Allocating Servers in Infostations for On-Demand Communications *

A.A. Bertossi
Dept. Comp. Science
University of Bologna, Italy
bertossi@cs.unibo.it

M.C. Pinotti and R. Rizzi
Dept. Comp. Science & Telecomm.
University of Trento, Italy
{pinotti,rrizzi}@science.unitn.it

P. Gupta
Dept. Comp. Science & Eng.
IIT Kampur, India
pg@cse.iitk.ac.in

Abstract

Given a set of service requests, each characterized by a temporal interval and a category, an integer k, and an integer h_c for each category c, the Server Allocation with Bounded Simultaneous Requests problem consists in assigning a server to each request in such a way that at most k mutually simultaneous requests are assigned to the same server at the same time, out of which at most h_c are of category c, and the minimum number of servers is used. Since this problem is computationally intractable, a 2-approximation on-line algorithm is exhibited which asymptotically gives a $\left(2 - \frac{h}{k}\right)$-approximation, where $h = min\{h_c\}$. Generalizations of the problem are considered, where each request r is also characterized by a bandwidth rate w_r, and the sum of the bandwidth rates of the simultaneous requests is bounded, and where each request is characterized also by a gender bandwidth. Such generalizations contain Bin-Packing and Multiprocessor Task scheduling as special cases, and they admit on-line algorithms providing constant approximations.

1 Introduction

An infostation is an isolated pocket area with small coverage (about a hundred of meters) of high bandwidth connectivity (at least a megabit per second) that collects information requests of mobile users and delivers data while users are going through the coverage area. The available bandwidth depends on the distance between the mobile user and the center of the coverage area, increasing with decreasing distance [4, 8, 9]. Infostations could be located along roadways, at airports, in campuses, and they can provide access ports to Internet and/or access to services managed locally.

The infostation system may retrieve the requested data from remote gateways, may provide Internet access, and may home local services (such as building access, credit card transactions, and map downloads). The mobile user connection starts when it first senses the infostation's presence and finishes when it leaves the coverage area. Depending on the mobility options, three kinds of users are characterized: *drive-through*, *walk-through*, and *sit-through*. According to the mobility options, the bit-rate connection is high variable for drive-through, low variable for walk-through, and fixed for sit-through. In addition to radio broadcast communication, infostations create opportunities to deliver new wireless information services dedicated to single-users, which could be supported for example by infrared technologies.

Each mobile user going through the infostation may require a data service out of a finite set of possible service categories available. The admission control, i.e., the task of deciding whether or not a certain request will be admitted, is essential. In fact, a user going through an infostation to obtain a (toll) service is not disposed to have its request delayed or refused. Hence, the service dropping probability must be kept as low as possible. For this purpose, many admission control and bandwidth allocation schemes for infostations maintain a pool of servers so that when a request arrives it is immediately and irrevocably assigned to a server thus clearing the service dropping probability. Precisely, once a request is admitted, the infostation assigns a temporal interval and a proper bandwidth for serving the request, depending on the service category, on the size of the data required and on the mobility kind of the user, as shown in Table 1 for a sample of requests with their actual parameters. Moreover, the infostation decides whether the request may be served locally or through a remote gateway. In both cases, a server (either in the infostation or in the gateway) is allocated on demand to the request during the assigned temporal interval. The request is immediately assigned to its server without knowing the future, namely with no knowledge of the next request. The server, selected out of the predefined server pool, serves the requests on-line, that is in an ongoing manner as they become available. Moreover, each server may serve more than one request simultaneously but it is subject to some architecture constraints. For example, no more than k requests could be served simultaneously by a local server supporting k infrared channels or by a gateway server connected to k infostations. Similarly, no more than h services of the same category can be delivered simultaneously due to access constraints on the original

*This work is partially supported by MIUR-RealWine Research Program.

Category	Kbps	Seconds	
		Low rate	High rate
FTP download	10000	100	10
Video streams	5000	50	5
Audio streams E-mail attachments	512	5	.5
E-mails Web Browsing	64	.6	.06

Table 1. Example of actual time intervals required to serve different kinds of requests.

data, such as software licenses, limited on line subscriptions and private access.

In this paper, a particular problem arising in the design of infostation systems is faced which consists in finding scheduling algorithms for allocating the minimum number of servers to the user requests in such a way that the temporal, architectural and data constraints are satisfied. In details, a service *request* r will be modeled by a service *category* c_r and a temporal *interval* $I_r = [s_r, e_r)$ with *starting time* s_r and *ending time* e_r. Two requests are simultaneous if their temporal intervals overlap. The input of the problem consists of a set R of service requests, a bound k on the number of mutually simultaneous requests to be served by the same server at the same time, and a set C of service categories with each category c characterized by a bound h_c. The output is a mapping from the requests in R to the servers that uses the minimum possible number of servers to assign all the requests in R subject to the constraints that the same server receives at most k mutually simultaneous requests at the same time, out of which at most h_c are of category c. Such a problem is called in this paper *Server Allocation with Bounded Simultaneous Requests*.

It is worthy to note that, equating servers with bins, and requests with items, the above problem is similar to a generalization of Bin-Packing, known as *Dynamic Bin-Packing* [1], where in addition to size constraints on the bins the items are characterized by an arrival and a departure time, and repacking of already packed items is allowed each time a new item arrives. The problem considered in this paper, in contrast, does not allow repacking and has capacity constraints also on the bin size for each category. Furthermore, equating servers with processors and requests with tasks, the above problem becomes a generalization of deterministic multiprocessor scheduling with task release times and deadlines [6], where in addition each processor can execute more than one task at the same time.

In Section 2, it is shown that Server Allocation with Bounded Simultaneous Requests is computationally intractable. Section 3 shows that the problem cannot be α-approximated with $\alpha < \frac{4}{3}$. Moreover, a 2-approximation on-line algorithm is exhibited which asymptotically gives a $\left(2 - \frac{h}{k}\right)$-approximation, where $h = \min_{c \in C} h_c$. In Section 4, a generalization of the problem is considered where each request r is also characterized by an integer *bandwidth rate* w_r, and the bounds on the number of simultaneous requests to be served by the same server are replaced by bounds on the sum of the bandwidth rates of the simultaneous requests assigned to the same server. For this problem, on-line and off-line algorithms are proposed which give a constant approximation. Other two generalizations are proposed in Section 5 in which each request is characterized either by a multi-dimensional bandwidth rate or by both a bandwidth rate and a gender bandwidth associated to the category of the requests.

2 Computational Intractability

The Server Allocation with Bounded Simultaneous Requests problem on a set $R = \{r_1, \ldots, r_n\}$ of requests can be formulated as a coloring problem on the corresponding set $I = \{I_1, \ldots, I_n\}$ of temporal intervals.

Problem 1. (Interval Coloring with Bounded Overlapping) *Given a set I of intervals each belonging to a category, an integer k, and an integer h_c for each category c, assign a color to each interval in such a way that at most k mutually overlapping intervals receive the same color, at most h_c mutually overlapping intervals all having category c receive the same color, and the minimum number of colors is used.*

In order to prove that Problem 1 is computationally intractable, the following simplified decisional formulation is considered, where there is a bound $h_c = 1$ for each category c.

Problem 2. (Interval Coloring with Unit Bounded Categories) *Given a set I of intervals each belonging to a category, and two integers k and b, decide whether b colors are enough to assign a color to each interval in such a way that at most k mutually overlapping intervals receive the same color and no two overlapping intervals with the same category receive the same color.*

The graph coloring problem below is well-known to be NP-complete [2], even if $b \geq 3$.

Problem 3. (Chromatic Number) *Given an integer b and an undirected graph $G = (V, E)$, decide whether the nodes in V can be colored with b colors in such a way that adjacent nodes receive different colors.*

Recall that a *stable set* of a graph $G = (V, E)$ is a subset S of nodes in V such that no two nodes in S are adjacent.

Problem 4. (Balanced Coloring) *Given two integers b and k, and an undirected graph $G = (V, E)$ of kb nodes, decide whether V can be partitioned into b stable sets each of size k.*

Lemma 2.1. *Balanced Coloring is NP-complete, even if $b \geq 3$.*

Proof: Chromatic Number is reduced in polynomial time to Balanced Coloring as follows. Given an instance of Chromatic Number, namely b and $G = (V, E)$, let $k = |V|$ and $G' = (V', E')$ be the graph obtained by considering b disjoint copies of G. Clearly, G can be colored with b colors if and only if V' can be partitioned into b stable sets each of size k. $\square$

Theorem 2.2. *Interval Coloring with Unit Bounded Categories is NP-complete, even if $b \geq 3$.*

Proof: Given an instance of Balanced Coloring, that is b, k, and $G = (V, E)$, an instance of Problem 2 is constructed in such a way that there are as many categories as there are nodes in V, a subset of intervals correponds to each node in V so that all such intervals are forced to receive the same color, and some pairs of intersecting intervals belong to the same category if and only if their corresponding nodes in V are adjacent.

Specifically, let $kb = |V|$ and $m = |E|$. Take the set of categories as $C = \{1, 2, \ldots, kb\}$. Construct $2kb+2m$ intervals as follows. For each generic node $v \in V$, let $v_1, \ldots, v_\ell$ be the neighbours of v (indexed, for the sake of simplicity, so that $v_i < v_{i+1}$). The following $\ell + 2$ intervals correspond to node v:

- $I_v^0 = [0, 2v - 1)$ with category $c_v^0 = v$,
- $I_v^v = [2v, 2kb + 1)$ with category $c_v^v = v$,
- $I_v^{v_i} = \left[2v - 1 + \frac{i-1}{\ell}, 2v - 1 + \frac{i}{\ell}\right)$ with category $c_v^{v_i} = v_i$, for $i = 1, \ldots, \ell$.

If a node v is isolated, that is it has no neighbour, then a single interval $I_v = [0, 2kb + 1)$ with category $c_v = v$ corresponds to it.

Note that $[v, u] \in E$, with $v < u$, if and only if both the following conditions hold:

1. $I_v^u \cap I_u^0 \neq \emptyset$ and $c_v^u = c_u^0 = u$, and
2. $I_u^v \cap I_v^v \neq \emptyset$ and $c_u^v = c_v^v = v$.

Thus, adjacency between two nodes u and v in G is coded by two pairs of overlapping intervals in I with categories u and v, which cannot be colored the same.

As an example, Figure 1 depicts all the $2kb + 2m$ intervals of I corresponding to a graph $G = (V, E)$ with $kb = 6$ nodes and $m = 10$ edges (in such a figure,

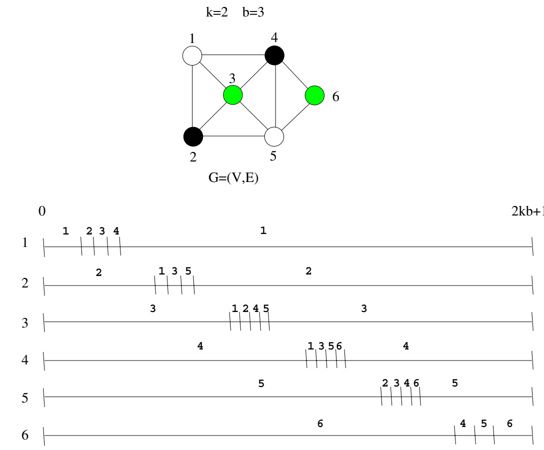

Figure 1. Example of reduction from Balanced Coloring to Interval Coloring with Unit Bounded Categories (categories are represented by numbers above the intervals).

$k = 2$ and $b = 3$). Assume that G can be colored with b colors. Assign to all the intervals $I_v^0, I_v^v, I_v^{v_1}, \ldots, I_v^{v_\ell}$ the same color that node v has in G. Clearly, if nodes v and u in G are colored the same, then they are not adjacent. Thus all the intervals corresponding to v and u can receive the same color. Moreover, since the same color appears in G exactly k times, exactly k mutually overlapping intervals of I receive the same color.

Conversely, assume that all the intervals in I can be colored with b colors. Observe that for each v, all the intervals $I_v^0, I_v^v, I_v^{v_1}, \ldots, I_v^{v_\ell}$ must be colored the same (otherwise, by construction, at least $b + 1$ colors would be required). Therefore, assign such a color also to the node v of G. If $I_v^0, I_v^v, I_v^{v_1}, \ldots, I_v^{v_\ell}$ and $I_u^0, I_u^u, I_u^{u_1}, \ldots, I_u^{u_p}$ have the same color, then nodes v and u are not adjacent in G, and thus they can receive the same color. Moreover, since k mutually overlapping intervals are colored the same, the same color is used in G exactly k times. $\square$

3 Interval Coloring with Bounded Overlapping

An α-approximation algorithm for a minimization problem is a polynomial-time algorithm producing a solution of value $appr(x)$ on input x such that, for all the inputs x,

$$appr(x) \leq \alpha \cdot opt(x),$$

where $opt(x)$ is the value of the optimal solution on x. In other words, the approximate solution is guaranteed to never be greater than α times the optimal

0-7695-1926-1/03/$17.00 (C) 2003 IEEE

solution [2]. For the sake of simplicity, from now on, *appr*(x) and *opt*(x) will be simply denoted by *appr* and *opt*, respectively.

Corollary 3.1. *The optimization version of Interval Coloring with Unit Bounded Categories admits no α-approximation algorithm for $\alpha < \frac{4}{3}$.*

Proof: By the reduction in the proof of Theorem 2.2, if there is an α-approximation algorithm with $\alpha < \frac{4}{3}$, then there is a decision algorithm for Balanced Coloring with $b = 3$. $\square$

Assume that the intervals in I arrive one by one, and are indexed by non-decreasing starting times. When an interval I_i arrives, it is immediately and irrevocably colored, and the next interval I_{i+1} becomes known only after I_i has been colored. An algorithm that works in such an ongoing manner is said *on-line* [5]. On-line algorithms are opposed to *off-line* algorithms, where the intervals are not colored as they become available, but they are all colored only after the entire sequence I of intervals is known.

A simple polynomial-time on-line algorithm for the more general Interval Coloring with Bounded Overlapping can be designed based on the following greedy strategy:

Algorithm 1 GREEDY (I_i)

- To color interval I_i use, if possible, a color already used for previous intervals, otherwise use a brand new color.

Theorem 3.2. *Algorithm Greedy provides a 2-approximation and, asymptotically, a $\left(2 - \frac{h}{k}\right)$-approximation, where $h = \min_{c \in C} h_c$, for Interval Coloring with Bounded Overlapping.*

Proof: Let *appr* $= \phi$ be the solution given by the algorithm and assume the colors $1, \ldots, \phi$ have been introduced in this order. Let $I_r = [s_r, e_r)$ be the first interval colored ϕ. Let Ω_1 be the set of intervals in I containing s_r and let Ω_2 be the set of intervals in I containing s_r and with category c_r. Hence, $\Omega_2 \subseteq \Omega_1$. Let $\omega_1 = |\Omega_1|$ and $\omega_2 = |\Omega_2|$. Clearly, $opt \geq \lceil \frac{\omega_1}{k} \rceil$ and $opt \geq \lceil \frac{\omega_2}{h_{c_r}} \rceil$. Color ϕ was introduced to color I_r because, for every $\gamma \in \{1, \ldots, \phi - 1\}$, at least one of the following two conditions held:

1. at least k intervals in Ω_1 have color γ;
2. at least h_{c_r} intervals in Ω_2 have color γ.

For $i = 1, 2$, let n_i be the number of colors in $\{1, \ldots, \phi - 1\}$ for which condition i holds (if for a color both conditions hold, then choose one of them arbitrarily). Hence, $n_1 + n_2 = \phi - 1$ or, equivalently, $appr = \phi = n_1 + n_2 + 1$. Clearly, $\omega_1 \geq kn_1 + h_{c_r} n_2 + 1$ and $\omega_2 \geq h_{c_r} n_2 + 1$. Therefore:

$$opt \geq \max\left\{\left\lceil \frac{\omega_1}{k} \right\rceil; \left\lceil \frac{\omega_2}{h_{c_r}} \right\rceil\right\} \geq$$

$$\max\left\{\left\lceil \frac{kn_1 + h_{c_r} n_2 + 1}{k} \right\rceil; \left\lceil \frac{h_{c_r} n_2 + 1}{h_{c_r}} \right\rceil\right\} \geq$$

$$\max\left\{n_1 + \left\lceil \frac{h_{c_r} n_2 + 1}{k} \right\rceil; n_2 + 1\right\} \geq$$

$$\max\left\{n_1 + \frac{h}{k} n_2; n_2 + 1\right\}$$

where $h = \min_{c \in C} h_c$.

If $n_2 + 1 \geq n_1 + \frac{h}{k} n_2$, then:

$$\frac{appr}{opt} \leq \frac{n_1 + n_2 + 1}{n_2 + 1} \leq$$

$$\frac{n_2(1 - \frac{h}{k}) + 1 + n_2 + 1}{n_2 + 1} = 2 - \frac{h}{k} \frac{n_2}{n_2 + 1} \leq 2$$

If $n_2 + 1 \leq n_1 + \frac{h}{k} n_2$, then:

$$\frac{appr}{opt} \leq \frac{n_1 + n_2 + 1}{n_1 + \frac{h}{k} n_2} \leq$$

$$\frac{n_1 + n_1 + \frac{h}{k} n_2}{n_1 + \frac{h}{k} n_2} = 1 + \frac{n_1}{n_1 + \frac{h}{k} n_2} \leq 2$$

Therefore, Algorithm 1 gives a 2-approximation.

To achieve the asymptotic approximation, first observe that $opt \geq \max\left\{n_1 + \frac{h}{k} n_2; n_2\right\}$. Moreover, when $opt \to \infty$, also $\phi \to \infty$, and hence $\phi' = \phi - 1 \to \infty$, too. The ratio $\frac{appr}{opt}$ is maximum when *opt* is minimum, that is, since $opt \geq n_2$ and $opt \geq n_1 + \frac{h}{k} n_2$, for $opt' = n_2 = n_1 + \frac{h}{k} n_2$. Thus,

$$\frac{appr}{opt} \to \frac{\phi'}{opt'} = \frac{n_1 + n_2}{n_2} = \frac{(1 - \frac{h}{k}) n_2 + n_2}{n_2} = 2 - \frac{h}{k}$$

Hence, asymptotically, Algorithm 1 gives a $\left(2 - \frac{h}{k}\right)$-approximation. $\square$

The following Corollary 3.3 shows that the $\left(2 - \frac{h}{k}\right)$-approximation is the best possible for Algorithm 1, even in the case that $h = 1$, $k = 2$, and no interval properly contains another interval.

Corollary 3.3. *Algorithm Greedy admits no α-approximation with $\alpha < 2 - \frac{1}{k}$ for Interval Coloring with Unit Bounded Categories.*

Proof: It is shown that there is an instance for which the $(2 - \frac{1}{k})$-approximation is achievable. Consider the particular input instance consisting of the k^2 mutually overlapping intervals $I_1, I_2 \ldots, I_{k^2}$, where the i-th interval $I_i = [i, i + k^2)$ has category

$$c_i = \begin{cases} i & \text{if} \quad 1 \le i \le k^2 - k \\ k^2 - k + 1 & \text{if} \quad k^2 - k + 1 \le i \le k^2 \end{cases}$$

The Greedy algorithm colors the interval I_i as soon as it becomes available, that is at time i, thus assigning color 1 to $I_1, \ldots, I_k$, color 2 to $I_{k+1}, \ldots, I_{2k}$, and so on. In particular, color j is assigned to $I_{(j-1)k+1}, \ldots, I_{jk}$, and color $k-1$ is given to $I_{(k-2)k+1}, \ldots, I_{(k-1)k}$. Moreover, for the remaining intervals $I_{k^2-k+1}, \ldots, I_{k^2}$, k additional colors are employed, one for each interval. Overall, $2k - 1$ colors are used. However, an optimal off-line algorithm, that knows in advance the entire sequence of intervals, uses k colors assigning color 1 to intervals $I_1, \ldots, I_{k-1}$ and interval I_{k^2-k+1}, color 2 to intervals $I_k, \ldots, I_{2(k-1)}$ and I_{k^2-k+2}, and so on. In particular, color j is assigned to intervals $I_{(j-1)(k-1)+1}, \ldots, I_{j(k-1)}$ and I_{k^2-k+j}, while color k is given to $I_{(k-1)(k-1)+1}, \ldots, I_{k(k-1)}$ and I_{k^2}. Therefore, $\frac{appr}{opt} = \frac{2k-1}{k} = 2 - \frac{1}{k}$. □

4 Interval Coloring with Weighted Overlapping

Consider now a generalization of Server Allocation with Bounded Simultaneous Requests, where each request r is also characterized by an integer *bandwidth rate* w_r, and the bounds on the number of simultaneous requests to be served by the same server are replaced by bounds on the sum of the bandwidth rates of the simultaneous requests assigned to the same server. Such a problem can be formulated as a weighted generalization of Problem 1 as follows.

Problem 5. (Interval Coloring with Weighted Overlapping) *Given a set I of intervals, with each interval I_r characterized by a category c_r and an integer weight w_r, an integer k, and an integer h_c for each category c, assign a color to each interval in such a way that the sum of the weights for mutually overlapping intervals receiving the same color is at most k, the sum of the weights for mutually overlapping intervals of category c receiving the same color is at most h_c, and the minimum number of colors is used.*

More formally, denote by

- $I[t]$ the intervals *active* at instant t, that is, $I[t] = \{I_r \in I : s_r \le t \le e_r\}$;
- $I[c]$ the intervals belonging to the same category c, that is $I[c] = \{I_r \in I : c_r = c\}$; and
- $I(\gamma)$ the set of intervals colored γ.

Moreover, let $I(\gamma)[t] = I(\gamma) \cap I[t]$ be the intervals colored γ and active at instant t. Finally, let $I(\gamma)[t][c] = I(\gamma)[t] \cap I[c]$ be the intervals of category c, colored γ, and active at instant t.

The constraints on the sum of the weights for mutually overlapping intervals receiving the same color can be stated as follows:

$$\sum_{I_r \in I(\gamma)[t]} w_r \le k \quad \forall \gamma, \forall t \tag{1}$$

$$\sum_{I_r \in I(\gamma)[t][c]} w_r \le h_c \quad \forall \gamma, \forall t, \forall c \tag{2}$$

An approximation on-line algorithm for Problem 5 (which contains Bin-Packing [1] as a special case) is presented below.

Algorithm 2 FIRST-COLOR (I_i)

- To color interval I_i use the lowest possible indexed color among those already used for previous intervals. If no such color exists, use a brand new color.

Theorem 4.1. *Algorithm First-Color asymptotically provides a constant approximation for Interval Coloring with Weighted Overlapping.*

Proof: Assume the on-line First-Color algorithm employs ϕ colors. Consider the first interval $I_r = [s_r, e_r)$ which is colored ϕ. At time s_r, I_r cannot be colored with any color in $\{1, \ldots, \phi - 1\}$ since otherwise at least one of the constraints (1) and (2) would be violated. Two cases may occur, depending on whether w_r is smaller or larger than $\frac{h}{2}$, where $h = \min_{c \in C} h_c$.

Case 1: Suppose $w_r \le \frac{h}{2}$. Let Ω_1 be the set of intervals in I containing s_r and let Ω_2 be the set of intervals in I containing s_r and having category c_r. Let $w(\Omega_1)$ (resp., $w(\Omega_2)$) be the sum of the weights of the intervals in Ω_1 (resp., Ω_2). Clearly, $opt \ge \lceil \frac{w(\Omega_1)}{k} \rceil$ and $opt \ge \lceil \frac{w(\Omega_2)}{h_{c_r}} \rceil$.

Color ϕ was used for I_r because, for every $\gamma \in \{1, \ldots, \phi - 1\}$, at least one of the following two conditions held:

1. $w(\Omega_1(\gamma)) \ge \frac{k}{2}$, where $\Omega_1(\gamma)$ are the intervals in Ω_1 already colored γ,

0-7695-1926-1/03/$17.00 (C) 2003 IEEE
Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS'03)

2. $w(\Omega_2(\gamma)) \geq \frac{h_{c_r}}{2}$, where $\Omega_2(\gamma)$ are the intervals in Ω_2 with category c_r already colored γ.

For $i = 1, 2$, let n_i be the number of colors in $\{1, \ldots, \phi - 1\}$ for which i holds (if for a color both conditions hold, then choose one of them arbitrarily). Hence, $n_1 + n_2 = \phi - 1$. Clearly, $w(\Omega_1) \geq n_1 \frac{k}{2} + n_2 \frac{h_{c_r}}{2}$ and $w(\Omega_2) \geq n_2 \frac{h_{c_r}}{2}$. Therefore,

$$opt \geq \max\left\{ \left\lceil \frac{w(\Omega_1)}{k} \right\rceil ; \left\lceil \frac{w(\Omega_2)}{h_{c_r}} \right\rceil \right\} \geq$$

$$\max\left\{ \left\lceil \frac{n_1 k + n_2 h}{2k} \right\rceil ; \left\lceil \frac{n_2}{2} \right\rceil \right\}$$

If $\frac{n_2}{2} \geq \frac{n_1 k + n_2 h}{2k}$, then:

$$\frac{appr}{opt} \leq 2 \frac{n_1 + n_2 + 1}{n_2} \leq$$

$$2 \frac{n_2(2 - \frac{h}{k}) + 1}{n_2} \rightarrow 2\left(2 - \frac{h}{k}\right)$$

If $\frac{n_2}{2} \leq \frac{n_1 k + n_2 h}{2k}$, then:

$$\frac{appr}{opt} \leq \frac{n_1 + n_2 + 1}{\frac{n_1 k + n_2 h}{2k}} =$$

$$\frac{2(n_1 + n_2 + 1)}{n_1 + \frac{h}{k} n_2} \leq 4$$

by the bound proved in Theorem 3.2.

Case 2: Suppose $w_r > \frac{h}{2}$. Two further subcases may come up.

Case 2.1: For each color $\gamma \in \{\lceil \frac{\phi}{2} \rceil, \ldots, \phi - 1\}$ at least one of the two following conditions holds:

$$\sum_{I_r \in I(\gamma)[s_r]} w_r \geq \frac{k}{2} \tag{3}$$

$$\sum_{I_r \in I(\gamma)[s_r][c]} w_r \geq \frac{h_c}{2} \quad \text{for some } c \tag{4}$$

Therefore, the sum of the weights of the intervals active at time s_r is at least $\frac{h}{2}\lfloor \frac{\phi}{2} \rfloor$. Thus,

$$opt \geq \left\lceil \frac{\frac{h}{2}\lfloor \frac{\phi}{2} \rfloor}{k} \right\rceil \quad \text{and} \quad \frac{appr}{opt} \leq \frac{\phi}{\frac{\frac{h}{2}\lfloor \frac{\phi}{2} \rfloor}{k}} \rightarrow \frac{5k}{h}$$

Case 2.2: There is a color $\overline{\gamma} \in \{\lceil \frac{\phi}{2} \rceil, \ldots, \phi - 1\}$ for which both conditions (3) and (4) do not hold. Thus, there is an interval $I_{\overline{r}} = [s_{\overline{r}}, e_{\overline{r}})$, colored $\overline{\gamma}$, of weight $w_{\overline{r}} < \frac{h}{2}$. When $I_{\overline{r}}$ was colored, $\overline{\gamma}$ was the lowest possible indexed color. Therefore, this subcase reduces to Case 1 above loosing a factor of 2. □

It is worthy to note that in the case there are no constraints on the total weight of mutually overlapping intervals of the same category, the above algorithm yields a 4-approximation. This can be easily checked assuming $h = k$ in the proof of Theorem 4.1 and observing that, in the Case 2.1, only condition (3) holds for all the colors γ.

Moreover, note that the worst approximation constant of Algorithm 2 is given by $\frac{5k}{h}$ when $\frac{k}{h} > \frac{8}{5}$, and by 8 otherwise. However, an off-line algorithm for Problem 5 can be proposed which guarantees an 8-approximation even in the case that $\frac{k}{h} > \frac{8}{5}$. The algorithm runs three passes over the entire input sequence I. Each pass scans the intervals in I by non decreasing starting times and delivers a new set of colors. Hence, denoted by ϕ_i the number of colors delivered in pass i, the total number of colors employed is $\phi = \phi_1 + \phi_2 + \phi_3$.

Algorithm 3 THREE-PASS-GREEDY (I)

PASS 1: Color all intervals I_r with $w_r > \frac{k}{2}$;

PASS 2: Color all the intervals I_r with $w_r > \frac{1}{2} h_{c_r}$;

PASS 3: Color all the remaining intervals.

Note that, in the first pass, the same color cannot be assigned to two overlapping intervals, and hence the problem reduces to coloring an interval graph, which can be done optimally in polynomial time [3]. Therefore, $\phi_1 \leq opt$. In the second and third passes, instead, Algorithm 1 is employed. For the second pass alone, an approximation of 3 can be shown by modifying the proof of Theorem 3.2 obtaining $\phi_2 \leq 3 \cdot opt$. For the third pass alone, the analysis on the approximation guarantee of Theorem 3.2 can again be easily adapted by loosing a factor of 2, thus having $\phi_3 \leq 4 \cdot opt$. As a consequence, $appr = \phi_1 + \phi_2 + \phi_3 \leq 8 \cdot opt$.

5 Further Generalizations

This section considers two generalizations of the Server Allocation with Bounded Simultaneous Requests problem, where each request r is characterized by real bandwidths, normalized in $[0, 1]$ for analogy with the Bin-Packing problem [1].

In the first generalization, which contains Multi-Dimensional Bin-Packing as a special case, each request r is characterized by a *k-dimensional* bandwidth rate $\mathbf{w}_r = (w_r^{(1)}, \ldots, w_r^{(k)})$, where the c-th component specifies the bandwidth needed for the c-th category and k is the number of categories, i.e. $k = |C|$. The overall sum of the bandwidth rates of the simultaneous requests of the same category assigned to the same server at the same time is bounded by 1, which implies that the total sum of the bandwidth rates over all the categories is bounded by k.

Problem 6. (Interval Coloring with Multi-Dimensional Weighted Overlapping) *Given a set I of intervals, with each interval I_r characterized by a k-dimensional weight $\mathbf{w}_r = (w_r^{(1)}, \ldots, w_r^{(k)})$, where $w_r^{(c)} \in [0,1]$, for $1 \leq c \leq k$, assign a color to each interval in such a way that the overall sum of the weights of the same category for mutually overlapping intervals receiving the same color is bounded by* 1.

More formally, according to the notations introduced in Section 4, the constraints on the sum of the weights of the same category for mutually overlapping intervals receiving the same color can be stated as follows:

$$\sum_{I_r \in I(\gamma)[t][c]} w_r^{(c)} \leq 1 \qquad \forall\gamma, \forall t, \forall c \tag{5}$$

Note that the above constraints, added up over all the categories in C, imply the following redundant constraints:

$$\sum_{c=1}^{k} \sum_{I_r \in I(\gamma)[t][c]} w_r^{(c)} \leq k \qquad \forall\gamma, \forall t. \tag{6}$$

Problem 6 can be solved on-line by Algorithm 2, introduced in the previous section.

Theorem 5.1. *The First-Color algorithm provides a $4k$-approximation for Interval Coloring with Multi-Dimensional Weighted Overlapping.*

Proof: Assume the on-line algorithm employs ϕ colors. Consider the first interval $I_{\overline{r}} = [s_{\overline{r}}, e_{\overline{r}})$ which is colored ϕ. At time $s_{\overline{r}}$, $I_{\overline{r}}$ cannot be colored with any color in $\{1, \ldots, \phi - 1\}$.

Consider two cases.

Case 1: There is a color $\overline{\gamma}$ among $\{\lceil\frac{\phi}{2}\rceil, \ldots, \phi - 1\}$ such that for each component c, with $1 \leq c \leq k$:

$$\sum_{I_r \in I(\overline{\gamma})[s_{\overline{r}}][c]} w_r^{(c)} \leq \frac{1}{2}$$

Let $I_{r'}$ be any interval in $I(\overline{\gamma})[s_{\overline{r}}][c]$. Clearly, $w_{r'}^{(c)} \leq \frac{1}{2}$, for all c. Consider now instant $s_{r'}$, when interval $I_{r'}$ was colored $\overline{\gamma} \geq \lceil\frac{\phi}{2}\rceil$. Since $I_{r'}$ cannot be colored with any color in $\{1, \ldots, \lceil\frac{\phi}{2}\rceil - 1\}$, then for every $\gamma \in \{1, \ldots, \lceil\frac{\phi}{2}\rceil - 1\}$ and for every c, with $1 \leq c \leq k$:

$$\sum_{I_r \in I(\gamma)[s_{r'}][c_{r'}]} w_r^{(c)} > \frac{1}{2}$$

Hence,

$$opt \geq \sum_{I_r \in I[s_{r'}]} \frac{1}{k} \sum_{c=1}^{k} w_r^{(c)} \geq$$

$$\frac{1}{k} \sum_{\gamma=1}^{\lceil\frac{\phi}{2}\rceil - 1} \sum_{I_r \in I(\gamma)[s_{r'}]} \sum_{c=1}^{k} w_r^{(c)} \geq$$

$$\frac{1}{k} \sum_{\gamma=1}^{\lceil\frac{\phi}{2}\rceil - 1} \frac{1}{2} \geq \frac{1}{k} \frac{\phi}{2} \frac{1}{2} \geq \frac{\phi}{4k}$$

Case 2: For every color γ in $\{\lceil\frac{\phi}{2}\rceil, \ldots, \phi - 1\}$, there is a category c, with $1 \leq c \leq k$, such that

$$\sum_{I_r \in I(\gamma)[s_r][c]} w_r^{(c)} > \frac{1}{2}.$$

By a reasoning analogous to Case 1, it follows that:

$$opt \geq \sum_{I_r \in I[s_{r'}]} \frac{1}{k} \sum_{c=1}^{k} w_r^{(c)} \geq$$

$$\frac{1}{k} \sum_{\gamma=\lceil\frac{\phi}{2}\rceil}^{\phi - 1} \sum_{I_r \in I(\gamma)[s_{r'}]} \sum_{c=1}^{k} w_r^{(c)} \geq$$

$$\frac{1}{k} \sum_{\gamma=\lceil\frac{\phi}{2}\rceil}^{\phi - 1} \frac{1}{2} \geq \frac{1}{k} \frac{\phi}{2} \frac{1}{2} \geq \frac{\phi}{4k}$$

Since $appr = \phi$, an approximation of $4k$ holds. □

The above problem, when considered as an off-line problem, is APX-hard since it contains a multidimensional bin-packing as a special case. Multidimensional bin-packing is APX-hard [7] already for $|C| = 2$.

In the second generalization, each request r is characterized by a *gender bandwidth rate* g_{r,c_r} associated to the category c_r and by a bandwidth rate w_r. The overall sum of the bandwidth rates of the simultaneous requests assigned to the same server at the same time is bounded by 1, as well as the overall sum of the gender bandwidth rates of the simultaneous requests of the same category assigned to the same server at the same time, which is also bounded by 1.

Problem 7. (Interval Coloring with Double Weighted Overlapping) *Given a set I of intervals, with each interval I_r characterized by a gender weight $g_{r,c_r} \in (0,1]$ associated to the category c_r and by a bandwidth weight $w_r \in (0,1]$, assign a color to each interval in such a way that the overall sum of the gender weights for mutually overlapping intervals of the same category receiving the same color is bounded by 1, the overall sum of the bandwidth weights for mutually overlapping intervals receiving the same color is bounded by 1, and the minimum number of colors is used.*

Formally, the constraints of Problem 7 are given below:

$$\sum_{I_r \in I(\gamma)[t]} w_r \leq 1 \qquad \forall \gamma, \forall t \tag{7}$$

$$\sum_{I_r \in I(\gamma)[t][c]} g_{r,c} \leq 1 \qquad \forall \gamma, \forall t, \forall c \tag{8}$$

Problem 7 can again be solved on-line by Algorithm 2, introduced in the previous section.

Theorem 5.2. *The First-Color algorithm provides a 10-approximation for Interval Coloring with Double Weighted Overlapping.* □

6 Conclusions

This paper has considered several on-line approximation algorithms for problems arising in infostations, where a set of requests characterized by categories and temporal intervals have to be assigned to servers in such a way that a bounded number of simultaneous requests are assigned to the same server and the number of servers is minimized. However, several questions still remain open. For instance, one could lower the approximation bounds derived for Problems 5, 6 and 7. Moreover, one could consider the scenario in which the number of servers is given in input, each request has a deadline, and the goal is to minimize the overall completion time for all the requests.

References

[1] E.G. Coffman, G. Galambos, S. Martello & D. Vigo, "Bin Packing Approximation Algorithms: Combinatorial Analysis", in *Handbook of Combinatorial Optimization*, DZ. Du & P.M. Pardalos (Editors), 1999, Kluwer, Dordrecht, pp. 151-207.

[2] M.R. Garey & D.S. Johnson, *Computers and Intractability*, 1979, Freeman, San Francisco.

[3] M.C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, 1980, Academic Press, New York.

[4] D.J. Goodman, J. Borras, N.B. Mandayam & R.D. Yates, "INFOSTATIONS: A New System Model for Data and Messaging Services", *IEEE Vehicular Technolgy Conference*, 1997, pp. 969-973.

[5] R.M. Karp, "On-Line Algorithms Versus Off-Line Algorithms: How Much is it Worth to Know the Future?", in *Proceedings of the IFIP 12th World Computer Congress. Volume 1: Algorithms, Software, Architecture*, Jan van Leeuwen (Editor), 1992, Elsevier Science Publishers, Amsterdam, pp. 416-429.

[6] E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, & H. Shmoys, *Sequencing and Scheduling: Algorithms and Complexity*, 1993, North-Holland, Amsterdam.

[7] G.J. Woeginger, "There is No Asymptotic PTAS for Two-Dimensional Vector Packing", *Information Processing Letters*, 1997, pp. 293-297.

[8] G. Wu, C.W. Chu, K. Wine, J. Evans & R. Frenkiel, "WINMAC: A Novel Transmission Protocol for Infostations", *IEEE Vehicular Technolgy Conference*, 1999, pp. 1-15.

[9] J. Zander, "Trends and Challenges in Resource Management Future Wireless Networks", *IEEE Wireless Communications & Networks Conference*, 2000.