USING GRAPHS AND GAMES TO GENERATE CAP SET BOUNDS

JOSH ABBOTT AND TREVOR MCGUIRE

Abstract. Let \( \mathbb{F}_3 \) be the field with 3 elements and consider the \( k \)-dimensional affine space, \( \mathbb{F}_3^k \), over \( \mathbb{F}_3 \). A line of length 3 in \( \mathbb{F}_3^k \) is a set of 3 \( k \)-tuples, \( \alpha, \beta, \gamma \) such that under componentwise addition, \( \alpha + \beta + \gamma = 0 \). We define a cap set as a subset of \( \mathbb{F}_3^k \) which contains no lines of length 3. We will interpret cap sets in terms of the card game Set. We will define Set on graphs and generate a class of weighted adjacency matrices to produce cap set bounds.

1. Introduction

Set is a card game involving 81 distinct cards. Each card has a variation of the following four characteristics:

(i) **Color**: Each card has red, green or purple

(ii) **Symbol**: Each card contains ovals, squiggles or diamonds

(iii) **Number**: Each card has one, two or three symbols

(iv) **Fill**: Each card is solid, open or striped

**Definition 1.1.** A set consists of a collection of three cards in which each characteristic is either the same on each card or is different on each card. That is to say, any characteristic in the collection of three cards is either common to all three cards, or is different from each card.

Date: April 26, 2008.
We can generalize a game of Set by allowing $k$ characteristics and $m$ values for each characteristic. Then there would be $m^k$ distinct cards in the game and a set would consist of $m$ cards following the prescribed requirements given above.

**Definition 1.2.** A line of length $m$ in $\mathbb{F}_m^k$, where $m$ is prime, is a collection of $m$ $k$-tuples, $\alpha_1, \ldots, \alpha_m$ such that $\sum_{j=1}^{m} \alpha_j = 0$.

**Definition 1.3.** A cap set is a subset of $\mathbb{F}_m^k$ which contains no lines of length $m$ where $m$ is prime.

Note: Unless otherwise stated, we will assume $m = 3$ as in the standard game of Set. Furthermore, we insist that $m$ be prime because, although the following constructions work for all integer values, finite fields only exist for prime power orders. Thus, we will make the constructions for general $m$, and use prime values to address the cap set problem.

2. *Set on Graphs*

**Definition 2.1.** Let $X$ be one of the $k$ characteristics. Let $\mathcal{C}_n^X$ be a simple graph on $n$ vertices. Let $V(\mathcal{C}_n^X)$ be a collection of $n$ cards from a $k$ characteristic game of Set. For $u, v \in V(\mathcal{C}_n^X)$, $uv \in E(\mathcal{C}_n^X)$ if and only if $u, v$ share the same value for the characteristic $X$. We call this an $X$-characteristic graph.
Example 2.1. Consider the following three cards from *Set*

![Image](image_url)

The Color-characteristic graph on these three cards is

![Color-characteristic graph](image_url)

Since all three cards are red, the Color-characteristic graph on these three cards is a $K_3$. We see that the Number-characteristic graph on these cards is the null graph since each card has a different value.

![Number-characteristic graph](image_url)
Theorem 2.1. (Classification of $C_n^X$) If $C_n^X$ is an $X$-characteristic graph, and $m$ is the number of values characteristic $X$ can take, then the following are true:

1. If $P_j \subseteq C_n^X$, where $P_j$ is a path of length $j$, then $\exists K_{j+1}$ such that $P_j \subseteq K_{j+1} \subseteq C_n^X$.
2. $\overline{C_n^X}$, the complement of $C_n^X$, contains no $K_{m+1}$.
3. The components of a $C_n^X$ are complete graphs.
4. $\kappa(C_n^X) \leq m$, where $\kappa$ counts the number of components of a graph.

Proof.

1. Note that $j = 1$ is the trivial case. For $j > 1$ we continue by contradiction. Suppose the statement is false. Then without loss of generality, there exists $v_1, v_2, v_3 \in V(C_n^X)$ such that $v_1v_2, v_2v_3 \in E(C_n^X)$, and $v_1v_3 \notin E(C_n^X)$. This violates the transitive property of equality since $v_1$ and $v_2$ have the same value, as well as $v_2$ and $v_3$, but not $v_1$ and $v_3$. Thus, paths of length two are contained in a $K_3$. This completes the proof since any $P_j \forall j > 2$ contains a $P_2$ subgraph.

2. Assume the statement is false. Then $\exists H \subseteq \overline{C_n^X}$ such that $|V(H)| = m + 1$ and $\forall v_i, v_j \in V(H), v_i \neq v_j$. However, there are only $m$ possible values, thus this violates the pigeon hole principle.

3. The first nontrivial case is $n = 3$. Assume the statement is false. Then there exists a $P_2 \not\subseteq K_3$. This contradicts Part 1.

4. Assume the statement is false. Then we have that $\kappa(C_n^X) > m$.

From part 3, we have that each component is complete. Thus, to
get a complete graph in $\overline{C}_n^X$, we need to choose exactly one vertex from each component of $C_n^X$ and the corresponding edges in $\overline{C}_n^X$.

This, however, is a contradiction, since, by part 2, we have no $K_{m+1} \subseteq \overline{C}_n^X$, thus, we can only have at most $m$ components in $C_n^X$.

□

Definition 2.2. Enumerate the collection of characteristics from 1 to $k$.

A *weighted* $X_i$-characteristic graph, $\mathcal{C}_n^{X_i}$, is a complete graph on $n$ vertices such that $E(\mathcal{C}_n^{X_i}) = E(C_n^X) \cup E(\overline{C}_n^X)$ and let $v_s v_t \in E(\mathcal{C}_n^{X_i})$, we define a binary edge weighting, $\Omega_i$, as follows:

$$\Omega_i(v_s v_t) = \begin{cases} p_{2i-1}, & v_s v_t \in E(\overline{C}_n^X) \\ p_{2i}, & v_s v_t \in E(C_n^X) \end{cases}$$

where $p_j$ is the $j^{th}$ prime.

That is, $\mathcal{C}_n^{X_i}$ is a complete graph on $n$ vertices such that each edge is weighted one of two consecutive primes. We can combine these $k$ distinct *weighted* $X_i$-characteristic graphs by taking the Hadamard (Schur) product of their adjacency matrices. The resulting matrix will be the weighted adjacency matrix of a new complete graph weighted with $2^k$ distinct numbers, fully representing a game of Set.

Definition 2.3. A *Set* graph, $S_n$, is a complete weighted graph, on $n$ vertices, whose weighted adjacency matrix is the Hadamard product of $k$ distinct weighted $X_i$-characteristic graphs.

Remark 2.1. Notice that the Hadamard product preserves equivalued $K_j$’s. That is, if there are $j$ vertices that have equivalued $K_j$’s on each
weighted $X_i$-characteristic graph, then under the Hadamard product, there will be an equivalued $K_j$ in the Set graph. If $j = m$, where $m$ is the number of values a characteristic can assume, then there is a set.

**Example 2.2.** Consider the three cards from a 2 characteristic game of Set.

\[
\begin{array}{c}
\text{0} \\
\text{00} \\
\text{0}
\end{array}
\]

The weighted Fill-characteristic graph and adjacency matrix are:

\[
\begin{pmatrix}
0 & 2 & 3 \\
2 & 0 & 3 \\
3 & 3 & 0
\end{pmatrix}
\]

The weighted Number-characteristic graph and adjacency matrix are:

\[
\begin{pmatrix}
0 & 7 & 5 \\
7 & 0 & 7 \\
5 & 7 & 0
\end{pmatrix}
\]
The Hadamard product is:

\[
\begin{pmatrix}
0 & 2 & 3 \\
2 & 0 & 3 \\
3 & 3 & 0
\end{pmatrix}
\times
\begin{pmatrix}
0 & 7 & 5 \\
7 & 0 & 7 \\
5 & 7 & 0
\end{pmatrix}
= 
\begin{pmatrix}
0 & 14 & 15 \\
14 & 0 & 21 \\
15 & 21 & 0
\end{pmatrix}
\]

Which corresponds to the Set graph

The result shows there is no set on these three cards.

**Lemma 2.1.** In a Set graph, \( S_n \), \( \not\exists v_s v_t \) where 
\[
\prod_{i=1}^{k} \Omega_i(v_s v_t) = \prod_{i=1}^{m} p_{2i-1}
\]
where \( v_s v_t \in E(S_n) \).

**Proof.** Assume the statement is false. Then there is exists two cards, \( v_s, v_t \in V(S_n) \) such that 
\[
\prod_{i=1}^{k} \Omega_i(v_s v_t) = \prod_{i=1}^{m} p_{2i-1}
\]
This means that for each characteristic, \( v_s \) and \( v_t \) have the same value, but this contradicts the distinctness of each card. \(\square\)
Corollary 2.1. In a Set graph, there are at most \(2^k - 1\) possible edge weights, where \(k\) is the number of characteristics in the game.

Proof. Each \(X_1\)-characteristic graph has two weights. In the Hadamard product, this yields \(2^k\) weights. By Lemma 2.1, at least one never occurs.

\[\square\]

3. No Sets and Cap Set Bounds

Definition 3.1. A no set is a collection of cards in which no set exists.

Definition 3.2. The maximal no set size, \(\mu_k\), in any \(k\) characteristic game of Set is the maximum number of cards that can be in play such that no set exists.

Lemma 3.1. In a 2 characteristic game of Set, if either \(\mathcal{C}_n^{X_1}\) has an equivalued \(K_3\) of value \(p_{2i-1}\), then there is a set.

Proof. Let \(p_j\) be a prime. By Lemma 2.1, if either \(\mathcal{C}_n^{X_1}\) has a \(K_3\) of value \(p_{2i-1}\), then the other cannot have any edges with value \(p_{2i}\) between any two of the three points in the \(K_3\). This means there is an equivalued \(K_3\) with edge value \(p_{2i+1}\). Thus, the Set graph has an equivalued \(K_3\). Since there is an equivalued \(K_3\) on the same three points of each characteristic graph, under the Hadamard product, there is an equivalued \(K_3\) in the Set graph. Thus, by Remark 2.1, there is a set.

\[\square\]

Remark 3.1. If we wish to find \(\mu_2\), we need only consider \(X\)-characteristic graphs with components of order 2 or less.
Proposition 3.1. The maximal no set size in a two characteristic game of Set is 4.

Proof. Lemma 3.1 and Remark 3.1 show that the weighted characteristic graphs must have components of order 2 or less. By Theorem 2.1.4, each $C_n^X$ contains at most three components. Since this is the case, we want to examine possible pairs of characteristic graphs on 5 points. The only possible graph that meets the above criterion is one composed of 2 $K_2$’s and a single vertex. There are only two possible placements for this single vertex, either it is the same on both characteristic graphs, or it is not. Thus, we must only consider these two pairs up to isomorphism.

Case 1:

\[
\begin{pmatrix}
0 & 3 & 3 & 3 & 2 \\
3 & 0 & 3 & 3 & 3 \\
3 & 3 & 0 & 2 & 3 \\
3 & 3 & 2 & 0 & 3 \\
2 & 3 & 3 & 3 & 0
\end{pmatrix} \ast \begin{pmatrix}
0 & 7 & 7 & 7 & 7 \\
7 & 0 & 5 & 7 & 7 \\
7 & 5 & 0 & 7 & 7 \\
7 & 7 & 0 & 5 & 3 \\
7 & 7 & 7 & 5 & 0
\end{pmatrix} = \begin{pmatrix}
0 & 21 & 21 & 21 & 14 \\
21 & 0 & 15 & 21 & 21 \\
21 & 15 & 0 & 14 & 21 \\
21 & 14 & 0 & 15 & 14 \\
14 & 21 & 21 & 15 & 0
\end{pmatrix}
\]

Note that all weights are 21 except in the path B-C-D-E-A. We will examine this further in a moment.
Case 2:

\[
\begin{pmatrix}
0 & 3 & 3 & 3 & 3 \\
3 & 0 & 3 & 3 & 2 \\
3 & 3 & 0 & 2 & 3 \\
3 & 3 & 2 & 0 & 3 \\
3 & 2 & 3 & 3 & 0
\end{pmatrix}
\begin{pmatrix}
0 & 7 & 7 & 7 & 7 \\
7 & 0 & 5 & 7 & 7 \\
7 & 5 & 0 & 7 & 7 \\
7 & 7 & 7 & 0 & 5 \\
7 & 7 & 7 & 5 & 0
\end{pmatrix}
= 
\begin{pmatrix}
0 & 21 & 21 & 21 & 21 \\
21 & 0 & 15 & 21 & 14 \\
21 & 15 & 0 & 14 & 21 \\
21 & 21 & 14 & 0 & 15 \\
21 & 14 & 21 & 15 & 0
\end{pmatrix}
\]

Note that in this Set graph, all weights are 21 except for the path B-C-D-E.

Since these are the only cases we must examine, we know all permutations of these graphs will contain either a 4 cycle containing no edges of weight 21, or a path of length 4 with no edges of weight 21. In the graphs with the path of length 4, removing that path yields:
This graph clearly has a monochromatic (equivalued) triangle on A, B, D.

Removing the 4 cycle from graphs that have it yields:

This graph has two monochromatic triangles, proving that $\mu_2 < 5$. We proceed now with an example to verify that $\mu_2 \geq 4$

We see here that there is no set on these 4 points, completing the proof. □
Definition 3.3. We say a Set Matrix is the weighted adjacency matrix of a Set graph, $S_n$.

A Set matrix contains valuable information about a collection of cards in play. A natural question to ask is if there is a way to find out if there is a set just by looking at the matrix. Another thought would be to generate bounds on cap sets by bounding the size of the Set matrices. This leads us to the following results.

Theorem 3.1. Consider a collection of $n$ cards from a $k$ characteristic game of Set. Let $S_n$ be the Set graph representing this collection. Suppose $G_x \subseteq S_n$ such that $V(G_x) = V(S_n)$ and $E(G_x) = \{v_s v_t \in E(S_n) | \prod_{i=1}^{k} \Omega_i(v_s v_t) = x\}$. Let $M$, $M_x$ be the weighted adjacency matrix of $S_n$, $G_x$, respectively, then the number of sets in the collection is exactly

$$\frac{1}{6} \sum_{x \in B} tr\left(\frac{1}{x} M_x\right)^3$$

where $B$ is the collection of distinct integers in $M$.

Proof. Recall that if 3 cards form a set, those three cards form an equivalued $K_3$ in $S_n$. $M_x$ represents the $n \times n$ matrix where each entry is either $x$ or 0, and $M_x$ contains every $x$ that was in $M$. The scalar $\frac{1}{x}$ converts $M_x$ into a matrix with entries either 0 or 1. Taking the trace of the cube of this scaled matrix yields the total number of $K_3$’s multiplied by 6. Scaling by $\frac{1}{6}$ gives the exact number of triangles. Summing this process over all distinct values in $M$ yields the number of equivalued $K_3$’s, and thus, the number of sets. \qed
Theorem 3.2. (Classification of Set Matrices) If $\mathcal{M}$ is a Set matrix with no sets, then it has the following properties:

1. $\mathcal{M}$ is symmetric about the main diagonal.
2. $\mathcal{M}$ has zeros on the main diagonal.
3. $\mathcal{M}$ has at most $2^{k-1}$ distinct non-zero square free entries with $p_{2k}$ the largest prime factor of any entry.
4. If $m_{s,r} = m_{s,t}$ then $m_{r,t} \neq m_{s,r}$.
5. Let $N = \{1 \ldots k\}, B \subseteq N$. If $p = \prod_{b \in B} p_{2b-1}$, $p|m_{s,r}$, and $p|m_{s,t}$, then $p|m_{r,t}$.

Proof.

1. Each weighted characteristic graph is symmetric about the main diagonal. The Hadamard product clearly preserves this property.
2. Each weighted characteristic graph is simple, and thus has zeros on the main diagonal.
4. Suppose not. If $m_{s,r} = m_{s,t} = m_{r,t}$ then we have an equivalued $K_3$, thus, there is a set. Then $\mathcal{M}$ would represent a game with a set.
5. Suppose $p_{2b_1-1}|p$, then we have that $p_{2b_1-1}|m_{s,r}$ and $p_{2b_1-1}|m_{s,t}$.

By Definition 2, this means that $st, sr \in E(C_n^{X'})$ and thus $rt \in$
using graphs and games to generate cap set bounds by theorem 2.1.1. thus, if \( p_{2b-1} | p \), the result holds. since \( p_{2b-1} | p \) \( \forall b \in B \), the result follows for \( p = \prod_{b \in B} p_{2b-1} \).

(6) we see that if \( m_{a,t} = qr \), then \( a \) and \( t \) have exactly \(|S|\) characteristics in common and differ in exactly \(|D|\) characteristics. we fix \( a \) and maximize the collection of \( t's \). let \( a, t_i \in \mathbb{Z}_3^k \). by construction, \( \Omega_s(at_s) = p_{2s-1} \forall s \in S \). i.e. \( a, t_s \) share each of the characteristics represented by \( S \). thus, it is sufficient to find a set in the projection of \( \mathbb{Z}_3^k \) onto \( \mathbb{Z}_3^{|D|} \), the \(|D|\)-dimensional space over \( \{0, 1, 2\} \) which corresponds to the \(|D|\) characteristics for which \( \Omega_d(at_d) = p_{2d} \). as such, we must only find a set in the \(|D|\)-tuples. let a prime indicate the \(|D|\)-tuple that is this projection. without loss of generality, let \( a' = (2, 2, \ldots, 2, 2). \quad \Omega_d(a't_d) = p_{2d} \forall d \in D \), thus, each component of the \( t'_i \)s differs from the corresponding component in \( a' \). as such, the \( t'_i \)s are over \( \{0, 1\} \). we see that addition of a set modulo 3 yields the \((0, 0, \ldots, 0) \) \(|D|\)-tuple, and since \( t'_i \neq t'_j \forall i \neq j \), we cannot have a set with the \( t'_i \)s alone. thus, if we have a set, it must contain \( a \). let

\[
f : \{b_i\} \to \{t'_i\}
\]

be defined as \( b_i = (1, 1, \ldots, 1) - t'_i \), then \( \{a', b_i, t'_i\} \) forms a set. clearly, \( f \) is one to one and onto, thus \( |\{b_i\}| = |\{t'_i\}| \). by construction, \(|\{b_i\}| + |\{t'_i\}| = 2^{|D|} \), thus, \(|\{t'_i\}| = 2^{|D|-1} = |T| \) maximally, as required.

\( \square \)
In Theorems 3.1 and 3.2, we have provided a criterion for a matrix representing a no set. We can proceed from here to generate cap set bounds. It is clear that if we have a maximum no set, then $\mathcal{M}$ is a $\mu_k \times \mu_k$ matrix. Thus, if we want to bound $\mu_k$, we need only bound $|T|$ from Theorem 3.2.6. Since the set $T$ deals with entries in the set matrix containing exactly $|D|$ primes of the form $p_{2d}$, we have exactly $\binom{k}{|D|}$ of them. Thus, we can say

$$\mu_k \leq \sum_{j=1}^{k} \binom{k}{j} |T|$$

In our case, we have bounded $|T|$ by $2^{|D|}-1$. Thus, finding better bounds for $|T|$ corresponds to finding better bounds for $\mu_k$. We can do better, though, by letting the set $T$ be the collection of more than one number in a row of $\mathcal{M}$ as it is in Theorem 3.2.6. Since the above bound for $T$ is the least upper bound, we left with the latter option for further research.

References

