PRECISION PLANNING FOR 5G ERA NETWORKS WITH SMALL CELLS
230.10.01

Precision planning for 5G Era networks with small cells

October 2019
Small Cell Forum accelerates small cell adoption to drive the wide-scale deployment of small cells and encourage the delivery of fully integrated HetNets.

We are a carrier-led organization. This means our operator members establish requirements that drive the activities and outputs of our technical groups.

We have driven the standardization of key elements of small cell technology including Iuh, FAPI, nFAPI, SON, services APIs, TR-069 evolution and the enhancement of the X2 interface.

Today our members are driving solutions that include small cell/Wi-Fi integration, SON evolution and automation, virtualization of the small cell layer, driving mass adoption via multi-operator neutral host, ensuring a common approach to service APIs to drive commercialization and the integration of small cells into 5G standards evolution.

The Small Cell Forum Release Program has now established business cases and market drivers for all the main use cases, clarifying market needs and addressing barriers to deployment for residential, enterprise, rural & remote, and urban small cells.

The Small Cell Forum Release Program website can be found here: www.scf.io

All content in this document including links and references are for informational purposes only and is provided ‘as is’ with no warranties whatsoever including any warranty of merchantability, fitness for any particular purpose, or any warranty otherwise arising out of any proposal, specification, or sample.

No license, express or implied, to any intellectual property rights is granted or intended hereby.

©2007-2017 All rights reserved in respect of articles, drawings, photographs etc published in hardcopy form or made available in electronic form by Small Cell Forum Ltd anywhere in the world.

If you would like more information about Small Cell Forum or would like to be included on our mailing list, please contact:

Email info@smallcellforum.org

Post Small Cell Forum, PO Box 23, GL11 5WA UK

Member Services memberservices@smallcellforum.org
About 5G Americas: The Voice of 5G and LTE for the Americas

5G Americas is an industry trade organization composed of leading telecommunications service providers and manufacturers. The organization’s mission is to advocate for and foster the advancement and full capabilities of LTE wireless technology and its evolution beyond to 5G, throughout the ecosystem’s networks, services, applications and wirelessly connected devices in the Americas.

5G Americas is invested in developing a connected wireless community while leading 5G development for all the Americas. The association prepares numerous white papers and reports on key initiatives, challenges and opportunities for the industry with a particular focus on those of greatest importance to North America and Latin America and the Caribbean. The reports are accessible on the association’s website www.5gamericas.org.

5G Americas’ Board of Governors members include: América Móvil, AT&T, Cable & Wireless, Cisco, CommScope, Entel, Ericsson, HPE, Intel, Kathrein, Mitel, Nokia, Qualcomm, Sprint, T-Mobile USA and Telefónica.

5G Americas collaborates with government and regulatory bodies, as well as standards organizations and other associations throughout the world to foster the successful development of wireless technologies.

- 5G Americas, in partnership with 5G PPP, 5GMF, 5G Forum and IMT-2020, participates in hosting the Global 5G Event on a bi-annual basis.
- 5G Americas is a Market Representation Partner at 3GPP.
- 5G Americas has agreements, memberships, relationships and/or Memorandums of Understanding with: 5G Forum, 5G PPP, ASETA, ASIET, ATIS, CASETEL, CINTEL, CITEL, ITU, NGMN Alliance, Small Cell Forum and the UMTS Forum with the purpose of cooperation to improve mobile communications.

Follow 5G Americas news on Twitter at @5GAmericas and Facebook at www.facebook.com/5gamericas.
Executive summary

The ever-increasing demand for mobile data is driving network densification with the deployment of small cells. Although lower cost than macro towers, the compact, low-power nature of small cells means they also have a smaller serving area. This in turn means they need to be located closer to demand hotspots in order to be effective in supplying traffic efficiently and delivering a good return on investment (ROI).

The benefit of any given small cell site essentially boils down to the amount of traffic it carries, whereas the cost is tied to installation, power, backhaul, rent, etc., and is largely independent of its utilisation. It follows that the value (or ROI) of a given site is affected by the precision with which it can be planned and sited. Small cells need to be located close to areas of high traffic demand that are not well served by the existing network, but they need to be placed with care. Recommended best practice is to locate the small cell within 20-40m of the ideal.

A key input to the precision planning process is geolocated measurements of network quality, which are used to build up a map of how well the existing network is serving traffic demand. Small cells should be deployed where there is high traffic demand but low signal quality.

The location accuracy of the quality reports contributes to the overall precision of the planning process. Averaging of observed time difference of arrival (OTDOA) network location estimates was found in one study to have a 60m median error, limiting its applicability to small cell planning. Smarter analysis of the locate data using a machine learning approach was found to reduce median location error down to 18m, which is within the required range.

Finally, an example of small cell design is shared. Using data and the goals of coverage and dominance, small cell placement is determined using an automated process. The example shared in this document provides improved coverage and reduced costs over a manual design.

Recommended best practice for precision planning

- For maximum ROI, small cells should be placed as close as possible to demand peaks; best practice is within 20-40m.
- MNOs would like equipment that estimates location of usage and quality reports to adopt smarter algorithms such as the machine learning approach demonstrated. Median locate errors less than 20m are expected for small cell planning purposes.
- Machine learning models should be part of any small cell design effort. Different inputs and assumptions will be factors in the resulting models that are generated.
Contents

1. **Introduction: Why precise location of small cells is important** ................................................................. 1
2. **Network planning methodology and accuracy requirements** ................................................................. 3
   2.1 Indoor/outdoor differentiation and 3D location ............................. 3
   2.2 Outdoor planning ................................................................. 4
   2.3 Indoor planning ................................................................. 14
   2.4 Accuracy of network quality data sources for planning ........... 19
3. **Impact of siting accuracy on ROI** ....................................... 20
4. **Enhancing the accuracy of UE location with machine learning** .............................................................. 21
   4.1 Types of UE location ............................................................. 21
   4.2 Using machine learning to create ‘enhanced’ timing-based solutions for improved location accuracy .................. 22
   4.3 Using ML to improve timing-based location techniques ...... 22
   4.4 Calculating the relative timing offset between cell sites for use in location calculations and effective inter-site synchronization ......................................................... 24
   4.5 Machine learning algorithms toward improved accuracy and results with OTDOA ................................................. 25
5. **Improving small cell placement using machine learning** ............................................................................ 28
   5.1 Factors included in ML model .................................................. 28
   5.2 Cost and coverage improvement using ML .............................. 28
6. **Acknowledgements** ............................................................... 31
References .................................................................................. 32

Figures

Figure 1–1  Matching small cell locations to hotspots ................................................................. 1
Figure 2–1  Differentiating between outdoor and indoor payload forecast in 2D ... 3
Figure 2–2  Vertical distribution of traffic within buildings ......................................................... 4
Figure 2–3  Per building traffic ................................................................................................. 4
Figure 2–4  Los Angeles DVHM (Digital Vector Height Model): submeter vertical and horizontal digital model of the environment, including elements that greatly impact the propagation of 5G signals, such as trees and buildings ................................................................................................. 5
Figure 2–5  Los Angeles Composite Demand Density Map. Demand is not uniformly distributed. Hotspots appear as redder areas, each building
presents a unique demand profile and roads are also characterized by mobility patterns ...........................................5

Figure 2–6 Los Angeles RSRP showing macro and small cells (AWS, Tier 1 operator) .................................................................6

Figure 2–7 Manhattan RSRP showing macro and small cells (AWS, Tier 1 operator) .................................................................7

Figure 2–8 Close up of Manhattan RSRP showing macro and small cells (AWS, Tier 1 operator) .....................................................7

Figure 2–9 3D view of Brooklyn DHVM (Digital Vector Height Model): submeter vertical and horizontal digital model of the environment, including elements that greatly impact the propagation of 5G signals, such as trees and buildings .................................................................9

Figure 2–10 Reston, VA RSRP. The interaction of the outdoor signal with the multiple elements of the environment must be precisely modelled, specially to understand its impact on indoor coverage at different heights ........................................................................................................10

Figure 2–11 Automated optimization of 5G cell reach: complex cumulative of connections at 5G mmWave frequencies (FWA and IAB). Red = lower latency; Green = higher latency .........................................................14

Figure 2–12 Indoor capacity coverage map .................................................................16

Figure 2–13 A 120º sector with seven narrow beams (16º each) .........................17

Figure 2–14 Composite coverage of a 120º sector with 35 narrow beams ..........17

Figure 2–15 Indicative location accuracy for different data sources .................19

Figure 3–1 Study of the effect of ROI with distance from the ideal [Source: Nokia] ..................................................................................20

Figure 4–1 Example of GPS accuracy degradation in downtown areas (in this case Dallas, Tx). Green = over 85% at 40m or better, Yellow / Orange = 50 to 85% < 40m, Red = less than 50% < 40m ..................21

Figure 4–2 Dallas metro collection area ................................................................22

Figure 4–3 One ‘set’ of locations derived from OTDOA pairs (blue circles), and the assumed known location (green circle) .............................................23

Figure 4–4 Percentage of time with fewer than N cell pairs producing location estimates at a given location ..................................................24

Figure 4–5 Patented approach toward synchronizing cell sites ..........................25

Figure 4–6 Location estimation when there are at least 10 cell pairs using simple averaging, AT&T’s machine learning algorithm, and a ‘genie’ (closest location estimate from the cell pairs to the known location) ........26

Figure 4–7 Example of using the improved OTDOA location accuracy in downtown areas (in this case Dallas, Tx). Green = over 85% at 40m or better, Yellow / Orange = 50 to 85% < 40m, Red = less than 50% < 40m ................................................27

Figure 5–1 Factors included in ML generation of Sites ....................................28

Figure 5–2 Manhattan Manual vs. Automated Designs .....................................29

Figure 5–3 Coverage (RSRP) in Manhattan in Manual vs. Automated Designs ...29

Figure 5–4 Dominance in Manhattan in Manual vs. Automated Design ..........30
1. Introduction: Why precise location of small cells is important

Network planning and optimization is becoming ever more important as small cells, placed in hotspots, are increasingly used to address capacity needs. But the location of such small cells is itself critically important. Where small cell placement was off by as little as half a cell radius from a given hotspot, the result – according to one operator – was that four times as many small cells were required to carry the same traffic. Location accuracy is even more challenging when higher frequency bands are used, reducing the practical serving radius of the cells.

Figure 1–1 illustrates the challenge of locating small cells to align with hotspots of demand not already served by the existing network, as indicated by the people on the upper graph. The purple line indicates the potential utilisation of the capacity of a small cell when placed at a given location.

![Figure 1–1 Matching small cell locations to hotspots](image)

A well-placed small cell has its serving area covering locations with high demand. As a result, most or all of its capacity will be highly utilised. This will represent a good return on investment for the company deploying the small cell. A poorly located small cell is too far away from the hotspot; its serving area does not cover an area of high demand. A further complication is that the efficiency with which any cell site can serve demand diminishes with distance from the antenna. This is indicated in the diagram by the fading colour of the serving area. The loftier objective of aligning the most spectrally efficient parts of the cell serving area to the peaks of demand further increases the required siting precision.

The size of this serving area is linked closely to the use case of the small cell. In urban scenarios small cells have to date been used to increase capacity in city squares, in a series of adjacent bus stops, or around shops, cafés and restaurants.

Getting the location right requires an understanding of the precision required in the planning process for small cells. This paper considers the accuracy of different parts of the process. It is arranged as follows:
Section 2 details the network planning process for small cells. This will begin with an overall understanding of where the network requires additional coverage and capacity. Depending on the location, the network may best be enhanced through either outdoor densification or indoor deployment. However, the considerations involved in each case will be quite different – as described in sections 2.2 and 2.3.

Section 3 shares the results of one study into the impact of siting accuracy on return on investment (ROI), and recommends best practice accuracy.

A key input to the planning process is an understanding of how well the existing network is serving demand – in effect the purple residual traffic demand curve in Figure 1–1. This understanding can be built up from information on geolocated traffic usage and from quality reports delivered by devices in the network. Such reports are available from probing network interfaces or from data collected from social media, mixed with network statistics. The accuracy of the associated locates determines the accuracy of the demand curve – and therefore the operator’s understanding of where hotspots really are. Section 4 considers this location accuracy – and, in particular, the use of machine learning to enhance the location accuracy of traffic usage and quality reports used in the planning process.

Even with highly accurate planning to identify the optimal location for a small cell, there are real-world limitations which mean they cannot be used in practice. Candidate sites for small cells – such as street furniture or indoor mounting – may not be available in ideal locations. Furthermore, traffic demand distributions vary daily, weekly and over years, so hotspots may move over time. Although optimisation through RAN configuration changes can help compensate for suboptimal locations, it cannot achieve the same performance as using the optimal sites in the first place. These challenges are discussed further in our paper [SCF195] [1] Small cell siting challenges and recommendations.
2. Network planning methodology and accuracy requirements

2.1 Indoor/outdoor differentiation and 3D location

A detailed description of the planning of heterogenous networks – networks that combine small and macro cells – can be found in an accompanying paper [SCF174] [2] *Capacity planning for HetNets*. In brief, however, planning starts with an overall understanding of where the network requires additional coverage and capacity. Depending on the location, the network may best be enhanced through either outdoor densification or indoor deployment. Figure 2-1 illustrates different views of network traffic differentiated into street level and individual buildings.

![Figure 2-1](image)

*Figure 2-1 Differentiating between outdoor and indoor payload forecast in 2D*

Although 2D is the norm for visualisation of UE location or carried payload, enhancement into 3D can reveal important insights which may help the operator to better identify effective small cell placement strategies. For instance, 2D mapping does not identify traffic variations on different floors of a building. This may lead to an expectation that multiple street-level small cells may be able to offload the majority of the indoor traffic. However, this is unlikely due to the physical nature of the overall site solution, including the antenna radiation pattern used, and could result in small cells being placed at street level, 5-10m from the ground with a directional antenna. In these deployments the small cell would not be able to serve traffic on the higher floors of the building.

A 2D-only view could therefore lead to an excessive number of street-level outdoor small cells. These would then need to be pruned back after the initial planning is completed, as part of a design review process.

The use of 3D location and visualisation, by contrast, provides an understanding of the distribution of traffic within the buildings, as shown in Figure 2-2. This helps operators decide whether indoor demand could be served using an outdoor-indoor approach, or whether an indoor deployment is needed.
Supplementary information can be summarised, such as the traffic per building, as shown in Figure 2–3. This would allow initial business cases to be created, or target solutions to small groups of floors, rather than the whole building.

Indoor and outdoor planning have different design considerations, which are detailed in the following sections.

2.2 Outdoor planning

2.2.1 Why does the entire system need to be examined holistically?

If an operator doesn’t build the best possible network – one that not only meets demand but maximises income and minimises capex and opex – it won’t get the best ROI. To ensure that it does, certain key elements have to be modelled precisely. These elements are environment (Figure 2–4), signal (Figure 2–5) and demand (Figure 2–6). This modelling process is a holistic one, taking data from a wide range of sources often gathered from a wide area.
Figure 2–4  Los Angeles DVHM (Digital Vector Height Model): submeter vertical and horizontal digital model of the environment, including elements that greatly impact the propagation of 5G signals, such as trees and buildings

Figure 2–5  Los Angeles Composite Demand Density Map. Demand is not uniformly distributed. Hotspots appear as redder areas, each building presents a unique demand profile and roads are also characterized by mobility patterns
This was not always the case. People once designed networks using teams – a coverage team, a capacity team, and a backhaul team, say. Each would work independently and feed some information to the other teams. The data input required was modest, as were the networks, which had a small number of base stations mainly used to provide voice communications.

Modern-day wireless networks are used not just for the tasks carried out by landline telephony but those offered by fixed broadband computing. They are influenced by a large number of independent factors and therefore have to be planned holistically, using multiple data inputs. Given the vast amount of inputs, the process needs to be automated – and it can be, using network design software tools.

The new normal

As 5G approaches, HetNets and densification are becoming the new normal for network rollout. A HetNet auto design process is therefore required to reduce planning time. The design software that enables this will, of necessity, learn and adapt in a machine learning process. But it needs something to learn and adapt from. It will need to include many inputs to inform the decisions that will be made – decisions on sector count, azimuth, fibre type and siting, street furniture, fronthaul, backhaul (wired or wireless) and dozens of other considerations that will be part of a design workflow.

But this only begins to illustrate the complexities that are faced when planning a modern communications network. Obviously since the advent of small cells and the introduction of a growing number of frequencies (with more – such as mmWave – to come), the challenges of multiple bands, spectrum and cell placement and interference also have to be taken into account. IoT, Wi-Fi, LoRa, WiGig and numerous other protocols, connectivity enablers and access technologies will soon be part of many operators’ capabilities or requirements.

The logic of HetNets is that they will be made up of many cells of many different sizes and strengths. Interference will therefore be a constant concern. Even if a macrocell is
many kilometres away, it might cause interference to small cells in the top floor of a building if it is positioned on a hill or a mountain offering no interruption to signal. At the same time buildings can block signals, affecting outdoor planning. This means that scale is also important: the area under consideration could be vast if the relevant data inputs are dispersed over a large area (Figure 2–7and Figure 2–8).

Figure 2–7  Manhattan RSRP showing macro and small cells (AWS, Tier 1 operator)

Figure 2–8  Close up of Manhattan RSRP showing macro and small cells (AWS, Tier 1 operator).

But a network is only as good as the efficiency of its service to end users. Thus a planning tool will need to answer many questions – questions like: Where is the
underserved demand? Where is the capacity? How efficient is network performance during busy hour? Can it be improved? What is the preferred cell radius? Where are the street poles needed for small cells and how should they be spaced? What is the signal to noise ratio (SNR)? How will SNR affect cell spectral efficiency – a critical part of modern network design? Can you extrapolate and predict future demand and densification roadmaps – and, if so, how far ahead?

There are multiple dimensions that have to be evaluated – and all of these different dimensions have to go into the planning process as early as possible to guarantee accuracy and shorten the time to deployment. At the same time, new data will need to be added over time, although this will of course help to reinforce the model and benefit planning.

**Inputs everywhere**

Luckily many data inputs are widely available that can offer useful information. In the US for example it is possible to connect to the FCC spectrum database to get a complete picture of US spectrum ownership.

But many other inputs, not directly cell-or signal-specific, can be used to assess user activity and movement in particular, which will influence cell siting. Most countries publish the results of their national census audit, offering residential and employment information along with growth estimates. Traffic estimates, GPS tracks and other data can help to establish vehicle density at different times of the day. Similarly GPS tracks, the mobility of social events, and a precise model of the environment can aid in calculating the mobility and indoor/outdoor status of user distribution.

Social media events can help too. For example, amenities such as shops and restaurants often involve the most concentrated amount of demand on a mobile network, and are the locations where the population distribution reaches its highest density. It is possible to estimate this via the quantity of check-ins found in social networks such as Facebook and foursquare.

Multiple data sources are useful but they are there to serve a purpose: assessing efficiency, coverage and capacity, which in turn need to be balanced against ROI. When an operator decides whether or not to place a site in a given position it has to balance the cost of putting in a cell and backhaul against demand and likely usage. What is the time to ROI? When does investment become profit? Network modelling is about revenue as well as efficiency.

All these factors – and many more – must be considered during the business case evaluation, since they both inform and define strategy. And that is why it is important to understand the holistic nature of network planning.

**2.2.2 How accurate does the computer model needed to mix powerful macros with less powerful picos and Wi-Fi APs etc have to be?**

The localised signal from a small cell inside a building can be overwhelmed by signals from more powerful external devices. It is, therefore, important for good indoor planning to ensure that the outdoor planning isn’t an obstacle. Not just proximity of macros to small cells but height will be part of this planning. Small cells at lower levels of a building are less likely to be affected by macro cells. But if a macro itself is at a great height (on a hill or mountain, say) its effect can be felt tens of kilometres away.
Of course, as small cells become part of the outdoor HetNet, interference in the outdoor network is also going to be more of a headache for planners. How then does planning software mitigate the threat of interference?

It will start off with the entire area being analysed. That could be a city of, say, 50 km across. Next it will create a computer model of the entire area under consideration: terrain, trees, roads, railroads, and buildings, for example. Buildings in particular are critical because they can block the signal. This modelling has to be carried out very precisely: the current state of the art is to model to an accuracy of one metre vertically and horizontally (Figure 2–9).

**Figure 2–9** 3D view of Brooklyn DHVM (Digital Vector Height Model): submeter vertical and horizontal digital model of the environment, including elements that greatly impact the propagation of 5G signals, such as trees and buildings

Network planning software should provide operators with different options for small cell placement or changes to the macro network. For example, they could downpower a macro, say, or reorient it – or reduce the signal in a building. For in-building coverage, the aim of planning software is to allow you to use a full indoor design – but it can also help you to minimise cost. Can you, for example, really afford small cells on each floor and then fibre to each one of them? Could there in fact be minimal use on specific floors, requiring minimal rollout?

You might then tell the planning tool just to optimise the lower floors and make sure there’s good coverage from a macro on the lower floors and less interference on the upper floors. A balance can be struck (Figure 2–10).
Remember too that small cell planning indoor and outdoor increases the need for accuracy of signal calculation since errors injected into the signal calculations intensify when computing the signal-to-interference ratio. If errors are present in the computer model (such as using very coarse clutter and terrain) this could result in an inability to predict the signal. And this brings us back to ROI: failure to predict the real return from a new deployment destroys the value of the planning process.

AI and machine learning will, necessarily, play a part in cell placement. Metaheuristic algorithms, of the sort used in today’s planning tools, aim to minimise the difference between input measurement data and predicted data. As happens with machine learning algorithms, input from previous analyses will help the planning tool decide what its next move is. For example was a site rotated in the past? What was learned? Can that experience influence the next situation where site rotation might be an option?

2.2.3 User activity is hyper-sensitive to network performance. How? Why?

Correct cell placement is not just about filling in coverage holes. In areas of very high demand operators will, of course, need more cells – but the signals from some cells may interfere with the signals from others, and this will influence the throughput users can get. Therefore an accurate estimation of interference is needed. That requires an accurate model of how a signal propagates – and, as noted before, that can only be achieved with a precise model of the environment.

In particular, the signal to noise ratio must also be considered. Strong signal to noise ratios due to bad placement, interference or other reasons will not just upset end users. They will require a robust modulation scheme and affect cell spectral efficiency. Bad cell performance will, in turn, mean less capacity and higher costs. Consumer distribution greatly affects cell spectral efficiency. A planning tool that mitigates this
will focus on accuracy in signal prediction and consumer distribution as an important part of the planning process.

The accuracy of the cell spectral efficiency calculation is directly relevant since it allows the network planning tool to choose the optimal location for new RF deployments. It is also important therefore to have an accurate model of the user distribution.

2.2.4 How are outdoor zones of underserved demand identified?

Where are the users? Operators can supply estimates of where they are and how much demand they are generating of course, but many other supporting sources are available and useful – social networks for example. Social events from Twitter and Instagram correlate with actual footfall gathered electronically. By far the most useful data for estimating population distribution in time are the social networks that allow a public stream that includes the location and time of an event.

Combining this with a vector land classification (clutter) model means that when events occur, they can be classified with greater accuracy. For example, did a tweet occur within a building or a park, or on a ferry?

Twitter features a high volume of such events. Similarly, Instagram and Flickr include the location of photo uploads and can include EXIF information that allows estimation of the height of an event.

The more accurate the model the more granular the information. The height of a building, together with demand on a building-by-building basis, can allow many different analytics to be performed on this information. Entire cityscapes can be analysed and the buildings that require the most capacity isolated.

Add the information mentioned earlier – censuses, jobs, traffic, terrain, social media, and even crime data – and it is possible to build a statistically significant indication of the population distribution – and thus, combined with operator inputs, actual and predicted demand. Planning software could even provide a growth model to be used for a densification roadmap, although the further forward this goes the greater the margin of error. Nevertheless, a view three years ahead is certainly possible, and with it a good estimate of when investment becomes profit.

2.2.5 What is the effect of optimal network designs that simultaneously minimise cost while maximising capacity?

This is a simple question to answer, but making optimal network design actually happen is difficult. The goal is to place sites with high spectral efficiency in high-demand areas. However, the set of things you, the operator, can choose from is massive. You can choose bands. You can choose technologies: Wi-FI, say, may be an option in places. Do you use wired or wireless backhaul? Are street poles or even (very cheap) strand-mounting available? Is the spacing of cells precisely modelled – not just to meet demand but to keep costs down?

Everything is associated with the cost. An operator is trying to maximise spectral efficiency to get the most capacity but also to maximise the return on investment by minimising capex and opex.

If you get all that right, then optimal network designs that simultaneously minimise cost while maximising capacity will, in theory, positively affect equipment purchase,
equipment efficiency, medium and even long-term-planning – and, above all, customer retention, now and over time.

2.2.6 How do spectrum and multiple bands play a role in controlling capacity and coverage?

Operators can already choose from a number of bands and often an operator’s needs will dictate their role. In some areas an operator may be better off using coverage spectrum like 700MHz and not considering small cells. In other places, where small cells are needed for capacity reasons, different spectrum will be necessary. There may also be a trade-off. Eventually mmWave rollout will give operators greater bandwidth but will make other demands – stricter line of sight requirements for example.

The quantity of spectrum also matters. If the quantity of spectrum is 20MHz and the cell spectral efficiency is one bit per second per hertz that means 20 megabits per second. That’s the total capacity of a well-performing site or sector on a site. Multiply that by the number of cells and you get the true capacity of the network.

But even with adequate spectrum, spectral efficiency and interference matter, especially when considering ROI. Thus spectrum and multiple bands are useful – but only if their use is planned correctly.

2.2.7 How do you plan for multi-RAT: mixing different technologies for different objectives (e.g. Wi-Fi, IoT)?

Everything must be taken into account in the planning process. The whole air interface is a mixture of different frequencies from lots of different devices, some of them interfering with one another, some of them isolated and not interfering. But it all has to be taken into account – because when an operator rolls out its business it needs to look at every single aspect of it. Again, this is a holistic process.

For example, planning for different objectives can be guided by demand models. High mobility, low mobility, indoor demand, outdoor demand, type of demand (voice, data or both), big demand in, say, stadiums or parks, how they are covered and the cost – all these considerations can inform technology deployment and whether an operator uses, for example, Wi-Fi, LTE, 5G, 700MHz, 1900MHz, 2500 MHz or 28GHz. If an operator is building a modern network – as EE (UK) or Verizon (USA) are – it probably has to consider all bands. All of this has to be predicted within a planning tool – another aspect of the holistic examination required for accurate network planning.

Relatively new technologies like IoT also have to be considered. IoT today is thought of as small amounts of information from smart metres, temperature gauges or sensors on factory equipment, but how will IoT evolve in use and network demand? Will information be aggregated through gateways? What does this mean for network design?

And then there’s backhaul. If you’re ten metres from the fibre it’s relatively inexpensive. If not, what are your options? Satellite or microwave backhaul are possible, but with 5GNR operators will be able to backhaul wirelessly using WiGIG – Wi-Fi at 60GHz. It’s unlicensed and therefore quite cheap. How should an operator plan for that?

Then there’s mmWave – game-changing for fixed wireless and eventually other forms of wireless: integrated access and backhaul and therefore more bandwidth. The
downside is that it is a short-range line-of-sight-only technology that can be affected not just by objects but even by rain.

Massive MIMO will offer its own challenges. But all these technologies and frequencies have to be part of the modelling process.

Finally, of course, the technology mix will also be affected by new or growing user profiles – travellers by road, rail and air, say. And what will be the requirements of autonomous vehicles?

2.2.8 How do design and rollout strategies affect ROI?

Once network design software has given an operator all the network modelling information needed for maximum capacity – and therefore maximum ROI for its opex – the operator can choose from a number of business plans knowing what (and when) the ROI is likely to be with only a small possible margin of error. An operator can do this with the confidence that, assuming the network design software has done its job, the model is as deployable as it can be because most eventualities that it is possible to predict have been included in the modelling. The operator knows that, even if one business model differs slightly from another, it’s going to get a return on this particular network design.

2.2.9 Why do 5G, and in particular mmWave, change traditional planning?

5G will bring a number of planning advantages – and challenges. Most notably it will mean more careful planning – not just because of densification, but because where everything is on the same band and the small cell is on the same frequency as other cells, small cells are more likely to experience interference from the macro.

Although it is a part of 5G, mmWave is important in its own right because it gives operators bandwidth – and when it’s all automated in 3GPP release 16 that means integrated access and backhaul. That integrated access and backhaul is what changes everything. mmWave allows operators to create antenna arrays that are the size of playing cards that can go into a box the size of a paperback book at what is considered a very low cost and, potentially, access to huge quantities of bandwidth and 64 users at a given time. The main short-term application for mmWave is fixed wireless access but that can, and probably will, expand into mobile applications (Figure 2–11).
Most importantly of all perhaps, 5G requires not just optimised HetNets but densification roadmaps, with all the added complexity that implies for cell placement.

2.2.10 Summary of outdoor planning

The first mobile network in London offered had about 18 sites. We can expect city-wide networks of 2,000 or more sites with multiple bands in the not-too-distant future, and that number will expand. 5G, in particular, will have an astonishing effect on demand and capacity but will make ever-more-complex demands on planning. But planning too will evolve. Planning tools will be soon able to make use of data derived from drones, context-aware equipment and self-deploying solutions, to name only three.

Whatever the future of planning brings, however, the underlying opportunity will be the same: place your cells where the users are – and avoid interference – and you maximise return on investment.

2.3 Indoor planning

Indoor networks are deployed in a variety of venues that have very few things in common. Public venues, such as large stadiums, convention centres, airports and subway stations, have favourable, mostly line-of-sight propagation conditions between access points and UEs. Enterprise venues – such as hospitals, hotels or corporate headquarters – will put many obstacles in the propagation path, resulting in mostly non-line-of-sight (NLOS) propagation, with reflection and diffraction off nearby surfaces. Some venues, like shopping malls, have a roughly equal amount of LOS and NLOS.

The amount of LOS vs NLOS is important if an indoor network is designed for coverage. A design for coverage specifies a certain signal level over a certain
percentage of a floorplan. For example, a typical coverage design would require an LTE RSRP (reference signal received power) signal level of -95 dBm over 95% of the floorplan for a 10 MHz LTE channel. The exact target signal level is a function of transmit reference signal power and the channel bandwidth. The percentage of the floorplan that needs to be covered is usually between 90 and 95%; 100% coverage is unrealistic because it would require too many access points. For coverage then, venues with mostly NLOS propagation need more access points per unit area than venues with mostly LOS. For example, a hospital floorplan would require more antennas than a garage floorplan.

When it comes to designing an indoor network for capacity, it is important to know the number of UEs and types of traffic they will demand. There are two limits at play here: a hardware (Layer 2) limit and a physical layer (Layer 1) limit.

In LTE, a hardware (Layer 2) limit for small cells may be up to 32 simultaneous connections (as an example). However, with the advent of C-RAN and O-RAN distributed networks, the baseband units, which are capacity sources, are separated from the amplifier and can be shared with many small cells, thus diminishing the importance of a hardware limit per individual small cell. On the other hand, physical layer limitation per cell is still important. Layer 1 capacity calculation implies calculating the amount of time needed to transmit file(s) to and from UEs during the peak, or busy hour, usage and comparing it with overall time during the busy hour. If too many UEs are connected and they all request large file transfers, the needed time to complete the transmission may be too long for an acceptable user experience. In areas where we expect heavy data use (such as 4K video files or gaming), more small cells are needed than in areas where light data use (such as wireless printing or file exchange) is expected.

A well-planned indoor network designed for capacity should identify areas where users tend to congregate (for example a lunch cafeteria at the corporate office, hotel bars, nightclubs or convention conference rooms), the type of traffic expected (streaming video download, Facebook video upload, email exchange or Twitter posting, say) and approximate file size download/upload per user during the peak hour. An indoor network design also needs a breakdown of the percentage of anticipated traffic between the existing Wi-Fi at the venue and the cellular traffic. The cellular traffic should also be broken down into MNOs according to their user penetration rate. In addition, within each MNO, the traffic needs to be broken down between technologies (3G, 4G and 5G).

When all this is addressed, we can finally place small cells on the floorplan. Care should be taken to position them near or at areas of heavy traffic, but also near or at VIP areas, such as C-level suites or penthouse apartments. We should also remember to cover other areas where traffic is not heavy in order to provide continuous coverage throughout the venue. Finally, capacity analysis should calculate the percentage of airtime usage for each small cell/node. The airtime usage can also be interpreted as ‘cell load’. A cell load of 80% means that, on average during peak use, a small cell is busy transmitting and receiving 80% of the time. A well-designed system should be designed for 50-80% loading, which represents good utilisation and also leaves room for growth before an upgrade is needed.

An example of a capacity coverage map is shown in the figure below. Here we assumed the maximum number of connections per node to be 32, and the target cell load to be 80%. Areas in dark and light red indicate nodes that experience hardware failure (more than 32 connections), and physical layer failure (cell load greater than 80%). Nodes that passed are coloured green.
Another key performance indicator is the maximum achievable data rate (MADR) coverage. The MADR is a function of signal to noise and interference ratio (SINR), which calculates the ratio of the serving cell signal to interference plus noise coming from non-serving small cells. In LTE networks, which mostly deploy omnidirectional and wide beam (60-90 degree) directional antennas, interference coming from non-serving small cells can be significant, which degrades SINR near the cell edge. Consequently, MADR is also low at the cell edge. With the advent of beamforming antennas, which will be widely deployed in 5G, the interference will be much reduced, because narrow beams mean that the target area is narrow, which reduces the probability that the serving and the interfering signal may arrive at the UE at the same time.

Here is an example that illustrates this point. Let’s assume that we have an 8x8x2 small cell panel antenna that creates individual 16° narrow beams. To cover a 120° sector, seven narrow beams are needed:
These seven beams cover a 120° horizontal sector, but only a 16° vertical sector. If we want to cover a ±40° vertical sector, then we need five rows of beams, which means that we need 7×5=35 beams. The composite panel coverage with 35 beams is shown below:

If we assume that the antenna array is analogue – which is a fair assumption given the current state of the art in millimetre wave small cells – then only one beam is transmitted at a time. If only one beam can be detected as interference at UE (again a fair assumption given that the beams are narrow), then interference can be detected only 1/35\textsuperscript{th} of the time. This leads us to the conclusion that interference is reduced by a factor of 10\log(1/35) = -15.4 \text{ dB} with respect to a full-sector 120-degree antenna. We can therefore see a significant interference reduction with beamforming.

Two more KPIs – signal strength and MIMO rank – should also be mentioned. For coverage planning, the main KPI is RSRP for LTE and SS-RSRP (synchronization signal reference signal received power) for 5GNRs. A MIMO rank indicator shows MIMO performance throughout the floorplan. For 2x2 MIMO, MIMO rank can be either 1 or 2. For 4x4 MIMO, the MIMO rank range is 1 through 4. The higher the rank, the better the MIMO throughput.
Lastly, a 5G network should be designed with backhaul throughput and delay in mind. A backhaul throughput is a function of capacity and predicted cell load. It tells us how much throughput is needed to provide connectivity to the outside network. An indoor network delay is delay caused by active elements and fibre and should be calculated from small cell to the main closet/MDF. This is important for edge computing, as many 5G applications are sensitive to delay.

To summarize, the important KPIs for indoor network design are

- Capacity and coverage,
- SS-RSRP
- MADR
- SINR
- MIMO rank
- Backhaul throughput requirements
- Indoor network delay

When speaking of millimetre wave indoor propagation, high-frequency signal range is rather poor, even in LOS conditions, while most indoor materials create a large penetration loss at that frequency band. Thus, indoor millimetre wave 5G applications will be limited to opportunistic deployment, rather than a building-wide coverage. Typical use cases for indoor millimetre wave would be hotspots, coffee shops, stadiums, convention centres and ballrooms.

Another factor to consider is the impact of the indoor network on the outdoor network. In most cases, this coverage creep should be kept to a minimum. Many operators specify the maximum allowed indoor signal level at a certain distance (usually 100 feet) from a building. This limit is not difficult to achieve in newer buildings with solar glass, which has a high penetration loss for both sunlight and radio waves. However, buildings with older types of glass, which have far less penetration loss, require careful network design. A common tactic is to use directional antennas near windows, pointing inside. Directional antennas with a high front-to-back ratio should be used, to further decrease the amount of radiation through the back lobes.

However, it should be pointed out that some indoor networks require expansion of the network outside. Specifically, stadiums experience heavy traffic before and after an event – at streets leading to a stadium, bus stops, train stations, and parking lots for example. While in most countries this traffic builds up and subsides 30-45 minutes before the event, in North America the build-up of traffic can last up to two to three hours before a sports event, notably when people congregate at a parking lot and stay there for a while (so-called 'tailgating'). Therefore, it is a common practice in that part of the world to extend indoor networks to parking lots surrounding large stadiums, to address the capacity needs before the game.
2.4 Accuracy of network quality data sources for planning

The accuracy of data sources available for planning is illustrated in Figure 2–15. Each of the data types has its own benefits and drawbacks; hence a pragmatic approach is required to realise the most appropriate solution. For instance, there is a wealth of cell statistics, but it cannot be resolved below the serving area of the cell. With minimisation of drive test data (MDT), the GPS location is reported for a subset of Android phones including an accuracy flag if the phone considers it to be a ‘stable location’. 2D and 3D geolocation using network probes has a relatively large variance of accuracy. Enhancing the accuracy of this geolocation is the subject of section 4 of this paper. Social media events are often tagged with GPS location. This is fine for rural or suburban networks but does start to ‘bounce around’ in street canyons of dense urban areas.
3. Impact of siting accuracy on ROI

Error! Reference source not found. Figure 3–1 shows the ROI (return on investment) of several different cost profiles of small cell deployment (plotted on the Y axis) against the distance from the ideal location (on the X axis), as given by a planning tool. There are 1500 possible candidate sites and 160 ideal locations chosen to meet the capacity and coverage criteria in this study. The ROI is based on the principle of income generated by carrying a GB of data – a value which is publicly estimated in many sources – and the cost of deployment of various site solutions. The technology deployed in this example is an LTE 5+5w small cell with a relatively small bandwidth.

![ROI vs Distance](image)

Where: ROI = profit / TCO over a five-year timeframe

**Figure 3–1  Study of the effect of ROI with distance from the ideal [Source: Nokia]**

The ROI is positive for all site solutions as long as the best location is secured during the build program. If the preferred site is not available, then there is some leeway before the ROI tends towards zero, but this depends on the cost profile of the site.

However, to maximise the true value of the small cell the accuracy should be within 20-40m of the ideal. This is because there many small cells where the roll-off in ROI is much steeper than the average shown here and a well-constructed planning process would want to capture all possible traffic, not just the average case. Another consideration is that, since this study was completed, the traffic-carrying capability of small cells has increased significantly. This will help to improve ROI, even with a falling retail price of a GByte of traffic. However, in all cases the extra cost of diligent location planning easily pays back.
4. Enhancing the accuracy of UE location with machine learning

Given the importance of obtaining accurate location for small cell planning, we have to understand the accuracy of the input data. While some mobiles (e.g. Android) may report accurate location in suburban or rural areas using assisted GPS (A-GPS), this method does not work well in urban areas. Other timing-based techniques like standard observed time difference of arrival (OTDOA) are also, typically, not accurate enough. It is, however, possible to apply a machine learning algorithm approach to improve the accuracy of OTDOA and other timing-based methods. The results are very good, achieving typical median location accuracies of about 18-25m in suburban/urban areas.

4.1 Types of UE locate

There have been several network-based approaches to location over the years, using signal levels or timing triangulation between cell sites [e.g. Gabe] that could obtain around 50m accuracy in dense urban areas. However, these approaches are more likely to have location accuracy of 200m+ in suburban areas and are not sufficient for most small cell placement strategies in these areas. In suburban and rural areas handset-based location solutions using A-GPS work well, with accuracies of around 10-15m. However, neither A-GPS nor other solutions have reliable location accuracies of much less than 50m in urban areas, since A-GPS from mobile devices is not able to ‘see’ enough satellites to get an accurate location.

Figure 4–1 Example of GPS accuracy degradation in downtown areas (in this case Dallas, Tx). Green = over 85% at 40m or better, Yellow / Orange = 50 to 85% < 40m, Red = less than 50% < 40m

One method to potentially overcome the lack of A-GPS in urban areas is to rely on handset-measured ‘timing measurements’ between cell sites – for example with
OTDOA measurements reported by the mobile stations on request. OTDOA is a standardized solution supported by nearly all mobile handsets today for emergency-class location requests (e.g. 911).

4.2 Using machine learning to create ‘enhanced’ timing-based solutions for improved location accuracy

Standard OTDOA measurements reported have marginal accuracy (more than 60-70m median accuracy) in multiple environments (urban, suburban, rural). However, we can show that, with machine learning, accuracy can be significantly improved, yielding median accuracies about 18-25m across urban, suburban and rural areas.

4.3 Using ML to improve timing-based location techniques

In order to determine the potential accuracy of any solution, measurements from a large number of handsets must be analysed. For the study used here, the Dallas metro area was chosen, and several thousand ‘sets’ of OTDOA-derived location measurements were collected, where a ‘set’ consists of OTDOA location measurements from multiple pairs of cell IDs.

![Map of Dallas metro area](image)

**Figure 4–2 Dallas metro collection area**

In addition, for each set of OTDOA location measurements from cell ID pairs, there is a measured GPS measurement as an assumed known location, as shown by the solid green circle in Figure 4–3.
Figure 4–3  One ‘set’ of locations derived from OTDOA pairs (blue circles), and the assumed known location (green circle).

As can be seen, there are a number of ‘location options’ (blue circles), one for each cell pair. The goal is to find the best estimated location as close to the known location (green circle), from the original OTDOA measurements (which produced these locate estimates), and/or by using the original OTDOA measurements and the location options. One would assume that the more location estimates from OTDOA pairs (blue circles), the better one’s estimate of the true location. The graph below illustrates how often fewer than ‘N’ cell pairs (or location estimates) were seen.
4.4 Calculating the relative timing offset between cell sites for use in location calculations and effective inter-site synchronization

Before attempting to use OTDOA-reported measurements from cell site pairs for location estimation, one must first take into account the relative timing offset between cell site pairs. A model can be created which shows – if one knows the location of a mobile (from A-GPS, say) – how one can calculate the cell site pair timing offset. This offset can then be used in subsequent OTDOA location estimations for devices where no A-GPS is known.
For any given instant that we collect OTDOA measurements, after, or in conjunction with, applying the inter-site timing offsets as described in the previous section, we can apply a variety of machine learning algorithms to estimate the actual location. The most simplistic algorithm may be to just ‘average’ all the individual location estimates from all of the cell pairs reporting at a given time. More sophisticated methods may work on both the OTDOA measurements themselves and the potential location set.

The figure below illustrates location accuracies over the entire test area in Dallas using simple averaging (green line), and AT&T’s patent-pending machine learning approach (red line), as compared to a ‘genie location estimate’ (black line), which picks the single closest location estimate from a given cell pair to the known location. In all cases, there is no change to the mobile stations and all computations are done in the cloud.
In this case, instances were chosen where at least 10 cell pairs were measured from the handset. Notice that the median (50th percentile) accuracy of the ML algorithm (red curve) is approximately 18m, with an 80th, and 90th percentile around 45m and 80m respectively. Additionally, the experience in comparing the accuracy of the actual reported OTDOA location from the standard OTDOA system was actually a bit worse than the ‘simple averaging’ solution shown in the figure above. Consequently, the accuracy of the ML approach at least improves on ‘simple averaging’.

Taking these results into account and looking again at the downtown Dallas area, but with enhanced OTDOA, gives us the outcome shown below.
Figure 4–7  Example of using the improved OTDOA location accuracy in downtown areas (in this case Dallas, Tx). Green = over 85% at 40m or better, Yellow / Orange = 50 to 85% < 40m, Red = less than 50% < 40m.

As can be seen, the location accuracy improvement (Fig. 8 vs. Fig. 1) is dramatic, as it shows good location accuracy in downtown urban areas now, where A-GPS struggled.
5. Improving small cell placement using machine learning

Given the factors of cost and technology as inputs to small cell placement, planning and execution of networks becomes complex. Manual planning of build-out plans can become cumbersome, even to the most knowledgeable and experienced engineers. Machine learning provides the opportunity to compare multiple scenarios with different factors to maximize coverage and minimize cost. AT&T has been piloting an effort to use ML for small cell placement.

5.1 Factors included in ML model

Our models start with the area to be covered (a geographic polygon), the traffic-weighted location data (ML enhanced per previous section), and the budgetary constraints associated with that area. Data is added to the model, including terrain and structures in the area, and the associated clutter classes and heights. Assumptions are made regarding the antennas (omni-directional, brand, etc.) Candidates are generated and classified and factored in the following priority order:

1. Existing LTE assets
2. Locations where we have agreement with site owners
3. Intersections identified by open street maps
4. Uniform hexagonal grid

The model uses the goals of coverage and dominance. An outdoor location is considered covered if the strongest signal at that location has an RSRP greater than -112 dBm. A location has satisfactory dominance if there is at least 5 dB of separation between the strongest and third strongest signal at that location. The figure below shows the factors included in our model.

![Factors included in ML generation of Sites](image)

Figure 5–1 Factors included in ML generation of Sites

5.2 Cost and coverage improvement using ML

With the model established and proper data identified and loaded, build-out plans can be generated. For this example, we used a section of Manhattan. The manual design identified 185 sites in the polygon that would need to be placed. The ML generated design created 111 sites. The figure below shows the differences between the two sites.
The automated design is able to provide coverage and dominance while reducing the number of sites required. The reduction of 74 sites provides a savings of 40% - but the model has also created optimized coverage for this lower cost. The following figures demonstrate the maintenance of coverage and improvement of dominance in the manual vs. automated plans.

**Figure 5–2  Manhattan Manual vs. Automated Designs**

**Figure 5–3  Coverage (RSRP) in Manhattan in Manual vs. Automated Designs**
Small cell design is an emerging application of ML – this example is one of the first design uses within AT&T. There are significant benefits to using this technology, both through reduced costs and improved customer experience. These models will continue to evolve as our assumptions, data, and other inputs continue to improve.

Figure 5–4  Dominance in Manhattan in Manual vs. Automated Design

83.8% dominance  94.7% dominance
6. Acknowledgements

This work is agreed by and represents the view of all members of the Small Cell Forum, but we would particularly like to thank the following companies that contributed content for this paper: AT&T, iBwave, Nokia and Keima.
References

1. [SCF195] ‘Small cell siting challenges and recommendations’, Small Cell Forum
2. [SCF174] ‘Capacity planning for HetNets’, Small Cell Forum