

Revisiting scalability of distributed wireless networks: A multi-hop communication perspective

R. Khalvandi, B. Sansò

G-2025-46

July 2025

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : R. Khalvandi, B. Sansò (Juillet 2025). Revisiting scalability of distributed wireless networks: A multi-hop communication perspective, Rapport technique, Les Cahiers du GERAD G- 2025-46, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (<https://www.gerad.ca/fr/papers/G-2025-46>) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: R. Khalvandi, B. Sansò (July 2025). Revisiting scalability of distributed wireless networks: A multi-hop communication perspective, Technical report, Les Cahiers du GERAD G-2025-46, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (<https://www.gerad.ca/en/papers/G-2025-46>) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2025
– Bibliothèque et Archives Canada, 2025

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2025
– Library and Archives Canada, 2025

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Télec. : 514 340-5665
info@gerad.ca
www.gerad.ca

Revisiting scalability of distributed wireless networks: A multi-hop communication perspective

Reza Khalvandi
Brunilde Sansò

GERAD & Département de génie électrique, Polytechnique Montréal, Montréal, (Qc), Canada, H3T 1J4

reza.khalvandi-ilezoole@polymtl.ca
brunilde.sanso@polymtl.ca

July 2025
Les Cahiers du GERAD
G–2025–46

Copyright © 2025 Khalvandi, Sansò

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract : Large-scale distributed wireless networks provide infrastructure-free and cost-effective connectivity, supporting applications from disaster recovery to global digital inclusion. However, multi-hop communication poses scalability challenges, as point-to-point (P2P) capacity decreases with the number of intermediate relays (*hop count*). Thus, we focus on the critical role of multi-hop communication and user interaction probability, which empirical evidence indicates decays as a power-law with geographic distance. We present a comprehensive analysis of network scalability, from capacity estimation to empirical evaluation of real-world interaction patterns. The capacity estimation problem is decomposed using a novel analytical methodology, along with symmetric topology selection and geometric partitioning, to overcome the analytical complexities inherent in previous models. The estimated P2P capacity bounds, derived from *expected* hop count, surpass previous benchmarks. Specifically, when the power-law exponent exceeds a critical threshold, the expected hop count remains stable and P2P capacity is sustained; otherwise, the hop count grows and capacity declines as the network scales. Thus, an analytical method is devised to relate real-world interaction patterns to the power-law exponent, quantified by the *contact distribution*. Then, analysis of multiple empirical datasets confirms that the exponent falls within a range that naturally supports *scalability*. Consequently, multi-hop communication does not fundamentally hinder the wide-scale deployment of distributed wireless networks. This capacity-based analysis provides a clear perspective on scalability under realistic assumptions and underscores the promising future of such networks, as well as their potential for widespread deployment.

Keywords : Large-scale distributed wireless networks, multi-hop communication, wireless network capacity, social interactions, distributed networking

1 Introduction

Distributed wireless networks rely on multi-hop communication, where each device acts as a relay to forward data across multiple wireless links. This decentralized approach allows long-range point-to-point (P2P) connectivity without relying on a central backbone. As the demand for ultra-dense connectivity is rapidly increasing [10], such networks could connect millions of users in urban-scale environments. These networks are also reasonable options with affordable costs for emerging applications such as autonomous vehicles [42, 58], unmanned aerial vehicles (UAVs) [17, 39], vehicular communication [19, 42, 49], smart cities [12], smart agriculture [11, 22, 38, 57], distributed computing [30, 46, 50], artificial intelligence [13, 40, 41, 43, 47], and the internet of things (IoT) [9]. Furthermore, they can provide vital self-organized communication among personal devices during natural disasters [1] to ensure connectivity, and can enable internet access to underserved regions, benefiting nearly 3 billion people currently offline [2]. However, realizing their full potential requires overcoming several key challenges, which have historically hindered scalable deployment [21]. Nevertheless, recent technological advancements, depicted in Figure 1, have significantly improved capacity, adaptability [39], and energy efficiency [18] and further enhance the feasibility of large-scale deployment. However, scalability is still challenged by the nature of multi-hop communication, where as the number of hops per connection increases, the available capacity per a P2P connection decreases. For instance, Figure 2 illustrates the impact of hop count on resource sharing between P2P connections in a linear arrangement of nodes, where each node establishes a multi-hop P2P connection spanning five hops to its forward neighbor. In this configuration, each wireless link is shared by five distinct connections. Consequently, the available capacity per P2P connection constitutes a fraction of the total wireless link resources, which is inversely proportional to the expected hop count (in this case, $1/5$).

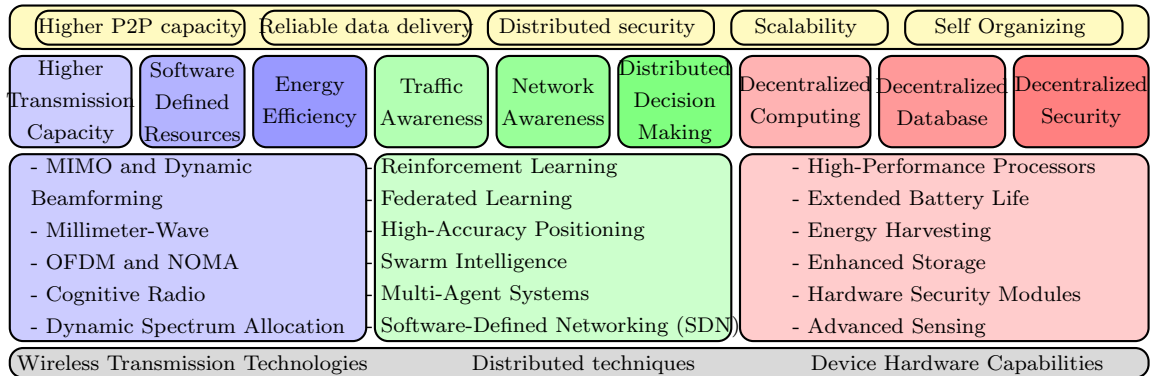


Figure 1: The promising horizon of Large-scale distributed networking concerning new technology advancements.

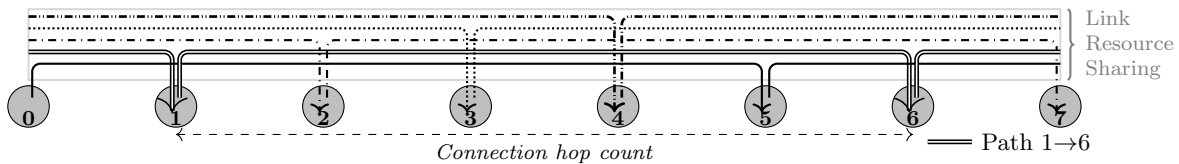


Figure 2: Multi-hop communication and wireless link resource sharing. Each node initiates a P2P connection to the node located 5 hops to its right (e.g., path $1 \rightarrow 6$). Resources are allocated inversely proportional to the P2P connections' hop count.

Although prior studies have investigated network capacity, the precise impact of multi-hop communication remains unresolved, as most existing work has focused on wireless technologies [23, 25, 26, 45] or node heterogeneity [3, 23]. However, the growth rate of the expected hop count, $\mathbf{E}(h)$, which depends on the underlying interaction model, plays a critical role in determining the average P2P capacity, C_{P2P} . We later show, in Eq. (1), that the upper bound on C_{P2P} is inversely proportional to $\mathbf{E}(h)$.

Early studies using a uniform user interaction model [23] predicted that $\mathbf{E}(h)$ scales as $\Theta(\sqrt{n})$ for a network of n nodes. However, subsequent studies [4, 36, 55] have emphasized that the interaction probability between two users separated by a distance d follows a power-law distribution proportional to $d^{-\alpha}$, where α is the power-law exponent. This model slows down the growth rate of $\mathbf{E}(h)$, which may effectively mitigate capacity degradation so that [4, 36, 55] suggested that for sufficiently large α , $\mathbf{E}(h)$ could remain bounded or grow at a reduced rate. Yet, predictions for α are uncertain, and thus $\mathbf{E}(h)$ varies from $\Theta(1)$ to $\Theta(\sqrt{n})$, casting doubt on whether distributed networks can truly maintain scalable P2P communication. To address this uncertainty in scalability, we present a comprehensive approach that first investigates, in theory, the impact of multi-hop communication on P2P capacity under a power-law interaction model, and then empirically extracts the power-law exponent from real-world data to apply it in network scalability analysis.

Our methodology isolates the role of interaction probability in scalability and decomposes the P2P capacity problem into independent analytical factors: link-level capacity, node heterogeneity, and expected hop count. By adopting a simplified *symmetric topology*, we eliminate the need to revisit well-studied link capacity and node heterogeneity, allowing to focus exclusively on hop count with minimal analytical complexity. Thus, this work 1) estimates *expected hop count trends* and *P2P capacity bounds* in a symmetric network where interaction probability follows a power-law distribution, 2) derives a method to extract the power-law exponent from empirical studies by analytically quantifying interaction patterns as the *contact distribution*, and 3) assesses scalability by integrating *theoretical capacity bounds* with *real-world datasets* of social interaction to evaluate whether the growth of $\mathbf{E}(h)$ fundamentally limits large-scale distributed networking or not. Overall, this paper makes the following key contributions:

Refined capacity bounds: We estimate C_{P2P} asymptotic behavior for a network of n nodes. For $2 < \alpha < 3$, these C_{P2P} bounds exceed previous estimates by a factor of $\ln(n)^{\alpha/2-1}$, and by $\sqrt{\ln(n)}$ for $\alpha > 3$ —enabled by a novel analytical framework and network scaling model.

A novel method for extracting the power-law exponent from empirical datasets: We quantify the interaction pattern as the contact distribution, and obtain α from the slope line value of the contact distribution over distance in the logarithmic scale. This provides a data-driven foundation for scalability analysis and clarifies the power-law exponent interpretation for capacity estimation, which reshapes conclusions about the scalability problem.

Scalability analysis through empirical evaluation of α : Empirical analysis of interaction patterns in diverse communication datasets consistently shows that α falls within the range that supports the scalability of distributed networking from multi-hop communication perspective, as a bounded $\mathbf{E}(h)$ supports sustainable P2P capacity.

The remainder of the paper is organized as follows: [Section 2](#) introduces key factors affecting capacity and reviews the existing literature. [Section 3](#) explains our methodology. [Section 4](#) presents the problem formulation and estimates $\mathbf{E}(h)$ and C_{P2P} . [Section 5](#) formulates the power-law exponent extraction method. [Section 6](#) evaluates the scalability of large-scale distributed networks. Finally, [Section 7](#) discusses the limitations, outlines future research directions, and concludes the paper.

2 Background

This section introduces key parameters influencing scalability and reviews foundational and recent studies. By analyzing the constraints on the total network transmission rate and the impact of multi-hop communication in [Theorem 1](#), we show for a distributed wireless network of n nodes operating under a TDMA communication model the P2P capacity C_{P2P} , which is the expected capacity available to any user at any given moment for sending data to a chosen destination, satisfies

$$C_{P2P} \leq \frac{\mathbf{E}(C_L)}{\mathbf{E}(h) \cdot \mathbf{E}(A_\rho)}. \quad (1)$$

Here $\mathbf{E}(C_L)$ is the expected capacity of a single active link. $\mathbf{E}(h)$ can be expressed as $\mathbf{E}(h) \propto \mathbf{E}(d)/r(n)$, where $\mathbf{E}(d)$ is the average P2P communication distance and $r(n)$ is the transmission range per link. Furthermore, $\mathbf{E}(A_\rho)$ can be interpreted as the expected normalized area, by node density ρ , dedicated to each active link. $\mathbf{E}(A_\rho)$ depends on the transmission range square and the desired link capacity. Therefore, by substituting $\mathbf{E}(A_\rho) \propto r^2(n)$ and $\mathbf{E}(h) \propto \mathbf{E}(d)/r(n)$ into Eq. (1), we have $C_{P2P} \propto \frac{\mathbf{E}(C_L)}{\mathbf{E}(d) \cdot r(n)}$ which demonstrates C_{P2P} asymptotic behavior dependency on the *key factors* when scaling up the network, can be summarized as:

$$C_{P2P} = \Theta\left(\frac{\mathbf{E}(C_L)}{r(n) \cdot \mathbf{E}(d)}\right). \quad (2)$$

Eq. (2) enables separate analysis of these key factors' impact on scalability and serves as a foundation for reviewing previous work. As highlighted in Eq. (2), minimizing the transmission range $r(n)$ is crucial for maximizing C_{P2P} . However, $r(n)$ must also guarantee network connectivity, particularly in random node arrangements where isolated nodes can arise. The foundational work of Gupta and Kumar [23] derived that the minimum transmission range required to ensure connectivity is $\Theta(\sqrt{\ln(n)})$ and established $\mathbf{E}(C_L)$ is proportional with available frequency resource W which mean total interference on a receiver even in a scale free network is bounded. Additionally, by assuming uniform interaction probability between users, they showed that $\mathbf{E}(d)$ scales as $\Theta(\sqrt{n})$. Thus, the resulting P2P capacity was calculated as $C_{P2P} = \Theta(W/\sqrt{n \ln(n)})$. This result, while foundational, suggests that as network size scales, the P2P capacity diminishes rapidly, making ad hoc networking feasible only for small networks with a few thousand nodes. However, studies in social science and social networks [5, 34] suggest that interaction probability between an individual with other is not uniform over the network coverage area, and it follows a power-law decay $d^{-\alpha}$, where d is distance and α is the power-law exponent. Using this, [36] estimated $\mathbf{E}(d)$ and C_{P2P} , where with $\alpha > 2$ in a one-dimensional network, $\mathbf{E}(d)$ is $\Theta(1)$. Similarly, Azimdoost et al. [4] examined the capacity bounds in the context of social networks, where each individual in the network has a finite number of contacts and the interaction between them follows the power-law distribution. Their research indicates that when $\alpha > 3$, C_{P2P} scales as $\Theta(W/\log(n))$. Fu et al. [20] expanded this by investigating the capacity for multicast communication scenarios. Wei et al. [55] determined the capacity of three-dimensional wireless social networks by considering advancements in aeronautical telecommunication and UAVs, where the wireless social network exhibits scalability for $\alpha > 4$. Hou et al. [26] investigated the capacity of hybrid networks comprising both ad hoc and cellular transmissions.

Node distribution also affects capacity trends. In grid-based networks $r(n)$ is $\Theta(1)$, but for randomly distributed nodes, $r(n)$ scales as $\Theta(\sqrt{\ln(n)})$ [23]. Alfano et al. [3] modeled clustering behavior using a shot-noise process. He et al. [24] introduced the car-following model to estimate and simulate the transmission capacity in vehicular ad hoc networks (VANETs). Zhou et al. [59] analyzed asymptotic capacity and delay in social-aware MANETs using a constrained mobility and rank-based social model. Cheng et al. [14] proposed a 3D cell-gridded wireless network model based on Zipf's law [61]. Qin et al. [45] evaluated full-duplex ad hoc networks with distance-limited communication pairs. Hou et al. [25] investigated beamforming techniques to enhance capacity, while Wang et al. [54] studied networks with Poisson-distributed nodes using multi-beam directional antennas.

Table 1 summarizes key theoretical models and results on capacity estimation under various assumptions. From the factors influencing C_{P2P} , the parameter $\mathbf{E}(C_L)$ reflects the wireless transmission characteristics, such as channel models, path-loss effects, transmission technologies, modulation schemes, noise, interference, and power constraints. Studies [23, 54, 55] have shown that $\mathbf{E}(C_L)$ is typically $\Theta(W)$ across various scenarios, which implies link level capacity does not degrade drastically even for a scale free network ($n \rightarrow \infty$). Furthermore, random node distributions affect capacity and lead to performance reduction scaling as $\Theta(1/\sqrt{\ln(n)})$, which is manageable even for large networks. Conversely, the communication distance $\mathbf{E}(d)$ remains a critical uncertainty. In the worst-case situation, it scales with network diameter: $\Theta(\sqrt{n})$. The primary source of this uncertainty is the lack of proper research on the interaction probability model and its parameters. This issue cannot be resolved

through capacity-bound estimation alone; rather, it requires empirical studies on real-world interaction data to accurately extract the model parameters and provide a clear answer to the scalability problem.

Table 1: Summary of Capacity Estimation Studies in Distributed Wireless Networks.

Reference	Network Setting	Communication Model	Interaction Model	$\mathbf{E}(C_L)$	$r(n)$	$\mathbf{E}(d)$	C_{P2P}
Gupta & Kumar (2000) [23]	Unit disk + Random Arrangement	TDMA	Uniform	$\Theta(W)$	$\Theta(\sqrt{\ln(n)})$	$\Theta(\sqrt{n})$	$\Theta(W/\sqrt{n \ln(n)})$
Li et al. (2001) [36]	1-D (Dimensional)	-	Power-law ($\alpha > 2$)	-	-	$\Theta(1)$	-
Azimdoost et al. (2012) [4]	Unit square + Random	TDMA	Power-law ($\alpha > 3$)	-	-	$\Theta(1)$	$\Theta(W/\log(n))$
Fu et al. (2016) [20]	Multi / Uni-cast Communication	TDMA	Rank based ($\alpha > 3/2$)	$\Theta(W)$	-	-	$O(W) - \Omega(W/\log(n))$
Wei et al. (2018) [55]	3-D (unit cube) + Random	TDMA	Power-law ($\alpha > 4$)	$\Theta(W)$	$\Theta(\sqrt{\ln(n)})$	$\Theta(1)$	$\Theta(1/\log(n))$
Hou et al. (2015) [26]	Hybrid (Ad hoc + Cellular)	Mixed Technologies	Power-law	$\Theta(W)$	-	-	Mixed parameters
Alfano et al. (2010) [3]	Unit square + Shot-noise Cox Process	TDMA	Uniform	$\Theta(W)$	Mixed	-	-
He et al. (2018) [24]	Vehicular (VANETs)	IEEE 802.11p	-	$\Theta(W)$	-	-	-
Zhou et al. (2020) [59]	Social-aware MANET	-	Rank-based power-law	-	-	Mixed	Mixed
Cheng et al. (2022) [14]	3-D Cell-gridded	TDMA	Zipf's Law-based	-	Mixed	Mixed	Mixed
Qin et al. (2016) [45]	Full-duplex Ad hoc	Full-duplex + TDMA	Distance-limited pairs	$\Theta(W)$	-	Limited	Bounded by design
Hou et al. (2018) [25]	2-D	Full-Duplex + MIMO	-	$\Theta(W)$	-	-	Improved by MIMO
Wang et al. (2021) [54]	2-D + Poisson Node Distribution	Multi-beam Antennas	Uniform	$\Theta(W)$	-	-	Improved by beam-forming

3 Methodology

To address the scalability question, a comprehensive yet effective solution is needed that is beyond the capacity estimation problem. This problem itself is inherently complex due to interactions across multiple network layers – including physical, data link, network, and transport. It complicates the analytical derivation of capacity bounds, often resulting in underestimated or misleading outcomes. For instance, in [4], a unit-square network area is divided into subareas based on transmission range $r(n)$, resulting in quantized distances. Consequently, the power-law interaction probability, dependent solely on source-destination distance, becomes intertwined with $r(n)$, causing parameter conflation, as detailed in Section 4.4. Similarly, [23] utilized the physically shortest path algorithm that favors longer links—contradicting both their own assumptions and Eq. (2)’s claim that minimizing $r(n)$ maximizes C_{P2P} , leading to pessimistic capacity estimations. It is essential to address these issues and, beyond that, to extract the power-law parameter directly from empirical datasets. Thus, we reassess and restructure the analytical approach used in state-of-the-art methodologies in the following aspects:

(1) Modular decomposition: We decompose the capacity estimation problem into independent analytical factors, including wireless link capacity, single-link transmission range, and the expected number of hops for a P2P connection, by deriving a stochastic analysis for a general network model as outlined in Theorem 1. This separation of key factors, formulated in Eq. (2), allows us to focus on the

primary research gap and bottleneck in scalability analysis: the role of multi-hop communication and interaction probabilities. Even when assuming fixed capacity across all wireless links, scalability is still fundamentally limited by multi-hop communication and interaction probability impact. Indeed, analyzing overly complex scenarios that closely model link-level communication and node heterogeneity is futile without first addressing the uncertainty of the power-law exponent. On the other hand, extensive prior work has shown that the scaling behavior of link-level capacity, $\mathbf{E}(C_L)$, under various network settings and communication models, consistently follows $\Theta(W)$, as summarized in Table 1. Therefore, to estimate C_{P2P} , we incorporate link-level capacity and node arrangement models from prior literature, leveraging the factorized framework established in Eq. (2). This approach eliminates the need to revisit detailed communication models for wireless links or account for heterogeneous node placement. Consequently, this separation allows us to adopt a deliberately simplified scenario, enabling accurate computation of the expected hop count $\mathbf{E}(h)$ without confounding effects from complex lower-layer modeling.

(2) Topology control: Previous studies focused primarily on limited-area scenarios such as unit discs [23], squares [4, 20, 26], or cubes [14, 55] and increased node density to model scaling behavior. However, real-world scenarios typically involve expanding network areas rather than increasing densities. Thus, we employ a scenario with constant node density and scaled network areas. Moreover, to isolate the effects of interaction probability, we adopt a *fixed transmission range* scenario by symmetrically positioning nodes at the vertices of a square lattice (see Figure 3). Besides, we do not recompute link-level capacity, and the network model only specifies how nodes are interconnected, abstracting away details of the communication model, interference, and transmission power. Links are assumed to have fixed capacity, and the focus is placed on computing the expected number of relays required for P2P communication. In this symmetric topology, both inter-node distances (immediate neighbors) and transmission range remain constant regardless of network size, ensuring that all links have equal metric and weight. Thus, the physically shortest path—which minimizes the number of relays for a P2P connection—serves as the optimal route for maximizing C_{P2P} . Furthermore, this symmetric setting facilitates mapping distance-based interaction probabilities to hop-count probabilities, thereby avoiding parameter conflation and preventing underestimated capacity bounds, as illustrated in Section 4.4.

(3) Analysis simplification: Utilizing the properties of power-law interaction probabilities, we simplify analytical complexity. Studies such as [4, 23] typically partition the network area into smaller equidistant subareas, which are effective for uniform interaction probabilities but inefficient for power-law distributions. A geometric partitioning, however, aligns naturally with the properties of the power-law distribution. In this distribution, $P(d) \propto 1/d^\alpha$, and $\delta P(d)/\delta d \propto -\alpha/d^{\alpha+1}$, indicating negligible probability variation at large distances ($d \gg 1$). Conversely, relative changes $\delta P(d)/P(d) = -\alpha \delta d/d$ fit well into a geometric progression of distances. With the geometric step change, the interaction probability of nodes in consecutive groups also follows a geometric progression, significantly simplifying analytical derivation.

(4) Comprehensive theoretical and empirical analysis: Recognizing limitations in prior research, we propose a fundamentally distinct analytical framework. We first theoretically identify critical factors influencing capacity. Next, we select an appropriate scenario that isolates key scalability factors and simplifies analytical complexity. This structured approach allows us to estimate $\mathbf{E}(h)$, derive capacity bounds, and extract interaction probability parameters from empirical data. Finally, we integrate multiple datasets of social interactions across various communication contexts, providing a comprehensive answer to network scalability.

4 Expected hop count and capacity estimation

We first formulate the problem model for the symmetric network setting shown in Figure 3, assuming that node interactions follow a power-law probability distribution. Then, we compute $\mathbf{E}(h)$ and derive asymptotic bounds for C_{P2P} .

4.1 Problem model

We assume a symmetric topology, where nodes are positioned at the vertices of a square lattice (see Figure 3). The physical link distance between neighboring nodes is constant and denoted as l . Thus, the node density is given by $1/l^2$, and the total network coverage is approximately $l^2 n$, where n represents the total number of nodes. The expected hop count of all possible P2P connections is defined as:

$$\mathbf{E}(h) = \sum_{k=1}^{k_{\max}} k \cdot P(h = k), \quad (3)$$

where $\mathbf{E}(h)$ represents the expected number of hops required to establish a P2P connection. This is calculated as the summation over all possible hop counts, denoted by k , each weighted by the probability $P(h = k)$ and ranges from the minimum possible hop count ($k = 1$) to the maximum possible hop count $k_{\max}$ for a P2P connection. While the hop count k reflects the number of relays required for a connection, it is the interaction probability between nodes that shapes how these connections are established. Thus, $P(h = k)$ is not directly accessible. However, due to the symmetric network configuration, the optimal routing that maximizes C_{P2P} naturally follows the shortest physical path. Consequently, there exists a linear relationship between the number of hops and the physical distance between nodes. This linear relation allows the original power-law interaction model, where

$$P(d) = \gamma/d^\alpha, \quad (4)$$

with γ as a constant, to be translated directly into a hop-based interaction model. Under this distribution, all nodes equidistant from a source node share the same interaction probability with that source. Hence, nodes can be grouped based on their distance from the source to compute $\mathbf{E}(h)$ for each group, and then summed across all groups. Each group can be represented as a ring of thickness l (link length), as shown in Figure A.2, with nodes in the i -th ring at distance $d_i = li$. However, for $i \gg 1$, the interaction probabilities for nodes at distances $d_i = il$ and $d_{i+1} = (i+1)l$ become nearly identical, as $i^\alpha/(i+1)^\alpha \approx 1$.

Since nodes at similar distances share nearly identical interaction probabilities, we approximate the continuous power-law distribution with a discrete tiered structure. Taking this into account, we implement multi-resolution grouping using a geometric progression, adjusting the step size from $1, 2, \dots, i$ to $1, \zeta^2, \dots, \zeta^i$ ($\zeta > 1$). This approach partitions the network into nested tiers of nodes around the source node, where all nodes of a same tier hold similar probability of interaction with that source node. In the first tier, nodes are enclosed within a square of area a (see Figure 3). Nodes located between distances $d_i = \sqrt{a}\zeta^{i-1}$ and $d_{i+1} = \sqrt{a}\zeta^i$ are assigned to the i^{th} tier, and so forth. To approximate the interaction probability for the i^{th} tier as P_i , we consider the $P(d)$ value at the side length of the square containing the nodes in that tier from Eq. (4):

$$P_i = P(d_i) = \frac{\gamma}{\sqrt{a}^\alpha \zeta^{\alpha(i-1)}}. \quad (5)$$

As the ratio of side lengths between consecutive tiers is ζ , interaction probabilities follow an exponential decay:

$$\frac{P_{i+1}}{P_i} = \frac{\frac{\gamma}{d_{i+1}^\alpha}}{\frac{\gamma}{d_i^\alpha}} = \left(\frac{d_i}{d_{i+1}} \right)^\alpha = \zeta^{-\alpha}. \quad (6)$$

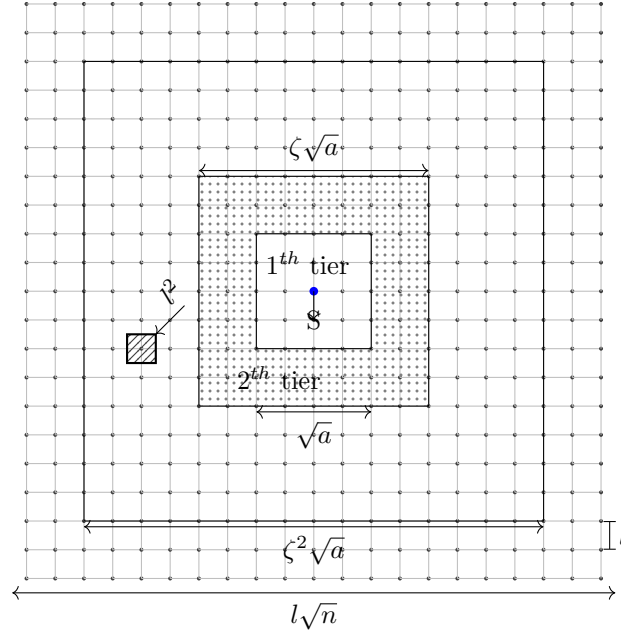


Figure 3: Symmetric lattice structure of the distributed network with concentric grouping with ratio $\zeta = 2$ and $\frac{\sqrt{a}}{l} = 4$.

Thus, if nodes located in the first tier around a source node S have an interaction probability P_1 , then nodes in subsequent tiers have their interaction probabilities scaled accordingly. Specifically, nodes in the second tier have an interaction probability $P_2 = \zeta^{-\alpha} P_1$, and more generally, for the i^{th} tier ($i \geq 2$), the interaction probability is given by:

$$P_i = P_1 \zeta^{-(i-1)\alpha}. \quad (7)$$

Since the sum of all interaction probabilities must equal one, P_1 is determined accordingly in the following subsection. To cover the entire network, we extend the tiers until the network's side length $l\sqrt{n}$ approximately matches the side length of the square enclosing the nodes in the last tier, $\sqrt{a}\zeta^{m-1}$. This allows us to express the number of tiers m as:

$$\sqrt{a}\zeta^{m-1} = l\sqrt{n} \quad \Rightarrow \quad m - 1 = \log_{\zeta} \left(l\sqrt{\frac{n}{a}} \right). \quad (8)$$

Next, we determine the number of nodes in each tier. The node density is $1/l^2$, so the number of nodes within the initial square (first tier) is approximately, neglecting edge effects:

$$N_1 = \frac{a}{l^2}. \quad (9)$$

For subsequent tiers, the second tier is enclosed by two squares with side lengths $\sqrt{a}$ and $\zeta\sqrt{a}$, covering $\zeta^2 a/l^2 = \zeta^2 N_1$ nodes. The number of nodes in the second tier alone is therefore:

$$N_2 = \zeta^2 N_1 - N_1 = (\zeta^2 - 1)N_1. \quad (10)$$

For higher tiers, the number of nodes follows a geometric progression:

$$N_{i+1} = N_i \zeta^2, \quad \text{for } i \geq 2. \quad (11)$$

Since $N_1 \gg 1$, we approximate $N_1 - 1 \approx N_1$. Finally, the general formula for N_i is:

$$N_i = \begin{cases} \frac{a}{l^2}, & i = 1 \\ \frac{a}{l^2}(\zeta^2 - 1)\zeta^{2(i-2)}, & i \geq 2. \end{cases} \quad (12)$$

Table 2: Parameters description.

Describe	Parameter
n	Number of nodes in network
d	Physical distance
$P(d)$	Interaction probability over distance d
α	Power-law distribution exponent
$\mathbf{E}(h)$	Expected hop count per connection
C_{P2P}	Expected point to point capacity per connection
$\Theta(\cdot)$	Asymptotic growth rate
C_L	Single hope transmission capacity (bps)
C_{net}	Total network transmission rate (bps)
$\mathbf{E}(C_L)$	The expected transmission rate per link (bps)
W	Available frequency (Hz)
ρ	Node density (number of nodes per unit area)
$\mathbf{E}(A_\rho)$	Expected normalized area per link by node density
$\mathbf{E}(d)$	Expected Communication distance
$r(n)$	Transmission range
l	Physical link distance
a	Area of the first tier
S	Source node
v	Node index
j	Link index
k	Hop index
i	Tire index
m	Number of tiers
P_i	Node's interaction probability of tire i^{th}
N_i	Number of nodes per tier
$\mathbf{E}_i(h)$	Average path length of tire i^{th}
$F(\cdot)$	Aggregate node hop count in a triangle
γ	Interaction probability constant
ζ	Exponential step size
$C(d)$	contacts distribution over distance d
$\mathbf{E}(C_i)$	Expected contact number per tier

4.2 Probability constant estimation

To determine the probability constant P_1 in Eq. (7), we ensure that when a source node S selects a destination, any other node v in the network, at distance d_v from S , has a nonzero probability $P(d_v)$ of being chosen. This probability must satisfy the condition $\sum_{v \neq S} P(d_v) = 1$. Since all nodes within the same tier have an equal probability of selection, we reformulate this condition in terms of the interaction probabilities, summing over all m tiers:

$$\sum_{i=1}^m N_i P_i = 1. \quad (13)$$

Depending on the position of the source node S , parameters of Eq. (13) may vary. Thus, we consider boundary cases: one where the source node is at the network center and another where it is at the network corner, representing the best-case and worst-case scenarios, respectively, as illustrated in Figure 4. In both cases, the first tier contains N_1 nodes with probability P_1 , while nodes in the i^{th} tier ($i \geq 2$) have probability $P_i = \zeta^{-(i-1)\alpha} P_1$. Substituting N_i and P_i from Eq. (12) and Eq. (7) into Eq. (13), and after some simplification, we obtain:

$$\frac{a}{l^2} P_1 + (\zeta^2 - 1) \frac{a}{l^2} \sum_{i=2}^m P_1 \zeta^{-\alpha} \zeta^{(2-\alpha)(i-2)} = P_1 \frac{a}{l^2} \left(1 + (\zeta^2 - 1) \zeta^{-\alpha} \sum_{i=2}^m \zeta^{(2-\alpha)(i-2)} \right) = 1. \quad (14)$$

Before proceeding, we establish in [Theorem 2](#) that *any interaction probability dependent solely on source-destination distance falls with at least $\Theta(1/d^2)$ over the distance from the source node*, which also applies to the power-law model. Thus, for $\zeta > 1$ and $\alpha > 2$, the summation $\sum_{i=2}^m \zeta^{(2-\alpha)(i-2)}$ forms a geometric series:

$$\sum_{i=0}^{m-2} \zeta^{(2-\alpha)i} = \frac{\zeta^{(2-\alpha)(m-1)} - 1}{\zeta^{(2-\alpha)} - 1}. \quad (15)$$

Based on [Eq. \(8\)](#), $m-1 = \log_{\zeta} \left(l\sqrt{n/a} \right)$ that gives us $\zeta^{(2-\alpha)(m-1)} = \zeta^{(2-\alpha)(\log_{\zeta}(l\sqrt{n/a}))} = 1/(l^2 n/a)^{\frac{\alpha-2}{2}}$. If $n \rightarrow \infty$, $1/(l^2 n/a)^{\frac{\alpha-2}{2}} \ll 1$, leading to:

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{m-2} \zeta^{(2-\alpha)i} = \frac{1}{1 - \zeta^{(2-\alpha)}}. \quad (16)$$

Finally, putting this in [Eq. \(14\)](#) results P_1 as:

$$P_1 = \frac{l^2}{a \left(1 + \frac{(\zeta^2 - 1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}} \right)}. \quad (17)$$

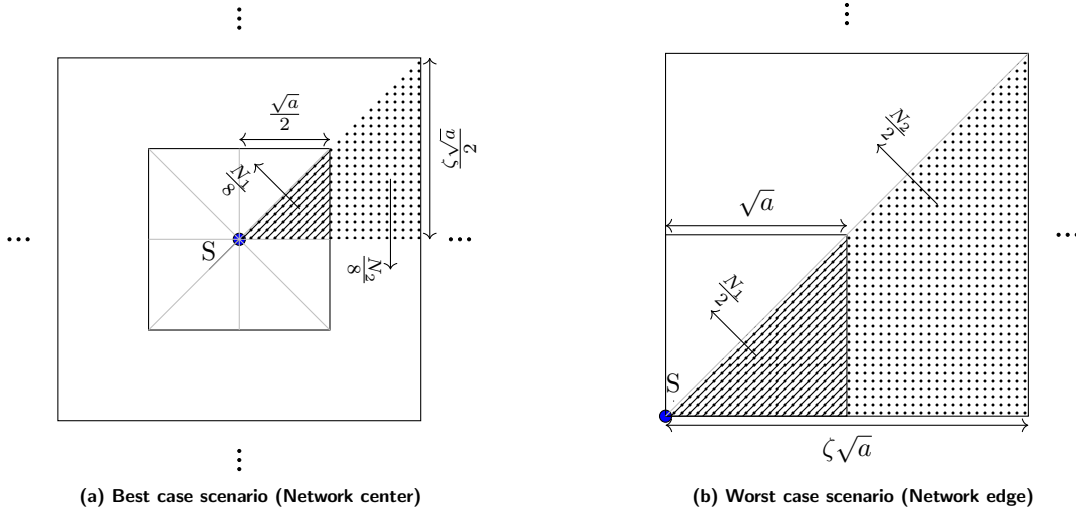


Figure 4: Boundary scenarios for expected path values. N_1 and N_2 are the number of nodes in the first and second tiers.

4.3 Expected hop count estimation

We now calculate the expected hop count, $\mathbf{E}(h)$, which is the sum of all possible hop counts weighted by their respective probabilities, as given in [Eq. \(3\)](#). However, as we grouped nodes into nested tiers, it is convenient to express $\mathbf{E}(h)$ as the sum of expected hop counts per tier:

$$\mathbf{E}(h) = \sum_{i=1}^m \underbrace{\mathbf{E}_i(h)}_{\text{Tier path length}} \cdot \underbrace{N_i P_i}_{\text{Tier probability}}. \quad (18)$$

In this formulation, we have already determined N_i and P_i , while $\mathbf{E}_i(h)$ represents the average path length per tier. $\mathbf{E}_i(h)$ is the total hop distance of all nodes from the source, divided by the number of nodes in that tier. Because all nodes within the same tier have the same probability P_i of interacting with the source node, $\mathbf{E}_i(h)$ characterizes the tier's contribution to the overall hop count. As $\mathbf{E}(h)$

depends on the source node's position in the network, we again consider boundary cases: the best-case scenario and the worst-case scenario, illustrated in Figure 4. They represent the minimum and maximum possible values of $\mathbf{E}(h)$. Now, to compute $\mathbf{E}_i(h)$ tier by tier, we start with the first tier in the best-case scenario, where nodes in this tier are arranged in eight symmetric isosceles triangles, as shown in Figure 4. Each triangle contains $N_1/8 = a/8l^2$ nodes. Using Lemma 1, the total hop distance of all nodes within the initial triangle is given by $F(\frac{\sqrt{a}}{2l})$, where each equal-length side contains $\frac{\sqrt{a}}{2l}$ nodes from the source. Thus, $\mathbf{E}_1^{\text{center}}(h)$ is computed as:

$$\mathbf{E}_1^{\text{center}}(h) = \frac{8l^2 F(\frac{\sqrt{a}}{2l})}{a} = \frac{8l^2 (\frac{\sqrt{a}}{2l})^3}{a} = \frac{\sqrt{a}}{2l}.$$

For the second tier of the best-case scenario, which is also divided into eight geometrically similar areas, each containing $(\zeta^2 - 1)N_1/8$ nodes, the path length is calculated over the region between two isosceles triangles with side lengths $\frac{\sqrt{a}}{2l}$ and $\frac{\zeta\sqrt{a}}{2l}$, as shown in Figure 4. Applying Lemma 1 for nodes' distance summation $F(\frac{\zeta\sqrt{a}}{2l}) - F(\frac{\sqrt{a}}{2l})$, we have:

$$\mathbf{E}_2^{\text{center}}(h) = \frac{8l^2 (F(\frac{\zeta\sqrt{a}}{2l}) - F(\frac{\sqrt{a}}{2l}))}{(\zeta^2 - 1)a} = \frac{(\zeta^3 - 1)\sqrt{a}}{2(\zeta^2 - 1)l}.$$

For the worst-case scenario, where the source node is positioned at the network edge, see Figure 4. Applying same approach and using Lemma 1, we could simply show average path length for the first and second tier follows:

$$\mathbf{E}_1^{\text{edge}}(h) = \frac{\sqrt{a}}{l}, \quad \mathbf{E}_2^{\text{edge}}(h) = \frac{(\zeta^3 - 1)\sqrt{a}}{(\zeta^2 - 1)l}.$$

Due to geometric similarity, the average distance increases by a factor of ζ with each subsequent tier in both cases as:

$$\mathbf{E}_{i+1}(h) = \zeta \mathbf{E}_i(h) \quad , \quad i \geq 2. \quad (19)$$

Since the number of nodes N_i and the interaction probabilities P_i remain identical in both cases, and given that the worst case $\mathbf{E}_i(h)$ is exactly twice the best case $\mathbf{E}_i(h)$, we conclude:

$$\mathbf{E}_i^{\text{edge}}(h) = 2\mathbf{E}_i^{\text{center}}(h). \quad (20)$$

Thus, the worst-case $\mathbf{E}(h)$ serves as an upper bound, while the general expected hop count is no lower than half this worst-case estimate. To maintain analytical simplicity, we pick the worst-case scenario $\mathbf{E}_i(h)$ to derive an upper bound for the $\mathbf{E}(h)$ estimation:

$$\mathbf{E}_i(h) = \begin{cases} \frac{\sqrt{a}}{l}, & i = 1 \\ \frac{(\zeta^3 - 1)\sqrt{a}}{(\zeta^2 - 1)l} \zeta^{i-2}, & i \geq 2. \end{cases} \quad (21)$$

Now, by substituting the values of P_i , N_i , and $\mathbf{E}_i(h)$ from Eq. (7), Eq. (12), and Eq. (21), respectively, into Eq. (18), we can express $\mathbf{E}(h)$ as:

$$\mathbf{E}(h) = \left(\frac{\sqrt{a}}{l}\right)^3 P_1 + \sum_{i=2}^m P_1 (\zeta^3 - 1) \left(\frac{\sqrt{a}}{l}\right)^3 \zeta^{-\alpha} \zeta^{-\alpha(i-2)} \zeta^{3(i-2)}. \quad (22)$$

Substituting probability constant P_1 from Eq. (17) for $\alpha > 2$, we get:

$$\mathbf{E}(h) = \frac{\sqrt{a}}{l(1 + \frac{(\zeta^2 - 1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}})} \left(1 + (\zeta^3 - 1)\zeta^{-\alpha} \sum_{i=2}^m \zeta^{(3-\alpha)(i-2)} \right). \quad (23)$$

For $m \gg 1$, the geometric series $\sum_{i=2}^m \zeta^{(3-\alpha)(i-2)}$ results could be computed, and then the expected hop count $\mathbf{E}(h)$ under a power-law distribution follows the generalized form:

$$\mathbf{E}(h) = \begin{cases} \frac{\sqrt{a}(1 + \frac{(\zeta^3-1)\zeta^{-\alpha}(1 - (\frac{l^2 n}{a})^{\frac{3-\alpha}{2}}))}{1 - \zeta^{\frac{3-\alpha}{2}})}, & \alpha > 3 \\ \frac{l(1 + \frac{(\zeta^2-1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}})}{\sqrt{a}(1 + (1 - \zeta^{-3}) \log_{\zeta}(l\sqrt{\frac{n}{a}}))}, & \alpha = 3 \\ \frac{\sqrt{a}(\zeta^3-1)\zeta^{-\alpha}(\frac{l^2 n}{a})^{\frac{3-\alpha}{2}}}{l(\zeta^{(3-\alpha)}-1)(1 + \frac{(\zeta^2-1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}})}, & 2 < \alpha < 3. \end{cases} \quad (24)$$

If we reduce the step grow rate to its minimum limit ($\zeta \rightarrow 1$), then $\lim_{\zeta \rightarrow 1} \mathbf{E}(h)$ computed as below, which minimize the effect of the discrete layering in $\mathbf{E}(h)$ approximation:

$$\lim_{\zeta \rightarrow 1} \mathbf{E}(h) = \begin{cases} \frac{\sqrt{a}(\alpha-2)(\alpha-3)(\frac{l^2 n}{a})^{\frac{3-\alpha}{2}}}{l\alpha(\alpha-3)}, & \alpha > 3 \\ \frac{\sqrt{a}(1+3\ln(l\sqrt{\frac{n}{a}}))}{3l}, & \alpha = 3 \\ \frac{(\frac{\sqrt{a}}{l})^{\alpha-2} 3(\alpha-2)n^{\frac{3-\alpha}{2}}}{\alpha(3-\alpha)}, & 2 < \alpha < 3. \end{cases} \quad (25)$$

Although the ultimate goal is the analysis of asymptotic behavior of $\mathbf{E}(h)$ and not its exact value, [Eq. \(25\)](#) could provide also a numerical approximation of the $\mathbf{E}(h)$ based on the network size n and power-law exponent α . Finally, $\mathbf{E}(h)$ asymptotic behavior could be derived from [Eq. \(25\)](#) as:

$$\mathbf{E}(h) = \begin{cases} \Theta(1), & \alpha > 3 \\ \Theta(\ln(n)), & \alpha = 3 \\ \Theta(n^{\frac{3-\alpha}{2}}), & 2 < \alpha < 3. \end{cases} \quad (26)$$

4.4 Capacity bounds and comparison

In [Section 2](#), we identify key factors influencing P2P capacity: link capacity $\mathbf{E}(C_L)$, transmission range $r(n)$, and expected hop count $\mathbf{E}(h)$, and establish their relationship with C_{P2P} . For a symmetric node arrangement with $r(n) = \Theta(1)$ and $\mathbf{E}(d)$ has have similar scaling behavior with $\mathbf{E}(h)$, as link lengths are constant. It is also well established that $\mathbf{E}(C_L) = \Theta(W)$ [23, 54, 55]. Thus, [Eq. \(2\)](#) using for the upper C_{P2P} bounds [Eq. \(2\)](#), the P2P capacity asymptotic behavior is expressed as:

$$C_{P2P} = \Theta\left(\frac{\mathbf{E}(C_L)}{r(n)\mathbf{E}(d)}\right) = \Theta\left(\frac{W}{\mathbf{E}(h)}\right). \quad (27)$$

Substituting $\mathbf{E}(h)$ asymptotic behavior from [Eq. \(26\)](#) into [Eq. \(27\)](#), the C_{P2P} asymptotic behavior can be formulated as:

$$C_{P2P} = \begin{cases} \Theta(W), & \alpha > 3 \\ \Theta\left(\frac{W}{\ln(n)}\right), & \alpha = 3 \\ \Theta\left(\frac{W}{n^{\frac{3-\alpha}{2}}}\right), & 2 < \alpha < 3. \end{cases} \quad (28)$$

To benchmark our findings against the state-of-the-art, we compare our results with the work of Azimdoost et al. [4], which establishes the following capacity bounds, with the same power-law interaction model:

$$\begin{cases} \Theta\left(\frac{n-q-1}{n^2 r^{\alpha-1}(n)}\right), & 2 \leq \alpha \leq 3, \quad q < \infty \\ \Theta\left(\frac{n-q-1}{n^2 r^2(n)}\right), & \alpha \geq 3, \quad q < \infty. \end{cases} \quad (29)$$

By incorporating the function $r(n) = \sqrt{\ln(n)/n}$, as suggested in [23] to ensure network connectivity, and assuming a contact number of $q = 1$, the capacity bounds in [4] are simplified to:

$$\begin{cases} \Theta\left(\frac{1}{n^{\frac{(3-\alpha)}{2}} \ln(n)^{\frac{(\alpha-1)}{2}}}\right), & 2 \leq \alpha \leq 3 \\ \Theta\left(\frac{1}{\ln(n)}\right), & \alpha > 3. \end{cases} \quad (30)$$

However, Azimdoost et al.'s model [4] does not necessarily assume a symmetric node arrangement. Gupta et al. [23] demonstrated that the achievable capacity bound for symmetric settings is $\sqrt{\ln(n)}$ times higher than in random configurations. To enable a fair comparison, we normalize the capacity bounds from [4] by multiplying them by $\sqrt{\ln(n)}$, yielding:

$$\begin{cases} \Theta\left(\frac{1}{n^{\frac{(3-\alpha)}{2}} \ln(n)^{\frac{\alpha}{2}-1}}\right), & 2 \leq \alpha \leq 3 \\ \Theta\left(\frac{1}{\sqrt{\ln(n)}}\right), & \alpha > 3. \end{cases} \quad (31)$$

Then, our capacity bounds in Eq. (28) outperform those presented in [4] as follows:

$$\begin{cases} \ln(n)^{\frac{\alpha}{2}-1}, & 2 < \alpha < 3 \\ \sqrt{\ln(n)}, & \alpha \geq 3. \end{cases} \quad (32)$$

This indicates that for $2 < \alpha < 3$, our capacity bound is superior by a factor of $\ln(n)^{\alpha/2-1}$, and for $\alpha > 3$, our bound is $\sqrt{\ln(n)}$ times better than [4]. This improvement arises primarily from two differences:

1. *Scaling model and network structure:* To analyze network scalability as the number of nodes n grows, [4] maintains a fixed unit square area, increasing node density, whereas our model expands the network size while preserving node density, which aligns more closely with empirical power-law interaction models. Moreover, our symmetrical scenario enables grouping nodes solely based on physical distances to calculate $\mathbf{E}(h)$, which allows us to convert a purely distance-dependent interaction probability into a hop count probability independently of $r(n)$ or other parameters.
2. *Analytical separation of key parameters:* our framework, consistent with Gupta [23], expresses the capacity as $C_{P2P} = \Theta(1/r(n))$ instead of $\Theta(1/r^2(n))$ [4], and $\mathbf{E}(h) = \mathbf{E}(d)/r(n)$, exhibiting only linear dependence on $r(n)$. Conversely, [4] divides the unit square area into a grid of cells with lengths proportional to transmission range $r(n)$, forming the basis for calculating the expected hop count. Thus, [4] directly ties interaction probability to $r(n)$, implying that changes in transmission range alter interaction probabilities even if physical distance is constant, which lacks physical justification. Consequently, the formulation in [4] includes higher-order terms such as $r^{\alpha-2}(n)$ and $r^2(n)$, which may lead to an underestimation of capacity and an implicit coupling between the transmission range $r(n)$ and the interaction probability parameter α .

5 Power-law exponent estimation

The ultimate goal of capacity estimation is to determine whether wireless distributed networking is scalable. Since the capacity bounds, given by Eq. (28), are influenced by the power-law exponent α , the scalability outcome depends on the value of α . However, none of the capacity estimation studies—including those focused on power-law interactions [4, 20, 36, 55]—have investigated how to measure the power-law exponent or what range of α values occur in real-world settings. Moreover, empirical studies on social ties that identified power-law patterns in human interactions offer diverse interpretations of this behavior depending on context. Most of these works, introduced in Section 6, do not examine social interaction datasets from a network perspective. Consequently, there is currently no

fundamental framework for directly applying empirical social data to capacity and scalability analysis. To bridge this gap, we introduce a structured mathematical approach to extract α directly from statistical data, ensuring consistency with both theoretical capacity bounds and real-world observations. Empirical studies [5, 16, 34, 44] consistently represent social ties quantified by the *contacts distribution* or the frequency of interaction over distance using a logarithmic curve, where both distance and contacts distribution follow exponential growth patterns. This scheme closely aligns with our network model in Section 4.1, where we partition the network into multi-resolution tiers that expand exponentially to estimate $\mathbf{E}(h)$. Under this scenario, the contacts distribution, within each tier i as $\zeta \rightarrow 1$, converges to the *contacts distribution* over distance, denoted as $C(d)$. This $C(d)$ matches the empirical representation of social ties presented in previous studies. Consequently, analyzing the slope of the log-log relationship between contact numbers and distance rigorously determines the parameters of the power-law model. This correspondence enables a robust extraction of the power-law distribution parameters directly from empirical datasets. Thus, we define the expected contacts distribution for a given node within each tier, $\mathbf{E}(C_i)$, as the product of the interaction probability and the total number of nodes in that specific tier, expressed as follows:

$$\mathbf{E}(C_i) = P_i N_i. \quad (33)$$

By substituting P_i and N_i from Eqs. (7) and (12) when $(i \geq 2)$, we can write:

$$\mathbf{E}(C_i) = P_1 \zeta^{-(i-1)(\alpha)} (\zeta^2 - 1) \frac{a}{l^2} \zeta^{2(i-2)} = P_1 \zeta^{-\alpha} (\zeta^2 - 1) \frac{a}{l^2} \zeta^{(2-\alpha)(i-2)}. \quad (34)$$

Now, we can define the average distance of each tier from the source node as $\bar{d}_i = \frac{d_i^{max} + d_i^{min}}{2} = \frac{\sqrt{a} \zeta^{i-1} + \sqrt{a} \zeta^i}{2} = \frac{\sqrt{a} \zeta^{i-2} (1 + \zeta)}{2}$. Therefore,

$$\zeta^{i-2} = \frac{2\bar{d}_i}{\sqrt{a} \zeta (1 + \zeta)}. \quad (35)$$

Substituting Eq. (35) into Eq. (34), $\mathbf{E}(C_i)$ is written as

$$\mathbf{E}(C_i) = P_1 \zeta^{-\alpha} (\zeta^2 - 1) \frac{a}{l^2} \left(\frac{2\bar{d}_i}{\sqrt{a} \zeta (1 + \zeta)} \right)^{(2-\alpha)}.$$

Based on Theorem 2, for any power-law distribution interaction probability, when the network size goes to infinity, we should have $\alpha > 2$. Hence, by using the value of P_1 from Eq. (17), where $\alpha > 2$,

$$\mathbf{E}(C_i) = \frac{1}{\left(1 + \frac{(\zeta^2 - 1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}}\right)} \zeta^{-\alpha} (\zeta^2 - 1) \left(\frac{2\bar{d}_i}{\sqrt{a} \zeta (1 + \zeta)} \right)^{(2-\alpha)}.$$

Therefore, if $\alpha > 2$ and $i \geq 2$, $\mathbf{E}(C_i)$ is

$$\mathbf{E}(C_i) = \frac{2^{2-\alpha} \sqrt{a}^{\alpha-2} (1 - \zeta^{-2})}{\left(1 + \frac{(\zeta^2 - 1)\zeta^{-\alpha}}{1 - \zeta^{(2-\alpha)}}\right) (1 + \zeta)^{2-\alpha}} \bar{d}_i^{2-\alpha}. \quad (36)$$

If we reduce the step size as much as possible ($\zeta \rightarrow 1$), $\bar{d}_i \approx d_i$ and $\lim_{\zeta \rightarrow 1} \mathbf{E}(C_i)$ can be written as Eq. (37):

$$\mathbf{E}(C_i) = \frac{2(\alpha - 2) \sqrt{a}^{\alpha-2}}{\alpha} d_i^{2-\alpha}. \quad (37)$$

By taking the logarithm on both sides of Eq. (37), we have:

$$\log(C(d)) = -(\alpha - 2) \log(d) + C_0. \quad (38)$$

Hence, if we define b as the slope of the $C(d)$ curve in the log-log scale, then $\alpha = 2 - b$. [Eq. \(38\)](#) enables the examination of network scalability using empirical observations.

[Eq. \(38\)](#) clarifies the interpretation of the measurement of power-law exponent from real empirical data. To illustrate it, we return to the interaction probability, which is defined as $P(d) = \gamma/d^\alpha$, and take the logarithm on both sides, we obtain:

$$\log(P(d)) = -\alpha \log(d) + \log(\gamma). \quad (39)$$

This equation shows that the slope of the log-log curve for interaction probability is $-\alpha$. This highlights a significant difference between *interaction probability* and *the contacts distribution* in [Eq. \(38\)](#). All empirical studies listed in [Table 3](#) have analyzed the relationship between distance and interactions by focusing on the *contacts distribution* or similar metrics, rather than the *interaction probability* itself. While their proposed power-law models accurately describe the empirical relationship, they do not directly correspond to the interaction probability used in theoretical analysis. To further illustrate this distinction, [Figure 5](#) presents simulation results showing the distribution of a node's contacts as a function of distance, under a power-law interaction probability model. These results demonstrate that the observed contact distribution—commonly measured in empirical studies—decays as $1/d^{\alpha-2}$, rather than as $1/d^\alpha$. However, in capacity estimation research, *contacts distribution has not been distinguished from the interaction probability* and they have been equated implicitly. Overlooking this critical distinction between social ties and interaction probability underestimates network capacity as illustrated in [Figure 6](#). Our work corrects this misinterpretation by introducing an analytical framework that explicitly links the computation of $\mathbf{E}(h)$, a measure of capacity, with $C(d)$, which reflects empirical interaction data—both within the same network structure. This model captures the logarithmic scaling of empirical data using a nested, exponentially growing arrangement and properly accounts for spatial dimensionality, which has been overlooked in prior studies. This structural approach creates a direct mathematical connection between capacity-bound derivations and empirical social tie estimation. Consequently, presenting a robust method for measuring the interaction probability parameter and clarifying its interpretation *enables more accurate predictions of wireless distributed network scalability, fundamentally reshaping conclusions about the scalability of these networks*, as detailed in [Section 6](#).

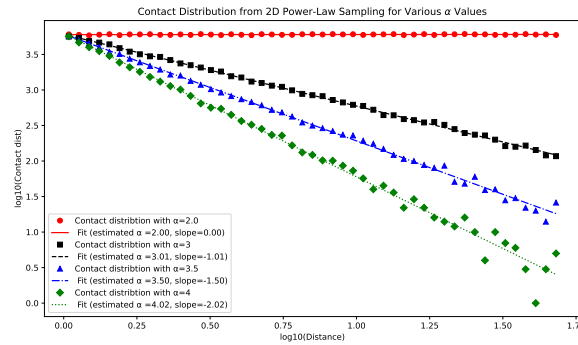


Figure 5: This plot shows the contact count as a function of distance for a node at the center of a 2-D plane, obtained via logarithmic binning, from a simulation of $n = 1.2 \times 10^8$ nodes randomly distributed in a 100×100 square. The probability that each node forms a contact with another node decays with their distance as $P(d) \propto 1/d^\alpha$.

6 Scalability of distributed wireless networks

Using the result from [Section 5](#), we can extract the power-law exponent α directly from empirical datasets. In this section, we first apply [Eq. \(38\)](#) to various empirical datasets covering diverse communication contexts, including social networks, email communication, video games, phone calls, and

physical interactions, to identify a realistic range of α . Subsequently, leveraging the capacity bounds established in Section 4, we analyze how these bounds scale within this identified range and provide a clear insight into the scalability of distributed wireless networks from a capacity perspective.

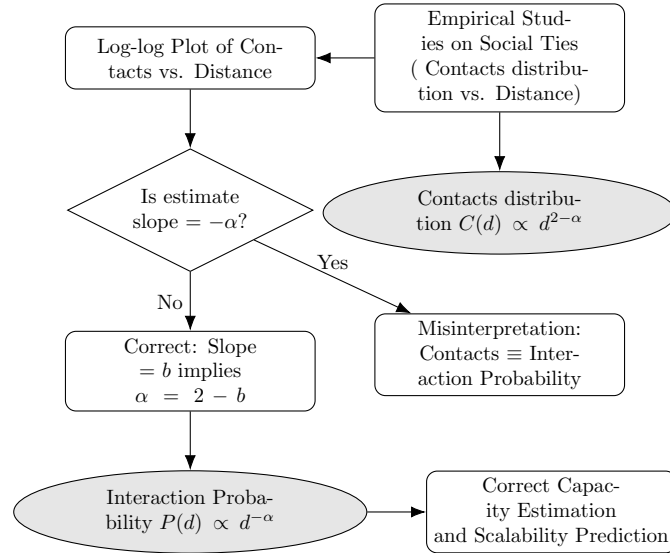


Figure 6: Correction of power-Law exponent interpretation in empirical analysis. The flowchart illustrates how applying the correct relation, $\alpha = 2 - b$, enables accurate derivation of network scalability from empirical data.

6.1 Empirical datasets

Latane [34] analyzed social interactions across different geographic regions, consistently finding a log-log slope near -1 , corresponding to $\alpha \approx 3$. This finding aligns with the social impact theory [34], which posits a fundamental relationship between social ties and distance, independent of technological advancements. More recent large-scale studies further validate these findings. Backstrom et al. [5] analyzed social interactions among 3 million Facebook users, again finding a slope of -1.05 ($\alpha = 3.05$). Similarly, [16] found α values between $[2.969, 2.996]$ in Australian community networks. [52] analyzed data from a massively multiplayer online game where players interact by communicating, trading, or fighting. They measured the frequency of exchanges versus distance on a log-log scale, with an approximated power-law exponent of $\alpha = 3.3$. [37] investigated friendship ties in a social network of 1 million bloggers on LiveJournal, which implies $\alpha = 3.2$. [35] studied social ties on Facebook and email communication, reporting $\alpha = 3.08$ and $\alpha = 2.99$, respectively. [32] studied 3.3 million mobile phone users and estimated the number of connected pairs versus distance, finding a power-law slope of -2 , yielding $\alpha = 4$. This behavior was explained using the gravity model. Similarly, [29] analyzed 2.5 million mobile phone users and showed that inter-city communication follows a gravity model, with interaction frequency proportional to the product of city sizes divided by the square of their distance, leading to $\alpha = 4$. [44] further confirmed this trend in mobile communication, observing α values of 3.58 and 3.49 for voice and text interactions, respectively.

[33] analyzed the online social ties of more than 10 million users of the Tuenti social network. They found that friendships were predominantly local, following a heavy-tailed power-law form. The study identified two distinct regimes: short-range links (under 300 km) with exponents $\alpha \in [3.1, 3.5]$ and long-range links (beyond 300 km) with a steeper exponent of approximately $\alpha = 4.4$. [6] introduced the Social Connectedness Index (SCI) using Facebook friendships in the United States. They observed that the elasticity of friendship ties with distance followed a power-law relationship, with exponents of $\alpha = 3.48$. *This confirms that, despite digital connectivity, geographic distance remains a dominant factor in social interaction.* In a later study, [7] analyzed social connectedness in urban areas, particularly New York City, and found an exponent of $\alpha = 3.23$. The case of NYC demonstrates that *even*

in a highly diverse and technologically advanced metropolis with widespread cellular network coverage (2020), geographic proximity remains a dominant factor influencing social ties. Despite the absence of significant political, cultural, or natural divisions, social connections within the city still cluster into ten distinct subcommunities aligned closely with geographic areas. Furthermore, their study on online friendship between European societies [8] calculated $\alpha = 3.38$ for Paris and identified 50 distinct regional communities in 20 countries.

Studies such as [48] and [15] reported power-law exponents below 3 (e.g., $\alpha = 2.8$ for location-based services and $\alpha = 2.7$ for Twitter data). However, in check-in-based services such as [48], users may travel to multiple cities and keep friends across locations, but their actual interactions remain largely local. This limitation suggests that such analyses may not accurately reflect meaningful social ties in terms of network capacity. Similarly, Twitter differs from typical P2P networks, as it is follower-based and interest-oriented rather than composed of direct social connections [31], where 10% of Twitter users are responsible for 92% of all tweets [51], and 20% of all users possess more than 96% of all followers [60]. Therefore, it has been proposed that interaction graphs should be extracted from broader social networks, such as Facebook, by considering only connections with a significant amount of explicit or implicit communication [28, 56]. Moreover, even in P2P networks, the distribution of friendships does not always align with the distribution of interactions, as [53] demonstrated that the frequency of messages exchanged over distance declines more strongly than the distribution of friendships. For example, people who created a Facebook profile 10 years ago may have migrated to another country. Although they still have many friends from home, over time, their interactions with them become increasingly infrequent.

In summary, the power-law model for P2P human interaction has been widely observed, leading to several theoretical explanations and justifications. These include Zipf's Law [61], Social Impact Theory [34], and the Gravity Model [32, 35, 52]. Additionally, it has been proposed that the spatial structure of social networks is scale-invariant, following a universal contact distribution of $1/d$ [27]. *A common aspect of all these models is that they predict an interaction probability decay of $\Theta(1/d^3)$, and empirical studies on P2P communication consistently support an exponent within this range that supports scalability.* Small variations in these values can be attributed to multiple factors. Part of the discrepancy arises from differences in communication contexts, while other variations stem from geographical, demographic, and political factors. Additionally, the resolution of location data used in analysis significantly impacts results. Many studies rely on approximate user locations, which can vary in accuracy—including city-level resolution [33], IP-based resolution, or zip-code-level resolution [32]—which introduce potential errors in the estimation of the power-law exponent. Finally, while friendships in social networks exhibit strong correlations, they do not necessarily translate into memorable interactions [34, 53]. Social connectedness can be viewed as an upper bound on long-distance interactions, whereas real-world interactions tend to be more localized [53]. This suggests that approximations based on social networks may vary depending on context.

6.2 Scalability analysis

The extracted values of α from multiple sources are given in Table 3. These homogeneous results Table 3 reinforce that social interaction probabilities consistently follow a power law with α values around 3 or exceeding it, which aligns with the *social impact theory* proposed by [34]. However, the cause of this behavior can also be justified by a capacity analysis perspective as the real-world values of α are within the range required for scalable distributed networking. To illustrate it, we consider real-world transportation through the streets of a city, which is a physical type of distributed network. Modeling real-world interaction using a power-law model, path lengths in physical transportation networks follow a similar trend as the expected hop count ($\mathbf{E}(h)$) in distributed networks (Section 4). Consequently, available P2P physical communication capacity closely follows Eq. (28), where W represents communication channel capacity, akin to street width. Under these conditions, if $\alpha > 3$, distributed networking remains scalable regardless of network size. For $\alpha \leq 3$, consider a numerical example

Table 3: Empirical work on social ties. α is calculated as the additive inverse of interaction frequency over distance plus two based on Eq. (38).

Research Work	Case Study	Slope Line (log-log scale)	α
Latane (1995) [34]	Memorable Interaction (US)[fig1]	-1.01	3.01
	Memorable Interaction (China)[fig2]	-1.05	3.05
	Social Sociologists Interaction[fig3]	-0.93	2.93
Backstrom (2010) [5]	Facebook Friendship [fig4]	-1.05	3.05
Onella (2011) [44]	Voice tie (phone call) [Figure 2]	-1.58	3.58
Onella (2011) [44]	Text tie [Figure 2]	-1.49	3.49
Daraganova (2012) [16]	Community network [table 4(λ column)]	[-1,-0.97]	[2.97,3]
Thurner (2015) [52]	Video game (communicating, trading, or fighting.)[Fig 2.]	-1.3	3.3
Lambiotte (2008) [32]	Phone call[Fig 2.]	-2	4
Krings (2009) [29]	Inter-city phone call intensity [Figure 3.b]	-2	4
Liben (2005) [37]	Livejournal bloggers[Fig. 3]	-1.2	3.2
Levy (2014) [35]	Facebook friendship [Fig. 2]	-1.08	3.08
Levy (2014) [35]	Email communication [Fig. 4]	-0.99	2.99
Laniado (2018) [33]	Tuenti Friendship [Fig. 6]	[-1.5,-1.1]	[3.1,3.5]
Bailey (2018) [6]	Facebook Friendship [Table 2 (Column 2)]	-1.48	3.48
(2020) [7]	Facebook Friendship [Table 1]	-1.23	3.23
(2020) [8]	Facebook Friendship [Table 1]	-1.38	3.38

involving two cities with populations $n_1 = 10^5$ and $n_2 = 10^{7.5}$ (≈ 31.6 million), where the real-world empirical distribution is $\Theta(1/d^{2.2})$ ($\alpha = 2.2$). From Eq. (28), the P2P capacity scales as

$$C_{P2P} \propto \frac{W}{n^{\frac{3-2.2}{2}}} = \frac{W}{n^{0.4}}.$$

If W_1 and W_2 are the average street widths in these cities, maintaining equal transportation capacity requires $W_1/n_1^{0.4} = W_2/n_2^{0.4}$. Thus, the larger city would need streets approximately 10 times wider, as:

$$\frac{W_2}{W_1} = \left(\frac{n_2}{n_1}\right)^{0.4} = (10^{2.5})^{0.4} = 10.$$

However, in reality, street widths remain relatively constant or increase only slightly in larger cities. Now, if $\alpha = 3$, keeping the same capacity would require the street width ratio to be:

$$\frac{W_2}{\ln(n_2)} = \frac{W_1}{\ln(n_1)} \Rightarrow \frac{W_2}{W_1} = \frac{\ln(n_2)}{\ln(n_1)} = \frac{7.5 \ln(10)}{5 \ln(10)} = 1.5,$$

which is much closer to reality. These results indicate *scalable distributed communication—whether physical or digital—requires an interaction probability exponent α close to or exceeding 3, ensuring an interaction probability of $\Theta(1/d^3)$ in real-world scenarios*. It serves as the capacity-based mathematical expression of the social impact theory [34].

Returning to our central question, our comprehensive analysis—integrating theoretical capacity bounds with empirical data—demonstrates that multi-hop communication **does not** fundamentally limit the scalability of distributed wireless networks. As shown in Table 3, real-world empirical values of the power-law exponent α typically range from [2.9, 4]. According to the derived expression for C_{P2P} (Eq. (28)), capacity remains stable for $\alpha > 3$, even as the network scales significantly. Furthermore, for values of α in the narrower range [2.9, 3], C_{P2P} decreases gradually without critically impacting scalability. To demonstrate this clearly, Figure 7 visualizes the scaling behavior of the capacity bounds using a base network size of $n = 10^5$. Here, *normalized capacity* is defined as the ratio of C_{P2P} at a given network size $n = y$ relative to $C_{P2P}^{n=10^5}$. As shown in Figure 7, by scaling up the network size even for networks with up to 10^9 nodes, normalized capacity only moderately decreases—by factors of 0.79, 0.63, and 0.57 for $\alpha = 2.95$, 2.9, and 3, respectively. However, significantly lower α values, such

as $\alpha = 2.5$, cause drastic reductions (a factor of approximately 10), potentially challenging scalability. Thus, our analysis confirms that, given realistic empirical ranges for α , even networks comprising millions or billions of users experience minimal reductions in C_{P2P} due to multi-hop communication.

Furthermore, Eq. (2) highlights the key factors influencing network capacity, noting that extensively studied wireless transmission characteristics and network topologies—quantified through parameters $\mathbf{E}(C_L)$ and $r(n)$ —do not fundamentally limit scalability, as discussed in Section 2. Historically, the primary uncertainty has involved the expected hop count $\mathbf{E}(h)$ or communication distance $\mathbf{E}(d)$, whose scalability impact remained unclear. Through rigorous theoretical derivation and empirical evaluation presented in this work around the role of interaction probability, we clarify that: *multi-hop communication impact on P2P capacity does not hinder the scalability of distributed wireless networks* as expected communication distance and number of relays of a connection remain bounded, even for very large networks. These findings have significant implications for the scalability of distributed wireless networks and suggest that large-scale multi-hop communication remains feasible without severe capacity degradation.

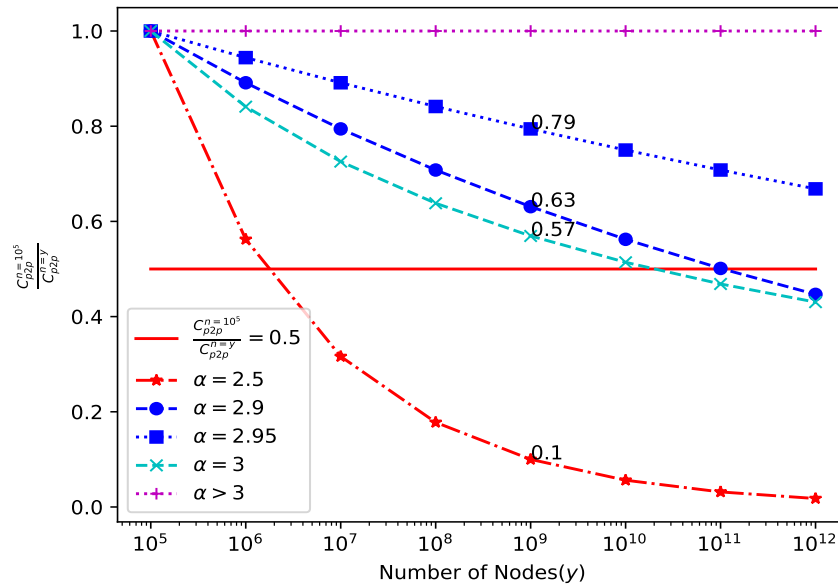


Figure 7: Scaling behavior under different values of the power-law exponent α . Starting from a base network with $n = 10^5$, as depicted in Figure 3, the network is scaled up by increasing the number of nodes while keeping the node density and network structure constant. The plot shows the ratio of asymptotic P2P capacity for a network of size $n = y$ to that of the base network, $C_{P2P}^{n=10^5}$. This ratio is presented as the normalized P2P capacity: $(C_{P2P}^{n=y} / C_{P2P}^{n=10^5})$.

7 Conclusion and future work

This study explored the scalability of distributed networking, where theoretical capacity bounds and statistical datasets on social ties support the potential for large-scale distributed networks. However, additional empirical studies are needed to refine our understanding of P2P traffic patterns and origin-destination communications. Moreover, in practice, the symmetric topology is uncommon; however, it simplifies routing and capacity analysis, providing a structured framework to focus on the role of the interaction probability and multi-hop communication. Although this assumption serves as a useful theoretical baseline for the *asymptotic behavior* of the P2P capacity, future research should explore more realistic assumptions to achieve a comprehensive *numerical performance analysis*. Specifically, the impact of real-world communication models (such as interference and fading), emerging technologies, and practical network topologies should be assessed to provide an accurate numerical evaluation of such networks. Future work should address the following key challenges.

- **Wireless network performance metrics:** Analyzing key performance indicators such as average capacity per user, delay, power consumption, and spectral efficiency concerning wireless transmission factors.
- **Impact of random node arrangements:** Evaluating non-symmetric network topologies to reflect real-world scenarios.
- **Routing and resource allocation:** Investigating efficient routing strategies and resource management to optimize network performance.
- **Empirical studies:** Expanding statistical analysis to better understand various P2P communication traffic patterns.

Moreover, the homogeneous behavior observed in empirical studies can be justified by theoretical work on capacity. This represents an interesting interdisciplinary topic that connects theoretical capacity analysis with social ties beyond digital communication.

Appendix

Theorem 1 (Upper bound on per-node P2P capacity). *Consider a distributed wireless network operating under a TDMA-based communication model, where n denotes the number of nodes and ρ is the node density. Assume that 1) the system is ergodic so that time averages equal ensemble averages, and 2) the total spatial area occupied by active links cannot exceed n/ρ . Let $\mathbf{E}(A)$ be the expected area allocated per active link, $\mathbf{E}(C_L)$ be the expected capacity of a single active link, and $\mathbf{E}(h)$ be the expected hop count of a P2P connection. Then, the P2P capacity C_{P2P} satisfies*

$$C_{P2P} \leq \frac{\mathbf{E}(C_L)}{\mathbf{E}(h) \mathbf{E}(A_\rho)},$$

where $\mathbf{E}(A_\rho) = \mathbf{E}(A) \rho$ is denoted to the expected area occupied by an active link normalized by network density.

Proof. We consider a TDMA-based communication model where, in each time slot, all active links utilize all frequency resources. Let the total network transmission rate in a given time slot be denoted as C_{net} . This total transmission rate is the summation of the capacities of all active links in that slot. Let n_L represent the number of active links in each slot, and let C_L^j denote the capacity of the j -th link. Then, C_{net} can be expressed as:

$$C_{\text{net}} = \sum_{j=1}^{n_L} C_L^j.$$

To ensure convergence when moving from summations to expected values, we assume ergodicity, which ensures that the time average equals the ensemble average for network parameters such as the number of active links and link capacities. The maximum number of active links, n_L , is constrained by the total spatial area required for each link to limit interference. Let A_j represent the area assigned to link j . The total area occupied by all active links must not exceed the total area of the network, given by the product of the number of nodes, n , and the inverse of the density of the nodes, ρ . Thus, we have:

$$\sum_{j=1}^{n_L} A_j \leq \frac{n}{\rho}.$$

Taking the expected value on both sides and applying the linearity of expectation:

$$\mathbf{E} \left(\sum_{j=1}^{n_L} A_j \right) = \mathbf{E}(n_L) \cdot \mathbf{E}(A) \leq \frac{n}{\rho},$$

where $\mathbf{E}(n_L)$ and $\mathbf{E}(A)$ are respectively the expected number of simultaneous active links and the expected area allocated to each active link. Rearranging this inequality provides an upper bound for $\mathbf{E}(n_L)$:

$$\mathbf{E}(n_L) \leq \frac{n}{\mathbf{E}(A) \cdot \rho}. \quad (\text{A.1})$$

The total network transmission rate, C_{net} , is the sum of the capacities of all active links. Taking the expected value:

$$\mathbf{E}(C_{\text{net}}) = \mathbf{E} \left(\sum_{j=1}^{n_L} C_L^j \right).$$

Using the linearity of expectation:

$$\mathbf{E}(C_{\text{net}}) = \mathbf{E}(n_L) \cdot \mathbf{E}(C_L),$$

where $\mathbf{E}(C_L)$ is the expected capacity of a single active link. Substituting the upper bound for $\mathbf{E}(n_L)$ from Eq. (A.1), we obtain:

$$\mathbf{E}(C_{\text{net}}) \leq \frac{n \cdot \mathbf{E}(C_L)}{\mathbf{E}(A) \cdot \rho}. \quad (\text{A.2})$$

On the other hand, the P2P capacity, C_{P2P} , represents the capacity allocated to a single P2P connection. If any node v transmits data on a P2P connection at an average rate of C_{P2P} and a hop count of h_v , the total concurrent transmission resources used by all P2P connections can be expressed as $\sum_{v=1}^n C_{P2P} h_v$. Taking the expectation, we have:

$$\mathbf{E} \left(\sum_{v=1}^n C_{P2P} h_v \right) = C_{P2P} \cdot \mathbf{E} \left(\sum_{v=1}^n h_v \right) = n C_{P2P} \cdot \mathbf{E}(h),$$

where $\mathbf{E}(h)$ is the expected number of hops for a P2P connection. The expected value of the total consumed resources is constrained by the total network transmission rate so that

$$n C_{P2P} \cdot \mathbf{E}(h) \leq \mathbf{E}(C_{\text{net}}).$$

Substituting the upper bound of $\mathbf{E}(C_{\text{net}})$ from Eq. (A.2),

$$n C_{P2P} \cdot \mathbf{E}(h) \leq \frac{n \cdot \mathbf{E}(C_L)}{\mathbf{E}(A) \cdot \rho},$$

By considering $\mathbf{E}(A_\rho) = \mathbf{E}(A) \rho$ as the the expected area occupied by an active link normalized by network density and simplifying the total number of nodes from both sides, the per-node P2P capacity is expressed as

$$C_{P2P} \leq \frac{\mathbf{E}(C_L)}{\mathbf{E}(h) \cdot \mathbf{E}(A_\rho)}. \quad (\text{A.3})$$

□

Lemma 1. *Let a source node be placed at the vertex of an isosceles triangle with κ nodes along each side of equal length, where $\kappa \gg 1$. The sum of all nodes' (minimum) hop counts from the source node within this triangle is approximated by:*

$$F(\kappa) = \frac{\kappa^3}{2}.$$

Proof. The triangle contains κ nodes along each side that can be further divided into κ layers, where layer x contains $x + 1$ nodes, as shown in Figure A.1. The shortest distance from the center node for a node v in layer x is $(x + v)$. Summing over all layers:

$$\sum_{i=1}^{\kappa} \sum_{v=1}^x (x + v) = \sum_{x=1}^{\kappa} \left(x^2 + \sum_{v=1}^x v \right) \approx \sum_{x=1}^{\kappa} \frac{3x^2}{2}.$$

Using the identity $\sum_{x=1}^n x^2 = n(n+1)(2n+1)/6$, we define:

$$F(\kappa) = \sum_{x=1}^{\kappa} \frac{3x^2}{2} = \frac{\kappa(\kappa+1)(2\kappa+1)}{4}.$$

For large κ , we approximate $\lim_{\kappa \gg 1} F(\kappa) = \kappa^3/2$. □

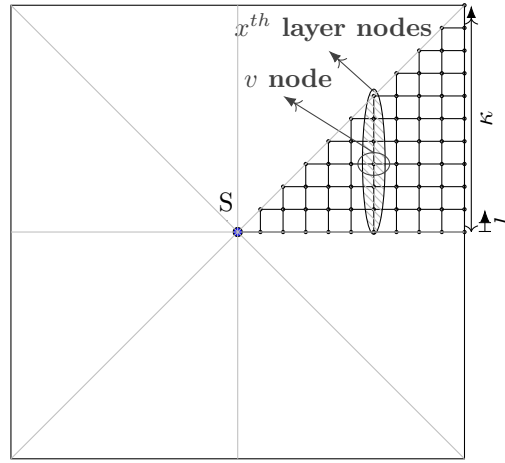


Figure A.1: Isosceles triangle and its node arrangement.

Lemma 2 (Exclusively distance-dependent distribution). *Consider a network with n nodes uniformly distributed in a planar (2D) area with node density $\rho = 1/l^2$. Suppose we sum the number of nodes multiplied by the interaction probability at any distance $d_i = li$ in nested tiers, as depicted in Figure A.2. Then, as $n \rightarrow \infty$:*

$$\sum_{i=1}^{\infty} P(d_i) d_i \quad \text{must converge}.$$

Proof. Suppose each node $v \neq S$ interacts with a source node S with probability $P(d_v)$, dependent exclusively on distance. Then, the total probability satisfies:

$$\sum_{v \neq S}^n P(d_v) = 1.$$

Since nodes at equal distances from the source node have equal interaction probabilities, we can sum the number of nodes multiplied by the interaction probability at any distance $d_i = li$ in nested tiers around the source node, giving:

$$\sum_{i=1}^{i_{\max}} P(d_i) N(d_i) = 1,$$

where $i_{\max}$ is the number of tiers and $N(d_i)$ is the number of nodes within the distance interval $d_i \in [l(i-0.5), l(i+0.5)]$ from the source node, as depicted in Figure A.2. Given the uniform distribution of nodes and approximating $d_i = il$, we have:

$$N(d_i) = \rho \cdot 2\pi [l^2(i+0.5)^2 - l^2(i-0.5)^2] = 2\pi i = \frac{2\pi d_i}{l}.$$

Furthermore, as the diameter of the network is $l\sqrt{n}$, the number of tiers $i_{\max}$ is proportional to $\sqrt{n}$. Thus, the condition becomes:

$$\sum_{i=1}^{\sqrt{n}} 2\pi P(d_i) d_i = 1.$$

To satisfy the condition that the sum of all interaction probabilities equals one, this simplifies to:

$$\sum_{i=1}^{\infty} P(d_i) d_i < \infty.$$

If this series diverges, the cumulative interaction probability would exceed 1 as $n \rightarrow \infty$, violating the assumption that each node must choose exactly one destination. Therefore, for the interaction probability distribution $P(d_i)$ to be valid in a large-scale uniformly distributed network, the series $\sum P(d_i) d_i$ must converge. $\square$

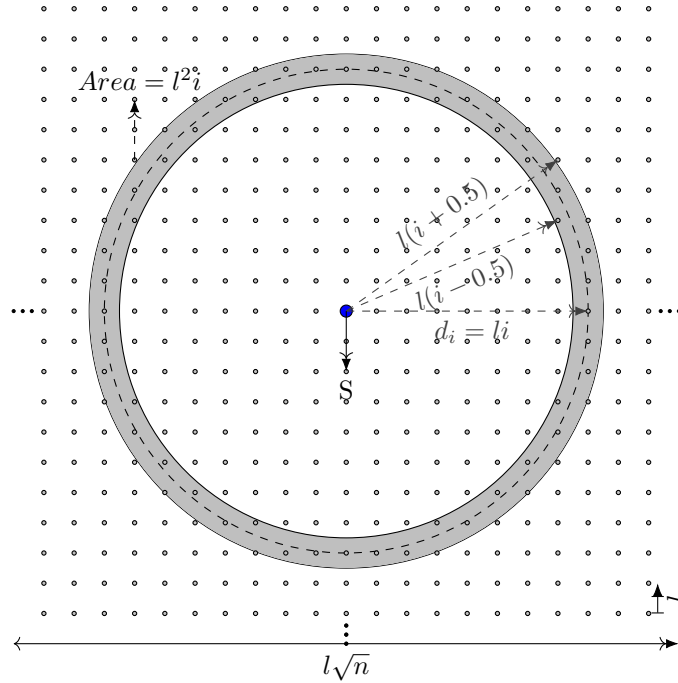


Figure A.2: Grouping nodes of any distance d_i .

Theorem 2. For any exclusively distance-dependent interaction probability, where n nodes are uniformly distributed; if $n \rightarrow \infty$, $P(d)$ should fall faster than $\Theta(1/d^2)$, so that $\lim_{d \rightarrow \infty} d^2 P(d) = 0$.

Proof. In Lemma 2, we showed that the necessary condition on the interaction probability over distance is that $\sum_{d=1}^{\infty} P(d)d$ converges. $P(d)$ is considered as $\Theta(P'(d)/d^\alpha)$, where $\lim_{d \rightarrow \infty} P'(d) = c < \infty$, and $\alpha = \{\max \alpha' \mid P(d) \text{ is } \Theta(P'(d)/d^{\alpha'})\}$. In order to obtain a finite sum for the series $\sum_1^{l\sqrt{n}} P'(d)d^{1-\alpha}$, we use the Cauchy condensation test, which states: if $\{a_n\}$ is a positive monotone decreasing sequence,

then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{k=1}^{\infty} 2^k a_{2^k}$ converges. By applying it to $\sum_1^{\sqrt{n}} P'(d) d^{(1-\alpha)}$, we get $\sum P'(2^d) 2^{d(2-\alpha)}$. If $\alpha > 2$ and $P'(2^d)$ is bounded, the series will converge, while it diverges when $\alpha < 2$, regardless of $P'(2^d)$. When $\alpha = 2$, $\sum P'(2^d)$ converges only if $\sum P'(d)$ is a series such as

$$\sum_n \frac{1}{\log d \cdot \log \log d \cdots \log^{o(k-1)} d \cdot (\log^{o_k} d)^\lambda},$$

where $\lambda > 1$. In other words, to satisfy $\sum_{v=1}^{n-1} P_v = 1$, $P(d)$ should fall faster than $\Theta(1/d^2)$, so that $\lim_{d \rightarrow \infty} d^2 P(d) = 0$.

□

References

- [1] Explainer: Why is it taking so long to re-establish power after the ice storm?, 2022. URL <https://montrealgazette.com/news/local-news/quebec-ice-storm-explainer-how-did-we-get-here-and-what-happens-next/>.
- [2] Ieee connecting unconnected, 2023. URL <https://ctu.ieee.org/>.
- [3] Giusi Alfano, Michele Garetto, Emilio Leonardi, and Valentina Martina. Capacity scaling of wireless networks with inhomogeneous node density: Lower bounds. *IEEE/ACM Transactions on Networking*, 18(5):1624–1636, 2010. doi: 10.1109/TNET.2010.2048719.
- [4] Bitu Azimdoost, Hamid R Sadjadpour, and JJ Garcia-Luna-Aceves. Capacity of wireless networks with social behavior. *IEEE Transactions on Wireless Communications*, 12(1):60–69, 2012. doi: 10.1109/GLOCOM.2016.7841651.
- [5] Lars Backstrom, Eric Sun, and Cameron Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*, pages 61–70, 2010. doi: 10.1145/1772690.1772698.
- [6] Michael Bailey, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–280, 2018. doi: 10.1257/jep.32.3.259.
- [7] Michael Bailey, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel. Social connectedness in urban areas. *Journal of Urban Economics*, 118:103264, 2020. doi: <https://doi.org/10.1016/j.jue.2020.103264>.
- [8] Michael Bailey, Drew Johnston, Theresa Kuchler, Dominic Russel, Bogdan State, and Johannes Stroebel. The determinants of social connectedness in europe. In *Social Informatics*, pages 1–14, Cham, 2020. Springer International Publishing. doi: 10.1007/978-3-030-60975-7_1.
- [9] C Balasubramanian and R Lal Raja Singh. Iot based energy management in smart grid under price based demand response based on hybrid fho-rernn approach. *Applied Energy*, 361:122851, 2024. doi: 10.1016/j.apenergy.2024.122851.
- [10] Mohammed Banafaa, Ibraheem Shayea, Jafri Din, Marwan Hadri Azmi, Abdulaziz Alashbi, Yousef Ibrahim Daradkeh, and Abdulraqeb Alhammadi. 6g mobile communication technology: Requirements, targets, applications, challenges, advantages, and opportunities. *Alexandria Engineering Journal*, 2022.
- [11] M Benzaouia, B Hajji, Adel Mellit, and Abdelhamid Rabhi. Fuzzy-iot smart irrigation system for precision scheduling and monitoring. *Computers and Electronics in Agriculture*, 215:108407, 2023. doi: 10.1016/j.compag.2023.108407.
- [12] Simon Elias Bibri, Alahi Alexandre, Ayyoob Sharifi, and John Krogstie. Environmentally sustainable smart cities and their converging ai, iot, and big data technologies and solutions: an integrated approach to an extensive literature review. *Energy Informatics*, 6(1):9, 2023. doi: 10.1186/s42162-023-00259-2.
- [13] Long Cheng, Yan Gu, Qingzhi Liu, Lei Yang, Cheng Liu, and Ying Wang. Advancements in accelerating deep neural network inference on aiot devices: A survey. *IEEE Transactions on Sustainable Computing*, 2024. doi: 10.1109/TSUSC.2024.3353176.
- [14] Zairan Cheng and Ying Liu. Performance study and optimization of 3d-manet:: A new analytical perspective based on zipf’s law. *Wireless Communications and Mobile Computing*, 2022, 2022. doi: 10.1155/2022/9904257.

- [15] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090, 2011.
- [16] Galina Daraganova, Pip Pattison, Johan Koskinen, Bill Mitchell, Anthea Bill, Martin Watts, and Scott Baum. Networks and geography: Modelling community network structures as the outcome of both spatial and network processes. *Social networks*, 34(1):6–17, 2012. doi: 10.1016/j.socnet.2010.12.001.
- [17] Hanqiang Deng, Jian Huang, Quan Liu, Tuo Zhao, Cong Zhou, and Jialong Gao. A distributed collaborative allocation method of reconnaissance and strike tasks for heterogeneous uavs. *Drones*, 7(2):138, 2023. doi: 10.3390/drones7020138.
- [18] Canh T. Dinh, Nguyen H. Tran, Minh N. H. Nguyen, Choong Seon Hong, Wei Bao, Albert Y. Zomaya, and Vincent Gramoli. Federated learning over wireless networks: Convergence analysis and resource allocation. *IEEE/ACM Transactions on Networking*, 29(1):398–409, 2021. doi: 10.1109/TNET.2020.3035770.
- [19] Mohammadali Farahpoor, Oscar Esparza, and Miguel Soriano. Comprehensive iot-driven fleet management system for industrial vehicles. *IEEE access*, 2023. doi: 10.1109/ACCESS.2023.3343920.
- [20] Luoyi Fu, Wentao Huang, Xiaoying Gan, Feng Yang, and Xinbing Wang. Capacity of wireless networks with social characteristics. *IEEE Transactions on Wireless Communications*, 15(2):1505–1516, 2015. doi: 10.1109/TWC.2015.2491278.
- [21] Andrea Goldsmith, Michelle Effros, Ralf Koetter, Muriel Médard, Asu Ozdaglar, and Lizhong Zheng. Beyond shannon: the quest for fundamental performance limits of wireless ad hoc networks. *IEEE Communications Magazine*, 49(5):195–205, 2011. doi: 10.1109/MCOM.2011.5762818.
- [22] Erika Pamela Silva Gómez and Sang Guun Yoo. Lorawan infrastructure design and implementation for soil moisture monitoring: A real-world practical case. In *International Conference on Computational Science*, pages 133–146. Springer, 2024. doi: 0.1007/978-3-031-63783-4_11.
- [23] Piyush Gupta and Panganmala R Kumar. The capacity of wireless networks. *IEEE Transactions on information theory*, 46(2):388–404, 2000. doi: 10.1109/18.825799.
- [24] Xinxin he, Weisen Shi, and Tao Luo. Transmission capacity analysis for vehicular ad hoc networks. *IEEE Access*, 6:30333–30341, 2018. doi: 10.1109/ACCESS.2018.2843333.
- [25] Jiancao Hou and Mohammad Shikh-Bahaei. Transmission capacity of full-duplex mimo ad-hoc network with limited self-interference cancellation. In *2018 IEEE Global Communications Conference (GLOBE-COM)*, pages 1–7. IEEE, 2018. doi: 10.1109/GLOCOM.2018.8647531.
- [26] Ronghui Hou, Yu Cheng, Jiandong Li, Min Sheng, and King-Shan Lui. Capacity analysis of hybrid wireless networks with long-range social contacts behavior. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 2209–2217. IEEE, 2015. doi: 10.1109/INFOCOM.2015.7218607.
- [27] Yanqing Hu, Yougui Wang, Daqing Li, Shlomo Havlin, and Zengru Di. Possible origin of efficient navigation in small worlds. *Physical Review Letters*, 106(10):108701, 2011. doi: 10.1103/PhysRevLett.106.108701.
- [28] Jing Jiang, Christo Wilson, Xiao Wang, Wenpeng Sha, Peng Huang, Yafei Dai, and Ben Y Zhao. Understanding latent interactions in online social networks. *ACM Transactions on the Web (TWEB)*, 7(4): 1–39, 2013. doi: 10.1145/1879141.1879190.
- [29] Gautier Krings, Francesco Calabrese, Carlo Ratti, and Vincent D Blondel. Urban gravity: a model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(07): L07003, 2009. doi: 10.1088/1742-5468/2009/07/L07003.
- [30] Heorhii Kuchuk and Eduard Malokhvii. Integration of iot with cloud, fog, and edge computing: A review. *Advanced Information Systems*, 8(2):65–78, 2024. doi: 10.20998/2522-9052.2024.2.08.
- [31] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600, 2010. doi: 10.1145/1772690.1772751.
- [32] Renaud Lambiotte, Vincent D Blondel, Cristobald De Kerchove, Etienne Huens, Christophe Prieur, Zbigniew Smoreda, and Paul Van Dooren. Geographical dispersal of mobile communication networks. *Physica A: Statistical Mechanics and its Applications*, 387(21):5317–5325, 2008. doi: 10.1016/j.physa.2008.05.014.
- [33] David Laniado, Yana Volkovich, Salvatore Scellato, Cecilia Mascolo, and Andreas Kaltenbrunner. The impact of geographic distance on online social interactions. *Information Systems Frontiers*, 20(6):1203–1218, 2018. doi: 10.1007/s10796-017-9784-9.

- [34] Bibb Latané, James H Liu, Andrzej Nowak, Michael Bonevento, and Long Zheng. Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin*, 21(8):795–805, 1995. doi: 10.1177/0146167295218002.
- [35] Moshe Levy and Jacob Goldenberg. The gravitational law of social interaction. *Physica A: Statistical Mechanics and its Applications*, 393:418–426, 2014. doi: 10.1016/j.physa.2013.08.067.
- [36] Jinyang Li, Charles Blake, Douglas SJ De Couto, Hu Imm Lee, and Robert Morris. Capacity of ad hoc wireless networks. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 61–69, 2001. doi: 10.1145/381677.381684.
- [37] David Liben-Nowell, Jasmine Novak, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences*, 102(33):11623–11628, 2005. doi: 10.1073/pnas.0503018102.
- [38] Marin Lujak, Elizabeth Sklar, and Frederic Semet. Agriculture fleet vehicle routing: A decentralised and dynamic problem. *AI Communications*, 34(1):55–71, 2021. doi: 10.3233/AIC-201581.
- [39] Maurilio Matraccia, Mustafa A Kishk, and Mohamed-Slim Alouini. Aerial base stations for global connectivity: Is it a feasible and reliable solution? *IEEE Vehicular Technology Magazine*, 2023. doi: 10.1109/MVT.2023.3301228.
- [40] Wided Moulahi, Imen Jdey, Tarek Moulahi, Moatsum Alawida, and Abdulatif Alabdulatif. A blockchain-based federated learning mechanism for privacy preservation of healthcare iot data. *Computers in Biology and Medicine*, 167:107630, 2023. doi: 10.1016/j.compbimed.2023.107630.
- [41] Luong Vuong Nguyen. Swarm intelligence-based multi-robotics: A comprehensive review. *AppliedMath*, 4(4):1192–1210, 2024. doi: 10.3390/appliedmath4040064.
- [42] Yuqiang Ning and Lili Du. Robust and resilient equilibrium routing mechanism for traffic congestion mitigation built upon correlated equilibrium and distributed optimization. *Transportation research part B: methodological*, 168:170–205, 2023. doi: 10.1016/j.trb.2022.12.006.
- [43] Lionel Nkenyereye, Kang-Jun Baeg, and Wan-Young Chung. Deep reinforcement learning for containerized edge intelligence inference request processing in iot edge computing. *IEEE Transactions on Services Computing*, 16(6):4328–4344, 2023. doi: 10.1109/TSC.2023.3320752.
- [44] Jukka-Pekka Onnela, Samuel Arbesman, Marta C González, Albert-László Barabási, and Nicholas A Christakis. Geographic constraints on social network groups. *PLoS one*, 6(4):e16939, 2011. doi: 10.1371/journal.pone.0016939.
- [45] Dongrun Qin and Zhi Ding. Transport capacity analysis of wireless in-band full duplex ad hoc networks. *IEEE Transactions on Communications*, 65(3):1303–1318, 2016. doi: 10.1109/TCOMM.2016.2640278.
- [46] Sita Rani, Meetali Chauhan, Aman Kataria, and Alex Khang. Iot equipped intelligent distributed framework for smart healthcare systems. In *Towards the Integration of IoT, Cloud and Big Data: Services, Applications and Standards*, pages 97–114. Springer, 2023. doi: 10.1007/978-981-99-6034-7_6.
- [47] Dukka Karun Kumar Reddy, Janmenjoy Nayak, HS Behera, Vimal Shanmuganathan, Wattana Viriyasitavat, and Gaurav Dhiman. A systematic literature review on swarm intelligence based intrusion detection system: past, present and future. *Archives of Computational Methods in Engineering*, pages 1–68, 2024. doi: 10.1007/s11831-023-10059-2.
- [48] Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, and Cecilia Mascolo. Socio-spatial properties of online location-based social networks. In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 329–336, 2011. doi: 10.1609/icwsm.v5i1.14094.
- [49] R Senthamil Selvan and A Senthil Kumar. Original research article wireless sensor network and iot based v2v connectivity, minimization of hazardous environment. *Journal of Autonomous Intelligence*, 7(5), 2024. doi: 10.32629/jai.v7i5.1604.
- [50] Sujie Shao, Juntao Zheng, Shaoyong Guo, Feng Qi, and I Xuesong Qiu. Decentralized ai-enabled trusted wireless network: A new collaborative computing paradigm for internet of things. *IEEE Network*, 37(2): 54–61, 2023. doi: 10.1109/MNET.002.2200391.
- [51] The Social Shepherd. 53 twitter statistics marketers must know in 2024, 2024. URL <https://thesocialshepherd.com/blog/twitter-statistics>. Accessed: 2025-03-11.
- [52] Stefan Thurner and Benedikt Fuchs. Physical forces between humans and how humans attract and repel each other based on their social interactions in an online world. *PLoS one*, 10(7):e0133185, 2015. doi: 10.1371/journal.pone.0133185.
- [53] Yana Volkovich, Salvatore Scellato, David Laniado, Cecilia Mascolo, and Andreas Kaltenbrunner. The length of bridge ties: structural and geographic properties of online social interactions. In *Proceedings*

- of the International AAAI Conference on Web and Social Media, volume 6, pages 346–353, 2012. doi: 10.1609/icwsm.v6i1.14271.
- [54] Yuhua Wang, Laixian Peng, Renhui Xu, and Shuai Cheng. The upper bound of capacity for multi-mode networking with multi-beam antenna array. In 2021 7th International Conference on Computer and Communications (ICCC), pages 2143–2148. IEEE, 2021. doi: 10.1109/ICCC54389.2021.9674633.
- [55] Zhiqing Wei, Huici Wu, Xin Yuan, Sai Huang, and Zhiyong Feng. Achievable capacity scaling laws of three-dimensional wireless social networks. *IEEE Transactions on Vehicular Technology*, 67(3):2671–2685, 2018. doi: 10.1109/TVT.2017.2772797.
- [56] Christo Wilson, Bryce Boe, Alessandra Sala, Krishna PN Puttaswamy, and Ben Y Zhao. User interactions in social networks and their implications. In *Proceedings of the 4th ACM European conference on Computer systems*, pages 205–218, 2009. doi: 10.1145/1519065.1519089.
- [57] Shiyang Wu, Andrew CM Austin, Ameer Ivoghlian, Akshat Bisht, and Kevin I-Kai Wang. Long range wide area network for agricultural wireless underground sensor networks. *Journal of Ambient Intelligence and Humanized Computing*, 14(5):4903–4919, 2023. doi: 10.1007/s12652-020-02137-1.
- [58] X. Yin, Y. Zhang, Y. Wang, et al. Multi-region asynchronous swarm learning for data sharing in large-scale internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 2023. doi: 10.1109/LCOMM.2023.3314662.
- [59] Yang Zhou, Yan Shi, and Shanzhi Chen. Capacity and delay analysis for large social-aware mobile ad hoc wireless networks. *Applied Sciences*, 10(5):1719, 2020. doi: 10.3390/app10051719.
- [60] Linhong Zhu and Kristina Lerman. Attention inequality in social media. *arXiv preprint arXiv:1601.07200*, 2016. doi: 10.48550/arXiv.1601.07200.
- [61] George Kingsley Zipf. *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio books, 2016.