A New Approach to the Directed Connectivity in Two-Dimensional Lattice Networks

Lei Zhang, *Member, IEEE,* Lin Cai, *Senior Member, IEEE,* Jianping Pan, *Senior Member, IEEE,* and Fei Tong, *Member, IEEE*

Abstract—The connectivity of ad hoc networks has been extensively studied in the literature. Most recently, researchers model ad hoc networks with two-dimensional lattices and apply percolation theory for connectivity study. On the lattice, given a message source and the bond probability to connect any two neighbor vertices, percolation theory tries to determine the critical bond probability above which a giant connected component appears. This paper studies a related but different problem, directed connectivity: what is the exact probability of the connection from the source to any vertex following certain directions? The existing studies in math and physics only provide approximation or numerical results. In this paper, by proposing a recursive decomposition approach, we can obtain a closed-form polynomial expression of the directed connectivity of square lattice networks as a function of the bond probability. Based on the exact expression, we have explored the impacts of the bond probability and lattice size and ratio on the lattice connectivity, and determined the complexity of our algorithm. Further, we have studied a realistic ad hoc network scenario, i.e., an urban VANET, where we show the capability of our approach on both homogeneous and heterogeneous lattices and how related applications can benefit from our results.

Index Terms—Connectivity, square lattices, bond percolation, ad hoc networks

1 Introduction

In ad hoc networks, nodes can be distributed in the space without any preexisting communication infrastructures. To deliver a message to its destination, short-range contacts and multi-hop wireless communications are needed. Such scenarios are termed as wireless, mobile, or vehicular ad hoc networks (W/M/VANETs) or delay/disruption-tolerant networks (DTNs). In general, the connectivity in ad hoc networks can be defined as the connection probability of any two nodes via multiple intermediate connections.

The connectivity problem in two-dimensional (2D) ad hoc networks has attracted lots of attention, most recently with geometrical probability, stochastic geometry, and percolation theories [1]–[3]. When the network can be adequately modeled as a 2D square[1] lattice (e.g., VANET in city blocks), percolation theory has been widely used. Initially in statistical physics, percolation theory studies the process of liquid filtering through porous materials. The *bond probability* indicates the connection probability of any two neighbor vertices. If the percolation directions are given, i.e., the connection between any two neighbor vertices is directed, such type of percolation is called directed percolation (DP).

In this paper, we study a related but different problem of DP: directed connectivity (DC), i.e., given a message source and the bond probability to connect neighbor vertices on a square lattice, what is the probability for the message to reach an arbitrary vertex following certain directions? The problem appears to be similar to DP at the macroscopic level, but different microscopically as DP only cares about the existence of a giant component, while DC has to determine the exact connectivity to each vertex, which is more relevant to network connectivity.

Despite the efforts in more than half a century, DP and DC are mostly studied on their convergence and critical behaviors with approximation results or numerically by simulation, from high dimensions [4] to two dimensions [5]–[7]. For DP, one of the most related work determined the critical probability analytically of a square lattice where the vertical bond probability is p_y and the horizontal probabilities are 1 and p_x interleaved at different layers [5]. Conceptually, DC problems are even harder than DP. For DC, authors in [6] discovered the relations between the connectivity and the percolation distance, the connectivity and percolation angle, respectively, but only with approximation. However, by extending our previous work on 2D ladder connectivity [8] and by proposing a new recursive decomposition approach, we have obtained the exact analytical expressions for the DC problem on square lattices. The approach can be extended to lattices with different horizontal and vertical bond probabilities, and different tessellation shapes.

This paper makes the following contributions. First, for the first time in literature to the best of our knowl-

An early preliminary version of this paper appears at the 32nd International Conference on Computer Communications (IEEE INFOCOM'13) in Turin, Italy.

- *L. Zhang, J. Pan and F. Tong are with the Department of Computer Science, University of Victoria, Victoria, BC, Canada, V8P 5C2. E-mail: {leiz, pan, tongfei}@uvic.ca.*
- *L. Cai is with the Department of Electrical and Engineering, University of Victoria, Victoria, BC, Canada, V8P 5C2. E-mail: cai@ece.uvic.ca.*

1. Following percolation notations, square lattice is a lattice of square cells. We use symmetric lattice to refer to the same number of horizontal and vertical cells in a 2D finite lattice, and strip lattice when the number of cells in one dimension is fixed but that in the other varies.

edge, it gives the exact analytical solution to the DC problem on square lattices and can quickly determine the network connectivity without lengthy simulation. Even though the results are based on square lattices, they can offer valuable insights when clustering and aggregation are possible in full 2D networks. Second, we explore the obtained analytical expressions and analyze the impact of the bond probability, and the lattice size and ratio on network connectivity, in addition to determining the complexity of the proposed approach. Third, we apply the approach to an urban VANET scenario to show the extensibility of the approach. Inspired by some existing work [8], [9], we carefully mapped the urban VANET message propagation to the DC problem. Both homogeneous and heterogeneous cases are discussed and valuable insights are discovered about how the applications can benefit from the results.

The remainder of this paper is organized as follows. In Section 2, we review the most related work to the connectivity problem in ad hoc networks and its relation to the directed connectivity. We present our analysis framework and derivation in Section 3. Further, we determine the algorithm complexity, verify the algorithm correctness and explore the analytical connectivity expressions in Section 4. Section 5 offers an extension of our approach to a real-world problem, i.e., directed connectivity in urban VANETs. Further discussion and conclusions are presented in Section 6 and 7, respectively.

2 Related Work

Connectivity of ad hoc networks has been extensively studied, mostly in the 2D Euclidean spaces [10], [11]. Without fixed communication infrastructures, nodes in ad hoc networks have to rely on their neighbors or leverage the mobility of them to deliver messages to the destination, often in a multi-hop manner through wireless communications and/or short-range contacts. As a result, the connectivity has to be characterized probabilistically. A wide variety of ad hoc networks exist, ranging from stationary (sensor networks [12]) to mobile ones. VANET is a special type of the latter, where vehicles are involved as the communication source, destination and relay [8], [13], [14]. The unique vehicle mobility introduces more challenges to connectivity, but also offers new opportunities. For example, along a highway, vehicles travel in one dimension, possibly also communicating with the vehicles on the reverse direction. In a city block scenario, a 2D square lattice is often used to approximate the road grid.

Analytical and algorithmic tools in graph theory and computational geometry have been widely used in the modeling and analysis of connectivity in ad hoc networks [15], together with geometrical probability, stochastic geometry, and percolation theories in recent years. For example, a connected dominating set is introduced in ad hoc networks to create a virtual backbone for the network [16]. Geometrical probability tools offer the characterization of distance distributions among nodes in and between different geometry shapes (e.g., triangles, rectangles and hexagons [17]–[19]), and stochastic geometry tools further introduce the time line in the random process of node coverage and connectivity [20]. Additional nodes can be deployed, sometimes even mobile ones, to improve the connectivity.

Most recently, percolation theory has found a wide range of applications in networking research, particularly on the connectivity in ad hoc networks [21]–[24]. Many networking scenarios can be adequately modeled as percolation on square lattices, either individually (e.g., VANET in city blocks) or after clustering and aggregation (cluster heads in wireless sensor networks). Although many 2D regular tilings are used (e.g., hexagons for cellular systems and rhombuses or triangles for cells with directional antennas), square lattices are most widely used. For messages with a given destination, or vehicles traveling in certain directions, geographical forwarding is often deployed to minimize the network overhead due to flooding [8]. Thus directed percolation becomes a premier model in such scenarios, and most existing work applies the results from isotropic or directed percolation on square lattices.

Motivated by the VANET connectivity in city scenarios, this paper studies the DC problem on square lattices. We try to establish the analytical expression for the directed connectivity from a given message source to any possible destinations in the network. The most related work is presented in [6], where $P_{\mathbf{R},p}$ indicates the probability of a pair of sites at $\mathbf{0}$ and $\mathbf{R}$ being connected with directed bond probability p. There exists a value p_c such that for $p < p_c$, $P_{\mathbf{R},p} \sim exp(-|\mathbf{R}|/\xi)$ with $\xi \sim (p_c - p)^{-\nu}$; p_c and ν can be estimated numerically. This approximation indicates an exponential decay of $P_{\mathbf{R},p}$ when $p < p_c$ and the exponent is determined by how far away the bond probability is from the critical bond probability. For $p > p_c$, percolation happens and there is always one (or more) infinite path (i.e., giant component). But this does not mean the sites $\mathbf{0}$ and $\mathbf{R}$ are connected. The work further discussed the impact of percolation angle on $P_{\mathbf{R},p}$, showing that the peculation can only happen within a "corn" area along the diagonal of the two connection directions.

Different from the existing work focusing on the convergence behaviors of directed connectivity with approximation results obtained either by simulation or numerically, we explore the exact connectivity expression between any two sites as a function of the bond probability. In addition, we extend our work to more general DC problems with variable bond probabilities, which is more desirable than most of the existing work [5], [6] which just consider either identical bond probability or different probabilities only between horizontal and vertical bonds.

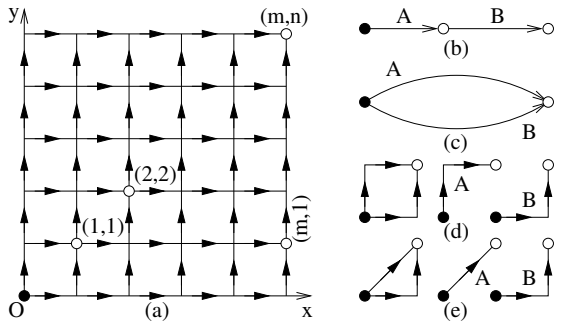


Fig. 1: System model and basic principles.

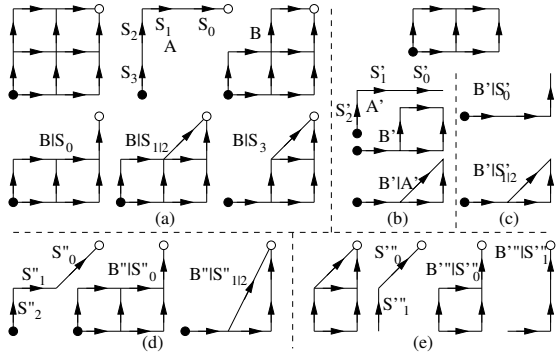


Fig. 2: The decomposition of a $2 * 2$ lattice.

3 Analysis Framework

In this section, we first give the system model for directed connectivity, and then present the analytical framework and derivation results for lattices with size varying from $1 * 1$ and $2 * 2$ to $m * n$.

3.1 System Model

As shown in Fig. 1(a), we consider a 2D lattice $L(m, n)$, with edges parallel to the x and y axes for notation convenience. A message is generated at the origin $O = (0, 0)$ at time $t = 0$, and propagated along the lattice edges in the directions indicated by arrows. Assuming the bond probability p of any two neighbor vertices, what we want to know is the connection probability from the origin to (m, n), as a function of p. The derivation of bond probability p varies in different applications. For instance, in Section 5, we consider p as the probability of any two adjacent road intersections in urban areas being connected and the detailed calculation could be referred to [8]. To be more general, bond probabilities in a lattice network could be different from edge to edge. But for the simplicity of presenting our approach, we assume a homogeneous p in this section. Heterogeneous case is presented as an extension in Section 5.

Even with the simplified model, it is still a hard problem to derive the connection probability at vertex (m, n), denoted as $P(m, n)$. For example, to reach (m, n), the message can go through $(m-1, n)$ or $(m, n-1)$ as the last hop. However, even if $P(m-1, n)$ and $P(m, n-1)$ were known, it is still difficult to derive $P(m, n)$, as the paths from $(0, 0)$ are not independent before they reach the last hop. A brute-force approach has to enumerate all possible paths and overlapping (i.e., when different paths share the same edges) and its complexity suffers the combinatorial explosion on the exponent. This is also the reason why the exact result of DC remains unsolved for so many years.

To facilitate the presentation, we also illustrate some basic principles and simple cases in Fig. 1. First, if there are two directed paths A and B connected by a common vertex *serially* as shown in Fig. 1(b), the end-to-end connectivity is $P(AB) = P(A)P(B)$, as A and B are always independent with directed edges. Here, we define $P(A)$ and $P(B)$ as the probabilities that path A and B are connected, respectively. Second, if there are two *parallel* paths A and B connecting the source and destination as shown in Fig. 1(c), the source-to-destination connectivity is $P(A + B) = P(A) + P(B) - P(AB)$ according to the principle of inclusion and exclusion (PIE). These two principles can be used to solve the $1 * 1$ lattice problem as shown in Fig. 1(d): $P(A) = P(B) = p \cdot p = p^2$, and $P(1, 1) = P(A) + P(B) - P(A)P(B) = 2p^2 - p^4$ as A and B are independent and not mutually exclusive. Later we find that we also encounter a triangular cell as shown in Fig. 1(e), and the end-to-end connectivity in this case is $P_T(1, 1) = p + p^2 - p \cdot p^2 = p + p^2 - p^3$. The cases become more complicated when A and B are also dependent.

3.2 $2 * 2$ Lattices

Following the same principles, we attempt to solve the $2 * 2$ lattice problem as shown in Fig. 2. Similarly, the top-leftmost path A and the union of all other paths, $\mathcal{B}$, are identified. But the difference from Section 3.1 is that they are no longer independent (as A and $\mathcal{B}$ have many overlapping edges). A naive approach is to consider each edge along A separately and check the impact of their connection status on $\mathcal{B}$. Depending on each edge being connected or not, there are 2^4 cases of a single path A and many more cases will be introduced in $\mathcal{B}$. If using the PIE principle, after the first level of decomposition, to further decompose $\mathcal{B}$, which has more than two layers, is very difficult, if not impossible. This is because most edges, other than the bottom-rightmost two, are shared by many paths. This also renders our previous approach [8] on one-layer ladders not applicable to multi-layer lattices. Observing A is a single path (i.e., no branches possible), we can have a simple partition on it. As shown in Fig. 2(a), define S_i as the event that the last i edges along A leading to the destination are all connected, but the last $(i + 1)$-th one is not, so $P(S_i) = p^i(1 - p)$ for $0 \leq i \leq m + n - 1$. For the origin and destination to be connected, we then have $m + n + 1$ mutually exclusive cases, including $\mathcal{B}|S_i$ and A being connected where $P(A) = p^{m+n}$. Define the probability that $\mathcal{B}$ is connected given S_i as $P(\mathcal{B}|S_i)$, we have

$$
\begin{aligned}
P(m,n) &= P(A+\mathcal{B}) = 1 - P(\overline{\mathcal{B}+A}) \\
&= 1 - P(\bar{\mathcal{B}}\bar{A}) \quad (1) \\
&= 1 - P(\bar{\mathcal{B}} \bigcup_{i=0}^{m+n-1} S_i) \quad (2) \\
&= 1 - \sum_{i=0}^{m+n-1} P(\bar{\mathcal{B}}|S_i)P(S_i) \quad (3) \\
&= 1 - \sum_{i=0}^{m+n-1} (1 - P(\mathcal{B}|S_i))P(S_i) \\
&= P(A) + \sum_{i=0}^{m+n-1} P(\mathcal{B}|S_i)P(S_i), \quad (4)
\end{aligned}
$$

where (1) is due to De Morgan's law, (2) due to $\bigcup_{i=0}^{m+n-1} S_i = \bar{A}$, (3) due to S_i being mutually exclusive, and (4) due to $\sum_{i=0}^{m+n-1} P(S_i) = P(\bar{A}) = 1 - P(A)$, i.e., A and S_i partition and constitute the entire event space in *total probability*.

For $L(2,2)$, given S_0, no end-to-end connection is possible via vertex $(0,2)$ or $(1,2)$, so we can discard the edges adjacent to the two vertices and have $\mathcal{B}|S_0$ as shown in Fig. 2(a). Given S_1, it implies that $(1,2)$ and $(2,2)$ are connected, and $\mathcal{B}$ does not include any edges from $(0,2)$, so we can merge $(1,2)$ with $(2,2)$ in $\mathcal{B}$ to obtain $\mathcal{B}|S_1$. Since S_1 and S_2 have the same effect, they are illustrated as $\mathcal{B}|S_{1|2}$ in Fig. 2(a). Given S_3, no connection is possible through $(0,1)$, so the edges adjacent to it have to be removed; it also implies that $(0,1)$, $(0,2)$, $(1,2)$ and $(2,2)$ are connected sequentially, so they can be merged, as $\mathcal{B}|S_3$ in Fig. 2(a).

After this decomposition, we have $\mathcal{B}|S_{0..3}$. Using the serial principle, $B|S_0$ can be decomposed into a $2*1$ lattice (or ladder) and an edge. Figure 2(b) shows how we further decompose the ladder into A', $\mathcal{B}'$ and $\mathcal{B}'|A'$ following the conditional probability approach that we previously proposed for ladders specifically [8], while Fig. 2(c) shows the new total probability approach with $\mathcal{B}'|S'_{0..2}$, which can both be solved directly using the serial principle, $P(1,1)$ and $P_T(1,1)$: the results are the same, but the new approach is simpler, especially when we have multi-layer lattices. Similarly for $\mathcal{B}|S_{1|2}$, they are decomposed in Fig. 2(d) to components of known connectivity (e.g., $\mathcal{B}''|S''_0$ is the same as $\mathcal{B}|S_0$), and part of $\mathcal{B}|S_3$ is decomposed in Fig. 2(e), where the serial principle and $P(1,1)$ can be applied. Using the total probability approach, the connectivity of the decomposed components can be reassembled,

$$
\begin{gathered}
P(A) = p^4,\ P(\mathcal{B}|S_0) = p^8 - p^7 - 2p^6 + 3p^4, \\
P(\mathcal{B}|S_{1|2}) = -p^9 + 3p^8 - 3p^6 - 3p^5 + 3p^4 + 2p^3, \\
P(\mathcal{B}|S_3) = p^7 - 2p^6 - p^5 + 2p^4 + p^3.
\end{gathered}
$$

Then $P(2,2)$ can be recovered as follows

$$
\begin{aligned}
P(2,2) &= P(A) + \sum_{i=0}^{3} P(\mathcal{B}|S_i)P(S_i) \\
&= p^{12} - 4p^{11} + 2p^{10} + 4p^9 + 2p^8 - 4p^7 - 6p^6 + 6p^4.
\end{aligned}
$$

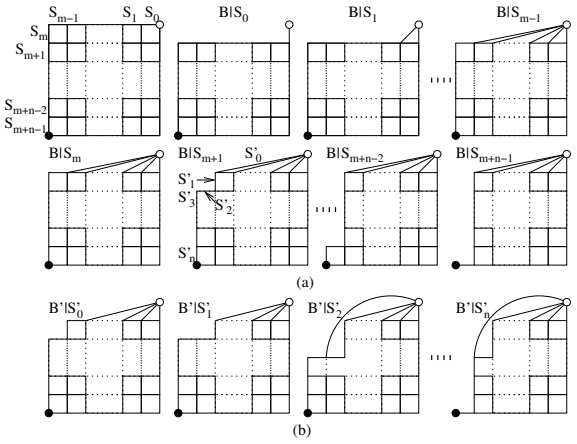


Fig. 3: The decomposition of an $m * n$ lattice.

3.3 $m * n$ Lattices

Following the new total probability approach, we attempt to solve the generic $m * n$ lattice problem as shown in Fig. 3(a). For clarity, we have omitted the arrow on edges in the following figures. Similar to $\mathcal{B}|S_i$ in Section 3.2, we can first remove the edges or merge the vertices on the top row of an $m * n$ lattice, as $\mathcal{B}|S_{0..m-1}$, and then remove the edges or merge the vertices along the leftmost column of the lattice, as $\mathcal{B}|S_{m..m+n-1}$, by considering the top-leftmost path A and events $S_{0..m+n-1}$, as well as their impacts on $\mathcal{B}$.

After this decomposition, similar to Section 3.2, we need to further decompose the components eventually to the ones of known connectivity. For example, $\mathcal{B}|S_0$ contains an $m*(n-1)$ lattice and an edge, which leads to a recursion among lattices. $\mathcal{B}|S_{m+1}$ is further decomposed in Fig. 3(b) and similarly for all the other $\mathcal{B}|S_i$s. Although all $\mathcal{B}|S_i$s and their decomposed components have different structures, we have found the similarities between these structures during the decomposition process, and we can introduce a generic structure called *Tower* to formulate the recursions among them.

Figure 4(a) shows the generic structure of Tower $\mathcal{T}$. All the decomposed components of the $m*n$ lattice, plus the lattice itself, can be regarded as a special case of $\mathcal{T}$. $\mathcal{T}$ has a layered structure, with more blocks near the bottom, as we remove edges and merge vertices gradually along the top-leftmost portion of the tower. The source is the bottom-leftmost vertex at $(0,0)$, and the destination is the top-rightmost one at (m,n). On each layer, there are two types of building blocks: *triangles* and *rectangles*. Each triangle $\triangle$, highlighted in blue on layer K in Fig. 4, is composed of two shortcuts to (m,n) and one or two ordinary lattice edges. A rectangle $\square$, highlighted in red, is the block originally in the lattice and not affected by the decomposition process yet. Depending on the number (t_i) of $\triangle$s and that (r_i) of $\square$s, we can denote layer i by (t_i, r_i), except for the base b which is represented by the number of the bottom edges. Taking into account all the layers in a configuration for the tower, we can

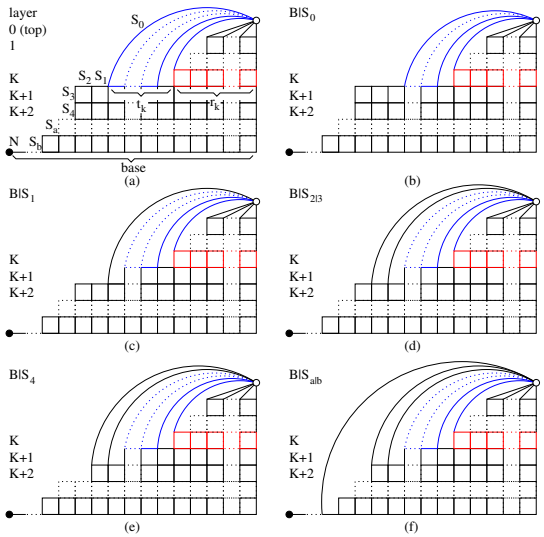


Fig. 4: The decomposition of a *Tower*.

denote it by $\mathcal{T}((t_0,0),\sqcup_{i=1}^{K}(t_i,r_i),\sqcup_{i=K+1}^{N}(0,r_i),b)$, where $\sqcup$ represents a series of layers. Be aware that we distinguish four types of layers: 1) the top layer 0 with $\triangle$s only; 2) the mixed ones of both $\triangle$s and $\square$s from layer 1 to K; 3) the ones with $\square$s only from layer $K+1$ to N; 4) the base b. For example, $\mathcal{T}(\sqcup_{i=0}^{n-1}(0,m),m)$ is the original $m*n$ lattice, $\mathcal{T}((1,0),\sqcup_{i=1}^{n-1}(0,m),m)$ is the $\mathcal{B}|S_1$ in Fig. 3, $\mathcal{T}((m-1,0),(0,m-1),\sqcup_{i=2}^{n-1}(0,m),m)$ is $\mathcal{B}|S_{m+1}$, and $\mathcal{T}((m-1,0),\sqcup_{i=1}^{n-1}(0,m-1),m))$ is $\mathcal{B}|S_{m+n-1}$.

For a generic tower as shown in Fig. 4, we can identify the top-leftmost path as A and a series of events S_i. Recall that S_i means the last i edges along A leading to (m,n), including the original lattice edges (either horizontal or vertical) and shortcut edges, are connected, but the last $(i+1)$-th one is not. Let s_i represent the last broken edge toward (m,n) and the edge s_i corresponds to the event S_i. It is important to recall that the decomposition happens serially from event S_0 to S_b, corresponding to Fig. 4(a). Each decomposition (e.g., S_i's) is performed on the tower obtained from the previous decomposition (i.e., S_{i-1}'s). Essentially we have four types of edges along A: 1) one shortcut edge (e.g., s_0 in Fig. 4(a)) on layer K; 2) horizontal edges (e.g., s_1 and s_2), of which there are $r_{i+1}-t_i-r_i$ on each layer for $i \geq K$; 3) at most one vertical and topmost corner edge (s_3 or s_a or s_b) on each layer for $i>K$ and $r_i > t_{i-1}+r_{i-1}$; 4) at most one vertical but not topmost edges (s_4) on each layer for $i>K$ and $r_i = r_{i-1}$. In the following, we will show how each type of S_i can reduce a tower to another of less complexity.

3.3.1 The shortcut edge along A

For example, s_0 is in A but not in $\mathcal{B}$, so whether it is broken or not does not affect $\mathcal{B}$, and $\mathcal{B}|S_0 = \mathcal{B}$ as shown in Fig. 4(b). Using the tower notation, $\mathcal{T}(\cdots,(t_K,r_K),\cdots) \xrightarrow{S_0} \mathcal{T}'(\cdots,(t_K-1,r_K),\cdots)$ with the absence of the shortcut edge at layer K, and no changes in other layers, so the tower complexity is reduced.

3.3.2 The horizontal edges along A

For s_1, it is in both A and $\mathcal{B}$, and if it is broken but s_0 is connected, it will remove all the horizontal edges left to it on layer K and introduce a shortcut to (m,n) directly on layer $K+1$, as shown in Fig. 4(c). Using the tower notation, $\mathcal{T}(\cdots,(t_K,r_K),(0,r_{K+1}),\cdots) \xrightarrow{S_1} \mathcal{T}'(\cdots,(t_K-1,r_K),(1,t_K+r_K-1),\cdots)$, i.e., one $\triangle$ on layer K is removed, but one $\triangle$ on layer $K+1$ is introduced. However, the number of $\square$s on layer $K+1$ has been reduced to r_K+t_K-1. Recall that $r_K+t_K \leq r_{K+1}$ for a valid tower, the tower complexity is reduced overall as well.

For s_2, if it is broken but s_1 and s_0 are connected, it will also remove all the horizontal edges left to it on layer K and introduce a shortcut on layer $K+1$, as shown in Fig. 4(d). In fact, all horizontal edges along A will have the same behavior, and since they always remove at least one $\square$ on the next layer and only introduce one $\triangle$ on the next layer, therefore, the tower complexity keeps decreasing with $s_{1|2}$-like edges.

3.3.3 The vertical and topmost corner edges along A

For s_3 on layer $K+1$, if it is broken but $s_{0..2}$ are connected, it will have the same effect as S_2, since s_3 is the topmost edge of a vertical path segment and there are no branches between s_2 and s_3, so the reduction is shown as $\mathcal{B}|S_{2|3}$ in Fig. 4(d).

3.3.4 The vertical but not topmost edges along A

For s_4 on layer $K+2$, if it is broken but $s_{0..3}$ are connected, it will remove a $\square$ from the same layer, without introducing any $\triangle$ in any layer, as shown in Fig. 4(e). Using the tower notation, $\mathcal{T}(\cdots,(0,r_{K+2}),\cdots) \xrightarrow{S_4} \mathcal{T}'(\cdots,(0,r_{K+2}-1),\cdots)$. In fact, all vertical but not topmost edges along A will have the same behavior, and since they always remove one $\square$ without introducing a $\triangle$, the tower complexity is further reduced.

3.3.5 The base

On the base line, serial and parallel principles can be applied to reduce the tower complexity. For example, as shown in Fig. 4(f), if layer N has t_N $\triangle$s and r_N $\square$s, it implies that the base layer has $b-(t_N+r_N)$ edges along a single path of connectivity $p^{b-(t_N+r_N)}$, so $P(\mathcal{T}(\cdots,(t_N,r_N),b)) = p^{b-(t_N+r_N)}P(\mathcal{T}'(\cdots,(t_N,r_N),t_N+r_N))$ using the serial principle. For the t_N $\triangle$s, each of them implies two parallel paths: one by the shortcut to the destination directly, and another through a horizontal edge and then a smaller tower. Since these two paths are mutually exclusive, the PIE principle applies as $P(\mathcal{T}'(\cdots,(t_N,r_N),t_N+r_N)) = p+(1-p)P(\mathcal{T}''(\cdots,(t_N-1,r_N),t_N+r_N))$. The PIE principle can be applied repeatedly until the base becomes $\triangle$ free. After that, another top-leftmost path A'

and layer K' can be identified and the above procedures can be repeated to further reduce the tower complexity, until the decomposition leads to the components of known connectivity.

3.3.6 The overall recursion

According to Fig. 3 and Eqn. (4), the recursion process can by summarized using the following theorem:

Theorem 1: For a lattice with size $m*n$, the connection probability at vertex (m,n),

$$P(m,n) = P(A) + \sum_{i=0}^{m+n-1} P(\mathcal{B}|S_i)P(S_i),$$

where $P(S_i) = p^i(1-p)$ and $P(\mathcal{B}|S_i)$ are

$$\begin{aligned}
P(\mathcal{B}|S_0) &= p \cdot P(m,n-1),\\
P(\mathcal{B}|S_1) &= P(\mathcal{T}((1,0), \sqcup_{i=1}^{n-1}(0,m), m)),\\
&\cdots\\
P(\mathcal{B}|S_{m-1}) &= P(\mathcal{T}((m-1,0), \sqcup_{i=1}^{n-1}(0,m), m)),\\
P(\mathcal{B}|S_m) &= P(\mathcal{T}((m-1,0), \sqcup_{i=1}^{n-1}(0,m), m)),\\
P(\mathcal{B}|S_{m+1}) &= P(\mathcal{T}((m-1,0), (0,m-1), \sqcup_{i=2}^{n-1}(0,m), m)),\\
&\cdots\\
P(\mathcal{B}|S_{m+n-2}) &= P(\mathcal{T}((m-1,0), \sqcup_{i=2}^{n-2}(0,m-1), (0,m), m)),\\
P(\mathcal{B}|S_{m+n-1}) &= P(\mathcal{T}((m-1,0), \sqcup_{i=2}^{n-1}(0,m-1), m)),
\end{aligned}$$

with termination conditions given in Section 3.1 and 3.2.

The derivation of the theorem is described before this subsection and its correctness will be proved in the following section. Note that the tower complexity is always reduced by each recursion, so the entire decomposition process will terminate for sure, and then the components can be reassembled, as well as the connectivity back to $P(m,n)$.

4 Performance Evaluation

In this section, we first offer the complexity analysis of the proposed method to obtain the connectivity expressions of $m*n$ lattices. We then verify its correctness by both symbolic analysis and simulation. The impact of bond probability and lattice size on the end-to-end connectivity is also discussed. In addition, we exhibit some connectivity expressions of lattices with various sizes. Because the connectivity expression is a function of bond probability, by analyzing the expressions, we uncover more insights into the impact of bond probability.

4.1 Computational Complexity

As being aforementioned, one existing approach to calculating the exact square lattice connectivity with size $m*n$ is to use the PIE principle. $P(m,n)$ can be obtained by enumerating all possible source-destination paths (i.e., $\binom{m+n}{n}$ paths of $m+n$ segments each), crosschecking their overlapping segments, and calculating the probabilities for each combination of them. The complexity is dominated by the total number of path combinations is

$$\sum_{i=1}^{\binom{m+n}{n}} \binom{\binom{m+n}{n}}{i} = 2^{\binom{m+n}{n}} - 1.$$

Because the source-destination connectivity on a lattice is symmetric along the diagonal, i.e., $P(x,y) = P(y,x)$, the total complexity of the PIE approach is $O(2^{\binom{m+n}{n}-1})$.

In our proposed approach, the $m*n$ lattice is decomposed into towers. Each tower is further decomposed into towers of smaller scales. Thus the total complexity of our approach is determined by the total number of components generated from the entire decomposition process. For each *Tower*, we define a *Stair* which is constructed with all $\square$s of the *Tower*. In other words, a *Stair* is the remaining part of a *Tower* when removing its top and all $\triangle$s. The total number of different *Stair*s with height i is $\left[\binom{m+i}{i} - \binom{m+i-1}{i-1}\right]$. Upon each *Stair*, there can be $\triangle$s with height from 1 to $n-i$. Thus for each *Stair* with a height i, there can be $(m-1)*(n-i)+1$ cases of $\triangle$s on top of the *Stair*. Each case, together with the corresponding *Stair*, forms a *Tower*. Thus for the total number N_t of the intermediate towers, which determines the complexity of the proposed algorithm, for an $m*n$ lattice,

$$\begin{aligned}
N_t &= \sum_{i=1}^{n-1} \left[\binom{m+i}{i} - \binom{m+i-1}{i-1}\right] \cdot [1 + (m-1)(n-i)]\\
&= \sum_{i=1}^{n-1} \binom{m+i-1}{i} \cdot [1 + (m-1)(n-i)].
\end{aligned}$$

For strip lattice cases where $m \gg n$, N_t can be calculated explicitly, e.g., when $n=2$ and $n=3$, the complexity can be expressed as

$$\begin{aligned}
N_t(n=2) &= m^2,\\
N_t(n=3) &= \frac{1}{2}m^3 + \frac{5}{2}m^2 - m.
\end{aligned}$$

which shows the feasibility of our approach for strip lattices with the drop of the complexity from exponential sense (i.e., $O(2^{\binom{m+n}{n}-1})$ of the PIE approach) to polynomial sense.

Because of the symmetry of the lattice structure and for the ease of the complexity expression, analysis and comparison, let $m=n$, then,

$$\begin{aligned}
N_t &= \sum_{i=1}^{n-1} \binom{n+i-1}{i} \cdot [1 + (n-1)(n-i)]\\
&= \sum_{i=1}^{n-1} \frac{\left(n^2 - (n-1)(i+1)\right)(n+i-1)!}{(n-1)!i!}\\
&= \frac{n(2n^2-n+1)(2n-1)! - (n^3+1)(n!)^2}{n!(n+1)!}\\
&= \frac{(2n^2-n+1)(2n)!}{2 \cdot n!(n+1)!} - (n^2-n+1).
\end{aligned}$$

Applying the big O notation, the complexity of our algorithm is $O(n^2 \cdot \frac{(2n)!}{n!(n+1)!})$. We can observe that $\frac{(2n)!}{n!(n+1)!}$ is

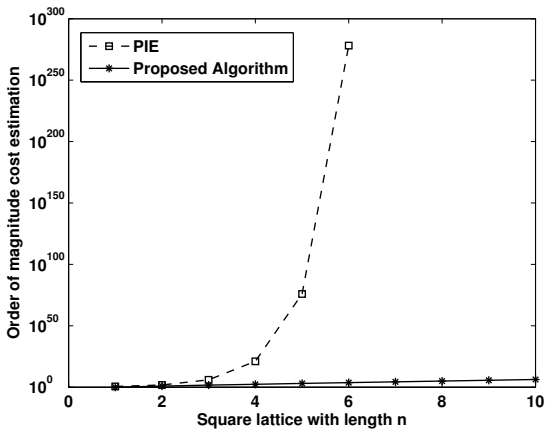


Fig. 5: The cost estimation for $n * n$ lattices.

the n^{th} $Catalan\ number$, which can be used to represent the number of monotonic paths along the edges of a lattice with $n*n$ square cells, without passing above the diagonal. Asymptotically, the n^{th} Catalan number grows as $C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}}$. Thus, we claim that the complexity of our algorithm to derive the directed connectivity expression for an $n*n$ lattice is $O(\sqrt{n}\cdot 4^n)$.

Verified by the simulation running time, our new approach is much more efficient than the PIE approach that has the combinatorial on the exponent. Shown in Fig. 5, for a $6*6$ lattice, the number of PIE basic operations, i.e., path enumerations, has already reached a magnitude over 10^{250}, while the total number of *Tower* enumerations that our approach needs has a magnitude below 10^4. Our approach is considered to be viable, especially in many engineering applications where one dimension of the lattice is limited, even though the other dimension can grow to a large number, e.g., in a VANET city block scenario with certain traffic flow directions. In addition, due to the recursive manner of our approach, when we obtain $P(m,n)$, we have also obtained all $P(x,y)$ for $x \le m$ and $y \le n$ as a byproduct, so the complexity should be amortized over all $m*n$ lattices. Further, during the recursion process, the connectivity of precalculated components can be stored for table lookup in new decomposition branches, which will greatly reduce the recursion depth and running time.

4.2 Symbolic Verification

4.2.1 2D Ladders

In [8], we derived the connectivity for 2D ladders, which is the connectivity from $(0,0)$ to $(x,1)$ on lattices, using another decomposition approach that is not extensible to lattices of more than one layer. However, we can use that approach to verify the new one. According to [8], the following recursive expressions can be defined for $P(x,1)$

$$\begin{aligned} P(x,1) &= p[p^x + P(x-1,1) - p^x\theta(x)],\ x \ge 1, \\ \theta(x) &= p[p + \theta(x-1) - p\theta(x-1)],\ x \ge 1, \end{aligned}$$

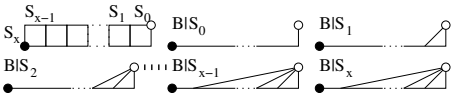


Fig. 6: The decomposition of a *Ladder*.

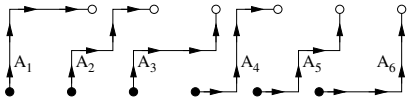


Fig. 7: All source-destination paths of a $2*2$ lattice.

where $P(0,1) = p$ and $\theta(0) = 0$. By simplifying these recursions, we can obtain the symbolic, non-recursive expression of the 2D ladder connectivity as follows

$$\begin{aligned} P(x,1) &= (p^{x+1}(-p^{x+3}(1-p)^{x+1} \\ &\quad -p(p((p-2)x+p-3) \\ &\quad +2(x+1))+x+1))/((p-1)p+1)^2. \quad (5) \end{aligned}$$

With the new approach, as shown in Fig. 6 for illustration purposes, we have the following recursions according to the decomposition process of towers (essentially of $\triangle$s)

$$\begin{aligned} P(\mathcal{T}((i,0),x)) &= p^{x-i}\cdot P(\mathcal{T}((i,0),i)), \\ P(\mathcal{T}((i,0),i)) &= p + p\cdot P(\mathcal{T}((i-1,0),i-1)) \\ &\quad -p^2\cdot P(\mathcal{T}((i-1,0),i-1)), \end{aligned}$$

with $P(\mathcal{T}((1,0),1)) = P_T = p + p^2 - p^3$. Then with the total probability in the new approach,

$$\begin{aligned} P(x,1) &= p^{x+1} + \sum_{i=0}^{x-1} P(\mathcal{T}((i,0),x))p^i(1-p) \\ &\quad +P(\mathcal{T}((x-1,0),x))p^x(1-p), \quad (6) \end{aligned}$$

which comes to the same expression as (5). For example, $P(0,1) = p$, obviously, $P(1,1) = 2p^2 - p^4$, the same as that obtained in Section 3.1 using PIE, and

$$\begin{aligned} P(2,1) &= p^7 - p^6 - 2p^5 + 3p^3, \\ P(3,1) &= -p^{10} + 2p^9 + p^8 - 2p^7 - 3p^6 + 4p^4, \\ &\cdots\ . \quad (7) \end{aligned}$$

4.2.2 2*2 Lattices

Because the approach used in [8] is not capable for lattices of more than one layer, we have to use the PIE principle. Here we use a $2*2$ lattice as an example. To use the PIE principle, we first identify all the paths from the source to the destination. For the case of a $2*2$ lattice, there are 6 paths in total, as shown in Fig. 7.

Using the PIE principle,

$$\begin{aligned} P(2,2) &= P(A_1 + A_2 + A_3 + A_4 + A_5 + A_6) \\ &= \sum_{k=1}^{6}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le 6}(A_{i_1}\cdot\ldots\cdot A_{i_k})\right) \quad (8) \\ &= p^{12} - p^{11} + 2p^{10} + 4p^9 + 2p^8 - 4p^7 - 6p^6 + 6p^4. \end{aligned}$$

where $(-1)^{k+1}$ indicates the inclusion and exclusion. Simplifying (8), we obtain the same result as that in Section 3.2 with the new approach.

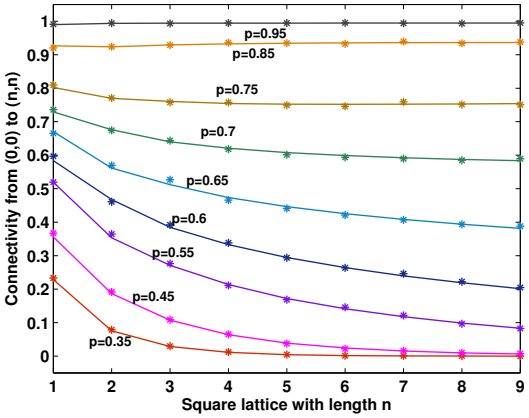


Fig. 8: The connectivity of $n * n$ lattices.

4.3 Simulation Verification

For ladders of more than one layer, however, there are no symbolic results in the literature, and the PIE complexity grows extremely quickly due to the combinatorial factor on the exponent. Thus we have to rely on the simulation results to verify the new approach. In addition, based on the calculation results, we try to understand the end-to-end connectivity from the lattice size and ratio perspective. In the next subsection, we will look at the connectivity problem from the viewpoint of bond probability by analyzing the connectivity expressions.

4.3.1 Symmetric Lattices

Figure 8 shows the connectivity of symmetric lattices whose length equals to their width. With the same bond probability, the connectivities of lattices with different size are plotted. The bond probabilities we choose here are from $p = 0.35$ to 0.95. For all the following figures in this subsection, the lines indicate the calculation results by using the obtained connectivity expressions, and the point markers show the results from the simulation. As shown in the figure, the new approach gives very accurate numerical results, which have a very good match with the simulation results, but without lengthy simulations.

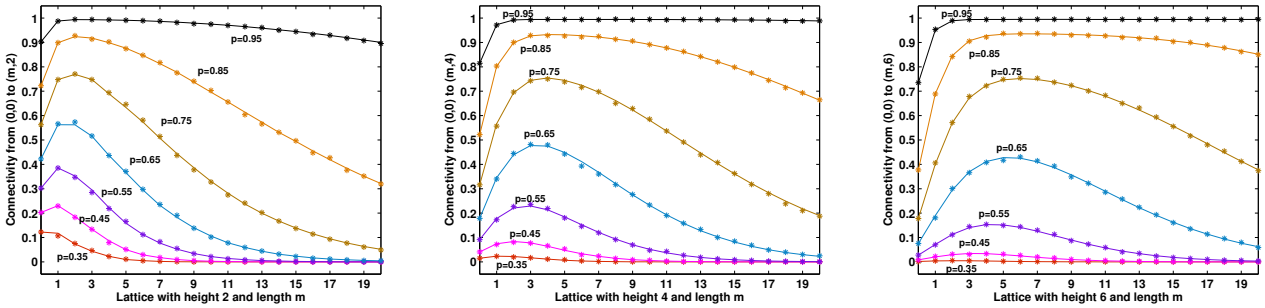
Obviously, for the same lattice, the higher the bond probability, the better the connectivity, because any adjacent vertices have a higher chance to be connected. With the same bond probability, however, the increase of lattice size does not have the same impact on the end-to-end connectivity. For small bond probabilities, i.e., from 0.35 to 0.65, a clear drop of connectivity can be observed. This actually corresponds to the conclusion in [6], where the end-to-end connectivity shows an exponential decay and the exponent is determined by how far away the bond probability is from the critical bond probability, i.e., around 0.6447 for the directed bond percolation on square lattices. But when the bond probability is reasonably large, i.e., higher than 0.65, the connectivity remains stable.

It is also interesting to investigate how the end-to-end connectivity increases with the rising of the bond probability p for the same size lattice. When p is small, e.g., increasing p from 0.35 to 0.45, only the connectivity for small $n*n$ lattices (e.g., $n < 3$) increases considerably, and the increase diminishes very quickly for larger n. However, when p is reasonably large, e.g., increasing from 0.55 to 0.65, even though the end-to-end connectivity still decreases with a large n, the increase due to an increased p actually amplifies as n increases. When p is further increased, e.g., from 0.65 to 0.75, the end-to-end connectivities are no longer to decrease with n (more obviously when $p = 0.85$ or 0.95). Recall that percolation occurs around $p = 0.6447$ on an infinite lattice, the end-to-end connectivity on a finite lattice also shows the deepest gradient when p is around 0.65, illustrated by the gap between curves of different bond probabilities in Fig. 8.

4.3.2 Strip Lattices

In many engineering fields (e.g., VANET in a city block scenario), we are more interested in propagating messages along certain directions (or traffic flows). In this sense, we shall focus more on the lattices with certain width, or strips, which are less computationally complex in terms of the connectivity expression derivation. Figure 9 shows the connectivity of lattices with different widths (n), when $n = 2$, 4 and 6 as examples. For any lattices with size $m*n$, we can observe that the higher the bond probability, the better connectivity. For each bond probability, with the fixed lattice width and increasing length, the connectivity first increases, followed by an eventual decrease. The wider the lattice, the further the peak will occur. These non-monotonic curves are very interesting to observe and very important in engineering fields to determine the optimal m, n and p for given applications. It shows that there is a trade-off between the total number of available paths and the length of each path. For a lattice with the given width, when the lattice length increases, the number of paths will increase, which brings more possibilities of connections between the source and destination. However, the length of these extra paths increases as well, leading to a lower probability to connect the source and destination along each path. For the overall end-to-end connectivity, the increase of path diversity has a positive effect, while the increase of path length has a negative one. Considering the curves shown in the figure, before the peak occurs, the positive effect of the path diversity is stronger than the negative one of path length increase, leading to the increase of connectivity probability. However, after the peak occurs, the negative effect of path length increase becomes dominating, which leads to the decrease of the overall connectivity. The peak occurs around the cases where the lattice length equals the width, which implies a symmetric $n * n$ lattice.

In all figures of this subsection, the numerical results from the new approach are very accurate when com-



(a) Lattice with $n = 2$ (b) Lattice with $n = 4$ (c) Lattice with $n = 6$

Fig. 9: The connectivity of $m * n$ lattices.

pared with the simulation ones. The analytical expressions obtained from the decomposition process can be used for further manipulation, e.g., derivatives, probability distribution functions and higher-order moments. Discussion on the expressions is presented in the next subsection.

4.4 Analysis on Connectivity Expressions

Our proposed approach provides us with the connectivity expressions of $m * n$ square lattices [25], which are polynomial functions of bond probability p, e.g.,

$$\begin{aligned} P(2,3) &= p^{17} - 7p^{16} + 15 * p^{15} - 6 * p^{14} - 9 * p^{13} - 5 * p^{12} \\ &\quad +11 * p^{11} + 8p^{10} + 4p^9 - 9p^8 - 12p^7 + 10p^5, \\ P(3,3) &= -p^{24} + 12p^{23} - 56p^{22} + 124p^{21} - 116p^{20} + 34p^{18} \\ &\quad +40p^{17} + 11p^{16} - 68p^{15} - 22p^{14} + 16p^{13} + 25p^{12} \\ &\quad +24p^{11} + 12p^{10} - 24p^9 - 30p^8 + 20p^6, \\ P(4,2) &= p^{22} - 10p^{21} + 37p^{20} - 58p^{19} + 23p^{18} + 20p^{17} \\ &\quad +15p^{16} - 34p^{15} - 16p^{14} + 6p^{13} + 15p^{12} + 16p^{11} \\ &\quad +7p^{10} - 16p^9 - 20p^8 + 15p^6. \end{aligned}$$

More are available at [25]. The first and second derivatives of the polynomials indicate the change rates of the connectivity and the first derivatives with regard to the bond probability, respectively. By computing the first and second derivatives of such connectivity expressions, we can reveal more insights quantitatively.

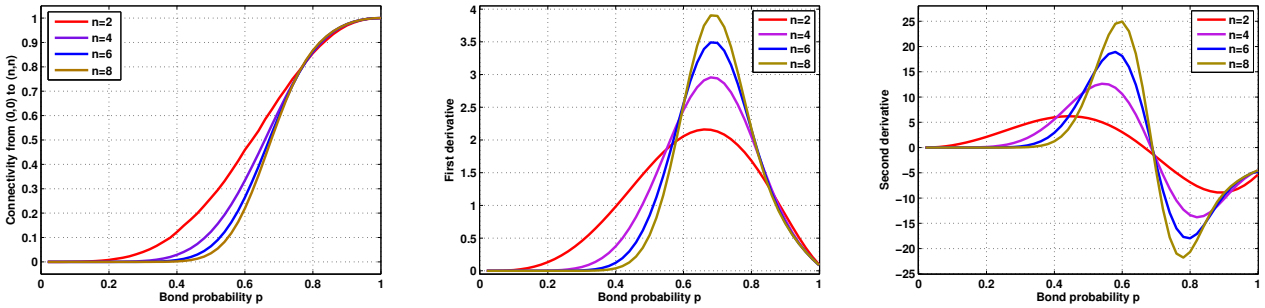
4.4.1 Symmetric Lattices

Figure 10(a) shows the $n * n$ lattice connectivity over different bond probabilities. For the same lattice, higher bond probabilities always help achieve better connectivity. For most values of p, smaller lattices always have higher connectivity. However, when the bond probability is large enough, e.g., $p > 0.8$, it is even possible for larger lattices to have higher connectivity due to many more paths.

All the connectivities increase significantly when the bond probability is between 0.4 and 0.8, where a slight increase of the bond probability will greatly increase the connectivity over lattices. The sharp transition corresponds to the same phenomenon in percolation on an infinite lattice, but here we have more microscopic results on the connectivity to any vertex on the lattice. For each curve in Fig. 10(a), the increasing slope, i.e., the first derivative of connectivity polynomial, is large when the bond probability is within a certain transition range. To quantitively define the transition range, we calculate the two inflection points of the first derivative curve. We call the range bounded by the two inflection points the *critical transition range* in this paper. Following the same legend as shown in Fig. 10(a), Fig. 10(b) and (c) show the first and second derivatives of connectivity expressions of $n * n$ lattices, where n is 2, 4, 6 and 8.

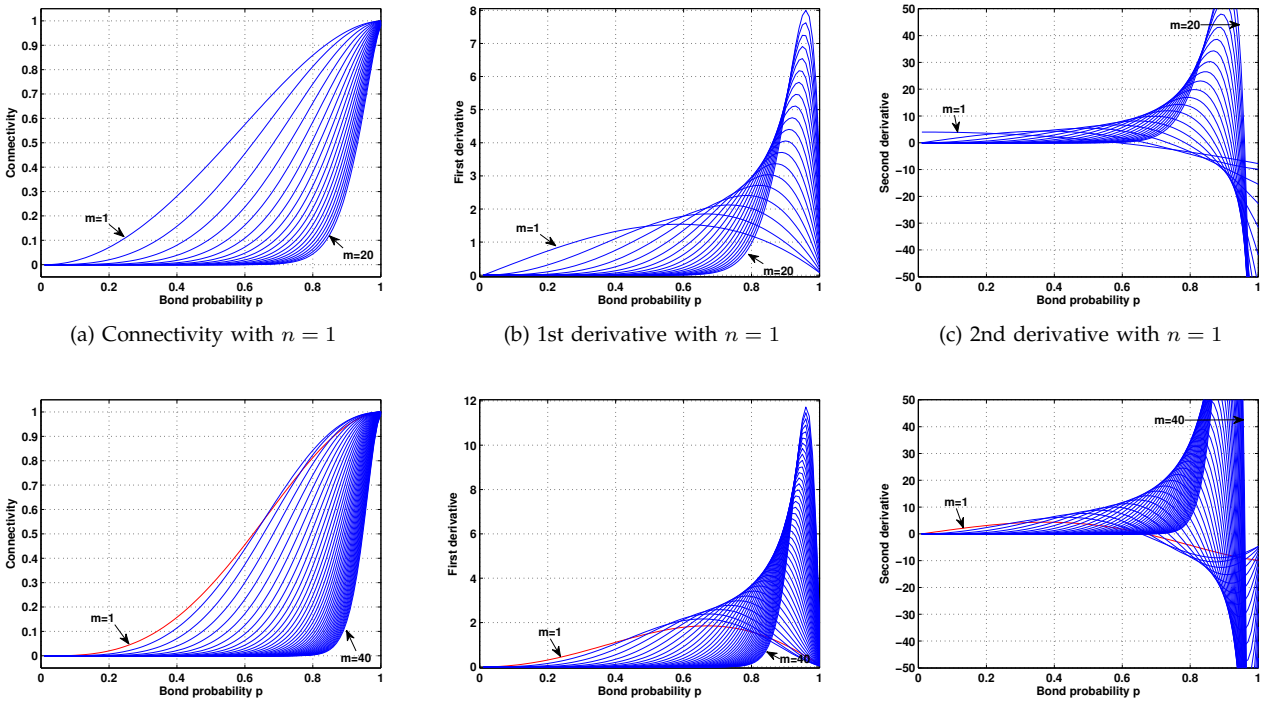
From Fig. 10(b), the difference of critical transition ranges for lattices of different size can be observed, e.g., for $2 * 2$ lattice, the connectivity has a significant increase when the bond probability falls in the range of $[0.43,\ 0.9]$. However, with the increment of the lattice size, the critical transition range shrinks, from $[0.43,\ 0.9]$ for a $2*2$ lattice to $[0.6,\ 0.79]$ for an $8*8$ lattice. Another property that can be noted from the figure is that, the larger the lattice is, the higher the connectivity increase amplitude it has, which means a sharper rise. These two observations imply that the larger the lattice size, the less sensitive the connectivity is to the low bond probability, i.e., $p \in [0,\ 0.4]$, but more sensitive to the reasonably high bond probabilities, i.e., $p \in [0.6,\ 0.8]$.

By calculating the solution of the second derivative equaling to 0, we can obtain a very important bond probability, i.e., $p_c(m,n)$, where the connectivity curve achieves the sharpest increase. This bond probability is also the probability where the maximum value of the first derivate curve occurs. By calculation, we find that the value of $p_c(m,n)$, is 0.645, 0.665, 0.668 and 0.669 for $2*2$, $4*4$, $6*6$ and $8*8$ lattices, respectively and the values are only subject to precision. An asymptotic behavior of bond probability could be observed since the difference between two consecutive critical bond probabilities reduces with the increment of the lattice size. The critical transition range can also be obtained from the second derivative, indicated by the range of the bond probability from the maximum second derivative to the minimum.



(a) Connectivity of $n * n$ lattice (b) 1st derivative of $n * n$ lattice connectivity (c) 2nd derivative of $n * n$ lattice connectivity

Fig. 10: Analysis of the $n * n$ lattice connectivity expressions.



(a) Connectivity with $n = 1$ (b) 1st derivative with $n = 1$ (c) 2nd derivative with $n = 1$

(d) Connectivity with $n = 2$ (e) 1st derivative with $n = 2$ (f) 2nd derivative with $n = 2$

Fig. 11: Analysis of the $m * 1$ and $m * 2$ lattice connectivity expressions.

4.4.2 Strip Lattices

Similarly, we investigate the connectivity expressions of strip lattices whose width is fixed, but with variable lengths, e.g., $m * 1$, $m * 2$, etc. In Fig. 11, the two rows of figures show the connectivity expressions and their first and second derivatives for lattices with width of 1 and 2, respectively. We choose these two lattice widths for illustration purposes and we believe the conclusion drawn from these figures could represent other strip lattices. For the first row where $n = 1$, 20 curves of lattices with length from 1 to 20 have been plotted. To make a fair comparison, we let the largest lattice in each row have the same length to width ratio, e.g., 20. Thus in the second row where $n = 2$, 40 curves of lattices with length from 1 to 40 have been plotted.

For lattices with the same width (i.e., in the same row), conclusions similar to $n * n$ lattices could be drawn. The connectivity increases as the bond probability increases. However, the connectivity critical transition range shrinks dramatically with the increase of the lattice length, while the connectivity increase amplitude rises significantly. In the second row of figures, the curves of lattices with $m = 1$ are plotted in red color. Specifically in Fig. 11(d), only the curve with $m = 1$ has intersections with other curves, which means that for smaller p, $P(2,1) > P(2,m)$, when $m > 1$ and for higher p, $P(2,1) < P(2,m)$. For any $m_1 > m_2 \geq 2$ with a given p, $P(2,m_1) > P(2,m_2)$ always holds. This could be further generalized that given a lattice width n and bond probability p, the monotonic increasing property of

TABLE 1: Critical bond probability for $m * n$ lattices.

$m * n$	$p_c(m,n)$	$m * n$	$p_c(m,n)$
5*1	0.8125	10*2	0.8283
10*1	0.9013	20*2	0.9061
15*1	0.9337	30*2	0.9359
20*1	0.9501	40*2	0.9514

$P(m,n)$, i.e., the connectivity from origin to (m,n), only exists when $m \geq n$.

The exact critical bond probability $p_c(m,n)$ can be calculated with the connectivity polynomial obtained. For lattices with width 1 and 2, the critical bond probabilities are shown in Table 1.

By studying the curves in Fig. 11 and numbers in Table 1, asymptotic behaviors can be observed as the lattice size grows and general properties for lattices with size $\varphi n * n$, can be summarized. When the lattice length to width ratio, i.e., φ, is fixed, the larger the n is, the narrower the critical transition range is but the larger the increase amplitude is. The critical bond probability $p_c(\varphi n * n)$ is also higher as shown in the table. Similar observations are found in the symmetric lattices as well, where φ could be considered as one. On the other hand, when n is fixed, the change of φ, implies the change of the percolation direction. We see with a larger φ, the critical bond probability $p_c(\varphi n * n)$ becomes higher.

5 Application in Urban VANETs

As a motivating example, we apply our approach to obtaining the connectivity of a realistic 2D network for urban VANETs, where a Manhattan-like road structure is considered and each road segment can be represented by an edge in the square lattice as shown in Fig. 13.

5.1 Problem Description

In an urban VANET system, many applications are based on message broadcasting, e.g., collision or traffic congestion messages can be propagated to notify drivers blocks away for them to detour well in advance; parking lots, hotels, and restaurants can advertise their availability to potential customers, reducing the extra time and fuel cost when the drivers are looking for empty spots. The efficiency of broadcasting is greatly impacted by the network connectivity with multi-hop relaying from the message source to the destination. Because message relaying only happens when the relay node falls within the transmission range of the transmitting node, the vehicle density along the road highly determines the multi-hop relaying efficiency.

We consider the most general Manhattan-like city road structure, which is composed of horizontal and vertical road segments. Such a structure can be modeled as a square lattice where each intersection is a site in the lattice and each road segment is the bond between sites. Vehicles move on each road segment between any two adjacent intersections. We assume that the distances between any two adjacent vehicles in the same lane, i.e., inter-vehicle distances, follow an independent and identical distribution (i.i.d.), which can be either derived mathematically or obtained empirically through measurement. Recent work [26], [27] made statistic analysis of the empirical data collected from the real world and found that an exponential distribution, taking vehicle density as the only parameter, can well capture the characteristics and variation of the vehicle traffic in terms of the vehicle inter-arrival time and inter-contact time [28]. Considering the nominal speed for vehicles in the same road segment, the exponential inter-arrival time is in fact equivalent to the exponentially distributed inter-vehicle distance.

By multi-hop transmissions of vehicles, the probability for a message reaching one intersection from its adjacent preceding intersection is denoted as p, i.e., the bond probability. The detailed derivation of p under the Manhattan-like city structure is given in the following subsection. Once we have p, the directed connectivity of any two intersections in such an urban VANET system can be explored, previously by simulation in [8], and now with the analytical expressions obtained by the new approach. Note that, in the real world, the bond probability p reflects the connection condition between two neighbor intersections. Besides the following mathematical derivation, p can be also obtained empirically with more realistic constraints.

Depending on the nature of a message, the message may be propagated among vehicles in the same street (e.g., expressway), which is the one-dimensional connectivity $P(x,0)$ on a line. If the message can be disseminated to all the intersections in the downtown area (e.g., parking availability), then the lattice connectivity $P(x,y)$ for all $x \leq m$ and $y \leq n$ applies. In this paper, we assume the message originates at the intersection $(0,0)$ and the directed connectivity to all other intersections will be obtained using our approach.

5.2 Bond Probability

To derive the bond probability p in the urban VANET system, we borrow the method from [8]. We start by investigating the size of the "connected vehicle cluster", within which all vehicles are connected via wireless transmissions, on the one-dimensional road. We denote the size of the cluster as a random variable (RV) C in the following derivation. Because we consider the one-dimensional case, the cluster is formed by vehicles distributed in a linear road. So, C is the sum of several inter-vehicle distance RVs, which are assumed to be exponentially distributed. It has been widely accepted that the sum of exponential distributed RVs follows a Gamma distribution, so the probability density function of C can be expressed as:

$$f_C(x) = x^{k-1}\frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, \quad \text{for } x > 0, \tag{9}$$

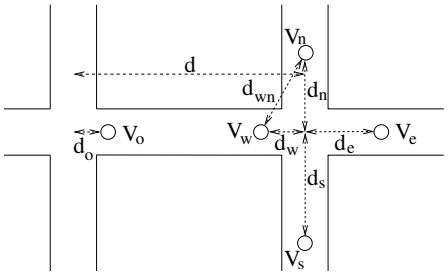


Fig. 12: Bond probability illustration.

where k and θ are the distribution parameters, $E[C] = k\theta$, $Var[C] = k\theta^2$, and $\Gamma(k)$ is the Gamma function evaluated at k. Therefore, the distribution parameters k and θ can be calculated with the first and second moments of C as:

$$k = (E[C^2]/E[C]^2 - 1)^{-1} \text{ and } \theta = E[C]/k.$$

Then the goal is left to obtain $E[C]$ and $E[C^2]$.

Let X_1 denote the inter-vehicle distance between the first and second vehicles in a cluster and R denote the communication range of vehicles. Then the expectation of the cluster size can be expressed as:

$$E[C] = E[C|X_1 < R] \times \Pr\{X_1 < R\}. \quad (10)$$

The conditional expectation $E[C|X_1 < R]$ can be expressed as:

$$E[C|X_1 < R] = E[X_1|X_1 < R] + E[C'],$$

where $E[C']$ is the expectation of the cluster size without counting the distance X_1. Because the cluster size is composed of i.i.d. RVs of the inter-vehicle distance, $E[C'] = E[C]$. We know the inter-vehicle distance X_1 follows the exponential distribution, i.e., $\Pr\{X_1 < R\} = 1-e^{-\lambda R}$. Let $\overline{X_1'}$ denote $E[X_1|X_1 < R] = \int_0^R \lambda x e^{-\lambda x}/(1-e^{-\lambda R})\mathrm{d}x$. Therefore, Eqn. (10) can be simplified as

$$E[C] = \left(\overline{X_1'} + E[C]\right) \times \Pr\{X_1 < R\},$$

and

$$\begin{aligned} E[C] &= \overline{X_1'} \times \frac{\Pr\{X_1 < R\}}{1 - \Pr\{X_1 < R\}} \\ &= \frac{1 - e^{-\lambda R}(\lambda R + 1)}{\lambda e^{-\lambda R}}. \end{aligned} \quad (11)$$

Similarly, for the second-order moment of RV C, we have

$$\begin{aligned} E[C^2] &= \Pr\{X_1 < R\} \times E[(C' + X_1)^2 | X_1 < R] \\ &= \frac{1 - e^{-\lambda R}}{e^{-\lambda R}} \times \left(2E[C]\overline{X_1'} + \overline{X_1'^2}\right), \end{aligned} \quad (12)$$

where $\overline{X_1'^2} = E[X_1^2|X_1 < R] = \int_0^R \lambda x^2 e^{-\lambda x}/(1-e^{-\lambda R})\mathrm{d}x$.

With $E[C]$ and $E[C^2]$ obtained, the Gamma approximation in Eqn. (9) can be derived.

p, as defined in previous subsection, is equivalent to the "bond probability" in percolation theory. In the urban VANET scenario, p represents the connection probability of any two neighbor intersections via wireless communication. Shown in Fig. 12, assume the distance between any two neighbor intersections is d. Let V_e, V_s, V_w, and V_n denote the vehicles which locate closest to the right intersection on their road segments, respectively. Their distances to the right intersection are d_e, d_s, d_w, and d_n, respectively. Assume the directed connection starts from the left intersection to the right one and V_o is the closest vehicle on its road segment to the left intersection with distance d_o. To make V_o connected to the left intersection, the distribution of d_o is a truncated exponential function $\lambda e^{-\lambda t}/(1-e^{-\lambda R})$, for $0 \le t \le R$.

In order to connect two neighboring intersections, the cluster, starting from V_o, should include at least one vehicle of V_e, V_s and V_n, in order to start new transmissions on the other road segments. Therefore, depending on whether V_e is connected to V_o, two disjoint cases need to be considered:

Case One: V_e is connected to V_o, which means the cluster originating from V_o has a size larger than $d-d_o$. Also with the consideration of V_o's location, the probability in this case is

$$p_1 = \int_0^R \int_{d-t}^{\infty} f_C(x)\mathrm{d}x \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda R}}\mathrm{d}t.$$

Case Two: V_e is not connected to V_o, which means the cluster size is smaller than $d - d_o$. Let V_w be the last vehicle of the cluster. To connect to the right intersection, either V_s or V_n or both of them need to connect to V_w. Denote the cluster size, i.e., the distance from V_o to V_w as x, d_o as t, and the distance from V_w to the right intersection as d_w, then at least one of d_n and d_s should be shorter than $\sqrt{(\eta R)^2 - (d-x-t)^2}$, where we consider $\eta \in (0,1)$ as the shadowing effect for signal transmissions to other perpendicular streets. Therefore, the probability that at least one of V_s and V_n is connected to V_w is $(1 - e^{-2\lambda\sqrt{(\eta R)^2-(d-x-t)^2}})$. Then the probability for this case is

$$p_2 = \int_0^R \int_{d-t-\eta R}^{d-t} (1-e^{-2\lambda\sqrt{(\eta R)^2-(d-x-t)^2}}) f_C(x)\mathrm{d}x \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda R}}\mathrm{d}t.$$

Considering the above two disjoint cases, p is given by

$$p = p_1 + p_2. \quad (13)$$

The above derivation has been verified through extensive simulation in our previous work [8]. The bond probability under different applications can have different definitions. Although the bond probability is not the main concern of this paper, it is the foundation of all the technical contribution and the analysis above shows the feasibility and practicability of our solution. We can also notice the above derivation does not take the time into consideration, which means it does not consider the change of p over time. This is because the change of bond probability depends on the change of the traffic, which is out of the scope of our main concern in this paper. For the same reason, we assume the bond probabilities of any neighbor road segments

are independent. We consider the traffic on the roads as a stationary ergodic random process, and the time average equals the ensemble average. Thus, the bond probability can represent both the percentage of time two intersections being connected and the likelihood that the two intersections being connected at a given time instant.

5.3 Connectivity of Heterogeneous Lattices

The vehicle density λ on each road segment plays an important role affecting the bond probability. In the simulation, we assume the distance between two adjacent intersections is 500 m, with the wireless transmission range of 200 m. We investigate two heterogeneous cases to study the impact of different vehicle density distributions on the lattice connectivity, as shown in Fig. 13(b) and (c), respectively. All road segments are categorized into different tiers, indicating the regions with different traffic "popularity". The darker the region, the more "popular" it is, implying a higher vehicle density, i.e., a larger λ. We assume the message origin always locates at the intersection $(0, 0)$. Two social spots, e.g., commercial area or transportation hub, are set at different locations with regard to the message origin in different cases, shown as the darkest regions in each figure. Different vehicle densities determine the heterogeneous bond probability of each road segment. We refer to [8] for the mapping between the vehicle density and bond probability, so road segments with different vehicle density have corresponding bond probabilities. For the four grey-scale regions in each heterogeneous figure, the vehicle densities are 0.02, 0.016, 0.012, and 0.01 vehicles per meter, respectively, from dark regions to light regions. And the corresponding bond probabilities are 0.89, 0.76, 0.59, and 0.47. With $1,332$ vehicles in total for each figure, Fig. 13(a) demonstrates the homogeneous case with the identical density 0.012 and identical bond probability over all road segments for comparison purposes.

Our approach can not only calculate the connectivity of lattices with homogeneous bond probabilities, it is also applicable to lattices with heterogeneous bond probabilities. The message is sent out from the origin to all other intersections within the lattice in a multihop manner by vehicles. The connectivity probabilities from $(0, 0)$ to every other intersection are calculated and plotted in Fig. 14(a), (b), and (c), corresponding to Fig. 13(a), (b), and (c), respectively. X and Y coordinates indicate the locations of intersections and the height of each bar represents the value of connection probability between the corresponding intersection and the source $(0, 0)$.

With the homogeneous vehicle density, the closer the intersection to the message origin, the higher the connection probability it has. The closer the intersection to the diagonal of the whole lattice, the higher chance it is connected to the message origin because there are more paths between itself and the source according to the nature of the directed propagation. Similar conclusions are discussed in Section 4.

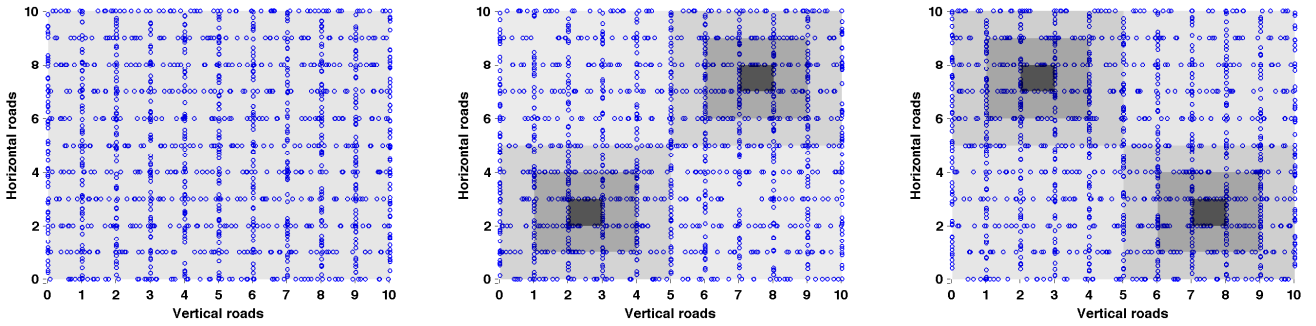
However, for the heterogeneous vehicle density cases, the connectivity distribution can be very different. For the first case, i.e., Fig. 13(b), the origin locates near one of the social spots and the other social spot locates on the diagonal of the lattice. Thus the connectivity of social spot areas is greatly increased. For the second case, however, the connectivities of all intersections are considerably low because both social spots locate far from either the origin or the diagonal. The overall connectivity is even worse than the homogeneous case, because the social spots attract more vehicles, leaving the other regions with smaller vehicle densities than the average level in the homogeneous case. Thus the high density of the social spots in this case does not help in improving the connectivity.

The connectivity analyses on heterogeneous bond probability distributions give us unique and valuable insights about some implementation details of the message propagation. First, the choice of the message source location can have a great impact on the end-to-end connectivity. As observed from the experiment, the closer the social spot is to the message source, the better the connectivity can be achieved; the closer the social spot locates to the diagonal of the two message propagation directions, the better connectivity can be achieved. This provides us a guideline where to choose the message source location, i.e., the location-fixed infrastructures or mobile vehicles which can determine the location to start broadcasting, for a better connectivity.

On the other hand, vehicles can help to increase the bond probability actively and wisely. Recall that the bond probability is affected by the vehicle transmission range. Therefore, vehicles can actively increase the transmission power to enlarge the transmission range, and further increase the bond probability. However, vehicles do not need to do so anytime anywhere which may cause more interference and a waste of energy. Only when a vehicle senses the environment and detects the bond probability of the current road segment is low, it can consider to tune up the transmission power. It can further consider the critical transition range of the bond probability discussed in the previous section. Only when the current bond probability falls within the critical transition range, it tunes up the transmission power in order to achieve a dramatic increase in the overall connectivity; otherwise, the improvement brought by the increased bond probability may not be significant.

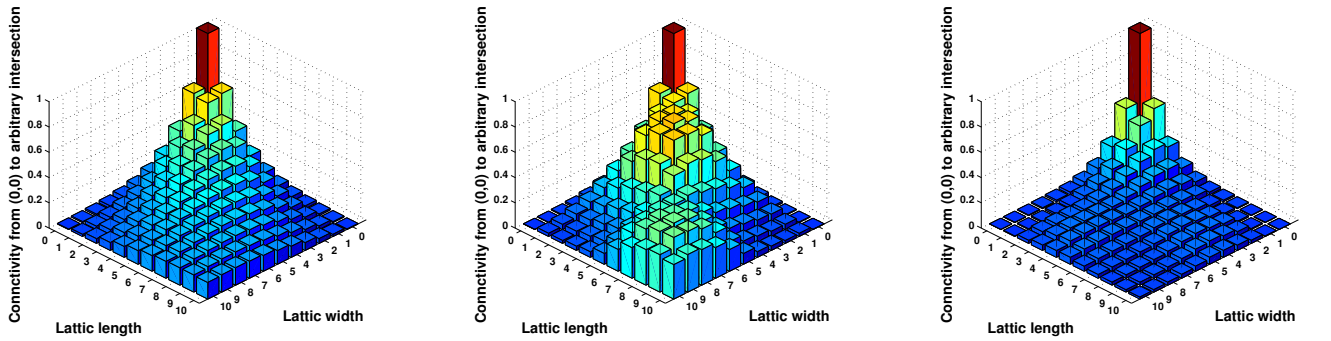
6 Further Discussion

In this paper, we illustrated a new decomposition approach following the mutually exclusive events and total probability with square lattices and the homogeneous bond probability. The approach itself is not limited by the size and shape of the lattice, as well as the bond probability on each edge. The keys are the introduction



(a) Homogeneous vehicle density (b) Heterogeneous vehicle density case 1 (c) Heterogeneous vehicle density case 2

Fig. 13: Vehicle density distribution of urban VANETs.



(a) Homogeneous vehicle density (b) Heterogeneous vehicle density case 1 (c) Heterogeneous vehicle density case 2

Fig. 14: The connectivity from $(0, 0)$ of urban VANETs.

of mutually exclusive events in the decomposition of the lattice, and the way being able to decompose all components recursively. It is our intention to apply the same technique to other directed connectivity problems on other regular tiling lattices, including triangles and hexagons used in many applications.

On the other hand, the new approach still encounters an exponential complexity (which is much better than the PIE principle that has the combinatorial factor on the exponent) of connectivity expressions, as there indeed exist $\binom{m+n}{n}$ distinct paths. When both m and n are large, this leads to a very large decomposition space. Dynamic programming approaches can be used to leverage the known connectivity of smaller components, but when the analytical expressions are reassembled, it will lead to extremely high-order polynomials (the order is as high as $2mn + m + n$). Fortunately, in most engineering problems, one dimension is often of limited size while the other dimension can grow, which keeps the exponential complexity manageable. Also depending on the needed precision, polynomial truncation can be used to limit the length and complexity of these expressions.

Although it is our goal to shed new light on the directed percolation problem, since the polynomial grows quickly if m and n go to infinity, so far it is not possible for us to obtain the exact expressions for arbitrarily large m and n. However, the polynomial expressions in terms of p for small m and n are readily available [25], and we are intended to explore more properties based on these analytical expressions. The future work will include how to verify the convergence behaviors mentioned in the literature using our exact results and exploring the math properties and coefficient patterns of the polynomials, which can reveal more insights of the connectivity.

7 Conclusions

In this paper, by proposing a new decomposition approach based on the mutually exclusive events and total probability, we presented a scheme to obtain the directed connectivity on arbitrary square lattices in a recursive manner. The results are given in polynomial expressions as a function of the bond probability on each edge. The approach and the obtained results are validated with the existing approaches and numerical results, which confirm the correctness of the new approach and the accuracy of the analytical results without lengthy simulation. Further analysis of the connectivity expressions and the possible application of these expressions in urban VANETs have been conducted to provide more insights into the directed connectivity in two-dimensional lattice networks. Our future work will focus on the issues discussed in the last section.

REFERENCES

[1] A. M. Mathai, "An introduction to geometrical probability: distributional aspects with applications," *Gordon and Breach Science*, 1999.

[2] F. Baccelli and B. Blaszczyszyn, "Stochastic geometry and wireless networks," *NOW Publishers*, 2009.

[3] C. Kurrer and K. Schulten, "Dependence of percolation thresholds on lattice connectivity," *Phys. Rev.*, E 48, pp. 614-617, 1993.

[4] R. Van der Hofstad and G. Slade, "Convergence of critical oriented percolation to super-Brownian motion above 4+ 1 dimensions," *Elsevier Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 39(2):413–485, 2003.

[5] L. Chen and F. Wu, "Directed percolation in two dimensions: An exact solution," *Differential Geometry and Physics, Nankai Tracts in Math.*, 10:160–168, 2006.

[6] E. Domany and W. Kinzel, "Directed percolation in two dimensions: numerical analysis and an exact solution," *Physical Review Letters*, 47(1):5–8, 1981.

[7] R. Durrett, "Oriented percolation in two dimensions," *JSTOR The Annals of Probability*, 999–1040, 1984.

[8] Y. Zhuang, J. Pan, and L. Cai, "A probabilistic model for message propagation in two-dimensional vehicular ad-hoc networks," in *Proc. ACM VANET'10*, pp. 31–40, 2010.

[9] N. Lu, T. Luan, M. Wang, X. Shen and F. Bai, "Capacity and delay analysis for social-proximity urban vehicular networks," in *Proc. IEEE INFOCOM'12*, pp. 1476–1484, 2012.

[10] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Comm Letters*, 6(10):437–439, 2002.

[11] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *Proc. IEEE INFOCOM'02*, pp. 1079–1088, 2002.

[12] H. Cai, X. Jia, and M. Sha, "Critical sensor density for partial connectivity in large area wireless sensor networks," in *Proc. IEEE INFOCOM'10*, 2010.

[13] S. Ukkusuri and L. Du, "Geometric connectivity of vehicular ad hoc networks: Analytical characterization," *Transportation Research Part C: Emerging Technologies*, 16(5):615–634, 2008.

[14] S. Shioda, J. Harada, Y. Watanabe, T. Goi, H. Okada, and K. Mase, "Fundamental characteristics of connectivity in vehicular ad hoc networks," in *Proc. IEEE PIMRC'08*, pp. 1–6, 2008.

[15] J. Gao and L. Guibas, "Geometric algorithms for sensor networks," *Philosophical Trans. of the Royal Society A*, 370(1958):27–51, 2012.

[16] P. Wan, K. Alzoubi, and O. Frieder, "Distributed construction of connected dominating set in wireless ad hoc networks," in *Proc. IEEE INFOCOM'02*, pp. 1597–1604, 2002.

[17] Y. Zhuang and J. Pan, "Random distances associated with equilateral triangles," *arXiv:1207.1511*, 2012.

[18] Y. Zhuang and J. Pan, "Random distances associated with rhombuses," *arXiv:1106.1257*, 2011.

[19] Y. Zhuang and J. Pan, "Random distances associated with hexagons," *arXiv:1106.2200*, 2011.

[20] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Commu.*, 27(7):1029–1046, 2009.

[21] I. Glauche, W. Krause, R. Sollacher, and M. Greiner, "Continuum percolation of wireless ad hoc communication networks," *Physica A*, 325(3):577–600, 2003.

[22] P. Santi, "Topology control in wireless ad hoc and sensor networks," *ACM Computing Surveys*, 37(2):164–194, 2005.

[23] D. Lu, X. Huang, P. Li, and J. Fan, "Connectivity of large-scale cognitive radio ad hoc networks," in *Proc. IEEE INFOCOM'12*, pp. 1260–1268, 2012.

[24] O. Dousse, "Percolation in directed random geometric graphs," in *Proc. IEEE ISIT'12*, pp. 601–605, 2012.

[25] L. Zhang, et al., Square Lattice Network Directed Connectivity Calculator, http://grp.pan.uvic.ca/~leiz/lattice_poly.html

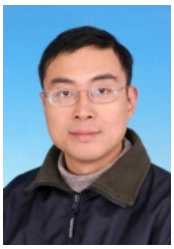
[26] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar and O. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE Journal on Selected Areas in Communications*, 25(8):1538–1556, 2007.

[27] S. Ukkusuri and L. Du, "Geometric connectivity of vehicular ad hoc networks: Analytical characterization," *Elsevier Transportation Research Part C: Emerging Technologies*, 16(5):615–634, 2008

[28] Y. Zhuang, J. Pan, Y. Luo, and L. Cai, "Time and location-critical emergency message dissemination for vehicular ad-hoc networks," *IEEE Journal on Selected Areas in Commu.*, 29(1):187–196, 2011.

Lei Zhang is currently a Ph.D. candidate in the Department of Computer Science at the University of Victoria, Victoria, Canada. He received his Bachelor's degree in information security from China University of Geosciences, Wuhan, China, in 2010. His current research interest mainly resides in advanced wireless networks, including user mobility modeling, social characteristic, and security and privacy concerns.

Lin Cai received the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Waterloo, Waterloo, Canada. She is now an Associate Professor in the Department of Electrical and Computer Engineering at the University of Victoria, Victoria, Canada. Her research interests span several areas in wireless communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic over wireless, mobile, ad hoc, and sensor networks. She has served as an Associate Editor for IEEE Transactions on Vehicular Technology, IEEE Transactions on Wireless Communications, Journal of Communications and Networks, EURASIP Journal on Wireless Communications and Networking, and International Journal of Sensor Networks.



Jianping Pan is currently an Associate Professor of computer science at the University of Victoria, Victoria, Canada. He received his Bachelor's and Ph.D. degrees in computer science from Southeast University, Nanjing, China, and he did his postdoctoral research at the University of Waterloo, Waterloo, Canada. He also worked at Fujitsu Labs and NTT Labs. His area of specialization is computer networks and distributed systems, and his current research interests include protocols for advanced networking, performance analysis of networked systems, and applied network security. He has been serving on the technical program committees of major computer communications and networking conferences including IEEE INFOCOM, ICC, Globecom, WCNC and CCNC. He is the Ad Hoc and Sensor Networking Symposium Co-Chair of IEEE Globecom 2012 and an Associate Editor of IEEE Transactions on Vehicular Technology.

Fei Tong received his Bachelor's degree in computer science and technology from South-Central University for Nationalities, Wuhan, China, in 2009, and received his M.S. degree in computer engineering from Chonbuk National University, Jeonju, South Korea, in 2011. He is currently a Ph.D. student in the Department of Computer Science, University of Victoria, Victoria, Canada. His research interests include protocol design and performance analysis for wireless ad hoc and sensor networks.