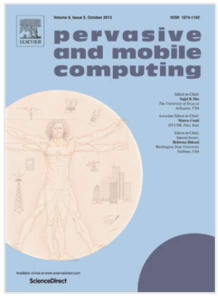
Accepted Manuscript



The role of congestion in probabilistic broadcasting for ubiquitous wireless multi-hop networks through
mediation analysis

D.G. Reina, S.L. Toral Marín, E. Asimakopoulou, F. Barrero,
N. Bessis

PII: S1574-1192(15)00121-2
DOI: http://dx.doi.org/10.1016/j.pmcj.2015.06.014
Reference: PMCJ 611

To appear in: *Pervasive and Mobile Computing*

Please cite this article as: D.G. Reina, S.L. Toral Marín, E. Asimakopoulou, F. Barrero, N. Bessis, The role of congestion in probabilistic broadcasting for ubiquitous wireless multi-hop networks through
mediation analysis, *Pervasive and Mobile Computing* (2015),
http://dx.doi.org/10.1016/j.pmcj.2015.06.014

THE ROLE OF CONGESTION IN PROBABILISTIC BROADCASTING FOR UBIQUITOUS WIRELESS MULTI-HOP NETWORKS THROUGH MEDIATION ANALYSIS

Authors: Reina, D. G.[1], Toral Marín, S.L.[1], Asimakopoulou, E.[2], Barrero, F.[1], Bessis, N.[2]

[1]*University of Seville (Spain)*

[2]*University of Derby (UK)*

[1] Escuela Superior de Ingenieros. Avda. Camino de los Descubrimientos, s/n

41092 Sevilla (Spain)

email: toral@esi.us.es, fbarrero@esi.us.es, dgutierrezreina@us.es

phone: +34 954 48 12 93; fax: +34 954 48 73 73

[2] School of Computing and Mathematics, Kedleston Road

Derby, DE22 1GB (UK)

email: eleana.asimakopoulou@googlemail.com, n.bessis@derby.ac.uk

phone: +44 1332 592108; fax: +44 1332 597741

Corresponding author: N. Bessis

THE ROLE OF CONGESTION IN PROBABILISTIC BROADCASTING FOR UBIQUITOUS WIRELESS MULTI-HOP NETWORKS THROUGH MEDIATION ANALYSIS

Abstract: Broadcast schemes play an important role in the performance of mobile ad hoc networks, which are a clear example of ubiquitous wireless multi-hop networks where nodes collaborate in a distributed way. They are widely used as a dissemination mechanism and as a part of the discovery phase of routing protocols. The simple flooding algorithm is the usual mechanism employed in mobile ad hoc networks, but its inefficiency has been demonstrated in congested scenarios due to the high number of collisions and contentions. However, these problems can be partially alleviated by using a probabilistic broadcast approach in which every node forwards the incoming packets according to a certain forwarding probability. In this paper, we use a simple probabilistic broadcast protocol to evaluate the effects of congestion on the performance of broadcasting in ad hoc networks through a mediation analysis. We hypothesize that the congestion mediates in the relationship between the forwarding probability (independent variable) and the output metric (dependent variable). We consider several output metrics according to the application of the broadcasting protocol such as reachability, broadcasting delay, packet delivery fraction and end to end delay. The simulation results show the existence of the mediating effects and how such effects may be counterbalanced depending on the target use of the probabilistic broadcast scheme.

Keywords Data Dissemination, Broadcasting, Routing protocols, ubiquitous wireless multi-hop networks, Mediation Analysis.

1. Introduction

Broadcasting protocols are generally used in ubiquitous wireless multi-hop networks such as Mobile Wireless Ad Hoc Networks (MANETs) and Vehicular Ad Hoc Networks (VANETs) [1][2] in two ways, as a stand-alone dissemination technique and as a part of the discovery process of routing protocols [3][4]. In the former case, the same information has to be transmitted to every node in the network as quick as possible. This is the case of disaster/emergency scenarios, which are typical applications of MANETs [5][6][7]. In the latter, nodes need to collect neighbour information in order to find a communication path between a source node and a destination node (unicast) or several destination nodes (multicast). In particular, many routing protocols designed for MANETs and VANETs rely on broadcasting as the mechanism to find an available route between two nodes [2]. As a consequence, an efficient broadcast mechanism is of paramount importance for establishing secure mobile ubiquitous communications among wireless electronic devices such as smartphones, laptops, tablets, etc.

In general, nodes use a simple broadcasting approach, namely flooding, in which every node retransmits once a given incoming packet. The simplest form of flooding has been demonstrated to be inefficient in terms of bandwidth and power consumption, causing the well-known Broadcast Storm Problem [8]. Alternative techniques to simple flooding can be found in [9]. A basic classification of broadcasting protocols divides them into two main categories, deterministic and probabilistic protocols. In deterministic protocols, a subset of nodes are selected as forwarders. Connected Dominating Sets (CDS) and Minimum Spanning Trees (MST) are some examples of deterministic broadcasting protocols [10]. However, such optimal solutions are NP-problems and may require global information, which is costly in terms of information exchange. Although there are distributed versions of CDS and MST [10], deterministic approaches may present some problems in mobile conditions since the algorithm used to select the forwarders has to be rerun continuously due to the topological changes. On the other hand, in probabilistic protocols, all nodes in the network have the opportunity of forwarding an incoming broadcast packet [11]. On receiving a new packet, a node forwards the received packet with probability p and it does not forward the packet with probability $1-p$. Although the simplicity of the probabilistic approach, it is not easy task to find the optimum forwarding probability which guarantees a high reachability in the network. Yet, probabilistic broadcast exhibits several advantages such as a good balance of power consumption among nodes and robustness against mobility and malicious nodes. The probabilistic broadcast schemes require an adequate selection and adaptation of the probability value depending on the specific conditions of the target scenario. Many heuristics have been proposed in the literature [11][12][13][14], which determine the forwarding probability based on parameters such as the position of nodes [12], the density of nodes [11] and the speed of nodes [13], among other parameters [14]. The main problem is that there are many interrelated variables that may cause undesirable effects when modifying the value of probability.

In this paper we focus on simple probabilistic broadcast scheme since the forwarding probability value allows us to easily tune the performance of probabilistic broadcast. In particular, we hypothesize that congestion is the key variable mediating the effects of the forwarding probability on the performance of a given output metric. To accomplish this goal we propose the use of a mediation analysis [15], which is a widely used technique in statistics to evaluate the effects of a mediator variable in the relationship between a dependent variable and an independent variable. This work is an extension of our previous work [16], considering delay-based output metrics and the implications of the obtained results for designing broadcasting and routing protocols based on probability.

The rest of this paper continues as follows: section 2 includes a related work which is divided into two parts, the first one introduce probabilistic broadcasting in ad hoc networks, and the second one presents the role of broadcasting in the congestion of ad hoc networks. Section 3 describes the proposed mediation analyses. Section 4 details the simulation results of the mediation analyses. Section 5 contains the discussion of the results presented in Section 4 and some future works. Finally, the conclusions of this paper are included in section 6.

2. Related Work

2.1. Probabilistic broadcasting in ad hoc networks

Probabilistic broadcasting has been an active research field in ubiquitous wireless multi-hop networks such as MANETs and VANETs for the last decade [4][11][12][13][14][17]. The main objective of probabilistic broadcasting is to find a forwarding probability value which guarantees that every node in the network receives a given broadcasting packet sent by a source node. The simplest solution may be simulating different forwarding probabilities until achieving the desired reachability, applying a try-and-error approach. However, it is not an efficient solution in extremely changeable environments such as MANETs and VANETs, where it is very difficult to find the same topology in two different moments.

Percolation theory and the phase transition phenomenon presented in random networks [18][19] seem to be a possible solution to find the desired optimal forwarding probability. A complex network like an ad hoc network exhibits the phase transition phenomenon, when the connectivity of the network changes abruptly, from a non-connected network to a connected network (giant component), by slightly varying a network parameter like the density (number of nodes) or the node's transmission range.

In the probabilistic broadcasting case, the idea is to check out if varying the forwarding probability value, we are able to observe such abrupt connectivity change beyond certain critical probability. In Sasson et al [18], the authors study the phase transition phenomenon in ad hoc networks and indicate some important differences between real-world MANETs and mathematical graphs such as random networks:

1) the real-world MANETs are not infinite so border effect may impact the system's behaviour,

2) nodes can leave and join the networks constantly modifying the network's density and consequently affecting the phase transition phenomenon and

3) collisions and contention also impact on the critical probability. The simulation results in [18] show that in real-world MANETs the phase transition is linear with the density of nodes for low network's densities so the network's connectivity does not change abruptly.

In addition, the results in [18] indicate that for networks with high density a low value of forwarding probability is sufficient to achieve a high connectivity. These simulation results are corroborated in the recent experimental results presented in [21], where the authors evaluate probabilistic broadcasting in a real-life testbed like DES-testbed [22] and indicate that collisions and noisy links are the main causes by which the phase transition phenomenon is not observed in ad hoc networks. Collitions are mainly caused by congesting the networks with messages since nodes should share and content for the wireless medium.

Since the difficulties in finding an optimal forwarding probability for every possible scenario, many heuristics based on topological parameters have been proposed in the last decade [11][12][13]. Parameters such as the number of neighbours [11], the number of duplicated packets, the speed of nodes [13], and the position of nodes [12], are suitable candidates to be used to determine the forwarding probability of nodes in a distributed and efficient manner. Although heuristic approaches cannot determine the optimal forwarding probability, they are suitable solutions, balancing cost and complexity.

In this paper, we make a step forward in evaluating probabilistic broadcasting, considering the congestion of the network as a primary parameter for selecting the forwarding probability. The results presented in this paper may help to improve the performance of some heuristic protocols found in the literature.

2.2. The effects of congestion in the performance of ad hoc networks: the role of broadcasting

In TCP-based networks like wired Internet, the congestion of the network is controlled by adjusting the buffer of nodes. Nodes adjust the congestion of the communication flows between two nodes by reducing the rate at which the packets are generated in the source. In wired networks, congestion is the only cause for the loss of packets. However, that is not the case in mobile wireless ubiquitous networks such as MANETs and VANETs in which mobility of nodes is also a primary reason. In such wireless networks, nodes have to share and content for the wireless medium. Thus, the MAC layer is responsible for enabling nodes to access the wireless medium in a coordinated mode. IEEE 802.11 DCF is de facto MAC standard protocol in MANETs and 802.11p is the one for VANETs [23]. In both cases, the performance of the MAC protocols is degraded in congested scenarios because of common issues in wireless networks such as the hidden terminal problem, the exposed terminal problem and the broadcast storm problem [8], which are even more aggravated in congested scenarios. The result is a worsening of the general performance of the ad hoc application.

In ubiquitous multi-hop networks, broadcasting plays an important role in the congestion state of the network. As a rule, nodes use massively broadcasting packets in the discovery phase of ad hoc routing protocols, occupying most of the available wireless medium. Notice that broadcasting packets have more priority in the sending buffers than application packets. As a result, the application packets may find difficult or impossible to access the medium, and consequently, reducing the delivery of packets from a source node to a destination node. Notice that in case that the application is broadcasting as a stand-alone dissemination of messages, it is also affected by the contention and the collisions of messages. To avoid or alleviate this problem, nodes should control the rate at which broadcasting packets are generated and/or retransmitted. The generation of broadcasting packets depends on the application, so it is difficult to control. Therefore, controlling the retransmission of broadcasting packets is a more feasible solution.

In this paper, we study how the forwarding probability can be used to control congestion by reducing the retransmission of broadcasting packets. We use a mediation analysis to quantify qualitatively and quantitatively the effects of congestion in the performance of MANETs when controlling the congestion by adjusting the forwarding probability.

2.3. Congestion control mechanisms for multi-hop ad hoc networks

The study of congestion control algorithms has been an active research topic for the last few years [15] [16]. In [15], the authors classify the congestion control techniques into four categories, such as a) traffic control, b) resource control, c) priority-aware congestion control schemes, and d) queue-assisted techniques. A similar classification can be found in [16], but in this case, the categories include traffic control, resource control, and traffic and resource control. We focus on resource control techniques since communication rates in multi-hop ad hoc networks depend in general on the target application. Consequently, we should not modify the data rates as required by traffic control mechanisms. For instance, in event-based or critical-mission applications, which are typical target application for multi-hop ad hoc networks like disaster scenarios, all the generated information should be delivered at the destination nodes as quick as possible. Alternatively, resource control mechanisms find alternative paths to route the data application among nodes in the network. In [17], the authors present HTAP a distributed framework for minimizing congestion and assuring reliable data transmissions in event based networks. In this approach, when congestion is about to occur alternative paths are selected to route data. These alternative routes are created by using non-congested nodes in the network. Although, this approach cannot guarantee the shortest path between the source and destination nodes, it alleviates the congestion of the network. In [18], LACAS an automata-based resource control mechanism is presented. As the previous control mechanism, when congestion is detected alternatives path are selected. However, the main feature of LACAS is that it uses a machine learning mechanism to learn from the past behaviours. In [19], the authors present CRP a congestion adapted routing protocol for MANETs. CRP uses a bypass mechanism to alleviate congested nodes. When a node detects that it is about to be congested, it informs the previous nodes in the routing path on this circumstance in order to find other alternatives routing paths. Then, the data in probabilistically routed through those alternative paths found. As a result, the congested node is bypassed.

In this context, the resource control mechanism is implemented by a probabilistic retransmission mechanism at network layer. The main idea is that by controlling the retransmission probability, we can control the number of different paths explored by a routing protocol, and therefore, we control the congestion of the network.

3. The proposed mediation analyses

3.1 Introduction to mediation analysis

In [15] mediation is defined as one way in which a researcher can explain the process or mechanism by which one variable affects another. The common mediation model is illustrated in Figure 1. This mediation model hypothesizes that the independent variable (X) affects the mediator variable (M), which in turn affects the dependent variable (Y). Therefore, the mediator variable is useful to clarify and identify the nature of the relationship between the independent and dependent variables [20].

In the general mediation model (Figure 1), e_1-e_3 represent the residuals of the regression models (described in the next section), and a, b, and c' are the regression coefficients. The direct effect, c' in Figure 1, measures the extent to which the dependent variable changes when the independent variable increases by one unit. In contrast, the indirect effect, which is the product of terms a and b in Figure 1, measures how dependent variable changes when the independent variable is held fixed and the mediator variable changes by the amount it would have changed with the independent variable increased by one unit.

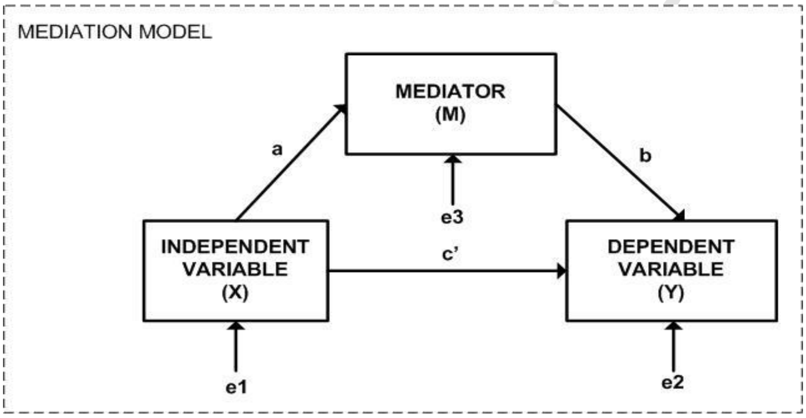


Fig. 1. General mediation model.

In our case, the objective is to analyze the effect of congestion on the relationship between an independent variable like the forwarding probability p, and a dependent variable which is the output metric. The used output metric will depend on the target objective of broadcasting. For this reason, the proposed mediation analysis is twofold:

- **Probabilistic broadcast as a standalone dissemination technique:** In this case, the independent variable is the forwarding probability p, the mediator is the congestion (C) which is measured as the number of packets sent by nodes in the network (considering routing and application packets) and we consider two independent variables (output metrics) such as the Reachability (*Re*) and the broadcasting delay (*D*). In this case, the mediation models are represented in Figure 2. Notice that *Re* is one of the most used metrics to evaluate broadcast schemes in MANETs [8]. In particular, *Re* is defined as the number of nodes that received a given broadcasting packet generated by a source node divided by total number of nodes in the network (connected network) or reachable from the source node (non-connected network). A high *Re* value is always desired in an efficient broadcasting protocol. Regarding the delay *D*, it is defined as the time elapsed since the source node generates the broadcasting packet until the last node receives the broadcasting packet. A low value of *D* is always a desirable feature for a broadcasting protocol.

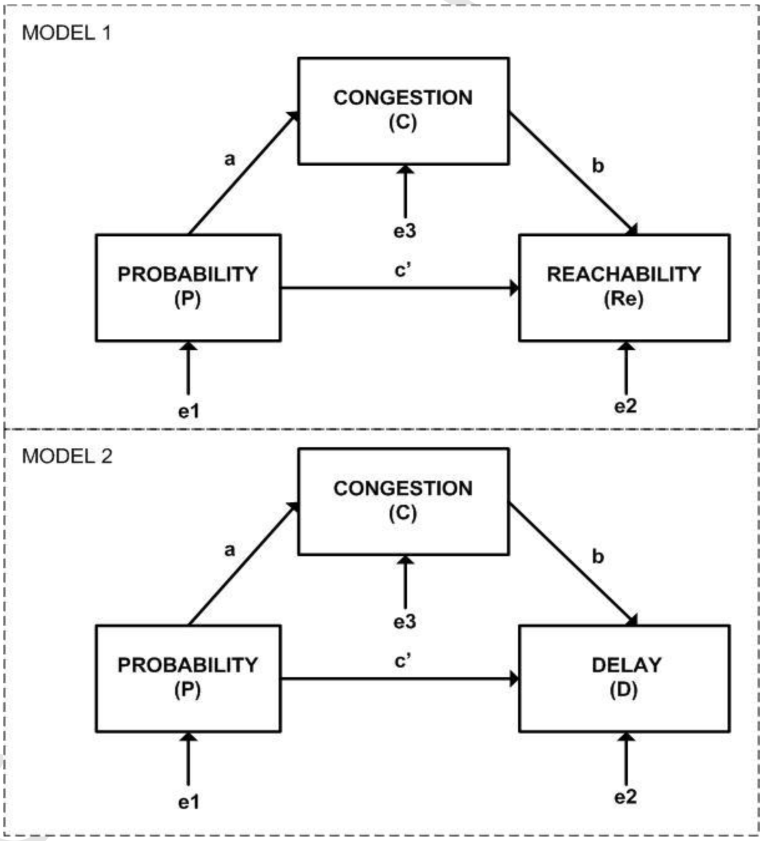


Fig. 2. Mediation models for probabilistic broadcast as a stand-alone dissemination technique.

- **Probabilistic broadcast as part of the discovery routing protocols:** In this case, the independent variable is again the forwarding probability *p*, the mediator is the congestion (*C*) measured as the number of packets sent by nodes in the network (considering routing and application packets), and we again consider two dependent variables such as the Packet Delivery Fraction (*PDF*) and the End to End delay (*E2E*). The *PDF* only takes into account application packets and is defined as the ratio between received and sent packets. Notice that similarly to *Re*, *PDF* is a widely used metric for assessing routing protocols in MANETs [14]. The main objective of an efficient routing protocol is to achieve a high *PDF*. Regarding *E2E*, it is defined as the time elapsed since the source node sent a broadcasting request (broadcasting packet) to establish a new route to a destination node until it receives the first broadcasting reply (broadcasting packet) from any of its neighbours. A low value of *E2E* is a desirable requirement for a routing protocol. In this case, the mediation models are illustrated in Figure 3.

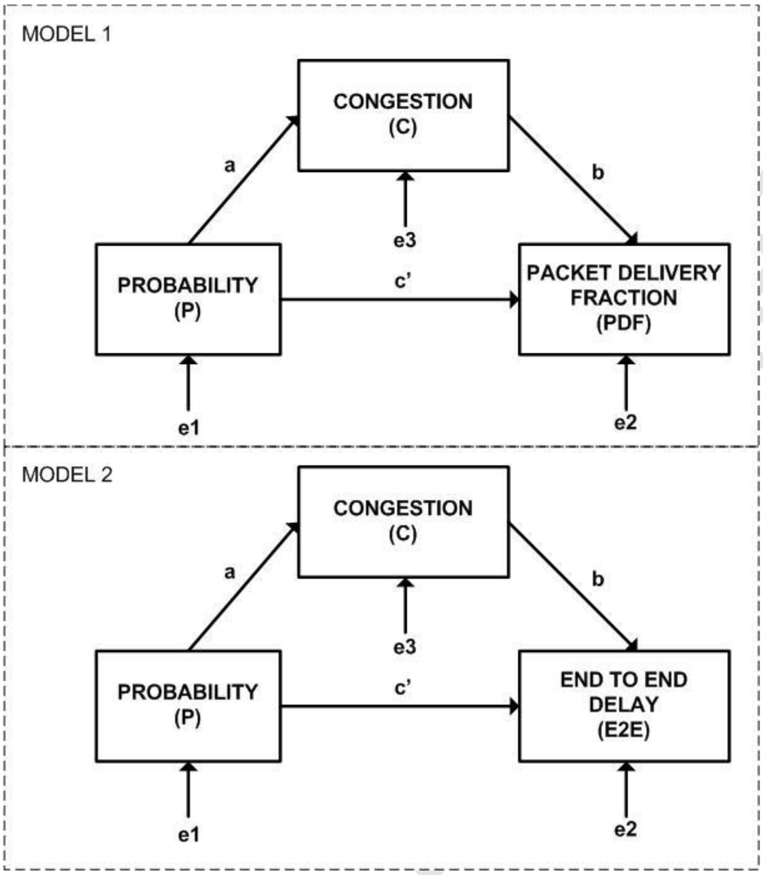


Fig. 3. Mediation models for probabilistic broadcast as a part of the discovery phase of routing protocols.

3.2 Methodology for assessing mediation effects

The causal steps approach described by Baron and Kenny [30] is used in this paper to assess the proposed mediation models. According to [30], the process is divided into four steps:

1) **Independent variable ~ Dependent variable:** Proving a significant relationship between the dependent and the independent variable is required. The regression model (1) should be used to test the relationship between variables (see Figure 1, where *X* represents the independent variable, *Y* represents the dependent variable, and *M* is the mediator). The coefficient *c* is also called the total effect. The terms i_1-

i_3 in the following equations (1), (2), and (3) represent the intercepts of the regression models.

(1)

2) **Mediator ~ Dependent variable:** To demonstrate a significant relationship between the independent variable and the hypothesized mediating variable. It means that the coefficient *a* in the regression model (2) must be significant.

(2)

3) **Independent variable ~ Dependent variable + Mediator:** To demonstrate a significant relationship between the mediator and the dependent variable when both the independent variable and the mediator are predictors of the dependent variable. Consequently, the term *b* in the regression model (3) must be significant. The coefficient *c'* is also called the direct effect, and the product *ab* is called the indirect effect.

(3)

4) **Sobel test:** To test the mediating effect. The objective of this step is to determine the significance of the indirect effect. The Sobel test is one of the most used methods to test the significance of the indirect effect [15].

4. Simulation results and discussion

4.1. Data Collection

The simulator NS-2.34 has been used to collect the data for the proposed mediation analysis. The nodes move following the Waypoint mobility model. In this mobility model nodes move randomly throughout the simulation scenario. When the target destination is reached, another random destination is chosen using a different random speed. The simulation scenario and the number of nodes both have been chosen to meet standard scenarios, as described by Kurkowski et al [31]. An Average Network Partitioning (ANP) of 0.03 (medium-high density) and the Average Shortest-path Hop Count (AspHops) of 4 have been chosen (it is recommended a number of hops higher than three [15]). These values characterize a standard MANET scenario with medium-high connectivity. These values of ANP and AspHops are equivalent to a scenario with 112 nodes and an area of 1696 x 1696 m^2. The routing protocol used is Ad Hoc On-demand Distance Vector (AODV) [32], which uses simple flooding during the discovery phase. The MAC layer is IEEE 802.11 Distributed Coordinate Function (DCF). Nodes have a coverage range of 250 m and the two-ray propagation model has been used. The simulation time is 300 s. Regarding the independent variable, random numbers following a Gaussian distribution have been generated through a Python script with a mean of 0.6 and a standard deviation of 0.15. The traffic pattern used is Constant Bit Rate (CBR) with User Datagram Protocol (UDP) as the transport protocol. The packet size is 512 bits and a rate of 4 packets per second has been chosen. The number of communications among nodes is 50 and they have been chosen randomly during the first 50 s of the simulation time. Up to 100 different broadcast processes have been considered for carrying the regression models involved in the four steps, implementing the Baron and Kenny procedure.

4.2. Assessing mediation effects

The mediation effects will be evaluated in two cases, 1) probabilistic broadcast as a stand-alone dissemination technique, and 2) probabilistic broadcast as a part of the discovery phase of routing protocols. In both cases, the same methodology described in section 2.2 has been used. The statistical software R has been used [33].

A. Probabilistic broadcast as a stand-alone dissemination technique

The objective of this section is to evaluate the models presented in Figure 2. We first evaluate the model that considers the *Re* as the output metric and then considers the second model with the delay *D* as the output metric.

1) Re as output metric:

Step 1 (Re ~ P): The objective is to demonstrate a significant relationship between the forwarding probability and *Re*.

Table 1. Results of the regression analysis Re ~ P.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.484	0.009	55.090	<2e-16***
Probability	0.355	0.014	25.150	<2e-16***

Significant codes: 0 ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘.’, 0.1 ‘ ’, 1.

The above significant codes (Table 1) are used for the null-hypothesis testing. A p-value lower than 0.05 or 0.01 is needed for rejecting the null-hypothesis at 0.05 or 0.01 significance level. According to the results included in Table 1, step 1 is accomplished and there is a significant relationship between the probability and *Re*. The total effect *c* is 0.355.

Step 2 (C ~ P): The objective is to demonstrate a significant relationship between the forwarding probability and the congestion of the network.

Table 2. Results of the regression analysis C ~ P.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	-495281	22979	-21.570	<2e-16***
Probability	2806390	36882	76.090	<2e-16***

According to the results included in Table 2, there is a significant relationship between the forwarding probability and the congestion. As the forwarding probability increases, the congestion of the network also increases.

Step 3 (Re ~ P + C): In this step, the congestion must be significantly related to *Re* when both the forwarding probability and the congestion are predictors of *Re*.

Table 3. Results of the regression analysis Re ~ P + C.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	5.962e-01	2.340e-02	25.470	<2e-16***
Probability	-2.803e-01	1.266e-01	-2.210	3.17e-2*
Congestion	2.263e-07	4.494e-08	5.035	7.45e-6***

As can be seen in Table 3, the congestion is significantly related to *Re*. However, the forwarding probability is not significantly related to the reachability. The direct effect is -2.803e-01 but it is lower (in absolute value) than the total effect 0.354 calculated in step 1. This result means that the impact of probability on *Re* is mainly caused by the indirect path through the congestion *ab*=0.635.

Step 4 (Sobel test): In this step, we demonstrate the significance of the indirect path *ab*. According to the results included in Table 4, the indirect path is equal to 0.635 and it is significant since the associated z-value is higher than 1.96. Notice that the z-value is the result of the z-test used in the Sobel test for determining the significance of the indirect effect. If z-value is higher than 1.96 (in absolute value), then the indirect effect can be considered significant at the 0.05 level.

Table 4. Results of the regression analysis Re ~ P + C.

Indirect effect	Std. Error	z-value
0.635	0.126	5.024

As a conclusion of this mediation analysis, we can state that congestion mediates positively in the relationship between the forwarding probability and *Re*. Consequently, if the forwarding probability is increased *Re* will be also increased.

2) Delay as output metric:

Step 1 (D ~ P): The objective is to demonstrate a significant relationship between the forwarding probability and *D*. The results (Table 5) show that there is a significant relationship between the forwarding probability and *D*. The total effect *c* is positive and is equal to 0.38833. It means that the delay increases as nodes increase their forwarding probability. Notice that it is a negative effect since the objective of a broadcasting algorithm is to spread out the information as quick as possible.

Table 5. Results of the regression analysis D ~ P.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.079	0.022	3.536	6.30e-04***
Probability	0.388	0.036	10.827	<e-16***

Step 2 (C ~ P): The results of this regression analysis are the same shown in Table 2 so the step 2 is accomplished.

Step 3 (D ~ P + C): In this step the congestion must be significantly related to *D* when both the forwarding probability and the congestion are predictors of *D*. As for the results in Table 7, the delay is significantly related to the congestion when both probability and *C* are considered. However, as happened with the *Re*, the probability is not significantly related to *D*. It means that the delay is mainly caused by the indirect effect through the congestion. The indirect effect *ab* is 0.411.

Table 7. Results of the regression analysis D ~ P + C.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	1.581e-01	3.784e-02	4.178	6.64e-5***
Probability	-2.291e-01	1.647e-02	-0.139	0.88
Congestion	1.422e-08	5.568e-08	2.554	0.01*

Step 4 (Sobel test): In this step we demonstrate the significance of the indirect path *ab*. We can observe in Table 8 a significant indirect effect. Consequently, if nodes increase the forwarding probability, the packets will have a higher *Re* but incurring in an increase of *D*. Notice that it is a negative feature for broadcasting algorithm.

Table 8. Significant test of indirect effect.

Indirect effect	Std. Error	z-value
0.411	0.161	2.550

B. Probabilistic broadcast in the discovery phase of routing protocols

The objective of this section is to evaluate the models presented in Figure 3. We begin with the model that considers the *PDF* as the output metric and then we continue with the model that considers the *E2E* delay.

1) PDF as output metric:

Step 1 (PDF ~ P): According to the results included in Table 9, a significant relationship between the forwarding probability and the *PDF* has been demonstrated.

The total effect *c* is -0.173. The negative value indicates that the *PDF* decreases as nodes increase their forwarding probability.

Table 9. Results of the regression analysis PDF ~ P.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.459	0.004	117.328	<2e-16***
Probability	-0.173	0.006	-27.589	<2e-16***

Step 2 (C ~ P): The results of this step are the same as those included in Table 2.

Step 3 (PDF ~ P + C): In this step the congestion must be significantly related to the *PDF* when both the forwarding probability and the congestion are predictors of the *PDF*.

Table 10. Results of the regression analysis PDF ~ P + C.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.408	1.037e-02	39.413	<2e-16***
Probability	0.112	5.613e-02	2.004	5.07e-2*
Congestion	-1.018e-07	1.992e-08	-5.114	5.71e-06***

According to the results shown in Table 10, the congestion is significantly related to the *PDF*. In fact, the congestion impacts negatively on the *PDF* value. The direct effect *c'* is 0.112, which is lower than the total effect *c*. It can be noticed that a change of sign in the relationship between the forwarding probability and the *PDF* (see Table 9 and Table 10). This change is known as inconsistent mediation and it is the result of the mediator variable acting as a suppressor [13]. In fact, although increasing the forwarding probability could lead to increasing the *PDF*, the mediating effect of congestion results in a negative impact on the *PDF*. Notice that this result is opposed to the case of *Re* as the output metric. The main implication of these results is that in congested networks lowering the forwarding probability during the discovery phase can improve the performance of routing protocols in terms of *PDF*. It means that when the network is congested the only way to increase the *PDF* is by reducing the forwarding probability.

Step 4 (Sobel test): According to the results of Table 11, the indirect effect is *ab* = -0.286 and it is significant. The congestion affects negatively the relationship between the forwarding probability and the *PDF*.

Table 11. Significant test of indirect effect.

Indirect effect	Std. Error	z-value
-0.286	0.056	-5.103

2) E2E as output metric:

Step 1 (E2E ~ P): First we need to demonstrate a significant relationship between the forwarding probability and the *E2E* delay. According to the results shown in Table 12, there is a significant relationship between the forwarding probability and the E2E metric. The total effect c = 1.043 is positive, so it means that increasing the forwarding probability causes an increment of *E2E*, which is a negative effect.

Table 12. Results of the regression analysis E2E ~ P.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	1.746	0.171	10.212	<2e-16***
Probability	1.043	0.275	3.793	2.63e-04***

Step 2 (C ~ P): The results of this step are the same as those included in above Table 2.

Step 3 (E2E ~ P + C): According to the results shown in Table 13, the congestion mediates the effect of the forwarding probability to the *E2E*. The forwarding probability is not significantly related to the *E2E*. It means that the indirect effect is the main cause for the increment of *E2E*. Since the indirect effect is positive, the *E2E* delay increases with an increment of the forwarding probability, which also causes an increment of congestion in the network.

Table 13. Results of the regression analysis E2E ~ P + C.

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	2.340	2.907e-01	8.047	2.69e-12***
Probability	-2.037	1.265	-1.609	1.10e-01
Congestion	1.065e-06	4.277e-07	2.490	1.45e-02*

Step 4 (Sobel test): The results of the Sobel test (Table 14) indicates that the indirect effect is significant. It is interesting that the indirect effect for the *E2E* model is higher than in the case of *D*. In both cases, the indirect effect is positive, but congestion affects more to the *E2E*. Consequently, the congestion impacts more severely to the application layer than the routing layer.

Table 14. Significant test of indirect effect.

Indirect effect	Std. Error	z-value
3.081	1.239	2.487

C. Extending the mediation analysis for different levels of congestion and implications for broadcasting and routing protocols design

The proposed mediation analyses have been extended for different levels of congestion. The number of communication flows have been varied from 25 to 75 in steps of 25. The obtained results are included in Table 15, Table 16, Table 17, and Table 18, and also represented in Figure 4, Figure 5, Figure 6 and Figure 7. The main findings are:

- The mediation effects are always positive in case of *Re*, *D*, *E2E* and are always negative in case of *PDF* (see Figure 4-7). Notice that a positive indirect effect is not a good feature for *D* and *E2E* model. Since it means that in both cases the packets will be delivered with high delays.
- Only *Re* improves with high values of forwarding probability (high congestion, see Figure 4).
- Both metrics of application layer *PDF* and *E2E* decrease notably with the increment of congestion (see Figure 6 and Figure 7). Consequently, the application layer is more affected by an increment of the forwarding probability.
- The mediation effects are always stronger as we increase the number of flows.

Table 15. Indirect effect for Re under different number of flows.

Flows	Indirect Effect	z-value
25	0.45	0.45
50	0.635	5.02
75	0.62	7.6

Table 16. Indirect effect for D under different number of flows.

Flows	Indirect Effect	z-value
25	0.06	3.91
50	0.16	2.55
75	0.81	2.88

Table 17. Indirect effect for PDF under different number of flows.

Flows	Indirect Effect	z-value
25	-0.40	-2.63
50	-0.28	-5.10
75	-0.22	-3.15

Table 18. Indirect effect for E2E under different number of flows.

Flows	Indirect Effect	z-value
25	3.55	5.68
50	3.08	2.48
75	3.03	2.86

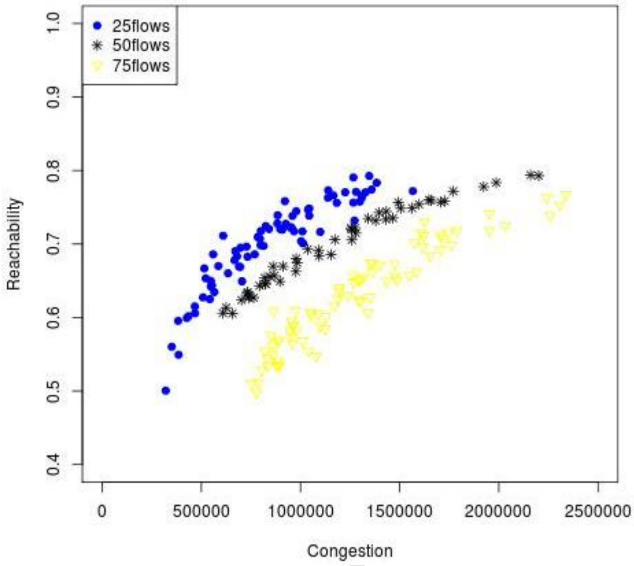


Fig. 4. Reachability vs Congestion for different number of flows.

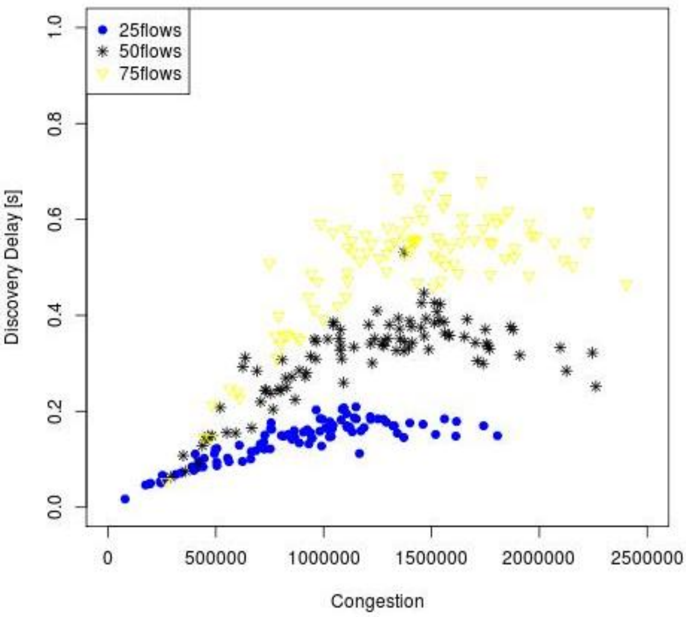


Fig. 5. Reachability vs Congestion for different number of flows.

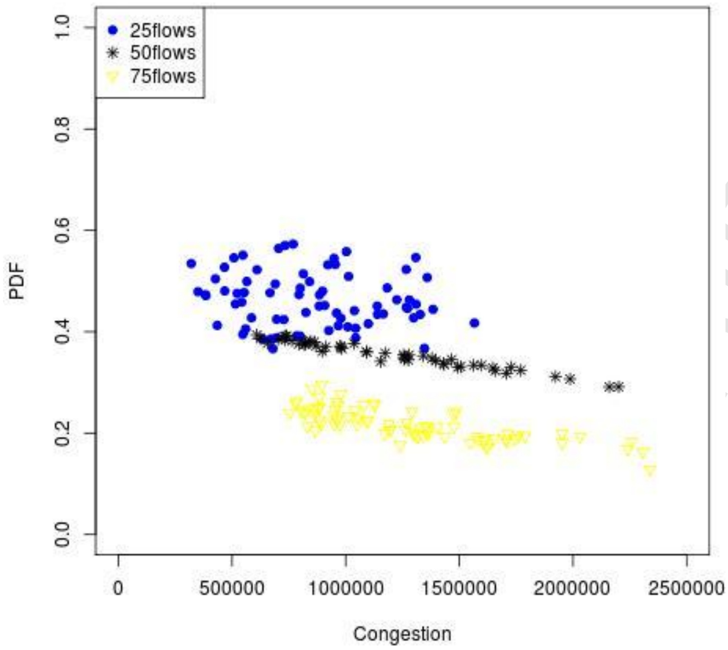


Fig. 6. PDF vs Congestion for different number of flows.

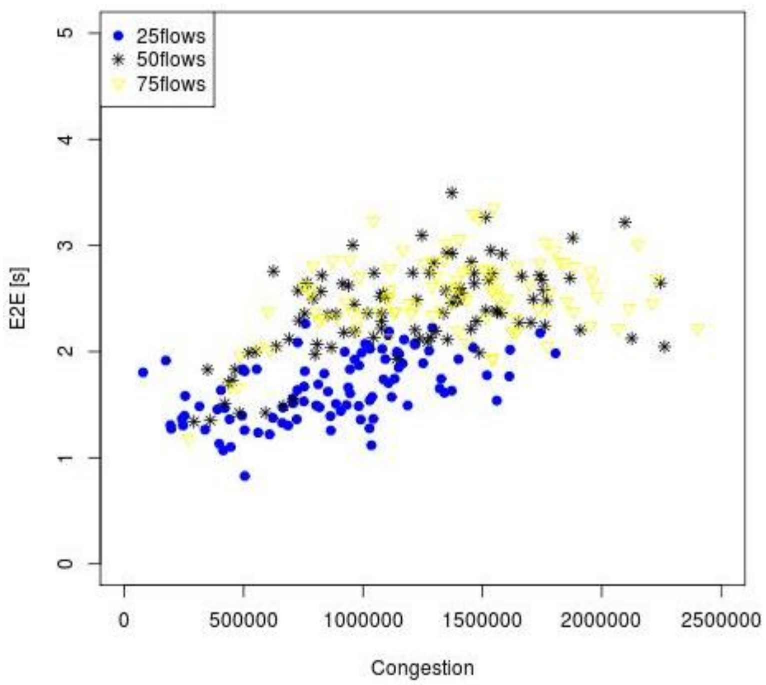


Fig. 7. E2E vs Congestion for different number of flows.

- If we consider *Re* and *PDF*, which measure data delivery at different network layers (*Re* for routing layer and *PDF* for application layer), it can be observed that both output metrics are counterbalanced (see Figure 8). Considering only *Re*, the optimal forwarding will be a forwarding probability of 1 (maximum). However, *PDF* worsens for high forwarding probabilities due to the negative mediation of congestion in the relationship between probability and *PDF*.

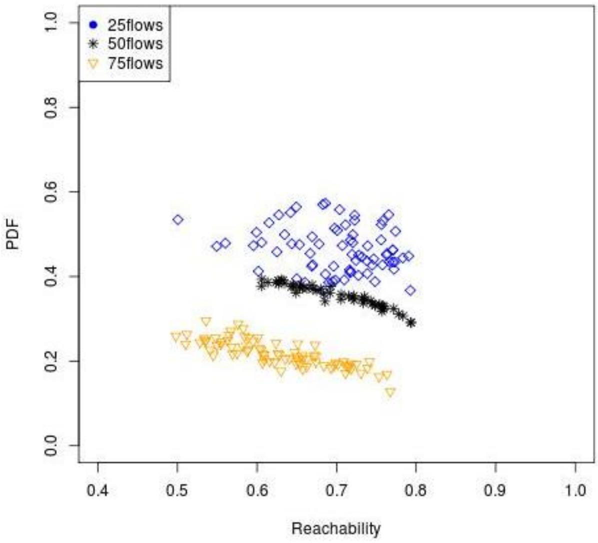


Fig. 8. PDF vs Reachability for different number of flows.

- If we consider *D* and *E2E*, which measure delivery delay at different communication layers (*D* for routing layer and *E2E* for application layer), both metrics decrease with the congestion if nodes increase their forwarding probability. When *D* decreases, *E2E* also decreases and vice versa (see Figure 9).

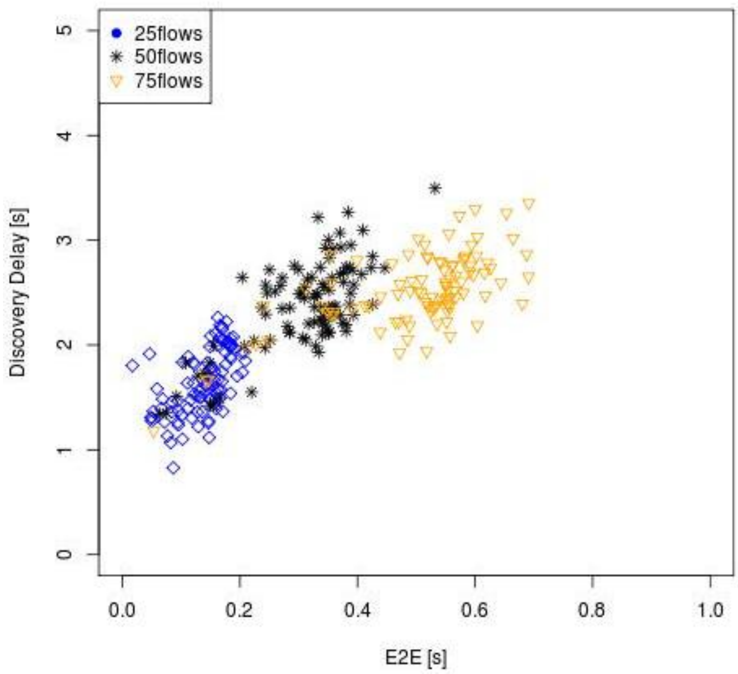


Fig. 9. PDF vs Reachability for different number of flows.

5. Discussion of results and future work

The implications of the presented mediation analyses for designing new broadcasting and routing protocols based on probability for ubiquitous multi-hop ad hoc networks are:

- To achieve a high *Re* is good to increase the forwarding probability. However, until certain point (critical probability of phase transition phenomenon if it happens), because increasing the forwarding probability further than this point will affect negatively the broadcasting delay. Observing Figure 4, the objective is to determine the forwarding probability at which the *Re* saturates. That is, when we notice a change in the increment rate of *Re*.

- Despite a high value of the congestion, the *Re* always increases. However, the increment rate of *Re* gets lower with the increment of congestion. It means that the critical probability predicted by phase transition will also depend on the congestion level of the network. This results corroborated those found in [24], where the author, state that a low value of forwarding probability suffices to achieve high *Re* in determined scenarios.
- Once determined the optimal probability to achieve a high *Re*, it may be good to reduce the forwarding probability to achieve a good performance of the application layer. We will reduce the performance of the routing layer but improving *PDF* and *E2E* as counter effects.

As a future work, we plan to propose a mechanism to adjust the forwarding probability based on the local congestion observed by each node. Consequently, each node determines its optimal forwarding probability according to the congestion in its neighbourhood.

6. Conclusions

We have conducted four mediation analyses considering the main applications of broadcasting in ubiquitous multi-hop networks like MANETs and VANETs (stand-alone dissemination technique and discovery phase of routing protocols) and we have extended these mediation analyses for different levels of congestion. The results indicate an important mediation effect of congestion in the relationship between the forwarding probability and several output metrics, such as *Re*, *D*, *PDF*, and *E2E*. The congestion only mediates in a positive way for *Re*. For *D*, *PDF* and *E2E* metrics, congestion always affects negatively in their values. These mediation effects are stronger with an increment of congestion. The findings of the proposed mediation analyses should be used as a framework for the design of new broadcasting and routing protocols based on probability for ubiquitous multi-hop ad hoc networks.

Acknowledgment

This work was supported in part by the University of Seville under the PhD grant PIF (Personal Investigador en Formación) of Daniel Gutiérrez Reina.

References

1. S. K. Sakar, T.G. Basavaraju, and C. Puttamadappa. Ad hoc mobile wireless networks: Principles protocols, and applications. Auerbach Publications. Taylor & Francis Group. pp. 24, 2008.
2. J. Bernsen and D. Manivannan. Unicast routing protocols for vehicular ad hoc networks: A critical comparison and classification. Pervasive and Mobile Computing, 5 (2009) 1-18.
3. M. Abolhasan, T. Wysocki, and E. Dutkiewicz. A review of routing protocols for mobile ad hoc networks. Ad Hoc Networks, 2 (2004) 1-22.
4. D. G. Reina, S. L. Toral, P. Johnson, and F. Barrero. Improving Discovery Phase of Reactive Ad Hoc Routing Protocols Using Jaccard Distance. Journal of Supercomputing, 67 (2014) 131-152.
5. D. G. Reina and S. L. Toral and F. Barrero and N. Bessis, and E. Asimakopoulou. Modelling and assessing ad hoc networks in disaster scenarios. Journal of Intelligent Ambient and Humanized Computing. 4 (2012) 571-579.

6. D. G. Reina, S. L. Toral, N. Bessis, F. Barrero, and E Asimakopoulou. An evolutionary computation approach for optimizing connectivity in disaster response scenarios. Applied Soft Computing, 13 (2013) 833-845.
7. D. G. Reina, J. M. Leon-Coca, S. L. Toral, E. Asimakopoulou, F. Barrero, P. Norrington, and N. Bessis. Multi-objective performance optimization of a probabilistic similarity/dissimilarity-based broadcasting scheme for mobile ad hoc networks in disaster response scenarios. Soft Computing, 18 (2014) 1745-1756.
8. Y. C. Tseng, S. Y. Ni, Y. S. Chen, and J. P. Sheu. The broadcast storm problem in a mobile ad hoc network. Wireless Networks, 8 (2002)153-167.
9. T. Camp and B. Williams. Comparison of Broadcasting Techniques for Mobile Ad Hoc Networks. Proceeding of the ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 194-205, 2002.
10. I. Stojmenovic, M. Seddigh, and J. Zunic. Dominating Sets and Neighbor Elimination-Based Broadcasting Algorithms in Wireless Networks. IEEE Transaction on Parallel and Distributed Systems, 13 (2002) 14-25.
11. Z. J. Haas, J. Y. Halpern, and L. Li. Gossip-based routing. IEEE/ACM Transaction on Networking, 14 (2006) 479-491.
12. N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar. Broadcast storm mitigation techniques in vehicular ad hoc networks. IEEE Wireless Communications. 14 (2007) 84-94.
13. Y. Mylonas and M. Lestas and A. Pitsillides and P. Ioannou, and V. Papadopoulou. Speed Adaptive Probabilistic Flooding for Vehicular Ad-Hoc Networks. IEEE Transactions on Vehicular Technology. DOI.10.1109/TVT.2014.2339316.

14. D. G. Reina, S. L. Toral, P. Johnson, and F. Barrero. Hybrid Flooding Scheme for Mobile Ad Hoc Networks. IEEE Communication Letters, 17 (2013) 592-595.
15. A. Ghaffari. Congestion control mechanism in wireless sensor networks: A Survey. Journal of Network and Computer Applications. DOI: 10.1016/j.jnca.2015.03.002.
16. C. Sergiou, P. Antoniou, V. Vassiliou. Congestion Control Protocols in Wireless Sensor Networks: A Survey. IEEE Communication Surveys and Tutorials, 16 (2014) 1839-1859.
17. C. Sergio, V. Vassiliou, P. Paphitis. Hierarchical Tree Alternative Path (HTAP) algorithm for congestion control in wireless sensor networks. Ad Hoc Networks, 11 (2013) 257-272.
18. S. Misra, V. Tiwari, M. S. Obaidat. Lacas: learning automata-based congestion avoidance scheme for healthcare wireless sensor networks. IEEE Journal of Selected Areas of Communications. 27 (2009) 466-479.
19. D. A. Tran, H. Raghavendra. Congestion Adaptive Routing in Mobile Ad Hoc Networks. IEEE Transaction on Parallel and Distributed Systems, 17 (2006) 1294-1305.
20. D. P. MacKinnon, A. J. Fairchild, and M. S. Fritz. Mediation analysis. Annual Review of Psychology, 58 (2007) 593-614.
21. D. P. MacKinnon. Introduction to Statistical Mediation Analysis (Multivariate Applications Series). Paperback. 2008.
22. D. G. Reina, S. L. Toral, E. Asimakopoulou, N. Bessis, and F. Barrero. A Dissemination Analysis in Mobile Wireless Ad Hoc Networks using Probabilistic Broadcast. Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS2014), pp. 110-114. 2014.

23. K. Paridel, T. Mantadelis , A. Yasar, D. Preuveneers, G. Janssens, Y. Vanrompay, Y. Berbers. Analyzing the efficiency of context-based grouping on collaboration in VANETs with large-scale simulation. Journal Ambient Intelligence Humanized Computing, 5 (2014) 475-490. 2014.

24. R. Albert and A. Barabási. Statistical mechanics of complex networks. Reviews of modern physics, 74 (2002) 47-97.

25. F. Palmieri. Scalable service discovery in ubiquitous and pervasive computing architectures: A percolation-driven approach. Future Generation Computer Systems. 29 (2013) 693-703.

26. Y. Sasson, D. Cavin, and A. Schiper. Probabilistic Broadcast for Flooding in Wireless Mobile Ad Hoc Networks. In IEEE Wireless Communications and Networking Conference, pp. 1124-1130. 2003.

27. B. Blywis, P. Reinecke, M. Günes, and K. Wolter. Gossip Routing, Percolation, and Restart in Wireless Multi-hop Networks. IEEE Wireless Communication and Networking Conference, pp. 3019-3023. 2012.

28. B. Blywis, M. Günes, F. Juraschek, O. Hahm, and N. Schmittberger. Properties and Topology of DES-Testbed (2nd Extended Revision). Telematic and Computer System. Freie Universitát Berlin. 2011.

29. A. J. Ghandour, M. Di Felice, H. Artail, and L. Bononi. Dissemination of safety messages in IEEE 802.11p/WAVE vehicular network: Analytical study and protocol enhancements. Pervasive and Mobile Computing 11 (2014) 3-18.

30. R. M. Baron and D. A. Kenny. The moderator-mediator variabledistinction in social pysichology research: conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51 (1986) 1173-1182.

31. S. Kurkowski, W. Navidi, and T. Camp. Constructing manet simulation scenarios that meet standards. In IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems (MASS 2007), pp. 1-9, 2007.

32. C. E. Perkins and E. M. Royer. Ad Hoc on-demand distance vector routing. In Procedding of IEEE Workshop on Mobile Computing System and Applications (WMCSA), pp. 1-11. 1999.

33. R: A Language and Environment for Statistical Computing. Available at http://www.R-project.org.