RESEARCH   **Open Access**

Asymptotic analysis of flooding in CSMA-based large scale ad-hoc wireless networks

Hamed Shah-Mansouri[1], Babak Hossein Khalaj[1], Seyed Pooya Shariatpanahi[1],
Javier Del Ser[2*] and Susana Pérez-Sánchez[2]

Abstract

In this article, we study the asymptotic behavior of flooding in large scale wireless networks. Specifically, we derive an upper bound on the coverage of flooding when the number of nodes n in the network goes to infinity. We consider two different regimes of transmission radii: first, the case of constant transmission radius r where the percentage of covered nodes scales as $O(nr^2 e^{-K_S nr^2})$ for a constant $K_S > 0$. In this case, as an important result, we observe that the percentage of covered nodes is upper bounded by a decreasing function, vanishing as the network size grows. Second, the case of vanishing r_n (i.e., r decreases as n increases) is considered where it is shown in the literature that the minimum value of r_n which maintains connectivity is $\sqrt{\log n / \pi n}$. In this case, a coverage percentage of at most $O(n^{-K'_S} \log n)$ is expected for a constant value of $K'_S > 0$, leading to an infinite number of covered nodes. In such case, the rate at which the network coverage is decreased can be controlled and be considerably reduced by a proper choice of network parameters (K'_S). Consequently, this result shows that flooding is a suitable strategy even for large networks.

Keywords: Asymptotic analysis, Scaling law, Wireless network

Introduction

Flooding is the simplest and most widely used form of broadcasting in wireless networks where each node simply retransmits received data once to its neighboring nodes [1,2]. Data dissemination and route discovery process in many applications are based on flooding [3-6]. Wide adoption of flooding for broadcasting traffic is mainly due to its simplicity. In addition, it does not require any centralized information about the network. In probabilistic flooding—a variant of flooding introduced in [7]—each node retransmits received data with a pre-set probability called forwarding factor in order to reduce unnecessary retransmissions. After the seminal study of Gupta and Kumar in [8], the scaling laws and fundamental limits of large wireless networks have received much attention. However, to the best of authors' knowledge, no study has been conducted in the literature with the goal of studying the scaling laws of flooding. The importance of such

analysis lies on the aforementioned popularity of flooding for data dissemination and route discovery in many routing protocols [3-6]. Based on this fact, the main motivation of this study is to effectively answer the question that whether flooding is suitable and applicable to large scale networks or not.

Previous related studies in [9-11] proposed a tight upper bound on the coverage of flooding in a multi-hop network using carrier sense multiple access (CSMA) as the MAC layer strategy for any finite network size. CSMA is a well-known multiple access algorithm for broadcasting in wireless networks [1]. It should be noted that implementing other MAC layer algorithms that require centralized information of the network nodes is not feasible in ad-hoc wireless networks. In addition, utilizing collision avoidance techniques in MAC layer for broadcasting traffic is not practical, since it generally requires a large amount of control message overhead to ensure that all neighbors successfully receive the broadcasted packet.

In this article, we take a step further by investigating the scaling law and asymptotic behavior of flooding in ad-hoc wireless networks based on the framework introduced in [9-11]. The major contribution of this article

---

*Correspondence: javier.delser@tecnalia.com
[2]Tecnalia Research & Innovation, 48170 Zamudio, Spain
Full list of author information is available at the end of the article

in comparison with the results presented in [9] is providing an analytic approach to asymptotic behavior of flooding. Although [9] provides important results on the coverage of flooding, its results are only obtained numerically through a recursive scheme. On the other hand, the analytic framework proceeded in this article leads to closed form representations of coverage upper bound in large networks. Therefore, the derived closed form results clearly demonstrate the behavior of coverage and provide deeper insight on the performance of flooding when the number of network nodes grows.

In order to analyze the coverage of flooding, first, we assume that only one of the nodes intends to broadcast its packets. Subsequently, a discussion on the case of multiple sources is presented. Our analysis includes two different cases: the case where the transmission radius is constant, and the case where the transmission radius r approaches zero as the number of nodes goes to infinity. Although flooding is in general known to be a simple yet inefficient broadcasting method [1], we obtain the following interesting results throughout the article:

- Under the assumption of constant transmission radius, network coverage is upper bounded by a decreasing function rapidly vanishing as the network size grows. Therefore, flooding is not a proper broadcasting scheme in such conditions.
- When the transmission radius approaches zero as network size grows, we show that a coverage bound (i.e., the percentage of covered nodes) close to 100% can be obtained in realistic large networks. This result, along with the tightness of the bound, demonstrates that flooding is a suitable and applicable broadcasting scheme in terms of network coverage even for large networks.

It should also be noted that, in addition to the aforementioned interesting results, the analytic framework presented in this article can provide proper basis for asymptotic analysis of other broadcasting schemes. A number of more complex broadcast methods are already proposed in the literature, such as counter-based, distance-based, location-based broadcasting [1], and multi point relaying (MPR) [12], with the goal of improving overall network efficiency. Although the analysis of such schemes appears to be more complicated than simple flooding, we believe that this study, as the first research in this field, can be extended to analyze them as well. It is important to note that no complete and comprehensive analytic model for performance analysis of such broadcasting schemes in the presence of a practical MAC layer protocol currently exists, even in the case of finite networks.

The rest of the article is organized as follows: in Section 'Notations and system model', we introduce the network

model, MAC layer model, and an exact definition of flooding and network coverage. Then, we restate briefly the results obtained in [9] in Section 'Coverage upper bound'. The asymptotic behavior of the upper bound on the coverage is derived in Section 'Asymptotic analysis' along with a discussion on the obtained results. Finally, Section 'Conclusions' concludes the article.

Notations and system model
Network model
Network nodes are distributed on a unit-radius sphere $\mathcal{R}$ according to a Poisson point process Φ with intensity measure $\Lambda = \lambda\mu$ representing the mean density of points, where $\lambda > 0$ is the intensity of the uniform Poisson point process and μ is the Lebesgue measure with total mass 4π. Figure 1 shows the spherical model adopted for our network area. It should be noted that symmetric properties of the sphere eliminate the edge-effect arising in planar models and, therefore, leads to a model in which all regions of the network will be governed by the same process. All derivations can then be applied to a two-dimensional network if the edge-effect in such network is ignored. Such model has been used by several articles in the literature, e.g., [8,13]. In large scale networks, we assume that $\lambda \to \infty$. Consequently, the mean number of nodes n goes to infinity. Let V and X_i, respectively, denote the set of nodes in the network and their locations. We will then have $\Phi = \{X_i; \forall X_i \in \mathcal{R}, \forall i \in V\}$. We refer to a given Φ as a network realization. Nodes i and j placed at
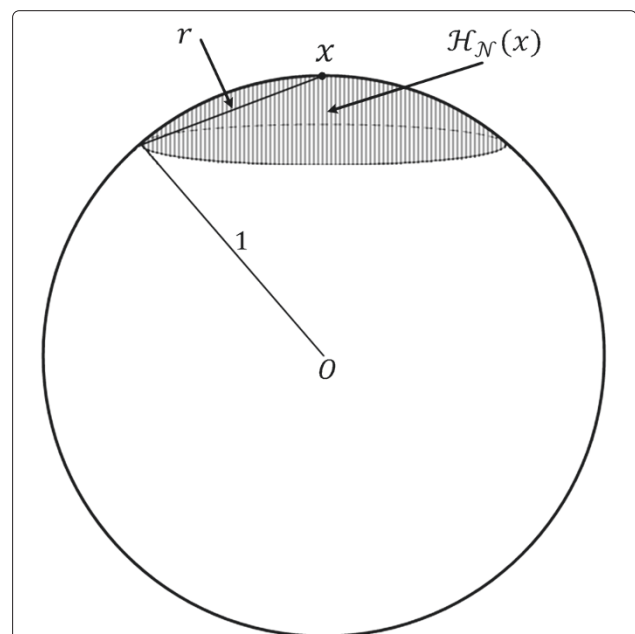

Figure 1 Network nodes are assumed to be distributed over a unit-radius sphere. The shaded area shows the neighboring region of point $\boldsymbol{x}$ [9].

points x and y are neighbors if their distance is less than the transmission range of the nodes, which is denoted by r. Since we have assumed that network nodes are distributed on a unit-radius sphere, r is the normalized value of the transmission radius. Let $\mathcal{H}_{\mathcal{N}}(x)$ denote the neighboring region around x, where $\mu(\mathcal{H}_{\mathcal{N}}(x)) = \pi r^2$ is the Lebesgue measure of this region on the unit sphere. The shaded area in Figure 1 shows the neighboring region of point x. The wireless transmissions are modeled according to the protocol model introduced in [8]. The interference region $\mathcal{H}_{I}(x)$ is defined as the region in which each point lies in a distance less than $r_{\Delta} = (1 + \Delta)r$ from x, where Δ is the constant factor of the protocol model. The transmission from node i to node j is successful if the following conditions are satisfied:

- The distance between nodes i and j is no more than r, i.e., j lies in the neighboring region of node i.
- For every other node k simultaneously transmitting over the channel, the distance between k and j is more than $(1 + \Delta)r$, i.e., k is located outside the interference region of node j.

It should be mentioned that for the sake of tractability, in this study all results are derived considering uniform transmissions, even though the proposed framework can be extended to analyze network scenarios with non-uniform transmissions. The non-uniform transmission assumption would affect neighboring and interference regions. It is further assumed that, under saturated condition, all nodes have packets for transmission and one of them intends to broadcast its packet. The source node is selected randomly.

MAC layer model

As mentioned earlier, CSMA is used as the MAC layer strategy, since collision avoidance techniques require a large amount of control message between transmitter and all neighboring nodes. Hence, in applications such as broadcasting which, by nature, involve a large number of transmissions, implementing collision avoidance techniques degrades the performance of MAC layer strategy. In the CSMA protocol, each silent node with packets ready for transmission senses the channel and transmits with probability p if the channel is idle, and postpones the channel sensing and transmission to the next time slot with probability $(1 - p)$. In the case of busy channel, the process will be postponed to the next time slot. Therefore, the probability of transmission from a node denoted by p' is equal to the probability that the channel is idle multiplied by p. The transmitted packet from a node placed at point x will be successfully received at a node placed at point y with probability $P_S(d_{x,y})$. Collisions between simultaneously transmitted packets reduce

network coverage, since CSMA does not benefit from any collision avoidance or re-transmission scheme. The mentioned parameters are carefully defined and computed in [9] with the help of the study presented in [14], which are as follows:

$$p' = p \cdot \text{Pr}\{\text{channel is idle at a given slot}\}$$
$$= \frac{\alpha p}{1 + \alpha - e^{-p'\overline{\mathcal{N}_I}}}, \tag{1}$$

$$P_S(d_{x,y}) = (1 - p')e^{p'\lambda C(d_{x,y})2\tau}e^{-p'\lambda\pi r_\Delta^2(2\tau+1)}, \tag{2}$$

in which $\alpha = r/c$ is the one-way propagation delay (with c denoting the speed of light), and $\mathcal{N}_I = \lambda\mu(\mathcal{H}_I(x))$ is the mean number of nodes in the interference region of x. In addition, $C(d_{x,y})$ is the intersection area of interference regions around points x and y, as determined in [14] for a planar network and in [9] for our sphere model. Figure 2 shows $C(d_{x,y})$ for a planar network. Moreover, we have $\tau = T/\alpha$, where T is the packet transmission time. For the sake of simplicity, T is assumed to be equal to 1, and all other time parameters are normalized with respect to T. Furthermore, $\overline{P_S}$ is defined as the average of $P_S(d_{x,y})$, and is equal to

$$\overline{P_S} = \frac{1}{\mu(\mathcal{H}_{\mathcal{N}}(x))} \int_{\mathcal{H}_{\mathcal{N}}(x)} P_S(d_{x,y})\mu(dy). \tag{3}$$

These parameters can be easily computed using the results presented in [9,14] and depend on the basic parameters of the network such as the CSMA transmission probability p, packet size τ, network size n, transmission radius r, and interference region factor Δ as well as network topology (planar versus sphere). The effect of transmission model is clear in calculating the MAC layer parameters. In Equations (1) and (2) the interference region is considered as the area from which the interference can make the channel busy or disturb the packet transmission. However, in calculation of $\overline{P_S}$, only transmissions from neighboring regions are considered potentially as successful transmissions.
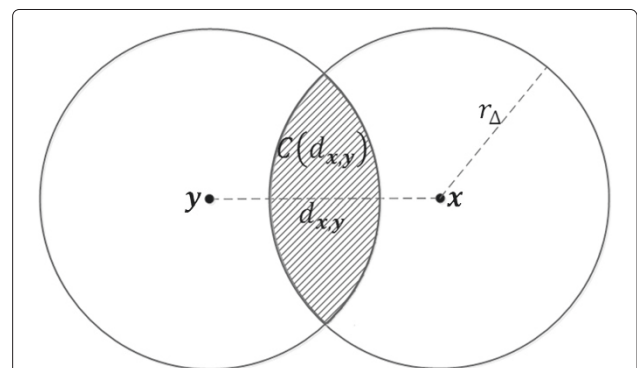

Figure 2 $C(d_{xy})$ shows the intersection area of interference regions around points x and y in a planar network.

It should be noted that the analysis framework used in this article only requires the two aforementioned parameters of MAC layer, i.e., transmission probability p' and successful transmission probability $\overline{P_S}$. Therefore, our model can be simply extended to any other MAC layer protocol or any different definition of CSMA protocol by recalculating these parameters for desired MAC model and substituting them in our analytic model.

Probabilistic flooding and network coverage

Before introducing the analytic framework, an exact definition of flooding and probabilistic flooding should be provided. In *probabilistic flooding*, flooding packets are generated by a source node. Each intermediate node receiving a new packet from a neighbor rebroadcasts it with a preset probability, which is generally called forwarding factor and denoted herein by P_{ff}. On the other hand, *simple flooding* is a special case of probabilistic flooding with $P_{ff} = 1$.

Our goal is to analyze the flooding scheme in terms of network coverage in large wireless networks. Network coverage is defined as the average percentage of nodes which successfully receive the flooded packet when with probability 1 the flooding phase is completed (i.e., none of the nodes has flooding packets waiting for rebroadcast). Although network coverage of 100 % is desired from a broadcasting scheme, packet collision in the shared wireless channel reduces the coverage. We denote the network coverage for network realization Φ by c_Φ. In addition, $\overline{c}$ denotes the average value of network coverage over all network realizations while $\overline{\mathcal{C}}$ is its upper bound.

Coverage upper bound

In this section, we briefly review the results obtained in [9]. We use this framework as a basis of our analysis later in this article[a]. In order to calculate the network coverage, we first introduce the analytic framework presented in [9]. We mark each node of the network realization Φ by a random variable denoted by ζ_{X_i} denoting whether node $i \in V$ has received the broadcasted packet or not. The marking random variable is as follows:

$$\zeta_{X_i} = \begin{cases} 1 & \text{with probability} \quad \psi(X_i), \\ 0 & \text{with probability} \quad 1 - \psi(X_i) \end{cases} \tag{4}$$

where $\psi(X_i)$ is the probability that node i placed at X_i receives the broadcasted packet when the flooding process is completed. The network coverage for network realization Φ can be determined by computing the ratio of covered nodes to total number of nodes as follows:

$$c_\Phi = \frac{\sum_{X_i \in \Phi} E\{\zeta_{X_i}\}}{n} = \frac{\sum_{X_i \in \Phi} \psi(X_i)}{n} \tag{5}$$

In (5), we have used the fact that $E\{\zeta_{X_i}\} = \psi(X_i)$ which shows the expectation of reception. This equation represents the ratio of average number of nodes that received the broadcasted packet to the total number of nodes for a given Φ. Let $\overline{\psi}(X_i)$ denote the average of $\psi(X_i)$ taken over the number and location of neighbors of node i. By averaging over mentioned random values and different network realizations Φ, we have

$$\overline{c} = \frac{\int_{\mathcal{R}} \overline{\psi}(\boldsymbol{x}) \Lambda(d\boldsymbol{x})}{n} = \frac{\int_{\mathcal{R}} \overline{\psi}(\boldsymbol{x}) \mu(d\boldsymbol{x})}{4\pi}. \tag{6}$$

Equation (6) presents a simple summation over the surface of sphere, where $\overline{\psi}(\boldsymbol{x})$ denotes the average probability of reception at point $\boldsymbol{x} \in \mathcal{R}$ and $n = 4\pi\lambda$ is the total number of nodes. In order to calculate $\overline{\psi}(\boldsymbol{x})$, we introduce another function $\overline{u}_h(\boldsymbol{x})$ which is defined over hop-counter h. We assume that each packet is marked by a hop-count parameter h. This value is set to zero upon generation at the source and is incremented by intermediate nodes after rebroadcasting the packet. Consequently, its value shows the number of rebroadcasting times experienced by the packet from its origination. When $h \rightarrow \infty$, the flooding process is completed with probability 1. The function $\overline{u}_h(\boldsymbol{x})$, namely the coverage function, determines the probability that a node placed at $\boldsymbol{x}$ receives the flooding packet with a hop-count equal to or less than h. Therefore, $\overline{u}_h(\boldsymbol{x})$ will be equal to $\overline{\psi}(\boldsymbol{x})$ when h goes infinity, i.e., when the flooding process is completed surely and we have

$$\overline{c} = \lim_{h \to \infty} \overline{c}_h = \lim_{h \to \infty} \frac{\int_{\mathcal{R}} \overline{u}_h(\boldsymbol{x}) \mu(d\boldsymbol{x})}{4\pi}, \tag{7}$$

where $\overline{c}_h$ is the network coverage obtained until the hth hop-count of the packet. In other words, we only count the number of nodes that receive the flooding packet by a hop-count equal to or less than h to calculate $\overline{c}_h$. Let $\overline{\mathcal{U}}_h(\boldsymbol{x})$ denote the upper bound on $\overline{u}_h(\boldsymbol{x})$, i.e., we have $\overline{u}_h(\boldsymbol{x}) \leq \overline{\mathcal{U}}_h(\boldsymbol{x}) \, \forall h > 0$ and $\forall \boldsymbol{x} \in \mathcal{R}$. Therefore, the upper bound on the network coverage denoted by $\overline{\mathcal{C}}$ is obtained as

$$\overline{\mathcal{C}} = \lim_{h \to \infty} \overline{\mathcal{C}}_h = \lim_{h \to \infty} \frac{\int_{\mathcal{R}} \overline{\mathcal{U}}_h(\boldsymbol{x}) \mu(d\boldsymbol{x})}{4\pi}, \tag{8}$$

where $\overline{\mathcal{C}}_h$ is the upper bound of $\overline{c}_h$. $\overline{\mathcal{U}}_h(\boldsymbol{x})$ can be calculated using (9) when the Hölder inequality[b] parameter b—which is independent of other parameters and is explained in detail in [9]—tends to infinity, i.e.,

$$\overline{\mathcal{U}}_h(\boldsymbol{x}) = 1 - (1 - I(\boldsymbol{x})) e^{-\lambda \mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})) P_{ff} \overline{P_S} \left(\int_{\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})} \overline{\mathcal{U}}_{h-1}^b(\boldsymbol{z}) \mu(d\boldsymbol{z}) \right)^{\frac{1}{b}}}, \tag{9}$$

where P_{ff} is the aforementioned forwarding factor of probabilistic flooding which is equal to 1 for simple flooding; and $\overline{P_S}$ is the average of $P_S(d_{\boldsymbol{x},\boldsymbol{y}})$, which is the probability of successful packet transmission from a node located at $\boldsymbol{x}$

to its neighbor located at $\boldsymbol{y}$. In Expression (9), $I(x)$ is equal to $P_S(d_{X_S,\boldsymbol{x}})$, where the source node is placed at X_S while $I(x) = 0$ for $\boldsymbol{x} \in \mathcal{R} \setminus \mathcal{H}_{\mathcal{N}}(X_S)$. Note that (9) represents a recursive function whose initial condition is $\overline{\mathcal{U}}_{-1}(\boldsymbol{x}) = 0$.

The Equation (9) is obtained using a rigorous analysis presented in detail in [9]. The simulation results show that the obtained upper bound is quite tight and reliable. In addition, the bound gets tighter when n increases. It should be noted that in this article, we only require this equation which is the final result of [9]. More carefully, we first calculate the MAC layer parameters in large networks and substitute the results in (9). Then, we focus on deriving the asymptotic behavior of (9) under two different regimes of transmission radius.

Asymptotic analysis
Main results

The main result of this article is the scaling law of the flooding coverage in large networks. We use asymptotic Knuth's notations [15] to show the coverage upper bound. In Knuth's notation, upper bound is denoted by O($\cdot$). We denote the upper bound of coverage by O(f) for any function f if a network coverage greater than f is not possible for flooding. In other words, the coverage obtained by the flooding is bounded above by and scales as O(f) if for any constant $1 < a < \infty$ and for any function f, a coverage of af is not achievable. We study the problem under two scenarios: in the first case, we assume that the transmission radius r is constant and independent of n, as denoted by $r = \Theta(1)$. In Knuth' notations, $\Theta(\cdot)$ is used for tight bound. Therefore, $r = \Theta(1)$ shows that r is a constant factor. In the second case, the transmission radius denoted by r_n is assumed to rapidly diminish as n increases. The following theorems characterize the scaling law for these two cases. The proofs are presented in the following sections.

Theorem 1. (Main result 1) *The network coverage of the flooding is upper bounded by* O($nr^2 e^{-K_S nr^2}$) *for a constant value of K_S when r is constant, i.e., $r = \Theta(1)$.*

Theorem 2. (Main result 2) *The network coverage of the flooding is bounded above by* O($1 - e^{-Knr^2 e^{-K_S nr^2}}$) *for constants K and K_S when $r_n \xrightarrow{n \to \infty} 0$. In order to keep connectivity, r_n should be $\sqrt{\log n / \pi n P_{\text{ff}}}$ or greater [8]. In the case of $r_n = \sqrt{\log n / \pi n}$ for simple flooding, the network coverage scales as* O($\log n / n^{K'_S}$). *Consequently, the number of covered nodes scales as* O($n \log n / n^{K'_S}$) *for a constant $K'_S > 0$.*

It can be concluded that the coverage bound and the number of covered nodes quickly approach zero when $r = \Theta(1)$. However, for the case of $r_n = \sqrt{\log n / \pi n}$, convergence to 0 is much slower and can be controlled

by a proper choice of the network parameters. In such condition, coverage bound greater than zero can be obtained for a large n by choosing the parameter set such that K'_S will be very small. In addition, an infinite number of nodes can be covered when $K'_S < 1$. As we will see in 'Proof of Theorem 2', K'_S is much smaller than 1. It should be noted that r_n is the normalized value of transmission radius since it is assumed that network nodes are distributed on a unit-radius sphere.

The following sections contain the proofs of these theorems. The main strategy of the proofs is similar for both theorems. First, the MAC layer parameters p' and $\overline{P_S}$ are determined asymptotically and substituted in (9). Then, Expression (9) is re-calculated by assuming that $n \to \infty$. Finally, the network coverage is analyzed asymptotically using the obtained results. The main difference between the proofs of the theorems is in calculating the integral term in the exponent part of (9).

Proof of Theorem 1

We should first determine the MAC layer parameters asymptotically. The transmission probability of nodes in CSMA protocol given by (1) will be

$$\lim_{n \to \infty} p' = \lim_{n \to \infty} \frac{\alpha p}{1 + \alpha - e^{-p' \overline{\mathcal{N}_I}}} = \frac{\alpha p}{1 + \alpha} \cong \alpha p = \frac{rp}{c}.$$

(10)

The parameter p' appears on both sides of the first equality in (10) which requires an iterative method to be calculated. However, the second equality is obtained based on the facts that $p' \geq \alpha p / (1 + \alpha) > 0$ and $\overline{\mathcal{N}_I} \to \infty$ when n goes infinity. Therefore, p' can be readily calculated in asymptotic regime. It should be noted that $\alpha \ll 1$. In (2), $C(d_{\boldsymbol{x},\boldsymbol{y}})$ is the intersection area of interference regions around points $\boldsymbol{x}$ and $\boldsymbol{y}$ and, therefore, is proportional to the interference region. Consequently, by letting $k_{\boldsymbol{x},\boldsymbol{y}} = C(d_{\boldsymbol{x},\boldsymbol{y}}) / (\pi r_\Delta^2)$ we can rewrite Expression (2) as

$$P_S(d_{\boldsymbol{x},\boldsymbol{y}}) \cong e^{-\frac{rp}{c} \frac{n}{4\pi} \pi r_\Delta^2 2\tau \left(1 - \frac{C(d_{\boldsymbol{x},\boldsymbol{y}})}{\pi r_\Delta^2}\right)} \triangleq e^{-k_S(1 - k_{\boldsymbol{x},\boldsymbol{y}})nr^2}.$$

(11)

By substituting (11) in (3), we have

$$
\begin{aligned}
\overline{P_S} &= e^{-k_S nr^2} \int_{\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})} \frac{e^{k_S nr^2 k_{\boldsymbol{x},\boldsymbol{y}}}}{\mu\left(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})\right)} \mu(d\boldsymbol{y}) \\
&= e^{-k_S nr^2} \int_{\mathcal{H}_{\mathcal{N}}^1(\boldsymbol{x})} e^{k_S nr^2 k_{\boldsymbol{x},\boldsymbol{y}}^1} \mu(d\boldsymbol{y}) \\
&\leq e^{-k_S nr^2} \left(\int_{\mathcal{H}_{\mathcal{N}}^1(\boldsymbol{x})} e^{k_{\boldsymbol{x},\boldsymbol{y}}^1} \mu(d\boldsymbol{y})\right)^{k_S nr^2} \triangleq e^{-K_S nr^2},
\end{aligned}
$$

(12)

where $\mathcal{H}_{\mathcal{N}}^1(\boldsymbol{x})$ and $k_{\boldsymbol{x},\boldsymbol{y}}^1$ are normalized values of $\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})$ and $k_{\boldsymbol{x},\boldsymbol{y}}$ when $r = 1$, and the inequality comes from the Cauchy-Schwarz inequality. In addition, $k_S = 0.5p(1+\Delta)^2$ which is very less than 1 in typical applications, and

$$K_S = k_S \left(1 - \ln\left(\int_{\mathcal{H}_{\mathcal{N}}^1(\boldsymbol{x})} e^{k_{\boldsymbol{x},\boldsymbol{y}}^1}\mu(d\boldsymbol{y})\right)\right) \ll 1 \qquad (13)$$

are constants. K_S can be readily calculated given that $k_{\boldsymbol{x},\boldsymbol{y}}^1 = C^1(d_{\boldsymbol{x},\boldsymbol{y}})/(\pi(1 + \Delta)^2)$, where $C^1(d_{\boldsymbol{x},\boldsymbol{y}})$ is the intersection of interference regions of nodes placed at distance $d_{\boldsymbol{x},\boldsymbol{y}}$ when it is assumed that $r = 1$. K_S is a constant factor that only depends on network topology (planar versus sphere).

It is proved that $\overline{\mathcal{C}}_h$ converges to $\overline{\mathcal{C}}$ when h goes to infinity, leading to convergence of $\overline{\mathcal{U}}_h(\boldsymbol{x})$ [9, Appendix C]. Therefore, there exists H such that $\forall h > H$, we have $\overline{\mathcal{U}}^n(\boldsymbol{x}) \triangleq \overline{\mathcal{U}}_h(\boldsymbol{x}) \cong \overline{\mathcal{U}}_{h-1}(\boldsymbol{x})$. Now, we replace $\overline{\mathcal{U}}^n(\boldsymbol{x})$ in both sides of (9), yielding

$$\overline{\mathcal{U}}^n(\boldsymbol{x}) = 1 - (1 - I(\boldsymbol{x}))e^{-\lambda\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x}))P_{\text{ff}}\overline{P}_S\left(\int_{\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})}\left(\overline{\mathcal{U}}^n(\boldsymbol{z})\right)^b \mu(d\boldsymbol{z})\right)^{1/b}}.$$
$$(14)$$

We simplify the exponent term in the right side of (14) by first approximating $\overline{\mathcal{U}}^n(\boldsymbol{z})$ over $\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})$ by $\overline{\mathcal{U}}^n(\boldsymbol{x})$. Since $r^2 \ll 1$ in typical applications and $\overline{\mathcal{U}}^n(\boldsymbol{z})$ is a smooth function, this is a valid approximation. Therefore,

$$\left(\int_{\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})}\left(\overline{\mathcal{U}}^n(\boldsymbol{z})\right)^b \mu(d\boldsymbol{z})\right)^{1/b} \approx \overline{\mathcal{U}}^n(\boldsymbol{x})\left(\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x}))\right)^{1/b}.$$
$$(15)$$

Since $\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})) > 0$, the above expression is equal to $\overline{\mathcal{U}}^n(\boldsymbol{x})$ as $b \to \infty$. Therefore,

$$\overline{\mathcal{U}}^n(\boldsymbol{x}) = 1 - (1 - I(\boldsymbol{x}))e^{-Knr^2\overline{P}_S\overline{\mathcal{U}}^n(\boldsymbol{x})}$$
$$\leq 1 - (1 - I(\boldsymbol{x}))e^{-Knr^2 e^{-K_S nr^2}}, \qquad (16)$$

where $K = P_{\text{ff}}/4$. The inequality comes from the fact that $\overline{\mathcal{U}}^n(\boldsymbol{x}) \leq 1$. Such simplification ensures upper bound condition and $\overline{\mathcal{U}}^n(\boldsymbol{x})$ is replaced by the right side of inequality in (16) as the new upper bound on the coverage function.

Now, based on the obtained results for the coverage function $\overline{\mathcal{U}}^n(\boldsymbol{x})$, we can calculate the network coverage $\overline{\mathcal{C}}$. By using (8), the network coverage can be simply equal to the average of $\overline{\mathcal{U}}^n(\boldsymbol{x})$ over the sphere. $\overline{\mathcal{C}}$ is also in the form of (16) since the integration in (8) does not change the scaling of Equation (16) and its relationship with n.

As $I(\boldsymbol{x}) = P_S(d_{X_S,\boldsymbol{x}})$ for $\boldsymbol{x} \in \mathcal{H}_{\mathcal{N}}(X_S)$ and zero for $\boldsymbol{x} \in \mathcal{R} \setminus \mathcal{H}_{\mathcal{N}}(X_S)$, it is clear that $\overline{\mathcal{U}}^n(\boldsymbol{x})$ is greater for $\boldsymbol{x} \in \mathcal{H}_{\mathcal{N}}(X_S)$ than for other nodes. This fact can be also concluded intuitively, since the nodes closer to source node have more chance of being covered. If we assume that the number of nodes in $\mathcal{H}_{\mathcal{N}}(X_S)$ is at most $A(n)$ and the

percentage of covered nodes in this area is $\overline{\mathcal{C}}_{\boldsymbol{x}\in\mathcal{H}_{\mathcal{N}}(X_S)}$, we have that

$$\overline{\mathcal{C}} \leq \frac{A(n)\overline{\mathcal{C}}_{\boldsymbol{x}\in\mathcal{H}_{\mathcal{N}}(X_S)} + (n - A(n))\,\overline{\mathcal{C}}_{\boldsymbol{x}\in\mathcal{R}\setminus\mathcal{H}_{\mathcal{N}}(X_S)}}{n}.$$
$$(17)$$

We know that $\overline{\mathcal{C}}$ has a form similar to (16). Consequently, $\overline{\mathcal{C}}_{\boldsymbol{x}\in\mathcal{H}_{\mathcal{N}}(X_S)}$ is also in the form of (16) when $I(\boldsymbol{x}) = P_S(d_{X_S,\boldsymbol{x}}) = \mathrm{O}(e^{-K_S nr^2})$ and $\overline{\mathcal{C}}_{\boldsymbol{x}\in\mathcal{R}\setminus\mathcal{H}_{\mathcal{N}}(X_S)}$ is in the form of (16) for $I(\boldsymbol{x}) = 0$. The following lemma determines $A(n)$.

Lemma 1. *The number of nodes in $\mathcal{H}_{\mathcal{N}}(X_S)$, denoted by $\mathcal{N}$, is at most $A(n) = A(e - 1)r^2 n/4$ for any $A > 1$ with high probability.*

Proof. Since nodes are Poisson distributed with mean $\lambda\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})) = nr^2/4$, by applying the Chernoff bound we have $\Pr\{\mathcal{N} > A(n)\} \leq \mathbf{E}\{e^{\mathcal{N}}\}/e^{A(n)}$. Since $\mathbf{E}\{e^{\mathcal{N}}\} = e^{n(e-1)r^2/4}$, as long as $A > 1$, $\Pr\{\mathcal{N} > A(n)\}$ tends to zero as n goes to infinity. $\square$

Therefore, the number of source's neighbor nodes is at most $A(e - 1)r^2 n/4$ with high probability. By substituting $A(n)$ from Lemma 1 into (17), network coverage bound scales as (18) and main result 1 is obtained as follows, using the fact that $e^y \cong 1 + y$ for small values of y:

$$\overline{\mathcal{C}} = \mathrm{O}\left(1 - \left(1 - \frac{A(e-1)r^2}{4}e^{-K_S nr^2}\right)e^{-Knr^2 e^{-K_S nr^2}}\right)$$
$$\cong \mathrm{O}\left(Knr^2 e^{-K_S nr^2} + \frac{A(e-1)r^2}{4}e^{-K_S nr^2}\left(1 - Knr^2 e^{-K_S nr^2}\right)\right)$$
$$\cong \mathrm{O}\left(nr^2 e^{-K_S nr^2}\right). \blacksquare$$
$$(18)$$

It should be noted that for large n, the last term in above expression approaches zero for any constant K_S.

Proof of Theorem 2
The analysis for this case is similar to the case presented in the previous section. The major difference is in the way that the term $\left(\int_{\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})}\left(\overline{\mathcal{U}}^n(\boldsymbol{z})\right)^b \mu(d\boldsymbol{z})\right)^{1/b}$ in (14) is simplified. In the case of the previous solution, $(\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})))^{1/b}$ equals 1 since $\mu(\mathcal{H}_{\mathcal{N}}(\boldsymbol{x})) > 0$. However, when $r_n \to 0$ this term (0^0) is undefined, thus we can not reuse the same procedure used in the proof of Theorem 1. Therefore, we should re-calculate this term when n goes to infinity by assuming that r_n approaches zero.

Lemma 2. *For a measurable function f, $||f||_\infty =$* $\lim_{b \to \infty} \left(\int |f|^b d\mu \right)^{1/b}$ *will be equal to the essential supremum of $|f|$ denoted by ess sup $|f|$.*

Proof. The proof of this lemma is presented in [16].

Essential supremum of f is the greatest value of f which occurs on a set that is not measure zero. Note that we always have ess sup $f \leq$ sup f since supremum can occur on a set of measure zero. In order to clarify the difference between supremum and essential supremum, the following example is useful. Consider a function $f(x)$. Assume that $f(x) = 1$ for $x = 0$ while it is zero elsewhere over $x \in (-\infty, \infty)$. The supremum of $f(x)$ is 1, whereas its essential supremum is zero since $f(x) = 1$ occurs on a zero-measure set.

Since $\overline{\mathcal{U}}^n(z)$ is piecewise continuous on $\mathcal{H}_\mathcal{N}(x)$ as proved in [9, Appendix C], we have that ess sup $\overline{\mathcal{U}}^n(z) =$ sup $\overline{\mathcal{U}}^n(z)$ and consequently,

$$\left(\int_{\mathcal{H}_\mathcal{N}(x)} \left(\overline{\mathcal{U}}^n(z) \right)^b \mu(dz) \right)^{1/b} = \sup \left\{ \overline{\mathcal{U}}^n(z) : z \in \mathcal{H}_\mathcal{N}(x) \right\}. \tag{19}$$

Note that $\overline{\mathcal{U}}^n(z)$ is also a decreasing function with respect to distance to the source. In other words, the coverage function on the points far from the source is lower than on the points closer. Therefore, sup $\{ \overline{\mathcal{U}}^n(z) : z \in \mathcal{H}_\mathcal{N}(x) \} = \overline{\mathcal{U}}^n(x_{sup})$ occurs at the point closest to the source node. Such point lies on the intersection of the boundary region of $\mathcal{H}_\mathcal{N}(x)$ and the great circle passing through X_S and x as shown in Figure 3. By considering the continuity of $\overline{\mathcal{U}}^n(z)$, we have $\overline{\mathcal{U}}^n(x_{sup}) \to \overline{\mathcal{U}}^n(x)$ when $r_n \to 0$. Therefore, by substituting $\overline{\mathcal{U}}^n(x) = \left(\int_{\mathcal{H}_\mathcal{N}(x)} \left(\overline{\mathcal{U}}^n(z) \right)^b \mu(dz) \right)^{1/b}$ in (14) we can use (16) for this case as well. It should be noted that in this case, $\left(\int_{\mathcal{H}_\mathcal{N}(x)} \left(\overline{\mathcal{U}}^n(z) \right)^b \mu(dz) \right)^{1/b}$ is equal to $\overline{\mathcal{U}}^n(x)$ and no approximation is made. The rest of the analysis is similar to the previous section and (17) and (18) are valid here. As $r_n \to 0$, we obtain

$$\overline{\mathcal{C}} = O \left(1 - e^{-Knr^2 e^{-K_S nr^2}} \right). \tag{20}$$

$\square$

It is clear that the above expression is increased by decreasing r_n. Although r_n can approach zero, the main necessary condition for connectivity of the network is that $r_n \geq \sqrt{\log n / \pi n P_{ff}}$ [8]. In the above expression, K_S is much less than 1. However, it can be decreased by reducing p. In typical applications, for example with the parameters given in Table 1 and used in [9], we have
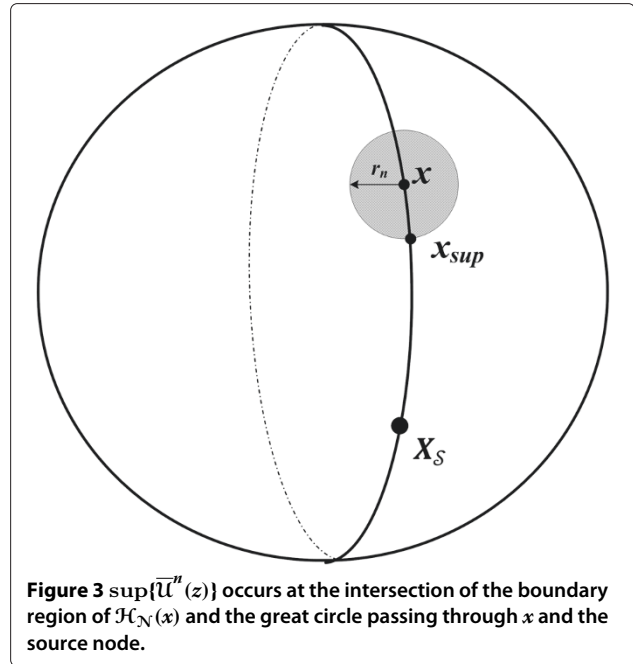

Figure 3 sup$\{\overline{\mathcal{U}}^n(z)\}$ occurs at the intersection of the boundary region of $\mathcal{H}_\mathcal{N}(x)$ and the great circle passing through x and the source node.

$K_S \approx 0.035$, which satisfies such condition. In this case, although $\overline{\mathcal{C}}$ ultimately converges to zero, it will be close to O(1) (a constant level of network coverage non-vanishing for some large values of n) for some large networks even with $n = 10^{70}$ nodes. Therefore, the decrease in network coverage begins in a very large network and the decrement rate is much less than the case of fixed r. Therefore, the decrement rate of network coverage can be controlled and considerably reduced. In these conditions, network coverage bound close to O(1) can be obtained in realistic large networks. In this way, flooding becomes a suitable and applicable broadcasting scheme for large wireless networks. The disadvantage of such condition is the increase in data dissemination delay. In addition, it is well known that flooding is not efficient in terms of energy consumption. Therefore, in applications where delay and energy are not critical issues, flooding is a proper choice of data dissemination even in large scale networks. It should be mentioned that the bound is tight and reliable as shown

Table 1 Model parameters

Parameter	Value	Description
p	0.02	Transmission probability
Δ	1	Protocol model factor
τ	85.33	Packet length in units of time slot
α	$1.17 \cdot 10^{-2}$	Normalized propagation delay

in [9]. However, when n goes to infinity in mathematical sense in (20), network coverage will be upper bounded by

$$\overline{\mathcal{C}} = \mathrm{O}\left(\log n/n^{K'_S}\right),\tag{21}$$

in the case that $r_n = \sqrt{\log n/(\pi n P_{\mathrm{ff}})}$ and where $K'_S = K_S/(\pi P_{\mathrm{ff}})$. Consequently, the number of covered nodes scales as $\mathrm{O}(n \log n/n^{K'_S})$ and if we choose network parameters such that $K'_S \leq 1$, an infinite number of nodes can be covered.

Further implications

Although the aforementioned results are valid for large values of n, we can use them for investigating the behavior of coverage for finite n as well. Figure 4 illustrates the curve for the first equality of Equation (18) in terms of n, for the choice of parameters given in Table 1 with $r = 0.3$ (the same parameters set of Table 2 in [9]). It is evident that there is a fixed value of N such that for $n < N$, the curve is an increasing function of n. However, for $n > N$ the coverage will rapidly decrease to zero. It should be noted that in this article, we focus on the tail of this curve and use its results implicitly for evaluating the coverage for different network sizes. In Figure six of [9], it has been claimed that the network coverage increases until it reaches 1 for large networks. We note that the tail of this curve has not been investigated in [9]. Heuristically, one can say that for some network sizes, the large number of neighbors increases the coverage. However, as n is increased further, the decrease in successful transmission probability of CSMA reduces the coverage.

By repeating the same procedure for the case of $r_n \to 0$ and letting $r_n = \sqrt{\log n/\pi n}$ for simple flooding, Figure 5

is obtained. We notice that N is very large in this case (about 10^{70}) and the network coverage is increasing and will be close to O(1) for $n < N$. In addition, the decrement rate of coverage in network sizes greater than N is much less than the other case. As mentioned earlier, r_n is the normalized value of transmission radius. In a realistic network, if we assume that power-limited wireless nodes are distributed on a large area growing when the number of nodes increases as used in [17], we can use the results of Theorem 2 for evaluating the performance of flooding. However, when the network area is finite, r_n cannot reach zero since the transmission power of realistic wireless nodes is lower bounded. In such cases, the results of Theorem 1 can predict the behavior of flooding.

As discussed in [9], the upper bound on network coverage is the same for one or multiple sources, since in the saturated condition, retransmission of packets related to a source at all the nodes is independent of the other broadcasting sources. Therefore, if we define the coverage as the percentage of the nodes that receive the broadcasted packet from each source, main results 1 and 2 are still valid for multiple sources. However, if we define the coverage as the percentage of nodes that receive broadcasted packets of all n_S source nodes, the network coverage will be $(\overline{\mathcal{C}})^{n_S}$, which can be readily calculated based on the results obtained in this article.

Conclusions

In this article, the scaling laws of the coverage of flooding in CSMA-based large wireless networks have been studied. An analysis has been performed for both cases of constant transmission radius r and vanishing radius r_n when the number of nodes n goes to infinity. For the case of
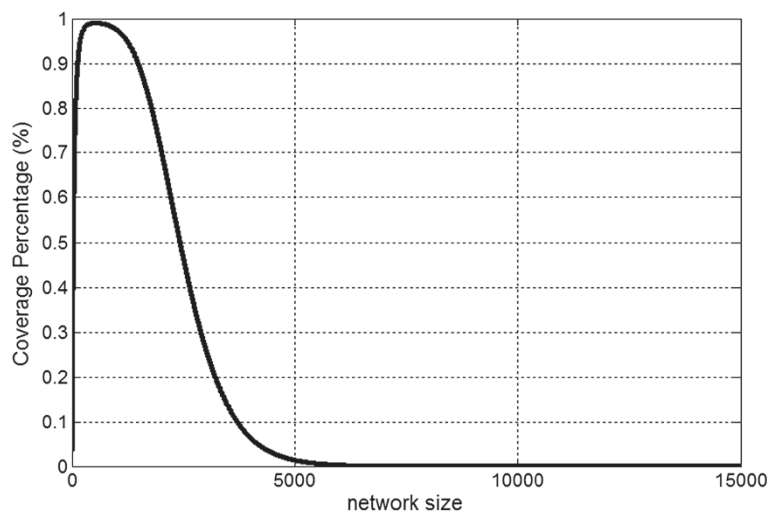

Figure 4 This curve illustrates the first term of (18) for different network sizes. Network coverage vanishes rapidly for large values of n in the case of constant r.
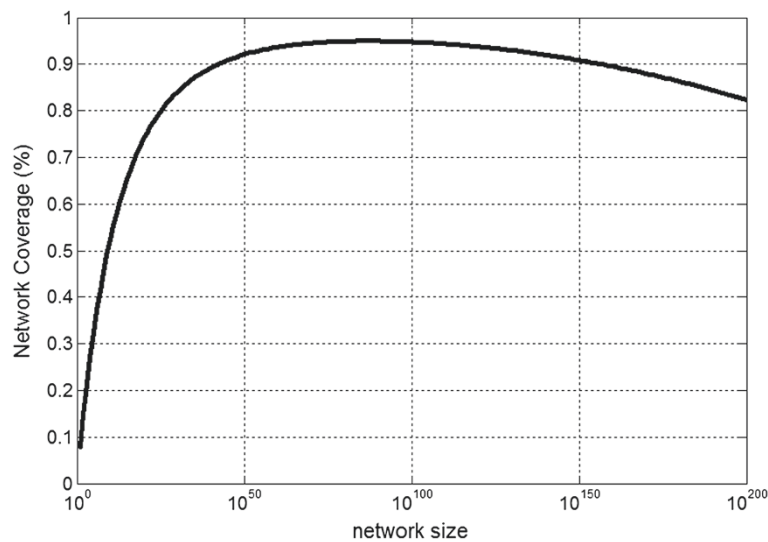
Figure 5 This curve illustrates the Equation (20) for different network sizes. The decrement in network coverage begins from 10^{70} and the decrement rate is much less after that in the case of vanishing r_n.

constant r, the results show that network coverage rapidly approaches zero when n increases. However, the rate of decrement of network coverage in case of vanishing r_n is much smaller compared to the case of constant radius. In addition, the rate can be controlled by proper choice of network parameters. Consequently, it makes flooding an appropriate and practical scheme for data dissemination, even in large networks. The number of covered nodes also approaches zero under fixed radius assumption while an infinite number of nodes can be covered in the case of $r_n \to 0$.

Endnotes

[a]In [9], we propose two upper bounds on the network coverage. Since the asymptotic analysis for the derivation of the two upper bounds is similar, we focus on deriving the asymptotic result for the first one in this article.

[b]The Hölder inequality is a fundamental inequality between integrals. Let $1 \leq a, b \leq \infty$ with $a^{-1} + b^{-1} = 1$. Consequently, for all measurable functions f and g on a measure space, $||fg||_1 \leq ||f||_a ||g||_b$. In [9], we have assumed that $a \to 1$ and consequently, we have that $b \to \infty$.

Author details
[1]Department of Electrical Engineering and Advanced Communication Research Institute (ACRI), Sharif University of Technology, Azadi Ave., Tehran, Iran. [2]Tecnalia Research & Innovation, 48170 Zamudio, Spain.

References
1. S Ni, Y Tseng, Y Chen, J Sheu, in *ACM MOBICOM conference*. The broadcast storm problem in a mobile ad hoc network, vol. 1 Washington, 1999), pp. 151–162
2. H Lim, C Kim, Flooding in wireless ad hoc networks. Comput. Commun. **24**(3-4), 353–363 (2001)
3. C Perkins, E Belding-Royer, S Das, Ad hoc on-demand distance vector (AODV) routing. IETF RFC 3561. Retrieved on 2010-06-18 (2003)
4. DB Johnson, in *IEEE Workshop on Mobile Computing Systems and Applications*. Routing in ad hoc networks of mobile hosts, vol. 1 Santa Cruz, CA, 1994), pp. 158–163
5. DB Johnson, DA Maltz, J Broch, in *Ad Hoc Networking*, ed. by CE Perkins. DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks (Addison-Wesley, Boston, MA, 2001), pp. 139–172
6. M Abolhasan, T Wysocki, E Dutkiewicz, A review of routing protocols for mobile ad hoc networks. Elsevier Ad Hoc Netws. **2**, 1–22 (2004)
7. Y Sasson, D Cavin, A Schiper, in *IEEE WCNC conference*. Probabilistic Broadcast for Flooding in Wireless Mobile Ad Hoc Networks, vol. 2 New Orleans, LA, 2003), pp. 1124–1130
8. P Gupta, PR Kumar, The capacity of wireless networks. IEEE Trans. Inf. Theory. **46**(2), 388–404 (2000)
9. H Shah-Mansouri, MR Pakravan, BH Khalaj, Analytical modeling and performance analysis of flooding in CSMA-based wireless networks. IEEE Trans. Veh. Technol. **60**(2), 664–679 (2011)
10. H Shah-Mansouri, MR Pakravan, in *IEEE ICC conference*. An upper bound on the performance of non-repetitive flooding over CSMA in wireless ad-hoc networks, Dresden, Germany, 2009), pp. 5297–5301
11. H Shah-Mansouri, MR Pakravan, in paper presented at the *IEEE PIMRC conference*. Performance analysis of flooding over CSMA in wireless ad-hoc networks, Cannes, France, 2008), pp. 2644–2648
12. A Qayyum, L Viennot, A Laouiti, in *the 35th Annual Hawaii International Conference on System Sciences, HICSS*. Multipoint relaying for flooding broadcast messages in mobile wireless networks, Island of Hawaii, USA, 2002), pp. 3866–3875
13. J Mammen, D Shah, Throughput and delay in random wireless networks with restricted mobility. IEEE Trans. Inf. Theory. **53**(3), 1108–1116 (2007)
14. L Wu, P Varshney, Performance analysis of CSMA and BTMA protocols in multi-hop networks: part 1-single channel case. Inf. Sci. Int. J. Inf. Comput. Sci. **120**(14), 159–177 (1999)

15. DE Knuth, Big omicron and big omega and big theta. ACM SIGACT News. **8**, 18–24 (1976)
16. M Adams, V Guillemin, *Measure Theory and Probability* (Cambridge, Birkhäuser, MA, 1996)
17. R Zheng, Asymptotic bounds of information dissemination in power-constrained wireless networks. IEEE Trans. Wirel. Commun. **7**(1), 251–260 (2008)