Discriminating codes in bipartite graphs

E. Charbit [1], I. Charon [2], G. Cohen [3] and O. Hudry [4]

Computer and Networks Department
ENST-CNRS
Paris, France

Abstract

We study some combinatorial and algorithmic properties of discriminating codes in bipartite graphs. In particular, we provide bounds on minimum discriminating codes and give constructions. We also show that upperbounding the size of a discriminating code is NP-complete.

Keywords: Coverings, bipartite graphs, identifying codes, complexity.

1 Definitions

Let $G = (I \cup A, E)$ be a bipartite, undirected graph. For any vertex v of G, let $N(v)$ denote the neighborhood of v. A subset C of A is said to be a *discriminating code* of G if:

- $\forall i \neq j \in I : C \cap N(i) \neq C \cap N(j)$, and
- $\forall i \in I : C \cap N(i) \neq \emptyset$.

For instance, I can be viewed as a set of *individuals* and A as a set of *attributes*, with an edge between $i \in I$ and $a \in A$ if i owns a; a discriminating code is then a set of attributes sufficient to distinguish all the individuals (for a related notion, see [6]).

Discriminating codes are closely related to *locating-dominating codes* [3] and to *identifying codes* [5] (see also [1] for references).

[1] Email: `emmanuel.charbit@polytechnique.org`
[2] Email: `irene.charon@enst.fr`
[3] Email: `gerard.cohen@enst.fr`
[4] Email: `olivier.hudry@enst.fr`

The differences are that

- "codewords" must belong to a prescribed subset (namely, A),
- only some vertices (namely, the elements of I) must be distinguished.

We study some combinatorial and algorithmic properties of discriminating codes. In particular, we provide bounds on minimum discriminating codes and give constructions. We also show that upperbounding the size of a discriminating code is NP-complete.

2 Upper bounds

Two individuals are called **twins** if their neighborhoods are equal; the graph $G = (I \cup A, E)$ is said to be **twin-free** if no elements of I are twins. The following is an obvious characterization of bipartite graphs admitting a discriminating code:

Proposition 2.1 *A necessary and sufficient condition for a bipartite graph to possess a discriminating code is that it be twin-free.*

Indeed, if there exist twins, they cannot be discriminated even by the whole set A; otherwise $C = A$ fits.

We now characterize minimality with respect to inclusion.

Proposition 2.2 *Consider a twin-free bipartite graph $G = (I \cup A, E)$. A discriminating code B is minimal for inclusion if and only if:*

$$\forall b \in B, \exists\ i \text{ and } j \in I \ :\ (N(i) \cap B)\Delta(N(j) \cap B) = \{b\},$$

where Δ denotes the symmetric difference.

Proof Sketch: If some $b \in B$ violates the condition, then it can be removed without destroying the discrimination, thus B is not minimal.

Define the incidence matrix of such a B, also denoted by B, to be the $M \times N$ binary matrix whose rows are indexed by the $b_i \in B$ and columns by the elements of I, with $|I| := N$ and $|B| := M$.

We now use Proposition 2.2 to upperbound the size M of a minimal discriminating code B.

Proposition 2.3 *If B is minimal for inclusion, its rows are independent; thus $M = rk(B) \leq rk(A) \leq N$.*

Proof sketch: By Proposition 2.2, for any row b, there exist two columns, say i and j, of B differing exactly on b. Thus, any linear combination (over an arbitrary field) of rows involving b projects to a nonzero couple on coordinates $\{i, j\}$.

Definition 2.4 Define the density Δ of the discriminating code C by:

$$\Delta := |C|/N.$$

Then, by the previous result, a minimum code has $\Delta \leq rk(A)/N \leq 1$.

Example 2.5

- Consider the Fano plane: $|A| = 7, N = 7$. Its incidence matrix A has rank equal to 4 over GF(2), thus $M \leq 4$. In fact, $M = 4$, and any 4 rows of A can be chosen to form a minimal B.
- On the other hand, consider a graph G wich is a perfect matching. In this case, A is the only discriminating code and we have: $\Delta = |A|/N = 1$, which shows that the inequality $\Delta \leq 1$ is tight.

3 The covering lower bound

As B satisfies $\forall i \in I, N(i) \cap B \neq \emptyset$, B is a covering of I.

Setting $\eta = \max_{a \in A} deg(a)$, the classical covering bound reads:

$$M \geq \tfrac{|I|}{\eta}, \text{ i.e., } \Delta \geq 1/\eta.$$

Note that the maximal degree can be replaced by the average degree α.

4 An improvement on the covering bound

The previous bound reads: $M \times \eta \geq N$. Assume for simplicity that G is attribute-regular, i.e., that $\eta = \alpha$. If B is discriminating, one can associate distinct lists of attributes to all individuals; at most M individuals can get lists of size 1, and the remaining $N - M$ will get lists of size at least 2. Thus, a double-counting of edges gives: $M \times \alpha \geq M + 2(N - M)$, i.e.

$$M \geq \frac{2}{\alpha + 1} \times N. \quad (1)$$

Let's call *perfect* a code meeting (1) with equality.

For any bipartite graph $G = (I \cup A, E)$, let $G^c = (I \cup A, F)$ be the *bipartite complementary graph* of G: an edge $\{i, a\}$ with $i \in I$ and $a \in A$ belongs to F if and only if it does not belong to E.

Proposition 4.1 *If B is a discriminating code of G, then B is also a discriminating code of G^c.*

Moreover, notice that G is α-attribute-regular if and only if G^c is $(N-\alpha)$-attribute-regular. From the previous proposition and the above, we get:

Theorem 4.2 *Let B be a discriminating code of G. Then we have:*

$$|B| \geq max(\lceil\ (\frac{2}{\alpha+1} \times |I|)\ \rceil, \lceil\ (\frac{2}{|I|-\alpha+1} \times |I|)\ \rceil).$$

5 Lower bounds based on the discrimination

The classical information-theoretic lower bound yields:

$$M \geq \log_2 N.$$

This is tight: pick $B = [1, M], \quad I = \mathcal{P}(\mathcal{B})$.

There is a simple improvement when the graph is attribute-regular, with $deg(a) := \omega N$:

$$M \geq (\log_2 N)/h(\omega),$$

where $h(x) := -x \log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function. This is also tight: pick again $B = [1, M], \quad I = \binom{[1,M]}{\omega M}$.

6 Infinite graphs

We present results on three classical examples of infinite bipartite graphs. The proofs will appear in the full paper.

First, the p-ary complete trees ($p \geq 1$), for which we have:

Theorem 6.1 *Let p and d be any positive integers. For the complete p-ary tree (p fixed) of depth d:*

- *if $p = 1$ (infinite path), then $\lim_{d\to\infty} \Delta = \frac{2}{3}$.*
- *if $p \geq 2$, then $\lim_{d\to\infty} \Delta = \frac{p^2}{p^2+1}$.*

Remark. Notice that when p grows, Δ tends to 1, the upper bound of Proposition 2.3.

The second example is the square grid (a regular graph with $\alpha = 4$), for which the improved covering bound (1) gives:

$$\Delta \geq 2/(1+\alpha) = 2/5.$$

This can be in fact constructively achieved, so that we have:

Theorem 6.2 *The square grid admits a perfect discriminating code of density* $\Delta = 2/5$.

The last infinite graph is the hexagonal mesh (a regular graph with $\alpha = 3$), for which the improved covering bound (1) gives:

$$\Delta \geq 2/(1+\alpha) = 1/2.$$

Here also, this can be constructively achieved:

Theorem 6.3 *The hexagonal mesh admits a perfect discriminating code of density* $\Delta = 1/2$.

7 Attribute-regular graphs

We know from Section 2 that the size of a discriminating code of minimum cardinality cannot exceed N and that this bound can be reached. We may be more specific for attribute-regular graphs.

Theorem 7.1 *Let* α *be a positive integer and let* G *be* α*-attribute-regular graph with* $|I| = N$. *Let* B *be a discriminating code of minimum cardinality. The only graphs* G *with* $|B| = N$ *are the ones with* $\alpha = 1$ *or* $\alpha = N-1$ *and admitting a matching of cardinality* N.

Similarly, for $\alpha = 2$ or $\alpha = N-2$, we can characterize the α-attribute-regular graphs for which $|B| = N-1$ holds for any discriminating code of minimum cardinality. On the other hand, for $3 \leq \alpha \leq N-3$, it is possible to build an infinite family of α-attribute-regular graphs such that any discriminating code of minimum cardinality has size equal to $N-2$.

8 Nonconstructive upper bound

Consider a random incidence $(M \times N)$ matrix B, every entry being chosen independently equal to 1 with probability p (Bernoulli distribution).

Proposition 8.1 *There exists a discriminating code B with* $M > (2\log_2 N)/h(p)$.

Remark. This is twice the lower bound of Section 5.

Proof sketch: Denote by π_{ij} the probability that the two columns i and j be equal. Then $\pi_{ij} = 2^{-Mh(p)}$. By the Union Bound, we get for the probability π of occurrence of identical columns (twins): $\pi \leq \binom{N}{2}2^{-Mh(p)}$.

For $M > (2\log_2 N)/h(p)$, $\binom{N}{2}2^{-Mh(p)}$ is less than 1. Hence $\pi < 1$, and there exists a discriminating code of size M.

9 Complexity results

The following problem has been shown to be NP-complete in [2] (for references and notations on the theory of complexity, see [4]):

Problem Π (**Identification**)
Instance: A graph $G = (V, E)$, an integer K.
Question: Is there an identifying code in G of size at most K?

The problem addressed in this paper is the following:

Problem Π' (**Discrimination**)
Instance: A *bipartite* graph $G' = (I' \cup A', E')$, an integer K'.
Question: Is there a discriminating code in G' of size at most K'?

Proof sketch. Let $G = (V, E)$ be the graph of any instance of Π. Then, set $I' = V$ and $A' = \{N(v) : v \in V\}$, where $N(v)$ still denotes the neighborhood of v, and define E' by: $E' = \{\{j, N(i)\}$ if $j \in N(i)$ in $G\}$. Last, set $K' = K$.

It is easy to show that G admits an identifying code of size at most K if and only of G' admits a discriminating code of size at most K' (more precisely, if B is an identifying code of G, then $\{N(b) : b \in B\}$ is a discriminating code of G', and conversely).

This yields the reduction $\Pi < \Pi'$, thus implying the NP-completeness of Π'.

References

[1] http://www.infres.enst.fr/∼lobstein/bibLOCDOMetID.html

[2] Cohen G., I. Honkala, A. Lobstein and G. Zémor, *On identifying codes*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **56** (2001) 97-109.

[3] Colbourn C.J., P.J. Slater and L.K. Stewart, *Locating dominating sets in series parallel networks*, Congr. Numer. **56** (1987), 135-162.

[4] Garey M. and D. Johnson, "Computers and Intractability", Freeman and Co, 1979.

[5] Karpovsky M.G., K. Chakrabarty and L.B. Levitin, *On a new class of codes for identifying vertices in graphs*, IEEE Trans. Inform. Theory **44** (1998), 599-611.

[6] Kogan A., *On the essential tests of discrete matrices*, Discrete Applied Mathematics **60** (1995), 249-255.