Embeddings of weighted graphs in Erdős-type settings

David M. Soukup

April 26, 2018

Abstract

Many recent results in combinatorics concern the relationship between the size of a set and the number of distances determined by pairs of points in the set. One extension of this question considers configurations within the set with a specified pattern of distances. In this paper, we use graph-theoretic methods to prove that a sufficiently large set E must contain at least $C_G|E|$ distinct copies of any given weighted tree G, where C_G is a constant depending only on the graph G.

1 Introduction

Many questions in combinatorics involve the behavior of the distance set $\Delta(E)$ of a set E, defined as $\Delta(E) = \{d(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in E\}$ for some distance function d. For instance, Erdős' celebrated distinct distances problem conjectured that for finite sets $E \subset \mathbb{R}^2$, $|\Delta(E)| \gtrsim |E|^{1-\epsilon}$ for any positive ϵ. This conjecture was proven in this form by Guth and Katz in [3]. Distance problems where the ambient space is a finite vector space have also been a subject of much research [2, 4, 5, 7].

A natural question follows: under what conditions on E can we find not just pairs of points a specified distance apart, but groups of points with some specfied pattern of distances? In other words, given some weighted graph G where the edge weights correspond to distances between points, when can we find a "copy" of G inside E? Bennett, Chapman, Covert, Hart, Iosevich, and Pakianathan proved a result in this direction for the Euclidean distance in $\mathbb{Z}_p^d$ and path graphs in [1], and McDonald gives a similar result for dot products in $\mathbb{Z}_p^d$ in [6]. In this paper, we give an answer to this question for wide classes of graphs, distances, and ambient sets. Moreover, we show that these questions, and other similar ones, are part of a much more general framework.

Definition 1.1. *Given an ambient set X and some set D of possible distances, a **symmetric distance function** is a function $d : X \times X \rightarrow D$ such that $d(x_1, x_2) = d(x_2, x_1)$ for all $x_1, x_2 \in X$. Such a function is **K-surjective** if for every $E \subset X$ with $|E| \geq K$, the restriction of d to $E \times E$ is surjective. In other words, every set of size at least E determines every distance.*

Some motivation for this definition is provided by the following theorem, which will allow us to apply the results of this paper to a common setting for distance problems. Throughout, we will use $\mathbb{F}_q$ to refer to the unique finite field with q elements for some prime power q, and we will let $\mathbb{F}_q^d$ be a d-dimensional vector space over this field.

Theorem 1.2 (A. Iosevich and M. Rudnev (2007) [4]). *Let $X = \mathbb{F}_q^d$, $D = \mathbb{F}_q$, and define $d(\{x_i\}, \{y_i\}) = \sum (x_i - y_i)^2$. Then d is K-surjective with $K = Cq^{(d+1)/2}$ for some constant C independent of q.*

Now we just need to define a graph embedding, which is done in the natural way:

Definition 1.3. *Suppose we have a space X, a set of distances D, and a symmetric distance function d. Then for a weighted graph G with edge weights in D, an **embedding** of G into X is an injective function $f : V(G) \to X$ such that for every edge $(v_1, v_2) \in E(G)$ with weight t,*

$$d(f(v_1), f(v_2)) = t$$

*We will typically identify such an embedding with its image in X. A collection of such embeddings $\{f_i\}_{i \in I}$ is **disjoint** if all its images are disjoint subsets of X.*

1.1   Results

The constant in our result will depend on an invariant of the graph G:

Definition 1.4. *Let G be a finite nonempty connected graph; let the degrees of the vertices of G be $d_1, d_2, \ldots, d_n$, ordered so that $d_1 \geq d_2 \geq \cdots \geq d_n$. Then the **stringiness** of G, denoted $\sigma(G)$, is defined to be*

$$\sigma(G) = (d_1 + 1) \prod_{i=2}^{n} d_i.$$

For example, the stringiness of the Petersen graph is $4 \cdot 3^9 = 78732$. The following estimate gives bounds on the stringiness of a tree:

Theorem 1.5. *Let G be a nonempty tree with n edges. Then*

$$n + 1 \leq \sigma(G) \leq 2^n.$$

The lower bound is sharp (the star graph $K_{1,n}$ attains it) while the upper bound is sharp up to a constant (the path graph P_{n+1} has stringiness $3 \cdot 2^{n-2}$).

Now we are ready to state our main theorem:

Theorem 1.6 (Main theorem). *Let X be a set with a symmetric distance function d to a set of distances D, let d be K-surjective, and let $E \subseteq X$ with $|E| = rK$. Then for any weighted tree G with edge weights in D, there exists a disjoint collection A_G of embeddings of G into E with*

$$|A_G| \geq \left(\frac{r}{\sigma(G)} - 1 \right) K$$

Corollary 1.7. *If $|E| \geq 2\sigma(G)K$, then there is at least one embedding of G into E.*

Combining this result with Theorem 1.2 gives the following result in $\mathbb{F}_q^d$:

Corollary 1.8. *Let G be a weighted tree with edge weights in $\mathbb{F}_q$. Then there exists a constant C independent of q and G such that every subset E of $\mathbb{F}_q^d$ such that $|E| \geq C\sigma(G)q^{\frac{d+1}{2}}$ contains an embedding of G.*

We compare this result to the following previous result:

Theorem 1.9 (Bennett, Chapman, Covert, Hart, Iosevich, Pakianathan, [1])**.** *Let G be a weighted path or star graph with edge weights in $\mathbb{F}_q$, and suppose G has k edges. Then there exists a constant C independent of q and G such that every subset E of $\mathbb{F}_q^d$ such that $|E| \geq Ckq^{\frac{d+1}{2}}$ contains an embedding of G.*

The idea of this paper is to show that results of the type of Theorem 1.9 are instances of a more general phenomenon. Corollary 1.8 is an example of this phenomenon applying to the standard distance on $\mathbb{F}_q^d$; it allows extending these results to general trees for which no results existed. We also note that proof of Theorem 1.6 is very modular, so it is possible to (for instance) use Theorem 1.9 to get slightly better bounds for embeddings of more complicated trees in $\mathbb{Z}_q^d$.

We proceed as follows: for illustrative purposes, we first state and prove Theorem 2.2, a weaker version of Theorem 1.6, using Lemma 2.1. We then give the very similar proof of Theorem 1.6, which relies on the analagous Lemma 2.5. Finally we prove Theorem 1.5.

2   Graph embeddings

2.1   An easier case of 1.6

The proof of 2.2 is by induction; it is convenient to state the base case as a separate lemma.

Lemma 2.1. *Let X be a set with a symmetric distance function d to a set of distances D, let d be K-surjective, and let $E \subseteq X$ with $|E| = rK$. Then for any fixed $t \in D$, there are at least $\frac{r-1}{2}K$ disjoint pairs $\{e_i, f_i\}_{i \in I}$ in E such that $d(e_i, f_i) = t$ for all $i \in I$.*

Proof. Since E is finite, let I be maximal. Let F be the union of all the pairs. Then by maximality of I, $E - F$ cannot contain any pair of points with distance t. By K-surjectivity, this means $|E - F| < K$; so $|F| \geq (r-1)K$, meaning $|I| = |F|/2 \geq \frac{r-1}{2}K$ as required. $\qquad\square$

Theorem 2.2. *Let X be a set with a symmetric distance function d to a set of distances D, let d be K-surjective, and let $E \subseteq X$ with $|E| = rK$. Then for any weighted nonempty tree G with edge weights in D, suppose G has n edges. Then there exists a disjoint collection A_G of embeddings of G into E with*

$$|A_G| \geq \left(\frac{r+1}{2^n} - 1 \right) K.$$

Proof. We proceed by induction on n. The case $n = 0$ is tautological, and the case $n = 1$ is equivalent to Lemma 2.1.

So assume $n \geq 2$, which means G must have a leaf. Fix any leaf; call it v, its associated edge e, the unique vertex adjacent to it w, and the edge weight t. Then $G - v$ is a tree with strictly fewer edges; so by the inductive hypothesis we have a disjoint collection A_{G-v} of embeddings of $G - v$ into E with

$$|A_{G-v}| \geq \left(\frac{r+1}{2^{n-1}} - 1 \right) K.$$

Consider the set $W = \{f(w) | f \in A_{G-v}\}$. By Lemma 2.1, there exist at least $\frac{|W|/K-1}{2}$ disjoint pairs of points $(\{f(w), f'(w)\}$ in W with $d(f(w), f'(w)) = t$. But for each of these pairs we can consider the function $g : V(G) \to E$ given by

$$g(x) = \begin{cases} f'(w) & x = v \\ f(x) & x \neq v \end{cases}$$

By construction, these are disjoint embeddings of G into E, and there are at least $\frac{|W|/K-1}{2} K$ of them; but by disjointness of A_{G-v}, $|W| = |A_{G-v}|$ so there are at least

$$\frac{|A_{G-v}|/K - 1}{2} K \geq \frac{\left(\frac{r+1}{2^{n-1}} - 1 \right) - 1}{2} K = \left(\frac{r+1}{2^n} - 1 \right) K$$

disjoint embeddings of G as required.

$\square$

Corollary 2.3. *If $|E| \geq 2^{n+1} K$, there is at least one embedding of G into E.*

Note that this proof would have worked equally well even if d were not symmetric, which will not carry over to Theorem 1.6.

2.2 Proof of Theorem 1.6

Analagously to the proof of Theorem 2.2, we will state a governing lemma (Lemma 2.5); the structure will be identical except that we are building our graph G out of stars instead of working purely with edges. For technical reasons, the application of the K-surjectivity assumption is more difficult in this case, so we will first state an auxilliary lemma, Lemma 2.4.

Lemma 2.4. *Let X be a set with a symmetric distance function d to a set of distances D, let d be K-surjective, and let $E \subseteq X$ with $|E| = rK$. Then for any fixed $t \in D$, $s \in \mathbb{N}$, there are at most sK points $e \in E$ for which there are not at least s other distinct points $e_1, e_2, \ldots, e_s \in E$ such that $d(e, e_i) = t$ for all i.*

Proof. Create a graph H whose vertices are the points of E and where two vertices are connected by an edge if and only if the corresponding points of E are distance t. Then consider the subgraph H^* of H generated by only

4

those vertices of degree $< s$. By construction, the maximum degree of vertices in H is less than s, which means H^* can be s-colored, that is, partitioned into s independent sets. Since the K-surjectivity condition guarantees that an independent set in H (and thus in H^*) has size $< K$, it follows that $|H^*| < sK$. This means that at most sK of the vertices of H have degree $< s$, proving the lemma. $\qquad\square$

Lemma 2.5. *Let X be a set with a symmetric distance function d to a set of distances D, let d be K-surjective, and let $E \subseteq X$ with $|E| = rK$. Then for any weighted nonempty star graph $G \cong K_{1,n}$ with edge weights in D, E contains at least $\frac{r-n}{n+1}K$ disjoint embeddings of G.*

Proof. As in the proof of Lemma 2.1, let I be a maximal set of embeddings of G into E, and define F to be the union of all the embeddings contained in I. Then $E - F$ must have no embeddings of G.

Suppose the set of edge weights of G is $\{t_1, t_2, \ldots, t_a\}$ appearing with multiplicities $\{m_1, m_2, \ldots m_a\}$ respectively. Then by Lemma 2.4, for each fixed i there are at most $m_i K$ points of $E - F$ which are not distance t_i from at least m_i other points of $E - F$. Summing over i we get that there are at most

$$\sum_{i=1}^{a} m_i K = nK$$

points of $E - F$ which are not distance t_i from at least m_i other points of $E - F$ for every i. But if $x \in E - F$ is in fact distance t_i from at least m_i other points of $E - F$ for every i, then x forms the root of an embedding of the graph G. Thus $|E - F| \leq nK$; so $|I| = |F|/(n+1) \geq \frac{rK-nK}{n+1} = \frac{r-n}{n+1}K$ as required. $\qquad\square$

Note that when $n = 1$, this is exactly the statement of Lemma 2.1, but when $n \geq 2$ we may have to deal with repeated edge weights, which adds the extra complexity.

Now we are ready to prove the main theorem, Theorem 1.6:

Proof. The proof proceeds by strong induction on the number of edges in G. If G contains no edges, $\sigma(G) = 1$ and the theorem is clearly true; if G is a star graph $K_{1,n}$, then $\sigma(G) = n + 1$ and the theorem is just Lemma 2.5.

So assume G is not a star graph. Then we can always find a vertex $w \in G$ such that:

1. w is not a leaf of G (equivalently, $\deg_G w \geq 2$).

2. All but one of the vertices connected to G are leaves (call these leaves $v_1, v_2, \ldots, v_y$, and the associated distances $t_1, t_2, \ldots, t_y$, where $y = \deg_G w - 1$.)

3. There exists a vertex $v \neq w$ in G such that $\deg_G v \geq \deg_G w$.

To see this, let L be the set of leaves of G. Then since G is a tree which is not a star graph, $G - L$ is a tree which contains at least one edge; any leaf of $G - L$ satisfies conditions 1 and 2, and since $G - L$ has at least two leaves, at least one of these must satisfy condition 3.

Define the graph H to be $G - \{v_1, v_2, \ldots v_y\}$. By conditions 1 and 2, H is a tree with fewer edges than G; by condition 3, $\sigma(H) = \frac{\sigma(G)}{y+1}$ (since we can reorder the product in Definition 1.4 to make v correspond to d_1, and all we lose by deleting these leaves is a factor of $\deg_G w$). By the inductive hypothesis we have a disjoint collection A_H of embeddings of H into E with

$$|A_H| \geq \left(\frac{r}{\sigma(H)} - 1\right) K.$$

Consider the set $W = \{f(w) | f \in A_H\}$. By Lemma 2.5, there exist at least $\frac{|W|/K - n}{n+1}$ disjoint sets of points $\{f(w), f_1(w), \ldots, f_y(w)\}$ contained in W with $d(f(w), f_i(w)) = t_i$ for every i (i.e. embedddings of a particular star graph). But for each of these sets we can consider the function $g : V(G) \to E$ given by

$$g(x) = \begin{cases} f_i(w) & x = v_i \\ f(x) & x \notin \{v_1 \ldots v_y\} \end{cases}$$

By construction, these are disjoint embeddings of G into E, and there are at least $\frac{|W|/K - n}{n+1} K$ of them; but by disjointness of A_H, $|W| = |A_H|$ so there are at least

$$\frac{|A_H|/K - n}{n+1} K \geq \frac{\left(\frac{r}{\sigma(H)} - 1\right) - n}{n+1} K$$
$$= \left(\frac{r}{(n+1) \cdot \sigma(H)} - 1\right) K$$
$$= \left(\frac{r}{\sigma(G)} - 1\right) K.$$

disjoint embeddings of G as required.

$\square$

Note that in view of Theorem 1.5, this is stronger than Theorem 2.2.

2.3   Proof of Theorem 1.5

Proof. For the lower bound on $\sigma(n)$, a simple inductive argument suffices; if $n = 1$ then $\sigma(n) = 2$ since there is only one possible graph with one edge; if $n > 1$ then deleting a leaf must decrease the stringiness since the degree of the other vertex decreases; so for a graph with n edges, $\sigma(G) > \sigma(G - v) \geq n$; so $\sigma(G) \geq n + 1$ (since $\sigma(G) \in \mathbb{Z}$).

For the upper bound, write the degrees of the vertices of G as $d_1, d_2, \ldots d_{n+1}$; without loss of generality say $d_{n+1} = 1$. Then by the arithmetic-geometric mean inequality:

$$((d_1 + 1) \cdot d_2 \cdot d_3 \cdot \cdots \cdot d_n)^{1/n} \leq \frac{d_1 + 1 + d_2 + d_3 + \ldots d_n}{n}$$

$$\sigma(G)^{1/n} \leq \frac{\sum d_i}{n}$$

$$\sigma(G)^{1/n} \leq \frac{2n}{n}$$

$$\sigma(G) \leq 2^n.$$

$\square$

3   Acknowledgements

References

[1] M. Bennett, J. Chapman, D. Covert, D. Hart, A. Iosevich, J. Pakianathan. *Long paths in the distance graph over large subsets of vector spaces over finite fields.* J. Korean Math. Soc. 53:1 (2016). 115-126. arxiv: /abs/1406.0107

[2] J. Bourgain, N. Katz, T. Tao. *A sum-product estimate in finite fields, and applications.* Geom. Func Anal. 14 (2004), 27-57. arXiv: math/0301343

[3] L. Guth, N. H. Katz. *On the Erdős distinct distances problem in the plane.* Annals of Mathematics 181:1 (2015), 155-190. arXiv:1011.4105.

[4] A. Iosevich, M. Rudnev. *Erdős distance problem in vector spaces over finite fields.* Transactions of the American Mathematical Society 359:12 (2007), 6127-6142. arXiv:math/0509005

[5] D. Koh, C.-Y. Shen. *The generalized Erdős-Falconer distance problems in vector spaces over finite fields.* Journal of Number Theory 132:11 (2012), 2455-2473.

[6] B. McDonald. *Hinges in $\mathbb{Z}_p^d$ and applications to pinned distance sets* (Undergraduate thesis). University of Rochester (2016). http://web.math.rochester.edu/people/faculty/iosevich/ mcdonaldhonorsthesis16.pdf

[7] V. Vu. *Sum-product esimates via directed expanders.* Math. Res. Lett. 15:2 (2008), 375-388.