

Inducing Spatial Clustering in MAC Contention for Spread Spectrum Ad Hoc Networks *

Xiangying Yang and Gustavo de Veciana

Department of Electrical and Computer Engineering
the University of Texas at Austin
Austin, TX 78712
{yangxy,gustavo}@ece.utexas.edu

ABSTRACT

This paper proposes a new principle for designing MAC protocols for spread spectrum based ad hoc networks – inducing spatial clustering in contending transmitters/receivers. We first highlight the advantages of spread spectrum in handling quality of service (QoS) requirements, enhancing energy efficiency, and enabling spatial multiplexing of bursty traffic. Then, based on stochastic geometric models and simulation, we show how idealized contention resolution among randomly distributed nodes results in clustering of successful transmitters and receivers, in turn leading to efficient spatial reuse. This motivates the central idea of the paper which is to explicitly induce clustering among contending nodes to achieve even better spatial reuse. We propose two distributed mechanisms to realize such clustering and show substantial capacity gains over simple random access/ALOHA-like and even RTS/CTS based protocols. We examine under what regimes such gains can be achieved, and how clustering and contention resolution mechanisms should be optimized to do so. We propose the design of ad hoc networks supporting hop-by-hop relaying on different spatial scales. By allowing nodes to relay beyond the set of nearest neighbors using varying transmission ranges (scales), one can reduce the number of hops between a source and destination so as to meet end-to-end delay requirements. To that end we propose a multi-scale MAC clustering and power control mechanism to support transmissions with different ranges while achieving high spatial reuse. The considerations, analysis and simulations included in this paper suggest that the principle of inducing spatial clustering in contention has substantial promise towards achieving high spatial reuse, QoS, and energy efficiency in spread spectrum ad hoc networks.

*This work was supported in part by NSF Grant CNS 0435307 and NSF Grant 0225448.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobiHoc'05, May 25–27, 2005, Urbana-Champaign, Illinois, USA.
Copyright 2005 ACM 1-59593-004-3/05/0005 ...\$5.00.

Keywords

Ad hoc networks, spread spectrum, MAC, clustering.

1. INTRODUCTION

Recently there has been extensive research towards understanding the asymptotic ‘capacity’ scaling of fixed and mobile ad hoc networks see e.g.[7][6]. In practice, however, maximizing capacity is but one of a myriad of possible design goals. At least two other objectives are critical in some ad hoc network applications: energy efficiency and quality of service(QoS). In order to design such networks one must be able to appreciate tradeoffs among the various figures of merit. One should not, for example, consider system capacity without tying it to energy efficiency, or, for some applications, consider ‘capacity’ without an understanding of the quality of service, e.g., delays, one will incur. In this paper we consider spread spectrum based ad hoc networks. Our premise is that such networks are well suited to meet quality of service and energy efficiency requirements while still providing good overall network capacity if suitable MAC and routing algorithms are devised. Our focus will be on ad hoc networks that are fairly *dense*, i.e., the number of nodes within a typical node’s transmission range is high, and in which an individual traffic load generated by a node may be bursty but requires only a fraction of the total system ‘capacity’, e.g., data and voice traffic. Below we discuss our motivation for considering this scenario in more detail along with related work in this area.

Is nearest neighbor routing optimal? One of the basic insights provided by recent work on the capacity scaling of ad hoc networks is that it is maximized by relaying traffic along nearest neighbor paths to a destination. Indeed it turns out to be better to maximize the density of concurrent transmissions in order to achieve a maximum amount of forward progress, i.e., bits-meter/sec. However, long range relaying may still be appealing considering the following aspects:

End-to-end delay. In practice when nearest neighbor routing is used, a packet may need to be relayed by a large number of nodes prior to reaching its destination. Each intermediate node would typically incur a delay, depending on the MAC protocol’s contention overheads, making it difficult to meet end-to-end delay requirements.

Energy. Using nearest neighbor routing, intermediate nodes would typically be switching among transmit, receive and idle modes, further decreasing the amount of time they can spend in the sleep mode. Depending on the actual energy

characteristics of the nodes the first three nodes can be fairly energy hungry[16]. An alternative would be to permit nodes to use longer transmission distances and relay through nodes that are in a larger ‘neighborhood.’ This would allow for paths with fewer hops between the source and destination, possibly enabling a larger number of nodes to dwell in the energy efficient sleep mode.

Load balancing. More nodes are available within longer transmission range as candidates for relaying and thus routing protocols may exploit this to achieve load balancing, which in turn results in better delay performance and energy savings.

However, long range relaying might require higher transmission power levels for successful reception, and increase interference, possibly compromising network capacity, since more nodes are affected by the associated higher power levels – this is particularly the case in *dense* networks. This could be very inefficient when traffic is bursty, as additional contention is required to serialize transmissions, and might lead once again to poor delay performance. An approach which enables concurrent overlapping of transmissions and flexible resource allocation among traffic is thus desirable.

Why use spread spectrum? In a regime where (some) nodes use longer transmission ranges, with a view on meeting QoS service requirements and/or better energy efficiency, a CDMA based physical layer has some key advantages – some of which are akin to those already exploited in cellular networks. Specifically consider a direct sequence CDMA (DS-CDMA) system with spreading factor m . When a receiver de-spreads its received signal using the associated code, it can recover the transmitted signal with a reduced interference from other transmissions. It roughly sees only $1/m$ of the interference before de-spreading.¹ As long as the received signal-to-interference-plus-noise ratio (SINR) after de-spreading exceeds a certain threshold β , the transmission will be successful. This ability of a receiver to decode a signal in the presence of a substantial number of *concurrent* transmissions is referred as *interference averaging* and has the following advantages:

‘Spatial multiplexing’. Because receivers can tolerate such fluctuations in interference, the network can statistically multiplex concurrent overlapping bursty traffic, e.g., on/off voice streams. A general example of such ‘spatial multiplexing’ is shown in Fig. 1, where applications sharing the network may have different QoS requirements and possibly require different relay scales, e.g., delay-sensitive applications may prefer to use long relay distances and high transmission power to achieve reduced end-to-end delays while best effort traffic can use shorter relay distances and low transmission power allowing for enhancing overall throughput. Such multiscale spatial multiplexing achieves efficient spatial/spectrum reuse, even with heterogeneous flows.

Power control. The ability to average interference, can also help in managing spatial inhomogeneities in nodes’ locations, and permit distributed power control with graceful degradation – e.g., when the power levels among various transmit nodes are not optimally selected due to the lack of central management/global information in an ad hoc network.

Robustness. Critical problems such as hidden or exposed

¹We consider codes generated by PN sequences which are quasi-orthogonal and conservatively assume that PN code cross-correlation is $\frac{1}{m}$ [4].

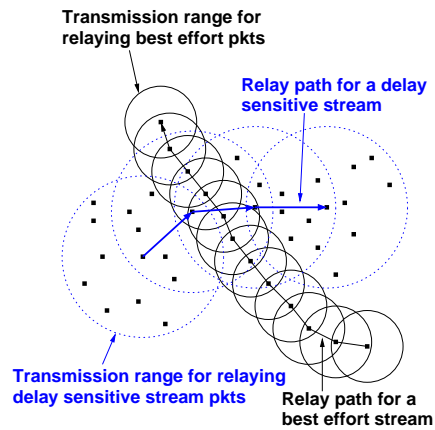


Figure 1: An example of ‘spatial multiplexing’, i.e., heterogenous traffic with different QoS requirements, using different transmit ranges, and co-existing on an ad hoc network. Note that adjacent hops can not be active at the same time because we assume nodes can only either transmit or receive.

terminals in narrow band systems are no longer as prominent in spread spectrum based ad hoc networks, while the ‘near-far’ problem can be mitigated via spreading gain and proper power control. These advantages may potentially simplify MAC design and operation.

What are the design challenges for efficient MAC?

In addition to a spread spectrum based physical layer, supporting applications with different QoS requirements and achieving energy savings also require MAC design to address several challenges. First, to allow nodes to use ‘long’ transmission ranges without significantly compromising capacity requires a MAC that achieves particularly efficient spatial reuse. Indeed, it has been shown that the fundamental scaling of capacity decreases quadratically with transmission range [8][1][19]. Second, nodes may use power control in order to realize energy savings. Yet heterogeneous power levels cause carrier sensing, commonly used in MAC for collision avoidance purposes, to be unreliable since a transmitter with strong power may fail to detect an existing transmission using small transmission power and thus severely interfere with that transmission. In this paper we investigate fundamental strategies to devise such MAC protocols and exploit the tradeoffs among QoS, energy efficiency and spatial reuse in spread spectrum ad hoc networks. Detailed implementations are not in the scope of this paper.

Related work. There has been a history and recently a resurgence of interest in using spread spectrum in ad hoc networks [12][14], suggesting some of the advantages mentioned above.

MAC protocol design for spread spectrum ad hoc networks has been studied in recent work. In [3] joint power control and scheduling are considered so as to achieve optimal spatial reuse but assuming a centralized scheduler to solve a global optimization problem. Such centralized implementation may not be feasible in a distributed system. In [14] a random access/ALOHA-like protocol, which is simple to implement but inefficient from a spatial reuse perspective, is proposed. This will be shown later in this paper. MAC designs using local signaling among nodes [5][9][11], e.g., exchanging request-to-send(RTS)/(clear-to-send)CTS mes-

sages over a common code/frequency control channel, realize transmitter-receiver hand-shaking and may improve the performance over random access schemes. However, they usually assume signaling itself in the control channel is contention free, which may not be true in a practical system.

Main contributions and organization of the paper.

In Section 2, we introduce an appropriate model and notation, which will serve to highlight the role interference plays in a spread spectrum system. In Section 3, we review some basic MAC designs and discuss their ability to realize trade-offs among QoS, energy and capacity. Then, by studying an idealized contention resolution mechanism, we identify a key feature of efficient MAC protocols for spread spectrum based ad hoc networks – clustered packing of concurrent transmissions. In Section 4, we motivate the study of a novel MAC design approach: inducing spatial clustering among contending nodes to further enhance spatial reuse. To this end, we propose and analyze two distributed algorithms that achieve such clustering. Through analysis and simulation we are able to show substantial capacity gains over simple random access/ALOHA-like and RTS/CTS based protocols, when clustering mechanisms are properly optimized. We propose a multi-scale MAC clustering and power control mechanism to enable transmissions with different ranges and support heterogeneous applications QoS requirements, while still achieving efficient spatial reuse. We briefly discuss some implementation issues in Section 5 and conclude in Section 6.

To our knowledge there is no previous work suggesting the benefits of clustered contention processes *at the MAC layer* of spread spectrum ad hoc networks nor how they might be optimized to realize its substantial promise towards achieving high spatial reuse, QoS and energy efficiency.

2. MODELING INTERFERENCE AND SPATIAL REUSE IN SPREAD SPECTRUM SYSTEMS

The capacity of an ad hoc network is fundamentally constrained by interference among concurrent transmissions. An efficient MAC design requires one to achieve a high degree of spatial reuse while maintaining acceptable interference levels among scheduled transmissions. In this section, we define a network and traffic model, to capture the nature of interference in spread spectrum ad hoc networks and propose a novel ‘dumbbell’ model as a means to visualize spatial reuse in such systems.

Network and traffic model. We begin by introducing, and then elaborating on, a simple stochastic geometric model for transmitters and receivers in an ad hoc network. The simplicity is key to allowing tractable analysis, yet the salient characteristics are still captured. We assume that a set of transmit nodes (including nodes relaying packets) are spatially distributed according to a homogenous Poisson point process with intensity λ [18]. Nodes are interchangeably referred to/by their locations. Each transmitter is assumed to be sending to a receiver, which is modeled as being at a random location a distance d away. For simplicity, we will assume receive nodes, are always available at these randomly selected locations. For reasons explained earlier, our initial focus and model are geared at investigating the scenario where the transmit range d exceeds the typical nearest neighbor distance. The model captures a homogenous

offered load where packets are typically relayed along hops with a transmission range d , leading to a homogenous distribution of transmitters. We shall further assume there is no mobility in the time scale of transmissions, and that transmissions are synchronous, or at least approximately so. As discussed in [17] an approximate synchronization provides significant advantages. See also [14][13] for representative protocols based on synchronized contention.

Interference and outage probability. We capture the spatial attenuation of signal power using a basic path loss model where if a transmitter uses a power level ρ the receive power at a distance d is given by $\rho_r(d) = \rho \times d^{-\alpha}$. The path loss exponent α is typically assumed to be between 3 and 5. A receiver, in our model, sees the degraded powers from other concurrent transmitters, as interference, albeit reduced by spread spectrum processing gain. Outage happens when the SINR at the receiver does not exceed a certain threshold, resulting in an unsuccessful transmission.

To evaluate the outage probability we condition on a typical transmitter at the origin O giving what is known as the Palm distribution for transmitters on the plane [18]. By Slivnyak’s Theorem [18] this conditional distribution is also a homogenous Poisson point process with intensity λ with an additional point at the origin. Now shifting this entire point process so that the receiver associated with the desired transmitter lies at the origin, we have that conditional distribution of interferers (excluding the transmitter of interest) is a homogenous Poisson point process with intensity λ . We shall denote this shifted point process of *interferer* locations by $\Pi = \{X_i, i \in \mathbb{N}\}$. As mentioned earlier an outage event occurs when, after de-spreading, the SINR is below some threshold, β . Thus mathematically, for conventional DS-CDMA, the outage probability for a typical receiver $p_o(\lambda, d)$, which depends on the intensity of transmitters and transmission range, is given by

$$p_o(\lambda, d) \approx \mathbb{P} \left(\frac{\rho_r(d)}{\sum_{X_i \in \Pi} \rho |X_i|^{-\alpha}} \leq \frac{\beta}{m} \right), \quad (1)$$

where $|X_i|$ denotes the Euclidean distance from interferer i to the receiver located at the origin.² Note we have neglected the role of ambient noise since the capacity of a dense network is mostly interference constrained. The outage probability for a transmission is a critical performance metric, because it not only represents a failed transmission, but also, wasted energy, violation in QoS requirements, e.g., possibly by requiring retransmission, and thus inducing additional overall load on the network.

Critical interference range. The interference term defined in (1), i.e., $\sum_{X_i \in \Pi} \rho |X_i|^{-\alpha}$ does not generally have a closed-form distribution. This makes the outage probability difficult to compute. Note the special case where $\alpha = 4$ does have a closed-form solution, see [15], which we will use later. There are, however, more intuitive ways to understand interference and outage in this model. In particular, let us consider the rough geometry of transmission and interference ranges.

Note that the aggregate interference is the sum of non-

²Such attenuation law has a simple form but is unrealistic when $|X_i| < 1$. However it will not change the analytical results since there will be an outage anyway. One can use more realistic attenuation functions like $|X_i + 1|^{-\alpha}$ and the analysis will basically remain the same.

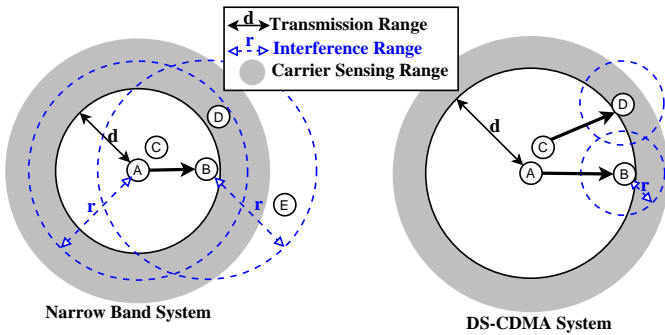


Figure 2: The transmission range, interference range and carrier sensing ranges for an idealized narrow-band and spread spectrum system. Since $A \rightarrow B$ requires no concurrent transmission in the critical interference range r around a B , $C \rightarrow D$ is not allowed in a narrow-band system, but may be allowed in the CDMA system.

i.i.d. random variables. Interferers in the ‘near’ field of a typical receiver contribute very strong interference. ‘Far’ field interferers may still potentially contribute enough aggregate interference to cause an outage although each transmitter only contributes a small amount due to path loss. To quantify near/far fields, we define the *critical interference range* r to be the largest distance from a receiver at which a *single* interferer could be located and still cause an outage, i.e., the largest r such that

$$\frac{\rho d^{-\alpha}}{\rho r^{-\alpha}} \leq \frac{\beta}{m} \implies r = \left(\frac{\beta}{m}\right)^{\frac{1}{\alpha}} d. \quad (2)$$

Thus roughly a disc of radius r around each successful receiver should contain no transmitter. Note that the suppression range of radius r is only a *necessary* condition for successful reception because far field interferers may still contribute an aggregate interference high enough to cause an outage. However as discussed in the sequel, a suppression range of radius r is indeed close to a sufficient condition.

There is a fundamental difference in considering interference in a narrow band versus a spread spectrum system. As shown in Fig. 2, r is usually larger than the transmission range d in a narrow-band system. By contrast, in a spread spectrum system, r would be smaller than d if the spreading factor m were large. This allows one to schedule concurrent overlapped transmissions, as long as there is no interfering transmitter within a range r of another receiver. Note that r depends on the transmission range and associated power control strategies and the level of allowable overlapping depends on m .

We can show that near-field interference as captured by nodes within a distance r from a receiver, see (2), provides a reasonable approximate abstraction for the relevant source of interference, particularly when α is large. To see this, let $\mathbf{B}(x, r)$ denote a ball centered at x with a radius r . We let the event \mathbf{E}_1 denote the occurrence that *at least* one interferer is within $\mathbf{B}(O, r)$ which in turn would necessarily cause an outage for a receiver at the origin. It follows that the outage probability for a typical receiver $p_o(\lambda, d)$ is such that

$$p_o(\lambda, d) \geq \mathbb{P}(\mathbf{E}_1), \quad (3)$$

For a Poisson point process with intensity λ , the probability of \mathbf{E}_1 is given by

$$p_o(\lambda, d) \geq \mathbb{P}(\mathbf{E}_1) = 1 - e^{-\lambda\pi r^2} = 1 - e^{-\lambda\pi\left(\frac{\beta}{m}\right)^{\frac{2}{\alpha}} d^2}.$$

As shown in Fig. 3, in the low outage regime, the outage lower bound provided by $\mathbb{P}(\mathbf{E}_1)$ is very accurate. Indeed, one can analytically show this bound has a slope of $\pi d^2 \lambda \left(\frac{\beta}{m}\right)^{\frac{2}{\alpha}}$ at the point of zero outage, which equals to the slope of exact outage probability solution when $\alpha = 4$ [15]. The slope, i.e., the linearization of the outage probability/lower bound is also shown in Fig. 3. The point of this analysis is to support the intuitive abstraction for the interference and outage in a network where nodes randomly distributed: nearby interferers within interference range r contribute most outage and thus only considering near-field interference, is reasonably accurate. This abstraction allows simple consideration on how contention among transmitters-receivers occurs.

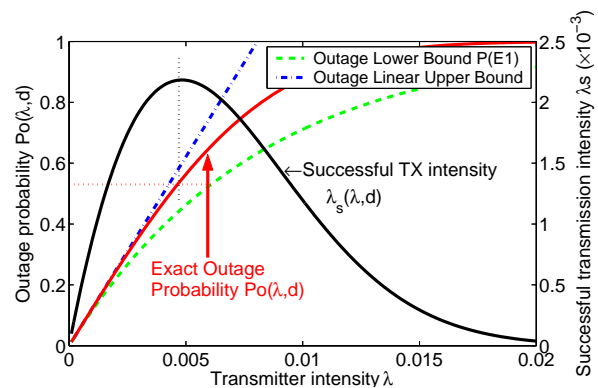


Figure 3: Effective capacity λ_s is maximized when $p_o \approx 0.5$. Also shown are our outage lower bound and its linear approximation, which actually serves as a tight outage upper bound. All bounds/approximations are close to the exact analytical result in the low outage regime for $p_o < 0.5$.

Dumbbell model: modeling concurrent transmissions and spatial reuse in spread spectrum ad hoc networks. Consider a spatially distributed set of transmitters. We assume that a transmitting node suppresses all receive nodes within a critical range r – see (2). Indeed the high interference from such nodes would preclude successful reception. Similarly, a receiver must have no interfering transmitters within a distance r of itself. As discussed earlier, these are only necessary conditions nevertheless good approximations to ensure successful transmissions. Equivalently, as shown in the left panel in Fig. 4, each transmission corresponds to a dumbbell with disks of radius $r/2$ at the transmitter and receiver, connected by a bar of length d . A successful transmission is modeled by a dumbbell without prohibited overlaps, i.e., no transmit disc overlaps with a receive disc on either end.³ Thus, among the three contending transmissions on the left panel in Fig. 4, only two transmissions can be successful. This is shown on the right panel in Fig. 4. Spatial reuse corresponds to realizing a high spatial density of dumbbells subject to *at least* satisfying the

³Note that, if two such disks overlap, then the associated nodes are within r of each other.

rules on overlaps. We shall use this dumbbell model to illustrate contention and later clustering phenomena among transmissions. Subsequent MAC designs will not be based on this dumbbell model and our simulations will factor the actual interference seen by receiver, i.e., both near and far field contributions.

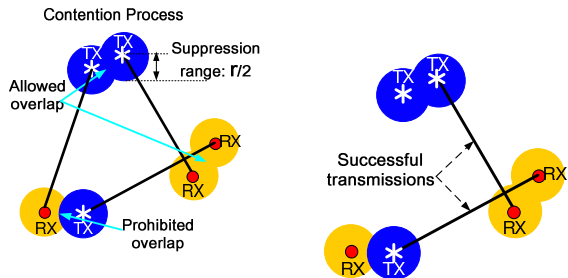


Figure 4: On the left contentions among three concurrent transmissions. On the right, after contentions, only those two transmissions whose receivers do not have prohibited overlap survive.

3. CHARACTERISTICS OF MAC PROTOCOLS

Let us reconsider the previous analysis results to understand basic MAC protocols including random access and contention resolution schemes, from the perspective of the overall system ‘capacity’, outage probability, and associated overheads. We will show that they tend to be very inefficient. Further by considering an idealized contention resolution mechanism, we will motivate a novel approach based on inducing clustering among MAC layer contenders.

3.1 Random access MAC protocol

Maximizing capacity. We consider a simple model for a random access protocol with no carrier sensing or hand-shaking akin to those in [14][15][1][19]. The basic premise is that spatially distributed nodes randomly (or based on a predefined pseudo-random sequence) choose to be transmitters and do so without contending/signaling with each other. The lack of a coordination phase among concurrent transmitter-receivers reduces overheads but increases the likelihood of failed transmissions. One might ask, what is the spatial intensity of contenders that maximizes the intensity of successful transmissions given a fixed transmission distance d .

Fig. 3 exhibits the tradeoff between the outage probability $p_o(\lambda, d)$, the intensity of successful transmissions $\lambda_s(\lambda, d) = \lambda(1 - p_o(\lambda, d))$ and the intensity of contenders λ . The figure shows numerical results for $\lambda_s(\lambda, d)$ using the exact analysis of outage probability in [15], and the approximate linear upper bound discussed earlier, i.e.,

$$\lambda_s(\lambda, d) \geq \lambda \left(1 - \pi d^2 \lambda \left(\frac{\beta}{m} \right)^{\frac{2}{\alpha}} \right). \quad (4)$$

One can in principle determine the λ which maximizes $\lambda_s(\lambda, d)$, say λ^* , which according to (4) is roughly given by

$$\lambda^* = \frac{1}{2\pi d^2} \left(\frac{m}{\beta} \right)^{\frac{2}{\alpha}} \quad \text{and} \quad \lambda_s(\lambda^*, d) = \frac{1}{4\pi d^2} \left(\frac{m}{\beta} \right)^{\frac{2}{\alpha}}. \quad (5)$$

Note from Fig. 3 to achieve a maximal capacity, one incurs a high outage probability, roughly 0.5. This observation, also

holds for the exact analysis in [1], wherein transmitters use an exponentially distributed transmission power and outage probability at the optimal contender intensity maximizing the capacity is roughly $1 - e^{-1} \approx 0.63$. The key observation here is that maximizing capacity based on a random access MAC will require a high density of transmitters resulting in high spatial reuse but also a high likelihood of outage. Outage in turn may be associated with poor energy efficiency, causing retransmissions and thus increased loads, or congestion collapse, and further compromise packet delay or service quality.

Maximizing capacity subject to an outage constraint.

If energy and/or system performance are a concern it is reasonable to limit outage, say not to exceed a small constant ϵ . This in turn places a limit on the intensity of contenders λ and intensity of successful transmissions λ_s . The following result found in [19] summarizes relationship of interest between an outage constraint ϵ and achievable intensity of successful transmissions:

THEOREM 1. For a small outage constraint ϵ , lower and upper bounds on the effective transmission capacity for DS-SS-CDMA when transmitters employing a fixed transmission power ρ , fixed transmission range d , and interference model (1) are given by :

$$\frac{1}{2} \frac{\epsilon(1-\epsilon)}{\pi d^2} \left(\frac{m}{\beta} \right)^{\frac{2}{\alpha}} \leq \lambda_s \leq \frac{\epsilon(1-\epsilon)}{\pi d^2} \left(\frac{m}{\beta} \right)^{\frac{2}{\alpha}} \quad (6)$$

In other words for small ϵ , say 10^{-2} , one obtains an intensity of successful transmissions which is roughly linear ϵ , i.e., a fairly low capacity which is truly only reasonable for very low traffic loads. To summarize, as is well known, random access MAC protocols tend to be suitable for low load situations. Here we stress that for MAC protocols based purely on spatial random access can only achieve a moderate spatial reuse but will incur a high outage probability and thus poor energy efficiency.

3.2 Simple contention resolution and hand shaking reduces data collisions but does not increase spatial reuse

The random access MAC model considered above is quite crude, in the sense that transmitter nodes send data without prior signaling and thus some transmissions are lost due to outages. More sophisticated protocols **introduce carrier sensing** and contention/signaling phases prior to data transmission [5][9][17]. Only the ‘winners’ of the contention/signaling process subsequently transmit. The goal of contention is to eliminate/defer certain contending transmitters, which will cause excessive interference and likely outage to other receivers if all transmit at the same time. The goal of signaling is to do hand-shaking between a transmitter and receiver to ensure that the intended receiver is indeed available. For example, typical signaling mechanisms use three way hand-shaking. A transmitter sends a request-to-send (RTS) message to its intended receiver. A receiver which successfully gets a RTS message replies back with a clear-to-send (CTS) message, which is hopefully in turn received by the transmitter and suppresses other neighboring transmitters with high probability. After a successful hand-shake, the data is transmitted. With such handshaking in place one can ensure receivers are available, and possibly dramatically reduce the likelihood of outage during actual data transmissions. Signaling messages can be quite

small relative to data packets and thus such mechanisms are worthwhile to reduce outages on data transmission and save on energy even if this is at the cost of failures in signaling during the contention/hand-shaking process. An analysis of the capacity in this case need only factor the density of successful RTS/CTS exchanges one can achieve, with the outage probability during data transmissions being negligible. Thus, assuming CTS signals are always successful, the spatial density of successful transmitters is roughly the same to the random channel access case, i.e., such signaling scheme may not help much to improve spatial reuse. More improvement on spatial reuse can be realized by proper back-off strategies which are effective under light loads. However, considering the much reduced interference range when using CDMA, see Fig. 2, simple variations of RTS/CTS or carrier sensing based narrow-band MAC protocols are too conservative, leading to a poor spatial reuse.

3.3 Idealized contention resolution

Intuitively MAC contention resolution schemes ‘remove’ transmissions with prohibited overlaps. For example, consider the realization of the contenders on the top left panel of Fig. 5. In this case two receivers have a prohibited overlap and only one transmission will be successfully scheduled by the RTS/CTS handshaking scheme discussed earlier – see the middle figure on the top of Fig. 5. Yet a sophisticated contention resolution mechanism could achieve a better spatial reuse. For example, an *idealized* contention resolution process in place might allow at least one of a set of prohibited overlaps to survive – i.e., remove dumbbells with prohibited overlaps one at a time until no such overlaps were left. A possible result with two successful transmissions of such an idealized scheme is shown on the top right panel of Fig. 5. A simulation of such an idealized scheme is shown at the bottom of Fig. 5: on the left is a realization of contenders while on the right is a subset of successful transmissions. Successful transmitters and receivers tend to be clustered. Note that this idealized scheme is much better in terms of spatial reuse than simply removing all transmitter-receiver pairs with prohibited overlaps, yet it would not be straightforward to implement.⁴

Our simulations systematically exhibited two key aspects of contention-based MAC that have perhaps not been fully appreciated to date. First, as shown in our earlier analysis to achieve a high density of successful transmissions, i.e., a dense packing of dumbbells, one needs to have a high density of contenders. As a result, a significant number of transmitters will need to defer due to contention resolution or see a high outage probability under random access MAC protocol—roughly 50%. The second observation, is that successful transmitters and receivers are *clustered*, in particular when efficient spatial reuse is achieved – see right panel in Fig. 5. This is a unique property of spread spectrum-based ad hoc networks where receivers are capable of interference averaging and thus can tolerate certain level of neighboring interference. This suggests that by explicitly inducing spatial clustering in contention mechanisms, one might further improve spatial reuse. We consider this next.

⁴Implementing such contention resolution in a distributed system requires substantial signaling. Note that signaling also involves contention. Thus achieving successful signaling in turn causes significant overheads per transmission.

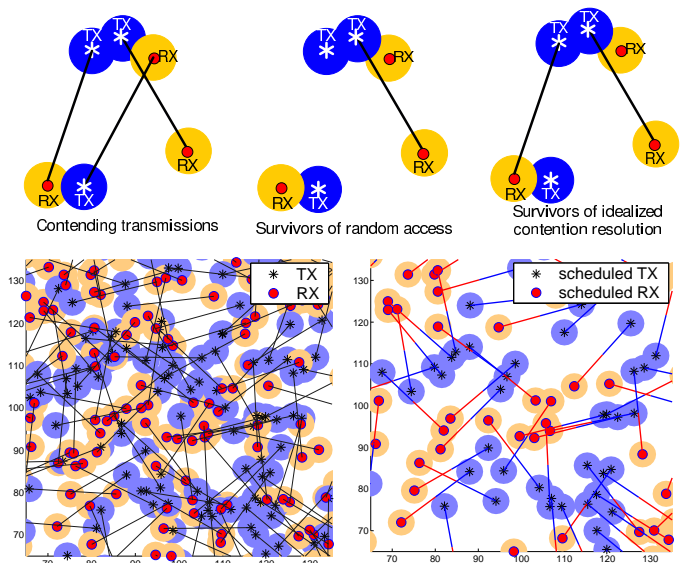


Figure 5: On the top left panel, an initial contending pattern of 3 transmissions. On the top middle panel, only one successful transmissions under random access protocol. On the top right panel, two successful transmissions under idealized contention resolution scheme. On the bottom left panel a realization of contending transmissions in our simulation, and on the bottom right transmissions surviving an *idealized* contention resolution of prohibited overlaps.

4. MAC PROTOCOLS BASED ON INDUCING SPATIAL CLUSTERING

One can consider inducing spatial clustering of transmitters in many ways. To intuitively show the benefit of inducing clustering we first consider an idealized deterministic placement. Then we propose two approaches. The first assumes that nodes contend synchronously and are aware of their locations. These capabilities are used to directly generate a spatial clustering of transmitters. The second assumes nodes are able to monitor interference levels to roughly infer relative locations of nodes and use signaling stages to achieve clustering of transmissions. We use these two representative distributed mechanisms to exhibit the benefits of *inducing spatial clustering* in a practical system.

4.1 Idealized deterministic clustering.

Idealized deterministic clustering. We begin by considering an idealized deterministic clustering of transmitters and receivers. The pioneering work of [10] showed that the optimal spatial reuse can be achieved by placing transmitters/receivers on a regular grid. We shall extend their result to a spread spectrum ad hoc network where optimal spatial reuse is achieved by clustering.

Assume we are free to select *both* the locations and states, e.g., transmitter or receiver, of nodes. Specifically, as shown on the left in Fig. 6, we assume tight clusters of transmitters and receivers are placed at the centers of cells in a regular square grid of size d^2 according to a checkerboard pattern. Each transmitter is assumed to transmit to a distinct receiver in one of the four neighboring cells and so has a fixed transmission range d . The number of nodes n within each

cluster will be determined to ensure all transmissions are successful. Let π_n^d denote a set of locations in \mathbb{R}^2 , corresponding to n -clustered transmitters in this checkerboard configuration, and where the origin corresponds to the center of a receive cluster cell. As seen earlier, in order for a transmitter at X_0 to successfully send to a receiver a distance d located at origin one must have that

$$\frac{\rho d^{-\alpha}}{\sum_{X_i \in \pi_n^d \setminus \{X_0\}} \rho |X_i|^{-\alpha}} \geq \frac{\beta}{m}$$

Fact 4.1 states this constraint in terms of a maximum allowable number of transmit nodes per cluster.

FACT 4.1. *Under the clustered grid placement of transmit nodes π_n^d , and $\alpha > 2$, a maximum number of nodes per cluster of*

$$\lfloor \frac{m/\beta + 1}{k(\alpha)} \rfloor, \text{ where } k(\alpha) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{4}{((2i)^2 + (2j+1)^2)^\alpha}$$

can be placed while ensuring no outage. This gives a density of successful transmissions of $\lambda_s = \frac{1}{2d^2} \lfloor \frac{m/\beta + 1}{k(\alpha)} \rfloor$.

This fact is shown through a brute force calculation of the aggregate interference offered by transmitter clusters in π_n^d at various ranges from the origin, which equals $nk(\alpha)d^{-\alpha}\rho$. Since $k(\alpha) \approx 4$ this suggests we need only consider the interference due to the $4n$ nodes which are a distance d from the origin. Thus in the Section 4.2 we will also focus on interference from the nearest clusters.

Comparing the best achievable spatial density of successful transmissions for randomly distributed MAC versus our ideally clustered grid, i.e., (5) to Fact 4.1, we have an approximate gain factor of $\frac{\pi}{2} \left(\frac{m}{\beta}\right)^{1-\frac{2}{\alpha}}$, e.g., when $\alpha = 4$, $m = 512$ and $\beta = 10\text{dB}$ a 10-fold increase in capacity.

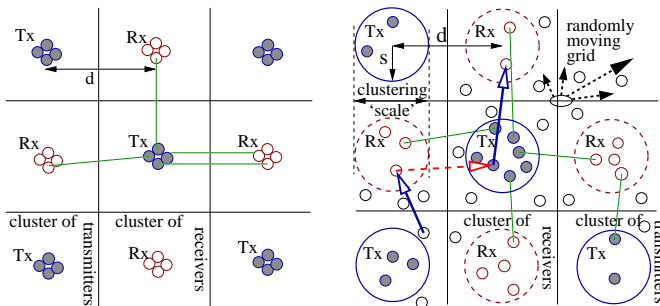


Figure 6: On the left an idealized deterministic placement. On the right a clustering of randomly located nodes in a random virtual grid with an example of typical routing patterns marked with hollow-headed arrows.

4.2 Inducing clustering based on a virtual grid.

In practice, one can not choose the placement of transmitters and receivers, i.e., it is unrealistic to expect that one can arrange for such a checkerboard clustered pattern. Yet for a homogeneous distribution of transmitters wishing to send a distance of roughly d one can approximate this pattern. Suppose, for example, that nodes are location aware and can determine their location relative to a known *virtual grid* of span d whose location evolves in a ‘random’ but known

manner with time. This can be achieved with GPS-capable synchronized nodes, that share a randomization seed driving the evolution of the grid – [14][13][2] have used this shared seed idea to allow nodes to infer other nodes’ states. Given this information and an a priori convention, a node can determine if it lies within a current transmitter/receiver cluster. As shown on the right of Fig. 6, we assume for now that nodes within a transmitter cluster transmit/relay to receivers in a neighboring receiver cluster.

Furthermore we let the parameter s determine the spatial scale of clustering and thus proximity of clustered nodes. Note that nodes that do not fall in either a transmitter or a receiver cluster region can defer, e.g., enter the sleep mode, unless they are sources or destinations.⁵ ‘Random moving’ of the grid happens in relatively large time scale and is mainly for balancing long-term energy consumption. If s is too small, each cluster will contain but a few transmitters and we may under-utilize the available capacity. If it is too large, there may be too many transmitters and/or interference variability (due to increased proximity), resulting in outages at receivers. So one may consider what is a good choice for s depending on the intensity λ of the Poisson point process Π of active nodes.

Let us first evaluate the outage probability of a receiver at the center of a receiver cluster and use this as an approximate estimate for the outage probability of a typical receiver. As for the deterministic placement, we will focus on the nearest 4 transmit clusters as the source of interference. Using Campbell’s theorem [18], we can evaluate the mean and variance of the interference as follows.

FACT 4.2. *Let Y denote the aggregate interference power level from the four transmit clusters closest to origin, i.e.,*

$$Y = \sum_{X_i \in \Pi} \mathbf{1}(X_i \cap A(s) \neq \emptyset) \times \rho |X_i|^{-\alpha}$$

where $A(s) = \bigcup_{i=-1,1} \bigcup_{j=-1,1} \mathbf{B}(i \times d, j \times d, s)$ is the union of these transmit cluster discs of radius s which are closest to the origin. Then

$$E[Y] = \lambda \int_{A(s)} \rho |x|^{-\alpha} dx, \quad \text{Var}(Y) = \lambda \int_{A(s)} \rho^2 |x|^{-2\alpha} dx.$$

□

Assuming Y is approximately Gaussian the outage probability for a receiver located at the origin is given by

$$p_o(\lambda, d, s) = \mathbb{P}\left(\frac{\rho d^{-\alpha}}{Y} < \frac{\beta}{m}\right) \approx \frac{1}{2} - \Phi\left(\frac{\frac{m}{\beta} - E[Y]}{\sqrt{\text{Var}(Y)}}\right). \quad (7)$$

Suppose in fact each transmitter finds a distinct receiver and thus there are no collisions due to concurrent transmissions to a receiver. Then for a fixed λ and d one can consider optimizing the cluster scale s so as to maximize the mean number of successful transmitters per cluster. Let n^* denote maximal mean number of successful transmissions per cluster, i.e., given by

$$n^* = \max_s \{\lambda \pi s^2 (1 - p_o(\lambda, d, s)) \mid s > 0\}. \quad (8)$$

⁵This requires a routing protocol to give special considerations to the first and last hops, e.g., the first hop of the typical route on the right of Fig. 6, which should also require a proper power control.

One may consider two regimes corresponding to different node density.

Regime 1: High node density. In this regime, the clustering scale $s \ll d$, i.e., the chosen nodes are clustered closely around the center of each cell. This is akin to the deterministic placement considered earlier. However each cluster will have a random, Poisson distributed, number of nodes with mean $\lambda\pi s^2$. In this regime Fact 4.2 gives

$$E[Y] = 4\lambda\pi s^2 \rho d^{-\alpha} \quad \text{and} \quad \text{Var}(Y) = \sigma_Y^2 = 4\lambda\pi s^2 \rho^2 d^{-2\alpha}.$$

We can in turn estimate the outage probability using (7). Note this optimization problem depends only on $z = \lambda\pi s^2$, the mean number of contending nodes per cluster.

As shown on the left panel of Fig. 7, in regime 1, the capacity achieved is close to the case of idealized deterministic placement. As shown on the right panel in Fig. 7, in this high density regime, we again obtain a transmission capacity that grows roughly linearly in $\frac{m}{\beta}$ as in the case of idealized deterministic placement, leading to a significant improvement over random access/ALOHA protocols, on the order of $\left(\frac{m}{\beta}\right)^{1-\frac{2}{\alpha}}$ or around 700% when $\frac{m}{\beta} = 50$, $\alpha = 4$.

Regime 2: Low node density. In general if the intensity of active nodes is not extremely high, the optimal choice for s will become comparable to d , i.e., one needs to increase the cluster scale so that a sufficient number of nodes can be scheduled. In this case the statistics of the interference, given in Fact 4.2, are affected by both the variability of the number of nodes per cluster, and their spatial variability within the cluster and relative to the origin. As shown on the left of Fig. 7 in Regime 2, if the spatial intensity λ is too small and there is not a sufficient number of nodes inside each cluster of size s , this negatively impacts the capacity. As λ grows larger, the capacity improves but the improvement eventually is marginal. Even in this regime, the achieved capacity is still significantly larger than random access/ALOHA. As shown on the right of Fig. 7, the capacity improves in $\frac{m}{\beta}$ sub-linearly but closer to linearly when λd^2 is larger.

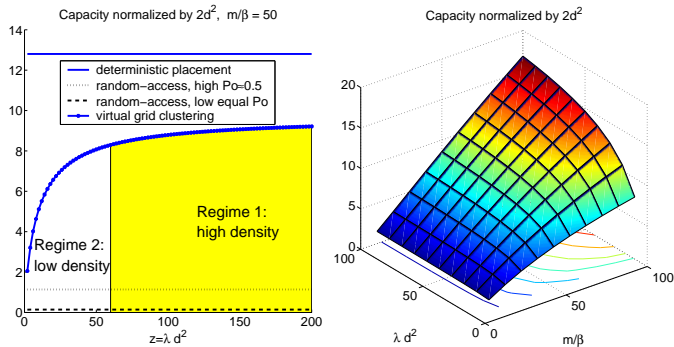


Figure 7: On the left, a comparison of capacities of different schemes. Here ‘random access with low equal p_o ’ is obtained by setting the outage constraint ϵ in (6) equal to the outage probability in the virtual-grid clustering protocol. On the right, the max capacity improves in both λd^2 and $\frac{m}{\beta}$.

Summary - virtual grid mechanisms. Note that (not shown in Fig. 7) for both high density and low density regimes, the outage probability is significantly lower at the

operating point achieving the highest transmission capacity, e.g., only about 5% for $\frac{m}{\beta} = 30$ compared with the outage probability of 50% in random access/ALOHA-like protocols, and this improves in m . When $\alpha = 2$, the capacity gain becomes marginal and independent of m . However, the benefit of low outage remains significant and improves with m .

An ad hoc network may have a non homogenous spatial density of nodes or traffic. Thus it is desirable to let the clustering scale adapt to such inhomogeneities. A straightforward approach would be to modify the our virtual grid mechanism such that each cell i has different cluster scale s_i . In this case the choice of s_i might be selected based on our analysis, i.e., $s_i^2 \approx \frac{1}{\pi\lambda_i} c(\frac{m}{\beta})$, where λ_i is the local node intensity in virtual cell i .

The virtual grid approach achieves a good tradeoff between capacity and energy efficiency. For example, nodes that are not covered by the transmit or receive cluster areas, can put themselves to sleep, until the grid moves. Furthermore, the overheads are low because nodes can infer locality of traffic and thus contention or signaling are not required, except that collisions where two or more transmitters send to the same receiver must be avoided. Collisions should be unlikely however, since the corresponding routing protocol should take advantage of long relay distance to achieve load balancing. The achievable capacity is close to that achieved by an ideal deterministic placement. Note that relatively low spatial intensity compared with spatial scale d may negatively impact the overall capacity because it prevents this mechanism from effectively inducing clustering.

4.3 Inducing clustering via multistage contention.

As seen in Section 3.2, protocols for ad hoc networks based on contention resolution and handshaking induce weak clustering among scheduled nodes – see Fig. 5. Since this is particularly beneficial to enhance the capacity of spread-spectrum based ad hoc networks, one might ask how these mechanisms might be optimized. Below we consider a representative approach of this type, starting with the case where nodes use a common transmission distance/power and then connecting to the case where they use heterogeneous transmission distance/power.

4.3.1 A multistage contention scheme.

Let us consider inducing clustering through a modified synchronous multi-stage RTS/CTS mechanism. Consider a two-stage example with the timing diagram shown in Fig. 8.

Stage 1 handshaking: In Stage 1 a subset of transmitters perform the three-way handshaking with their intended receivers, i.e., RTS, CTS, followed by an additional ‘confirmation’ RTS message. Only transmitter-receiver pairs who successfully exchange the three messages survive the first stage. These survivor pairs serve as ‘seeds’ for clusters in the subsequent handshaking stage(s).

Stage 1 monitoring: During Stage 1 contention, potential transmitters and receivers⁶ not participating in the first stage handshaking process synchronously monitor interference levels, for which they can indeed distinguish RTS and CTS time slots. Doing so permits them to evaluate their proximity to surviving Stage 1 transmitters and receivers.

⁶Those who will not be active at this cycle do not need to monitor, which is more efficient than [11] in which all nodes have to do consistent monitoring.

Stage 2 handshaking: In Stage 2, transmitters that sensed a ‘strong’ (see below) CTS signal in Stage 1 do not participate in Stage 2, i.e., are suppressed since they would likely interfere with the a successful Stage 1 receiver. Similarly a Stage 2 receiver which successfully receives an RTS from a transmitter, will only send back a CTS, if during Stage 1 it did no sense a ‘strong’ confirmation RTS signal. Thus the role of the Stage 1 ‘confirmation’ RTS is to signal receivers in the Stage 2 that they will be interfered with and thus to suppress their CTS.

This process can be carried out through multiple stages for a higher level of spatial reuse and might be performed in different ways, e.g., survivors of Stage 1, might also concurrently participate in Stage 2, to permit estimation of aggregate interference, rather than simply local interactions.

As exhibited by the simulation in Fig. 9, nodes that survive Stage 1 only achieve ‘weak’ clustering while Stage 2 survivors are dense and clustered. One may naturally ask:

- How does Stage 1 (simple contention) realize initial weak clustering and interact with subsequent stage(s) to realize these gains?
- What is the capacity gain provided by multistage contention process and how should it be optimized?

In the sequel, we will use our dumbbell model to analyze the multistage contention process and answer these questions.

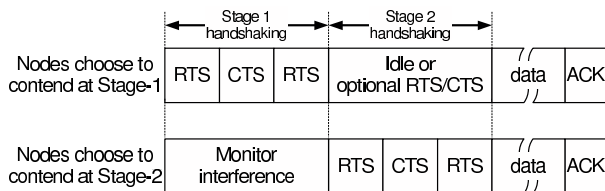


Figure 8: Timing diagrams of a two-stage contention MAC with the top for Stage 1 transmitter/receiver and the bottom for Stage 2 transmitter/receiver.

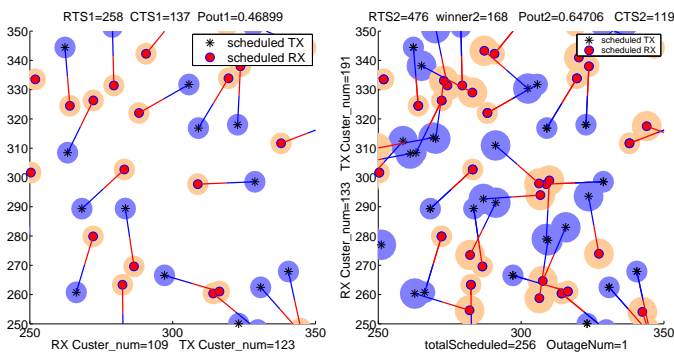


Figure 9: On the left, contention result of successful transmitter-receiver pairs, which serve as cluster ‘seeds’ for Stage 2. On the right the contention result after Stage 2, in which transmitters/receivers are indeed closely clustered with stage-1 transmitters/receivers and this significantly increases the overall clustering level.

First, we need to formally define thresholds which are used to decide when signals should be deemed strong enough to result in suppression. The critical range analysis, see

(2), suggests that a single interferer will cause outage for a transmitter-receiver pair using transmit power level ρ over a transmission range d , if the interference, as seen at the receiver exceeds $\frac{\rho d^{-\alpha} m}{\beta}$. To tolerate measurement uncertainty in the interference, we introduce a backoff factor c , where $0 < c \leq 1$ and thus a signal will be deemed strong if it exceeds $c \times \frac{\rho d^{-\alpha} m}{\beta}$. Note that c should be close to 1 otherwise we may be too conservative in utilizing available capacity. For purposes of visualizing this with ‘dumbbells’ and analytically studying clustering phenomenon later, we calculate the clearance range r_c around transmitters and receivers to be $r_c = c^{-\alpha} \left(\frac{\beta}{m}\right)^{-\alpha} d$.

‘Weak’ clustering from Stage 1. To quantify the clustering effect after Stage 1, we will consider given a successful receiver at the origin, what is the intensity of other successful receivers around the origin after Stage 1. ‘Clustering’ means the intensity of successful receivers should be higher close to the origin. Let $\lambda^{(1)}$ and $\lambda^{(2)}$ denote the intensity of contending transmitters in Stage 1 and 2 respectively. Consider a receiver that succeeds Stage 1 and suppose it is located at the origin O . Since it was successful during Stage 1 it must have cleared a disc of radius r_c of transmitters around it. Now conditioning on this receiver contending transmitters *outside* the disc are still homogenously distributed with intensity $\lambda^{(1)}$. Specifically let us evaluate the intensity of successful receivers within the ball $B(O, r_c)$ centered at the origin, Fact 4.3 summarizes the results in this regard.

FACT 4.3. *During Stage 1 contention of the multi-stage mechanism described above, successful receivers tend to cluster. Conditioning on a successful receiver at the origin O , we have the following results.*

1. *Let $p_s(x)$ denote the probability that a receiver a distance x from the origin, where $0 < x < r_c$, is suppressed by a Stage 1 transmitter, then*

$$p_s(x) = 1 - e^{-\lambda^{(1)} a(x)},$$

where

$$a(x) = x^2 \left[\omega \sin \theta + \omega^2 (\pi - \theta) + (\omega^2 + 1 - 2\omega \cos \theta) \times \left(\pi - \arccos \frac{1 - \omega \cos \theta}{\sqrt{\omega^2 + 1 - 2\omega \cos \theta}} \right) \right] - \pi r_c^2,$$

with $\omega = r_c/x$ and $\cos \theta = \frac{x}{2r_c}$ is the area of $B(x, r_c) \setminus B(o, r_c)$ as shown on the left of Fig. 10.

2. *The intensity of other successful receivers $\lambda_s^{(1)}(x)$ within the disc $B(O, r_c)$ at a distance x from the origin is roughly given (upper-bounded) by $\lambda_s^{(1)}(x) = \lambda^{(1)}(1 - p_s(x))$. The clustering of successful receivers is because the success probability of a receiver $1 - p_s(x)$ decreases sharply in its distance to the origin, shown on the right of Fig. 10.*

PROOF. The argument given in Item 1 is similar to that used to compute the outage lower bounds. Consider a receiver located a distance x from the origin. The receiver will survive Stage 1, if it has no transmitters within a ball of radius r_c of itself. As shown in Fig. 10 part of this ball has already been cleared of transmitters, since a successful Stage 1 receiver lies at the origin. Thus our candidate receiver will be suppressed if there are no transmitters within the region $B((x, 0), r_c) \setminus B(O, r_c)$ whose area is given by $a(x)$, with the

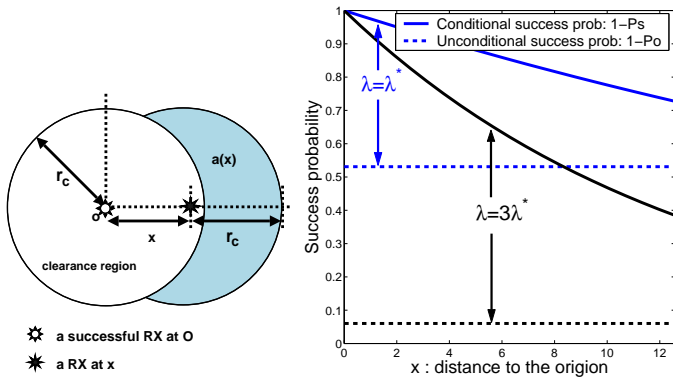


Figure 10: On the left the area $a(x)$ for obtaining outage lower bound conditioning on a successful receiver at O . On the right given the intensity λ of contending transmitters, the upper bounds for $1 - p_s(x)$, the success probability of a receiver at distance x conditioning on a successful receiver at O , and the success probability $1 - p_o(\lambda, d)$ without conditioning. For $p_o(\lambda, d)$ and λ^* see (1)(5).

probability of this occurring given by 1 minus the probability that a homogenous Poisson point process places no points in a region of area $a(x)$. Item 2 then follows that the intensity of Stage 1 receivers within $B(O, r_c)$ which are sent RTS's by transmitters outside this disc is also homogenous and has intensity $\lambda^{(1)}$ as long as $r_c < \frac{d}{2}$. \square

Fact 4.3 shows that *given* a successful receiver at the origin, the probability $p_s(x)$ that another Stage 1 receiver at distance x away from it is successful decreases quickly with distance. Furthermore, the *conditional* intensity of other *successful* receivers within $B(O, r_c)$ forms a non-homogenous Poisson process with intensity $\lambda_s^{(1)}(x)$ – the graph on the right in Fig. 10 exhibits this decay in intensity as a function of x . This decay is more significant when intensity is larger. Note that the above analysis was carried out by conditioning on a random event - ‘a successful receiver at O ’, with probability $1 - p_o$ decreasing in $\lambda^{(1)}$, and thus the clustering in Stage 1 does not really increase significantly in $\lambda^{(1)}$.

Optimizing the additional clustering realized by Stage 2. An intuitive explanation for the substantial clustering after Stage 2 is that survivors of Stage 1 serve as seeds to Stage 2 by creating areas around transmitters and receivers where they suppress receivers and transmitters respectively, in order to avoid close-by interference and further enhance clustering in subsequent stages. In fact these results point to the robustness of the proposed two stage signaling. Because the first stage ‘seeds’ the Stage 2 clusters, the second stage is able to handle fairly high density of contenders achieving a high success rate. Overall, unlike Stage 1, the system capacity is not as sensitive to the intensity of Stage 2 contenders. Furthermore only transmitters which know they will not severely interfere a Stage 1 receiver would in fact attempt Stage 2 signaling.

Fact 4.3 can be used to approximately study the additional clustering induced by Stage 2. After Stage 1 transmitter-receiver pairs are already scheduled in space, and will suppress nearby Stage 2 receiver and transmitters respectively. Thus, the process of Stage 2 transmitters which contend depends on the process of successful Stage 1 receivers and

thus is no longer Poisson. Nevertheless successful Stage 2 receivers will most likely cluster around Stage 1 receivers. Considering such a receiver at the origin and *assuming* that the Stage 2 transmitters which actually contend correspond roughly to a homogenous Poisson process with intensity $\lambda^{(2)}$ outside $B(O, r_c)$, we can reuse the results in Fact 4.3. Specifically the expected number of successful Stage 2 receivers clustered within the ball $B(O, r_c)$ of a Stage 1 receiver is approximately

$$\int_0^{r_c} \lambda^{(2)} e^{-\lambda^{(2)} a(x)} 2\pi x dx. \quad (9)$$

To maximize (9), i.e., maximize spatial reuse, we solve for the optimal $\lambda^{(2)} \approx \frac{1.75}{2r_c^2 \alpha}$, which is about $1.75\pi c^2$ times of the optimal contending intensity for Stage 1, see (5). Subsequently, the mean number of successful Stage 2 receivers around a successful Stage 1 receiver is roughly 0.93, i.e., we get roughly a 93% improvement on the capacity at the second stage and the expected number of successful receivers per cluster to be roughly 1.93. Our simulation results in Fig. 9 match this analysis quite well.

4.3.2 Handling multi-class or non-homogenous traffic and node distributions.

A realistic network may support transmissions with different relay distances for at least two reasons. First, spatial intensity of nodes may be heterogenous and nodes may need to use different distances to maintain connectivity. Second, applications sharing the network may have different QoS requirements and possibly require different relay strategies, e.g., relaying on different spatial scales, see the discussion of ‘spatial multiplexing’ and Fig. 1 in Section 1. In this case nodes should use power control to choose transmit power levels corresponding to the desired relay distances. Inducing spatial clustering to achieve high spatial reuse faces additional difficulties in this context because monitoring nodes can not correctly infer what are the interference regions of contending nodes in the previous stage(s) given a mixture of heterogeneous transmissions with different power levels.

One approach to deal with this problem is via a multi-class and multi-stage contention process. The basic idea is to allow transmissions with higher transmission power to perform handshaking first so as to enable transmitters and receivers in subsequent stages to detect their RTS/CTS and correctly estimate interference regions. Specifically, consider a network where nodes use one of k possible relay distances $d_i, i = 1, \dots, k$ satisfying $d_1 > d_2 > \dots > d_k$ each with an associated transmit power level $\rho_i^t, i = 1 \dots k$. Suppose these power levels are known, and the associated ranges are such that typically the receive power are the same, e.g., $\rho_i^t = \rho d_i^\alpha$. Note that this is unrealistic and we will discuss imperfect power control vs. relay distance later in Section 5. In the sequel we refer to nodes which choose relay distance d_i to be of class i . Our new contention protocol is a variation on the multi-stage RTS/CTS/RTS process considered earlier. In particular we assume that only class i nodes perform handshaking at Stage i based on monitoring interference levels for stages $1, \dots, i - 1$ and thus inferring whether they will interfere with or be interfered by contenders in previous stages, by taking into account predefined power levels used at each stage. We do not cover implementation details here due to space constraint.

The intuition for this choice is that by allowing only a

given class to contend at a particular stage, nodes monitoring the process can obtain reasonable estimates of the proximity of contenders based on a priori knowledge of their transmitting power levels and the received interference levels. Furthermore the ordering in which classes contend (based on transmission ranges) is enforced because if $i < j$ then the packing achieved by class i is likely to be less dense than that of class j . This ensures that large range transmissions are effectively packed prior to committing to short range ones. This ordering also ensures a contender will hear the signaling of all relevant contenders with higher transmission power in previous stages. This approach, achieves a multi-scale clustering and high spatial reuse of successful transmissions. Fig. 11, exhibits a realization of a two-class scenario. Transmissions with long relay distances and larger interference regions (dumbbell size) are scheduled in Stage 1, while transmissions with shorter relay distance are scheduled in Stage 2. As can be seen, in addition to clusters of receivers for both classes of transmitters, additional fine scale clustering of short range transmission fill the void area remaining after Stage 1.

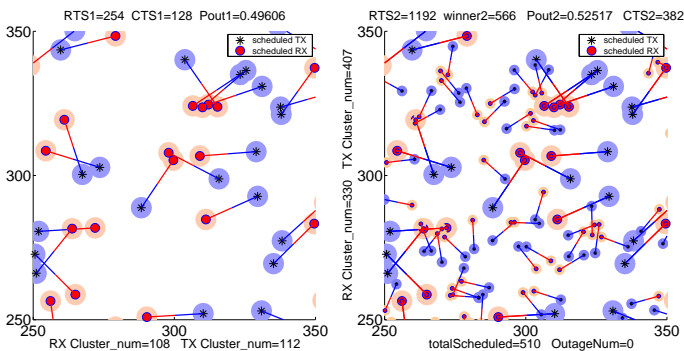


Figure 11: On the left, the resulting transmitter-receiver pairs of a multi-stage multi-class contention protocol’s Stage 1 contention among nodes relaying delay sensitive traffic, i.e., using longer relay distances and thus having larger large interference ranges. On the right the resulting transmitter-receiver pairs for Stage 1 and 2 for a multi-class multi-stage contention protocol. Note how the shorter range transmissions cluster around Stage 1 receivers as well as independently in voids left during Stage 1.

Summary - multi-stage contention and clustering. First, when the contention parameters are properly configured, multistage contention can achieve close-to optimal capacity with 2 or 3 stages. See [20] for more performance evaluations. Second, the associated overhead for each successful transmission, is fairly close to the simple RTS/CTS mechanism with additional overhead to monitor local interference levels. Finally the previous clustering and passive monitoring of the contention process benefit the next round of contention reducing the sequential failure rate of contention significantly, e.g., by 65% relative to homogenous Poisson point process in our simulations. For example, assuming nodes switch between transmit and receive modes, successful transmitters in a given round tend to cluster, making themselves likely to be successful in the next round as receivers.

5. PRACTICAL DESIGN AND IMPLEMENTATION CONSIDERATIONS

The previous section suggests a mechanism for inducing spatial clustering on multiple scales has some significant benefits in terms of meeting QoS requirements and enabling power savings while achieving efficient spatial reuse. This section gives some thoughts on design and implementation considerations. Due to space constraints, we will only briefly introduce some of the problems.

Synchronization. Throughout this paper, we have assumed a slotted system with synchronous contention and transmission. We believe synchronization is critical for efficient MAC scheduling in ad hoc networks. First, data transmission and ACK are well protected after handshaking, which eliminates the need for maintaining states, e.g., NAVs in 802.11 and [11]. Second, synchronous contention provides better priority access and thus better QoS support than asynchronous implementations[17]. Although synchronization incurs extra overheads, such as inter-frame spacing, similar MAC inefficiency also exist for asynchronous contention resolution in which the required carrier sensing usually causes conservative back-off both spatially and temporally. Therefore benefits of synchronization will warrant these overheads.

Spreading factor. Our clustering approach requires a system with a large spreading factor, i.e., $\frac{\beta}{m} \ll 1$. The performance gains of clustering are more significant when the spreading factor m is large. Note that we conservatively assume correlation among quasi-orthogonal codes is $\frac{1}{m}$ while some in the literature use $\frac{1}{3m}$, which would allow clustering to work well even for systems with relatively moderate spreading.

Transmission power and range. To be effective, our multi-stage contention MAC requires pre-defined power levels. This is actually a realistic model since real devices typically only do discrete power control. Yet the transmission range can continuously change rather than fixed per class as assumed in this paper. This leads to a variable interference range r even at the same transmitting power level. Thus a back-off criterion should always conservatively assume monitored transmissions are of the maximal transmission range allowed by the power levels associated with each class. This causes a penalty in the capacity, but only marginal in particular when path loss is large.

Node spatial distribution. We have assumed homogeneous Poisson point process throughout this paper. What if the spatial density is not homogeneous? For the multi-stage contention, existing clustering patterns may indeed help the performance and this scheme is very robust for inducing clustering among contenders for different scenarios. By contrast, the virtual-grid scheme, works well in a ‘dense’ network, i.e., when there are many nodes within each virtual cell of size d^2 . In a ‘dense’ network, our virtual-grid scheme is robust to fluctuations of spatial density.

Routing protocol. The importance of routing has been mentioned in previous items. In particular, the challenge in the new design paradigm proposed in this paper is to enable long relay distance. Maintaining the information of all the nodes in a large neighborhood may not be feasible. Thus the actual design may try to discover preliminary routes consisting small hops and associated power budget. Then a source node may recompute a new route by skipping some

relay points on preliminary routes. At the same time, using spread spectrum results in a reduction in the bandwidth per channel, while enabling multiple quasi orthogonal channels. Thus a routing algorithm for such networks should carefully spatially balance traffic loads across available nodes (channels) to achieve efficient operation.

6. CONCLUSION

In this paper we have considered MAC design for ad hoc networks based on a spread spectrum physical layer. Our interests in considering such networks are multiple, and involve effective ways of addressing quality of service, energy efficiency, and efficient spatial multiplexing of bursty traffic, etc. We are specifically interested in the regime where the transmission range of a node exceeds nearest neighbor scales and thus is likely to include a number of other potential transmitters. In other words the transmission range may in fact exceed several nearest neighbor hops. In this case a properly designed MAC and spread spectrum physical layer will permit overlaying of concurrent transmissions without resulting in outage. The key observation in this paper is that for such a physical layer technology an efficient packing of transmitter-receiver pairs will be one exhibiting clustering of receive and transmit nodes. As such we propose a novel approach to MAC design whereby clustering is induced in the contention process with a view on enhancing capacity, and sketch two distributed approaches that achieve this end.

7. REFERENCES

- [1] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler. A spatial reuse aloha MAC protocol for multihop wireless mobile networks. In *Proc. of Annual ALLERTON Conference on Communication.*, 2003.
- [2] L. Bao and J. Garcia-Luna-Aceves. Transmission scheduling in ad hoc networks with directional antennas. In *Proc. of ACM MOBICOM*, September 2002.
- [3] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad-hoc networks. In *Proc. of IEEE INFOCOM*, 2002.
- [4] K. S. G. et. al. On the capacity of a cellular CDMA system. *IEEE Trans. on Veh. Technology*, 40(2):303–12, May 1991.
- [5] J. Garcia-Luna-Aceves and J. Raju. Distributed assignment of codes for multihop packet-radio networks. In *Proc. of IEEE MILCOM*, 1997.
- [6] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. *IEEE Transactions on Networking*, 10(4):477–486, 2002.
- [7] P. Gupta and P. Kumar. Critical power for asymptotic connectivity in wireless networks. In *Stochastic Analysis, Control, Optimization and Applications*, Birkhauser, Boston, pages 547 – 566, 1998.
- [8] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.
- [9] M. Joa-Ng and I. Lu. Spread spectrum medium access protocol with collision avoidance in mobile ad-hoc wireless network. In *Proc. of IEEE INFOCOM*, 1999.
- [10] L. Kleinrock and J. Silvester. Spatial reuse in multihop packet radio networks. *Proceedings of the IEEE*, 751(1):156–167, January 1987.
- [11] A. Muqattash and M. Krunz. CDMA-based MAC protocol for wireless ad hoc networks. In *Proc. of ACM MOBIHOC*, 2003.
- [12] M. Pursley. The role of spread spectrum in packet radio networks. *Proceedings of the IEEE*, 75(1):116–134, 1987.
- [13] R. Rozovsky and P. Kumar. Seedex: A MAC protocol for ad hoc networks. In *Proc. of ACM MOBIHOC*, October 2001.
- [14] T. Shepard. A channel access scheme for large dense packet radio networks. In *Proc. of ACM SIGCOMM*, pages 219–230, 1996.
- [15] E. Sousa and J. Silvester. Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network. *IEEE Journal on Selected Areas in Communications*, 8(5):762–771, June 1990.
- [16] M. Stemm and R. H. Katz. Measuring and reducing energy consumption of network interfaces in hand-held devices. *IEICE Trans. on Communications – Special Issue on Mobile Computing*, E80–B(8):1290–1302, 1997.
- [17] J. Stine and G. de Veciana. A paradigm for quality of service in wireless ad hoc networks using synchronous signaling and node states. *IEEE JSAC Special Issue on Quality of Service Delivery in Variable Topology Networks*, 22(7):1301–1321, September 2004.
- [18] D. Stoyan, W. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester, 1995.
- [19] S. Weber, X. Yang, G. de Veciana, and J. Andrews. Transmission capacity of CDMA ad hoc networks. In *Proc. of IEEE ISSSTA*, September 2004.
- [20] X. Yang and G. de Veciana. Designing MAC protocols for spread spectrum ad hoc networks: Thinning versus spatial state-dependent packing. In *Proc. of CISS*, March 2005.