# The Critical Neighbourhood Range for Asymptotic Overlay Connectivity in Ad Hoc Networks

Sandrine Calomme and Guy Leduc
Research Unit in Networking
Electrical Engineering and Computer Science Department
University of Liège, Belgium
Email: {calomme,leduc}@run.montefiore.ulg.ac.be

March 6, 2006

**Abstract**

We first motivate the use of ad hoc overlays. In particular, we argue that overlay routing could play a role in the spreading of ad hoc networks.

We then define a simple criterion for neighbourhood: two overlay nodes are neighbours if and only if there exists a path between them of at most R hops, and R is called the (overlay) neighbourhood range. A small R may result in a disconnected overlay, while an unnecessarily large R would generate extra control traffic. We are interested in the minimum R ensuring overlay connectivity, the so-called critical R.

We study conditions on $R$ to achieve asymptotic connectivity of the overlay almost surely, i.e. connectivity with probability 1 when the number of nodes in the underlying ad hoc network tends to infinity (so-called dense networks) or when the size of the field tends to infinity (so-called sparse networks), under the hypothesis that the underlying ad hoc network is itself asymptotically almost surely connected.

For dense networks, we derive a necessary and sufficient condition on $R$, and for sparse networks we derive distinct necessary and sufficient conditions that are however asymptotically tight.

These conditions, though asymptotic, shed some light on the relation linking the critical $R$ to the number of nodes $n$, the field size $\ell$, the radio transmission range $r$ and the overlay density $D$ (i.e., the proportion of overlay nodes). These conditions can be considered as approximations when the number of nodes (resp. the field) is large enough. Since $r$ is considered as a function of $n$ or $\ell$, we are able to study the impact of topology control mechanisms, by showing how the shape of this function impacts the critical $R$.

**Keywords:** ad hoc networks, connectivity, overlay, topology control.

## 1 Introduction

Early wireless overlay propositions consisted of unified systems which incorporated a variety of different transmission technologies, exploiting the advantages of every involved media, in order to improve the global performance [1,2]. Then, overlay multicasting was broadly studied by the wired community, followed by the ad hoc one [3–8]. Similarly, with the increasing popularity of peer-to-peer networks, some studies recently appeared to build them in ad hoc networks.

However, the potential of overlaying in ad hoc networks or in hetereogeneous wireless environments including ad hoc networks, greatly surpasses the mere transposition of their advantages from the wired to the ad hoc world.

Unlike typical Internet applications, most applications of MANETs [9] involve one-to-many and many-to-many communication patterns [10]. Moreover, the grouping behaviour of the mobile users has been observed in actual field trials of local area wireless networks [11, 12]. We can thus consider ad hoc networks as made of one or several communities, many of them in a pervasive ad hoc environment, sharing at least one application over a common physical medium. As individuals have limited battery, and as the shared medium has a limited capacity, the use of a power control algorithm is highly recommandable [13]. This one however imposes to the member of a group to communicate with each other in a multi-hop fashion, through nodes that are not part of their community.

As the most appropriate routing protocol for a given community highly depends on the application it is running and on the network conditions, we cannot assume in spontaneous networks that every node has chosen the same routing solution, or knows a set of protocols that would meet everyone's needs [14].

Our proposition is to copy the layered approach of Internet [15]: agree only on a few unspecialized protocols at the physical, data link and routing layers, imposed by their proved qualities or de facto, and over this basic architecture, develop plenty of more specialized solutions, from routing to application. Any ad hoc network would provide a minimum network service, allowing a wide variety of communities to be incorporated into various realization of the complete protocols stack.

Overlay routing could promote the deployment of ad hoc networks, offering a very flexible ground for a variety of applications and underlying customized network protocols.

Overlay routing advantages come however at the expense of the overlays creation, usage and maintenance, that must be kept moderate. Consequently, a full mesh is probably not the most adapted overlay topology. A natural rule of thumb is to admit as overlay neighbours a set of close overlay nodes, the distance measure employed being the number of hops. Two approaches are possible. One can fix the cardinality of the set of neighbours or the maximum number of hops admitted between overlay neighbours. We adopt the latter one. In this case, the maximum distance between two neighbours is an integer value that must be sufficiently high to obtain a connected overlay but as low as possible to limit the amount of messages generated in the network by overlay nodes communication.

The parallel with topology control in ad hoc networks is obvious. To achieve connectivity, each ad hoc node could use its maximum transmission range, in order to reach many neighbours. However, mobile devices have a limited amount of battery power. Moreover, this would create a lot of interferences, reducing the overall capacity of the network. With a homogeneous topology control algorithm, all nodes adopt the same transmission range value. The critical transmission range problem consists of determining the minimum value that generates a connected network.

We have adopted a similar terminology for our problem: the maximum number of hops allowed between overlay neighbours is called the neighbourhood range and the determination of its best value the critical neighbourhood range problem.

This paper focuses on the necessary and/or sufficient conditions on the neighbourhood range to achieve asymptotic connectivity of the overlay almost surely, i.e. connectivity when the number of nodes in the basic graph tends to infinity (so-called dense networks) or when the size of the field tends to infinity (so-called sparse graphs).

We first demonstrate that in connected networks, as the network gets denser or larger, the shortest path between any pair of nodes draws close to the straight line. Thanks to this property, that we call the asymptotic path length theorem, we are able to derive an analytical solution to the critical neighbourhood range problem for dense networks. For a large class of sparse networks, we are also able to determine asymptotically tight bounds.

The main reason for addressing asymptotic connectivity is its mathematical tractability. We build on several asymptotic results on basic geometric graphs to derive properties of the overlays. Our asymptotic results can be seen as approximations of finite (real) networks either when the number of nodes (resp. the field) is large enough. Nevertheless our mathematical conditions already shed some light on the relation linking the number of nodes, the field size, the radio transmission range, the overlay density (i.e., the proportion of overlay nodes) and the overlay neighbourhood range to get a connected overlay.

We do take into account the potential use of a homogeneous topology control algorithm at the underlay level and allow the overlay density to evolve with the network size. In particular, if the overlay density diminishes, our results show how a compensation in neighbourhood range can keep the overlay still connected. They also point out that a more efficient topology control algorithm of an ad hoc network will require more traffic for the use and maintenance of overlays built on it.

This paper is structured as follows. In Sect. 2, we give an overview on previous related work over the critical transmission range. In Sect. 3, we precisely define the problem studied. In Sect. 4 and 5, we present analytical results, respectively for dense and sparse networks, and discuss some of their practical implications. We then conclude.

# 2 Related Work

In many realistic scenarios, node positions are not known in advance. Hence a probabilistic approach is used in every analytical study of the critical transmission range problem.

First studies of graph connectivity were developed in the context of the random graphs theory. A random graph is a graph generated by some random procedure [16]. In 1960, Erdos and Rényi [17] showed that for many monotone-increasing properties of random graphs, like connectivity, graphs of a size slightly less than a certain threshold are very unlikely to have the property, whereas graphs with a few more graph edges are almost certain to have it. This is known as a phase transition phenomenon.

In classical random graph models, there is no *a priori* structure. All vertices are equivalent and there is no correlation between different edges existence. In ad hoc and sensor networks, nodes are more likely to be direct neighbours if they are located close to each other. Therefore random geometric graphs are more suited to model them. Random geometric graphs are constructed by placing points at random according to some arbitrary specified density function on a $d$-dimensional Euclidean space and connecting nearby points [18]. Some of the geometric random graphs results can be applied in the study of connectivity in ad hoc and sensor networks [19]. Various transition phenomena can also be observed in geometric random graphs [20]. Monotone properties for this class of graphs have sharp treshold [21]. Asymptotically, as the network density tends to infinity, a critical value transmission range can thus be established [22], [23], [24].

In [25], the fixed radius model used in the geometric random graphs theory is extended by adding a new geometric parameter: the network deployment region size. The authors then use the occupancy theory [26] to obtain an asymptotic formula, when this parameter tends to infinity, for the critical transmission range in sparse as well as in dense networks.

In [27], the authors exploit the same model and, using a bin-covery technique, derive tighter bounds for the asymptotic connectivity.

We are not aware of any work related to the critical neighbourhood range problem for asymptotic overlay connectivity. In the following sections, we define it in more details, and we solve it using known results on the critical transmission range problem cited above.

# 3 Problem Definition, Notations and Discussion

We are interested in the asymptotic connectivity of overlay graphs built over asymptotically almost surely (a.a.s.) connected basic graphs.

These notions are defined in the following paragraphs. We then close this section with a discussion on the implicit assumptions we make in the problem and model specification.

## 3.1 Basic and Overlay Graphs

Consider an ad hoc network of $n$ nodes, deployed over a square field of length $\ell$, and where each node is assigned a transmission range of length $r$, the unit used for measuring $r$ and $\ell$ being identical. This network is modelled by a random geometric graph denoted $g(n, r, \ell)$ which has the following properties. The vertices of $g$ are uniformly and independently distributed on a square of size $\ell \times \ell$. They can either have been disseminated following the uniform distribution of $n$ points or by a spatial homogeneous Poisson point process of mean $n$. There exists an edge between each pair of vertices if and only if the Euclidean distance between them is not greater than $r$.

Let then $g(n, r, \ell)$ be a connected graph, $D$ be a real number with $0 \leq D \leq 1$ and $R$ be an integer with $R \geq 1$. An overlay graph $G(n, r, \ell, D, R)$ denotes a graph with the following properties. The $D$ parameter represents the overlay nodes density. The number of vertices of $G$ equals a proportion $D$ of the number of vertices of $g$. These are randomly and uniformly selected in the vertices set of $g$, which is called its basic graph. The parameter $R$ is called the neighbourhood range. There exists an edge between a pair of vertices $(v_1, v_2)$ if and only if the shortest path in $g$ from $v_1$ to $v_2$ contains less than or exactly $R$ hops.

In the following, in conjunction with the ad hoc and sensor networks terminology, the vertices of an overlay graph will be referred to as overlay nodes and the vertices of its basic graph as nodes.

## 3.2 Asymptotic Connectivity

We use two models for studying the asymptotic connectivity of random geometric graphs, that we qualify as dense and sparse.

### 3.2.1 Dense Model

**Definition** The field length $\ell$ is a constant. All other parameters are functions of the number of nodes. For example, $r(n)$ can be decreasing when $n$ increases, which is a desired behaviour for minimizing the capacity loss due to interferences.

**Notations** In this context, the $\ell$ parameter is assumed to be set to one and can be omitted in the notations. A basic graph can be denoted by $g(n, r(n))$ and an overlay graph by $G(n, r(n), D(n), R(n))$ or $G(g, D(n), R(n))$. We may generally simply write $g(n, r)$, $G(n, r, D, R)$ or, if $g(n, r)$ is given, $G(g, D, R)$.

**Asymptotic Connectivity** A dense graph is connected asymptotically almost surely if and only if the probability that it is connected tends to one as its *number of vertices* tends to infinity.

$$\text{Dense graph } \mathcal{G} \text{ is connected a.a.s.} \iff \lim_{n\to\infty} \mathrm{P}[\mathcal{G} \text{ is connected}]=1.$$

Note that for overlay graphs, the vertices are the overlay nodes. This means that $D(n)$ must be such that $\lim_{n\to\infty} D(n)n = +\infty$.

**Use** This model is only suited for studying the asymptotic behaviour of dense networks because the vertices density $\frac{n}{\ell^2}$ tends to infinity as $n$ does.

### 3.2.2 Sparse Model

**Definition** All parameters are functions of the field length $\ell$. This was the model used in [25] to study the critical transmission range in sparse networks.

**Notations** A basic graph should be denoted by $g(n(\ell), r(\ell), \ell)$ and an overlay graph by $G(n(\ell), r(\ell), \ell, D(\ell), R(\ell))$ or $G(g, D(\ell), R(\ell))$. We may also simply write $g(n, r, \ell)$, $G(n, r, \ell, D, R)$ or, if $g(n, r, \ell)$ is given, $G(g, D, R)$.

**Asymptotic Connectivity** A sparse graph is connected asymptotically almost surely if and only if the probability that it is connected tends to one as the *field length* tends to infinity.

$$\text{Sparse graph } \mathcal{G} \text{ is connected a.a.s.} \iff \lim_{\ell\to\infty} \mathrm{P}[\mathcal{G} \text{ is connected}]=1.$$

**Use** In this context, the node density $\frac{n}{l^2}$ might either converge to 0, or to a constant $c > 0$, or diverge as the size of the deployment region grows to infinity, depending on the relative values of $r$, $n$, and $l$. This model is thus more general than the dense one and suited for studying the asymptotic behaviour of sparse as well as of dense networks.

Notice that the sparse appellation, which was used to present this model in [25], could be a little confusing. However, as exposed in the following, the theorems based on the dense model are more convenient than the ones based on the sparse one. For this reason, the latter should only be used to analyze sparse networks. This justifies the model denomination.

## 3.3 Problem and model discussion

### 3.3.1 Connected basic graph

We consider only connected basic graphs. This seems reasonable to us as a disconnected basic graph will not provide connected overlays, whatever the neighbourhood range is, unless all the overlay nodes are concentrated in a connected part of it.

### 3.3.2 Asympotics

Many asymptotic properties of random geometric graphs have been demonstrated [18]. In particular, we mentioned in Sect. 2 several studies of the asymptotic connectivity of ad hoc networks, while the connectivity probability of a finite network, because of its complexity, has been the subject of very few analytical studies [28].

### 3.3.3 Homogeneous transmission range assignment

The transmission range is represented as a function of the number of nodes, directly in the dense case and indirectly, via its dependence to $\ell$, in the sparse case. This allows us to model a possible topology control protocol running on the ad hoc network, which would reasonably reduce the transmission range as the number of nodes increases, in order to conserve energy and global network capacity. We however implicitly limit ourselves to homogeneous topology control protocols, i.e. protocols which assign the same transmission range to all nodes.

This assumption greatly simplifies further mathematical developments and seems realistic in the context of our study. A common transmission range at each node provides some appealing features, that can be consulted in [29], such as the creation of bidirectional links only. Moreover, it is shown in [13] that, under a homogeneous spatial distribution, choosing a common transmission range can decrease capacity at most by a factor of $\sqrt{\ln n}$, where $n$ is the number of nodes, in comparison to allowing the flexibility of a different power level for each packet at each node [29]. This means that asymptotically a common power is nearly optimal in terms of network capacity [30]. Finally, as we use a uniform distribution of nodes and study an asymptotic property, more sophisticated topology control algorithms would intuitively lead to transmission range values converging in probability to a common function $r(n)$. All these reasons make us believe that a homogeneous transmission range assignment is both general and adapted.

# 4 Dense Networks

## 4.1 Known Results on Basic Graphs

Consider a basic graph $g(n, r)$. Let us build a graph $g\prime(n, r\prime)$ that has the same nodes set as $g$ and such that there is an edge between every pair of nodes. Let $M_n$ denote the longest edge length of the minimal spanning tree built on $g\prime$. In [31], Penrose demonstrated that the graph $g(n, r)$ is connected if and only if $r \geq M_n$ and [1]

$$\forall \alpha \in R : \lim_{n \to +\infty} P[n\pi {M_n}^2 - \ln n \leq \alpha] = \exp(-e^{-\alpha}) \tag{1}$$

This implies directly the following theorem.

**Theorem 4.1** *(Asymptotic connectivity of dense basic graphs)*
*A graph $g(n, r)$ with*

$$\pi r^2 = \frac{\ln n + k(n)}{n}$$

*is connected a.a.s. if and only if* $\lim_{n \to +\infty} k(n) = +\infty$.

The same result was demonstrated by Gupta and Kumar for a uniform distribution of nodes over the unit disk [22].

Note that for dense networks, a unique condition has been demonstrated to be both sufficient (when it is fulfilled, the graph is a.a.s. connected) and necessary (when it is not fulfilled, the graph is not a.a.s. connected).

## 4.2 Minimal Neighbourhood Range

**Theorem 4.2** *(Necessary condition for the asymptotic connectivity of dense overlay graphs)*
*An overlay graph $G(n, r, D, R)$ with*

$$\pi(Rr)^2 = \frac{\ln(\lceil Dn \rceil) + K(n)}{\lceil Dn \rceil} \tag{2}$$

*is not a.a.s. connected if* $\lim_{n \to +\infty} K(n) \neq +\infty$.

*Proof:* Let $G(n, r, D, R)$ be an overlay graph. Consider a graph $g\prime(\lceil Dn \rceil, Rr)$ such that the vertices sets of $G$ and $g\prime$ are identical. By definition, if there exists an edge in $G$ between two vertices $v_1$ and $v_2$ then the shortest path between them contains less than or exactly $R$ hops. As the distance between two consecutive nodes on a path cannot be longer than the transmission range $r$, the maximal distance between $v_1$ and $v_2$ is thus $Rr$ and this edge also exists in $g\prime$. Consequently, the edges set of $G$ is included in the edges set of $g\prime$. If $g\prime$ is not connected, then $G$ neither is.

Applying Theorem 4.1 to a graph $g\prime(\lceil Dn \rceil, Rr)$ , we obtain a necessary condition for the asymptotic connectivity of an overlay graph $G(n, r, D, R)$. ■

[1]Note that the theorem in [31] is more general. We isolated here the results that are of direct interest to us.

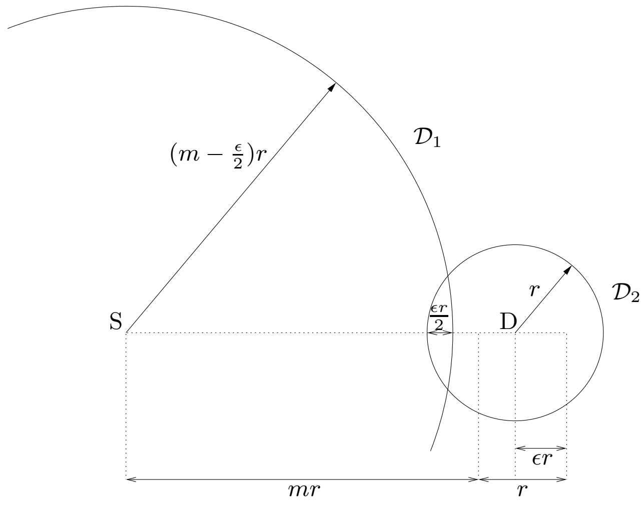


Figure 1:

## 4.3 Sufficient Neighbourhood Range

We start with the following lemma, whose proof is given in annex.

**Lemma 4.3** *Let $X_S$ be a random variable designating the number of nodes on a surface $S$ with $0 \leq S \leq 1$. For the uniform distribution of $n$ nodes, as for the Poisson two-dimensional spatial distribution of mean $n$ on the unitary square, $P[X_S = 0] \leq \exp(-nS)$.*

Exploiting this lemma, we can derive the following theorem on the asymptotic path length.

**Theorem 4.4** *(Asymptotic path length)*
*Let $g$ be an a.a.s. connected graph and $m$ be a strictly positive integer. Let $n_1$ and $n_2$ be two nodes of $g$. If the Euclidean distance between $n_1$ and $n_2$ is strictly less than $mr$, then there exists a.a.s. a path between them composed of less than or exactly $m$ hops.*

*Proof:* *(the asymptotic path length theorem in the context of dense graphs)*

We adopt an inductive approach. Assume that $n_1$ is located at point $S$ and $n_2$ at $D$. If $m = 1$, then the Euclidean distance between $S$ and $D$, denoted by $|SD|$, is strictly less than $r$. The nodes $n_1$ and $n_2$ are thus neighbours and there exists a path of one hop between them; the property is valid.

Let us now prove that if the property is valid for an integer $m$, then it is also valid for the integer $m+1$.

Assume that $|SD| = (m + 1 - \epsilon)r$ with $m > 0$ and $0 < \epsilon \leq 1$.

Let us draw a disk $\mathcal{D}_1$ centered on $S$ and of radius $(m-\frac{\epsilon}{2})r$ and another disk $\mathcal{D}_2$ centered on $D$ and of radius $r$, as in Fig. 1. The disks have a non-empty intersection, that we denote $\mathcal{I}$.

If the three following conditions are all satisfied, then there is a path of at most $m + 1$ hops between $n_1$ and $n_2$ :

1. there is a node $n_i$ in $\mathcal{I}$,
2. there is a path of at most $m$ hops between $n_1$ and $n_i$,
3. there is a path of length 1 between $n_i$ and $n_2$.

In terms of probabilities, this can be written:

$$\begin{aligned} &P[path\_len(n_1, n_2) \leq m+1] \\ &\quad \geq P[some\, n_i\, in\, \mathcal{I}] \\ &\quad\quad \times P[path\_len(n_1, n_i) \leq m \,|\, some\, n_i\, in\, \mathcal{I}] \\ &\quad\quad \times P[path\_len(n_i, n_2) = 1 \,|\, some\, n_i\, in\, \mathcal{I}] \end{aligned}$$

By geometric construction, if there is a node in $\mathcal{I}$ then this node is a neighbour of $n_2$.
Thus the third probability equals one and, asymptotically, we have:

$$\begin{aligned} &lim_{n\to\infty}\, P[path\_len(n_1, n_2) \leq m+1] \\ &\geq lim_{n\to\infty}\, P[some\, n_i\, in\, \mathcal{I}] \times P[path\_len(n_1, n_i) \leq m \,|\, some\, n_i\, in\, \mathcal{I}] \ . \end{aligned} \tag{3}$$

Figure 1 reveals that the value of $r$ is only a scaling factor; the area of $\mathcal{I}$, $A(\mathcal{I})$, is proportional to $r^2$, the proportional factor being a function of $m$ and $\epsilon$ only. This can also be checked by using the circle-circle intersection area formula, that we can for example find in [32].
Let $A(\mathcal{I}) = C(m, \epsilon)r^2$.
As $g(n, r)$ is by hypothesis an a.a.s. connected graph, we know by Theorem 4.1 that there exists a function $k(n)$ such that $\pi r^2 n = \ln n + k(n)$ and $\lim_{n\to+\infty} k(n) = +\infty$.
Thus $A(\mathcal{I}) = C(m, \epsilon)r^2 = C(m, \epsilon)\frac{\ln n + k(n)}{\pi n}$
Lemma 4.3 $\Rightarrow$

$$\begin{aligned} &\lim_{n\to+\infty}\ P[no\ node\ in\ \mathcal{I}] \\ &\quad \leq \lim_{n\to+\infty} \exp(-nA(\mathcal{I})) \\ &\quad \leq \lim_{n\to+\infty} \exp(-\tfrac{C(m,\epsilon)}{\pi}(\ln n + k(n))) = 0 \end{aligned}$$

Hence, whatever the value of $m$ and $\epsilon$,

$$lim_{n\to\infty}\, P[some\, n_i\, in\, \mathcal{I}] = 1 \ . \tag{4}$$

Moreover, as the Euclidean distance between $n_1$ and $n_i$ is strictly less than $mr$, by inductive hypothesis we have:

$$lim_{n\to\infty}\, P[path\_len(n_1, n_i) \leq m \,|\, some\, n_i\, in\, \mathcal{I}] = 1 \ . \tag{5}$$

Equations 3, 4 and 5 $\Rightarrow$

$$lim_{n\to\infty}\, P[path\_len(n_1, n_2) \leq m+1] = 1$$

The property is thus verified for $m + 1$.
By induction, the property is valid for any integer.

■

Using this theorem, we can derive the main result of this section.

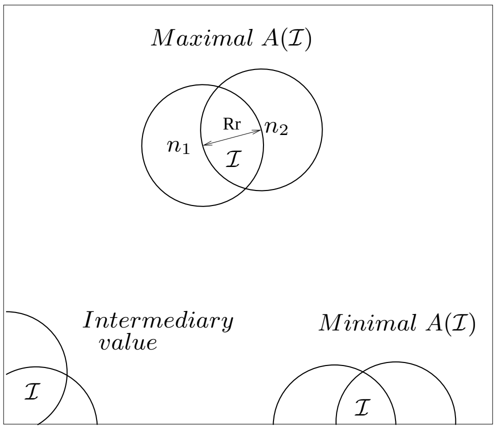


Figure 2:

**Theorem 4.5** *(Sufficient condition for the asymptotic connectivity of dense overlay graphs)*
*Consider an overlay graph* $G(g, D(n), R(n))$*. Assume* $g(n, r(n))$ *is a.a.s. connected and* $\lim_{n\to+\infty} Dn = +\infty$*. If*

$$\pi(Rr)^2 = \frac{\ln(\lceil Dn \rceil) + K(n)}{\lceil Dn \rceil} \tag{6}$$

*with* $\lim_{n\to+\infty} K(n) = +\infty$ *then* $G$ *is a.a.s. connected.*

*Proof:* Let us build a graph $g\prime(\lceil Dn\rceil, Rr)$ such that the vertices set of $g\prime$ and $G$ are the same.
Consider an edge of $g\prime$ linking two nodes $n_1$ and $n_2$.
By definition, the distance between $n_1$ and $n_2$, denoted $|n_1n_2|$, is less than or equal to $Rr$.
Let us first assume that $|n_1n_2| < Rr$. By Theorem 4.4, as $g$ is a.a.s. connected, the maximum number of hops between $n_1$ and $n_2$ is a.a.s. less than or equal to $R$. Hence, asymptotically, any edge of $g\prime$ of length strictly less than $Rr$ also exists in $G$.
Let us now assume that $|n_1n_2| = Rr$. We can draw two disks of radius $Rr$ respectively centered on $n_1$ and $n_2$. Let $\mathcal{I}$ denote the disks intersection and $A(\mathcal{I})$ its area. As shown in Fig. 2, $A(\mathcal{I})$ is minimal when $n_1$ and $n_2$ are both located on a border of the field.

Using the circle-circle intersection area formula [32], we obtain $A(\mathcal{I}) \geq C(Rr)^2$ with $C = \frac{1}{12}(4\pi - 3\sqrt{3})$, wherever $n_1$ and $n_2$ are located. By lemma 4.3, the probability that there is no overlay node in $\mathcal{I}$ is less than or equal to $exp(-C(\lceil Dn\rceil - 2)(rR)^2)$.

Assume that $\pi(Rr)^2 = \frac{\ln(\lceil Dn\rceil)+K(n)}{\lceil Dn\rceil}$ with $\lim_{n\to+\infty} K(n) = +\infty$.

Asymptotically, there exists almost surely an intermediary overlay node $n_i \in \mathcal{I}$.

The distances $|n_1n_i|$ and $|n_in_2|$ are strictly less than $Rr$ thus, by Theorem 4.4 with $m = R$, there exists a.a.s. two edges $(n_1, n_i)$ and $(n_i, n_2)$ in $G$.

Thus, for any edge $(n_1, n_2)$ of $g\prime$, there exists a path between the corresponding nodes in $G$.

As their vertices sets are the same, the asymptotic connectivity probability of $G$ is greater than or equal to the asymptotic connectivity probability of $g\prime$.

We assumed that $\pi(Rr)^2 = \frac{\ln(\lceil Dn \rceil)+K(n)}{\lceil Dn \rceil}$ with $\lim_{n\to+\infty} K(n) = +\infty$. Consequently, by Theorem 4.1, $g\prime$ is a.a.s. connected.

Thus $G$ is also a.a.s. connected. ■

## 4.4 Discussion

The first observation we can make about Theorems 4.2 and 4.5 is that, as for basic graphs, the necessary condition for the asymptotic connectivity of overlay graphs is also sufficient.

The following corollaries are meant to give an insight about the relationship between the neighbourhood range and the overlay density. For both of them, we consider an overlay graph $G(g, D, R)$ and make the assumptions that $g$ is a.a.s. connected and that $\lim_{n\to+\infty} Dn = +\infty$.

As we will extensively use the notations for the asymptotic behaviour of functions in the following, we recall them in appendix 7.1.

**Corollary 4.6** *If $DR^2 \geq 1$ then $G$ is a.a.s. connected.*

*Proof:* If basic graph $g$ is a.a.s. connected then there exists a funtion $k(n)$ such that $\pi r^2 n = \ln n + k(n)$ and that $\lim_{n\to+\infty} k(n) = +\infty$.

Thus $DR^2\pi r^2 n \geq \pi r^2 n = \ln n + k(n) \geq \ln(Dn) + k(n)$. By Theorem 4.5, the overlay graph is a.a.s. connected. ■

The sufficient condition $R > \frac{1}{\sqrt{D}}$ shows that a decreasing overlay density does not necessarily make the overlay graph a.a.s. disconnected. We can for example have $D = \frac{1}{\ln n}$ and $R = \sqrt{\ln n}$. It also confirms the intuitive idea that the lower $D$ is, the larger $R$ must be.

The advantage of the previous corollary is that we do not need any information about the basic graph, except that it is a.a.s. connected. However, lower values for the neighbourhood range could be obtained if the relationship existing between $n$ and $r$ is known.

**Corollary 4.7** *Let $\pi r^2 n = \ln n + k(n)$ with $k(n) \gg 1$. Assume $D$ is constant and $R$ is an integer with $R \geq 1$.*

1. *If $k(n) \gg \ln n$ then $G$ is a.a.s. connected for any $R$.*
2. *If $k(n) \geq a \ln n$ with $a > 0$ then $G$ is a.a.s connected for any $R > \frac{1}{\sqrt{D(1+a)}}$.*
3. *If $k(n) \ll \ln n$, $G$ is a.a.s. connected if and only if $R \geq \frac{1}{\sqrt{D}}$.*

*Proof:* Let $K(n) = \pi r^2 n DR^2 - ln(Dn)$. $G$ is connected if and only if $K(n) \gg 1$.

1. By definition, if $k(n) \gg \ln n$ then for every $M > 0$ and $n$ sufficiently large, $k(n) \geq M \ln n$. By hypothesis, $R \geq 1$ thus $K(n) \geq [D(1+M) - 1]\ln n - \ln D \geq [D(1+M) - 1]\ln n$. Let $M = \frac{1}{D}$. For $n$ sufficiently large, $K(n) \geq D \ln n$. By definition, $D > 0$ thus $K(n) \gg 1$.
2. If there exists $a > 0$ such that $k(n) \geq a \ln n$, then $K(n) \geq [DR^2(1+a) - 1]\ln n$. If $R > \frac{1}{\sqrt{D(1+a)}}$ then $K(n) \gg 1$.

3. For $k(n) \ll \ln n$, we know by Corollary 4.6 that $DR^2 \geq 1$ assures the asymptotic connectivity of $G$. Assume $DR^2 < 1$. By definition, if $k(n) \ll \ln n$ then for any $\epsilon > 0$ and $n$ sufficiently large, $k(n) < \epsilon \ln n$. This gives $K(n) < [(1+\epsilon)DR^2 - 1] \ln n - \ln D$. Let $\epsilon = \frac{1-DR^2}{DR^2}$. By hypothesis, $DR^2 < 1$ thus $\epsilon > 0$. For $DR^2 < 1$, there exists $\epsilon > 0$ such that for $n$ sufficiently large, we have $K(n) < -\ln D \Rightarrow \lim_{n\to+\infty} K(n) \neq +\infty$.

■

Concerning a basic graph, a function $k(n)$ that grows quickly just accelerates the convergence of the connectivity probability [25]. This function has a stronger impact on the neighbourhood range needed for connectivity. For example, for a constant overlay density $D$, it decides if $R$ can take any value or must be greater than a fixed threshold.

In particular, if the transmission range $r$ is kept constant while the number of nodes grows, we have $k(n) \gg \ln n$ which implies that $R = 1$ is sufficient to obtain an a.a.s. connected overlay. The overlay nodes do not need other intermediary nodes to forward their packet for communicating. The subnetwork composed of the overlay nodes only is a.a.s. connected. In fact, there is no need for building an overlay in this case. The overlay nodes can directly use their own routing protocol, with customized packet format.

Oppositely, if a topology control protocol is used for optimizing the transmission range, $R = 1$ can be too small to make the overlay a.a.s. connected. In this case, the subnetwork composed of the overlay nodes only is a.a.s. disconnected. It is necessary for some overlay nodes to communicate through intermediary non overlay nodes. Overlay techniques are required; the overlay nodes control and data packets must be encapsulated in packets that can be routed by all nodes.

# 5 Sparse Networks

## 5.1 Known Results on Basic Graphs

Combining results from [25] and [27], we can state the following conditions on the connectivity of a basic graph $g(n, r, \ell)$. Refer to Sect. 3 for notations.

**Theorem 5.1** *(Necessary condition for the asymptotic connectivity of sparse basic graphs)*
*Let $r$ be strictly less than $\sqrt{2}\ell$. If $r^2 n = O(\ell^2)$ then $g(n, r, \ell)$ is not a.a.s. connected. If $r = O(\ell^\epsilon f(\ell))$ with $0 \leq \epsilon < 1$ and $f(\ell)$ a function that grows strictly slower than any function of type $\ell^\gamma$ where $\gamma > 0$ and if $r^2 n < \frac{1}{2}(1 - \epsilon)\ell^2 \ln \ell$ then $g(n, r, \ell)$ is a.a.s. not connected.*

**Theorem 5.2** *(Sufficient condition for the asymptotic connectivity of sparse basic graphs)*
*If $r \geq \sqrt{2}\ell$, then $g(n, r, \ell)$ is a.a.s. connected. If $r = \Omega(\ell)$ and $r^2 n = \Omega(\ell^2 \ln \ell)$, then $g(n, r, \ell)$ is a.a.s. connected. If $r = \Omega(\ell^\epsilon f(\ell))$ with $0 \leq \epsilon < 1$ and $f(\ell)$ a function that grows strictly slower than any function of type $\ell^\gamma$ where $\gamma > 0$ and if $r^2 n \geq 4(1 - \epsilon)\ell^2 \ln \ell$ then $g(n, r, \ell)$ is a.a.s. connected.*

## 5.2 Minimal Neighbourhood Range

**Theorem 5.3** *(Necessary condition for the asymptotic connectivity of sparse overlay graphs)*
*Let $Rr$ be strictly less than $\sqrt{2}\ell$. If $(Rr)^2\lceil Dn\rceil = O(\ell^2)$ then $G(g, D, R)$ is not a.a.s. connected. If $Rr = O(\ell^\epsilon f(\ell))$ with $f(\ell)$ a function that grows strictly slower than any function of type $\ell^\gamma$ where $\gamma > 0$ and if $(Rr)^2\lceil Dn\rceil < \frac{1}{2}(1-\epsilon)\ell^2 \ln \ell$ then $G(g, D, R)$ is a.a.s. not connected.*

*Proof:* As for dense networks, if a graph $g\prime(\lceil Dn\rceil, Rr, \ell)$ is not connected a.a.s., $G(n, r, \ell, D, R)$ cannot be connected a.a.s. Applying Theorem 5.1 to a graph $g\prime(\lceil Dn\rceil, Rr, \ell)$, we obtain the above conditions on the asymptotic connectivity of $G$. ■

## 5.3 Sufficient Neighbourhood Range

We use the same techniques as for dense graphs, the only difference being that the probability for a node to be located on a surface $S$ equals $\frac{S}{\ell^2}$, instead of $S$.

We first demonstrate that Theorem 4.4, called the asymptotic path length theorem and stated in Sect. 4, still holds for sparse graphs.

*Proof: (the asymptotic path length theorem in the context of sparse graphs)* If $r \geq \sqrt{2}\ell$, every overlay node can reach any other overlay node in one hop. Assume $r < \sqrt{2}\ell$ and let $P_S$ denote the probability that there is no node on a surface $S = c\pi r^2$. For any constant $c > 0$, by lemma 4.3, $\lim_{\ell\to+\infty} P_S \leq \lim_{\ell\to+\infty} \exp(-n\frac{c\pi r^2}{\ell^2})$. Graph $g$ is a.a.s. connected, thus, by Theorem 5.1, $r^2 n \gg \ell^2$. This implies that $\lim_{\ell\to+\infty} P_S = 0$ and, using the same technique as for dense graphs (see proof of Theorem 4.4 in Sect. 4), we can demonstrate that if $|n_1 n_2| < mr$ then there exists a path between $n_1$ and $n_2$ composed of at most $m$ hops. ■

Using this theorem, we can derive the main result of this section.

**Theorem 5.4** *(Sufficient condition for the asymptotic connectivity of sparse overlay graphs)*
*Let $g(n, r, \ell)$ be a.a.s. connected. If $r \geq \sqrt{2}\ell$, then $G(g, D, R)$ is a.a.s. connected. If $Rr = \Omega(\ell)$ and $(Rr)^2\lceil Dn\rceil = \Omega(\ell^2 \ln \ell)$, then $G(g, D, R)$ is a.a.s. connected. If $Rr = \Omega(\ell^\epsilon f(\ell))$ with $0 \leq \epsilon < 1$ and $f(\ell)$ a function that grows strictly slower than any function of type $\ell^\gamma$ where $\gamma > 0$ and if $(Rr)^2\lceil Dn\rceil \geq 4(1-\epsilon)\ell^2 \ln \ell$ then $G(g, D, R)$ is a.a.s. connected.*

*Proof:* If $r \geq \sqrt{2}\ell$, every overlay node can reach any other overlay node in one hop and $G$ is connected whatever parameters $D$ and $R$ are. The hypotheses imply, by Theorem 5.2, that a graph $g\prime(\lceil Dn\rceil, rR, \ell)$ is a.a.s. connected.

As for dense graphs, exploiting Theorem 4.4, we can prove that if a graph $g\prime(\lceil Dn\rceil, rR, \ell)$ is a.a.s. connected then $G$ is a.a.s. connected.

Consequently, the hypotheses imply that $G$ is a.a.s. connected. ■

## 5.4 Discussion

For sparse networks, no condition for the asymptotic connectivity has been demonstrated to be both necessary and sufficient. Note however that the bounds for basic graphs given by Theorems 5.1 and 5.2 are asymptotically tight and that they have remained close for overlay graphs.

As for dense graphs, values for the neighbourhood range can be obtained if the relationship existing between $n$ and $r$ is known. For example, if $r = \sqrt{a\ell \ln \ell}$ with $a > 0$ and $n = 2\ell$, a sufficient condition for the overlay graph to be a.a.s. connected is: $R \geq 1$ and $(Rr)^2 Dn \geq 4(1 - \frac{1}{2})\ell^2 \ln \ell$, which is fulfilled if $R \geq \lceil \frac{1}{\sqrt{aD}} \rceil$.

**Corollary 5.5** *Let $g(n, r, \ell)$ be an a.a.s. connected graph. Assume $D$ is constant. If $r^2 n \gg \ell^2 \ln \ell$ then $G$ is a.a.s. connected for any $R \geq 1$.*

*Proof:* If $D$ is constant, $R \geq 1$ and $r^2 n \gg \ell^2 \ln \ell$ then $(Rr)^2 Dn \gg \ell^2 \ln \ell$ and, by Theorem 5.4, G is a.a.s. connected. ■

In particular, if the node density is kept constant while the field length grows, and if the transmission range is such that $r \gg \sqrt{\ln \ell}$, $R = 1$ is sufficient to obtain an a.a.s. connected overlay. As explained in Sect. 4.4, there is no need for building an overlay in this case.

Oppositely, if a topology control protocol is used for optimizing the transmission range, the basic graph can be a.a.s. connected with a transmission range only proportional to $\sqrt{\ln \ell}$, while the subnet composed of the overlay nodes only can be a.a.s. disconnected. As also explained in Sect. 4.4, overlay techniques are then required.

# 6 Conclusions

We first motivated the study of overlays built over ad hoc networks.

We then presented and analyzed the critical neighbourhood range problem.

We demonstrated that in connected networks, as the network gets denser ($n \to +\infty$) or larger ($\ell \to +\infty$), the shortest path between any pair of nodes draws close to the straight line. This sets an upper bound on the number of hops between any pair of nodes, knowing the distance between them and the nodes transmission range $r$.

Thanks to this property, that we called the asymptotic path length theorem, and known works on the critical transmission range problem, we were able to derive an analytical solution to the critical neighbourhood range problem for dense networks. For a large class of sparse networks, we were able to determine asymptotically tight bounds.

The mathematical conditions obtained do take into account the potential use of a homogeneous topology control algorithm and allow the overlay density $D$ to evolve with the network size ($n$ or $\ell$). In particular, if $D$ diminishes, they show how a compensation in $R$ can keep the overlay still connected.

The analysis of these results provides, among others, the following properties for overlays built on ad hoc networks.

Whatever the characteristics of the underlying network are, an overlay built on a dense connected network with $DR^2 \geq 1$ is asymptotically almost surely connected. We conjecture that this still holds in sparse networks.

In many cases, if the relationship between $n$ and $r$ is known, one can set $R$ to a lower value than $\lceil \frac{1}{\sqrt{D}} \rceil$ and still obtain asymptotic overlay connectivity.

For constant $D$, depending on the network degree of connectivity, the minimal value of $R$ for asymptotic overlay connectivity can either be equal to one, or to a higher fixed threshold, or be an unbounded function of the network size.

In particular, in dense networks, if $D$ and $r$ are kept constant while the number of nodes increases, the overlay nodes can asymptotically use their own routing protocol, bypassing the network routing protocol common to all nodes. This is also the case in sparse networks if the node density is kept constant and $r$ increases with the field length $\ell$ so that $r \gg \sqrt{\ln \ell}$.

Oppositely, if the transmission range value is optimized, using a topology control protocol, the network composed only of the overlay nodes can be asymptotically disconnected.

## Acknowledgements

This work has been partially supported by the Belgian Science Policy in the framework of the IAP program (MOTION P5/11 project) and by the European Commission (FP6 IST E-NEXT NoE).

## References

[1] H. Wu, C. Qiao, S. De, and O. K. Tonguz, "Integrated cellular and ad hoc relaying systems: icar," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 2105–2115, 2001.

[2] A. Zadeh, B. Jabbari, R. Pickholtz, and B. Vojcic, "Self-organizing packet radio ad hoc networks with overlay (soprano)," *IEEE Commun. Mag.*, vol. 40, no. 6, pp. 149–157, 2002.

[3] K. Chen and K. Nahrstedt, "Effective location-guided tree construction algorithms for small group multicast in manet," in *Proc. Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, New York, NY, June 2002, pp. 1180–1189.

[4] J. Xie, R. Talpade, A. Mcauley, and M. Liu, "Amroute : Ad hoc multicast routing protocol," *Mobile Networks and Applications*, vol. 7, no. 6, pp. 429 – 439, Dec. 2002.

[5] C. Gui and P. Mohapatra, "Efficient overlay multicast for mobile ad hoc networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC'03)*, Mar. 2003.

[6] L. Xiao, A. Patil, Y. Liu, L. Ni, and A.-H. Esfahanian, "Prioritized overlay multicast in mobile ad hoc environments," *IEEE Computer*, vol. 37, no. 2, pp. 67–74, Feb. 2004.

[7] M. Ge, S. Krishnamurthy, and M. Faloutsos, "Overlay multicasting for ad hoc networks," in *Proc. of the Third annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net'04)*, Bodrum, Turkey, June 2004.

[8] A. Patil, Y. Liu, L. Xiao, A.-H. Esfahanian, and L. Ni, "Solonet: Sub-optimal location-aided overlay network for manets," in *Proc. of the First IEEE Int. Conf. on Mobile Ad hoc and Sensor Systems (MASS'04)*, Fort Lauderdale, FL, Oct. 2004.

[9] [Online]. Available: http://www.ietf.org/html.charters/manet-charter.html

[10] P. Mohapatra, C. Gui, and J. Li, “Group communications in mobile ad hoc networks,” *IEEE Computer*, vol. 37, no. 2, pp. 52–59, Feb. 2004.

[11] K. H. Wang and B. Li, “Efficient and guaranteed service coverage in partitionable mobile ad-hoc networks,” in *Proc. Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM’02)*, New York, NY, June 2002.

[12] D. Tang and M. Baker, “Analysis of a local-area wireless network,” in *Proc. of the Sixth Annual International Conference on Mobile Computing and Networking (MOBICOM’00)*, Boston, MA, Aug. 2000.

[13] P. Gupta and P. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[14] E. Royer and C.-K. Toh, “A review of current routing protocols for ad hoc mobile wireless networks,” *IEEE Personal Commun. Mag.*, pp. 46–55, Apr. 1999.

[15] D. Clark, “The design philosophy of the DARPA Internet protocols,” *Computer Communication Review*, vol. 18, no. 4, pp. 106–114, Aug. 1988.

[16] B. Bollobas, *Random Graphs.* London, England: Academic Press, 1985.

[17] P. Erdos and A. Rényi, “On the evolution of random graphs,” *Hungarian Academy of Science*, vol. 5, pp. 17–61, 1960.

[18] M. Penrose, *Random Geometric Graphs.* England: Oxford University Press, 2003.

[19] M. D. Penrose, “On the $k$-connectivity for a geometric random graph,” *Random Structures and Algorithms*, vol. 15, no. 2, pp. 145–164, 1999.

[20] B. Krishnamachari, S. Wicker, and R. Bejar, “Phase transition phenomena in wireless ad-hoc networks,” in *Proc. IEEE Global Conference on Telecommunications (Globecom’01), Symposium on Ad-Hoc Wireless Networks*, San Antonio, TX, Nov. 2001.

[21] A. Goel, S. Rai, and B. Krishnamachari, “Sharp thresholds for monotone properties in random geometric graphs,” in *Proc. of the thirty-sixth annual ACM symposium on Theory of computing (STOC’04)*, Chicago, IL, June 2004, pp. 580–586.

[22] P. Gupta and P. Kumar, “Critical power for asymptotic connectivity in wireless networks,” in *Stochastic Analysis, Control, Optimization and Applications, A Volume in Honor of W. H. Fleming*, W. McEneaney, G. Yin, and Q. Zhang, Eds. Boston: Birkhäuser, 1999.

[23] X.-Y. Li, Y. Wang, P.-J. Wan, and C.-W. Yi, “Fault tolerant deployment and topology control in wireless ad hoc networks,” *Journal of Wireless Communications and Mobile Computing (WCMC)*, vol. 4, no. 1, pp. 109–125, Feb. 2004.

[24] P.-J. Wan and C.-W. Yi, "Asymptotic critical transmission ranges for connectivity in wireless ad hoc networks with Bernoulli nodes," in *Proc. of the 2005 IEEE Wireless Communications and Networking Conference (WCNC 2005)*, New Orleans, Louisiana, Mar. 2005.

[25] P. Santi and D. M. Blough, "The critical transmitting range for connectivity in sparse wireless ad hoc networks," *IEEE Trans. Mobile Comput.*, pp. 25–39, January-March 2003.

[26] V. Kolchin, B. Sevast'yanov, and V. P. Chistyakov, *Random Allocations*. Washington D.C.: V.H. Winston and Sons, 1978.

[27] S. Muthukrishnan and G. Pandurangan, "The bin-covering technique for thresholding random geometric graph properties," Rutgers and Purdue Universities, USA, Tech. Rep. DIMACS-TR-2003-39, Nov. 2003.

[28] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Commun. Lett.*, vol. 10, no. 6, pp. 437–490, Oct. 2002.

[29] V. Kawadia and P. Kumar, "Principles and protocols for power control in wireless ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 76–88, Jan. 2005.

[30] S. Narayanaswamy, V. Kawadia, R. Sreenivas, and P. R. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the compow protocol," in *Proc. of European Wireless 2002. Next Generation Wireless Networks: Technologies, Protocols, Services and Applications*, Florence, Italy, Feb. 2002, pp. 156–162.

[31] M. D. Penrose, "The longest edge of the random minimal spanning tree," *The Annals of Applied Probability*, vol. 7, no. 2, pp. 340–361, 1997.

[32] E. W. Weisstein, *Circle-Circle Intersection*, From MathWorld–A Wolfram Web Resource, 1999. [Online]. Available: http://mathworld.wolfram.com/Circle-CircleIntersection.html

# 7 Appendix

## 7.1 Mathematical Notations

Let $f$ and $g$ be functions of the same parameter $x$.

1. $f(x) = O(g(x))$ iff
   $\exists\, x_0, C > 0 :\ x \geq x_0 \Rightarrow |f(x)| \leq Cg(x)$
2. $f(x) = \Omega(g(x))$ iff $g(x) = O(f(x))$
3. $f(x) = \Theta(g(x))$ iff
   $f(x) = O(g(x))$ and $g(x) = O(f(x))$
4. $f(x) \ll g(x)$ iff $lim_{x\to+\infty} \frac{f(x)}{g(x)} = 0$
5. $f(x) \gg g(x)$ iff $g(x) \ll f(x)$

## 7.2 Proof of Lemma 4.3

*Proof:* [Lemma 4.3] We start with the uniform distribution.

If $n$ nodes are distributed uniformly and independently on the unitary square, then the probability that a node lies on a surface $S \leq 1$ equals $S$.

Let $X_S$ be a random variable designating the number of nodes on a surface $S$ with $0 \leq S \leq 1$.

$$P[X_S = k] = S^k (1-S)^{(n-k)}$$

Thus the probability that there is no node on $S$ is

$$\begin{aligned} P[X_S = 0] &= (1-S)^n \\ &= \exp(n \ln(1-S)) \end{aligned}$$

with

$$\ln(1-x) = -x(1 + \frac{x}{2} + \frac{x^2}{3} + \frac{x^3}{4} + ...)$$

$$\Rightarrow P[X_S = 0] = \exp(-nS(1 + O(S))) \leq \exp(-nS)$$

Let us now focus on the Poisson distribution The spatial Poisson point process has mean $n$.

The probability of having $k$ nodes on a surface $S$ is

$$P(X_S = k) = \frac{(nS)^k}{k!} \exp(-nS)$$

$$\Rightarrow P(X_S = 0) = \exp(-nS)$$

Thus, for both distributions

$$P[X_S = 0] \leq \exp(-nS) \tag{7}$$

■