



A graph theoretical analysis of the number of edges in k -dense graphs

Linda Eroh^a, Henry Escudro^b, Ralucca Gera^c, Samuel Prahlow^d, Karl Schmitt^e

^aDepartment of Mathematics, University of Wisconsin-Oshkosh, Oshkosh, WI, USA

^bDepartment of Mathematics, Juniata College, Huntingdon, PA, USA

^cDepartment of Applied Mathematics, Naval Postgraduate School, Monterey, CA, USA

^dDepartment of Mathematics and Statistics, Valparaiso University, Valparaiso, IN, USA

^eDepartment of Mathematics and Statistics, Department of Computer and Information Sciences, Valparaiso University, Valparaiso, IN, USA

eroh@uwosh.edu, escudro@juniata.edu, rgera@nps.edu, {samuel.prahlow, karl.schmitt}@valpo.edu

Abstract

Due to the increasing discovery and implementation of networks within all disciplines of life, the study of subgraph connectivity has become increasingly important. Motivated by the idea of community (or sub-graph) detection within a network/graph, we focused on finding characterizations of k -dense communities. For each edge $uv \in E(G)$, the **edge multiplicity** of uv in G is given by $m_G(uv) = |N_G(u) \cap N_G(v)|$. For an integer k with $k \geq 2$, a **k -dense community** of a graph G , denoted by $DC_k(G)$, is a maximal connected subgraph of G induced by the vertex set

$$V_{DC_k(G)} = \{v \in V(G) : \exists u \in V(G) \text{ such that } uv \in E(G) \text{ and } m_{DC_k(G)}(uv) \geq k - 2\}.$$

In this research, we characterize which graphs are k -dense but not $(k + 1)$ -dense for some values of k and study the minimum and maximum number of edges such graphs can have. A better understanding of k -dense sub-graphs (or communities) helps in the study of the connectivity of large complex graphs (or networks) in the real world.

Keywords: k -dense subnetworks (or k -dense subgraph), k -dense community, k -dense graph, k -core, k -core subnetwork

Mathematics Subject Classification : 05C69, 05C75, 05C62

DOI: 10.5614/ejgta.2016.4.1.4

Received: 3 December 2014, Revised: 5 March 2016, Accepted: 10 March 2016.

1. Motivation and definitions

This section covers definitions and notation that will be relevant for the whole study. Please refer to [1] for notation and terminology that are not discussed in this paper. In this study, all graphs are simple (i.e., no loops and no multiple edges). Recall that $\delta(G)$ denotes the minimum degree and $\Delta(G)$ denotes the maximum degree among all vertices in G . Also, a graph G is k -edge connected, if for any set S of fewer than k edges $G - S$ is connected, and there is a set T of k edges such that $G - T$ is disconnected.

Also recall that a *clique* in a graph G is a complete subgraph of G . The order of a largest clique is the *clique number* $\omega(G)$. Note that the maximum clique in a graph may not be unique. It was shown by Karp in [2] that the decision problem of finding a clique of some integer size k in a given graph G is NP-complete.

For a graph G and vertex $v \in V(G)$, the *open neighborhood* of v , denoted by $N_G(v)$, is the set of all neighbors of v , namely, $N_G(v) = \{u : uv \in E(G)\}$. If we restrict our attention to a subgraph H of G , we then have $N_H(v) = \{u : uv \in E(H)\}$. The *closed neighborhood* of v is $N[v] = N(v) \cup \{v\}$.

In this paper, we study the idea of k -dense subgraphs, introduced by Saito, Yamada and Kazama in [4] as an alternative to cliques in detecting communities in a graph. We believe this is an alternative to possibly detecting cliques in graphs, since the inferred graphs may not always contain the complete information, such as inferring all router connections in a network. Thus, a clique in the real topology may be a k -dense community in the inferred and studied graph. The following definitions were introduced in [4] but we rephrase them here in graph theoretical terms.

Definition 1. Let G be a graph. For each edge $uv \in E(G)$, the **edge multiplicity** of uv in G is given by $m_G(uv) = |N_G(u) \cap N_G(v)|$.

Definition 2. For an integer k with $k \geq 2$, a **k -dense community** of a graph G , denoted by $DC_k(G)$, is a maximal connected subgraph of G induced by the vertex set

$$V_{DC_k(G)} = \{v \in V(G) : \exists u \in V(G) \text{ such that } uv \in E(G) \text{ and } m_{DC_k(G)}(uv) \geq k - 2\}.$$

In other words, $DC_k(G)$ is a connected subgraph of G with k or more vertices in which every two adjacent vertices have at least $k - 2$ neighbors in common in $DC_k(G)$. Note that a k -dense community might contain a connected proper subgraph $G[S]$ where $S \subseteq V(G)$ and for every pair of vertices u, v in S that are adjacent in $G[S]$, we have $m_{G[S]}(uv) \geq k - 2$. Here, $G[S]$ denotes the subgraph of G induced by S . In this case, we call S a **k -dense sub-community** of G . For example, consider the graph G obtained from two copies of K_5 , by identifying 3 of the vertices of both copies (thus $G = K_3 + 2K_2$). Each of the original copies of K_5 is a 5-dense sub-community of G while the graph G itself is a 5-dense community.

As was shown in [4], a k -clique in a graph G is also a k -dense sub-community in G . However, the converse is not true. Consider the graph $H = G[\{v_1, v_2, \dots, v_6\}] = C_4 + 2K_2$ shown in Figure 1. Note that H is a 4-dense community in G that is not K_4 nor contains K_4 as a subgraph.

One can observe that just like cliques in graphs, a given graph might have more than one k -dense community. In this paper, we say that the union of all k -dense communities of G is the *k -dense subgraph of G* , a concept that was introduced in [4] under the name of k -dense subnetwork.

Definition 3. For an integer $k \geq 2$, the **k -dense subgraph** of a graph G , denoted by $D_k(G)$, is the union of all the k -dense communities in G . If $D_k(G) = G$, then we say that G is a **k -dense graph** or simply that G is **k -dense**. If $D_k(G) = G$ and $D_{k+1}(G) \neq G$, then we say that G is a **k^* -dense graph** or simply that G is **k^* -dense**.

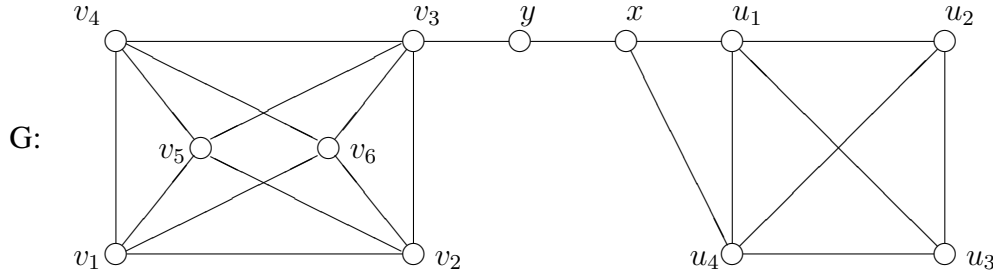


Figure 1. $G[\{v_i, u_j | 1 \leq i \leq 6, 1 \leq j \leq 4\}]$ is the 4-dense subgraph of G , $G[\{v_j | 1 \leq j \leq 4\}]$ is a 4-dense community that does not contain a 4-clique, $G[V(G) - \{y\}]$ is the 3-dense subgraph of G , while G itself is 2-dense.

Note that the k -dense subgraph, $D_k(G)$, of a graph G is unique but need not be connected. Also, a graph G is k^* -dense if and only if G is k -dense but not $(k + 1)$ -dense.

In [4], Saito et al. introduced the idea of the k -dense subgraph of a graph G as an alternative way to detect communities in graphs with many vertices and edges (viewed as large-scale complex networks). Earlier methods used k -cliques (complete subgraphs that contain k vertices) and k -cores (maximal induced subgraphs whose minimum degree is $k - 1$). By applying their algorithm to detect close-knit communities using the k -dense method on various real-life networks, Saito et al. found that their method is almost as efficient as the k -core method. Moreover, the communities found using the k -dense method proved to be comparable to those found using the k -clique method. Saito et al. indicate that a hierarchy of k -clique $\subseteq k$ -dense $\subseteq k$ -core exists, without providing a formal proof. They also state that the complete graph on k vertices is k -dense (and a k -core) and that if G is a k -dense graph with k vertices, then G must be a k -clique (complete). We restate the last claim as Proposition 3.1 and present a proof for it.

While the authors of [4] analyzed case studies on the concept of the k -dense subgraph, this paper pursues a graph theoretical study on the topic. In particular, we focus on determining lower and upper bounds for the number of edges in a graph with n vertices that is k^* -dense. In Section 3 we present general bounds and a realization result on the number of edges in k^* -dense graphs. In Section 4 we present the minimum number of edges in 2^* -dense, 3^* -dense and 4^* -dense graphs, followed by the maximum number of edges in a k^* -dense graph in Section 6. Finally, in Section 7, we present complete results for 2^* , 3^* -dense and 4^* -dense graphs.

2. General observations

Using the definitions from Section 1, we have the following quick proposition.

Proposition 2.1. A graph G is k -dense if and only if every edge of G appears in at least $k - 2$ different triangles.

The next result tells us that the minimum degree of a k -dense graph is at least $k - 1$.

Proposition 2.2. *Let G be a nontrivial graph that is k -dense. If $v \in V(G)$, then $\deg(v) \geq k - 1$. Furthermore, if $\deg(v) = k - 1$, then v is in a subgraph that is isomorphic to K_k .*

Proof. Suppose $v \in V(G)$ for some nontrivial graph G that is k -dense. Since G is nontrivial and has no isolated vertices, v must be incident with an edge uv . Since u and v have at least $k - 2$ common neighbors, it follows that $\deg(v) \geq k - 1$. Now, suppose $\deg(v) = k - 1$, say $N(v) = \{v_1, v_2, \dots, v_{k-1}\}$. Then for each i , $1 \leq i \leq k - 1$, the vertices v and v_i have $k - 2$ common neighbors. Since $\deg(v) = k - 1$, those common neighbors must be v_j for $1 \leq j \leq k - 1$, $j \neq i$. Thus, $v, v_1, v_2, \dots, v_{k-1}$ induces a complete subgraph. \square

If G is a k -dense graph, then $G - v$ is not necessarily k -dense. The following result tells us that $G - v$ is at least $(k - 1)$ -dense.

Proposition 2.3. *Suppose G is a k -dense graph, where $k \geq 3$, and $v \in V(G)$. Then $G - v$ is at least $(k - 1)$ -dense.*

Proof. Suppose G is k -dense and $v \in V(G)$. Let $u, w \in V(G - v)$ such that $uw \in E(G - v)$. Then u and w are adjacent in G as well. Hence, u and w have $k - 2$ common neighbors in G , one of which may be v . This means that u and w will have at least $k - 3$ common neighbors in $G - v$. \square

To see that the graph $G - v$ need not be $(k - 1)^*$ -dense if G is k^* -dense, consider the graph G obtained from a 5-clique, whose vertices are v_1, v_2, \dots, v_5 , and a 6-clique, whose vertices are u_1, u_2, \dots, u_6 , by identifying v_5 with u_5 (that is, $v_5 = u_5$ in G). Then G is 5^* -dense (G is not 6-dense since v_1 and v_2 have only 3-common neighbors) and $G - u_6$ is 5^* -dense as well.

The next proposition gives us a sufficient condition for a graph G on n vertices to be k -dense.

Proposition 2.4. *If G is a graph of order n and $\delta(G) \geq \frac{n+k}{2} - 1$, then G is k -dense.*

Proof. Let u and v be any two adjacent vertices in G . Since $\delta(G) \geq \frac{n+k}{2} - 1$, there are at least $2 \left[\frac{n+k}{2} - 1 \right] - 2 = n + k - 4$ edges between u and v and the remaining $n - 2$ vertices in the graph. Let x be the number of common neighbors of u and v . Then $n + k - 4 - x \leq n - 2$. This implies that $x \geq k - 2$, and so u and v have at least $k - 2$ common neighbors. \square

We now look at the edge connectivity of a k -dense connected graph.

Proposition 2.5. *If a graph G is k -dense and connected, then G is at least $(k - 1)$ -edge-connected.*

Proof. Suppose that G is k -dense and connected. If $k = 2$, then since G is connected, G is 1-edge-connected. If $k \geq 3$, consider any set of $k - 2$ edges e_1, e_2, \dots, e_{k-2} . We claim that $G - \{e_1, e_2, \dots, e_{k-2}\}$ is connected. Let $u, v \in V(G)$. Since G is connected, there exists a $u - v$ path in G , say $u = u_0, u_1, u_2, \dots, u_p = v$. For each i , $0 \leq i \leq p - 1$, if $u_i u_{i+1} = e_j$ for some j , then notice that u_i and u_{i+1} have at least $k - 2$ common neighbors. Since there are only $k - 3$ remaining edges e_j , one of the common neighbors of u_i and u_{i+1} in G is still a common neighbor in $G - \{e_1, e_2, \dots, e_{k-2}\}$, so we can replace the edge $u_i u_{i+1}$ with a path of length 2 through this common neighbor. Thus, there is still a $u - v$ walk in $G - \{e_1, e_2, \dots, e_{k-2}\}$. \square

The converse of Proposition 2.5 is not true. Moreover, there is no similar result for vertex connectivity. These are shown in the two results that follow.

Proposition 2.6. *For every positive integer t , there is a graph that is $2t$ -connected and $2t$ -edge connected that is 2^* -dense.*

Proof. The cartesian product of t copies of C_4 is a graph that is connected and contains no triangles. Hence, the graph is 2-dense but not 3-dense. A minimum vertex cut consists of the $2t$ neighbors of a vertex, and a minimum edge cut consists of the $2t$ edges incident with a vertex. \square

Proposition 2.7. *For every positive integer $k \geq 3$, there exists a connected graph G that is k -dense but not 2-connected.*

Proof. Take two copies of K_k and identify a vertex. The resulting graph is k -dense yet has a cut-vertex. \square

3. Bounds, realization, and characterizations

For any integer k with $k \geq 2$, the complete graph K_k is k -dense but not $(k + 1)$ -dense. Hence, we know that a k^* -dense graph exists for each integer $k \geq 2$. A more interesting question is that of determining the values of k and n for which there is a graph on n vertices that is k^* -dense. We will answer this question in this section and then classify which graphs on n vertices are 2^* -dense, $(n - 1)^*$ -dense, and n^* -dense.

We begin by giving bounds for k in terms of the number of vertices of the graph. Since the edge multiplicity of every edge in a graph on n vertices is between 0 and $n - 2$, we have the following.

Observation 1. *If G is k^* -dense graph on n vertices, then $2 \leq k \leq n$.*

The next result shows that the bounds given in Observation 1 are sharp and characterizes which graphs satisfy the given bounds.

Proposition 3.1. *Let n be an integer such that $n \geq 3$ and let G be a graph on n vertices. The following are true.*

- (a) *The graph G is 2^* -dense if and only if G has no isolated vertices and there is an edge uv in G such that $N(u) \cap N(v) = \emptyset$.*
- (b) *The graph G is n^* -dense if and only if $G = K_n$.*

Proof. The first part follows readily from the definition of k^* -dense given in Definition 3 with $k = 2$.

Let us now establish the second part. Since any two vertices in K_n are adjacent, it follows that every edge in K_n has edge multiplicity $n - 2$ and so K_n is n -dense. By definition, no graph on n vertices is $(n + 1)$ -dense. It follows that K_n is n^* -dense. We now establish the converse by proving its contrapositive. Let G be a graph on n vertices such that $G \neq K_n$. Thus, there exists two vertices u and v that are not adjacent in G . Since any graph on n vertices with isolated vertices is not n -dense, we may assume that u is adjacent to some vertex w in G . Since $v \notin N(u)$, it follows that the edge uw has edge multiplicity at most $n - 3$. Thus, G is not n -dense and hence, is not n^* -dense. \square

We are now ready to determine for which values of k and n there exists a graph G on n vertices that is k^* -dense.

Proposition 3.2. *If k and n are integers such that $2 \leq k \leq n$, then there is a k^* -dense graph on n vertices.*

Proof. Let q and r be the unique pair of integers such that $n = kq + r$ where $0 \leq r \leq k - 1$. If $r = 0$, let $G = qK_k$ and if $0 < r \leq k - 1$, let $G = (q - 1)K_k \cup (K_k + K_r)$. One can verify that in both cases, G is a k^* -dense graph on n vertices. \square

The next result characterizes the graphs on n vertices that are $(n - 1)^*$ -dense.

Proposition 3.3. *If n is an integer such that $n \geq 3$, then G is $(n - 1)^*$ -dense if and only if $G \cong K_n - e$.*

Proof. Observe that every edge in $K_n - e$ has edge-multiplicity at least $n - 3$. Thus, $K_n - e$ is $(n - 1)$ -dense but not n -dense and so $K_n - e$ is $(n - 1)^*$ -dense. From Proposition 3.1, K_n is n^* -dense. Now, let G be a graph on n vertices such that $G \neq \{K_n, K_n - e\}$. It follows that there exist distinct $uv, wx \in E(G)$ for some vertices (not necessarily all distinct) $u, v, w, x \in V(G)$. We may assume that G has no isolated vertices for if G does, then G is not $(n - 1)$ -dense. Suppose first that $\{u, v\} \cap \{w, x\} \neq \emptyset$, say $u = w$. It follows that $v \neq x$ and each of v and x is not in $N(u)$. Moreover, every edge incident to u will have edge multiplicity at most $n - 4$. Thus, G is not $(n - 1)$ -dense. Suppose now that u, v, w and x are distinct vertices; that is, $\{u, v\} \cap \{w, x\} = \emptyset$. If $w \in N(u)$, then the edge uw will have edge multiplicity at most $n - 4$ and G is not $(n - 1)$ -dense. Using similar arguments, one can show that G is not $(n - 1)$ -dense if either $x \in N(u)$ or $\{w, x\} \cap N(v) \neq \emptyset$. Thus, we may assume that $G[\{u, v, w, x\}] = \overline{K_4}$. Since G has no isolated vertices, there is an edge incident to u in G . Now every edge incident to u in G has edge multiplicity at most $n - 5$ and so G is not $(n - 1)$ -dense. Consequently, $K_n - e$ is the only graph on n vertices that is $(n - 1)^*$ -dense. \square

4. Minimum number of edges in 2^* -dense, 3^* -dense and 4^* -dense connected graphs on n vertices

In Section 3, we characterized all graphs on n vertices that are k^* -dense for $k = 2, k = n - 1$ or $k = n$. We now investigate lower bounds for the number of edges in a graph on n vertices that are k^* -dense for some small values of k . We begin with a definition.

Definition 4. *For $k, n \in \mathbb{N}$, let*

$$e(k, n) = \min\{|E(G)| : G \text{ is a connected graph on } n \text{ vertices that is } k^* \text{-dense}\}.$$

Let us now consider a connected graph G on n vertices. Proposition 3.1 says that G is 2^* -dense if and only if G has an edge that does not belong to a triangle in G . Observe that any connected graph is 2-dense and that trees have the minimum number of edges among all connected graphs on n vertices. Since trees do not have triangles, we have the following result.

Theorem 4.1. *For all $n \in \mathbb{N}$ with $n \geq 2$, the minimum number of edges in a connected 2^* -dense graph is $e(2, n) = n - 1$, and this is achieved by any tree on n -vertices.*

The next result tells us what $e(3, n)$ is for all $n \geq 3$.

Theorem 4.2. For all $n \in \mathbb{N}$ with $n \geq 3$, the minimum number of edges in a connected 3^* -dense graph is $e(3, n) = \lceil \frac{3}{2}(n - 1) \rceil$.

Proof. We consider two cases according to whether n is odd or even.

Case 1 : n is odd.

Let $n = 2\ell + 1$ where $\ell \in \mathbb{N}$. We need to show that $e(3, 2\ell + 1) = 3\ell$. Consider ℓ copies of K_3 all identified at one vertex w . Since this is a 3^* -dense graph with n vertices and 3ℓ edges, we know $e(3, 2\ell + 1) \geq 3\ell$.

By Proposition 3.1, the result is true when $\ell = 1$. Let us assume that $e(3, 2r + 1) = 3r$ for some $r \in \mathbb{N}$ with $r \geq 1$. We need to show that $e(3, 2r + 3) = 3(r + 1) = 3r + 3$. Assume to the contrary that $e(3, 2r + 3) \leq 3r + 2$. Let G be a connected graph with $2r + 3$ vertices that is 3^* -dense such that G has $e(3, 2r + 3)$ edges.

We first show that G has at least two vertices of degree 2. Recall that by Proposition 2.2, every vertex in G has degree at least 2. If G has at most one vertex of degree 2, then $e(3, 2r + 3) = |E(G)| \geq \frac{1}{2}(2 + 3(2r + 2)) = 3r + 4$. But this contradicts our assumption that $e(3, 2r + 3) \leq 3r + 2$.

Now let $x, y \in V(G)$ such that $\deg_G x = 2 = \deg_G y$. Let us first assume that $xy \in E(G)$. It follows from Definition 1 that $|N(x) \cap N(y)| = 1$, say $N(x) \cap N(y) = \{u\}$. Since $\deg_G x = 2 = \deg_G y$, we know that each of the edges xy, xu , and yu has edge multiplicity 1 in G (that is, x and y belong to only one triangle in G , namely $G[x, y, u]$). This means that $G - \{x, y\}$ is a connected 3^* -dense graph with $2r + 1$ vertices and $e(3, 2r + 3) - 3$ edges. From our assumption that $e(3, 2r + 3) \leq 3r + 2$ it follows that $G - \{x, y\}$ has at most $(3r + 2) - 3 = 3r - 1$ edges. But this means that $e(3, 2r + 1) \leq 3r - 1$ which contradicts our inductive assumption that $e(3, 2r + 1) = 3r$.

Let us now assume there exists no vertices x and y which are adjacent in G if $\deg_G x = 2 = \deg_G y$. Observe that $G - \{x, y\}$ is a connected graph with $2r + 1$ vertices that has at most $(3r + 2) - 4 = 3r - 2$ edges. Since $G - \{x, y\}$ has less than $e(3, 2r + 1)$ edges, it follows that $G - \{x, y\}$ is not 3^* -dense.

Since G is 3^* -dense, each of x and y must be in exactly one triangle, say x, u and v form a triangle and y, w , and z form a triangle. Furthermore, since $G - \{x, y\}$ is not 3^* -dense, either uv or wz has edge multiplicity 0 in $G - \{x, y\}$, say without loss of generality that uv has edge multiplicity 0 in $G - \{x, y\}$. Let H be the graph obtained from $G - x$ by deleting the edge uv and then identifying the vertices u and v . Since uv had edge multiplicity 0, this will not affect the multiplicity of any other edge in $G - x$. Note that, by construction, H is a 3^* -dense graph on $2r + 1$ vertices that has $e(3, 2r + 3) - 3$ edges. This means that H has at most $(3r + 2) - 3 = 3r - 1$ edges and so $e(3, 2r + 1) \leq 3r - 1$. But again, this contradicts our inductive assumption that $e(3, 2r + 1) = 3r$. We therefore conclude that $e(3, 2r + 3) = 3r + 3$. Consequently, $e(3, 2\ell + 1) = 3\ell$ for all $\ell \in \mathbb{N}$.

Case 2 : n is even.

Let $n = 2\ell$ where $\ell \in \mathbb{N}$ with $\ell \geq 2$. We need to show that $e(3, 2\ell) = 3\ell - 1$. Consider $\ell - 2$ copies of K_3 and one copy of $K_4 - e$ all identified at one vertex w . This is a 3^* -dense graph with n vertices and $3\ell - 1$ edges, so we have $e(3, 2\ell) \geq 3\ell - 1$.

By Proposition 3.3, the result is true when $\ell = 2$. Let us assume that $e(3, 2r) = 3r - 1$ for some $r \in \mathbb{N}$ with $r \geq 2$. We need to show that $e(3, 2r + 2) = 3(r + 1) - 1 = 3r + 2$. Assume to the contrary that $e(3, 2r + 2) \leq 3r + 1$. Let G be a connected graph with $2r + 2$ vertices that is 3^* -dense such that G has $e(3, 2r + 2)$ edges. We first show that G has at least one vertex of degree

2. If $\deg v \geq 3$ for all $v \in V(G)$, then $e(3, 2r + 2) = |E(G)| = \frac{1}{2}(3(2r + 2)) \geq 3r + 3$. But this contradicts our assumption that $e(3, 2r + 2) \leq 3r + 1$. Let $x \in V(G)$ such that $\deg x = 2$. Consider the graph $G - x$. Observe that $G - x$ is a connected graph with $2r + 1$ vertices and $e(3, 2r + 2) - 2$ edges. Thus, $G - x$ has at most $(3r + 1) - 2 = 3r - 1$ edges. Since $e(3, 2r + 1) = 3r$ from Case 1 above, we know that $G - x$ is not 3^* -dense. This means that there exist vertices w and z such that $G[w, x, z]$ is a triangle in G such that the edge wz has edge multiplicity 0 in $G - \{x\}$ (that is, $N_{G-\{x\}}(w) \cap N_{G-\{x\}}(z) = \emptyset$). Let H be the graph obtained from $G - \{x\}$ by deleting the edge wz and then identifying the vertices w and z . Note that, by construction, H is a 3^* -dense graph on $2r$ vertices that has $e(3, 2r + 2) - 3$ edges. This means that H has at most $(3r + 1) - 3 = 3r - 2$ edges and so $e(3, 2r + 1) \leq 3r - 2$. But this contradicts our inductive assumption that $e(3, 2r) = 3r - 1$. We therefore conclude that $e(3, 2r + 2) = 3r + 2$. Consequently, $e(3, 2\ell) = 3\ell - 1$ for all $\ell \in \mathbb{N}$ with $\ell \geq 2$. \square

We now determine $e(4, n)$ for all $n \geq 4$.

Theorem 4.3. For every positive integer $n \geq 4$, the minimum number of edges in a connected 4^* -dense graph is

$$e(4, n) = \begin{cases} 2n - 2 & \text{if } n \equiv 1 \pmod{3} \\ 2n - 1 & \text{otherwise.} \end{cases}$$

Proof. For convenience in the proof, let

$$F(n) = \begin{cases} 2n - 2 & \text{if } n \equiv 1 \pmod{3} \\ 2n - 1 & \text{otherwise.} \end{cases},$$

so that we wish to prove that $e(4, n) = F(n)$. Notice that $F(n - 3) = F(n) - 6$ for all n .

First, we show that $e(4, n) \leq F(n)$ using an inductive construction. For $n = 4$, let $G = K_4$. For $n = 5$, let $G = K_5 - e$. For $n = 6$, we have the graph $G = K_2 + 2K_2$ that is a 4^* -dense graph with 6 vertices and 11 edges. Suppose we have a graph G with n vertices and $F(n)$ edges. Associate one vertex of the complete graph K_4 with a vertex of G to obtain a 4^* -dense graph with $n + 3$ vertices and $F(n) + 6 = F(n + 3)$ edges.

Next, we must show that $e(4, n) \geq F(n)$. We proceed by induction on n . For $n = 4$ and $n = 5$, the result follows from Propositions 3.1 and 3.3. Suppose we have a 4^* -dense graph on 6 vertices. From Proposition 2.2, we know that every vertex has degree at least 3, and furthermore, any vertex with degree 3 is in a subgraph isomorphic to K_4 . If every vertex has degree 4 or more, then the graph has 12 edges. Suppose there is a vertex u with $\deg u = 3$. Then u lies in a subgraph isomorphic to K_4 , so there are vertices v, w and x such that u, v, w and x induce a complete subgraph, with 6 edges. The remaining two vertices in the graph, say y and z , each have degree at least 3. If yz is an edge, then there must be at least 4 more edges between $\{u, v, w, x\}$ and $\{y, z\}$, for a total of $6 + 1 + 4 = 11$ edges. If yz is not an edge, then there are at least 6 edges between $\{u, v, w, x\}$ and $\{y, z\}$, for a total of at least $6 + 6 = 12$ edges. Thus, a connected 4^* -dense graph on 6 vertices must have at least 11 edges.

Consider an integer $n \geq 7$. Assume, for all k such that $4 \leq k < n$, we have $e(4, k) = F(k)$. We wish to show that $e(4, n) = F(n)$. Suppose, to the contrary, that there is a connected 4^* -dense graph G with n vertices and at most $F(n) - 1$ edges. Thus, G has at most $2n - 2$ edges. From

Proposition 2.2, the minimum degree in G is at least 3. If every vertex in G has degree at least 4, then G would have at least $\frac{4n}{2} = 2n$ edges, which is a contradiction. Thus, there is some vertex u in G such that $\deg u = 3$. Let $N(u) = \{v, w, x\}$. From Proposition 2.2, we know that $\{u, v, w, x\}$ induces a complete subgraph.

Notice that in $G - u$, each of the edges vw , wx , and xv is in the triangle v, w, x , and hence has multiplicity at least 1. If all three edges have edge multiplicities equal to 2, then $G - u$ is 4^* -dense.
Case 1 : $G - u$ is 4^ -dense.*

In this case, $G - u$ has $n - 1$ vertices and at most $(F(n) - 1) - 3 = F(n) - 4$ edges. If $n \equiv 0 \pmod 3$, then $G - u$ has at most $(2n - 1) - 4 = 2n - 5$ edges. Since $n - 1 \equiv 2 \pmod 3$, we have $F(n - 1) = 2(n - 1) - 1 = 2n - 3$. If $n \equiv 1 \pmod 3$, then $G - u$ has at most $(2n - 2) - 4 = 2n - 6$ edges. Since $n - 1 \equiv 0 \pmod 3$, we have $F(n - 1) = 2(n - 1) - 1 = 2n - 3$. If $n \equiv 2 \pmod 3$, then $G - u$ has at most $(2n - 1) - 4 = 2n - 5$ edges. Since $n - 1 \equiv 1 \pmod 3$, we have $F(n - 1) = 2(n - 1) - 2 = 2n - 4$. In each case, $G - u$ has fewer than $F(n - 1)$ edges, which contradicts the inductive hypothesis that $F(k) = e(4, k)$ for all k with $4 \leq k < n$.

Case 2 : The edges vw , vx , and wx all have edge multiplicity 1 in $G - u$.

Notice that the edge multiplicity 1 in each case comes from the triangle v, w, x . Thus, we can delete the edges vw , vx , and wx without changing the edge multiplicity of any of the other edges in $G - u$ and then identify all three vertices to obtain a new graph H with $n - 3$ vertices and at most $(F(n) - 1) - 6 = F(n) - 7$ edges. Notice that H is connected and 4^* -dense by construction. Since $F(n - 3) = F(n) - 6$ for all $n \geq 7$, the graph H has order $n - 3$ and fewer than $F(n - 3)$ edges, which contradicts the inductive hypothesis.

Case 3 : Exactly one of the edges vw , vx , and wx , say vx , has edge multiplicity 1 in $G - u$.

We may assume that vw and wx each have edge multiplicity at least 2 in $G - u$. Thus, v and w have at least one common neighbor, say y , other than x , and w and x have at least one common neighbor, say z , other than v . Notice that if $y = z$, then y is a common neighbor of v and x , which contradicts our assumption that vx has edge multiplicity 1 (that w is the only common neighbor of v and x).

We may delete the edge vx from $G - u$ without changing the edge multiplicity of any other edge except for vw and wx , and then identify the vertices v and x , merging the two edges vw and wx into one new edge. The endpoints of this new edge have at least two common neighbors, y and z . Thus, the resulting graph H is 4-dense and connected, with $n - 2$ vertices and at most $(F(n) - 1) - 5 = F(n) - 6$ edges.

If $n \equiv 0 \pmod 3$, then $F(n) - 6 = (2n - 1) - 6 = 2n - 7$. Since $n - 2 \equiv 1 \pmod 3$, we have $F(n - 2) = 2(n - 2) - 2 = 2n - 6$. If $n \equiv 1 \pmod 3$, then $F(n) - 6 = (2n - 2) - 6 = 2n - 8$. Since $n - 2 \equiv 2 \pmod 3$, we have $F(n - 2) = 2(n - 2) - 1 = 2n - 5$. If $n \equiv 2 \pmod 3$, then $F(n) - 6 = (2n - 1) - 6 = 2n - 7$. Since $n - 2 \equiv 0 \pmod 3$, we have $F(n - 2) = 2(n - 2) - 1 = 2n - 5$. In each case, we have a connected 4^* -dense graph with $n - 2$ vertices and fewer than $F(n - 2)$ edges, which contradicts the inductive hypothesis.

Case 4 : Exactly two of the edges vw , vx , and wx , say vx and vw , have edge multiplicity 1 in $G - u$.

We may delete vw and vx in $G - u$ without changing the edge multiplicity of any other edge in the graph except for wx , and then identify v with w , to ensure that the resulting graph H is connected. The edge multiplicity of wx in H is at least 1. If the edge multiplicity of wx is at least

2, then H is a connected 4^* -dense graph with $n - 2$ vertices and at most $(F(n) - 1) - 5 = F(n) - 6$ edges. This contradicts the inductive hypothesis, as in Case 3.

We may assume that wx has edge multiplicity 1 in H . Let y be the unique common neighbor of w and x in H . Since yw and yx each have multiplicity at least 2 in G and in H , we know that x and y have a common neighbor z with $z \neq w$, and w and y have a common neighbor z' such that $z' \neq x$. If $z = z'$, then w and x have common neighbors y and z , which contradicts our assumption that wx has multiplicity 1. If $z \neq z'$, then we will delete the edge wx and identify the vertices w and x , merging the edges xy and wy into a single new edge. The endpoints of this new edge have common neighbors z and z' , so the new edge has multiplicity at least 2. We also add the edge zz' . Notice that z and z' have at least two common neighbors, y and the vertex formed from identifying w and x , so zz' has edge multiplicity at least 2. The edge multiplicity of the remaining edges is not decreased, so the new graph J has $n - 3$ vertices and at most $(F(n) - 1) - 7 + 1 = F(n) - 7$ edges. We can readily check that $F(n - 3) = F(n) - 6$ for all n , so this contradicts our inductive hypothesis. \square

Notice that in Theorem 4.1, Theorem 4.2, and Theorem 4.3,

$$e(k, n) = q \binom{k}{2} + \binom{r}{2} + r(k - r),$$

where q and r are the unique integers such that $n - 1 = q(k - 1) + r$ and $0 \leq r < k - 1$. The next section shows that this need not be the case for $k \geq 8$. We thus have the following conjecture.

Conjecture 1. *Let G be a k -dense connected graph with n vertices, where $k \leq 7$. If q and r are the unique integers such that $(n - 1) = q(k - 1) + r$ and $0 \leq r < k - 1$, then*

$$e(k, n) = q \binom{k}{2} + \binom{r}{2} + r(k - r).$$

5. Upper bounds on the minimum number of edges in k^* -dense graphs on n vertices

In Section 3, we characterized all graphs on n vertices that are 2^* -dense, $(n - 1)^*$ -dense, or n^* -dense, while in Section 4, we found the minimum number of edges in connected graphs on n vertices that are 2^* -dense, 3^* -dense or 4^* -dense. In this section, we present a result that gives us an upper bound on the minimum number of edges in graphs (not necessarily connected) on n vertices that are k^* -dense, for all k such that $2 \leq k \leq n$.

Proposition 5.1. *Let k and n be integers such that $2 \leq k \leq n$. If q and r are the unique integers such that $n = kq + r$ and $0 \leq r < k$, then there exists a k^* -dense graph on n vertices that has*

$$q \binom{k}{2} + \binom{r}{2} + r(k - r)$$

edges.

Proof. The graph G ($G = qK_k$ if $r = 0$ and $G = (q - 1)K_k \cup K_k + K_r$ if $0 < r < k$) given in the proof of Proposition 3.2 is a k^* -dense graph on n vertices that has $q\binom{k}{2} + \binom{r}{2} + r(k - r)$ edges. \square

In Proposition 5.1, we did not require the k^* -dense graph to be connected. If we consider connected graphs only, then we have the following.

Proposition 5.2. *Let k and n be integers such that $2 \leq k \leq n$. If q and r are the unique integers such that $n - 1 = (k - 1)q + r$ and $0 \leq r < k - 1$, then there exists a k^* -dense connected graph on n vertices that has*

$$q\binom{k}{2} + \binom{r}{2} + r(k - r)$$

edges.

Proof. Let $H_1 = K_k + K_r$ and for $i = 2, 3, \dots, q$, let $H_i = K_k$. For $i = 1, 2, \dots, q$, let v_i be a vertex of H_i . If G is the graph obtained by identifying the vertices v_1, v_2, \dots, v_q , then G is a k^* -dense connected graph on n vertices that has $q\binom{k}{2} + \binom{r}{2} + r(k - r)$ edges. \square

It follows from Proposition 5.2 that, in general, $e(k, n) \leq q\binom{k}{2} + \binom{r}{2} + r(k - r)$, where k, n, q , and r are as given above. However, for certain values of k and n , we can improve this upper bound for $e(k, n)$ as we now show.

Observation 2. *Let*

$$H = K_{n-k} \cup \left(\frac{k}{2}\right) K_2$$

if k is even and let

$$H = K_{n-k} \cup \left(\frac{k-1}{2}\right) K_2 \cup K_1$$

if k is odd, where $n - k \geq 2$. The graph $G = \overline{H}$ is a connected graph on n vertices that is k^* -dense.

Let us consider particular examples of the graphs described in Observation 2. If $n = 26$ and $k = 23$, then $H = K_3 \cup 11K_2 \cup K_1$; that is H is the union of one triangle, 11 independent edges, and an isolated vertex and the graph G is the complement of H . In this construction, G is a connected graph with 26 vertices that is 23^* -dense and has

$$\binom{26}{2} - 3 - 11 = 325 - 13 = 311$$

edges. Note that $25 = 1 \cdot 22 + 3$. Hence, the construction of Proposition 5.2 gives us a connected graph with 26 vertices that is 23^* -dense and has

$$1 \cdot \binom{23}{2} + \binom{3}{2} + 3 \cdot (23 - 3) = 316$$

edges. As another example, if $n = 26$ and $k = 24$, then $H = 13K_2$; that is H is the union of 13 independent edges and the graph G is the complement of H . In this construction, G is a connected graph with 26 vertices that is 24^* -dense and has

$$\binom{26}{2} - 13 = 325 - 13 = 312$$

edges. Note that $25 = 1 \cdot 23 + 2$. Hence, the construction of Proposition 5.2 gives us a connected graph with 26 that is 24^* -dense and has

$$1 \cdot \binom{24}{2} + \binom{2}{2} + 2 \cdot (24 - 2) = 321$$

edges.

For many values of k and n , the upper bound for $e(k, n)$ given by Proposition 5.2 is better than the one given by Observation 2 (say $n = 10, k = 6$ or $n = 10, k = 7$). But the examples above show that the opposite is true for some values of k and n ($n = 26, k = 23$ or $n = 26, k = 24$).

6. Maximum number of edges in k^* -dense connected graphs on n vertices

In this section, we determine the maximum number of edges in a k^* -dense connected graph on n vertices. We start by giving a definition that is analogous to $e(k, n)$.

Definition 5. For $k, n \in \mathbb{N}$, let

$$E(k, n) = \max\{|E(G)| : G \text{ is a connected graph on } n \text{ vertices that is } k^* \text{-dense}\}.$$

Theorem 6.1. If k and n are integers such that $2 \leq k \leq n$, then $E(k, n) = n + k - 3 + \binom{n-2}{2}$.

Proof. Suppose first that $n \geq k + 2$ and $k \geq 2$ and let G be a graph on n vertices with $n + k - 2 + \binom{n-2}{2}$ edges. Let u and v be any two adjacent vertices in G . Note that there are at most $\binom{n-2}{2}$ edges that are not incident with either u or v (consider the edges in the subgraph induced by the vertex set $V(G) - \{u, v\}$.) Thus, there are at least $n + k - 3$ edges between u and v and the vertex set $V(G) - \{u, v\}$. By the pigeonhole principle, u and v must have at least $k - 1$ common neighbors. Since this is true for any adjacent pair u and v , it follows that G is $(k + 1)$ -dense. This means that G is not k^* -dense. It follows that $E(k, n) \leq n + k - 3 + \binom{n-2}{2}$.

To see that this is sharp, consider the graph G formed by starting with a complete graph K_{n-2} . Now add two vertices u and v , with an edge between u and v . Partition the vertices of the K_{n-2} into three sets A, B , and C such that $|A| = \lfloor \frac{n-k}{2} \rfloor, |B| = k-2$ and $|C| = \lceil \frac{n-k}{2} \rceil$. Join u to every vertex in A and B and join v to every vertex in B and C . We can check that G has exactly $n + k - 3 + \binom{n-2}{2}$ edges. Since u and v are adjacent and have exactly $k - 2$ common neighbors, specifically the vertices of B , we know that G is not $(k + 1)$ -dense. Now, any two vertices distinct from u and v are in the subgraph K_{n-2} , so they have at least $n - 4 \geq k - 2$ common neighbors. The vertices u and $a \in N_G(u)$, where $a \neq v$, have at least $\lfloor \frac{n-k}{2} \rfloor + k - 3 \geq \frac{2}{2} + k - 3 = k - 2$

common neighbors (similarly for the vertices v and b where $b \in N_G(v)$ and $b \neq u$). Hence, G is k -dense. Consequently, G is k^* -dense.

Now, if $n = k$, then $n + k - 3 + \binom{n-2}{2} = \binom{k}{2}$. We know that the only k^* -dense graph on k vertices is K_k by Proposition 3.1. If $n = k + 1$, then $n + k - 3 + \binom{n-2}{2} = \binom{k+1}{2} - 1$ and by Proposition 3.3, the only k^* -dense graph on $k + 1$ vertices is $K_{k+1} - e$. \square

7. Complete Results for 2*-dense, 3*-dense and 4*-dense Connected Graphs

By Theorem 4.1 and Theorem 6.1, we have the bounds in the following corollary.

Corollary 7.1. *Let n be an integer with $n \geq 4$. If G is 2*-dense graph with n vertices and m edges, then*

$$n - 1 \leq m \leq \binom{n-2}{2} + n - 1. \tag{1}$$

Moreover, for every integer a with $n - 1 \leq a \leq \binom{n-2}{2} + n - 1$, there is a 2*-dense graph G with n vertices and a edges.

Proof. Equation (1) follows from Theorem 4.1 and Theorem 6.1.

For the realization result, let u_1 and u_2 be two adjacent vertices of a tree on n vertices, such that $\deg u_1 = 1$, then add $a - n + 1$ more edges not incident to u_1 . Since $m(u_1, u_2) = 0$, this graph is 2-dense. \square

By Theorem 4.2 and Theorem 6.1, we have the bounds in the following corollary.

Corollary 7.2. *Let n be an integer with $n \geq 4$. If G is 3*-dense graph with n vertices and m edges, then*

$$\frac{3}{2}(n - 1) \leq m \leq \binom{n-2}{2} + n. \tag{2}$$

Moreover, for every integer a with $\frac{3}{2}(n - 1) \leq a \leq \binom{n-2}{2} + n$, there is a 3*-dense graph G with n vertices and a edges.

Proof. Equation (2) follows from Theorems 4.2 and 6.1.

For the realization result, let us first consider the case where n is odd. Then the graph G obtained from

1. $\frac{n-1}{2}$ copies of K_3 all identified at one vertex w (so that $\deg w = n - 1$ and all other vertices have degree 2),
2. if u_1 and u_2 are two adjacent vertices of degree 2, then add $a - \frac{3(n-1)}{2}$ more edges such that $N(u_1) \cap N(u_2) = \{w\}$.

Since $m(u_1, u_2) = 1$, it follows that G is an n -vertex graph that is 3*-dense, with a edges.

For the case where n is even, use the same construction as above, but replace one of the triangles by a $K_4 - e$, such that $\deg w = n - 1$ is still valid. \square

Similarly, by Theorem 4.3 and Theorem 6.1, we have the bounds in the following corollary.

Corollary 7.3. *Let n be an integer with $n \geq 4$. If G is 4^* -dense graph with n vertices and m edges, then:*

$$\begin{cases} 2n - 2 & \text{if } n \equiv 1 \pmod{3} \\ 2n - 1 & \text{otherwise} \end{cases} \leq m \leq n + 1 + \binom{n-2}{2}. \quad (3)$$

Moreover, for all values a between the lower and upper bounds of Equation 3, there is a 4^* -dense graph G with n vertices and a edges.

Proof. Equation (3) follows from Theorems 4.3 and 6.1.

For the realization result, let us first consider the case where $n \equiv 1 \pmod{3}$. Then the graph G is obtained from

1. $\frac{n-1}{3}$ copies of K_4 all identified at one vertex w (so that $\deg w = n - 1$ and all other vertices have degree 3),
2. if u_1, u_2 and u_3 are three mutually adjacent vertices of degree 3, then add $a - \binom{4}{2} \frac{(n-1)}{3}$ more edges such that $N(u_1) \cap N(u_2) = \{u_3, w\}$.

Since $m(u_1, u_2) = 2$, it follows that G is an n -vertex graph that is 4^* -dense, with a edges. If $n \equiv 2 \pmod{3}$, use the same construction as above, but replace one of the copies of K_4 by a $K_5 - e$ such that $\deg w = n - 1$ is still valid. If $n \equiv 0 \pmod{3}$, use the same construction as above, but replace two of the copies of K_4 by two copies of $K_5 - e$ such that $\deg w = n - 1$ is still valid. □

8. Conclusion

In this paper, we introduced the topic of k -dense to the graph theory community, and we studied the minimum number of edges, $e(k, n)$, and the maximum number of edges, $E(k, n)$, that a connected k^* -dense (k -dense but not $(k + 1)$ -dense) graph of order n can have. We established a formula for $E(k, n)$ and found an upper bound for $e(k, n)$ for all $2 \leq k \leq n$ that is sharp for small values of n . For small values of k (namely 2, 3, 4), we showed that there exists a k^* -dense connected graph G on n vertices that has m edges for each m satisfying $e(k, n) \leq m \leq E(k, n)$. Using a computer to assist us in our investigation (see Appendix), we believe that the formula for $e(k, n)$ that we found for $k = 2, 3, 4$ generalizes to all $k \leq 7$. However, for $k \geq 8$ the formula for $e(k, n)$ that we found does not hold. We conclude the paper with the following open problem.

Problem 1. *Let k and n be integers such that $5 \leq k \leq n$. Determine $e(k, n)$; that is, find the minimum number of edges that a graph G on n vertices can have if G is k^* -dense.*

Acknowledgement

The authors would like to thank the referees for their comments and suggestions.

References

- [1] G. Chartrand, L. Lesniak and P. Zhang, *Graphs and Digraphs* (5th Edition), CRC Press, Taylor & Francis Group (2010).
- [2] R. Karp, Reducibility among combinatorial problems, In Raymond Miller and James Thatcher, editors, *Complexity of Computer Computations*, (1972), 85–103.
- [3] B. D. McKay and A. Piperno, Practical Graph Isomorphism, II, *J. Symbolic Computation* **60**, (2013), 94–112. <http://dx.doi.org/10.1016/j.jsc.2013.09.003>. Preprint version at arxiv.org.
- [4] K.Saito, T. Yamada and K. Kazama, Extracting Communities from Complex Networks by the k -Dense Method, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **91** (11), (2008), 3304–3311.
- [5] W. A. Stein et al., *Sage Mathematics Software (Version 6.2)*, The Sage Development Team, 2014, <http://www.sagemath.org>.
- [6] V. Zlatic, D. Garlaschelli and G. Caldarelli, Complex Networks with Arbitrary Edge Multiplicities, *EuroPhysics Letters* **97** (2), (2011).

Appendix: Sage Code

For checking graphs with $k \geq 5$, we used Sage[5]. Specifically, we generated graphs within the range of anticipated edge-counts for a given n , then tested each graph using the function found in Figure 2.

```
def is_kdense(g,k):
    #Function tests a graph (g) for density (k).
    #Returns true if the graph is k-dense, and false if it is not k-dense
    k_dense=True
    #Check two basic k-dense properties
    if min(g.degree())<k-2:
        k_dense=False
    if g.size() < floor(g.order()*(k-1)/2):
        k_dense=False
    if k_dense:
        #Iterate through all the edges in the graph
        for i in list(g.edge_iterator(labels=False)):
            #A temporary neighbor list from node start
            tmpset=set(g.neighbors(i[0]))
            #Intersect with neighbor list from node end
            common=tmpset.intersection(g.neighbors(i[1]))
            #If any edge fails ..
            if len(common) < k-2:
                k_dense=False
                break
    return k_dense
```

Figure 2. Sage code to test a graph for k -density

The actual graphs we tested were generated using the ‘Nauty Geng’ package created by McKay and Piperno [3], which is included in Sage by permission. A sample of our code for testing $n = 9$ is given in Figure 3.

```
#generator for n=9
n9graphs=set()
for g in graphs.nauty_geng("9 -c 27:32 -d6"):
    gtmp=g.copy(immutable=True)
    if is_kdense(gtmp,7):
        n9graphs.add(gtmp)
```

Figure 3. Sample Sage code for testing graphs of $n = 9$ for density of $k = 7$.

Note that while the sample code given here works, it has not been optimized or carefully coded for efficiency.