Abstract—Stochastic models of base station infrastructure deployment by multiple mobile operators can be an invaluable tool for deriving fundamental results about wireless network sharing. In this paper, we study stochastic geometry models for a shared cellular network consisting of base stations deployed by multiple mobile operators, based on real cellular network data coming from three European countries. Relying on a statistical approach as well as the evaluation of wireless network performance metrics, we find that, in the city center areas, operators tend to deploy their antennas in close proximity, while in other areas this is not the case and we find some level of repulsion between antennas of different operators. As we show, the log-Gaussian Cox process provides the most compelling fitness results with real multi-operator base station deployment patterns and a potentially attractive model that offers some degree of analytical tractability. Moreover, we observe that the behaviour which can be modelled with the help of these processes occurs over and over again for similar areas in different countries, which suggests universality of the proposed models.

Keywords—mobile networks, network sharing, infrastructure sharing, stochastic geometry, log-Gaussian Cox process.

1 INTRODUCTION

Modelling and simulating the performance of shared wireless networks is of high importance to the research community and even more so to the cellular networks industry. From a technical perspective, sharing has been ongoing for a while now, predominantly in the form of sharing masts and antenna sites between mobile operators, which is often referred to as passive sharing [1]. More recently, we have been witnessing an increase in active forms of sharing, which involve some level of resource virtualization that allows individual operators to jointly utilize base stations and/or the core network [2]. We see examples of this approach appearing more and more in response to the large costs associated with the provisioning of wireless cellular infrastructure. Some commercial examples of this type of sharing include the merger of T-Mobile and Orange networks in Poland\(^1\), the joint venture of two mobile operators in Denmark\(^2\) or the recent deal between Eircom and Three in Ireland\(^3\).

The fundamental gains and limitations of such shared networks are still an open question, and the research to date has been rather conservative in treating the topic. To quantify these gains and limitations, there is the need for accurate models that are representative of the superposition of radio access network deployments made by multiple mobile operators. Herein, we make a contribution to the state-of-the-art in wireless network modelling by providing a quantitative analysis of stochastic models that may be used to represent shared cellular network deployments.

In general, wireless network deployment can be modelled as a collection of points distributed on a two-dimensional plane. Each point, representing a base station (BS) transceiver, is assigned some wireless network related properties, such as downlink transmit power and operational frequency. Then, the average performance experienced by a user of such a network may be obtained either using stochastic geometry or through computer simulations. In stochastic geometry a typical working assumption is to use a homogeneous Poisson point process.

\(^1\) The process is comprehensively described in the following slides: http://www.slideshare.net/zahidtg/deployment-of-ran-sharing-in-poland (accessed 22.10.2014).

\(^2\) http://www.fiercewireless.com/europe/story/teliasonera-and-telenor-given-ok-network-sharing-jv/2012-03-02

\(^3\) http://www.fiercewireless.com/europe/story/eircom-agrees-long-term-network-sharing-deal-3-ireland/2014-08-29
In this paper, we identify point processes that model a cellular network consisting of base stations belonging to multiple operators (a multi-network), operating in urban environments. We base our work on BS deployment data coming from Ireland, Poland and the UK. What we observe is that multi-networks cannot be simply modelled with a superposition of independent realizations of a point process describing a single-operator deployment (a single-network). In fact, as we show later, there is a high degree of dependence between the infrastructure deployment made by multiple operators. This quantitative result supports intuition which suggests that commercial networks are designed and deployed to satisfy the demand unevenly spread across geographic areas by competing operators offering similar services.

In order to find the model that best fits multi-networks, we apply a spatial statistics approach [6]. Namely, we calculate the empirical estimates of the pair-correlation function, which is a second-order characteristic, and we fit the point processes to those estimates using the methods of maximum pseudolikelihood, whenever possible, and minimum contrast [7]. Subsequently, we evaluate goodness-of-fit of our estimated models by comparing empirical estimates of some other summary statistics obtained from the real data against maximum and minimum values of the same statistics calculated for our estimated models. In the course of our analysis we find that the clustered point processes best represent base station clustering occurring in multi-networks, with the log-Gaussian Cox process (LGCP) providing the most compelling fitness results and a potentially attractive option for some degree of analytical tractability. Moreover, we observe that this clustering behaviour, which can be modelled with the help of clustered point processes, occurs over and over again for similar areas in different countries, which indicates some degree of universality of the proposed models.

## 2 Related work

A new wireless network roll-out involves three phases [8]: dimensioning, planning, and optimization. In the first phase, a rough estimate of the network layout and elements is prepared based on technical demands (such as an area to provide coverage and capacity for), forecasted subscriber base and service usage characteristics [9]. Next, network element configuration is performed, whereby base station site equipment is selected based on power budget calculations, followed by detailed topology planning, which includes coverage and capacity planning based on pathloss information and prediction models (such as the modified Hata model), and dynamic models provided by commercial network planning and optimization tools [10]. In addition, for 3G and 4G systems, mobile traffic information is taken into account as it affects capacity requirements and acceptable interference levels. At this stage the cellular network is designed and ready to be deployed. During the deployment the positions of base station sites, which have been established according to the pre-specified models, become distorted by a number of non-technical aspects, such as [10]: site acquisition (availability of a given site), site legalization (license for radio emission from a given site), and site preparation (installation and availability of mast, power supply, air conditioning system, etc.).
Fig. 1: Illustration of different types of two-operator BS deployment patterns for the city of Dublin (the Dublin city area is represented here as a superposition of the smallest unit areas provided in [5], and located within a radius of 5 km from the city’s center), generated from: a) a hexagonal lattice, b) a Poisson point process, c) a clustered point process, and d) real data. The contours represent the Dublin city area, whereas the dots and triangles stand for BS site locations of two different operators. The real BS pattern is clustered in the central area, with irregular empty spaces forced by environmental obstacles and, most likely, coverage planning. Qualitatively, the clustered point process realization exhibits the closest match to the real BS pattern, with a number of smaller clusters distributed throughout the area of interest and relatively many empty spaces, while the Poisson point process pattern rather uniformly fills the Dublin area and the hexagonal pattern exhibits high regularity with little resemblance to the actual BS deployment.

Moreover, the incumbent mobile network operators, as a natural cost minimization strategy, will tend to reuse existing sites [8]. Effectively, even initial coverage provisioning deployment may be allocated in a way that is far from idealistic lattice-like scenarios. Furthermore, as the service demand grows, and large capacities become essential, a microcellular topology is deployed, typically characterized by much lower transmit powers and rooftop level antennas, with additional pico-/small-cell systems to extend the capacity to indoor offices or places like shopping malls.

Despite these practical insights on the cellular network roll-out process, there seems to be no consensus on how to model and analyse the resulting BS deployments. In the studies to date the choice of a model is typically dependent on the type of tool applied, either stochastic geometry or computer simulation.

Stochastic geometry allows us to derive analytical expressions for various performance parameters, such as outage probability or data rate, under predefined wireless channel conditions, averaged over many realizations of a pre-specified stochastic model. For analytical tractability, much of the existing work modelling radio access network deployments (e.g., [11], [12], [13], [14]) relies on the homogeneous PPP model, which assumes no correlation in the analysed patterns. The usage of the PPP model can be analytically justified under sufficiently high log-normal shadowing, see [15]. However, spatial statistical analysis of the available mobile network deployment data to date suggests that single-operator mobile networks can be better represented using regular processes. For example, in [3] the fitted Strauss process turns out to provide the closest match to a single-operator mobile network deployment. In [16] the determinantal point process (DPP) is found to accurately represent regularities present in single-operator base station deployments. Analytical insights into these point processes are limited to some specific cases, see [17]. Yet, heuristic approaches, such as the one in [18], may be applied. Clustering in real mobile networks, as observed in large-scale single-operator deployments [19], can be modelled using cluster processes. However, analysis of these processes seems to be even harder and some level of analytical tractability has been achieved only for certain classes of cluster processes, see [20].

Computer simulation-based approach typically relies on scenarios and models defined and published by standardization bodies, such as the third generation partnership project (3GPP), as a part of technical reports, for example, [21]. In those scenarios macro- and microcell base station deployments are modelled using lattices, most frequently adopting a hexagonal lattice, which creates repetitive structures to simplify inter-cell interference management [8], and “ensure optimal coverage over a given area” [10]. Even if initial deployments were performed according to a lattice-based model, the evidence to date shows that the existing networks are far from these idealistic models. Hence, to account for real-world irregularities a constant deviation may be added to a lattice-based model [10]. However, these perturbed lattices, as they are called, provide less satisfactory results than some other stochastic models (see
Moreover, we denote \( \Lambda(\Phi) \) as the number of points in the realization \( \Phi \) in a set \( R^2 \). Using the random counting measure formalism [26], we also denote \( \varphi \) as a realization of \( \Phi \), and \( \varphi(W) \) as the number of points in the realization \( \varphi \) distributed over a finite observation window \( W \subset R^2 \). Moreover, we denote \( \Lambda(W) \) as the intensity measure of \( \Phi \), which, informally, is the mean number of points of a point process \( \Phi \) in a set \( W \). In consequence, \( \lambda(x) \) will denote the intensity function that describes the point density around location \( x \in R^2 \).

There are three general classes of point processes: uniform, clustered and regular. To say that a point process is uniform is equivalent to assuming that no inter-point interactions can be observed in a point pattern coming from that process. Clustering means that some form of positive interaction (attraction) exists between points, leading to clustered patterns, while regularity stands for some form of negative interaction (repulsion) between points in the process. Since multi-network patterns exhibit, at least qualitatively, some form of clustering we focus our attention on point processes that allow us to model this phenomenon. Namely, in our work we rely on three point processes: the log-Gaussian Cox process (LGCP), the Matern cluster process (MCP), and the Thomas process (TP). However, in order to get the full picture of the kinds of interactions one may expect in such BS deployments we also test whether the observed point patterns follow either the homogeneous PPP or possibly a superposition of a number of independent realizations of the Strauss process (SP), which was shown to provide a good fit to single-networks [3]. While Guo and Haenggi in [3] have already introduced the PPP and the SP, in the following we briefly introduce the Cox processes and the cluster processes. More comprehensive description of all these point processes can be found in [6], [7], [26].

### 3.1 Cox processes

Cox processes describe spatial point patterns which exhibit some forms of aggregation or clustering, typically caused by some random environmental phenomenon [7]. The Cox process is obtained as a generalization of the inhomogeneous PPP, where the intensity measure of the point process becomes random. Conversely, conditioned on the intensity measure the resulting point process is Poisson, and thus the Cox process is also referred to as the doubly stochastic Poisson process.

**Log-Gaussian Cox process** We focus on a subclass of Cox processes, where the logarithm of the intensity function is a Gaussian process [27]. This subclass is called the log-Gaussian Cox process (LGCP), with the intensity function denoted as \( \lambda(x) = \exp(Y(x)) \), where \( Y = \{Y(x) : x \in R^2\} \) is a real-valued Gaussian process with mean \( \mu \) and covariance function \( \sigma^2 f(r) \), where \( \sigma^2 \) is the variance and \( f(r) \) is some predefined spatial correlation function. For the LGCP to be well-defined the spatial correlation function has to be positive semi-definite [28]. This condition is met by, for example, the exponential function \( f(r) = \exp(-r/\beta) \) and the stable function \( f(r) = \exp(-\sqrt{r/\beta}) \), where \( \beta > 0 \) is a scale parameter (other examples of functions that meet this condition can be found in [28]). The LGCP is fully characterized by \( \mu \), \( \sigma^2 \) and the parameter(s) of \( f(r) \), which are easy to estimate due to the existence of closed-form expressions for the summary statistics. The process is also relatively easy to simulate based on the location dependent rejection of points generated according to the homogeneous PPP with \( \lambda^2 = \max(\lambda(x)) \). Moreover, the intensity of the process is expressed as \( \lambda = \exp(\mu + \sigma^2/2) \).

### 3.2 Cluster processes

Generally, cluster processes can be viewed as realizations of two hierarchical point processes: the parent point process, which distributes clusters, and the daughter point process, which distributes points around the parent point location. Typically the parent process does not lead to the generation of points, therefore the cluster process consists of the union of all daughter points.

**Matern cluster process** The Matern cluster process is a doubly Poisson cluster process, i.e., parent points are generated according to a Poisson process with intensity \( \kappa \), and daughter points, with the mean number of them \( \mu \), are uniformly distributed inside the circle of radius \( R \) centered at each parent point. The intensity of the daughter process takes the following form:

\[
\lambda_d(r) = \frac{\mu}{v(b(o,R))} \mathbb{1}_{[b(o,R)]}(r),
\]

where \( \mathbb{1}_{[b(o,R)]}(r) \) is an indicator function, \( b(o,R) \) denotes a circle centered at parent point \( o \) with radius \( R \), and \( v(\cdot) \) is the Lebesgue measure, which in the Euclidean space corresponds to the surface area. The intensity of the process is \( \lambda = \kappa \mu \).

**Thomas process** Similarly to the MCP, the Thomas process is a doubly Poisson cluster process, where the parent points are distributed according to a Poisson process.
with intensity \( \kappa \), while the daughter points are normally distributed, i.e., according to the intensity function:

\[
\lambda_d(r) = \frac{\mu}{2\pi\sigma} \exp\left(-\frac{r^2}{2\sigma^2}\right),
\]

where \( \mu \) is the mean number of daughter points and \( \sigma^2 \) is the variance. The intensity of the process is again \( \lambda = \kappa\mu \).

### 3.3 Summary characteristics

In spatial statistics there are numerous summary characteristics that describe and quantify local groupings of points, typically through their relative positions or inter-point distances. In the following we briefly introduce the metrics that we will subsequently use for model fitting and goodness-of-fit tests.

Generally, a point process that drives multi-network patterns is unknown. Therefore, to obtain each of the following metrics we are required to utilize non-parametric estimators with appropriate methods to mitigate finite observation window distortions that occur at the edges. Moreover, since each of the analysed deployments has some local inhomogeneities, we normalize each characteristic obtained against inhomogeneities of the corresponding deployment.

**Empty space function** Empty space function \( F(r) \) describes the distribution of distance \( r \) from a point \( o \in W \) (for stationary point processes this corresponds to the origin) to the nearest point of the point process \( \Phi \):

\[
F(r) = P (|o - \Phi| \leq r) = P (\varphi(b(o,r)) > 0), \quad r \geq 0.
\]

(3)

We use the empty space function as the first-order goodness-of-fit statistic.

**J-function** One way to investigate the interaction between two point patterns, each with a different type of points, observed within the same window, is to utilize the cross-type J-function [29]. The cross-type J-function is the ratio between the distributions of the distances from an arbitrary fixed point of type \( i \) to the nearest type \( j \) point, therefore it can be used to quantify clustering or repulsion between two sets of points. The cross-type J-function is given as:

\[
J_{ij}(r) = \frac{1 - G_{ij}(r)}{1 - F_{j}(r)},
\]

(4)

where \( G_{ij}(r) \) is the nearest neighbour distance function, which describes the distribution of distances from the points of type \( j \) to the points of type \( i \), and \( F_{j}(r) \) is the empty space function of the point pattern of type \( j \). In case the two analysed point patterns are independent \( J_{ij} = 1 \), as is the case with any homogeneous PPP realizations. \( J_{ij} < 1 \) suggests that the points of different types aggregate, while \( J_{ij} > 1 \) suggests repulsion. However, the cross-type J-function taking value 1 for all \( r \) is not a sufficient characterization of independence between the two patterns [29]. It is also worth noting that \( J_{ij} \) is not symmetric in \( i \) and \( j \).

**L-function** The L-function is defined as \( L(r) = \sqrt{K(r)/\pi} \) for \( r \geq 0 \), where \( K(r) \) is a second-order summary characteristic which describes the mean number of points within radius \( r \) from points in point process \( \Phi \). The L-function is particularly useful in the estimation of model parameters, such as the interaction distance (maximum inter-point distance at which interaction force is at work) or interaction force (the force that describes mutual attraction or repulsion between points in a point process). For the PPP the L-function takes a convenient linear form \( L(r) = r \). Hence, for some pattern \( \varphi \), \( L(r) < r \) will indicate regularity, while \( L(r) > r \) will indicate clustering. We use the L-function to examine fitness of the estimated models with respect to second-order characteristics.

**Pair-correlation function** The pair-correlation function \( g(r) \) for a stationary point process \( \Phi \) with intensity function \( \lambda \) and second order product density \( \rho^{(2)}(r) \), which is the joint probability that there are two points of \( \Phi \) in infinitesimally small areas surrounding two locations separated by distance \( r \), is defined as follows [26]:

\[
g(r) = \frac{\rho^{(2)}(r)}{\lambda^2}.
\]

(5)

Informally, the pair-correlation function characterizes the frequency of inter-point distances in the process. Herein, we use the pair-correlation function for fitting clustered point processes to our data. The fitting method we use requires an analytical expression for the summary statistic in question. The expressions for the pair-correlation function of the analysed clustered point processes are summarized in Tab. 1.

### 4 Datasets and model fitting

#### 4.1 Datasets

The aim of our study is to find a stochastic model that characterizes multi-networks aggregated for the purpose of infrastructure sharing. The input data we use consists of either real BS information or radio license information (for simplicity, we will assume that an assigned radio license is equivalent to a BS deployed in the given location), which includes spatial coordinates of the BS sites and unique cell identifiers, for GSM and UMTS technologies in three countries: Ireland, Poland, and the United Kingdom. The data was extracted from publicly available information collected by the national telecommunications regulators. Each of the regulators collects data supplied by the local mobile network operators (MNO), who also ensure accuracy and keep the information updated.

The Irish database contains BS site information updated by the Irish MNOs and maintained by the Irish communications regulator ComReg. Based on the quarterly report from ComReg (Q4 2013), Ireland had four
TABLE 1: Analytical expressions for the pair-correlation function of the analysed clustered models, where \( h(z) = \frac{16}{\pi (z \arccos(z) - z^2 \sqrt{1 - z^2})} \) for \( z \leq 1 \) and \( h(z) = 0 \) otherwise.

<table>
<thead>
<tr>
<th>Model</th>
<th>LGCP</th>
<th>MCP</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g(r) )</td>
<td>( \exp(\sigma^2 f(r)) )</td>
<td>( 1 + (4\pi R \kappa)^{-1} h(\frac{r}{\sigma}) )</td>
<td>( 1 + (4\pi \kappa \sigma^2)^{-1} \exp(-\frac{r^2}{2\sigma^2}) )</td>
</tr>
</tbody>
</table>

major MNOs (Vodafone, eircom Group Mobile, O2 and Three) together sharing 93.9% of the Irish mobile market. We deem the Irish dataset as the most reliable, as ComReg requires mobile infrastructure providers to report all the masts and structures that contain “mobile base stations”. Therefore, we use the Irish dataset to confirm the reliability of the other two datasets. The database for Poland contains spatially allocated radio license information from all MNOs operating in the Polish mobile market, available through the national regulator UKE\(^5\). By 2013 there were four major MNOs in Poland, namely Orange, Play, Plus and T-Mobile, which had a combined 98.7% of the subscriber share and 99.7% of the market revenues. The database for the UK was compiled by OfCom (UK’s Office for Communications) based on BS site information voluntarily supplied by the local MNOs\(^6\). As of 2012 there were five major local MNOs: O2, Orange, T-Mobile, Three, and Vodafone, all of them together having a higher than 90% market share.

From each of the databases we have selected a number of city areas. For each city area we have focused on BS site information for two radio technologies: GSM (900 MHz) and UMTS (2100 MHz), and macro- and microcells only (if the transmit power or cell radius information was available). For each city area we have extracted this information for an observation window of size 10-by-10 km with arbitrarily selected central locations\(^7\). The choice of the observation window was driven by our desire to capture features of BS deployments at local scales, and to correctly assess environmental forces that underlay point distributions for multi-network deployments. In our study, we have simplified each co-location to a single BS site, which is a reasonable working assumption given that inter-operator co-location accounts on average for only 10% of all base stations. Tab. 2 summarizes basic information on all the selected locations.

Since the analysis covers different geographical areas as well as various statistical metrics, for clarity we have decided to focus our presentation on the study of Dublin. The results for other areas (listed in Tab. 2) are substantially the same and they are included in the estimated point process model, which we discuss at the end of this section.

7. In order to test whether our results are affected by the choice of the observation window we also looked at observation windows of size 2.5-by-2.5 km, and 5-by-5 km. In the main, the results are consistent across those three observation windows.

4.2 Fitting method

A desirable model is one that is easy to interpret, estimate and simulate, while still preserving features of the observed BS patterns. Herein, we provide the results of fitting PPP, SP, MCP, TP and LGCP, with exponential and stable correlation functions, to the multi-network patterns under study. To fit the first two we have used the maximum pseudolikelihood method, described in [3]. When the pseudolikelihood function is not easily tractable, which is typically the case with clustered point processes, we apply the minimum contrast method (MCM) [7]. In general, the MCM seeks parameter values for the analytical expression of a summary statistic \( C \) which minimize the difference between \( C \) and its non-parametric estimate obtained from the analysed patterns \( C_0 \), i.e.:

\[ D_0 = \int_{r_{\text{min}}}^{r_{\text{max}}} |\widehat{C}(r) - C_0(r)|^p dr \] (6)

where \( C_0 \) is the theoretical expression for the summary characteristic parameterized with \( \theta \), while \( p \) and \( q \) are simulation parameters. Applying the MCM to the analysed processes is natural, as for each of those processes we can obtain closed-form expressions for the \( n \)-th order summary characteristics. Specifically, we use the pair-correlation function, for which closed-form expressions are given in Tab. 1. Setting \( p = 2 \), and \( q = 1 \) (for the MCP and TP), allows us to obtain closed-form expressions for the parameters of the fitted processes, which we summarize in Tab. 3. We derive these expressions in Appendix A.

In our analysis we look at a range of \( r \) values, where \( r_{\text{min}} = \min_{i \neq j} \{||x_i - x_j||\} \) (as suggested in [28]) and \( r_{\text{max}} \) is set as the quarter of the side-length of the square observation window, which, according to our simulations, provides a good balance between observation of the local features and computational effectiveness. We perform our analysis using open-source statistical computing software R [30], and, in particular, we use functions from the \textit{spatstat} library [31]. For the goodness-of-fit tests we use the envelope test, which can be interpreted as the significance test [6], with a significance level in a one-sided test equal to \( 1/(k+1) \), where \( k \) is the number of trials. For the envelope test we utilize both the L-function (second-order statistic) and the empty space function (first-order statistic), as suggested in [6] to minimize the bias caused by testing our model against the data that was used to estimate model parameters.
Before we turn to the multi-network analysis, we investigate the L-function results for single-networks to confirm observations made in [3]. In Fig. 2 we observe the L-function results for selected GSM and UMTS deployments, along with 99 realizations of a Poisson fit (corresponding to a significance level of 0.01), and its well-known closed-form representation. The results somewhat confirm the observations made in [3], as the observed point patterns for various operators in different areas exhibit either some mild forms of inhibition or uniformity for the whole range of inter-point distances.

We now turn our attention to the relationship between multi-operator deployments. If single-networks were realizations of independent Poisson processes, we would expect that the observed multi-network patterns should not exhibit any forms of regularity. However, as noted previously the resulting multi-network point patterns clearly exhibit clustering. This fact and the single-operator fit would suggest that the realizations of networks of different operators are not independent from each other, which also makes sense intuitively, as one would expect that BS placements account for local phenomena, such as high availability of potential subscribers. In order to test this hypothesis, we look into the cross-type J-function results for multiple operators in Dublin. Since the J-function involves comparison between different patterns, we drop for a moment the assumption of simplicity and allow for BSs of different operators to co-locate. Both the results for GSM (Fig. 3a) and UMTS (Fig. 3b) confirm that there is significant clustering for the majority of inter-point distances, with some repulsive behaviour for Three-Meteor/Meteor-Three and Vodafone-Meteor pairs at larger distances, which may reflect the structure of the Irish mobile market and its existing infrastructure sharing among operators. Note that these results, as any results in empirical statistics, are merely an indication of clustering among the observed BS patterns rather than a formal proof.

4.3 Fitness results - Dublin case

Let us now move on to describing the results of fitting a model to multi-networks. As stressed previously, we present in more details the results for Dublin, while noting that the fitting results for the remaining cities follow similar patterns, as can be seen in Fig. 4. In the figure we can see that real multi-networks are characterized by

---

**TABLE 2: Basic information on the extracted point patterns**

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Name</th>
<th>Center (longitude,latitude)</th>
<th>Number of extracted BSs (UMTS/GSM)</th>
<th>Inter-operator co-location (UMTS/GSM) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ireland</td>
<td>Dublin</td>
<td>53.3478°N;6.2597°W</td>
<td>447/409</td>
<td>5.0/4.0</td>
</tr>
<tr>
<td>Poland</td>
<td>Kraków</td>
<td>50.0614°N;19.9383°E</td>
<td>348/273</td>
<td>8.0/8.0</td>
</tr>
<tr>
<td></td>
<td>Poznań</td>
<td>52.4°N;16.9167°E</td>
<td>400/362</td>
<td>10.0/14.0</td>
</tr>
<tr>
<td></td>
<td>Warszawa</td>
<td>52.2333°N;21.0167°E</td>
<td>812/916</td>
<td>12.0/8.0</td>
</tr>
<tr>
<td>UK</td>
<td>Birmingham</td>
<td>52.4831°N;1.8936°W</td>
<td>281/144</td>
<td>5.0/0.0</td>
</tr>
<tr>
<td></td>
<td>Leeds</td>
<td>53.7997°N;1.5492°W</td>
<td>238/115</td>
<td>16.0/0.0</td>
</tr>
<tr>
<td></td>
<td>Liverpool</td>
<td>53.4°N;3°W</td>
<td>266/122</td>
<td>7.0/0.0</td>
</tr>
<tr>
<td></td>
<td>London</td>
<td>51.5072°N;0.1275°W</td>
<td>1442/1513</td>
<td>3.0/1.0</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
<td>53.4667°N;2.2333°W</td>
<td>303/155</td>
<td>7.0/0.0</td>
</tr>
</tbody>
</table>

**TABLE 3: MCM expressions for model parameters**

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Model</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>LGCP</td>
<td>$\bar{\mu} = \log(\lambda) - \sigma^2/2$</td>
<td>$\hat{\sigma}^2 = \frac{B_{L}(\hat{\sigma})}{\Lambda_{L}(\sigma)}$</td>
<td>$\hat{\beta} = \arg\max\frac{B_{L}(\sigma)^2}{\Lambda_{L}(\sigma)}$</td>
<td></td>
</tr>
<tr>
<td>MCP</td>
<td>$\bar{\mu} = \frac{1}{\lambda}$</td>
<td>$R = \arg\max\frac{B_{M}(R)^2}{\Lambda_{M}(\sigma^2)}$</td>
<td>$R = \arg\max\frac{B_{M}(R)^2}{\Lambda_{M}(\sigma^2)}$</td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>$\bar{\mu} = \frac{1}{\lambda}$</td>
<td>$\hat{\sigma}^2 = \arg\max\frac{B_{P}(\sigma^2)^2}{\Lambda_{P}(\sigma^2)}$</td>
<td>$\hat{\sigma}^2 = \arg\max\frac{B_{P}(\sigma^2)^2}{\Lambda_{P}(\sigma^2)}$</td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 2: The L-function results for Meteor GSM and UMTS deployments in Dublin (marked with triangles) with the envelope of the fitted Poisson model at a significance level of 0.01 (grey area), and the theoretical value for the Poisson process (marked with circles).
order to show that multi-network patterns have strong clustering features, which cannot be captured by processes describing single-networks, we also provide fitting results for other point processes: the PPP (to disprove the complete spatial randomness hypothesis [6]), and the superposition of independent SP realizations. The figures also include envelopes, which represent the maximum and minimum of 99 realizations of the fitted model, which are used to assess the goodness-of-fit of the model in the envelope test [6].

As we can see from Fig. 5 and Fig. 6 both the PPP and the superposition of independent realizations of the SP can be rejected as candidate models, since they do not produce sufficiently large empty spaces, both in the GSM and UMTS case. Moreover, these point processes are unable to model short-range clustering as can be seen from the results obtained for the L-function in Fig. 7 and Fig. 8. Doubly Poisson clustered models (MCP and TP), seen in Fig. 5 and Fig. 6, do not pass the envelope test for the F-function, as they under-represent the occurrence of empty spaces in the patterns. Moreover, when second-order characteristics are taken into account, these point processes seem to insufficiently capture clustering, while completely missing out on the long-range repulsion, see Fig. 7 and Fig. 8.

By far the most satisfying results we have observed come from the LGCP models, which in the UMTS case passed the F-function envelope test for both models, see Fig. 6, and for the GSM case passed the F-function test for the model with the stable correlation function, see Fig. 5. Apparently, the trade-off for achieving better fitness is related to increased variance in the occurrence of empty spaces prevalent in the modelled patterns. When it comes to second order statistics, we can see in Fig. 7 and Fig. 8 that both the LGCP models are able to account for the short-range clustering, while also being able to capture long-range repulsion in the patterns. In our simulations we have also tried other than the exponential and stable spatial correlation functions (for example, Gauss or spherical); however, fitting against a model based on either of the two functions gave us the most satisfying results.

### 4.4 Fitness results - other areas

When fitting the models to other patterns from our dataset, we have observed similar results, i.e., we could model the short-range clustering with any of the clustered point process, while the long-range repulsion was best represented with the LGCP model. For some areas, however, the LGCP was insufficient to capture all the inter-point interactions; in those cases, in order to achieve fits one could adjust the intensity function of the process with some additional environmental covariates or inject an inter-point interaction force into the density function of the LGCP model. Both of these, however, destroy any potential analytical tractability of the model, with the risk of over-fitting, which goes against our goal of model universality. In addition to a number of urban

![Fig. 3: The cross-type J-function results for combinations of pairs of major MNOs in Dublin: Meteor (Me), Three (Th) and Vodafone (Vo). Recall that by definition $J_{ij}$ is not symmetric and also $J_{ij} < 1$ indicates clustering (while $J_{ij} > 1$ inhibition) between the BS locations of any two MNOs.](image)

![Fig. 4: The pair-correlation function for real deployments in various cities; Dublin marked with triangles and the envelope of the fitted Poisson model (at a significance level of 0.01) marked with solid lines.](image)
cases, we have looked also at some areas with lower density of base stations, such as Athlone, Ireland. In order to obtain meaningful results for rural areas we had to increase the observation window size to 25-by-25 km. However, even increasing the observation window left us with relatively small patterns that contained significant amount of empty spaces. In those patterns we could still observe some mild forms of clustering at smaller distances, which in most cases was enough to reject the PPP hypothesis. The considered clustered point processes were able to capture this clustering as well as the following uniformity in the pattern, although the fitted models had significant variance. And since these results were not entirely consistent across the other rural areas we have analysed, we have decided to focus on modelling urban locations.

Based on the fitness results we have concluded that the most promising point process model to model multi-networks in urban areas is the LGCP. Furthermore, we have observed that we can almost unambiguously classify the parameters of the models according to the type of deployment, i.e., urban, and dense urban. Tab. 4 summarizes the two classes of models we have found. It is interesting that for the urban scenario the standard deviation is relatively low (around 10%) for both shape parameters, i.e., $\sigma^2$ and $\beta$, which suggests similar shape of the intensity function $\lambda(x)$ for each of the geographic areas considered, with only scale parameter $\mu$ (precisely $\exp(\mu)$) having higher deviation to account for the different number of BSs in each area. For the two dense urban areas (London and Warsaw) the estimated model parameters differ significantly, which is presumably the result of much higher number of base stations and much richer structure of the area (higher concentration of BSs, larger number of neighbourhoods with concentration of BSs).

5 Performance analysis

Having evaluated the statistical match between our proposed point process models and real multi-network deployments, we now turn our attention to wireless network-relevant evaluation based on the radio access network sharing scenarios of: infrastructure, spectrum full network sharing, as studied in [14]. As a metric for our evaluation we use the coverage probability [26], i.e., the complementary cumulative distribution function (CCDF) of SINR, which we obtain for each estimated deployment model and each scenario via Monte Carlo simulations.

In the following we assume that the power of the signal received by a typical user located in the origin $(0,0)$ is affected by two factors: pathloss $l(x)$ and power fading $h_x$, where $x$ denotes the location of the serving BS. The pathloss function $l : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ is of the form $l(x) = |x|^{-\alpha}$, where $\alpha$ is the pathloss exponent, and the power fading between the user and the serving transmitter $x$ is spatially independent and exponentially distributed (e.g., Rayleigh fading) with mean $\mu_x$ which corresponds to the transmit power of a BS located in $x$, i.e., $h_x \sim \exp(1/\mu_x)$. In the following we analyse two approaches to assigning transmit powers to BS transmitters: initially, we set $\mu_x = 1$ for all $x \in \Phi$, and, subsequently, we extend our model and assign transmit powers according to the scheme we proposed in [32]. In addition, we make the following assumptions: (i) our results apply to the downlink only, (ii) worst-case interference is considered, i.e., all transmitters transmit simultaneously, (iii) we assign individual BSs to operators uniformly at random, (iv) noise is unitary for all the receivers, and (v) $\alpha = 4$. For each realization of the estimated model we generate 2000 user locations, over which we average our results. We also apply edge-effect correction based on an enlarged window, i.e., we assume that the user locations lie inside a square window centered at the analysed area with the side equal to half the side of the window containing the original area.

Coverage probability - unitary transmit power

Fig. 9 presents the coverage performance of our models for each of our sharing scenarios and technologies when BS transmit power is unitary. In each case the LGCP model with the stable correlation function gets the closest to the coverage provided by a real multi-operator network. In the case of full network sharing our models significantly over-estimate the amount of coverage in the network (approx. 2 dB vertical shift for the “LGCP stable” model). However, in the case of infrastructure sharing we observe that the “LGCP stable” model provides a very tight match to a shared network performance. The spectrum sharing scenario provides significantly worse coverage, which should come as no surprise as we assumed worst-case interference. Interestingly, in the case of spectrum sharing we observe a cross-over point, which we have reported also in [14], which is related to the amount of clustering in the evaluated patterns. Clearly, a real multi-network provides higher clustering than any of our models can produce. It is also worth noting that, while statistical tests have shown that a superposition of independent realizations of the Strauss process provide a reasonably good match to the real data, according to our coverage analysis this way of modelling multi-operator patterns fails to reproduce coverage properties of multi-networks in wireless network sharing scenarios.

Coverage probability - transmit power assigned as proposed in [32]

In [32] we studied the distribution of maximum downlink transmit power in macro- and microcell base stations using spatial statistical analysis of 3G radio access network deployment data for the UK. Our major finding was that “downlink macro- and microcell transmit power allocations in a real 3G network, both in single and multi-operator cases, can be considered independent and identically distributed”, which has led us to propose a random assignment from a fixed probability distribution as an assignment method for base station downlink
transmit powers. The first thing we observe in Fig. 10 is that BS transmit power assigned from a model based on real data has a negligible impact on the coverage of the analysed network model (the coverage probability is degraded by roughly 1% for each SINR point), which is consistent with our prior observations [32]. This is true also when the noise level of -130 to -90 dBW is included at the receiver. This result is good news as it allows us to use, without loss of generality, a simplified model for multi-network that includes unitary transmit power.

What we can conclude from the above results is that the chosen point processes provide a reasonable representation of real base station deployment when sharing scenarios are considered, even when exact information about operator-BS assignments is not available. Furthermore, similar results hold even when transmit powers are assigned from a model based on real data [32], which also suggests that this type of analysis may be robust to certain operator settings.

### 6 Discussion and Conclusion

This paper is, to our knowledge, the first to study multi-operator base station deployments using spatial statistics and stochastic geometry. What we have shown here is that networks deployed by independent operators tend to cluster at shorter distances, which may correspond to areas of high demand, and repulse at long-distances, which would presumably correspond to coverage provisioning and low availability of shared sites. This result implies that if one wants to model (or simulate) the performance of a generalized shared infrastructure between \( N \) mobile operators, one cannot simply model the entire network as a superposition of \( N \) independent realizations representing single-operator networks. Instead, the approach we propose is to model such a network as a single clustered point process and indeed, as we have shown in our results, clustered point processes that represent well such networks can be found. What is even more appealing is that similar models may be used to represent networks of multiple operators in different countries (herein, we examined Ireland, Poland and the UK), which suggests that using such a model may come useful when deriving fundamental results about wireless network sharing, or evaluating the performance of protocols and management strategies designed for shared wireless infrastructure.

As part of our future work, we are studying, using stochastic geometric models, the impact of spatial distribution of base stations on the performance of infrastructure and spectrum sharing between multiple mobile operators. Our initial study, see [14], analyses the trade-off between infrastructure and spectrum sharing. However, one may extend our work also in other ways. One such possible extension includes modelling base station deployments as a multivariate LGCP, see [28], whereby additional variates describe, for example, whether a base station is private or shared, or some network relevant parameter such as the antenna height. These parameters, however, would have to be subject to thorough spatial statistical analysis. Obviously our methodology may also be applied to fit other point process models which may, for some applications, provide additional benefits related to, for example, more in-depth insights into the correlations present in multi-operator radio access network deployments.

### Acknowledgment

This material is based upon work supported by the Science Foundation Ireland under grants no. 10/CE/11853 and 10/IN.1/3007.

### References


### Table 4: Fitted model parameters

<table>
<thead>
<tr>
<th>Location type</th>
<th>Model</th>
<th>( \sigma^2 )</th>
<th>( \beta )</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>LGCP</td>
<td>3.9040 ± 0.3699</td>
<td>0.03 ± 0.0031</td>
<td>−0.5634 ± 0.2956</td>
</tr>
<tr>
<td>Dense Urban (London and Warszawa)</td>
<td>LGCP</td>
<td>2.0561, 2.7228</td>
<td>0.054, 0.0288</td>
<td>1.2665, 1.9477</td>
</tr>
</tbody>
</table>


Jacek Kibilda is currently a PhD candidate at CONNECT, Trinity College, The University of Dublin, Ireland. He received the M.Sc. degree from Poznań University of Technology, Poland, in 2008. Subsequently, he worked at Nokia Siemens Networks and EIT+ ICT Research Center, both in Wrocław, Poland, before in 2012 he joined Trinity College. His research focuses on architectures and modelling for future mobile networks.

Boris Galkin was awarded a BAI and MAI in computer & electronic engineering by Trinity College Dublin in 2014 and since then he has been working toward the Ph.D. degree at CONNECT, Trinity College, The University of Dublin, Ireland. His research interests include small cell networks, unmanned airborne devices and cognitive radios.

Luiz A. DaSilva holds the chair of Telecommunication at Trinity College Dublin. He also has a research appointment in the Bradley Department of Electrical and Computer Engineering at Virginia Tech, USA. His research focuses on distributed and adaptive resource management in wireless networks, and in particular wireless resource sharing, dynamic spectrum access, and the application of game theory to wireless networks. He is currently a Principal Investigator on research projects funded by the National Science Foundation in the United States, the Science Foundation Ireland, and the European Commission under Horizon 2020 and Framework Programme 7. He is a Co-principal Investigator of CONNECT, the Telecommunications Research Centre in Ireland. Prof DaSilva is an IEEE Fellow, for contributions to cognitive networking and resource management for wireless networks, and an IEEE Communications Society Distinguished Lecturer.
Fig. 5: GSM - Empty space function (F-function). The estimated F-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.1 (grey area), and the theoretical realization of the PPP (marked with circles). As we can see from the figure the “LGCP exponential” and “LGCP stable” fitted models pass the envelope test for Dublin.

Fig. 6: UMTS - Empty space function (F-function). The estimated F-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.1 (grey area), and the theoretical realization of the PPP (marked with circles). As we can see from the figure “LGCP stable” passes the envelope test for Dublin.
Fig. 7: GSM - the L-function. The estimated L-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model corresponding to a significance level of 0.01 (grey area), and the theoretical realization of the PPP (marked with circles).

Fig. 8: UMTS - the L-function. The estimated L-function for the multi-network in Dublin (marked with triangles) with the envelope of the fitted point process model at a significance level of 0.01 (grey area), and the theoretical realization of the PPP (marked with circles).
Fig. 9: The coverage probability comparison for real deployment in Dublin (circle) and various fitted point process models, for each radio access network sharing scenario and radio technology when transmit power is unitary for all BS transmitters. Clearly, “LGCP stable” (hexagonal star) provides the closest fit to real network data.

Fig. 10: The coverage probability comparison for real deployment in Dublin (circle) and various fitted point process models, for each radio access network sharing scenario and radio technology when transmit power is assigned to BS transmitters according to the scheme proposed in [32]. Again, “LGCP stable” (hexagonal star) provides the closest fit to real network data.