Basic Analysis and Simulation of Ad-Hoc Networks

Technical Report 1

10-Dec-2001

Horst Hellbrück	Stefan Fischer
International University	TU Braunschweig
School of IT	IBR
Campus 2	Mühlenpfordtstr. 23
D-76646 Bruchsal	D-38106 Braunschweig
horst.hellbrueck@i-u.de	stefan.fischer@tu-bs.de

Abstract: In order to determine whether the deployment of Ad-Hoc networks in a certain region or field of application is feasible or reasonable, analytic considerations as well as simulations are helpful. In this technical report we present a simulation tool for such networks. It performs in a simple manner the investigation of different Ad-Hoc scenarios by entering some fundamental input parameters. At first we introduce the underlying basic model. In the process of the report we show that a thorough analytic approach is difficult and complex, and show the limitation of the analysis. Therefore we introduce our simulation tool.

1 Introduction and Background

Recent years have seen a tremendous growth in the usage of devices equipped with communication interfaces for Wireless LAN, Bluetooth or GSM. Hence the possibility of building Ad-Hoc Networks is within reach. The idea consists of forming a network of mobile stations in proximity to each other, where some of the devices act as relay stations. The desired result is that two stations, which are not in direct range, can communicate with each other, by forming longer routes using available intermediate stations.

However the question arises, when will it pay off to try communication over such networks? It is obvious that there is little sense to form a long-lasting communication connection if, on the one hand, the probability for the successful set-up is very small, and on the other hand, established connections were constantly interrupted by the mobility of the intermediate stations. Pre-estimations can be very useful for an investigation of the use of Ad-Hoc networks for certain applications in a specific geographical surrounding field.

For this purpose we developed the simulation tool „ANSim “(**A**d-hoc **N**etwork

Simulation), that will be introduced in more detail in this document. It serves as a tool for statistical simulation for practice-oriented Ad-Hoc scenarios. The user can determine the boundary conditions of the simulation by the input of some base parameters. These include size and shape of the geographical area, range of the stations, etc. The user receives as result e.g. the probability that two selected stations are connected.

The rest of the paper is organised as follows: Section 2 presents the basic communication model used for the further investigations. Before the main part, which covers the simulation of Ad-Hoc networks, Section 3 introduces some analytic results of Ad-Hoc networks. It will turn out that a correct analysis represents a very complex and not practically solvable task. For this reason we shift to simulations for further investigations. Section 4 presents the current version of the simulation tool developed at the International University, and illustrates some typical simulation scenarios. Section 5 presents first basic results including interpretations. This paper concludes with a conclusion and future extensions (Section 6).

2 Transmission Model for Analysis and Simulation

For simplification we ignore some aspects of the Ad-Hoc Networks in the first step, such as node movement and path determination. We limit ourselves to a simple model in the geometry of two dimensions $\Re^2$ like illustrated in Figure 1.

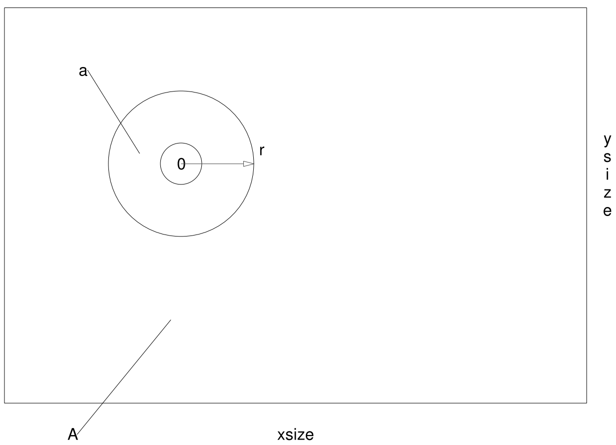

Figure 1: Illustration of the basic communication model

We select a rectangular area A as an illustration for possible positions of mobile nodes. Within this area N nodes are located. The nodes are numbered N_0 to

N_{N-1}. The transmission range is indicated as a circle with radius r. With the parameters given in the figure, the areas are calculated as follows: $a = \pi r^2$ and $A = xsize \cdot ysize$.

Edge effects, such as shading (by obstacles), Reflection (at big surfaces), dispersion (at small surfaces) and diffraction (at sharp edges), are neglected and only the free-space loss is taken into account. This model, as also used in [3], enforces a simple digital decision:

- if the distance d between two nodes is smaller than the transmission radius r, communication is possible.
- no transmitting can take place outside the transmission range.

We neglect mobility and uniformly distribute nodes with digital transmitting characteristics within the area A. The positions $P_A = (x_a; y_a)$ and P_B of two arbitrary nodes N_A and N_B are uncorrelated and independent of each other, i.e., $P_B \neq f(P_A)$.

3 Analytic Investigations and Approximations

We first perform calculations with the above model. We distribute 2 nodes N_0 and N_1 in the area A.

The probability p_v that the two nodes are in connection is equal to the probability that the point $P_1 = (x_1; y_1)$ is located in the area $a(x_0, y_0)$ spanned by N_0: $p_v = p(P_1 \in a(x_0, y_0))$. The node takes each position (is statistically uniformly distributed). The probability that it is located in any infinitesimal area is calculated as $\frac{dx \cdot dy}{A}$. We compute the probability $p_v(x_0, y_0)$ for a certain position of N_0 by integrating (summation) over the area A.

$$p_v(x_0, y_0) = \iint\limits_A \frac{(P_1 \in a(x_0, y_0)}{A} \, dxdy = \frac{a'(x_0, y_0)}{A} \tag{1}$$

where $a'(x_0, y_0) = a(x_0, y_0) \cap A$ (see Figure 2).

Since N_0 is uniformly and independently distributed within the area, we can describe the probability p_v finally by an area integral over all $p_v(x, y)$.

$$p_v = \iint\limits_A \frac{p_v(x, y)}{A} \, dxdy = \frac{1}{A} \iint\limits_A \frac{a'(x, y)}{A} \, dxdy \tag{2}$$

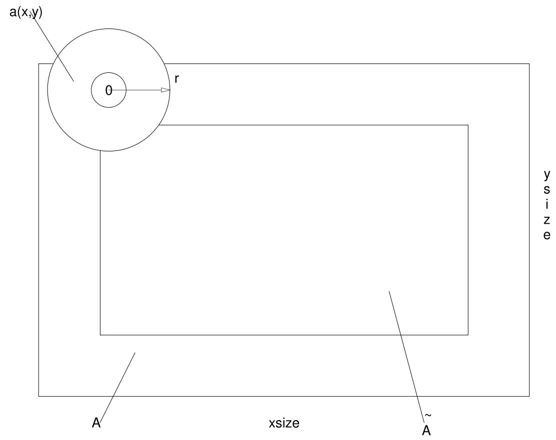

Figure 2: varying size of $a'(x, y)$ by intersection of $a(x, y)$ and A

A closed solution of the integral is only possible when we provide the shape and size of the area. The quantity of the area $a'(x, y)$ depends on x and y, as a will be cut off by the intersection with A at the edges. With the prerequisite of a rectangular surface and $a \ll A$, the integral in Equation 2 can be divided into two areas (see Figure 2). The first part integrates over the core surface $\tilde{A}$, in which node N_0 is situated in such a way that the circle is not cut at all. The second part is the boundary region, within which the circle is cut at the edges by the intersection with A.

$$p_v = \frac{1}{A^2} \left(\iint\limits_{\tilde{A}} a'(x, y)\, dxdy + \iint\limits_{A \backslash \tilde{A}} a'(x, y)\, dxdy \right) \tag{3}$$

In the first part of the integral the size of the area $a'(x, y)$ is equal to a. In the second part of the integral $a'(x, y)$ varies between $\frac{1}{4}a$(corner of A) and a (corner of $\tilde{A}$) depending on the position of N_0. More precise investigation, which exceeds the framework of this article, shows that $\frac{3}{4}a$ is a good approximation of the area if the core area $\tilde{A}$ forms a single closed area, i.e. the area A does not degenerate. This results in the approximation

$$p_v \approx \frac{a}{A^2} \left(\frac{3}{4}A + \frac{1}{4}\tilde{A} \right) \tag{4}$$

where:

$$a \ll A, \qquad e.g. \qquad r < \frac{1}{2}xsize \qquad and \qquad r\frac{1}{2}ysize. \tag{5}$$

An accurate solution of this problem can be always achieved by restriction of the generality (specification of further details concerning the area A). As an example for the use of Equation 4 we select one of the standard Ad-Hoc scenarios of the project MANET [5], which is used e.g. in [4] and [2]. Here nodes with a range of 250m are distributed within an area of 1000m × 1000m. The probability that two arbitrary placed nodes are located in direct connection to each other, results in

$$p_v \approx \frac{\pi r^2}{(xsize \cdot ysize)^2}\left(\frac{3}{4}xsize \cdot ysize + \frac{1}{4}(xsize - 2r) \cdot (ysize - 2r)\right) \tag{6}$$

$$= \frac{\pi 62500\mathrm{m}^2}{(1 \cdot 10^{12}\mathrm{m}^4}\left(0,75 \cdot 10^6\mathrm{m}^2 + 0,25 \cdot 250 \cdot 10^3\mathrm{m}^2\right) \tag{7}$$

$$= 0,159 \tag{8}$$

Simulations performed with the tool ANSim presented in Section 4 confirm the above approximation. Simulations with identical input parameters provide $p_v = 0,157$. As already stated in the introduction of this paper, the principle of Ad-Hoc networks is that intermediate nodes operate as relay stations for nodes which are not in direct contact to each other. Therefore especially scenarios with the condition $a \leq \frac{1}{4}A$ are of great interest. The approximation of Equation 4 results in an accuracy of one-digit percent range compared to ANSim in these scenarios, which is more than suitable for our further considerations.

We insert now a third node acting as a relay station into our calculation scenario. The additional node can support the connection of two nodes (Hop by Hop). The connection probability p_{02} between the nodes N_0 and N_2 as calculated in Equation 2 for direct connection increases by p_{012} when introducing the intermediate node.

$$p_{ges} = p_{02} + p_{012} \tag{9}$$

The proportion p_{012} can be derived in accordance with the sketch in Figure 3.

We neglect all edge effects, which occur at the boundaries of the area A for simple calculation. To reach connectivity, N_1 must be located in the circle spanned by node N_0. The location of node N_2 is restricted to the grey highlighted area $a'(x_1, y_1)$. Other positions lead either to a direct connection of node N_0 and N_2 or to the nodes not being connected at all. The maximum size of the grey area is obtained when node N_1 is located straight on the edge of the circle stretched by node N_0. The area can be computed as the subtraction of the entire circular area a from the two segments as_0 plus as_1. The two segments are of equivalent size, since both

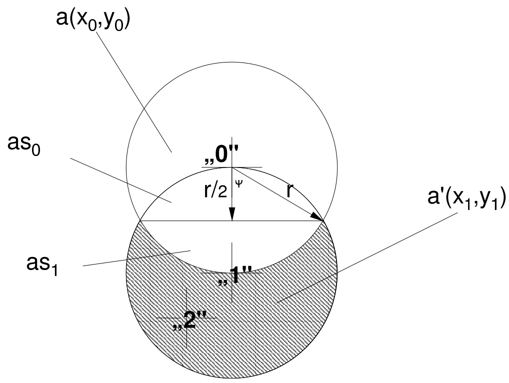

Figure 3: Hop-by-Hop-Connection including details for calculation of connection probability

circles have the same radius r. The area of the segments is calculated according to [6] with $\varphi = 2 \cdot \sin^{-1}(0,5r/r) = 120°$ (see isosceles triangle in Figure 3) as,

$$\begin{aligned} as_0 + as_1 &= 2\frac{r^2}{2}\left(\frac{\pi\varphi}{180°} - \sin\varphi\right) && (10) \\ &= r^2\left(\frac{2\pi}{3} - 0,866\right) && (11) \\ &\approx \frac{2}{5}\pi r^2 && (12) \\ &= \frac{2}{5}a && (13) \end{aligned}$$

The maximum size of $a'(x_1, y_1)$ is $\approx \frac{3}{5}a$. The minimum size of $a'(x_1, y_1)$ becomes 0 if the circles overlap completely. P_1 is uniformly distributed in A, hence obviously also within $a(x_0, y_0)$. As a result of numeric integration over all positions of N_1, the average value of the area is as $\approx \frac{2}{5}a$, i.e. the circular areas overlap on average by 60%. The fraction of the probability p_{012} that node 0 is connected to node 2 via node 1, becomes

$$p_{012} = p_{01} \cdot p_{12} \approx \frac{2}{5}p_v^2 \qquad (14)$$

whith p_v according to Equation 4. We insert Equation 14 into Equation 9 to get the total probability for a connection of node 0 with node 2.

$$p_{ges} = p_{02} + p_{012} \approx p_v \left(1 + \frac{2}{5} p_v\right) \tag{15}$$

The constraints for the above formula are already given in Equation 5. For the illustration of the proportion p_{012} of one intermediate node in Equation 9 we refer again to the MANET standard scenario in the beginning of this section. The value of p_{012} is $0,4 \cdot 0,157^2 = 0,0098 = 6,28\% p_v$. One additional node, as an intermediate node, increases the absolute value of the connectivity by approximately 1%. We conclude that in order to reach connectivity above 90% there is a need for a large number of relay stations,. The analytic results achieved so far can be extended to 4 and more nodes without any further detailed explanation.

$$\begin{aligned} p_{ges,N=4} &= p_{03} + p_{013} + p_{023} + p_{0123} + p_{0213} \\ &\approx p_v + 2 \cdot 0,4 \cdot p_v^2 \cdot (1 - p_v^2) + 2 \cdot 0,4^2 \cdot p_v^3 \end{aligned}$$

$$\begin{aligned} p_{ges,N=5} &= p_{04} + p_{014} + p_{024} + p_{034} + p_{0124} + \ldots + p_{01234} + \ldots \\ &\approx p_v + 3 \cdot 0,4 \cdot p_v^2 \cdot (1 - 2 \cdot p_v^2) + 6 \cdot 0,4^2 \cdot p_v^3 \cdot (1 - p_v^3) + 6 \cdot 0,4^3 \cdot p_v^4 \end{aligned}$$

...

Using the above equations to calculate the connection probability p_{ges} by however, results in decreasing accuracy with rising number of nodes, as previously neglected edge effects occur. E.g. The contribution of the probability $p_{0123 \ldots N-1}$ that all intermediate nodes serve as relay stations becomes 0 very soon. With increasing number N the limitation of the area A does not allow such long chains. This effect and others are neglected in the above equations.

We conclude the section by stating that a closed solution of the problem of 2 nodes by means of integral calculus is very complex, and becomes solvable only under further assumption of the shape and the size of the area A by means of multiple integrals. Computations of scenarios with a high number of nodes (relay stations) quickly show the limitations of the analytic view.

4 ANSim Simulations

The limitations of the analytic considerations have been pointed out, and further investigations based on discrete simulations are now considered. There are two different approaches for simulations in general.

- **Statistical simulation.** N Nodes are distributed on a point grid via a random process (called scenarios). Two nodes are selected randomly and examined whether these are connected. With a rising number of scenarios the accuracy of the results increases.
- **Excessive simulation.** Scenarios with all possible positions of nodes are systematically analysed (comparable with the numeric integration in the previous section). This leads, e.g. with a grid of 1000×1000 (see MANET standard scenario of Section 3) to $1000000 = 1 \cdot 10^6$ positions per node. N nodes give $1 \cdot 10^{N \cdot 6}$ different scenarios. The accuracy of the results is very good, and they are very well reproducible. For a large number of nodes this approach is limited due to the "exploding" computing times even with the powerful machines available today.

Due to the expected large number of nodes we use statistical simulation exclusively. For this purpose the tool „ANSim“ (**A**d hoc **N**etwork **Sim**ulation) was developed, and implemented in the programming language Java. The graphical version uses AWT, and is available as standalone application or applet. The simulations can also be performed by command line calls in combination with shell scripts. A current version of ANSim as applet is available at http://www.i-u.de/schools/hellbrueck/ansim/start.htm.

The input parameters of the simulation are

- description of the area extension (here $xsize$ and $ysize$)
- description of the shape of the area (here rectangle or oval)
- description of the range r of a station
- number of nodes distributed within the area N
- number of scenarios which can be examined (*only in command line version*)
- distribution in the area (Gaussian distribution or uniformly distribution)
- extension of the area to a multiple of the original area, in order to avoid edge effects (*extension*-parameter)
- communication mode (peripheral - two arbitrarily placed nodes or central – one partner is always situated in the focal point of the area)

The output parameters calculated by the simulation are

- total probability p_{ges} that any two nodes are connected
- average distance $dist$ of the two examined nodes
- average number of directly attainable nodes $neighbours$
- mean value of the hops to reach the target and the appropriate variance
- individual portions of probabilities p_n that the connection consists of n hops (only in command line version).

p_{ges} is calculated statistically as the quotient of the successful scenarios with connectivity and the total number of examined scenarios.

$$p_{ges} = \frac{\{\text{Scenarios} \mid (\text{Sender \& Receiver connected})\}}{\text{Number of examined Scenarios}} \tag{16}$$

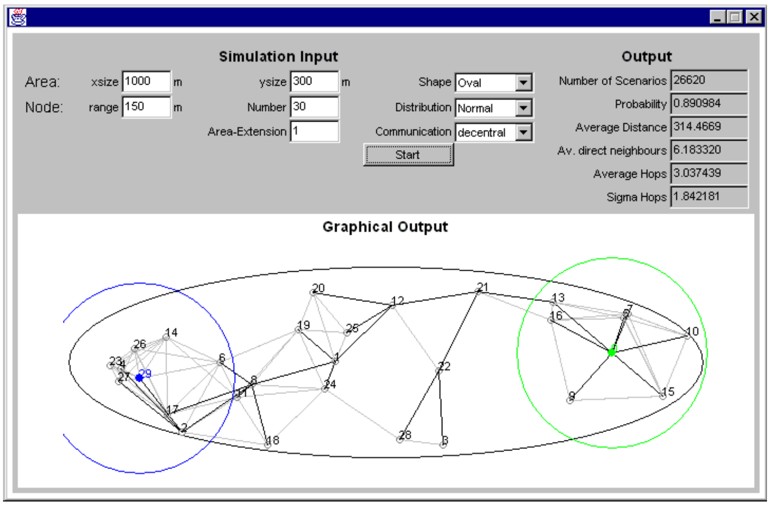

Figure 4: Graphical User Interface of the simulation tool ANSim

Figure 4 illustrates some parameters and introduces the user interface of the graphical simulation tool. On the left side one finds an input block with the simulation parameters to be entered. On the right side the textual output block displays the numeric results calculated by the simulation. Within the lower region (graphical output block) individual analysed scenarios are visualized. The figure shows a screen shot with the input parameters xsize=1000m, ysize=300m, shape=oval, range=150m, N=30, distribution=normal, extension=1, communication=decentral. The sending and receiving nodes are highlighted in the graphical output. The transmitting node N_0 (in green) tries to establish the shortest possible connection (Spanning Tree according to Dijkstra [1]) to N_{29} (in blue). The algorithm ends as soon as the receiving node is reached (connectivity achieved), or the tree cannot be further extended (no connectivity). The black lines show the spanning tree (the connections found on the way from green to blue). All the existing connections between the nodes are marked in grey. In Figure 4 the connection is established via 7 hops. The two circles of the transmitting and receiving node indicate the transmission range.

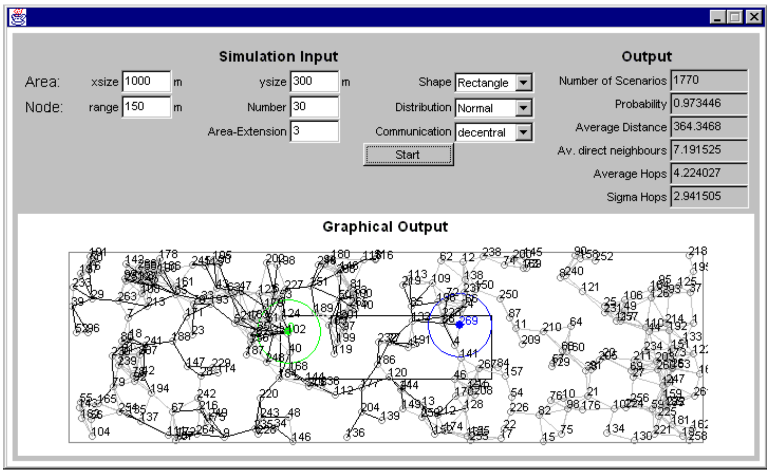

Figure 5: Impact of the extension-parameter

Figure 5 demonstrates the mode of operation of the parameter *extension* (here set to the value 3). This entry causes ANSim to fill an area of $3 \times xsize$ and $3 \times ysize$ with $3 \times 3 \times N$ nodes. Extending the area by the help of this parameter, means adjusting a node density with the parameters $xsize$, $ysize$ and N. The radius of the circle shrinks optically to a third of the original size. The transmitting and receiving nodes however are always located in the core area indicated by the black border. As one can see in the scenario, a connection can now be established between N_0 (green) and N_{269} (blue) by means of intermediate nodes outside of the core area.

In the next section we attempt to find a closed formula that describes, at least approximately, the system behaviour on the basis of some simulation results.

5 Results and Interpretations

Having introduced to the input and output parameters of the simulation in the previous section, we want to perform some simulations and investigate the results.

All following simulations are based on the MANET standard scenario presented in Section 3, where the simulation series are performed by increasing number of nodes N, beginning from 2, until full connectivity is reached ($p_{ges} = 1$).

We use the following set of parameters: $xsize$ = 1000m; $ysize$ = 1000m; $shape\ r$ = rectangular; $range\ r$ = 250m; numberOf Nodes N = <variabel>; $distribution\ d$ = normal (uniformly distributed); $extension\ e$ = 1; $communication\ c$ = decentral

The parameter $distribution$, which influences the random placement of the nodes, is kept fixed to normal (=uniformly distributed) for the rest of the section. The interpretation of the results concentrates on the „interesting“ range $p_{ges} \geq 80\%$. Thus the cases in which the connection between two any nodes is at least possible in 4 out of 5 cases. The range smaller 80% is of little interest for the operation of Ad-Hoc networks, since the relation between expenditure and benefit becomes very unfavourable.

5.1 First Simulation Series

The purpose of this test series is the investigation of the relationship between the input parameters of the area A and the connection probability p_{ges}, qualitatively and quantitatively, in order to find an approximation for p_{ges} that allows for first rough estimations. The input parameters, particularly the shape and size of A, were varied in a wide range

- from $ysize \ll r$ and $xsize > r$ (modelling of narrow streets) very narrowly
- over $ysize \approx 3 \cdot r$ and $xsize \gg r$ (modelling of wide streets, long-drawn-out places) moderately broad but long
- to $ysize, xsize \gg r$ (modelling of very large areas) large in x and y-direction compared with the range r of the nodes

As expected, dependencies of the connection probabilities on the standardized size (A/r) and shape $((U/r))$ of the area as well as the number of nodes N showed up. Deriving of the approximation is done in 2 steps:

1. change the scaling of x-axis so that all curves $p_{ges}(A, U, r, N)$ overlap approximately
2. find a mathematical description (parameters of the function) of the resulting approximation curve

After extensive empirical investigations, approximate coverage of all curves is reached by the following scaling of the x axis (see Figure 6)

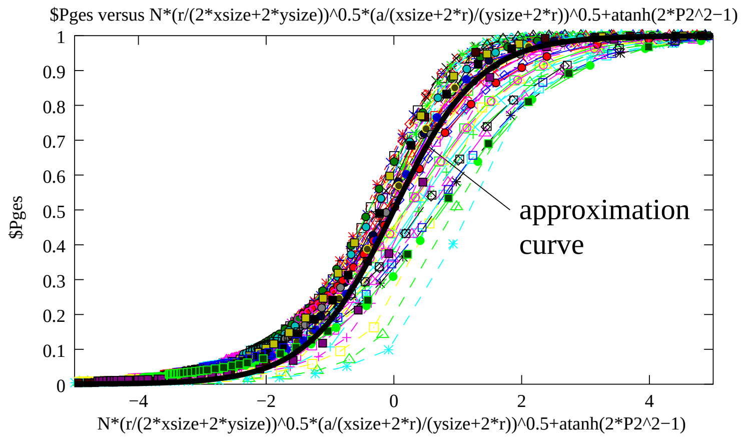

Figure 6: Crowd of simulation scenarios with approximation function

$$x = N \cdot \sqrt{\frac{r}{U}} \sqrt{\frac{a}{A_{ext}}} + atanh(2p_v^2 - 1) \qquad (17)$$

where U equals to the perimeter of A , A_{ext} equals the size of the area A extended by the radius r, and p_v according to Equation 2. The mathematical description of the resulting function (approximation curve of the connection probability p_{ges}) by the use of x is

$$p_{ges} \approx \frac{1}{2}\left(tanh(\frac{3}{4}x) + 1\right) \qquad (18)$$

Calculations based on the approximation curve, however, are only suitable for rough estimations, as the max. deviation from the simulation (as can be seen in figure 9) reaches up to 50%.

5.2 Standard MANET Scenario

At the end of this section we apply the first obtained conclusions to the standard MANET scenario. As previously mentioned, an area of 1000m × 1000m is covered by 50 nodes with a range of 250m. Equations 17 and 18 result in a connectivity of 94,3%. The simulation run shows a connection probability of 98% with an average distance between sender and recipient of approx. 520m. On average, each node has 7,7 direct neighbours. The scenario is balanced, and represents a good compromise between connection probability and expected perturbation of communication between neighbouring nodes.

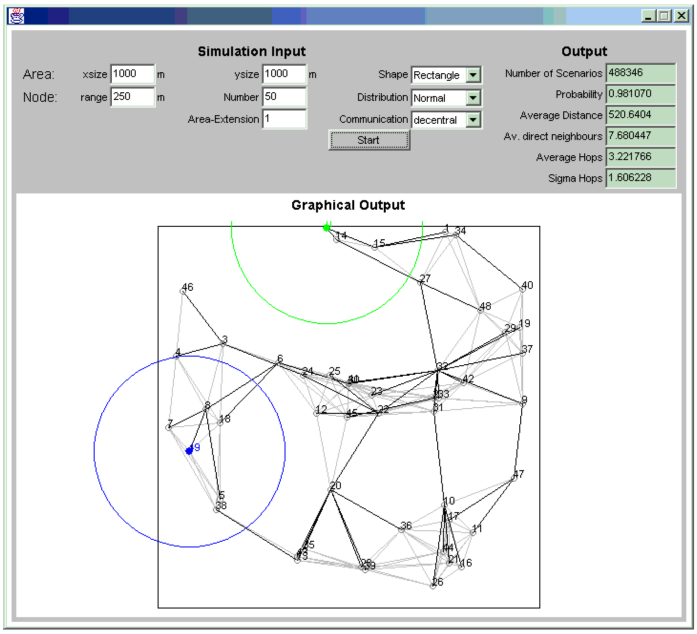

Figure 7: Results of the MANET Scenarios 50 nodes on a 1000×1000m area with a range of 250m

6 Conclusion and Future Work

Based on a simple theoretical model, we calculated connection probabilities of Ad-Hoc networks. We have proved the deduced formulas by validating them through simulations. A basic finding is the fact that the increase in connection probability is very small by bringing in intermediate nodes. An unexpected large number of nodes are necessary in order to form a completely interconnected network structure. The limitations of the analytic approach were pointed out.

Due to the too high computation times of excessive simulation a tool performing statistical simulations was developed under Java. With the help of a graphical user interface, most of the important parameters of an Ad-Hoc network can be entered. The tool calculates the statistical characteristics of Ad-Hoc networks and illustrates scenarios. The simulation tool ANSim has acknowledged all previously carried out calculations. Designers of Ad-Hoc networks as well as researchers can benefit from this tool by deriving scenarios for detailed investigations of protocols e.g. by means of *ns-2*.

The investigation on the number of neighbouring Nodes and the effects on the throughput of the network are target of further investigations, particularly in the comparison with an Ethernet LAN, in which all nodes are connected directly via a hub.

Finally, an important future project is extensions to the simulation tool as there are throughput calculations, comparison of different transmission models, which consider side effects such as shading, reflection, dispersion and diffraction, and most interesting the influence of mobility on existing connections and their simulation during a longer period. These extensions are in development.

References

[1] E. Dijkstra. A note on two problems in connection with graphs, 1959.

[2] T.D. Dyer and R.V. Boppana. A comparison of tcp performance over three routing protocols for mobile ad hoc networks. ACM Mobihoc, October 2001.

[3] P. Gupta. *Design and Performance Analysis of Wireless Networks.* PhD thesis, Coordinated Science Labatory, University of Illinois, 2000.

[4] Elizabeth M. Royer and Charles E. Perkins. Transmission range effects on aodv multicast communication. Swedish Workshop on wireless Ad-Hoc Networks, March 2001.

[5] J. Macker S. Corson. Routing protocol performance issues and evaulatin considerations. Technical Report RFC2501, MANET Working Group, January 1999.

[6] Horst Stöcker. *Taschenbuch mathematischer Formeln und moderner Verfahren.* Deutsch, Frankfurt, 1999.