SIAM J. DISC. MATH.
Vol. 2, No. 1, pp. 68-72, February 1989

008

ON THE SIZE OF SYSTEMS OF SETS EVERY t OF WHICH HAVE AN SDR, WITH AN APPLICATION TO THE WORST-CASE RATIO OF HEURISTICS FOR PACKING PROBLEMS*

C. A. J. HURKENS† AND A. SCHRIJVER‡

Abstract. Let $E_1, \cdots, E_m$ be subsets of a set V of size n, such that each element of V is in at most k of the E_i and such that each collection of t sets from $E_1, \cdots, E_m$ has a system of distinct representatives (SDR). It is shown that $m/n \leqq (k(k-1)^r - k)/(2(k-1)^r - k)$ if $t = 2r - 1$, and $m/n \leqq (k(k-1)^r - 2)/(2(k-1)^r - 2)$ if $t = 2r$. Moreover it is shown that these upper bounds are the best possible. From these results the "worst-case ratio" of certain heuristics for the problem of finding a maximum collection of pairwise disjoint sets among a given collection of sets of size k is derived.

Key words. packing, system of distinct representatives, worst-case ratio, heuristics

AMS(MOS) subject classifications. 05C65, 05A05, 90C27

1. Introduction. We prove the following theorem, where m, n, k, and t are positive integers, with $k \geqq 3$.

THEOREM 1. *Let $E_1, \cdots, E_m$ be subsets of the set V of size n, such that we have the following*:

(1) (i) *Each element of V is contained in at most k of the sets $E_1, \cdots, E_m$;*
(ii) *Any collection of at most t sets among $E_1, \cdots, E_m$ has a system of distinct representatives.*

Then, we have the following:

$$(2)\quad \text{(i)}\ \frac{m}{n} \leqq \frac{k(k-1)^r - k}{2(k-1)^r - k} \quad \text{if } t = 2r - 1;$$

$$\text{(ii)}\ \frac{m}{n} \leqq \frac{k(k-1)^r - 2}{2(k-1)^r - 2} \quad \text{if } t = 2r.$$

Note that by the König–Hall Theorem, condition (1)(ii) can be replaced by the following:

(3) For any $s \leqq t$, any s of the sets among $E_1, \cdots, E_m$ cover at least s elements of V.

We give a proof of Theorem 1 in § 2. We also show that the bounds given in (2) are best possible in the following sense.

THEOREM 2. *For any fixed k, t (with $k \geqq 3$), there exist m, n and $E_1, \cdots, E_m \subseteq V$ (with $|V| = n$) satisfying* (1) *and having equality in the appropriate line of* (2).

The proof of Theorem 2 is based on a construction using regular graphs of large girth (see § 3).

Finally, in § 4 we apply these results to derive the worst-case ratio of certain heuristic algorithms for the problem of finding a largest family of pairwise disjoint sets among a given family of sets of size k (this problem is NP-complete for any $k \geqq 3$).

---

* Received by the editors April 1, 1988; accepted for publication July 13, 1988. Part of this research was performed while the authors were visiting the Rutgers Center for Operations Research, Rutgers University, New Brunswick, New Jersey 08903.

† Department of Econometrics, Tilburg University, P. O. Box 90153, 5000 LE Tilburg, the Netherlands. The research of this author was supported by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) through the Stichting Mathematisch Centrum.

‡ Department of Econometrics, Tilburg University, P. O. Box 90153, 5000 LE Tilburg, the Netherlands, and Mathematical Centre, Kruislaan 413, 1098 SJ Amsterdam, the Netherlands.

2. Proof of Theorem 1. To show Theorem 1, we first give a lemma. Let $E_1, \cdots, E_m$ be a collection of finite nonempty sets, which we order so that $|E_1|, \cdots, |E_h| \geqq 2$ and $|E_{h+1}| = \cdots = |E_m| = 1$, for some $h \leqq m$. We define a new collection as follows. Let

$$W := E_{h+1} \cup \cdots \cup E_m. \tag{4}$$

Let for each $i = 1, \cdots, h$, X_i be a set of size $|E_i| - 2$, disjoint from $E_1 \cup \cdots \cup E_m$ and so that if $i \neq j$ then $X_i \cap X_j = \varnothing$. Let $X_1 \cup \cdots \cup X_h =: \{y_1, \cdots, y_q\}$. Then the *derived* collection of sets is formed by the following sets:

$$(E_1 \backslash W) \cup X_1, \cdots, (E_h \backslash W) \cup X_h, \{y_1\}, \cdots, \{y_q\}. \tag{5}$$

Furthermore, we define a collection $E_1, \cdots, E_m$ to have the *t-SDR-property* if any t sets among $E_1, \cdots, E_m$ have a system of distinct representatives.

LEMMA. *For $t \geqq 3$, if $E_1, \cdots, E_m$ has the t-SDR-property, then the derived collection (5) has the $(t-2)$-SDR-property.*

Proof. Suppose (5) does not have the $(t-2)$-SDR-property. Then there exists a collection Π of p sets among (5) covering at most $p-1$ elements, for some $p \leqq t-2$. Assume we have chosen p minimal. This immediately implies the following:

(6) (i) $|\bigcup \Pi| = p - 1$;
(ii) Each element in $\bigcup \Pi$ is covered by at least two sets in Π.

From (6)(ii) we directly have for any $i = 1, \cdots, h$ and $x \in X_i$:

$$\{x\} \in \Pi \Leftrightarrow (E_i \backslash W) \cup X_i \in \Pi. \tag{7}$$

Without loss of generality, all sets $(E_1 \backslash W) \cup X_1, \cdots, (E_h \backslash W) \cup X_h$ belong to Π (as we can delete all sets E_j from $E_1, \cdots, E_h$ for which $(E_j \backslash W) \cup X_j \notin \Pi$), and without loss of generality, $(E_1 \cup \cdots \cup E_h) \cap W = E_{h+1} \cup \cdots \cup E_m$.

Note the following:

$$q = |X_1 \cup \cdots \cup X_h| = \sum_{i=1}^{h} (|E_i| - 2), \qquad p = h + q, \tag{8}$$
$$\left| \bigcup_{i=1}^{h} (E_i \backslash W) \right| = |\bigcup \Pi| - q = (p-1) - q = h - 1.$$

So,

$$\left| \bigcup_{i=1}^{m} E_i \right| = \left| \bigcup_{i=1}^{h} (E_i \cap W) \right| + \left| \bigcup_{i=1}^{h} (E_i \backslash W) \right| = (m-h) + (h-1) = m-1. \tag{9}$$

Moreover, by (6)(ii), $\sum_{i=1}^{h} |E_i \backslash W| \geqq 2 \cdot |\bigcup_{i=1}^{h} (E_i \backslash W)|$, and hence

$$\begin{aligned} m = h + \left| \bigcup_{i=1}^{h} (E_i \cap W) \right| &\leqq h + \sum_{i=1}^{h} |E_i \cap W| = h + \sum_{i=1}^{h} |E_i| - \sum_{i=1}^{h} |E_i \backslash W| \\ &\leqq h + \sum_{i=1}^{h} |E_i| - 2 \cdot \left| \bigcup_{i=1}^{h} (E_i \backslash W) \right| = h + 2h + \sum_{i=1}^{h} (|E_i| - 2) - 2(h-1) \\ &= h + 2h + q - 2(h-1) = h + q + 2 = p + 2 \leqq t. \end{aligned} \tag{10}$$

Inequalities (9) and (10) contradict the fact that $E_1, \cdots, E_m$ has the t-SDR-property. □

Proof of Theorem 1. We prove Theorem 1 by induction on t.

Case 1. $t = 1$. Then we have that each of $E_1, \cdots, E_m$ is nonempty, and hence $m \leqq \sum_{i=1}^{m} |E_i| \leqq kn$, by (1)(i).

Case 2. $t = 2$. Then we have that each of $E_1, \cdots, E_m$ is nonempty, and that no two of the singletons among $E_1, \cdots, E_m$ are the same. Without loss of generality, let $E_{h+1}, \cdots, E_m$ be the singletons among $E_1, \cdots, E_m$. Then $m - h \leqq n$, and

$$m+h=2h+(m-h)\leqq \sum_{i=1}^{h} |E_i| + \sum_{i=h+1}^{m} |E_i| = \sum_{i=1}^{m} |E_i| \leqq kn \tag{11}$$

(by (1)(i)). Hence $2m = (m - h) + (m + h) \leqq (k + 1)n$, and (2) follows.

Case 3. $t \geqq 3$. Then consider the derived collection $E'_1, \cdots, E'_{m'}$ on $V' := \bigcup_{i=1}^{m'} E'_i$ as in (5). Note that $m' = h + q$ and $n' := |V'| = n - |W| + q$. Denote the right-hand side term in (2) by $\varphi(k, t)$.

As by the lemma above, $E'_1, \cdots, E'_{m'}$ has the $(t - 2)$-SDR-property, and as trivially each element of V' is in at most k of the sets $E'_1, \cdots, E'_{m'}$ we have by induction that $m' \leqq \varphi(k, t - 2)n'$. That is,

$$h+q\leqq\varphi(k,t-2)(n-|W|+q). \tag{12}$$

Writing the terms in different order, we have

$$\varphi(k,t-2)|W|+h-(\varphi(k,t-2)-1)q\leqq\varphi(k,t-2)n. \tag{13}$$

Moreover, as $E_1, \cdots, E_m$ cover any element at most k times:

$$|W|+2h+q=|W|+2h+\sum_{i=1}^{h}(|E_i|-2)=|W|+\sum_{i=1}^{h}|E_i| = \sum_{i=1}^{m}|E_i|\leqq kn. \tag{14}$$

Hence,

$$\begin{aligned} m &= h+|W| \\ &= \frac{1}{2\varphi(k,t-2)-1}(\varphi(k,t-2)|W|+h-(\varphi(k,t-2)-1)q) \\ &\quad + \frac{\varphi(k,t-2)-1}{2\varphi(k,t-2)-1}(|W|+2h+q) \\ &\leqq \frac{1}{2\varphi(k,t-2)-1}\varphi(k,t-2)n + \frac{\varphi(k,t-2)-1}{2\varphi(k,t-2)-1}kn \\ &= \frac{(k+1)\varphi(k,t-2)-k}{2\varphi(k,t-2)-1}n = \varphi(k,t)n. \end{aligned} \tag{15}$$

The last equality follows directly by substituting the corresponding right-hand side of (2). □

3. Proof of Theorem 2. To prove Theorem 2 we use a result of Erdös and Sachs [1]:

(16) For every k and γ there exists a k-regular graph of girth γ.

As a consequence of (16) we have the following:

(17) For every k, s, and γ there exists a bipartite graph of girth at least γ, with color classes U and W, say, such that each vertex in U has degree k, and each vertex in W has degree s.

(To see that (17) follows from (16), let H be a $2ks$-regular graph of girth γ. Consider any Eulerian orientation of the edges of H (i.e., one for which all indegrees and outdegrees equal ks). Split each vertex v into $k + s$ vertices $v_1, \cdots, v_k, w_1, \cdots, w_s$ and divide the arcs entering v equally over $v_1, \cdots, v_k$ and divide the arcs leaving v equally over $w_1, \cdots, w_s$. Forgetting the orientations, we obtain a bipartite graph with the required properties.)

Now choose k, t. Let $r := \lfloor \frac{1}{2}t \rfloor$. Consider the tree T, with vertices $1, 2, \cdots, 1 + (k-1) + (k-1)^2 + \cdots + (k-1)^{r-1}$, so that for $i < j$, vertices i and j are connected by an edge, if and only if $(k-1)i \leqq j \leqq (k-1)i + (k-2)$. So each vertex has degree k, except for vertex 1, which has degree $k-1$, and for the vertices $1 + (k-1) + \cdots + (k-1)^{r-2} + 1, \cdots, 1 + (k-1) + \cdots + (k-1)^{r-1}$, which have degree one.

First let t be even. Let G be a $(k-1)^r$-regular graph of girth $t + 1$ (cf. (16)). Let G have p vertices: $v_1, \cdots, v_p$. Consider p copies $T_1, \cdots, T_p$ of T (denoting the copy of vertex i in T_j by i_j). For each $j = 1, \cdots, p$, partition the set of $(k-1)^r$ edges of G incident to v_j (arbitrarily) into $(k-1)^{r-1}$ classes of size $k-1$, and connect them to the $(k-1)^{r-1}$ vertices i_j in T_j of degree one. So the final graph $H = (W, F)$ has all degrees equal to k, except for the vertices $1_1, \cdots, 1_p$, which have degree $k-1$. Let $E_1, \cdots, E_m$ be the collection $F \cup \{\{1_1\}, \cdots, \{1_p\}\}$. This collection clearly satisfies (1)(i), and direct counting shows equality in (2)(ii). To see that the collection satisfies (1)(ii), let $E_1, \cdots, E_s$ form a subcollection with $|E_1 \cup \cdots \cup E_s| < s$ and s as small as possible. Suppose $s \leqq t$. As $E_1, \cdots, E_s$ must form a connected hypergraph, it contains at most one singleton (since any path between 1_i and 1_j in H contains at least $t - 1$ edges). So assume $E_2, \cdots, E_s$ are edges of H. Then they do not contain any circuit (as each T_i is a tree and as G has girth $t + 1 > s$). So $|E_2 \cup \cdots \cup E_s| \geqq s$, a contradiction.

Next let t be odd. Let G be a bipartite graph, of girth at least $t + 1$, so that in one color class U each vertex has degree $(k-1)^r$ and in the other color class W each vertex has degree k. Let $U =: \{u_1, \cdots, u_p\}$. Consider again p copies $T_1, \cdots, T_p$ of T, as above. For $j = 1, \cdots, p$ partition the set of $(k-1)^r$ edges of G incident to u_j (arbitrarily) into $(k-1)^{r-1}$ classes of size $k-1$, and connect them to the $(k-1)^{r-1}$ vertices i_j in T_j of degree one. Again, the final graph $H = (W, F)$ has all degrees equal to k, except for the vertices $1_1, \cdots, 1_p$ that have degree $k-1$. Let $E_1, \cdots, E_m$ be the collection $F \cup \{\{1_1\}, \cdots, \{1_p\}\}$. Similarly, as above, we show that this collection satisfies (1) and has equality in (2)(i).

4. Application to the worst-case ratio of heuristics. The problem of finding a largest collection of pairwise disjoint sets among a given collection $X_1, \cdots, X_q$ of k-sets is NP-complete, for any $k \geqq 3$. Call any collection of pairwise disjoint sets a *packing*.

For any fixed s, we can apply the following heuristic algorithm H_s. Start with the empty packing. If we have found a packing $Y_1, \cdots, Y_n$ from $X_1, \cdots, X_q$, we could select $p \leqq s$ sets among $Y_1, \cdots, Y_n$, and replace them by $p + 1$ sets from $X_1, \cdots, X_q$, so that the arising collection is a packing with $n + 1$ sets. Repeating this, the algorithm terminates with a collection $Y_1, \cdots, Y_n$ so that

(18) For each $p \leqq s$, the union of any $p + 1$ pairwise disjoint sets among $X_1, \cdots, X_q$ intersects at least $p + 1$ sets among $Y_1, \cdots, Y_n$.

This defines heuristic H_s, which is, for any fixed s, a polynomial-time algorithm—however it clearly need not lead to a largest packing. We might ask how far the packing found with H_s is from the largest packing.

To this end, consider a largest packing $Z_1, \cdots, Z_m$ from $X_1, \cdots, X_q$. We claim that m/n satisfies the bounds given in (2), taking $t := s + 1$, and that these bounds are best possible. That is, the "worst-case ratio" of the heuristic is given in (2).

Indeed, let

(19) $V := \{Y_1, \cdots, Y_n\}$ and $E_i := \{Y_j \mid Y_j \cap Z_i \neq 0\}$ for $i = 1, \cdots, m$.

Then by (18), $E_1, \cdots, E_m$ satisfy (1), and hence we obtain the bounds given in (2).

In turn, it is not difficult to see that for any collection $E_1, \cdots, E_m$ of sets of size at most k, containing any point at most k times, we can assume they are of form (19) for certain packings $Y_1, \cdots, Y_n$ and $Z_1, \cdots, Z_m$ of k-sets. Thus starting with $E_1, \cdots, E_m$ as described in § 3 above, making these $Y_1, \cdots, Y_n, Z_1, \cdots, Z_m$, and taking $\{X_1, \cdots, X_q\} := \{Y_1, \cdots, Y_n, Z_1, \cdots, Z_m\}$, we obtain a system of sets attaining the worst-case ratio. (That is because we may assume that H_s selects the sets $Y_1, \cdots, Y_n$ in the first n iterations.)

Note that we may assume even that the sets $Y_1, \cdots, Y_n, Z_1, \cdots, Z_m$ form the collection of all cliques of size k in a graph. Hence, we cannot obtain a better worst-case ratio by restricting the collections of sets to collections of k-cliques.

REFERENCE

[1] P. Erdős and H. Sachs, *Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl*, Wiss. Z. Univ. Halle, Math.-Nat., 12 (1963), pp. 251–258.