Multicast Network Coded Flow in Grid Graphs

John Gormley and Eric Manley
Department of Mathematics and Computer Science
Drake University
Des Moines, IA 50311

March 29, 2014

Abstract

Network coding, a relatively new paradigm for transmitting information through communication networks, allowing intermediate nodes in the network to combine data received on separate incoming channels before transmitting on outgoing channels. When compared with traditional routing paradigms, network coding can result in benefits such as higher throughput, fault-tolerance, and security.

In this paper, we focus on studying network coding properties on a specific class of graphs called grid graphs. Network coding properties are related to well-known Steiner properties in graphs. Specifically, Steiner properties of grid graphs are studied, because they model VLSI layout design, which makes network coding properties a natural extension of this investigation.

In particular, we looked at the maximum size of a communication group that is possible in a grid graph, given a specific desired transmission rate. Letting ρ_k (G) be the maximum fraction of nodes in graph G that can be included in a network coded multicast group with an integral flow of size k, we prove that $\rho_2(G) < 1$, $\rho_3(G) < \frac{1}{2}$, and $\rho_4(G) < \frac{1}{3}$. In the first two cases, we construct families of communication groups on grid graphs which approach these bounds. In the latter case, we present a family of communication groups approaching a density of $\frac{1}{4}$.

1 Introduction to Paper

Given a graph G, with a set of vertices V and edges E, one classical problem is to discover subgraphs of G which maintain certain connectivity properties among some subset of the nodes $S \subseteq V$. Variations of this problem include the minimum Steiner tree problem, which seeks a minimum subgraph keeping the nodes in S connected, and the Steiner packing problem, which seeks to find several edge-disjoint trees all individually connecting S. These Steiner problems are interesting, because they are NP-Complete. Also, they have many applications in computer science, including multicast (one-to-many) routing in communication networks.

One such interesting concept is network coding, a relatively new paradigm for transmitting information through communication networks, allows intermediate nodes in the network to combine data received on separate incoming channels before transmitting on outgoing channels. When compared with traditional routing paradigms, network coding can result in benefits such as higher throughput, fault-tolerance, and security. Furthermore, many network coding problems can be cast as a relaxation of a related Steiner problem, so studying network coding properties in graphs can shed light on an entire spectrum of multicasting problems. For a few references on similar work using Steiner trees and network coding, see [2], [3], [4].

In this paper, we focus on studying network coding properties of a specific class of graphs called grid graphs. The nodes of the graph can be viewed as being laid out on an $m \times n$ grid, with edges between adjacent nodes in any given column or row of the grid. Steiner properties of grid graphs are often studied, because they model problems in VLSI design. This makes network coding properties a natural extension of this investigation.

In particular, we looked at the maximum size of a communication group that is possible in a grid graph, given a specific desired transmission rate. Letting $\rho_k(G)$ be the maximum fraction of nodes in graph G that can be included in a network coded multicast group with an integral flow of size k, we found that $\rho_2(G) < 1$, $\rho_3(G) < \frac{1}{2}$, and $\rho_4(G) < \frac{1}{3}$. Thus, we have upper bounds on the maximum density of a communication group in any grid graph.

Also, we constructed a family of communication groups for any grid graph which gives us a lower bound on the maximum density of a communication group. For a flow of 2 or 3, these lower bounds approach the upper bounds in the limit, and for a flow of 4, the communication group approaches a density of $\frac{1}{4}$. As a corollary, this gives us bounds on the related Steiner problems and may have applications to the communication and layout problems. From here, we will introduce a few key definitions to our paper. Then, the proof for the maximum density is broken into three subsections, one for each k in $\rho_k(G)$.

2 Definitions

Before getting into the proofs for each $\rho_k(G)$, let us first discuss some terms, definitions, and notation.

A *graph* is a structure consisting of a vertex set, V, and an edge set, E, with at most one edge between two vertices. Two vertices are *adjacent vertices* if they are endpoints of the same edge. A *directed graph* is a graph that has a function assigning each edge to an ordered pair of vertices. A specific set of ordered pairs is called an *orientation* of the graph.

A *path* is a sequence of vertices and edges that appear in a list. A u-v-path is a path with endpoints u and v. A pair of *edge-disjoint paths* are a pair of paths in the same graph, G, such that neither shares the same edge, e, in their list.

The *in-degree* of a vertex is the number of edges that are directed towards the vertex (the head of the edge points to the vertex). The *out-degree* of a vertex is the number of edges that are directed away from the vertex (the tail of the edge points to the vertex). A *source vertex* is the origin vertex of a transmission in a graph which models a network. A *destination vertex* is the destination of a transmission.

According to Ahlswede, et al., the maximum multicast throughput is the minimum max flow among source-destination pairs [1]. In a graph with unit capacities and integral flow, max flow is equal to the maximum number of edge-disjoint source-destination paths. We refer to such routes and coding schemes as a Multicast Coded Flow. The Multicast Coded Flow rate in a given transmission is the number of edge-disjoint paths used between the source and each destination vertex for that transmission.

A *horizontal axis* on a grid graph is the set of edges which share the same row coordinate.A *vertical axis* on a grid graph is the set of edges which share the same column coordinate. The *source horizontal axis* is the horizontal axis containing the source vertex. The *source vertical axis* is the vertical axis that contains the source vertex.

The *communication set* of a given transmission is the cset containing the source vertex and destination vertices. The *density*, ρ, of a communication set is defined as the number of vertices in the communication set divided by the total number of vertices in the graph.

3 Theorem

In this section, we will state the main theorem. Then, in the next sections, we will prove the theorem.

Theorem 1. *$\rho_2(G) < 1$, $\rho_3(G) < \frac{1}{2}$, and $\rho_4(G) < \frac{1}{3}$, and examples can be constructed showing $\rho_2(G)$ approaches 1 and $\rho_3(G)$ approaches $\frac{1}{2}$ with a large enough graph. Furthermore, in the case of a coded flow of 4, there exists examples approaching a density of $\frac{1}{4}$.*

4 Supporting Lemmas

Before we start with the details of the proof subsections, we will introduce two useful lemmas.

Lemma 1. *In an oriented grid graph with a fixed communication set, if some non-source vertex has in-degree of 0 and an out-degree ≥ 2, then there exists another orientation of the grid graph in which this vertex has in-degree ≥ 1 that supports the same Multicast Coded Flow.*

Proof. Since all of the edges on the vertex are outgoing, none of these edges can possibly be used in the edge-disjoint paths between the source vertex and any destination vertices, because there is no way for a path to enter this vertex. Thus, without loss of generality, we can flip any of the directions of these edges without affecting the original source-destination paths. □

Lemma 2. *In a graph, let r be the number of destination vertices, $|E|$ the number of edges in the graph, $|V|$ the number of vertices in the graph, and k the multicasted coded flow number. Then:*

$$r \leq \frac{|E| - |V| + 1}{k - 1}$$

Proof. Given a graph, $G = (V, E)$, and a Multicast Coded Flow rate k, there are: 1 source vertex, and r destination vertices. This means that there must be $|V| - r - 1$ other vertices, where $|V|$ is the total number of vertices in graph G. Each destination vertex must have at least k incoming edges in any orientation for there to be k edge-disjoint paths going to that vertex. Also, without loss of generality, by Lemma 1, there must be at least 1 incoming edge to each of the $|V| - r - 1$ other vertices.

We can bound the size of k by counting edges according to which vertex they are oriented towards. Therefore, we can count at least $k \cdot r + |V| - r - 1$ edges, since this cannot exceed the total number of edges in the graph. So, we have $k \cdot r - |V| - r - 1 \leq |E|$. This can be rewritten as $r \leq \frac{|E| - |V| + 1}{k - 1}$. □

5 Proofs for Theorem 1

We now prove Theorem 1, looking at each case for ρ_2, ρ_3, and ρ_4 in the following subsections.

5.1 Multicast Coded Flow of 2

In this subsection, we will look at ρ_2 and prove upper and lower bounds on it when the Multicast Coded Flow rate is 2.

5.1.1 Finding an Upper Bound

Theorem 2. *In an $m \times n$ grid graph, when the Multicast Coded Flow rate is 2, the maximum size of a communication group is at most $m \cdot n - m - n + 1$. This leads to a density, $\rho_2 \leq 1 - \frac{1}{n} - \frac{1}{m} + \frac{1}{n \cdot m}$.*

Proof. Applying Lemma 1, we can see:

$$r \leq \frac{(2 \cdot m \cdot n - m - n) - (m \cdot n) + 1}{2 - 1}$$

The number of edges $(2 \cdot m \cdot n - m - n)$, and the number of vertices is $(m \cdot n)$. This simplifies to:

$$r \leq m \cdot n - m - n + 1$$

We can calculate the density of this graph by dividing the size of the communication group by the size of the entire grid graph. We just showed that the maximum size of a communication group is $m \cdot n - m - n + 1$. Also, we know that the number of total vertices in the grid graph are $m \cdot n$. This will get us a density of:

$$\rho_2 = \frac{m \cdot n - m - n + 1}{m \cdot n}$$
$$= 1 - \frac{1}{n} - \frac{1}{m} + \frac{1}{m \cdot n}$$

□

Corollary 1. *As m and n become large, the upper bound on the density of the communication set approaches 1.*

5.1.2 Constructing the Graph

Let us now construct a communication set of size $m \cdot n - m - n + 2$ in an $m \cdot n$ grid graph, G. In fact, we can construct a communication group of that size regardless of where the source vertex is in the graph. So, let (p, q) be the location of the source vertex.

To start off, one must orient the edges along source axes so that they flow away from the source vertex. Orient all other horizontal edges away from the source vertical axis and orient all other vertical edges away from the source horizontal axis. This will create a graph that resembles Figure 1.

5.1.3 Paths from Source to Destination Vertices

Let (x, y) be some vertex not on the source axes. Suppose $x \geq p, y \geq q$. The two paths to get from (p, q) to (x, y) are:

$$(q, p), (q+1, p), (q+2, p), \ldots, (x, p), (x, p+1), (x, p+2), \ldots, (x, y);$$
$$(q, p), (q, p+1), (q, p+2), \ldots, (q, y), (q+1, y), (q+2, y), \ldots, (x, y)$$

An example is provided in Figure 1. These paths are edge disjoint. Similarly, there exist two edge disjoint paths in the other three cases when $(x < q, y < p)$; $(x < q, y > p)$; $(x > q, y < p)$.

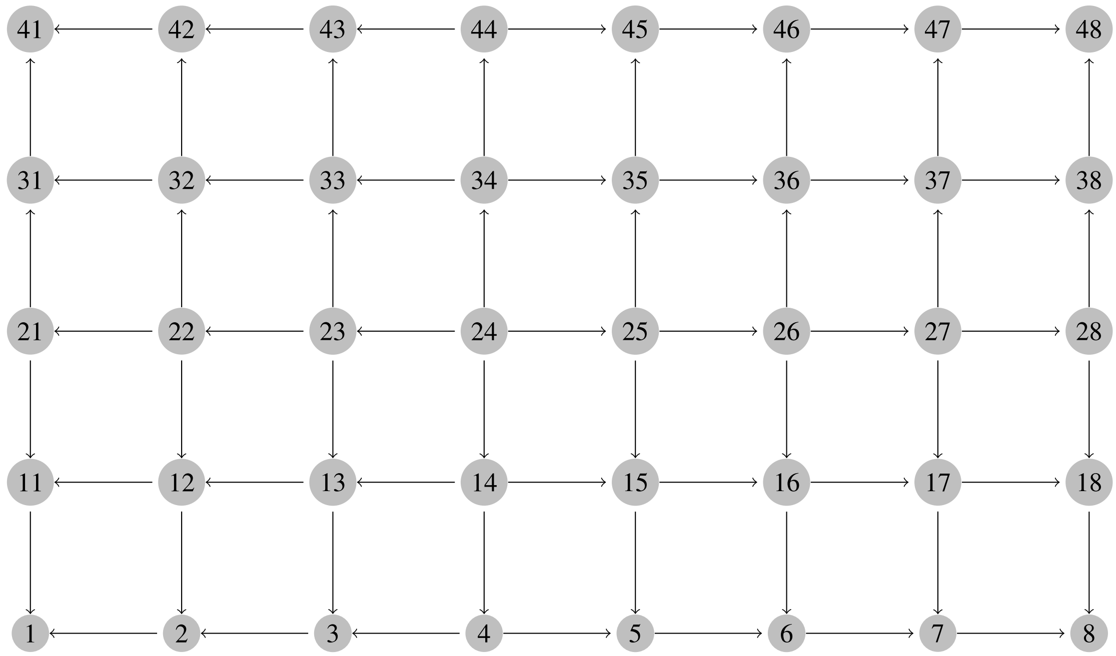

Figure 1: A grid graph with Multicast Coded Flow rate of 2 and a source at vertex 24.

5.1.4 Calculating number of Destination Vertices

With this orientation, all vertices not on the source axes may be in the communication group. The number of destination vertices on any non-source horizontal axis would be $m-1$. The number of destination vertices on a vertical axis would be $n-1$. So, the total number of vertices in the communication set would be $(m-1)\cdot(n-1)$. We know $(m-1)\cdot(n-1) = m\cdot n-m-n+1$. Since this is the upper bound in Theorem 2, we have shown that $m\cdot n-m-n+1$ is a tight upper bound.

5.2 Multicast Coded Flow of 3

We will start out this section much like the last. We will first show the upper bound on the density using Lemma 1. After, we will show a construction of a grid graph with Multicast Coded Flow rate of 3. We will show that a communication group of size $\frac{m\cdot n}{2}-\frac{3\cdot m}{2}-n+1$ can exist for any source vertex (p,q) with degree ≥ 3. This will cause the proof to break up into four cases that depend on the parity of p and $n-p$. However, for brevity, we will only show the case for the smallest density; the other three cases are in Appendix A.

5.2.1 Finding an Upper Bound

Theorem 3. *In an $m\times n$ grid graph, G, when the Multicast Coded Flow rate is 3, the density, $\rho_3 \leq \frac{1}{2}-\frac{1}{2\cdot n}-\frac{1}{2\cdot m}-\frac{1}{2\cdot m\cdot n}$.*

Proof. Looking at Lemma 1, we can see that an upper bound for r would be:

$$r \leq \frac{(2\cdot m\cdot n-m-n)-(m\cdot n)+1}{3-1}$$

This simplifies to:

$$r \leq \frac{m\cdot n-m-n+1}{2}$$

We can calculate the density of this graph by dividing the size of the communication group by the size of the entire grid graph. We just showed that the maximum size of a communication group is $\frac{m\cdot n-m-n+1}{2}$. Also, we know that the number of total vertices in the grid graph are $m\cdot n$. This will get us a density of:

$$\begin{aligned}\rho_3 &= \frac{\frac{m\cdot n-m-n+1}{2}}{m\cdot n}\\ &= \frac{1}{2}-\frac{1}{2n}-\frac{1}{2m}+\frac{1}{2mn}\end{aligned}$$

□

Corollary 2. *The density of a grid graph will approach $\frac{1}{2}$ as m and n become large.*

5.2.2 Constructing the Graph

The first step, with an $m \times n$ sized grid graph, G, is to pick an arbitrary non-corner vertex (q, p) to be the source vertex. After picking the vertex, orient all edges on the source axes away from the source vertex. The rows incident to (above and below) the horizontal source axis will also need the same orientation. Next, the columns incident to (to the right and left of) the vertical source axis will need the same orientation as the vertical source axis. This will create a three row, or column, thick *power axis*, as we will call it.

After orienting the *power axes*, the edges on the outside of the grid graph need to be oriented so that they form a cycle around the outside of the graph. Finally, all that is left are the interior columns and rows.

For the columns, start at the original horizontal source axis and orient the vertical edges incident to it such that they point towards the axis. The next vertical edges will orient away from the horizontal axis. This pattern continues along the column for every column, which will be columns 2 through $q-2$ and $q+2$ through $m-1$.

For the rows, the first rows closest to the horizontal power source axis, $p+2$ and $p-2$, should be oriented such that the horizontal edge incident to column q is oriented away from the vertical source axis. The rest of the edges in the row should be oriented towards the vertical source axis. This will orient every edge in such a way that a variant of Figure 2 will form, depending on the parity of p and $n-p$.

5.2.3 Listing the Paths from Source to Destination Vertices

Assuming (x, y) is a destination vertex where $x > q$ and $y > p$, the three paths to any destination vertex (x, y) are:

$$(q,p),(q+1,p),(q+1,p+1),...,(qq+1,y-1),(q+2,y-1),...,(x,y-1),(x,y);$$

$$(q,p),(q,p+1),...,(q,y),(q+1,y),(q+1,y+1),(q+2,y+1),...,(x,y+1),(x,y);$$

$$(q,p),(q,p-1),(q+1,p-1),(q+2,p-1),...,(m,p-1),(m,p),...,(m,y),(m-1,y),...,(x,y)$$

Similarly, there exist three edge disjoint paths in the other three cases when $(x < q, y < p)$; $(x < q, y > p)$; $(x > q, y < p)$.

5.2.4 Calculating number of Destination Vertices

Now, we will count the number of vertices that could be destination vertices in the communication group under this orientation. In order to count the communication group, we will split the grid graph up into seven sections. There are four quadrants and three axes that form. The four quadrants are counted up first, starting with the Upper Left. Then, we move to the axes, and sum the parts. At any point, Figure 2 is a good reference to look at when figuring out how the parts are summed.

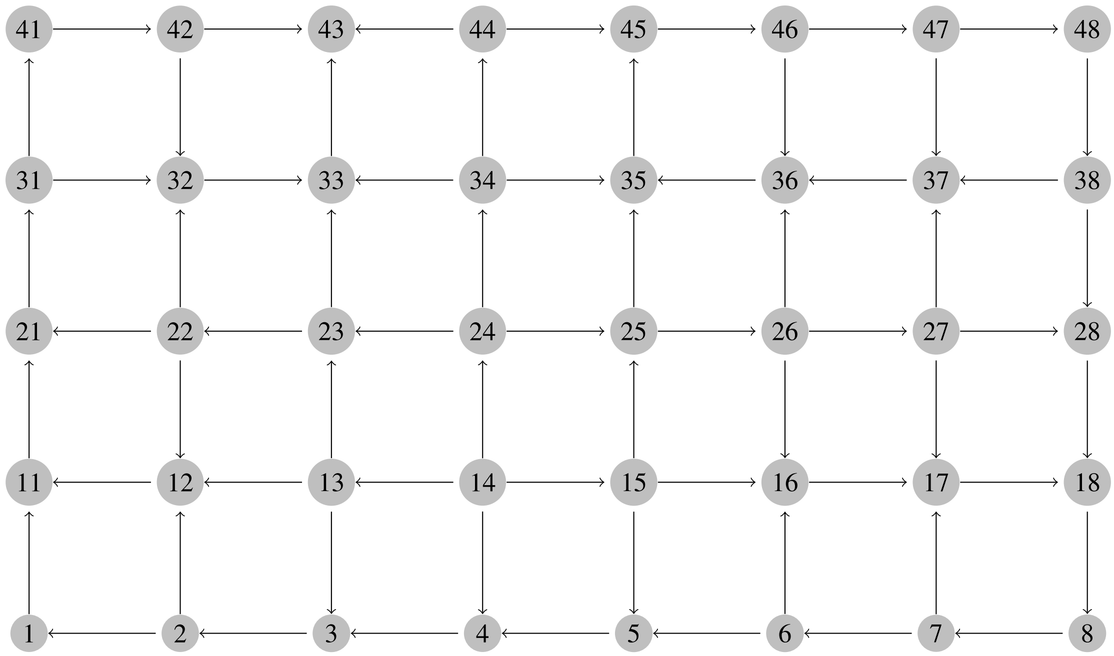

Figure 2: A grid graph with Mutlicast Coded Flow rate of 3 and a source vertex at vertex 14.

Calculating the Upper Left Quadrant

The Upper Left Quadrant will count from row 2 to row $p-2$, (starting with row 3), where exactly half are in the communication set, because there are an even number of rows. This spreads from column 2 to column $q-2$. This leads to a communication group of size:

$$\frac{(p-2)\cdot(q-3)}{2}$$

Calculating the Upper Right Quadrant

The Upper Right Quadrant will have the same counting argument, encompassing rows 2 through $p-2$,(again, starting with row 3), and columns $q+2$ through $m-1$. This leads to a communication group of size

$$\frac{(m-q-2)\cdot(p-3)}{2}$$

.

Calculating the Lower Left Quadrant

The Lower Left Quadrant spans from row $p+2$ through row $m-1$ (starting with $p+2$). It also spans from column 2 through column $q-2$. This leads to a communication group of size:

$$\frac{(q-3)\cdot(n-p-2)}{2}$$

Calculating the Lower Right Quadrant

The Lower Right Quadrant counts the same rows as the Lower Left Quadrant ($p+2$ through $n-1$). It counts from column $q+2$ through $m-1$. This leads to a communication group of size:

$$\frac{(m-q-2)\cdot(n-p-2)}{2}$$

Calculating the Power Axis

We will group two of the columns in the North Power Axis so they form one column with $p-3$ vertices in the communication set. The third column in the North Power Axis will get half of the vertices spanning from row 2 to row $k-2$. This leads to a communication group of size $\frac{p-3}{2}$ for that column. The total number of vertices in the North Power Axis communication group will be:

$$\frac{3\cdot(p-3)}{2}$$

The South Power Axis has a similar counting argument, which contains rows $p+2$ through $n-1$. This leads to a communication group of size:

$$\frac{3\cdot(n-p-3)}{2}$$

This is because two columns will combine for $n-p-3$ and the third column will have a communication group of half of that. The East-West Power Axis will yield $m-4$ potential destinations.

Summing the Quadrants and Axes

In order to find the number of destination vertices, we need to find the sum of the four quadrants and the three sections of the axes. Summing up all of these parts will lead to a total communication group of size:

$$\frac{m\cdot n}{2}-\frac{3\cdot m}{2}-n+1$$

Corollary 3. *The density of the grid graph in Case 2 would be* $\frac{1}{2}-\frac{3}{2\cdot n}-\frac{1}{m}+\frac{1}{m\cdot n}$.

As one can see, this is pretty close to what was obtained for the upper bound using Lemma 1 in Theorem 3.

5.3 Multicasted Coded Flow of 4

We will start out this section the same way we did the last two. We will first show the upper bound on the density using Lemma 1. After, we will show a construction of a grid graph with Multicast Coded Flow rate of 4. Even though such an example can be constructed regardless of source placement, to save space, we will show it only for the source at vertex $(2,2)$.

However, for space, we will show an example case where $(q=2, p=2)$. Then, we will show that a communication group of size $\frac{m\cdot n-m-n+1}{4}+1$ can exist for any source vertex (q,p) with degree ≥ 4.

5.3.1 Finding an Upper Limit

Theorem 4. *An upper bound for the size of a communication group, r, on an $m\times n$ grid graph, G, with Multicast Coded Flow rate of four is* $\frac{m\cdot n-m-n+1}{3}$. *This leads to a denisty of* $\rho_4\leq\frac{1}{3}-\frac{1}{3\cdot m}-\frac{1}{3\cdot n}+\frac{1}{3\cdot m\cdot n}$.

Proof. Applying Lemma 1, one can see that the upper bound for r will be:

$$r\leq\frac{(2\cdot m\cdot n-m-n)-(m\cdot n)+1}{4-1}$$

This simplifies to:

$$r\leq\frac{m\cdot n-m-n+1}{3}$$

We can calculate the density of this graph by dividing the size of the communication group by the size of the entire grid graph. We just showed that the maximum size of a communication group is $\frac{m\cdot n-m-n+1}{3}$. Also, we know that the number of total vertices in the grid graph are $m\cdot n$. This will get us a density of:

$$\begin{aligned}\rho_4&\leq\frac{\frac{m\cdot n-m-n+1}{3}}{m\cdot n}\\&\leq\frac{1}{3}-\frac{1}{3n}-\frac{1}{3m}+\frac{1}{3mn}\end{aligned}$$

□

5.3.2 Creating the Graph

In this section, we will show how to construct a graph with a source at $(2,2)$. A graph similar to the one described below can be constructed for any arbitrary source, though.

To start off, we place the source at $(2,2)$. Next, we orient the edges along the source axes away from the source (this will be row 2 and column 2). Similarly, for rows/columns 1 and 3, the edges should be oriented away from the source.

The vertical edge in column 4 that connects rows 1 and 2 must be oriented toward row 1 (the top of the graph). The vertical edge in column 5, however, must be oriented towards row n (the bottom of the graph). This pattern continues with the edge in column 6 being oriented toward row 1 and the edge in column 7 being oriented toward row n.

Starting in row 4, the horizontal edges that connect columns 1 and 2 must be oriented such that the first edge points towards column 2. Then, with row 5, the edge points towards column 1. This pattern continues along those edges such that the edge in row 6 points towards column 2. The horizontal edges that connect rows 2 and 3 should all be oriented away from the vertical source axis.

The first destination vertex will start at $(4,4)$. Since, in the case with a Multicast Coded Flow rate of 4, the destination vertices need 4 edge-disjoint paths, this means that the edges adjacent to the vertex will need to point towards it. This will be true for all destination vertices.

Starting in column 5, the remaining vertical edges in every odd column must be oriented away from the horizontal source axis. Similarly, for odd rows starting at row 5, the horizontal edges must be oriented away fro the vertical source axis.

Finally, all that is left to do is orient the last row and last column away from their respective source axes.

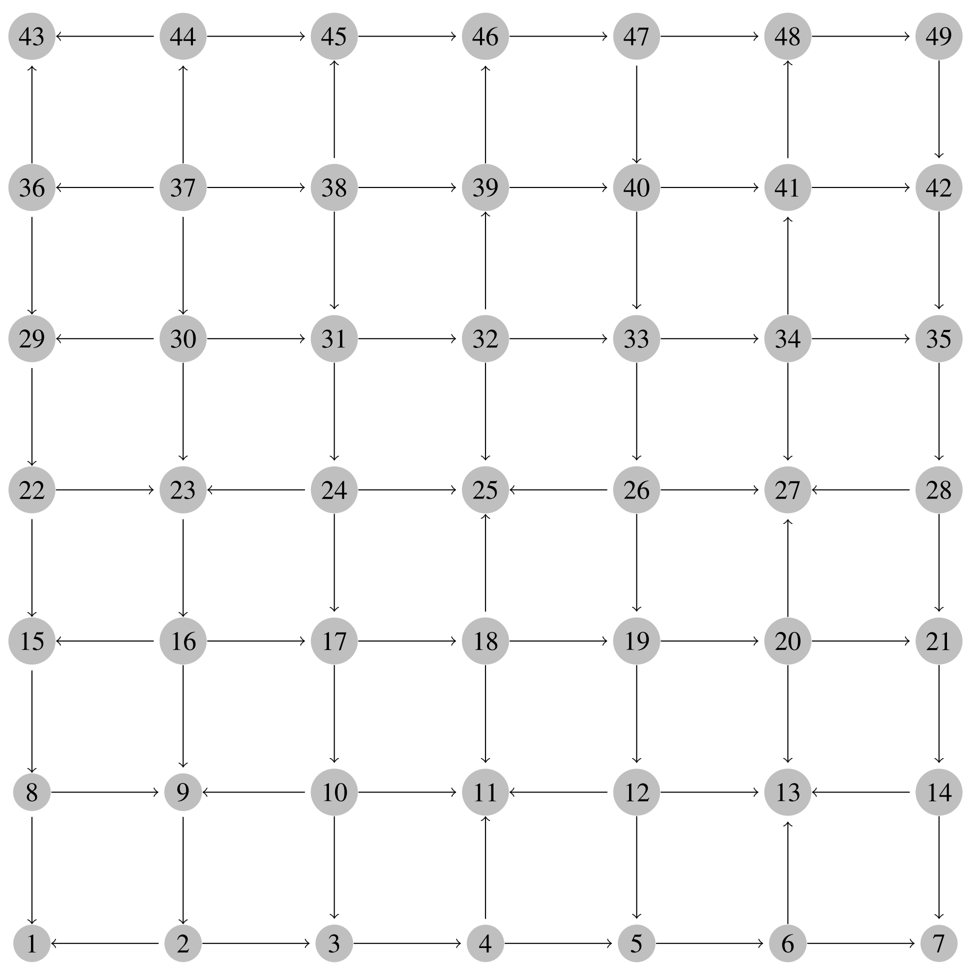

Figure 3: A grid graph with Multicast Coded Flow rate of 4 and a source vertex at vertex 37.

5.3.3 Listing the Paths from Source to Destination Vertices

Let us assume that (x,y) is some destination vertex under the construction shown above. The four edge-disjoint paths from (q,p) to (x,y) are:

$$(q,p),(q+1,p),...,(x-1,p),(x-1,p+1),...,(x-1,y),(x,y);$$
$$(q,p)(q-1,p),(q-1,p+1),(q-1,p+2),(q,p+2),(q,p+3),...,(q,y+1),(q+1,y+1),...,(x,y+1),(x,y);$$
$$(q,op,(q,p-1),(q+1,p-1),(q+2,p-1),...,(x+1,p-1),(x+1,p),...,(x+1,y),(x,y);$$
$$(q,p),(q,p+1),(q+1,p+1),(q+2,p+1),...,(x-1,p+1),(x-1,p+2),...,(x-1,y),(x,y)$$

5.3.4 Counting the Number of Vertices

When looking at Figure 4, one can see that the destination vertices will be in every other row and column starting with vertex $(4,4)$. So, counting up the destination vertices along row 4 will lead to getting:

$$\frac{m-4}{2}$$

This is because the transmission will never reach the first three vertices, nor the last one. Using a similar argument, one can count up the destination vertices along column 4 will lead to getting:

$$\frac{n-4}{2}$$

In order to find the total number of destination vertices, one must multiply the destination vertices along the row times the number of destination vertices along the column. This will lead to a total number of destination vertices:

$$\frac{m-4}{2}\cdot\frac{n-4}{2}$$
$$=\frac{m\cdot n}{4}-m-n+4$$

In order to find the density of this graph, we divide $\frac{m\cdot n}{4}-m-n+4$ by $n\times m$. This leads to a density of:

$$\rho_4=\frac{\frac{m\cdot n}{4}-m-n+4}{n\cdot m}$$
$$=\frac{1}{4}-\frac{1}{n}-\frac{1}{m}+\frac{4}{m\cdot n}$$

6 Conclusion

We have found upper bounds on the density of a communication group in grid graphs for network coded integral flow rates of 2, 3, and 4. We have also presented constructions of families of communication groups with density approaching the upper bound for flow rates of 2 and 3. We believe that the upper bound of $\frac{1}{3}$ for the flow rate of 4 could be lowered to $\frac{1}{4}$, and we intend to continue investigating this in the future. In the future, we would also like to extend this to additional classes of graphs, including multilayer grid graphs and triangular grid graphs.

References

[1] R Alswede, Ning Cai, Y R. Li, S, and R.W Yeung. Network information flow. *Information Tehory, IEEE Transactions on*, (4):1204–1216, Jul 2000.

[2] Nam-Dng Hong and Thorsten Koch. Steiner tree packing revisited. *Mathematical Methods of Operations Research*, 76(1):95 – 123, 2012.

[3] Shang Song Pu and Jing Tong. Steiner minimal trees in rectilinear and octilinear planes. *Acta Mathematica Sinica*, 23(9):1577 – 1586, 2007.

[4] Martin Zachariasen and Andr Rohe. Rectilinear group steiner trees and applications in vlsi design. *Mathematical Programming*, 94(2/3):407, 2003.