three dimensional ray-tracing code for indoor wireless systems in enclosed spaces. The second is also an indoor wireless prediction tool described by Ji [4].

More recently authors have implemented indoor resource optimisation techniques such as Bongyong [12] using a beam-forming space division algorithm, Abbasi [1] implemented a greedy select optimisation algorithm with beam-formed and Nebel [8] implemented a MIMO based spatial capacity solution, however none of the recent solutions produce a very discrete optimisation such as that provided by Rappaport [10] for a specific number of users given very specific capacity requirements.

The novelty in this paper will be in the application of ray-tracing methods within optimisation algorithms, the optimisation technique itself which uses signal to interference ratios rather than pathloss algorithms, and a fast numerical solution to provide the mean value of the signal power at a point.

III. INDOOR RESOURCE OPTIMISATION

In a telecommunications system the downlink is the connection from many servers in a network to many receivers. These receivers are usually called user equipment (UE) because they range from voice links on mobile phones and tablet PCs to laptops and multimedia phones. The uplink is the connection from the UE back to a receiver in the server.

When a simulation is performed to obtain the optimal number of users required to be attached to the system, we usually perform what is known as a downlink and uplink optimisation. The downlink optimisation requires a number of user equipments with a particular capacity requirement assigned to each, to be connected to the system. The uplink optimisation determines how many can connect to the system assuming that each UE is possibly interfering with one another and that they are randomly distributed in the building.

A. Downlink Optimisation Requirements

Section III-B to III-E describe prerequisites of the system that need to be defined before addressing the downlink optimisation in III-F.

B. The Environment

The environment will be a multi-story indoor building defined exactly as in Kenny [5]. The walls, doors, windows, floors and ceilings will consist of filled convex spaces having permittivity, permeability and conductivity parameters associated with them. For experimental purposes the filled convex spaces can be perfect electric conductors (PEC), lossless or lossy.

C. Traffic

User Equipment (UE) can be a mobile phone, lap-top or desktop computer. The receiver positions in the building will serve as positions where UEs can be placed. The capacity required by the UEs can be defined as the traffic of the system. Each user will require a service suitable for their needs. For instance if a person requires a mobile call they will require a voice link at 12.2 kbps. Whereas if a person is using a laptop they may require a high speed connection rate of 128 kbps or higher.

The capacity of the system is a function of the area, number of users and capacity requirements of each user of the system. This can be expressed numerically as:

$$\gamma = \sum_{i=1}^{N} \gamma(x_i, y_i, z_i)$$

where $\gamma$ is the total capacity requirement in the system, and $\gamma(x_i, y_i, z_i)$ is the capacity requirement of a user at position $(x_i, y_i, z_i)$, where there exists $N$ users.
In a modern mobile communication system the statistics of interest is not the path-loss, rather it is the signal interference ratio (SIR). The target signal to interference ratio is related to the energy per bit per noise factor written $E_b/N_0$. At each user position we need to generate a measure of required signal level $S_L$ defined as follows:

$$S_L = \frac{I_m F_m R_s}{P_g}$$

where $I_m$ is the interference margin, $F_m$ is the fade margin, $R_s$ is the receiver sensitivity, and $P_g$ is the processing gain.

The receiver sensitivity is calculated using the following formula:

$$R_s = k T B N_f E_b/N_0$$ (3)

where $k = 1.380658 \times 10^{-23}$ J/K is the Boltzman constant, $T = 290$ K is the temperature measured in Kelvin of the antenna, $B = 3.84$ MHz is the bandwidth and $N_f = 5$ dB for base stations and 7 dB for receivers is the noise factor defined by Laiho [6].

### Table I

<table>
<thead>
<tr>
<th>Direction (receive diversity)</th>
<th>Service</th>
<th>$E_b/N_0$</th>
<th>Fast Fade Margin (dB)</th>
<th>Processing Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downlink</td>
<td>12.2 kbps voice</td>
<td>9.4</td>
<td>5.5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>128 kbps real-time data</td>
<td>11.7</td>
<td>5.5</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>128 kbps non real-time data</td>
<td>6.7</td>
<td>3.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Uplink</td>
<td>12.2 kbps voice</td>
<td>1.7</td>
<td>6.3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>128 kbps real-time data</td>
<td>1.0</td>
<td>6.3</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>128 kbps non real-time data</td>
<td>0.3</td>
<td>3.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Uplink (no receive diversity)</td>
<td>12.2 kbps voice</td>
<td>8.6</td>
<td>6.3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>128 kbps real-time data</td>
<td>8.7</td>
<td>6.3</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>128 kbps non real-time data</td>
<td>6.4</td>
<td>3.4</td>
<td>2.4</td>
</tr>
</tbody>
</table>

We will see that in practice we can remove the fade margin from the calculations if we can define a mean signal value over a small area. We will see later that this can be achieved by the use of a mean random value of the signal value supplied by a ray-tracing model.

Table I shows the $E_b/N_0$, fade margin and processing gain based on measurements made by Laiho [6] defined in a book on the subject of UMTS.

In fact the target signal to interference ratio (tSIR) required for our problem will be defined as

$$tSIR = \frac{E_b/N_0}{P_g}$$ (4)

The interference margin is usually included in the SIR formula, but in this case is not defined because all interferers are known in the system.

**D. Convex Space Ray Tracing**

Figure 1 shows a simple example of a building constructed with planes that lie in the $x$, $y$ and $z$-axes. Given multiple $T_x$ transmitters and multiple $R_x$ locations, any iteration of an optimisation algorithm will perform six changes of direction of each transmitter $T_x$, indicated in the above figure by arrows pointing outwards. The step size will be dictated by the line search algorithm that requires each transmitter $T_x$ to be moved in the $\{-x, +x\}, \{-y, +y\}$ and $\{-z, +z\}$ search directions. The signal power received at each $R_x$ from each $T_x$ must be computed for the line search in each iteration of the optimisation algorithm.

The computational complexity increases dramatically when there are $M$ transmitters, $N$ receivers and $P$ iterations of the algorithm since $6 M N P$ ray-path calculations must be performed to produce convergence. This is why empirical methods as used by Rappaport [11] were favoured over semi-deterministic approaches such as ray-tracing.

Standard visibility algorithms with polar sweep will not cope well with such a problem since the visibility space must be computed with every change of transmitter location. The tetrahedron method [16] of Yun will improve the situation by using a fast spatial technique that divides the building into tetrahedrons, thus removing the requirement for visibility algorithms. However, Kenny [5] attempts to go further by splitting the building structure of rooms, doors, corridors and such like into convex spaces/boxes describing the air in rooms as free (or empty) convex spaces and building materials in terms of distinct interconnected convex spaces/boxes. This results in a fast Method of Moments algorithm for calculating the reflection and diffraction paths in the ray-trace. The method hinges on the utilisation of the property that a ray travelling through a convex space, entering at one point and exiting at another, resulting in a fast traversal of the building. This method is computationally efficient for optimisation techniques since the interconnectivity of the objects in the building acts like a visibility algorithm that is calculated only once for a whole building and then stored for future executions of the algorithm.
As the number of convex spaces in a building increases an oct-tree method can be employed to speed up the spatial discretisation technique even further.

E. System Constraints

Each base transceiver station (BTS) for the system contains a number of servers. Each server contains one antenna system. In practice the antenna can fall into one of the following four categories: it could be a dipole, be sectorized (usually into 90° sectors), consist of an array or a smart antenna. For the optimisation tool we will assume that the antenna is omnidirectional or sectorized for simplication purposes. A more complicated antenna type leads to a much larger number of unknowns in the problem.

Usually there is a cost function, in our case this will be related to the capacity requirements of the users only. The economic cost is not taken into account. The economic cost is related to the cost of certain antennas as well as the cost associated with where they can be positioned in a building. The economic cost can contain parameters such as installation, maintenance, health and appearance.

Power control at the BTS and the UEs is always discrete with a maximum permissible power level available which we will define as \( W_{\text{max}} \). We define the power \( P_i \) in the \( i^{th} \) BTS to be an element of the discrete set \( \{ W_1, W_2, \cdots, W_{\text{max}} \} \). The maximum permissible power is defined as:

\[
W_{\text{max}} = R_s I_m R_{\text{min}} \tag{5}
\]

where \( R_s \) and \( I_m \) are already defined, and \( R_{\text{min}} \) is the minimum signal power to another base station from the current server in the current base station.

The receiver sensitivity is defined to be the signal required to provide a defined error output from the receiver. It is obvious from its definition that the receiver sensitivity is related directly to the signal to interference ratio (SIR). In the case of an uncooperative system we use the term acceptable interference level instead of receiver sensitivity. An uncooperative system is another system that is set up separately from the current system in use, and the telecommunication has no control over how it is run.

The interference margin is a statistic which states a value above the interference signals that the received signal must exceed. In the case of uncooperative systems we call this the protection ratio. It is a value below the receiver signal level, that the interferers signals may not exceed.

F. Non-linear Downlink Optimisation Algorithm

The algorithm described by Rappaport [11] employs a non-linear optimisation technique which places a number of base stations in a building and then uses a gradient technique to iterate to an optimal set of locations for the base stations subject to a number of path-loss constraints for a given distribution of UE’s and an objective function consisting of a linear combination of minisum and minimax objective functions.

The optimisation technique requires a starting solution, a method to obtain the descent direction, a search direction to be chosen and a stopping criteria. The starting point \( x^{(0)} \) is defined to be the locations of the base stations each of which is allocated at one point at the centre of gravity of a hyper-rectangle. These hyper-rectangles must be predetermined. The descent direction can be obtained by using a number of methods such as the conjugate gradient, Newton or Hookes and Jeeves method. The stopping algorithm is usually related to a measure of the gradient change.

There are 7 main changes to the original algorithm discussed in each of the following sub-sections III-F1 to III-F7.

1) Allocation of Hyper-Rectangles where Base Stations Reside:

The method of Rappaport [11] uses a centre of gravity approach to find the location of a hyper-rectangle where the initial position of the base stations is allocated, in the vicinity of receivers \((x_i, y_i, z_i)\) each having an associated priority weight \( w_i \), so that the centre of gravity \( c_j \) in the \( j^{th} \) hyper-rectangle will be calculated using the formula:

\[
c_j = \left( \frac{\sum_{i=1}^{N} w_i x_i}{N}, \frac{\sum_{i=1}^{N} w_i y_i}{N}, z_i \right) \tag{6}
\]

Rappaport does not suggest a physical interpretation for the weights, but we do know that they must be set by the user of the system. In this implementation the weights \( w_i = \gamma_i \) are related to the capacity \( \gamma_i \) in the system. So the starting location for our base stations will be the centre of capacity \( \Gamma_j \):

\[
\Gamma_j = \left( \frac{\sum_{i=1}^{N} \gamma_i x_i}{N}, \frac{\sum_{i=1}^{N} \gamma_i y_i}{N}, z_i \right) \tag{7}
\]

This is exactly what you would expect, since the base station needs to be in close vicinity to the high capacity links and is not required to be as close to the lower capacity links.

Rappaports method for calculating the hyper-rectangles is given as follows:

1) The building is stored inside one hyper-rectangle.
2) Find the centre of gravity in all hyper-rectangles.
3) Split across the longest dimension in the \( x \) or \( y \) direction generating a new set of two hyper-rectangles.
4) Continue the algorithm as in steps 2-3 splitting the hyper-rectangles until the required number of base stations is obtained.

In the case of 3G systems provided at the time of Laiho’s writing [6], we know that the capacity of the servers was \( s_j = 2Mb/s \), so that the server contains receivers whose capacity when added together will not exceed the capacity of the server. Therefore we can implement the constraint that

\[
\sum_{i=1}^{N} \gamma_i \leq s_j \tag{8}
\]

We define the splitting algorithm as follows:

1) The building is stored inside one hyper-rectangle.
2) Find the centre of capacity in all hyper-rectangles.
3) Split across the \( x \) and \( y \) direction to obtain two hyper-rectangles in each case. The hyper-rectangles that contain the least number of servers will be the chosen splitting direction for the hyper-rectangles.
4) Continue the algorithm as in steps 2-3 until all hyper-rectangles meet the server capacity constraint.
2) **Objective Function and Constraints:** The objective function \( f \) of Rappaport consists of a linear combination of a minimium objective function \( f_1 \) and maximin objective function \( f_2 \) subject to a number of placement constraints. These constraints as stated above are the boundaries of the hyper-rectangles. The minimium objective function has the drawback that it might ignore some remotely located UEs, while obtaining a good total weighted coverage. The minimin objective function on the other hand has the drawback that it concentrates on the worst case situation at the expense of the overall averaged weighted coverage. For this reason the linear combination of the two methods is implied. Rappaport uses a method based on signal power while we know in practice that the signal to interference ratio is the statistic of real significance. Rappaports non-linear optimisation problem is stated as:

\[
\begin{align*}
\text{minimize} \quad & f(X, Y, Z) = \phi f_1(X, Y, Z) \\
& + (1 - \phi) f_2(X, Y, Z)
\end{align*}
\]

subject to
\[
\begin{align*}
0 \leq x_j \leq h_{i1}, & \quad 0 \leq y_j \leq h_{ij}, \\
0 \leq z_j \leq h_{j3} & \quad \text{for all } j = 1, 2, \ldots, n
\end{align*}
\]

where
\[
\begin{align*}
f_1 &= \frac{1}{m} \sum_{i=1}^{m} w_i p_i(X, Y, Z) \\
&+ \mu_1 \max \{0, p_i(X, Y, Z) - s_i\} \\
f_2 &= \frac{1}{n} \sum_{j=1}^{n} \max_{i \in S_j} w_i [p_i(X, Y, Z) \\
&+ \mu_1 \max \{0, p_i(X, Y, Z) - s_i\}]
\end{align*}
\]

and \( X = (x_1, x_2, \ldots, x_n), Y = (y_1, y_2, \ldots, y_n), Z = (z_1, z_2, \ldots, z_n). \) \( (x_i, y_i, z_i) \) is the current position of the base station in the \( i \)th hyper-rectangle \( H_i \) of the building. \( s_i \) is the required signal power at a UE location and \( p_i \) is the calculated signal power at \( m \) receiver points using the 914MHz path-loss model described in Rappaport [10]. \( \mu_1 \) are the penalty functions defined at each UE location.

When using a signal to interference based formula there is a significant change of constraints and of the functions \( f_1 \) and \( f_2 \).

Firstly the path-loss model is replaced with a ray electric field signal power model. \( p_i(X, Y, Z) = (\|R(r_1 - r_{ik})\|, \|R(r_2 - r_{ik})\|, \ldots \|R(r_m - r_{ik})\|) \) is defined to be the ray field strength from \( m \) transmitters to the receiver point \( r_{ik} \) that has best server \( k \). The constraints \( p_i > s_i \) are replaced with signal to interference ratio constraints of the form \( SIR > tSIR \) where \( tSIR \) is a target interference ratio to be met, described in Section III-E. The modified objective function \( f_1 \) and \( f_2 \) take the form:

\[
\begin{align*}
f_1 &= \frac{1}{m} \sum_{i=1}^{m} \gamma_i [p_i(X, Y, Z) \\
&+ \mu_1 \max \{0, tSIR_i(X, Y, Z) - SIR_i\}] \\
f_2 &= \frac{1}{n} \sum_{j=1}^{n} \max_{i \in S_j} w_i [p_i(X, Y, Z) \\
&+ \mu_1 \max \{0, tSIR_i(X, Y, Z) - SIR_i\}]
\end{align*}
\]

For the case of an omni-directional antenna the constraint equation is given by:

\[
SIR_i = \frac{P_i \|R(r_i - r_{ij})\|}{\sum_{k \neq i} P_k \|R(r_k - r_{ij})\| + \eta_t} > tSIR_i
\]

where \( r_i \) is the location of the \( i \)th base station, \( r_{ij} \) is the \( j \)th UE position that has best server \( r_i, P_i \) is the power transmitted from the \( i \)th server. \( \|R(r_k - r_{ij})\| \) is the total ray-based field strength prediction obtained by calculating a vector norm on all ray-path field strengths between server \( r_k \) and the point \( r_{ij} \). The power \( P_i \) is set to be constant obtained by ensuring that the base stations do not interfere with one another. The ray-tracing technique is exactly the same as the one defined in Kenny [5]. The power control derivation is described in the next section.

3) **Power Control:** In the case of an omni-directional antenna the formula for obtaining the transmitting power is defined by equation (5), where \( R_{min}^i \) is defined to be the minimum field strength between the \( i \)th base station and every other base station written as:

\[
P_i = R_i I_m R_{min}^i
\]

\[
R_{min}^i = \min_{\forall j \neq i} \|R(r_i - r_j)\|
\]

If the antenna is sectorized with \( m \) equally sized sectors \( a_k^i \in \{a_1^i, a_2^i, \ldots, a_m^i\} \) then the power is multiplied by a gain \( G_k \) in the \( k \)th sector \( a_k^i \) of the \( i \)th base station to define the transmit power for that sector. The minimum field strength for each sector is then defined as:

\[
R_{min}^{ik} = \min_{\forall j \neq i} \|R(r_i - r_j)\| \|a_k^i
\]

where the ray-path based field strength prediction is restricted to the sector \( a_k^i \).

We wish to set the gains \( G_k \) so that the power from each sector \( a_k^i \) are given by \( P_{ik} = R_i I_m G_k R_{min}^i \) are divided equally amongst each sector. The formula for the gain can then be defined as

\[
G_k = \frac{R_i}{m R_{min}^{ik}}
\]

where \( R_i = \sum_{k=1}^{m} R_{min}^{ik} \) is the total field strength for the base station. This has the effect that each sector emits \( P_i/m \) of the power where \( P_i \) is the total power output of the server. In the case of a sectorized antenna the signal to interference ratio constraint on the optimisation changes to:

\[
SIR_i = \frac{\sum_{k=1}^{m} G_k P_i \|R(r_i - r_{ij})\| \|a_k^i\|}{\sum_{k=1}^{m} \sum_{l \neq i} G_k P_l \|R(r_i - r_{ij})\| \|a_k^i\| + \eta_t} > tSIR_i
\]

4) **Stopping and Line Search Criterion:** The stopping criterion is usually related to the gradient of the objective function as already stated. In this case however an extra piece of information is added to the loop controlling the stopping criteria. We say that if the number of receivers reaching their target signal to interference ratio is above a certain percentage, then we will stop the algorithm. In practice it was found that 90% of receivers reach the required target SIRs.
Also we found that when performing the line search, we usually half the step length until we get convergence in the gradient calculation. We found that if you get an increased number of receivers reaching the target signal to interference ratio, then we should stop the line search at this point, and use the new positions as the next point in our continuing optimisation.

In the line search algorithm it is necessary to calculate finite difference derivatives to obtain the steepest descent path for the gradient method. The derivatives take the form:

$$\xi_{ij} = \frac{f(r_i + \delta e_j) - f(r_i)}{\delta}$$  \hspace{1cm} (20)

giving a steepest descent direction \(V_{sd} \):

$$V_{sd} = \max_{i,j} \xi_{ij}$$  \hspace{1cm} (21)

where \(e_j \) is a unit vector pointing in the \(x, y \) or \(z \) direction and \(\xi_{ij} \) is calculated for every base station position \(r_i \). An exception to this rule is when the steepest descent direction has the effect of causing the server capacity to be exceeded. In this case the next best direction is assumed.

Rappaport used a value of 0.5 for \(\delta \) for his methods and halved this twice in the next two steps to get the algorithm to converge nicely. The value was then reset again and the process would start all over again. We noted that this value should be set to be something highly physical when moving from one position to another. We know that a signal can fall sharply from one position to another by just moving centimetres at a time, so because of the sampling theorem we guessed that \(\delta = \lambda/2 \) would be a good starting point for the value of \(\delta \). In practice it was found that the method could become unstable if too small a value of delta was used since that would result in a line search algorithm that is moving around in fast fades. The results turned out to be meaningless in this case.

5) **Annealing:** It is presumed that the algorithm is in a local maximum if there is no significant change in the number of receivers meeting the signal to interference ratio, and if this happens, a step length of \(\delta = 2\lambda \) is taken and the position with the highest number of receivers meeting the SIR is assumed to be the new starting point for the algorithm.

In some cases all directions for the steepest descent calculation result in the server capacity being exceeded. In this case the annealing step must be introduced to correct the problem.

6) **Hard Handover:** It is possible in some cases to find a position of the base stations where one server has exceeded its capacity requirement and then use a hard handover to another server to balance the servers. It was something that we looked briefly at, but could make for a more promising algorithm in the future. Currently in the optimisation tool it is ignored.

7) **Mean Field Strength Values:** The mean of the total field strength for a point to point link can be calculated using the formula for the mean of a random walk. This involves taking each complex ray path \(E_{ij} \) separately between the \(i^{th} \) base station and \(j^{th} \) receiver and substituting them into a random walk formula. It should be noted that taking the random walk result yields a far more constrained system than just taking the sum of the complex ray paths. This can be seen from looking at the SIR formula:

$$\text{SIR}_t = \frac{P_i \sum E_{ij}}{\sum_{k \neq i} P_k \sum E_{kj} + n_t} > \text{tSIR}_t$$  \hspace{1cm} (22)

The sum of complex numbers in the denominator tend to cancel out in phase giving a smaller absolute value of the field strength, and in turn the target signal to interference ratio tends to be met more easily. We cannot assume this is true in reality, since there may exist building measurement error, incorrect permittivity or neglected complexity of the building and therefore we cannot exactly predict the phase of the complex number. The random walk mean value turns out to be physically more meaningful.

The total power at the mobile terminal (UE) is obtained by setting some norm on the individual ray-path field strength predictions. Assuming the use of the Euclidean norm is not always correct. For instance, in a real environment there are effects due to scattering from walls that we cannot predict exactly. They can be assumed to be random, so that the phase of the field strength leaving any scattering object is altered. If we want a measure of the received power at the mobile terminal, we can say that the result is obtained from a set of random variables. A good measure of the power is the mean power obtained at the receiver including gaussian random phase. An analogous way of looking at this is that we want the average power about some point so that we move the mobile terminal randomly around about its current position to avoid fading. Takahashi [13],[14] describes a random phase summation applied to ray-tracing which is applicable for our problem.

The power loss along any ray path is a complex number of the form:

$$x_k + jy_k = E_k \exp(j\phi_k)$$  \hspace{1cm} (23)

and \(r \) is the sum of \(N \) such random variables given by

$$r = \sum_{k=1}^{N} E_k \exp(j\phi_k)$$  \hspace{1cm} (24)

where \(E_k = \sqrt{x_k^2 + y_k^2} \) is the amplitude of the power and \(\exp(j\phi_k) \) is the phase, \(\phi_k \) is a set of independently and identically distributed uniform random variables in the range \([0, 2\pi]\).

The probability density function giving rise to the mean of the sum of the random variables is of the form

$$W_N(r) = r \int_0^{\infty} \xi J_0(\xi r) \prod_{k=1}^{N} J_0(\xi E_k) d\xi$$  \hspace{1cm} (25)

The mean of the random variable is then given by

$$\tau = \int_0^{\infty} r W_N(r) dr$$  \hspace{1cm} (26)

It can be seen from the above equations (25) and (26) that the mean of the random variable is a doubly infinite integral which is quite time consuming to solve numerically using a quadrature rule with a set discretisation size. The alternative
is to calculate the random variable \( r \) given by equation (24) \( n \) times, where \( n \) is very large and then apply the formula

\[
r = \frac{1}{n} \sum_{i=1}^{n} r_i
\]  

(27)

where \( r_i \) is the result of the \( i^{th} \) random variable calculation. It turns out that this procedure is even longer to calculate than the doubly infinite integral because it requires that \( n \) is of the order of 100000 so that the error in the random mean is guaranteed to be 1%.

We found that if you define a numerical quadrature rule of order 2, that is the recursive composite trapezoidal rule, we can apply this formula to the calculation of the \( W_N(r) \) integral, converging to the correct solution quickly because of the recursive nature of the rule (not requiring recalculation of any discretization points already included in the formula).

Now the computation of the outer integral must be obtained. The algorithm is as follows:

1) Set a step size \( d_1 \) for the numerically finite integral \( I_n \) in the range \([0, nd_1]\) which will approximate the infinite integral equation (26) and assume an euler rule so that \( I_n \) takes the form

\[
I_n = d_1 \sum_{i=1}^{n} r_i W_N(r_i)
\]

(28)

2) At each step of the calculation of the euler rule, we much calculate \( r_i W_N(r_i) \) for \( r_i = \{d_1, 2d_1, \ldots, nd_1\} \) until

\[
\frac{r_n W_N(r_n)}{I_{n-1}} < \text{tol}
\]

(29)

When the tolerance is met, the integral has reached its desired accuracy and stops at that point.

This method is confirmed by comparing the numerical result with the result of equation (27).

G. Uplink Optimisation Requirements

The uplink optimisation is a simulation resulting in a number of users being able to connect to the closest base station (best server) and achieving their full capacity requirement or some percentage of their requirement.

It was decided that two types of simulations of a working system would be required. One where there is an initial set of users and more are connected or one where there are no users on the system and each user is connected one at a time. It turns out that the former is a subset of the latter by inspection so it was assumed that the population starts at nil and increases one at a time.

For any UE the following information is required:

- **Number** - An index to the mobile terminal.
- **Type** - The type of terminal either Voice, real-time data or non real-time data, will inform the user of the maximum data rate achievable for that type of use, and the height above ground of the user. For instance a PC is approximately one metre above the floor level, while a voice link is approximately 1.5 metres above floor level.

- **Reference Location** - This is an \((x, y, z)\) coordinate which tells where in a particular story the terminal exists.
- **Final Location** - If the terminal moves then this is the position at its final state.
- **Maximum/Minimum Power** The maximum power or minimum power output from the UE must be known before commencing the initial population selection.

H. Uplink Optimisation Algorithm

The algorithm for choosing the population is as follows:

- The reference location \( P_i(x, y, z) \) for the UE is generated by setting \( x, y \) and \( z \) to be independent uniformly random variables. This location will be an item which if successfully connected will be pushed onto a linked list, otherwise it will be dropped and disregarded.
- Generate the UE position, if it is contained in a free convex space, i.e. not inside a wall, then it is a valid position at its final state.
- The maximum power or minimum power output from the UE must be known before commencing the initial population selection.
- Determine the best server for the current UE.
- The reference location \( P_i(x, y, z) \) for the UE is generated by setting \( x, y \) and \( z \) to be independent uniformly random variables. This location will be an item which if successfully connected will be pushed onto a linked list, otherwise it will be dropped and disregarded.
- Generate the UE position, if it is contained in a free convex space, i.e. not inside a wall, then it is a valid position at its final state.
- The maximum power or minimum power output from the UE must be known before commencing the initial population selection.
- Determine the best server for the current UE.
- The reference location \( P_i(x, y, z) \) for the UE is generated by setting \( x, y \) and \( z \) to be independent uniformly random variables. This location will be an item which if successfully connected will be pushed onto a linked list, otherwise it will be dropped and disregarded.
connected to that best server as can be seen from the missing term in equation (32).

- Now that all the information is available, such as position of the UE, maximum power and SIR information it is important to check that the capacity of the system can be met. The following steps need to be adhered to:
  1) If the signal to interference ratio is met then proceed to step 2, otherwise stop.
  2) If the total capacity in the current server is not exceeded and the total capacity will be met after adding the new UE, then proceed to step 3. Otherwise half the capacity requirement and try to meet the total capacity upper bound again. Failing this an attempt is made to use the second best server and if the capacity requirement is met, again proceed to step 3.
  3) The capacity is met and connection is possible to the server so the UE is then connected to the system.
- Once the server is connected to the system, an attempt is made to add new connections, until all UEs are added.

This algorithm was run $N$ times for a large number $N$ to ascertain an average number of connections. So if the number of connections was $x_k$ on the $k^{th}$ attempt then the average was calculated using the formula:

$$\tau = \frac{\sum_{k=1}^{N} x_k}{N}$$  \hspace{1cm} (33)

I. Computational Savings

Another point that is important to consider is that as the number of UE connections increases, the calculation time of the signal to interference ratio in equation (31) and (32) will greatly increase, if the values are recalculated every time. Avoiding this overhead is quite simple if the signal to interference ratio is stored. The interference part of the SIR can be extracted and updated as new connections are established. Assuming that the $i^{th}$ connection has just been added to the system each existing SIR$_k$ for location $r_k$ with best server $r_l$ must be updated for $k = 1,2,\cdots,i-1$. This can be achieved as follows:

$$\text{interference}^{\text{old}} = \frac{P_{kl}R_{kl}}{\text{SIR}_k} - n_t$$  \hspace{1cm} (34)

$$\text{interference}^{\text{new}} = \text{interference}^{\text{old}} + P_{ij}P_{il}$$  \hspace{1cm} (35)

$$\text{SIR}_k = \frac{P_{kl}R_{kl}}{\text{interference}^{\text{new}} + n_t}$$  \hspace{1cm} (36)

Also the interference to the current server must be stored as SIR$_i$. The total capacity in each server and number of connections currently to each server may also be stored to increase efficiency in the algorithm.

IV. RESULTS

A. Experiment 1: Open Foyer Problem with Three Transmitters

Figure 2 illustrates a foyer with 3 high capacity areas requiring data links (26 UEs) marked in yellow and 95 low capacity links marked in red. This experiment was used to rigorously test the correctness of the optimisation algorithm. For this simple configuration with a total capacity of 5 megabits, the splitting algorithm of Rappaport splits across the centre of the $x$-axis first, to give two hyper-rectangles each containing approximately 2.5 megabits, and then splits each of these across the $y$-axis to give a total of 4 hyper-rectangles each containing about 1.25 megabits. All antennae are assumed to be omni-directional.

![Fig. 2. Three high capacity areas marked with large yellow circles (26 128kbps connections), with 95 low capacity voice links marked with red smaller circles (12.2kbps) in a single building foyer.]

![Fig. 3. Three hyper-rectangles and centres of capacity generated using new algorithm.]

The new algorithm described in Section IV-E1 determines that initially split horizontally and yield 4 base station hyper-rectangles while splitting vertically initially will yield 3 base stations. The smaller number is preferred and therefore chosen. The result of the splitting is shown in Figure 3.

This example was specifically chosen to show how the method will define a smaller number of base stations to start with. For many tests without exception we found that this is true. Three areas of high capacity where chosen as shown in Figure 2 because the splitting algorithm is non-trivial in...
this case. If a symmetric problem was chosen the method of Rappaport would almost certainly give the same answer as the updated algorithm. These three areas represent hot-spots where we want to try and guarantee coverage with high data rates. It is apparent when looking at the uplink optimisation part of the algorithm that users are not necessarily restricted to using high data rates in these areas. They may require high data rates elsewhere. The downlink positions of high capacity serve as a guide to how efficiently the system will meet the customers needs.

Next the downlink optimisation can be performed to achieve the optimal positions of the base transceiver stations (BTS). The chosen frequency for the simulation is 2GHz which is used for current Bluetooth and WIFI technologies. The technique was able to obtain 92.5% of the required coverage for the user capacity requirements shown in Figure 2. The ray-tracing was set initially to have reflections of order 1, and then using the optimal positions obtained from this optimisation the algorithm was rerun with reflections of order 2. This is a useful way to run the algorithm, because it acts as a fine tuning in each increment. The best servers and their coverage are shown in Figure 4. The orange, light orange and yellow circles correspond to the coverage points of servers 1, 2 and 3 respectively. The black smaller circles are the receiver positions that did not meet their SIR requirements.

Although the starting position of the transmitters reaches 82% of the capacity requirement it was noted that it is possible to meet 92.5% of the capacity requirement using the algorithm described in this paper. As expected the centre of gravity hyperspaces give a good starting solution. It is particularly notable as shown in Figure 3, that the transmitters attempt to move away from their centres of gravity towards the nearer wall in the case of the bottom two transmitters. This is because the optimisation algorithm is attempting to reduce the affect of the signal to interference ratio (SIR). This open foyer problem was chosen specially because of its lack of inner wall boundaries that reduce the effects of signal to interference ratio. The moving away of transmitters from each other is to be expected and physically explainable in light of the SIR effects.

To find the absolute theoretical optimal location for the base stations would require an order \(O(N^3)\) computation where \(N\) is the number of possible base station locations defined on a regular grid. This computation could take a matter of days to run, whereas the downlink optimisation algorithm runs in just under 34 minutes on an Intel i7-3770 processor.

The overall computational time is not dominated by the ray-tracing calculation which only account for 1/8 of the computational time. The random mean calculation at each point although numerical faster than calculating the doubly infinite integral is still relatively slow in the overall computation.

**B. Experiment 2: Unrealistic Oversaturated Environment**

The downlink and uplink optimisation problem is set up exactly the same as in Experiment 1 except this time it was applied to a real building, the Printing House at Trinity College Dublin. As in the related work by Kenny [5] the building data was input by hand for each convex space and taken from draftsman drawings. A fictitious and over subscribed user capacity requirement of 92 voice data users and 60 data users.
is displayed in Figure 5 is set for the building that results in the requirement of 8 antennas. All antennae are assumed to be omni-directional. If the result of the first experiment are correct we would then assume that we will not be able to achieve even 90% of the user capacity requirement when finishing at the most optimal set of best server locations. This is indeed the case given the initial positions as defined in Figure 6. In fact, it was only possible to evolve to a solution where 46/152 = 30.26% of the capacity requirements are met. At this point the optimisation algorithm becomes completely stuck and cannot evolve to a better solution even with annealing applied. One way to circumvent this problem is to assume that the users only require 60% or 40% of their capacity requirements for data and assume a 100% requirement for voice links. But this is not the point of the algorithm. We are trying to plan a system where the requirements are met given the initial constraints and see how the system performs in the scenario provided.

An unrealistic set of user capacity requirements results in the optimisation of 8 antenna positions in a small vicinity. This problem was designed to test the assumption that an evolution to a set of best servers will improve greatly if the planning scenario is well founded to begin with.

C. Experiment 3: Realistic Capacity Requirement

In this experiment as in Experiment 2, the Printing House at Trinity College Dublin was used as the environment, but this time the user capacity requirement was a realistic set of values with 2-3 data rate access points per room and 2-6 voice data points in the same rooms or adjoining rooms as shown in Figure 7. In total, there is a requirement for 44 users of the system with capacity for 19 mobile phones using 2G voice and 25 mobile phones using 3G data rates (see Figure 7). All antennae are assumed to be omni-directional.

The optimisation algorithm was found to converge nicely to an optimal solution in the case of the single reflection and double reflection ray-tracing model within the optimisation technique. In both cases 97.7% (or 43 of the 44 points) can obtain their target capacity. Two antennas are required to meet the target capacity requirements for this sample problem. One antennae centre of gravity is in the bottom left room in the top right corner, whilst the other is a good way along an open corridor as shown in Figure 7. It is notable that the antennae do not move that much but they do move away from each other as the optimisation algorithm proceeds. The left most antenna moves towards the outer extremity of the building whilst the second one moves towards the right most end of the corridor. This essentially is allowing the antennae to meet all the capacity requirements whilst not interfering with each other. Also worth noting is the movement upwards of the transmitters towards the ceilings during the optimisations line search algorithm. By moving upwards the transmitters are clearly combating the signal to interference ratio problems more easily. This was not immediately obvious to us as designers of the system, but makes perfect sense on closer inspection.

Also, worth noting is the fact that the centre of capacity starting points only result in 15/44 = 34.1% of the capacity requirements being met, so the improvement obtained by the optimisation algorithm is drastic in this use case.

It was interesting to observe that running the optimisation in the following order did not achieve convergence in a realistic environment:

1) Place transmitters at centres of gravity for capacity requirements \(\{r_i^0\}_{i=1}^r\).
2) Perform a 1\(^{st}\) order ray-trace based optimisation: iterate optimisation with 1 reflection and stop at point of convergence and target signal to interference ratio \(\{r_i^1\}_{i=1}^r\).
3) Perform a 2\(^{nd}\) order ray-trace based optimisation: starting at the transmitter positions \(\{r_i^1\}_{i=1}^r\) (from step 2) rather than centres of gravity \(\{r_i^0\}_{i=1}^r\) (from step 1), iterate the optimisation algorithm with 2 reflections until the stopping criterion is met.

It was found that the above method meets the stopping criterion step 2, but never converges for step 3. However, if in step 3, the centre of gravity is chosen as the starting point for the technique, it was found that the method always converges, albeit more slowly. The only rational explanation for this is that the local minimums obtained from step 2 are too difficult to escape from and that the moving transmitters even with annealing applied return to the same points for the 2\(^{nd}\) order optimisation.

Another seemingly troubling aspect of the algorithm in its initial implementation was that 88.6% (38 of 44) of the capacity requirements were met at the stopping criteria of the optimisation algorithm. The figure is lower than expected given that we would expect the annealing algorithm to help the optimisation to escape from the local minimums. To combat this problem the negative steepest descent direction was applied just once to escape the local minima. Such a problem can arise in highly non-linear problems such as the one addressed here. The negative direction strategy was previously implemented by such authors as Moiseev [7] when an optimisation direction is proceeding quickly towards a boundary from which there is no escape.

The setup in Experiment 1 does not have the same convergence issues as in Experiment 3 due to the simplicity of the building geometry. A foyer is a simple convex space whereas a more realistic building, such as in Experiment 3 consists of quasi-convex spaces with hard borders that strongly effect the signal power and signal to interference ratio. A transmitter approaching and passing through the boundary of a filled convex space results in large changes in received signal power. These large fluctuations in signal power were not observed in Experiment 1.

V. DISCUSSION

One clear observation of the optimisation algorithm is the difference between using the absolute sum and the mean value of the complex ray-paths. In the case of near saturation, the mean value calculation will tend to result in more receivers achieving the capacity requirement, whereas a less crowded network will result in the absolute sum yielding a higher result. In physical terms we would expect that the mean value of
a random signal is more reliable since we are dealing with building draftman errors and clutter in the building such as desks, chairs and cabinets that are not accounted for in the current model.

The objective function of the non-linear program is related to a measure of signal to interference ratio. In systems such as CDMA with FDD, a path-loss based objective function similar to that of Rappaport’s could be used, however, parameters described by Nokia or some similar parameters would still be required to define realistic system parameters.

It is interesting to note that the optimisation method using a 2nd order reflection ray-tracing method could not meet its capacity requirements when transmitters are placed at the best server points obtained from the 1st order reflection ray-tracing method in some situations. This is likely due to transmitters falling into local minimums which they cannot escape from. Placing transmitters at the centres of gravity in increasing higher order ray-traces does yield reliable convergence in most cases, but the computational time does increase quickly also.

Use of more complicated antenna designs would result in an increase in the computational cost of the algorithm as per equations (17)-(19) due to the inclusion of additional summations. The computational cost in the line search algorithm will not be affected since the number of search directions does not change.

VI. CONCLUSIONS

The novelty in the optimisation technique used in this paper is in the initial placement of the base stations, the writing of the problem as a signal to interference problem, the use of realistic parameters, the use of a better norm on the field strength values (random mean field strength), the introduction of more complicated antenna types and the increased refinement in the number of reflection/diffraction paths as the transmitters evolve to their optimal positions.

Using a ray-tracing algorithm whose visibility algorithm is defined by a spatial decomposition of a building and does not depend on transmitter location allows the optimisation algorithm to run with reasonable computational times. If the computational times were too slow for the field strength calculations then empirical methods such as path-loss models could be used instead. Such empirical models do not account for destructive interference and summing of ray-paths to arrive at a higher signal power. This was not the case and this paper shows conclusively that ray-tracing has a place in future optimisation techniques.

The initial experiments confirm that the technique works well and show improvement over the centre of gravity initial
start points converging well in realistic planning environments. The final transmitter locations are not always immediately intuitively understood but are explainable upon closer inspection.

ACKNOWLEDGMENT

We would like to acknowledge the invaluable advice and help of the late Dr. Gerry Butler in this project.

The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

REFERENCES