Multivariate Analysis of the Cross-Layer Interaction in Wireless Networks Simulations

Jean-Michel Dricot
Computer & Network Engineering Dpt.
Université Libre de Bruxelles
Av. Franklin Roosevelt, 50
1050 Brussels,Belgium
Email: jdricot@ulb.ac.be

Philippe De Doncker
Waves & Signals Dpt.
Université Libre de Bruxelles
Av. Franklin Roosevelt, 50
1050 Brussels,Belgium

Esteban Zimànyi
Computer & Network Engineering Dpt.
Université Libre de Bruxelles
Av. Franklin Roosevelt, 50
1050 Brussels,Belgium

Abstract— **While there exist many papers that compare the performances of different routing protocols for wireless ad-hoc networks, these analysis are often realized using tools from descriptive statistics (curves drawing, means and variances computation, etc.). In this paper we propose the use of multivariate statistics to unveil and characterize the interaction between the input variables of a wireless network simulation. The ANalysis Of VAriance (ANOVA) tool helps us in getting the impact of the four variables that we have studied: routing algorithm, propagation conditions, nodes density, and mobility scheme. Using our methodology we are able to show that not only a single layer of the protocol stack can affect the network operation but also the interactions between two or more inputs of the simution.**

An important implication of the study is that the efficiency of the routing algorithm is strongly correlated with the environment (indoor/outdoor) and that the performance analysis of the lower levels of the OSI stack should be conducted by focusing on a single layer. In the same way, wireless networks simulators cannot be considered accurate if they neglect a realistic-enough implementation of the physical layer (i.e., propagation, interference and modulation)

I. Motivation and Contribution

WIRELESS ad-hoc network simulation is among one of the most hot topics of these ten last years. Since it is difficult to conduct experiments with real equipment and a large number of nodes, computer-based simulations have been used from the beginning in order to investigate and to validate new routing schemes as to conduct performance analysis on future wireless network[14] [15]. However, in [12], the authors report that the use of too simplistic models for the particular aspects of wireless networks (physical layer model, interference, modulation schemes, etc.) are more and more lowering the credibility of the studies based on simulations. In a companion paper [1] we have shown that the choice of physical layer instance (e.g., shadowing, free space, or ray-tracing) has a significant influence on the performance of the network. Moreover, it can lead to radically different conclusions on the protocol efficiency if the conlusions are subsequent to the analysis of some metrics like routing overhead or packet delivery fraction.

We conducted a simulation-based experimental analysis to characterize the cross-layer interaction. Our research has been motivated by earlier work by Barrett *et al.* [10] who has set up a statistical analysis of the intraction between routing and MAC, Takai et al. [7] and, results of Royer and al. that emphasized in [8] *"...the critical need for studying interaction between protocol layers..."*.

In this paper, we introduce the use of multivariate statistical analysis in order to reveal the existence of complex interactions between different factors of indoor wireless networks. To our best knowledge, only one paper [10] has used this statistical approach. However, this paper was more focused on the interaction between RTR, MAC and injection rate in a more general context. We present here a study on the interaction between the routing algorithm, the deployment scenario, the nodes density, and the propagation model in an indoor environment.

II. Experimental Setup

The statistical analysis using ANOVA requires a set of input variables with a discrete range of values and allows to conduct analysis on ouput variables with continuous values. Apart from the physical layer model, we use as input variables the routing algorithm, the number of nodes and the deployment scenario. Our evaluation criteria consists of the following widely-used metrics [9]:

1) the routing overhead, defined as the ratio of the number of packets spent for routing purposes on the total number of packets sent in the entire network
2) the packet delivery fraction which is the ratio of packets successfully delivered to the destination
3) the end-to-end delay which is the mean time spent for the delivery of the packets that were successfully received.

In our study, we used the NS-2, a discrete event simulator used by the networking research community [13]. It implements several routing algorithms and some of the most widely-used MAC algorihms and propagation models. As a previous study has already focused on the interaction between the MAC and RTR, we have used here the same 802.11 DCF MAC for each simulation.

1. **Deployment:** medium-sized office floor ($40m \times 40m$) in a concrete and brick building.
2. **Connections and Traffic**: 7 connections between 5 different nodes over the whole simulation. CBR traffic with an injection interval of $0.001s$. The transport agent is UDP. The connection and traffic schemes are the same for all of the simulations.
3. **Routing Protocols**: Two reactive (AODV, TORA) and one active (DSDV) routing protocols. The three levels of this discrete variable are denoted as $\{RTR_i\} = \{AODV, DSDV, TORA\}$.
4. **Nodes Density**: 10, 20 and 30 nodes were randomly deployed all over the office floor. The associated levels are denoted as $\{NODES_i\} = \{Nodes10, Nodes20, Nodes30\}$. Each scenario has a different random deployment.
5. **Scenarios**: 12 different scenarios modeling nodes deployment and mobility. For each scenario, the nodes are spread uniformely accross the 2D floor. Nodes are moving along straight lines connecting halls and office. Pause time is limited to one minute between the arrival and the departure from a point.
6. **Simulator Used**: ns-2.27 compiled, validated and running on MacOSX. The simulation time was 600 seconds of simulated operation.
7. **Propagation**: channel frequency set to 2.4 Ghz with 11 Mbp/s data rate. The other parameters, such as the carrier threshold and the signal strength are left to their default values (Lucent-class card). Three propagation models are refered as $\{PROPA_i\} = \{FreeSpace, Shadowing, Raytracing\}$.
8. **MAC Protocol**: 802.11 DCF.
9. **PHY layer**: we have written a subclass of NS-2's WirelessPhy class in order to take into account interferences and collision. This feature is not fully implemented in the current distribution of NS-2.
10. **Modulation**: BPSK. Once again we had to implement it as it is not present in the simulator.

Fig. 1. Summary of variables used in the ANOVA statistics

Propagation model

The channel model in NS-2 is quite simple. The simulator computes the received power each time a packet transmission takes place by using an user-selected propagation model. Three radio propagation models are used here: free-space model, shadowing model and ray-tracing. In the first model, the world is assumed to be flat and a direct path always exists between the trasmitter and the receiver. While this model is inherently limited, it is reported in [12] to be often used.

The shadowing model is an improved version of the free-space model. It considers the existence of direct and inderect rays between the two nodes and it introduces a crossover distance d_0 after which the reflecting rays may destructively interfere with the direct ray and drasticaly reduce the field strength. This model also adds statistical fluctuations of the signal over time by the means of a zero-mean Gaussian variable. The power is then computed as follows:

$$P_{Rx}[dBm] = P_{Tx,Friis}(d_0)[dBm] - 10.n.log(\frac{d}{d_0}) + N(0, \sigma)$$

$$P_{Tx,Friis}(d0)[Watt] = P_{Tx}[Watt].\frac{G_T.G_R.\lambda^2}{(4\pi d_0)^2.L}$$

where n is called the *path-loss exponent* and σ the *shadow variability*. Typical values for n range from 2 (free space) to 4 (indoor) and σ is often set to four. λ is the wavelength, G_T and G_R are the antennas gain, L the system loss.

Finally, the last model we use is based on the ray-tracing technique [16]: rays are drawn from the emitter to the receiver and Maxwell's laws are applied every time a reflexion on a wall, diffraction or refraction takes place. It is not present in NS-2 and has been implemented for the purpose of this study. Further details on this implementation can be found in [1].

The scenario variable

The scenario variable refers to deployment and mobility aspects. Deployment is important in that the whole network connectivity can drastically vary over time, especially indoors, with respect to the density of nodes and their repartition. In this work, 12 initial deployments were generated by using a uniformly random position over a 2D floor map. The mobility of the nodes was limited to the halls and offices: they follow a straight line between two particular points of the map (offices or hall), then sleep for maximum time of one minute and run again towards a new target point. The need for this type of real-life mobility scheme has been motivated by an earlier study [6] that showed the impact of mobility scheme over the performance of the network.

Routing algorithms

Three routing algorithms were injected in the simulation: AODV [3], DSDV [4], and TORA [5]. DSDV is a proactive protocol (routes are discovered and maintained all-time long) while AODV and TORA are reactive protocols (routes discovery is triggered only when needed).

III. The Analysis of Variance as a Network Design Tool

A system (or a model) is referred as *multivariate* when there are multiple dependant input variables. Multivariate systems present a much more complex behaviour as the global performance of the system can be affected by the values of one input variable *and* the interaction between two or more variables. By using multivariate tests, such as ANOVA [18], it is possible to get information about the strength of the relationship of the input variables. For instance, let us consider a system with 2 discrete-values input variables referred to as

α and β. The output of the system p_{ij} can be mathematically expressed as

$$p_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

where p_{ij} is the measurement of the output variable (e.g., the packet delivery fraction of the network) for the i^{th} level of α (e.g., the routing protocol) and the j^{th}level of β (e.g., the propagation model). α_i (resp. β_j) is the effect of the first (resp. second) variable and $(\alpha\beta)_{ij}$ captures the effect of the interaction between the two input variables. ϵ_{ij} is the residual random error.

The output of the ANOVA tests gives, for each input variable and each n-way interaction between the variables, the *F-statistic* and *p-value*. Shortly put, the *F-statistic* represents the level of signifiance of the variable (or interaction). The *p-value* is the probability to make a mistake by considering that this variable has a significant impact on the system: the *p-value* has to be kept as low as possible.

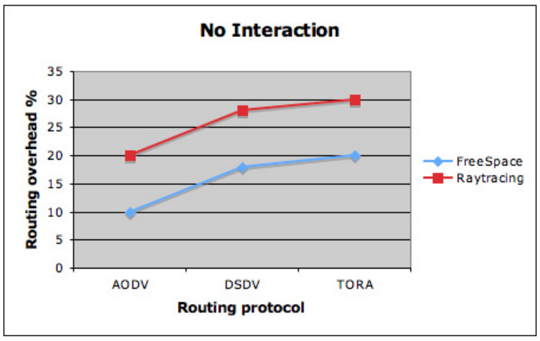

Fig. 3. Non-significant interaction between the two input variables.

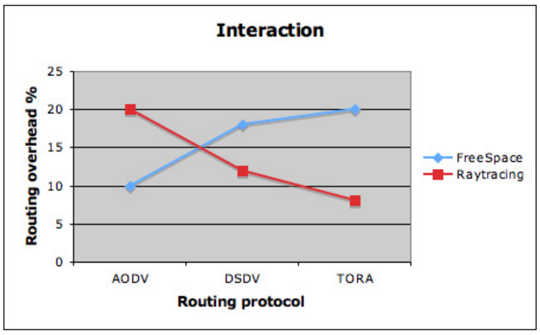

Fig. 4. Significant interaction between the two input variables.

A variable (or a n-way interaction between n variables) is considered to be significant if the variation of its level has a real impact on the system. Figures 3 and 4 show an example of the concept of significant interaction. In this example, we consider two input variables: the routing protocol and the propagation model. The output is the routing overhead. On Figure 3, changing from FreeSpace to Raytracing has no *significant* impact: the curve is slightly the same and it just raises the value of the ouput variable: the two variables are non-significantly correlated. On the other hand, on Figure 4, varying from FreeSpace to Raytracing changes *significantly* the shape of the curve for each level of the routing algorithm variable. The two variables are then refered as "variables interacting significantly".

IV. Summary and Interpretation

Figure 2 shows the table containing the ANOVA results. Each time a *p-value* is < 0.05 (confidence of up to 95%), the variable (or the n-way interaction) is considered to have a significant impact on the performance observed. A high level of *F-statistic* ensures that this variable has a high influence on the output. We give in the following lines an interpretation of these results and we will more particularly focus on the consequences of the n-way interactions between the protocol layers.

A. Routing overhead

The subset of significant interaction between input variables and routing overhead is {NODES, RTR, PROPA, NODES*RTR, RTR*PROPA}. Firstly, the NODES variable has a clear impact on the system in that a higher amount of nodes leads to more routing messages exchanged, especially in reactive routing protocols (AODV and TORA) where requests are flooded all over the network before establishing a route. While the analysis of the impact of the two next variables is trivial, it is important here to note the presence of two 2-way interactions: RTR*PROPA and NODES*RTR. The NODES*RTR can be explained easely in that some routing algorithms keep a constant message overhead while the number of nodes is growing. This result can be seen in [1] where DSDV keeps a quite low routing overhead while AODV has a linear-growing number of messages. Figure 5 shows this interaction.

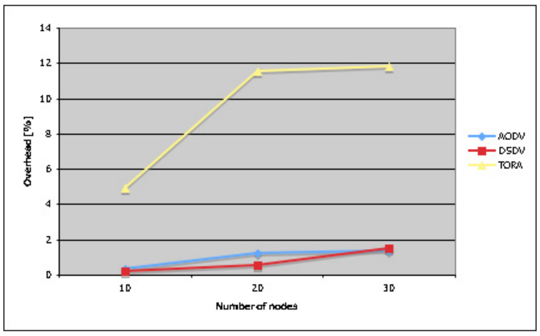

Fig. 5. Significant interaction between the number of nodes and the routing algorithm. All other variables are fixed (i.e., fixed scenario, propagation is raytracing)

An important observation about the results in Figure 2 is that the interaction between RTR and PROPA is significant for the routing overhead and the packet delivery delay. This result, often set out but never statistically explained, suggests that *the design of wireless indoor protocols should take into account the propagation conditions.* It also emphasizes that a rigorous study of MANET needs to rely on a good propagation and interference model, as pointed out in [1] and [11]. Figure

Input variables		Performance measures					
Interaction	Source	Routing overhead		Packet delivery fraction		end-to-end delay	
		F-test	p-value	F-test	p-value	F-test	p-value
1-way	NODES	10.5	0.0	60.3	0.0	35.3	0.0
	RTR	83.0	0.0	38.47	0.0	86.4	0.0
	PROPA	14.8	0.0	722.4	0.0	69.2	0.0
	SCEN	1.4	0.2	30.68	0.0	10.37	0.0
2-way	NODES*RTR	6.9	0.0	1.65	0.18	10.15	0.0
	NODES*PROPA	1.7	0.20	6.37	0.0	7.73	0.0
	NODES*SCEN	1.17	0.33	16.40	0.0	4.76	0.0
	RTR*PROPA	8.8	0.0	8.86	0.0	4.95	0.0
	RTR*SCEN	1.4	0.19	1.35	0.20	1.07	0.41
	PROPA*SCEN	1.98	0.06	11.46	0.0	4.08	0.0
3-way	NODES*RTR*PROPA	1.36	0.26	4.30	0.0	1.15	0.35
	NODES*RTR*SCEN	1.06	0.43	1.59	0.07	1.54	0.08
	NODES*PROPA*SCEN	1.13	0.6	6.29	0.0	3.44	0.0
	RTR*PROPA*SCEN	1.76	0.06	0.88	0.62	0.95	0.54

Fig. 2. Results of the ANOVA test. A *p-value*< 0.05 (confidence of 95%) indicates a significant result

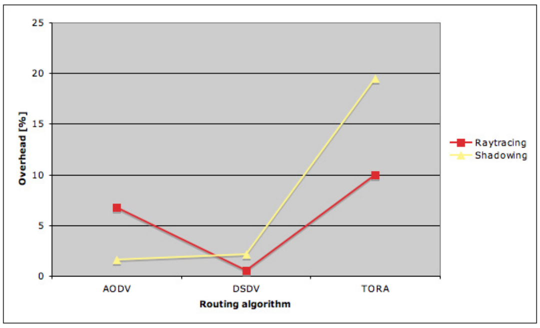

Fig. 6. Significant interaction between routing and the propagation. All other variables are fixed (i.e., fixed scenario, 20 nodes)

6 illustrates the interaction between routing and propagation model when the others variables of the simulation (i.e., amount of nodes and scenario) are constant. It is easely understandable that with active protocols like DSDV the routing overhead is constant, as it sends routes updates continuously. However, routes might be invalidated between two routes updates without triggering a new route lookup leading to a higher packet loss level (see below). This is unlike in AODV and TORA where a new route is searched when packets are losts or destination is no more reachable, leading to a much higher routing overhead in indoor (raytraced) propagation environment.

B. Packet Delivery Fraction

The subset of significant variables for the packet delivery fraction is {NODES, RTR, PROPA, SCEN, NODES*PROPA, NODES*SCEN, RTR*PROPA, PROPA*SCEN, NODES*RTR*PROPA, NODES*PROPA*SCEN}. The initial deployment (NODES, SCEN) and the mobility of the nodes has a clear impact on the fraction of successfully delivered packets. This is coherent with the results found in [6] and the we will omit here in-depth explaination due to lack of space. Instead, if we lock the values of the SCEN and NODES variables, we can see that PROPA plays a key role in the performance of the network and that PROPA and RTR are once more interacting. The role of the PROPA variable is a consequence of the scarce propagation conditions that exist in raytracing: the most realistic propagation conditions raises the number of packets loss due to between-nodes interference and signal variation over time (signal shadowing). Figure 7 indicates that, even if DSDV has lower routing overhead, it delivers succesfully less packets that reactive algorithms like AODV and TORA. This can be considered as a side effect of the periodic route beaconing that does not allow to reactively trigger a route reconstruction when some packets are lost.

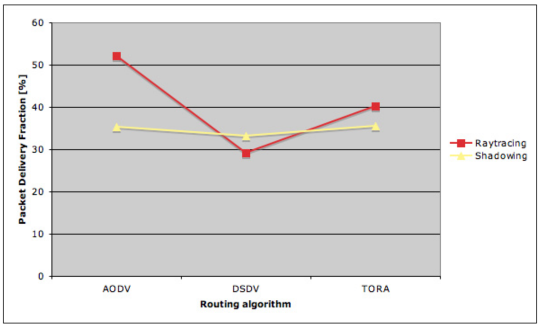

Fig. 7. Significant interaction between routing and propagation for packet delivery.

C. End-to-end delay

The subset of significant variables for the end-to-end delay delivery fraction is {NODES, RTR, PROPA, SCEN, NODES*RTR, NODES*PROPA, NODES*SCEN, PROPA*SCEN, NODES*PROPA*SCEN}. One can observe that for this metric, the most importants variables (those who have higher *F-test* and are interacting with the others) are those representing the density of nodes and their mobility scheme.

Other variables, like the 2-ways interactions between routing and propagation model and the interaction between routing and mobility are not significant. *It suggests that the end-to-end delay is mostly related to the interaction of the density of nodes with the other variables*. These variables (mainly routing parameters) should be thus adjusted accordingly to the number of nodes they will have to handle in order to get best performance.

Finally, the 2-way interaction between NODES and SCEN indicates that the distribution of the mobile nodes along time may strongly affect the delays and dropped packets in the network. Though, from a network designer's point of view it is impossible to control these two variables as they are modeling the fact that people move around with their mobile terminals, disturbing networking graph connectivity or creating temporary aggregates. Globally, we could say the SCEN variable suggests that the routing algorithm should adapt quickly to the large mobility of nodes, on requirement in which reactive routing algorithms behave better that active ones. The NODES variable indicates that a high density of nodes is required in order to have good connectivity but it raises the amount of routing messages exchanged (routing overhead) as it raises the number of hops required to reach a destination host (end-to-end delay).

V. Concluding Remarks and Future Work

We undertook a detailled and formal statistical study to characterize the effect of interaction between MANET protocols, network topology, and propagation conditions. The study emphasizes the significant aspects of network design and operation. It highlights the way wireless routing protocols can be deeply affected by propagation conditions and nodes density. The ANOVA method provides with a formal approach to find out what are the complex interactions that take place during network operation. This general method can also be used in order to improve routing algorithms by getting the set of parameters (e.g., beaconing frequence, route reconstruction strategy) that leads to a *real* modification of the performance (either gain or loss).

This work unveils that, when running an ad-hoc network simulation, the propagation model has a large impact on the network performance metrics. Moreover, the propagation model interacts with nearly every input variable entered in the simulator. It is now obvious that accurate simulations require a development effort in order to have good simulator code that covers today's wireless technologies (802.11a/b/g, 802.16, MIMO, etc.).

The undertaken work aim at the improvement of actual routing algorithms in order to develop a sort of "propagation-enhanced routing model". This adaptive algorithm will have to take into account the environment of operation (indoor/outdoor which is equivalent to shadowing/raytraced). ANOVA-based simulations will be used to demonstrate the real impact of these improvements and to provide a feedback on the way the parameters of the algorithm are useful to modify the behaviour of the network.

VI. Acknowledgements

The authors would like to thank Katia Lecocq for her statistical support and contstant help in making this work possible.

References

[1] J.-M. Dricot, Ph. De Doncker. High-accuracy physical layer model for wireless network simulations in NS-2. Proc. of the Int. Workshop on Wireless Ad-hoc Networks, IWWAN'04, Oulu, Finland, May-June 2004.
[2] J.-M. Dricot, Ph. De Doncker, E. Zimányi, F. Grenez, Impact of the physical layer on the performance of indoor wireless networks In Proc. of the Int. Conf. on Software, Telecommunications and Computer Networks, Split, Croatia, October 2003.
[3] Charles E. Perkins and Elizabeth M. Royer. "ad-hoc On-Demand Distance Vector Routing." Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, February 1999, pp. 90-100.
[4] C. Perkins and P. Bhagwat. Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In Proc. of the ACM SIGCOMM, October 1994.
[5] V. Park and S. Corson. A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks. Proc. IEEE INFOCOM ?97, Kobe, Japan (April 1997).
[6] M. Bhatt, R. Chokshi, S. Desai, S. Panichpapiboon, N. Wisitpongphan, and O.K. Tonguz, "Impact of Mobility on the Performance of ad-hoc Wireless Networks", IEEE VehiculTechnology Conference (VTC 2003), Orlando, FL, October 2003.
[7] Mineo Takai and Jay Martin and Rajive Bagrodia. Effects of wireless physical layer modeling in mobile ad hoc networks. MobiHoc '01: Proceedings of the 2nd ACM international symposium on Mobile ad-hoc networking & computing. Long Beach, CA, USA, 2001.
[8] E. Royer, S. Lee and C. Perkins. The Effects of MAC Protocols on ad-hoc Network Communications. Proc. IEEE Wireless Communications and Networking Conference, Chicago, IL, September 2000.
[9] J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva. Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. Proc. 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking, ACM, Dallas, TX, October 1998.
[10] Christopher L. Barrett, Martin Drozda, Achla Marathe, and Madhav V. Marathe. Analyzing Interaction Between Network Protocols, Topology and Traffic in Wireless Radio Networks. Proc. IEEE Wireless Communications and Networking Conference (WCNC'03), New Orleans, Louisiana, IEEE Press, 2003.
[11] A. L. Cavilla, et al., "Simplified Simulation Models for Indoor MANET Evaluation Are Not Robust," Proc. SECON 2004.
[12] K. Pawlikowski, H.-D.J Jeong, and J.-S.R. Lee. On credibility of simulation studies of telecommunication networks. IEEE Communications, 40(1):132 139, January 2002.
[13] The Network Simulator. http://www.isi.edu/nsnam/ns/
[14] Per Johansson, Tony Larsson, Nicklas Hedman, Bartosz Mielczarek, and Mikael Degermark. Scenario-based performance analysis of routing protocols for mobile adhoc networks. In Proceedings of the Fifth Annual International Conference on Mobile Computing and Networking, pages 195 205. ACM Press, 1999.
[15] J. Broch, D. A. Maltz, D. B. Johnson, Y. C. Hu, and J. Jetcheva. A Performance Comparison of Multi-Hop Wireless ad-hoc Network Routing Protocols. In Proc. of the ACM/IEEE MobiCom, October 1998
[16] Catedra, Manuel F., Cell Planning for Wireless Communications, Artech House, Inc, 1999
[17] Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu, and Jorjeta Jetcheva. A Performance Comparison of Multi-Hop Wireless ad-hoc Network Routing Protocols. In Proceedings of the Fourth Annual International Conference on Mobile Computing and Networking (Mobi-Com'98), ACM, Dallas, TX, October 1998.
[18] Howell David. Statistical Methods for Psychology 5th ed. June 2001. Thomson Learning editor